The EMBO Journal vol.5 no.1 pp.9-13, 1986

Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework

George A.Karlin-Neumann and Elaine M.Tobin

Biology Department and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

Communicated by M.van Montagu

We have identified three major blocks of amino acid homology shared by the transit peptides of two nuclear-encoded chloroplast proteins, the light-harvesting chlorophyll a/b-protein (LHCP) II of the thylakoid membrane and the small subunit (SSU) of ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) of the stroma. These previously unrecognized homology blocks lie at the beginning, middle and end of both transit sequences, and are separated by differing lengths of unshared (interblock) sequence in the two proteins. These interblocks may be dispensible or they might confer a specific property on the individual proteins, such as facilitating proper compartmentalization within the chloroplast. We propose that these three shared sequence elements form a common framework in transit-bearing chloroplast precursors which mediates the common functions performed by each transit peptide. Ferredoxin, the only other such nuclear-encoded protein for which a published transit sequence exists, conforms to the predictions of this hypothesis. These findings stand in contrast to mitochondrial leader sequences and the well-studied signal peptides of secretory and certain integral membrane proteins in which no such framework has been observed.

Key words: light-harvesting chlorophyll a/b-binding proteins/nuclear-encoded chloroplast proteins/precursor protein import/small subunit of ribulose 1,5-bisphosphate carboxylase-oxygenase/transit peptide homology

Introduction

Numerous chloroplast proteins are encoded by the nuclear genome (Ellis, 1981). These gene products are synthesized on cytosolic ribosomes and then post-translationally imported into the chloroplast (Gilham *et al.*, 1978). Their ultimate functional locations within the chloroplast include the envelope membranes, the stroma, the thylakoid membranes and the lumen (Chua and Schmidt, 1979; Haehnel *et al.*, 1981; Grossman *et al.*, 1982; Ortiz *et al.*, 1985). The structural features of these proteins and the mechanisms which are responsible for their targeting to the proper organelle, their uptake and processing by it, and their proper intra-chloroplast localization are little understood.

Studies have shown that a number of these nuclear-encoded chloroplast proteins are synthesized as larger mol. wt. precursors containing an amino-terminal sequence which is cleaved from the mature protein during or after transport into the chloroplast (reviewed by Chua and Schmidt, 1979; Schmidt *et al.*, 1982; Grossman *et al.*, 1982; Smeekens *et al.*, 1985). Chua and Schmidt (1979) proposed calling this amino-terminal extension a 'transit peptide' both to distinguish it from the more hydrophobic signal peptide which mediates the co-translational transport of secretory protein precursors (von Heijne, 1983; Walter *et al.*, 1984) and to suggest its involvement in post-translational transport of chloroplast precursors. This suggestion is supported by experiments showing that at least two different precursor polypeptides – the small subunit (SSU) of ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and the light-harvesting chlorophyll a/b-protein (LHCP) II – can be synthesized *in vitro* and subsequently taken up by isolated chloroplasts and processed to their final functional forms (Chua and Schmidt, 1978; Schmidt *et al.*, 1982). More recently it has been shown that the SSU transit peptide is necessary and sufficient for the import of proteins into the chloroplast (Mishkind *et al.*, 1985; Van den Broeck *et al.*, 1985).

The best-studied of these extra-chloroplast encoded proteins, SSU, has been sequenced from a number of plant species including monocots (*Lemna* and wheat), dicots (pea, soybean and tobacco) and an alga, *Chlamydomonas*. Comparisons of the directly determined or deduced amino acid sequences have shown all of these to contain highly conserved mature proteins with notably less well-conserved transit sequences. The latter vary in length from 44 amino acids (*Chlamydomonas*) to 57 amino acids (pea, *Lemna* and tobacco). Despite the poorer conservation of SSU transit peptide sequence, substantial regions of amino acid homology and conservation of charged residues have been noted for it among the species compared (Broglie *et al.*, 1983; Stiekema *et al.*, 1983; Mishkind *et al.*, 1985).

Furthermore, the variability in size and amino acid sequence of the SSU transit peptide (see Figure 1b) does not seem to reflect divergence of its functionally important residues or the chloroplast components which recognize and act upon them. *In vitro* reconsitution experiments in which SSU precursor protein (pS) was fed to heterologous isolated chloroplasts have shown that the import and processing of pS to its mature size occurs by mechanisms conserved in the higher plants examined (Chua and Schmidt, 1978; Coruzzi *et al.*, 1983b; Mishkind *et al.*, 1985). Thus, pea or spinach pS can be taken up, cleaved to its mature size and complexed into RuBisCO holoenzyme by isolated chloroplasts from either plant (Chua and Schmidt, 1978). Similar results were obtained for the uptake of pS from monocots (wheat and barley) by isolated dicot chloroplasts (pea and spinach, respectively) (Coruzzi *et al.*, 1983b; Mishkind *et al.*, 1985). These results are complemented by *in vivo* transformation experiments in which pea transit sequences were able to direct an adjoining polypeptide into tobacco or petunia chloroplasts (Van den Broeck *et al.*, 1985; Schreier *et al.*, 1985; Broglie *et al.*, 1984). Thus, at least for the SSU precursor, both the organellar components necessary for its proper uptake, processing and localization within the chloroplast, and the transit peptide with which they interact are functionally well conserved among distantly related species.

Results and Discussion

In this work we have considered whether or not the transit sequences of all known nuclear-encoded chloroplast proteins have structural similarities which can ultimately be related to com-

mon mechanisms of import and processing. It has been generally concluded from the limited data available that this is not the case, although a few similarities have been noted (Cashmore, 1984; Mishkind *et al.*, 1985; Dunsmuir, 1985; Lamppa *et al.*, 1985; Smeekens *et al.*, 1985). We have recently completed the sequences of light-harvesting chlorophyll a/b-proteins from two additional species (*Lemna*, Karlin-Neumann *et al.*, 1985; Kohorn *et al.*, 1986; *Arabidopsis*, Leutwiler *et al.*, in preparation). Analysis of these, in conjunction with the other available data, demonstrates that in fact three distinct blocks of amino acid homology are shared by the LHCP II and SSU transit peptides, and also by that of a third chloroplast protein, ferredoxin.

LHCP II and SSU transit peptides share three major blocks of amino acid homology

Figure 1a and b shows the LHCP II and SSU transit peptides to be well conserved among diverse species. Figure 1a presents a compilation of LHCP II transit sequences as well as a consensus sequence drawn from five plant species, including two monocots (*Lemna* and wheat) and three dicots (*Arabidopsis*, pea and petunia). Where mutliple gene sequences from a single species were available, a consensus sequence is shown, with the exception of *Lemna*. Here, two very dissimilar transit sequences are shown separately. Figure 1b shows a compilation of SSU transit sequences from two monocots, three dicots and an alga, and a consensus sequence drawn from all of these. Amino acids outlined by dotted boxes indicate that a majority of the residues are identical at that position: these positions comprise ~90% of the LHCP II transit sequence and ~75% for that of SSU.

If one compares the most conserved regions (lying at the beginning, middle and end) of the LHCP II and SSU transit sequences, three major homology blocks are found which are shared by these two proteins. These are enclosed in boxes with solid lines and labeled as I, II and III. It is seen that identically numbered boxes of the SSU and LHCP II consensus sequences (Figure 1a and b) each share a series of residues (shaded); these homology blocks appear in the same order in both proteins. Additionally, two other nearly invariant residues, proline (P) and serine (S), are found in similar positions (relative to the adjacent major blocks) in interblock 2 (Figure 1d) of both transit sequences. The shared residues within each block may be separated by one or more non-shared residues, and the internal spacing between shared residues of each block varies slightly between the LHCP II and SSU transit peptides. The amino acid sequence characteristic of each of the three common sequence blocks appears in Figure 1d as part of a larger shared framework.

We propose the term 'framework' to suggest that these common sequence elements, either individually or in concert, are essential for the common events mediated by the transit peptides during import of chloroplast precursors. These may include binding to receptors in the chloroplast envelope (Cline *et al.*, 1985), transport through the envelope and processing to the mature-sized product. Since the mature LCHP II and SSU proteins are situated in different locations within the chloroplast, it would not be expected that these common sequence elements are themselves responsible for final localization of the chloroplast precursors. If the transit peptide also plays a role in this event, its contribution would be expected to lie elsewhere.

In this regard, a significant feature of the framework may lie in the non-shared interblocks (1 and 2, see Figure 1d). The dissimilarity of these regions between the SSU and LHCP II transit peptides might reflect their participiation in a class-specific event, such as compartmentalization within the chloroplast or, alternatively, it might reflect their expendability. If either is involved in any as yet hypothetical class-specific role(s), interblock 2 appears to be the more likely candidate: there are a number of well-conserved residues in each protein's interblock 2, but few of these are shared, and the size of this region differs greatly between the two proteins. Several features of this interblock, however, are common to both proteins and may reflect some common structural requirements even for this region: (i) each bears the conserved 'P' residue an identical distance from block II, and the conserved 'S' residue within several amino acids of block III; (ii) both proteins usually contain another pair of 'S' or 'TS' (T = threonine) residues on the N-terminal side of the shared 'S' residue; and (iii) adjacent to this area the transit peptides of both bear the single acidic residue found along their entire lengths – glutamic acid (E) in LHCP II and aspartic acid (D) in SSU. It is noteworthy that as previously observed for mitochondrial leader sequences and the transit peptide of SSU (Horwich *et al.*, 1984; Broglie *et al.*, 1983), the basic character of the LHCP II transit peptide is due not only to a somewhat greater frequency of basic residues (R or K) than found in the mature protein (1/11 compared with 1/14), but even more so to the strikingly lower frequency of acidic residues in the transit peptide portion (1/34 compared with 1/10). This suggests a general selection against acidic residues in these transit peptides and, perhaps further, that the single conserved acidic residue may be involved in a significant interaction(s) with one or more of the basic residues found in both of these transit peptides.

In contrast to the suggested importance of interblock 2 for the proper functioning of the LHCP II and SSU transit peptides, interblock 1 is wholly absent from LHCP II and appears to be largely unnecessary in SSU. In the latter, beyond the general variability in length and sequence of this region (Figure 1b), 12 of its 18 residues are absent in the wheat SSU consensus sequence. Despite this absence, wheat pS is properly imported and processed by pea chloroplasts (Coruzzi *et al.*, 1983b). There are, however, some well-conserved amino acids at the margins of interblock 1 – notably an invariant alanine (A) adjacent to block I and an invariant valine (V) adjacent to block II. Thus, although several of the marginal residues of interblock 1 may be important, much of the sequence between blocks I and II seems to be generally dispensible for pS uptake and processing by chloroplasts. It will be interesting to see if more concise transit peptides can substitute for the longer native sequences in homologous systems.

Do other nuclear-encoded chloroplast precursors also share the framework?

If our postulation of the functional importance of the framework sequence is correct, we would expect the transit sequences of other nuclear-encoded chloroplast proteins to share these sequence elements. At present, there is only one other such protein for which a published sequence exists and against which this hypothesis can be measured: this is for ferredoxin, a thylakoid-associated stromal-facing protein, from *Silene pratensis* (Smeekens *et al.*, 1985). As can be seen in Figure 1c, each of the three major blocks is present in this 48-amino acid long transit sequence, albeit with different spacing between them than found in either LHCP II or SSU. One or both of the threonines in block I of ferredoxin may be functionally conservative substitutions for serines found in similar positions in the other two proteins. Smeekens and co-workers also note the presence of the 'GRV' triplet of block III in both the ferredoxin and SSU transit peptides and speculate on its involvement in processing. The ferredoxin transit peptide, too, has a higher frequency of basic

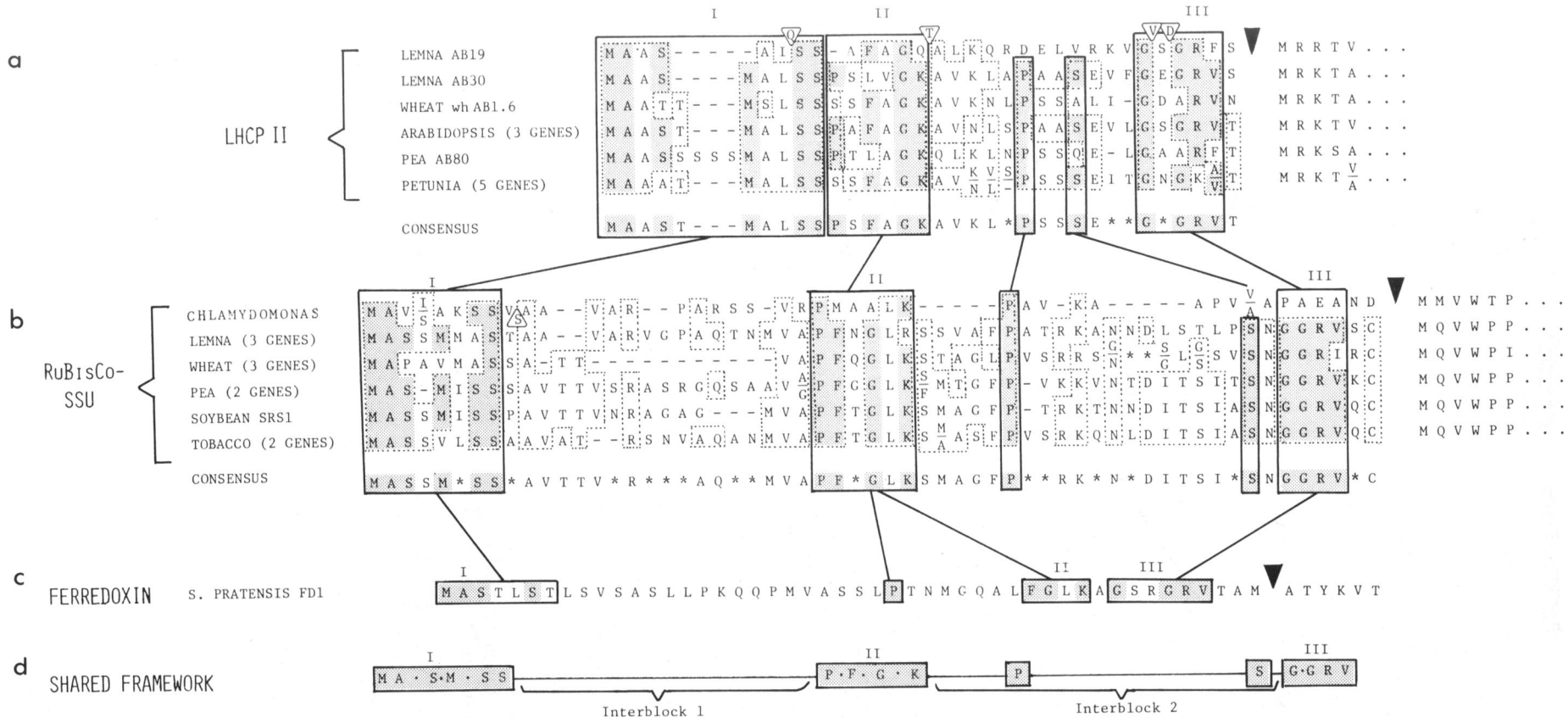

Fig. 1. Comparison of transit peptide sequences. **(a)** LHCP II sequences (*Lemna* AB19, Karlin-Neumann *et al.*, 1985; *Lemna* AB30, Kohorn *et al.*, 1986; wheat, Lamppa *et al.*, 1985; *Arabidopsis*, Leutwiler *et al.*, in preparation; pea, Cashmore, 1984; petunia, Dunsmuir, 1985). **(b)** RuBisCO-SSU sequences (*Chlamydomonas*, Schmidt *et al.*, 1979; *Lemna*, Stiekema *et al.*, 1983; wheat, Broglie *et al.*, 1983; Smith *et al.*, 1983; pea, Coruzzi *et al.*, 1983a; Cashmore, 1983; soybean, Berry-Lowe *et al.*, 1982; tobacco, Pinck *et al.*, 1984; Mazur and Chui, 1985). **(c)** Ferredoxin sequence (*Silene*, Smeekens *et al.*, 1985). Species for which multiple sequences were available are presented as a consensus of these. Dotted lines enclose amino acids shared by a majority of the sequences at the position. Solid lines enclose residues or segments containing residues common to both LHCP II and SSU transit peptides. Shaded residues correspond to those found in the shared framework (see **d**). I, II and III denote the major homology blocks. Residues inside triangles represent additional amino acids present in a single sequence at that position. A dash indicates that no residue is present at the position; an asterisk indicates a non-conserved residue at the position; the arrowheads denote the site (or presumed site) of cleavage separating transit peptide from mature protein. **(d)** Framework sequence shared by LHCP II, SSU and ferredoxin transit peptides. A dot indicates the possible presence of an unshared residue(s) at this position.

residues (1/12) compared with that found in the mature protein (1/20) and an unusually low occurrence of acidic residues (0/48) compared with mature ferredoxin (1/5). It is also interesting that there is no interblock 2 in this transit sequence. Assuming this sequence represents a functional transit peptide, this might reflect an expendability of even interblock 2 (see above). Alternatively, perhaps necessary structure present in this region in both LHCP II and SSU transit peptides may be unnecessary for ferredoxin or permutable to another region(s) of the transit sequence. Although some further similarities can be seen between the ferredoxin and SSU transit sequences (e.g., interblock 1), more complete comparison will have to await the accumulation of ferredoxin transit sequences from multiple species.

Possible functions served by the homology blocks

Possible functional roles for the framework elements are suggested by this and other work. We propose that homology blocks I and/or II are essential and perhaps sufficient for mediating recognition, binding and uptake of precursors into the chloroplast. This is based on the following observations: (i) *Chlamydomonas* pS is both imported and partially processed by isolated spinach and pea chloroplasts, but the absence of either the entire or just the N-terminal portion of this transit peptide through the middle of block II is sufficient to block *in vitro* uptake of the truncated precursor (Mishkind *et al.*, 1985); (ii) the *Chlamydomonas* transit peptide contains homology blocks I and II, and they are largely within the N-terminal portion necessary for import; it lacks most of interblock 2 and homology block III, and as would be predicted (see below), it is not cleaved at the appropriate nearby junction between transit and mature sequences (Mishkind *et al.*, 1985); and (iii) as argued above, interblock 1 appears to be largely or entirely unnecessary for proper import and processing of either SSU or LHCP II precursors. It has been previously proposed that the net positive charge of the SSU transit peptide may be important in binding to the negatively charged outer surface of the chloroplast (Broglie *et al.*, 1983), presumably facilitating recognition and subsequent import of the precursor. The present analysis suggests this basic character may be a general property of transit peptides. However, this is also a property exhibited by the leader sequences of cytoplasmically synthesized mitochondrial precursors (Horwich *et al.*, 1984) and would, therefore, not seem to permit discrimination by itself of chloroplast from mitochondrial precursors.

Homology block II may be important for an intermediate processing event. pS from both an alga and a vascular plant is first cleaved to an intermediate form (iS) during two-step processing to mature SSU, as has been found for several nuclear-encoded mitochondrial precursors (Hay *et al.*, 1984). In the best-characterized of these intermediates, *Chlamydomonas* pS imported by isolated spinach and pea chloroplasts was found by micro-sequencing to be cleaved between the two alanines (A) of homology block II (Mishkind *et al.*, 1985). Intermediate processing was also demonstrated in a homologous system by Robinson and Ellis (1984b) where a partially purified protease from pea chloroplasts generates a discrete intermediate during *in vitro* production of mature-sized SSU from pea pS. The site of this latter cleavage, however, is as yet unknown.

It seems probable that homology block III is important for cleavage at the mature processing site of transit-bearing chloroplast precursors from vascular plants. Although there is some uncertainty about the exact cleavage site for the LHCP II precursor, evidence is consistent with it occurring on one side or the other of the methionine (Cashmore, 1984). In any case, all of the cleavages occur within one to three residues of homology block III in each of the three proteins analyzed here (Figure 1a–c; Cashmore, 1984; Schmidt *et al.*, 1979; Smeekens *et al.*, 1985) and, aside from the presence of methionine (M) adjacent to each of these junctions, homology block III is the only uniformly well-conserved sequence in this region. It is also noteworthy that the only transit sequence lacking homology block III, that of *Chlamydomonas* pS, fails to undergo cleavage at the mature processing site when imported by vascular plant chloroplasts (see above). The 'GGRV' sequence of block III has further structural properties which may argue for its involvement in the neighboring cleavage. This sequence contains several glycine (G) residues which, along with proline, are the strongest disrupters of ordered secondary structure (Chou and Fasman, 1978). It also contains a basic residue, arginine (R). It is thus likely that this region is maintained in a random coil configuration and kept superficial by the positive charge. Geisow and Smyth (1980) have suggested that such characteristics make a region susceptible to proteolytic cleavage. In this context, it is interesting that homology block II, which also contains a basic residue, lysine (K), and several disrupters of secondary structure (P and G), has been shown to be a site of intermediate cleavage in pS from *Chlamydomonas* (see above).

A common operator on transit sequences may already have been identified

The work of Robinson and Ellis (1984a) offers support for the framework hypothesis in another important way. The partially purified pea proteolytic activity which processes the SSU precursor can also process the precursors of plastocyanin from wheat and barley to their mature sizes, but does not affect non-chloroplast proteins. Thus, if a single protease is responsible for cleaving both pS and pre-plastocyanin to their mature forms, it may represent one of the postulated components of a common chloroplast uptake and processing apparatus. It would be predicted, then, that such a protease would recognize similar structures in the two transit peptides. It will be interesting to see if homology block III ('GGRV'), as well as the other major blocks, are found in the plastocyanin transit sequence.

No comparable framework is observed in mitochondrial leader sequences or in signal peptides

Although amino-terminal pre-sequences have been found to be sufficient to direct fusion proteins into mitochondria (Hurt *et al.*, 1984; Horwich *et al.*, 1985) and the endoplasmic reticulum (Lingappa *et al.*, 1984), analyses of both mitochondrial leader sequences (Horwich *et al.*, 1984; Morohashi *et al.*, 1984) and numerous signal sequences (von Heijne, 1983, 1984) have shown only some general structural similarities in either of these. In neither leader nor signal sequences has any evidence been found for a framework of shared amino acids such as we see for the chloroplast transit sequences.

In summary, this analysis supports the proposition that transport of nuclear-encoded precursors to and into the chloroplast, as well as processing to their mature forms, is mediated by common molecular features in their transit peptides and in the cellular components which recognize and act on these signals. Importantly, it identifies common sequence elements and their linear relationships to one another which are shared by all presently characterized chloroplast protein transit peptides. These elements, either individually or in concert, may be responsible for the associated events of precursor recognition by the chloroplast, uptake and cleavage to the mature polypeptide. Fuller testing of this hypothesis will have to await the accumulation of sequence data

for other chloroplast precursors, as well as direct experimental challenge of its predictions. Identification of these conserved blocks will provide a rational basis for dissecting the multiple roles which the transit peptide plays in the uptake and processing of nuclear-encoded chloroplast precursors.

Acknowledgements

We gratefully acknowledge the use of transit sequence data from Bruce Kohorn (*Lemna* AB30) and Leslie Leutwiler (*Arabidopsis*) prior to its publication. Also, we thank Bruce Kohorn, Jane Silverthorne, Leslie Leutwiler and J. Philip Thornber for helpful criticisms of this manuscript and Loraine Kohorn for expert assistance in preparation of the figure. This work was supported by NIH grant GM23167 and USDA grant 83-CRCR-1-1319. G.K.-N. is an ARCO fellow.

References

Berry-Lowe,S.L., McKnight,T.D., Shah,D.M. and Meagher,R. (1982) *J. Mol. Appl. Genet.*, **1**, 493-498.

Broglie,R., Coruzzi,G., Lamppa,G., Keith,B. and Chua,N.-H. (1983) *Biotechnology*, **1**, 55-61.

Broglie,R., Coruzzi,G., Fraley,R.T., Rogers,S.G., Horsch,R.B., Niedermeyer, J.G., Fink,C.L., Flick,J.S. and Chua,N.-H. (1984) *Science*, **224**, 838-843.

Cashmore,A.R. (1983) in Kosuge,T., Meredith,C.P. and Hollaender,A. (eds.) *Genetic Engineering of Plants*, Plenum Press, NY, pp. 29-38.

Cashmore,A.R. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2960-2964.

Chou,P.Y. and Fasman,G.D. (1978) *Annu. Rev. Biochem.*, **47**, 251-276.

Chua,N.-H. and Schmidt,G.W. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 6110-6114.

Chua,N.-H. and Schmidt,G.W. (1979) *J. Cell Biol.*, **81**, 461-483.

Cline,K., Werner-Washburne,M., Lubben,T.H. and Keegstra,K. (1985) *J. Biol. Chem.*, **260**, 3691-3696.

Coruzzi,G., Broglie,R., Cashmore,A. and Chua,N.-H. (1983a) *J. Biol. Chem.*, **258**, 1399-1402.

Coruzzi,G., Broglie,R., Lamppa,G. and Chua,N.-H. (1983b) in Ciferri,O. and Dure,L.,III (eds.), *Structure and Function of Plant Genomes*, Plenum Press, NY, pp. 47-59.

Dunsmuir,P. (1985) *Nucleic Acids Res.*, **13**, 2503-2518.

Ellis,R.J. (1981) *Annu. Rev. Plant Physiol.*, **32**, 111-137.

Geisow,M.J. and Smyth,D.G. (1980) in Freedman,R.B. and Hawkins,H.C. (eds.), *The Enzymology of Post-translational Modification of Proteins*, Academic Press, London, pp. 259-287.

Gilham,N.W., Boynton,J.E. and Chua,N.-H. (1978) *Curr. Top. Bioenerg.*, **8**, 211-260.

Grossman,A.R., Bartlett,S.G., Schmidt,G.W., Mullet,J.E. and Chua,N.-H. (1982) *J. Biol. Chem.*, **257**, 1558-1563.

Haehnel,W., Berzborn,R.J. and Andersson,B. (1981) *Biochim. Biophys. Acta*, **637**, 389-399.

Hay,R., Bohni,P. and Gasser,S. (1984) *Biochim. Biophys. Acta*, **779**, 65-87.

Horwich,A.L., Fenton,W.A., Williams,K.R., Kalousek,F., Kraus,J.P., Doolittle, R.F., Konigsberg,W. and Rosenberg,L.E. (1984) *Science*, **224**, 1068-1074.

Horwich,A.L., Kalousek,F., Mellman,I. and Rosenberg,L.E. (1985) *EMBO J.*, **4**, 1129-1135.

Hurt,E.C., Pesold-Hurt,B. and Schatz,G. (1984) *FEBS Lett.*, **178**, 306-310.

Karlin-Neumann,G.A., Kohorn,B.D., Thornber,J.P. and Tobin,E.M. (1985) *J. Mol. Appl. Genet.*, **3**, 45-61.

Kohorn,B.D., Harel,E., Chitnis,P.R., Thornber,J.P. and Tobin,E.M. (1986) *J. Cell Biol.*, in press.

Lamppa,G.K., Morelli,G. and Chua,N.-H. (1985) *Mol. Cell. Biol.*, **5**, 1370-1378.

Lingappa,V.R., Chaidez,J., Yost,C.S. and Hedgpeth,J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 456-460.

Mazur,B.J. and Chui,C.-F. (1985) *Nucleic Acids Res.*, **13**, 2373-2386.

Mishkind,M.L., Wessler,S.R. and Schmidt,G.W. (1985) *J. Cell Biol.*, **100**, 226-234.

Morohashi,K., Fujii-Kuriyama,Y., Okada,Y., Sogawa,K., Hirose,T. and Inayama,S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4647-4651.

Ortiz,W., Lam,E., Chollar,S., Munt,D. and Malkin,R. (1985) *Plant Physiol.*, **77**, 389-397.

Pinck,M., Guilley,E., Durr,A., Hoff,M., Pinck,L. and Fleck,J. (1984) *Biochimie*, **66**, 539-545.

Robinson,C. and Ellis,R.J. (1984a) *Eur. J. Biochem.*, **142**, 337-342.

Robinson,C. and Ellis,R.J. (1984b) *Eur. J. Biochem.*, **142**, 343-346.

Schmidt,G.W., Devillers-Thiery,A., Desruisseaux,H., Blobel,G. and Chua,N.-H. (1979) *J. Cell Biol.*, **83**, 615-622.

Schmidt,G.W., Bartlett,S.G., Grossman,A.R., Cashmore,A.R. and Chua,N.-H. (1982) *J. Cell Biol.*, **91**, 468-478.

Schreier,P.H., Seftor,E.A., Schell,J. and Bohnert,H.J. (1985) *EMBO J.*, **4**, 25-32.

Smeekens,S., van Binsbergen,J. and Weisbeek,P. (1985) *Nucleic Acids Res.*, **13**, 3179-3194.

Smith,S.M., Bedbrook,J. and Speirs,J. (1983) *Nucleic Acids Res.*, **11**, 8719-8733.

Stiekema,W.J., Wimpee,C.F. and Tobin,E.M. (1983) *Nucleic Acids Res.*, **11**, 8051-8061.

Van den Broeck,G., Timko,M.P., Kausch,A.P., Cashmore,A.R., Van Montagu, M. and Herrera-Estrella,L. (1985) *Nature*, **313**, 358-363.

von Heijne,G. (1983) *Eur. J. Biochem.*, **133**, 17-21.

von Heijne,G. (1984) *EMBO J.*, **3**, 2315-2318.

Walter,P., Gilmore,R. and Blobel,G. (1984) *Cell*, **38**, 5-8.

Note added in proof

After submission of this paper, the sequence of a fourth nuclear-encoded chloroplast protein, plastocyanin, was reported by Smeekens *et al.* (*Nature*, **317**, 456–458). The N-terminal sequence of this transit peptide (MATVTSS) resembles homology block I. Homology block II is present (PSFAGLK) and very nearly identical to those of LHCP II and SSU. It is separated from the first homology block by a short hydrophobic interblock 1 sequence of five amino acids. These authors have also noticed some general similarities in these regions among the same four transit peptides. Homology block III appears less well-conserved in this single example, but is probably present as the sequence GILAGNA adjacent to the inferred processing site in this species. This sequence bears the two proximate glycines (G) characteristic of block III, but seems to have replaced the basic residue, arginine (R), with asparagine (N) and the hydrophobic residue, valine (V), with another hydrophobic residue, alanine (A). In this block, as well as about the adjacent mature processing site, it is most similar in sequence to ferredoxin. Homology blocks II and III are separated by an unusually long interblock 2 (38 amino acids) which may, as these authors also suggest, be instrumental in compartmentalization of plastocyanin within the chloroplast. Finally, it is noteworthy that the plastocyanin transit peptide is also very basic, due primarily to the absence of acidic residues (1/66 in the transit sequence versus 1/7 in the mature protein). The single acidic residue in the plastocyanin transit sequence, glutamic acid (D), is found in its interblock 2, a feature shared by LHCP II and SSU transit peptides.