Mol Gen Genet (1987) 207:288–293

MGG

Nucleotide sequence of cDNA clones encoding the complete "33 kDa" precursor protein associated with the photosynthetic oxygen-evolving complex from spinach

A. Tyagi[1], J. Hermans[1], J. Steppuhn[1], Ch. Jansson[2], F. Vater[3], and R.G. Herrmann[1]

[1] Botanisches Institut der Ludwig-Maximilians-Universität, Menzinger Str. 67, D-8000 München 19, Federal Republic of Germany
[2] Arrhenius Laboratory, University of Stockholm, S-10691 Stockholm, Sweden
[3] Technische Universität Berlin, Institut für Biochemie, Franklinstr. 29, D-1000 Berlin 10, Federal Republic of Germany

Summary. Several cDNA clones encoding the "33 kDa" protein associated with the photosynthetic water oxidation activity of spinach were sequenced. A 1208 bp insert of one of the clones encodes the entire 331 amino acid residues of the precursor protein including 84 amino acids (8.5 kDa) of the amino-terminal transit peptide, 49 bp of the 5′ and 111 bp of the 3′ untranslated segment of the mRNA. The 3′ poly(A) tail starts 19 bp downstream from a putative polyadenylation signal, TATAAA. The hydrophilic mature protein consists of 247 amino acid residues corresponding to an M_r of 26.5 kDa, which is 6.5 kDa smaller than the value determined by SDS-polyacrylamide gel electrophoresis (33–34 kDa), and shows a certain degree of conservation with the putative Mn-complexing active sites of bacterial Mn-dependent superoxide dismutases. The anatomy of the unusually long transit sequence is discussed with regard to current concepts of protein import into and protein routein within the organelle.

Key words: Photosynthetic oxygen evolution – "33 kDa" protein – cDNA nucleotide sequence – Transit peptide – Spinach

Introduction

Recent evidence suggests that approximately half of the genes for the ca. 50 polypeptide species that constitute thylakoid membranes of higher plants are located in the plastid, the other half originate in the nucleus (Herrmann et al. 1985; Tittgen et al. 1986). The nuclear components, synthesized on cytosolic ribosomes as precursors with a transient covalently linked transit sequence, are imported into the organelle, post-translationally processed, and the mature proteins are assembled at their proper locations within the membrane (Chua and Schmidt 1979; Cashmore et al. 1985; Karlin-Neumann and Tobin 1986; Tittgen et al. 1986). This dual genetic origin of eukaryotic thylakoid membranes raises intriguing questions concerning intergenomic integration within the cell, specifically the regulation of expression of corresponding plastid- and nuclear-encoded genes, the modes of transport of the cytosolically made components into the organelle, the maturation of these proteins and their topogenesis into functional multisubunit complexes.

To clarify these questions, the nucleotide sequences of the genes and the primary structure of their products need to be determined. Significant progress has been made in defining plastome-encoded genes for thylakoid proteins (reviewed in Herrmann et al. 1985; see also Hennig and Herrmann 1986; Kirsch et al. 1986; Cozens et al. 1986; Steinmetz et al. 1986; Westhoff et al. 1986). On the other hand, equivalent information on the nuclear-encoded thylakoid protein complement is still limited, primarily due to technical problems in selecting discrete genes or cDNAs from the complex genome. So far, only three nuclear-encoded genes for thylakoid proteins have been characterized, viz. the chlorophyll *a*/*b* apoproteins of the light-harvesting complex associated with photosystem II (Coruzzi et al. 1983; Dunsmuir et al. 1983; Cashmore 1984; Karlin-Neumann et al. 1985; Kohorn et al. 1986), ferredoxin (Smeekens et al. 1985a) and plastocyanin (Herrmann et al. 1983; Smeekens et al. 1985b; Rother et al. 1986). Using expression cloning in phage lambda gt11 (Young and Davis 1983) we have recently succeeded in the isolation of cDNA-containing chimaeric plasmids for 20 nuclear-encoded thylakoid proteins (Tittgen et al. 1986 and unpublished results) including recombinant DNAs for three components of the photosynthetic water-splitting (oxygen-evolving) complex. The analysis of cDNAs for the largest of these proteins, the so-called lysine-rich "33 kDa" protein that may be involved in stabilizing the Mn atoms associated with the complex (Åkerlund and Jansson 1981; Kuwabara and Murata 1982; Andersson 1986; Miyao and Murata 1987), is the subject of this paper.

The oxygen-evolving unit of the photosynthetic electron transport chain is located in the thylakoid lumen and associated with the intrinsic core complex of the photosystem II reaction centre. It catalyses the reduction of a strong oxidant, produced through the light reactions of photosystem II, by the oxidation of water to molecular oxygen. The mechanism by which this fundamental process occurs is not known and the composition of the complex as well as the function of the individual components involved in the catalysis of water oxidation are just beginning to be elucidated (reviewed in Andersson 1986). Recent evidence suggests that the native complex is a composite of at least six subunit species with approximate molecular masses of 33, 24, 23, 22, 16 and 10 kDa (Ljungberg et al. 1986). Three major, hydrophilic polypeptides (33, 23 and 16 kDa) are synthesized from poly(A) RNA as larger precursors on cytosolic ribosomes, translocated into the thylakoid lumen by isolated chloroplasts and processed to their mature form (Westhoff et al. 1985). In this paper, we provide the nucleo-

Offprint requests to: R. Herrmann

tide sequence of cDNA clones that bear the complete 33 kDa precursor protein. Equivalent data for the 23 and 16 kDa polypeptides will be presented elsewhere (Jansen et al. 1987).

Material and methods

Restriction enzymes, T4 DNA ligase and Klenow fragment of DNA polymerase I were obtained from Boehringer (Mannheim) and used according to the instructions given by this company; α-[^{32}P]dNTPs were from Amersham (Braunschweig). All chemicals were of analytical reagent grade.

The N-terminal sequence of the 33 kDa protein was determined by automatic Edman degradation of the purified protein with a solid phase sequencer, and the residues were identified by high-performance liquid chromatography as described (Vater et al. 1986).

The isolation and characterization of the basic recombinant cDNA clones in pBR322 and phage lambda gt11 for the 33 kDa protein associated with the water-oxidizing complex have been described (Tittgen et al. 1986). Nick translation, Southern blot analysis, phage hybridization, DNA subcloning in pUC118 (Messing and coworkers, personal communication), pSP65 (Melten et al. 1984) and Bluescribe M13 (Vector Cloning Systems, San Diego, Calif.) vectors have also been described (Tittgen et al. 1986). Small- and large-scale plasmid or phage purification, 3′ fill-in synthesis with the Klenow fragment, and DNA sequence analysis according to the chemical modification and chain cleavage procedure (Maxam and Gilbert 1980) were as described by Tittgen et al. (1986). DNA and protein sequence analysis and alignment were performed using computer programs provided by Kröger and Kröger-Block (1982). Hydropathy analysis was performed according to Kyte and Doolittle (1982) using programs supplied by G. Rauscher.

Results and discussion

cDNA clone selection and characterization

The rationale for cloning DNA sequences complementary to the 33 kDa polypeptide associated with the oxygen-evolving complex has been described in Tittgen et al. (1986). The pBR 322 and lambda gt11 libraries, derived from poly(A)-containing RNA of etiolated spinach seedlings illuminated for 16 h, were screened by hybrid-selection translation and immunologically using a monospecific, polyclonal antiserum. Seventy-five positive recombinant DNAs were isolated from approximately 4×10^5 transformants and their inserts sized on polyacrylamide gels.

The two largest clones, l23Soc33-1 and l23Soc33-2 (inserts of 1489 and 1116 bp), of the initial screening cycle were sequenced and found to be incomplete. They contained information for the entire mature protein and an additional in-phase 78 or 82 amino acid residues, respectively, at the N-terminus but the open reading frame continued to the 5′ end without a terminal methionine residue. (l23Soc33-1 carries a long stretch of 514 nucleotides from the 3′ untranslated region.) In order to isolate more complete cDNA clones, the 5′ terminal part of the insert was prepared by digestion with *Sal*I and the resulting 350 bp *Eco*RI-*Sal*I fragment was used as a probe for re-screening the lambda gt11 library after end labelling. Of 23 positive

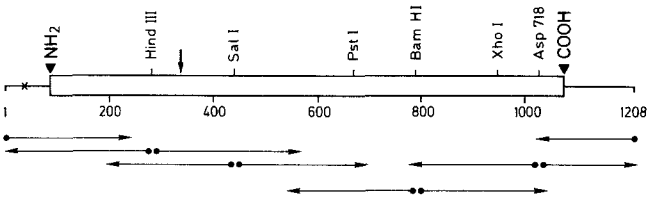


Fig. 1. Strategy for sequencing the cDNA insert of the recombinant plasmid p23Soc33-1208. The restriction map is drawn 5′ to 3′. The positions of the amino- and carboxyl-termini as well as the terminal processing site of the precursor protein are indicated by *arrowheads* and an *arrow*, respectively. The *cross* (x) marks the 5′ terminal complementary sequence interval similar to the mRNA 3′ end (see text). *Horizontal arrows* denote the direction and extent of individual sequence reactions

phage obtained, a clone (l23Soc33-3) containing a 1208 bp cDNA fragment was found to carry the complete precursor sequence for the 33 kDa protein and 49 nucleotides of 5′ and 111 nucleotides of 3′ untranslated sequence excluding a 20 nucleotide poly(A) chain (for restriction map and sequencing strategy see Fig. 1). The difference between the size of the mRNA as estimated by Northern blot analysis (re-determined to be ca. 1300 nucleotides; cf. Tittgen et al. 1986) and the size of this cDNA insert makes it unlikely that l23Soc33-3 constitutes a full-length cDNA copy of a mature spinach 33 kDA polypeptide mRNA. Computer-assisted translation of the nucleotide sequence revealed a single, large open reading frame of 993 nucleotides coding for a protein of 331 amino acid residues of a calculated mass of 35 kDa and ending with an ochre codon at position 994–996 (Fig. 2).

A complication arising from this analysis was that the 5′ untranslated 35 nucleotide segment of the p23Soc33-3 insert revealed a dT tail of 16 nucleotides followed by a stretch of 19 nucleotides identical to the non-coding strand at the 3′ end of this cDNA. Subsequently, the sequence diverged completely (Fig. 1). Similar 5′ poly(dT) tails and mRNA non-specific sequences have been observed in cDNAs for ferredoxin and plastocyanin precursors and interpreted as an artefact (Smeekens et al. 1985a, b). We, therefore, assume that the first 35 nucleotides of the 5′ end of the insert reflect a cloning error and are possibly derived from the 3′ end of another RNA or cDNA molecule for the 33 kDa protein. That the remaining 5′ untranslated sequence is specific to the mRNA is strengthened by the observations that (a) the subsequent sequence is completely unrelated to the 3′ sequence, (b) two other independent cDNA inserts, shorter by only 2–6 amino acids for the signal peptide, exhibit identical nucleotide sequences, and (c) that the typical consensus motifs for the translational initiation 'scanning model' in eukaryotes, e.g. A at position -3 and G at $+4$, are found at the assumed initiatory methionine residue (Kozak 1986). Moreover, the open reading frame extending into the 5′ untranslated region, contains an in-frame stop codon 15 triplets upstream of the assumed initiator codon.

The 3′ untranslated region of p23Soc33-3 includes a hexanucleotide motif, TATAAA, at position 1080–1085 which might serve as a polyadenylation signal 19 bp downstream. However, the 3′ untranslated region of various clones, e.g. of p23Soc33-1, showed length polymorphism and a variable pentanucleotide sequence (ATATT instead of AAAAA) just before the poly(A) tail in one of the

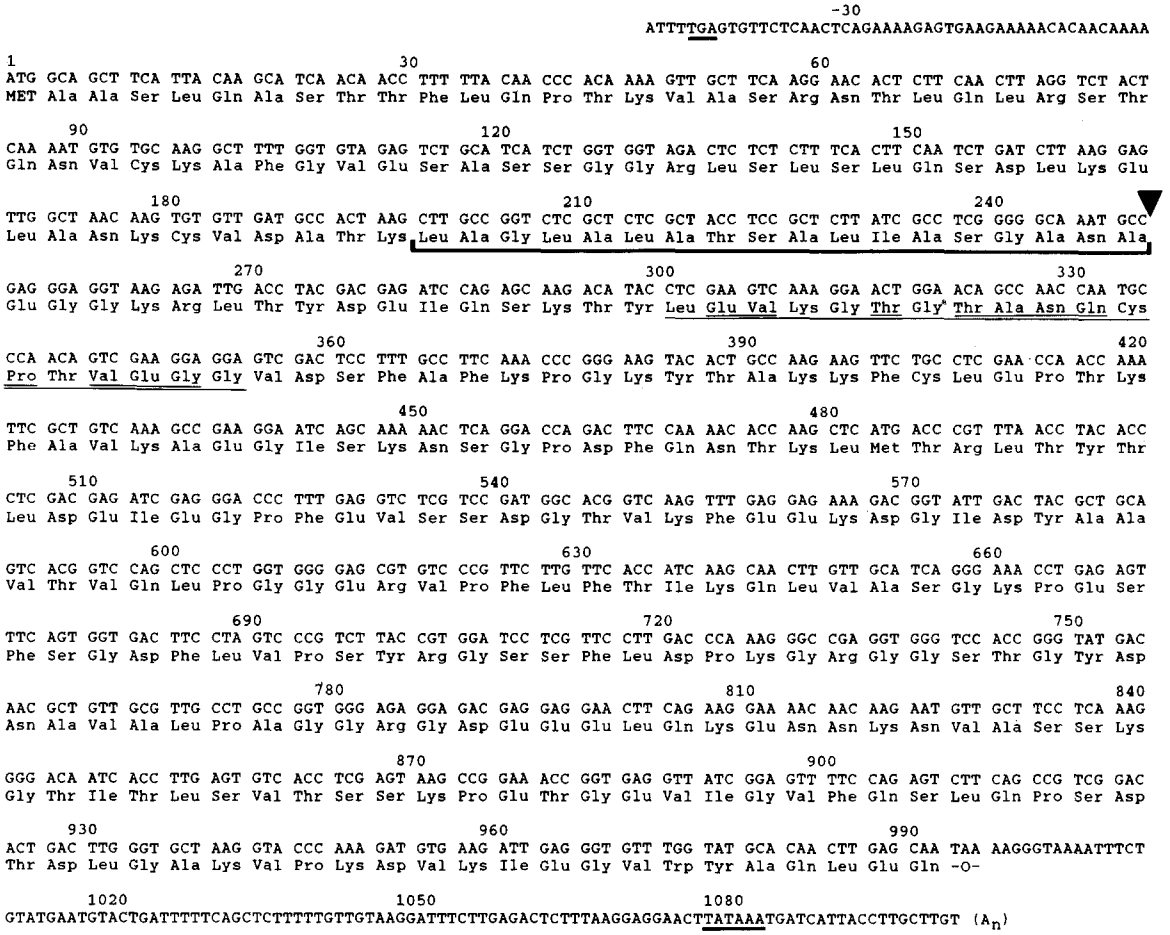


Fig. 2. Nucleotide sequence (coding strand) and predicted amino acid sequence of the cDNA insert of the chimaeric plasmid p23Soc33-3 carrying the complete "33 kDa" precursor protein associated with the photosynthetic water oxidation complex from spinach. The complementary 5′ duplication of the 3′ mRNA (see text and Fig. 1) is omitted. Nucleotide residues are numbered in the 5′ to 3′ direction beginning with the first residue encoding the putative initiatory methionine. A potential polyadenylation site is *underlined.* The terminal processing site of the precursor is marked by an *arrowhead,* and a putative h-domain (Von Heijne 1986b) indicated by a *horizontal bracket.* An interval of 18 amino acids (nucleotides 301–354) that exhibits similarity to Mn-binding sites in bacterial superoxide dismutases (see text) is *underlined*

four cDNA clones sequenced (data not shown). It remains to be clarified whether this represents multiple polyadenylation sites such as have been deduced from cDNA clones of other plant genes (Dean et al. 1986; Hernández-Lucas et al.; Rother et al. 1986) and/or the existence of a paucigene family which had been indicated by genomic Southern blots (Tittgen et al. 1986). We have isolated three independent genomic clones from an EMBL4 library of spinach nuclear DNA which show identical restriction patterns favouring the former alternative (J. Bichler, personal communication).

Protein characteristics

The deduced amino acid sequence from codon 85 onwards clearly confirms the identity of the cDNA clones. Apart from the fact that the determined and derived amino acid composition, corrected for the lower molecular weight of the protein (see below), agrees well (data not shown), both our own determination of 14 N-terminal amino acid residues and the recently published amino acid sequence for the mature protein (Oh-oka et al. 1986) match completely with the corresponding amino acid sequence deduced from the nucleotide sequence except at the C-terminus where only a single Gln residue (instead of two) is present. Oh-oka et al. (1986) have also found sequence variability at the C-terminus based on an analysis of tryptic and Staphylococcal protease peptides. In our hands all four cDNA studied ended with the same C-terminal residue. Our findings imply also that the terminal cleavage site of the precursor is Ala/Glu (nucleotide position 252/253; Fig. 2) which is consistent with the observation that transit sequences generally terminate in residues with small neutral side chains (e.g. Von Heijne 1986a). Processing ultimately yields a protein of 26.5 kDa (247 amino acid residues). This is 6.5 kDa less than the value calculated from SDS-polyacrylamide gels (cf. also Oh-oka et al. 1986; Andersson 1986). The reason for this difference is unknown; similar discrepancies have also been observed with various other thylakoid membrane proteins.

The hydropathy profile (Kyte and Doolittle 1982), calculated for a moving window of 11 amino acids, demonstrates as expected the hydrophilic nature of the protein (Fig. 3). Inspection of the amino acid sequence for the mature protein discloses three regions rich in charged amino acids, viz. residues 185–193, 262–274 and 308–322 (Fig. 2).

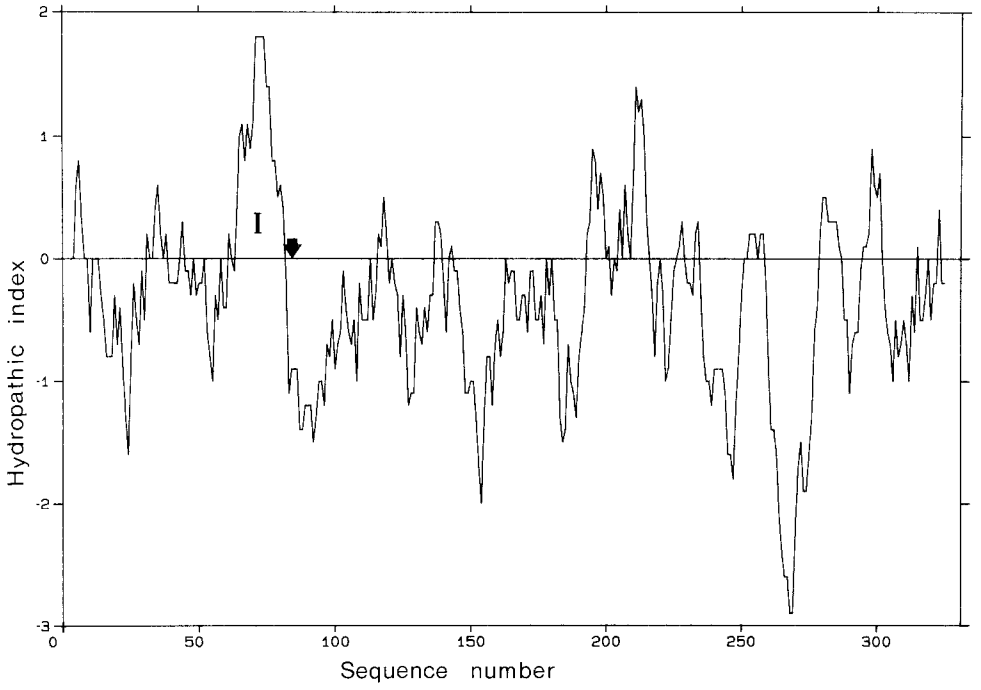


Fig. 3. Hydropathy profile of the "33 kDa" (precursor) protein associated with the photosynthetic water oxidation complex from spinach. The profile was calculated according to Kyte and Doolittle (1982) using an 11-point moving interval. The *arrowhead* indicates the terminal processing site of the precursor in the chloroplast. A possible h-domain of the transit sequence (Von Heijne 1986b) is marked by a *roman number* (I)

In all, about 25 % residues are charged with an approximately balanced number of acidic (35) and basic (29) residues. Of the latter 23 represent Lys residues. A certain degree of sequence conservation has been noted between proposed Mn ligands in bacterial Mn superoxide dismutases (Brook and Walker 1980; Steinmann 1978; Marres et al. 1985) and residues 101–118 of the 33 kDa protein (Oh-oka et al. 1986), but the putative ligand Asp in Mn superoxide dismutases has been replaced by a Cys residue at an equivalent position (112) in the mature 33 kDa protein (Fig. 2). It remains to be shown whether this segment is, in fact, of importance in retaining Mn atom(s).

The transit sequence

The first 84 N-terminal amino acids represent the transit peptide for the 33 kDa protein. This sequence must possess information for both, the selective transfer of the protein into the organelle and its routeing towards the inner surface of the thylakoid membrane. The specific import by isolated chloroplasts of plastid protein precursors including those for the oxygen-evolving complex (Westhoff et al. 1985; Tittgen et al. 1986) from cell-free translation assays of polyadenylated RNA is well established (Grossman et al. 1982; Cline et al. 1985). Seven classes of proteins distinct with respect to their intraorganelle location (outer and inner envelope, intermembrane space, stroma, membrane-extrinsic or -intrinsic, lumen) utilize the translocation and modification machinery, and each of the components examined reached its correct intraorganelle location (e.g. Grossman et al. 1982; Cline et al. 1985; Herrmann et al. 1985; Westhoff et al. 1985; Tittgen et al. 1986).

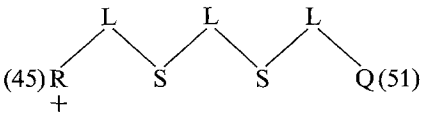
Although several attempts have been made to delineate structurally and functionally discrete regions of transiently attached signal or transit peptides (Hurt and von Loon 1986; Karlin-Neumann and Tobin 1986; Von Heijne 1986b) it is not yet possible to describe the relevant topogenic signals in terms of primary sequence. Comparison of the transit sequences for chloroplast-imported polypeptides (Karlin-Neumann and Tobin 1986), as for those of mitochondria (e.g. Hurt and von Loon 1986; Pfanner and Neupert 1986; Von Heijne 1986b), does not show any apparent homology at the primary sequence level indicating that proteases and receptors recognize either predominantly structural feature rather than primary sequence and/or that this reflects distinct import routes for various peptide classes. It is also not clear whether transit peptides alone or their coevolution with the native protein or other structural components are required for *efficient* translocation.

The concept that guides current work is that presequences are organized in a linear array of functionally distinct domains, e.g. targeting, stop-transfer and processing elements, that they are rich in charged and hydroxylated amino acid residues (n-domain), and frequently contain a relatively long stretch of uncharged amino acids (h-domain) that is followed by a cleavage site (c-domain). However, the two basic (non-exclusive) hypotheses, the block/interblock model (Karlin-Neumann and Tobin 1986), and the n-, h-, c-domain model (Hurt and van Loon 1986; Von Heijne 1986b) cannot easily be reconciled with transit sequences of lumen proteins since the 33 kDa precursor polypeptide, like the precursors for the 16 kDa and 23 kDa proteins of the oxygen-evolving unit or for plastocyanin, traverse three membranes, the two envelopes *and* the thylakoid membrane (cf. Rother et al. 1986; Smeekens et al. 1986). The block/interblock model is based primarily on the analysis of transit sequences of small subunits of the stromal enzyme ribulose bisphosphate carboxylase/-oxygenase and of the integral chlorophyll *a*/*b* apoproteins of light-harvesting complexes II that are available from different organisms (Karlin-Neumann and Tobin 1986). Therefore, by analogy with nuclear-encoded proteins present in the intermembrane space of mitochondria (e.g. Pfanner and Neupert 1986), processing of lumen proteins may generally occur in at least two steps and involve two different protease activities. This, in turn, should imply that the transit peptide for the lumen proteins is a composite of (at least) two discrete transit sequences, one each for the envelopes and the

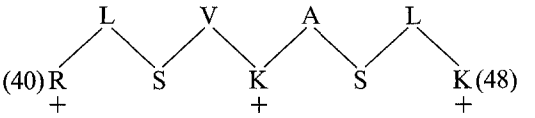
thylakoid membrane (Rother et al. 1986; Smeekens et al. 1986), and that each of these parts is internally structured into distinct domains. Several observations outlined below are consistent with this view.

The targeting sequences for the 33 kDa polypeptide and plastocyanin (Rother et al. 1986) show a consistent pattern of characteristic features that includes different elements known from other plastid and mitochondrial proteins of nuclear origin. The transit peptide of the 33 kDa polypeptide shares the presence of Arg and Lys residues, an abundance of the hydroxylated amino acids, Ser and Thr, and general design with mitochondrial transit peptides (e.g. Hurt and von Loon 1986; Pfanner and Neupert 1986). It differs with regard to length, by the presence of (four) acidic residues (the overall ratio of basic to acidic residues is 2:1 for the transit peptide, while in the mature protein both types of residues are approximately equal), in the moderate hydrophobicity of the hydrophobic domain, and in secondary structure predictions (see below). Like the transit sequences for the small subunit of ribulose biphosphate carboxylase and for the chlorophyll *a/b* apoproteins, it starts with a relatively long stretch of 15 uncharged amino acid residues (see Karlin-Neumann and Tobin 1986 and references cited therein), is followed by 51 amino acids interspersed with charged (8 basic/4 acidic) residues, then by 18 uncharged amino acids that form a relatively hydrophobic region (Fig. 3), and ends at the final cleavage site. The N-terminal sequence MAASLQAS exhibits a remote similarity to the homology box I, MA.S.M.SS, proposed by Karlin-Neumann and Tobin (1986), but no clear-cut homology to boxes II (P.F.G.K) and III (G.GRV) is detectable. Whether amino acids GGRL at position 43–46 or G_{59} G_{81} N_{83} A_{84} which are remarkably similar to sequence elements in the plastocyanin transit peptide (Rother et al. 1986) are related to two processing events and could serve any proposed function of homology box III remains to be shown (see below).

The 33 kDa precursor chain, like that of plastocyanin (Smeekens et al. 1986), appears to be removed by (at least) two successive proteolytic cleavages (unpublished observation) which are probably catalysed by different proteases (Robinson and Ellis 1984; Robinson, personal communication). The first processing step should occur during or after translocation into the organelle, similar to that of stromal or intrinsic proteins. The intermediate may then span the thylakoid membrane transiently, via its hydrophobic domain, with its N-terminus facing the matrix space, and the second cleavage should occur at the inner thylakoid surface. The presence of this latter processing domain ensures the release of the protein into the lumen. It is noteworthy that secondary structure predictions (Chou and Fasman 1978; Garnier et al. 1978; Link and von Jagow, personal communication) for precursor segments of lumen proteins show remarkably consistent features. For example, in the 33 kDa protein and plastocyanin (Rother et al. 1986) the region of the putative intermediate processing sites (residues 45–55 in the former, residues 42–48 in the latter case) possesses the potential of forming amphipathic β-sheets,



(33 kDa protein) and



(plastocyanin; Rother et al. 1986); the structure of this region is conserved (residue 39–45) in *Silene* (Smeekens et al. 1985b). These elements are followed by a predicted turn to a long α-helical segment including the hydrophobic domain and then by the terminal processing site (Figs. 2 and 3). The hydrophobic regions may, therefore, serve as transfer stop signals in translocation (cf. also Smeekens et al. 1986). The transit sequences of the lumen 16 kDa and, if hydrophobicity scales are corrected in terms of absolute values, of the 23 kDa polypeptide contain comparable elements at equivalent positions (data not shown).

The analysis of chloroplast transit sequences clearly deserves further study. A comparison of transit sequences for approximately ten nuclear-encoded thylakoid proteins including those for the four lumen proteins will be the subject of a separate paper.

Acknowledgements. The authors thank Drs. Link and von Jagow for advice on secondary structure evaluation. This work was supported by the Deutsche Forschungsgemeinschaft (grant He 693) and the Fonds der Chemischen Industrie. AT is a Visiting Academic Fellow under a special Gene Technology Programme of the German Academic Exchange Service.

References

Åkerlund H-E, Jansson C (1981) Localization of a 34000 and a 23000 Mr polypeptide to the lumenal side of the thylakoid membrane. FEBS Lett 124:229–232

Andersson B (1986) Proteins participating in photosynthetic water oxidation. Encycl Plant Physiol 19:447–456

Brook CJ, Walker JE (1980) Superoxide dismutase from *Bacillus stearothermophilus*: complete amino acid sequence of a manganese enzyme. Biochemistry 19:2873–2882

Cashmore AR (1984) Structure and expression of a pea nuclear gene encoding a chlorophyll a/b-binding polypeptide. Proc. Natl Acad Sci USA 81:2960–2964

Cashmore A, Szabo L, Timko M, Kausch A, Van de Broeck G, Schreier P, Bohnert H, Herrera-Estrella L, Van Montagu M, Schell J (1985) Import of polypeptides into chloroplasts. Bio/Technology 3:803–808

Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol 47:45–148

Chua N-H, Schmidt GW (1979) Transport of proteins into mitochondria and chloroplasts. J Cell Biol 81:461–483

Cline K, Werner-Washburne M, Lubben TH, Keegstra K (1985) Precursors to two nuclear-encoded chloroplast proteins bind to the outer envelope membrane before being imported into chloroplasts. J Biol Chem 260:3691–3696

Coruzzi G, Broglie R, Cashmore A, Chua N-H (1983) Nucleotide sequences of two pea cDNA clones encoding the small subunit of ribulose 1.5-bisphosphate carboxylase and the major chlorophyll a/b-binding thylakoid polypeptide. J Biol Chem 258:1399–1402

Cozens AL, Walker JE, Phillips AL, Huttly AK, Gray JC (1986) A sixth subunit of ATP synthase, an F component, is encoded in the pea choroplast genome. EMBO J 5:217–222

Dean C, Tamaki S, Dunsmuir P, Fareau M, Katayama C, Dooner H, Bedbrook J (1986) mRNA transcripts of several plant genes

are polyadenylated at multiple sites *in vivo*. Nucleic Acids Res 13:3179–3194
Dunsmuir P, Smith SM, Bedbrook J (1983) The major chlorophyll a/b binding protein of petunia is composed of several polypeptides encoded by a number of distinct nuclear genes. J Mol Appl Genet 2:285–300
Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97–120
Grossman AR, Bartlett S, Schmidt GW, Mullet JE, Chua N-H (1982) Optimal conditions for post-translational untake of proteins by isolated chloroplasts. J Biol Chem 257:1558–1563
Hennig J, Herrmann RG (1986) Chloroplast ATP synthase of spinach contains nine nonidentical subunit species, six of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement. Mol Gen Genet 203:117–128
Hernández-Lucas C, Royo J, Paz-Ares J, Ponz F, Garcia-Olmedo F, Carbonero P (1986) Polyadenylation site heterogeneity in mRNA encoding the precursor of the barley-toxin β-hordothionin. FEBS Lett 200:103–106
Herrmann RG, Westhoff P, Alt J, Winter P, Tittgen J, Bisanz C, Sears BB, Nelson N, Hurt E, Hauska G, Viebrock A, Sebald W (1983) Identification and characterization of genes for polypeptides of the thylakoid membrane. In: Cifferi O, Dure III (eds) Structure and function of plant genomes. Plenum Publ, New York, pp143–154
Herrmann RG, Westhoff P, Alt J, Tittgen J, Nelson N (1985) Thylakoid membrane proteins and their genes. In: van Vloten-Doting L, Groot GSP, Hall TC (eds) Molecular form and function of the plant genome. Plenum Publ, New York, pp233–256
Hurt EC, von Loon APGM (1986) How proteins find mitochondria and intramitochondrial compartments. Trends Biochem Sci 11:204–207
Jansen T, Rother Ch, Steppuhn J, Reinke H, Beyreuther K, Jansson C, Andersson B, Herrmann RG (1987) Nucleotide sequence of cDNA clones encoding the complete "23 kd" and "16 kd" precursor proteins associated with the photosynthetic oxygen-evolving complex from spinach. FEBS Lett (in press)
Karlin-Neumann GA, Tobin EM (1986) Transit peptides of nuclear-encoded chloroplast protein share a common amino acid framework. EMBO J 5:9–13
Karlin-Neumann GA, Kohorn BD, Thornber JP, Tobin EM (1985) A chlorophyll a/b-protein encoded by a gene containing an intron with characteristics of a transposable element. J Mol Appl Genet 3:45–61
Kirsch W, Seyer P, Herrmann RG (1986) Nucleotide sequence of the clustered genes for two P_{700} chlorophyll a apoproteins of the photosystem I reaction center and the ribosomal protein S14 of the spinach plastid chromosome. Curr Genet 10:843–855
Kohorn BD, Harel E, Chitnis PR, Thornber P, Tobin EM (1986) Functional and mutational analysis of the light-harvesting chlorophyll a/b protein of thylakoid membranes. J Cell Biol 102:972–981
Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283–292
Kröger M, Kröger-Block A (1982) A flexible new computer program for handling DNA sequence data. Nucleic Acids Res 10:229–236
Kuwabara T, Murata N (1982) Inactivation of photosynthetic oxygen evolution and concomitant release of three polypeptides in the photosystem II particles of spinach chloroplasts. Plant Cell Physiol 23:533–539
Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132
Ljungberg U, Åkerlund H-E, Andersson B (1986) Isolation and characterization of the 10-kDa and 22-kDa polypeptides of higher plant photosystem 2. Eur J. Biochem 158:477–482
Marres CA, van Loon APGM, Oudshoorn P, van Steeg H, Grivell LA, Slater EC (1985) Nucleotide sequence analysis of the nuclear gene coding for manganese superoxide dismutase of yeast mitochondria, a gene previously assumed to code for the Rieske iron-sulphur protein. Eur J Biochem 147:153–161
Maxam AM, Gilbert W (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol 64:499–560
Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promotor. Nucleic Acids Res 12:7035–7056
Miyao M, Murata N (1987) Oxygen evolving complex of photosystem II and photoinactivation. Topics in Photosynthesis 9 (in press)
Oh-Oka H, Tanaka S, Wada K, Kuwabara T, Murata N (1986) Complete amino acid sequence of 33 kDa protein isolated from spinach photosystem II particles. FEBS Lett 197:63–66
Pfanner N, Neupert W (1986) Biogenesis of mitochondrial energy-transducing complexes. Curr Top Bioenerget 15 (in press)
Robinson C, Ellis RJ (1984) Transport of proteins into chloroplasts, partial purification of a chloroplast protease involved in the processing of imported precursor polypeptides. Eur J Biochem 142:337–342
Rother C, Jansen T, Tyagi A, Tittgen J, Herrmann RG (1986) Plastocynin is encoded by an uninterrupted nuclear gene in spinach. Curr Genet 11:171–176
Smeekens S, van Binsbergen J, Weisbeek P (1985a) The plant ferredoxin precursor: nucleotide sequence of a full length cDNA clone, Nucleic Acids Res 13:3179–3194
Smeekens S, de Groot M, van Binsbergen J, Weisbeek P (1985b) Sequence of the precursor of the chloroplast thylakoid lumen protein plastocyanin. Nature 317:456–458
Smeekens S, Bauerle C, Hageman J, Keegstra K, Weisbeek P (1986) The role of the transit peptide in the routing of precursors toward different chloroplast compartments. Cell 46:356–375
Steinmann HM (1978) The amino acid sequence of mangano superoxide dismutase from *Escherichia coli* B. J Biol Chem 253:8708–8720
Steinmetz AA, Castroviejo M, Sayre RT, Bogorad L (1986) An additional component of photosystem II identified through its plastid gene in maize. J Biol Chem 261:2485–2488
Tittgen J, Hermans J, Steppuhn J, Jansen T, Jansson C, Andersson B, Nechushtai R, Nelson N, Hermann RG (1986) Isolation of cDNA clones for fourteen nuclear-ended thylakoid membrane proteins. Mol Gen Genet 204:258–265
Vater J, Salnikow J, Jansson C (1986) N-terminal sequence determination and secondary structure analysis of extrinsic membrane proteins in the water splitting complex of spinach. FEBS Lett 203:230–234
Von Heijne G (1986a) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683–4690
Von Heijne G (1986b) Towards a comparative anatomy of N-terminal topogenic protein sequences. J Mol Biol 189:239–242
Westhoff P, Jansson C, Klein-Hitpaß L, Berzborn R, Larsson C, Bartlett SG (1985) Intracellular coding sites of polypeptides associated with photosynthetic oxygen evolution of photosystem II. Plant Mol Biol 4:137–146
Westhoff P, Farchaus T, Herrmann RG (1986) The gene for the Mr 10000 phosphoprotein associated with photosystem II is part of the psbB operon of the spinach plastid chromosome. Curr Genet 11:165–169
Young RA, Davis RW (1983) Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA 80:1194–1198

Communicated by C.P. Hollenberg

Received November 25, 1986 / January 12, 1987