1993

The presequence of *Euglena* LHCPII, a cytoplasmically synthesized chloroplast protein, contains a functional endoplasmic reticulum-targeting domain

Ram Kishore
University of Nebraska - Lincoln

Umesh S. Muchhal
University of Nebraska - Lincoln

Steven D. Schwartzbach
University of Nebraska - Lincoln

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 11845–11849, December 1993
Cell Biology

The presequence of *Euglena* LHCPII, a cytoplasmically synthesized chloroplast protein, contains a functional endoplasmic reticulum-targeting domain

(microsomal processing/signal peptide/polyprotein)

RAM KISHORE*, UMESH S. MUCHHAL, AND STEVEN D. SCHWARTZBACH†

School of Biological Sciences, University of Nebraska, Lincoln, NE 68588

Communicated by Myron K. Brakke, September 20, 1993 (received for review May 5, 1993)

ABSTRACT **The precursor to the *Euglena* light-harvesting chlorophyll *a*/*b*-binding protein of photosystem II (pLHCPII) is unique; it is a polyprotein, synthesized on membrane-bound ribosomes and transported to the Golgi apparatus prior to chloroplast localization. A cDNA corresponding to the 5′ end of LHCPII mRNA has been isolated and sequenced. The deduced amino acid sequence of this cDNA indicates that *Euglena* pLHCPII contains a 141-amino acid N-terminal extension. The N-terminal extension contains three hydrophobic domains and a potential signal peptidase cleavage site at amino acid 35. Cotranslational processing by canine microsomes removed approximately 35 amino acids from an *in vitro* synthesized 33-kDa pLHCPII composed of a 141-amino acid N-terminal extension and a 180-amino acid partial LHCPII unit truncated at the beginning of the third membrane-spanning hydrophobic domain. Processed pLHCPII was degraded by exogenous protease, indicating that it had not been translocated to the microsomal lumen. Extraction with 0.1 M Na_2CO_3, pH 11.5, did not remove the processed pLHCPII from the microsomal membrane. A stop-transfer membrane anchor sequence appears to anchor the nascent protein within the membrane, preventing translocation into the lumen. Taken together, these results provide biochemical evidence for a functional cleaved signal sequence within the N-terminal extension of a *Euglena* cytoplasmically synthesized chloroplast-localized protein.**

The light-harvesting chlorophyll *a*/*b*-binding proteins of photosystem II (LHCPII) are a group of highly conserved 25- to 28-kDa proteins encoded by a nuclear multigene family composed of 3–20 members, depending upon the organism studied (1). Differences in gene structure and amino acid composition at 14 positions allow LHCPIIs to be classified into three types (2). What if any functional differences exist between LHCPII types remain unclear.

Chloroplast proteins are synthesized as precursors containing an N-terminal extension, the transit peptide, which contains the information required for the posttranslational uptake and proteolytic processing of the precursor by the chloroplast (3, 4). The precursor of LHCPII, pLHCPII, typically contains a 29- to 36-amino acid transit peptide (4). The pLHCPII transit peptide is required for chloroplast localization, while insertion of pLHCPII into the thylakoid membrane appears to utilize targeting signals contained within the mature protein (5–7). *Euglena* pLHCPIIs are unique polyproteins encoded by 6.6- and 9.5-kb mRNAs (8). A *Euglena* pLHCPII genomic clone (GC18) encodes five mature LHCPIIs separated by decapeptide linkers (9, 10). The LHCPIIs can be divided into three groups. There is greater than 90% amino acid sequence identity within a group and 50–70% identity between groups (9). Since GC18 comprises slightly more than half of the gene encoding the 6.6-kb LHCPII mRNA (9), the complete *Euglena* pLHCPII polyprotein probably contains 8 or 9 LHCPIIs.

A number of differences have been found between the synthesis of *Euglena* LHCPII and the synthesis of LHCPII in all other organisms. The light-induced synthesis of *Euglena* pLHCPII is regulated at the level of translational elongation (11) rather than at transcription. The half-life of *Euglena* pLHCPII is 20 min, in contrast to the extremely short half life of pLHCPII in other plants and algae (8, 12). *Euglena* pLHCPII is synthesized on membrane-bound ribosomes as found for proteins transported into the endoplasmic reticulum (ER) rather than on free ribosomes as normally found for cytoplasmically synthesized chloroplast proteins (4, 11). Immunogold electron microscopy localizes *Euglena* LHCPII to the Golgi apparatus and chloroplast when LHCPII is being synthesized, while at times when LHCPII is not being synthesized, an immunoreaction is seen only in the chloroplast (10, 13). It thus appears that *Euglena* pLHCPII is transported into the ER prior to chloroplast localization.

In contrast to post-translational transport of proteins into chloroplasts, transport of proteins into the ER is cotranslational and dependent upon a 15- to 30-amino acid N-terminal extension, the signal sequence (3). The signal sequence is cleaved on the lumenal side of the ER membrane (3, 14). Signal peptides differ in length and primary sequence (3, 15). All identified signal peptides contain a positively charged N-terminal region, a hydrophobic core, and a more polar C-terminal region (3, 15). The signal peptidase cleavage site can be predicted on the basis of the amino acid found at the −1 and −3 positions (15, 16). The precursor to a *Euglena* calcium-binding protein contains a typical eukaryotic three-domain signal peptide at its N terminus (17). This *Euglena* precursor is cotranslationally processed by and translocated into canine microsomes (17), demonstrating that canine microsomes can be used as an *in vitro* system for identifying functional *Euglena* signal peptides.

The precursors to two *Euglena* cytoplasmically synthesized chloroplast-localized proteins, the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (SSU) and porphobilinogen deaminase (PBGD), contain 134- and 139-amino acid N-terminal extensions (18, 19). The N-terminal extensions do not resemble chloroplast transit peptides, but the N-terminal 35–40 amino acids constitute a signal peptide domain with a potential signal peptidase cleavage site

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ER, endoplasmic reticulum; Lac, β-lactamase; LHCPII, light-harvesting chlorophyll *a*/*b*-binding proteins of photosystem II; pLHCPII, precursor to LHCPII (other precursors are named similarly); PBGD, porphobilinogen deaminase; SSU, ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit.

*Present address: Department of Cell Biology and Anatomy, New York Medical College, Valhalla, NY 10595.
†To whom reprint requests should be addressed.

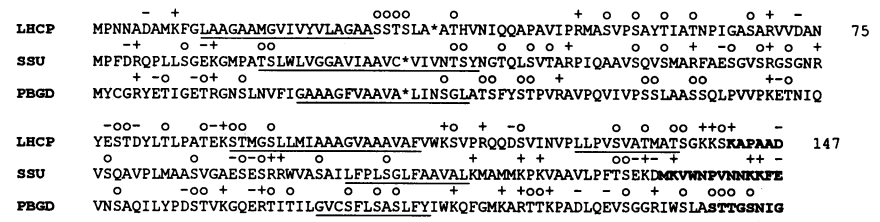

FIG. 1. Comparison of the deduced amino acid sequence of the N-terminal extension of *Euglena* LHCPII (LHCP), *Euglena* SSU (18), and *Euglena* PBGD (19) precursors. Hydrophobic domains (underlined), predicted N-terminal signal peptidase cleavage site (*), hydroxylated amino acids (○), charged amino acids (+, −), and the start of the mature protein (boldface) are indicated.

(18–20). The precursor of PBGD, pPBGD, was not processed by or transported into canine microsomes, suggesting that the signal peptide domain and signal peptidase cleavage site are nonfunctional (20). This paper characterizes the N-terminal extension of *Euglena* pLHCPII.‡ As found for pPBGD and pSSU, pLHCPII contains a 141-amino acid N-terminal extension having a signal-peptide-like domain within the first 35 amino acids. *Euglena* pLHCPII was cotranslationally processed by and inserted into canine microsomes. For the first time that we are aware of, biochemical evidence has been obtained demonstrating that a *Euglena* cytoplasmically synthesized chloroplast-localized protein contains a functional signal peptide. A brief report of this work has appeared (21).

MATERIALS AND METHODS

Construction of Clones Encoding Truncated LHCPII Polyproteins. NLH1, a cDNA clone complementary to the 5′ end of *Euglena* pLHCPII mRNA, was isolated by reverse transcriptase–PCR using two nested 3′ primers and a 5′ primer complementary to the trans-spliced sequence found at the 5′ end of most *Euglena* mRNAs (22). First-strand cDNA was synthesized by using the oligonucleotide 5′-CATGGCCAGGCGGCCGTTCTT-3′, which corresponds to a conserved region of *Euglena* LHCPII (9). Second-stand synthesis and amplification utilized the oligonucleotide 5′-CGAGCTCGAGTGTC$\begin{matrix}A\\T\end{matrix}$TTTTTTTTCG-3′, corresponding to the 5′ consensus trans-spliced sequence (22), and the oligonucleotide 5′-CCGTTCTTCACCTCCTTCACCTT-3′, corresponding to a nested conserved region within *Euglena* LHCPII (9). cDNA was synthesized from 10 μg of total RNA by using Superscript reverse transcriptase (GIBCO/BRL) as described by the supplier. PCR amplification was performed by using *Pfu* DNA polymerase (Stratagene) as described by the supplier for 30 cycles of 1 min at 95°C, 30 sec at 60°C, and 3 min (last cycle 15 min) at 75°C. The PCR product was gel purified and subcloned as a blunt-ended insert into the polylinker of the vector pBluescript II KS (+) (Stratagene), using standard protocols (23). The *Eco*RI–*Xba* I fragment of psAB80XD/4 (a gift of Kenneth Cline, University of Florida, Gainesville) encoding pea pLHCPII (7) was also subcloned in the polylinker of the vector pBluescript II KS (+). The transcript encoding *Escherichia coli* β-lactamase was obtained from Promega. DNA sequences were determined as described previously (9) and analyzed with the University of Wisconsin Genetic Computer Group DNA sequence analysis software (24).

***In Vitro* Processing Experiments.** Capped transcripts were synthesized by using T7 RNA polymerase (Promega) as described by the supplier. Transcripts (0.1 μg/25-μl reaction mixture) were translated for 60 min in a nuclease-treated rabbit reticulocyte lysate (Promega) containing [^{35}S]methionine in the presence or absence of 3.6 equivalents of canine pancreatic microsomal membranes (Promega) as described by the supplier. Translation was terminated by the addition of 0.1 vol of a solution containing 0.12 M methionine and 0.3 mM cycloheximide. Protease protection assays were performed by incubating the translation mix for 1 h on ice with proteinase K at 0.1 mg/ml and 10 mM $CaCl_2$ in the presence or absence of 0.5% Triton X-100 (25). Microsomal membranes were extracted with Na_2CO_3 (26) by incubation for 30 min on ice with 100 mM Na_2CO_3, pH 11.5, and 3 equivalents of microsomal membranes. The Na_2CO_3-extracted membranes were recovered by centrifugation for 15 min at 150,000 × *g*. Microsomal membranes were washed with potassium acetate (26) by adding an equal volume of 1 M potassium acetate and 3 equivalents of microsomal membranes to the translation mixture, layering 20 μl of membranes over a 100-μl cushion of 0.5 M sucrose/0.5 M potassium acetate/2 mM magnesium acetate and recovering the membranes by centrifugation for 10 min at 150,000 × *g*. All centrifugations were performed at 0–4°C. Translation products were immunoprecipitated with anti-LHCPII, proteins were separated on SDS/8–12% polyacrylamide gels, and [^{35}S]methionine-labeled proteins were visualized by fluorography as described previously (8). Densitometric analysis of the fluorographs was performed with a Visage 60 image analyzer (BioImage, Ann Arbor, MI) using whole band analysis software. The amount of protein recovered in the membrane pellet is expressed as a fraction of the total protein recovered in the pellet and supernatant fractions.

RESULTS

A cDNA clone, NLH1, corresponding to the 5′ end of the *Euglena* 6.6-kb LHCPII mRNA was obtained by reverse transcriptase–PCR. NLH1 encodes a 33-kDa protein composed of a 141-amino acid N-terminal extension (Fig. 1) and a 180-amino acid truncated LHCPII unit ending just prior to the third membrane-spanning hydrophobic domain (9). The N-terminal extension contains three hydrophobic domains. Amino acid 35 is the predicted signal peptidase cleavage site of a potential signal peptide (15, 16). The decapeptide linking LHCPII units within the polyprotein precursor has the consensus sequence XMXAXXGXKX (9, 10). The four conserved amino acids are probably processing protease recognition sequences. The N-terminal extension ends in the peptide ATMATSGKKS (Fig. 1). The conserved C-terminal linker sequence (GXKX) is present at the presequence–mature protein junction but the conserved N-terminal linker sequence (XMXA) is missing. This suggests that the C terminus of the decapeptide linker and the N-terminal extension are cleaved by the same processing protease.

Known chloroplast transit peptides are enriched in hydroxylated amino acids and have few acidic amino acids, and their C-terminal domain has a high potential for forming an amphiphilic β-strand (4). The *Euglena* pLHCPII N-terminal extension is enriched in hydroxylated amino acids. However, there are a large number of acidic amino acids (Fig. 1) and an amphiphilic β-strand-forming domain is absent. The overall

‡The sequence from which the amino acid sequence reported in this paper has been deduced has been deposited in the GenBank data base (accession no. U03392).

structure of the N-terminal extension is not indicative of a chloroplast transit peptide.

The N-terminal extensions of *Euglena* pLHCPII, pSSU (18), and pPBGD (19) are significantly longer than the 20-amino acid signal peptide found on the precursor to a *Euglena* ER-targeted calcium-binding protein (17). Alignment at the N-terminal methionine indicates little sequence homology between the three chloroplast presequences, although they do share a number of structural similarities (Fig. 1). Each presequence contains a potential signal peptidase cleavage site 32 (SSU, PBGD) to 35 (LHCPII) amino acids from the N terminus (Fig. 1). The regions flanking the second hydrophobic domain are enriched in charged amino acids. The second hydrophobic domain of the pLHCPII presequence is a potential signal peptide (16) and it is large enough to span the membrane (19 amino acids), suggesting that it could function as a stop-transfer membrane anchor sequence. The corresponding hydrophobic domains in the SSU and PBGD presequences are not sufficiently large to function as membrane-spanning domains (Fig. 1). Another difference is the absence of a third hydrophobic domain in the SSU and PBGD presequences (Fig. 1).

Canine microsomes provide an *in vitro* system for identification of functional signal peptides (25). The small size of the pLHCPII N-terminal extension (14 kDa) relative to the large size of the pLHCPII polyprotein (207 kDa) would make the molecular mass change due to presequence removal difficult to detect on SDS gels. To overcome this problem, *in vitro* processing studies utilized the truncated pLHCPII encoded by the cDNA clone NLH1.

In vitro transcripts were prepared from NLH1 and translated in a nuclease-treated rabbit reticulocyte lysate, and the products were immunoprecipitated with anti-LHCPII to eliminate endogenous reticulocyte translation products. A full-length 33-kDa protein, p-NLH1, is the major NLH1 mRNA translation product synthesized in the absence of microsomes (Fig. 2*A*, lane 1) and when canine microsomes were added after termination of translation (Fig. 2*A*, lane 5). A 30-kDa protein, i-NLH1, is the major NLH1 translation product synthesized when canine microsomes were present during translation (Fig. 2*A*, lane 2). The cotranslational conversion of p-NLH1 to i-NLH1 through removal of an approximately 3-kDa peptide demonstrates the presence of a functional cleaved signal peptide. The procedure of von Heijne (15, 16) predicts a signal peptidase cleavage site 35 amino acids from the N terminus.

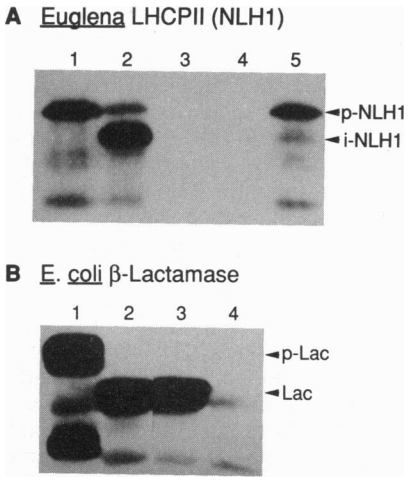

FIG. 2. Microsomal processing of a truncated *Euglena* pLHCPII and *E. coli* β-lactamase. mRNA encoding a truncated *Euglena* pLHCPII transcribed from clone NLH1 (*A*) or *E. coli* pβ-lactamase mRNA (*B*) was translated in the absence (lane 1) or presence (lanes 2–4) of canine microsomes. Translation was stopped and the translation mixtures were incubated with proteinase K (lane 3), proteinase K and Triton X-100 (lane 4), or microsomal membranes (lane 5). Translation products were immunoprecipitated with anti-LHCPII in *A* and in both cases were separated by SDS gel electrophoresis and visualized by fluorography. p-NLH1 and p-Lac, full-length translation products; i-NLH1 and Lac, processed translation products.

Signal peptide removal occurs on the lumenal face of the microsomal membrane (14). The cotranslational conversion of p-NLH1 to i-NLH1 demonstrates insertion of the nascent protein into the microsomal membrane and initiation of translocation into the lumen. In the absence of a membrane anchor stop-transfer sequence (3), i-NLH1 would be translocated into the microsomal lumen and the microsomal membrane would protect i-NLH1 from degradation by exogenous protease (25). To determine whether i-NLH1 was translocated through the microsomal membrane, microsome-containing translation mixtures were incubated with proteinase K at the completion of translation. As expected, the unprocessed translation product, p-NLH1, was completely degraded (Fig. 2*A*, lanes 3 and 4). Unexpectedly, the cotranslationally processed translation product, i-NLH1 was also completely degraded (Fig. 2*A*, lanes 3 and 4). This suggests that translocation into the microsomal lumen was prevented by a membrane anchor stop-transfer sequence, the microsomal membrane preparation was not translocation competent, or the microsomal membranes were not intact and thus were unable to protect the protein from exogenous protease.

The precursor to *E. coli* β-lactamase (p-Lac) is known to be cotranslationally processed by and transported into canine microsomes, providing a positive control for assessing the intactness and translocation competence of the canine microsomal preparation. p-Lac was synthesized *in vitro* and cotranslationally processed to β-lactamase (Lac) by the canine microsomal membrane preparation utilized for p-NLH1 processing (Fig. 2*B*, lanes 1 and 2). Lac was protected from proteinase K degradation in the absence but not in the presence of detergent (Fig. 2*B*, lanes 3 and 4), confirming the intactness and translocation competence of the microsomal membrane preparation.

Since the microsomal membranes were intact and translocation competent, the protease sensitivity of i-NLH1 suggests that it was inserted into but not fully translocated through the microsomal membrane. Depending on the number of start- and stop-transfer sequences, either a large C-terminal tail or multiple loops may be exposed on the outer surface of the membrane (3), resulting in protease-protected fragments which are too small to be resolved on the gels used. Na_2CO_3 extraction converts microsomal vesicles into open membrane sheets, releasing lumenal and peripheral proteins while integral membrane proteins remain associated with the membrane (27). Extraction with 0.5 M potassium acetate releases many but not all peripheral proteins, while lumenal proteins are not released (25). To determine if i-NLH1 is an integral membrane protein, NLH1 mRNA was translated in the presence of canine microsomes, membranes were extracted with potassium acetate or Na_2CO_3, and the extracted membranes were recovered by centrifugation.

Canine microsomes cotranslationally processed p-NLH1 to i-NLH1 (Fig. 3*A*, lanes 1 and 2). All of the p-NLH1 and i-NLH1 were recovered in the membrane pellet after potassium acetate extraction (Fig. 3*A*, lanes 5 and 6). On the other hand, a large amount of the i-NLH1 but only a small amount of the p-NLH1 pelleted with Na_2CO_3-extracted membranes (Fig. 3*A*, lanes 3 and 4). Densitometric analysis of the fluorograph indicated that 70% of the i-NLH1 and only 8% of the p-NLH1 were recovered in the membrane pellet after Na_2CO_3 extraction. The Na_2CO_3-resistant association of i-NLH1 with microsomes suggests that it is an integral membrane protein.

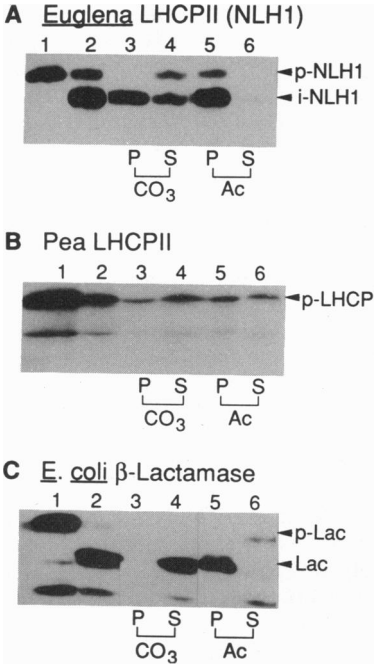

FIG. 3. Microsomal processing and intramicrosomal localization of a truncated *Euglena* pLHCPII, pea pLHCPII, and *E. coli* pβ-lactamase. mRNA encoding a truncated *Euglena* pLHCPII transcribed from clone NLH1 (*A*), pea LHCPII mRNA (*B*), or *E. coli* pβ-lactamase mRNA (*C*) was translated in the presence (lanes 2–6) or absence (lane 1) of canine microsomes. Translation was stopped, and the membranes were incubated with 0.1 M Na_2CO_3 (CO_3) or 0.5 M potassium acetate (Ac) and separated into pellet (P) and supernatant fractions (S). Translation products were immunoprecipitated with anti-LHCPII in A and in all three cases were separated by SDS gel electrophoresis and visualized by fluorography. p-NLH1, pLHCP, and p-Lac, full-length translation products; i-NLH1 and Lac, processed translation products.

The LHCPII portion of p-NLH1 and i-NLH1 is 65% identical to higher plant LHCPIIs and contains two of the three hydrophobic membrane-spanning domains that mediate the post-translational insertion of LHCPII into the thylakoid membrane (5–7, 9). Nonspecific alkali-resistant interactions between LHCPII and a variety of membranes, including microsomal membranes, have been reported (6). The Na_2CO_3-resistant i-NLH1 membrane association could thus be due to a nonspecific interaction between the hydrophobic domains of LHCPII and the Na_2CO_3-disrupted membranes. Alternatively, the Na_2CO_3 extraction conditions used may not fully disrupt the membrane, causing retention or trapping of lumenal and other soluble proteins. As positive controls for the release of nonspecifically bound hydrophobic proteins and trapping of soluble lumenal proteins, the cotranslational association of pea pLHCPII and *E. coli* β-lactamase with canine microsomes was studied.

p-Lac synthesized *in vitro* by the reticulocyte lysate was cotranslationally processed to Lac (Fig. 3*C*, lanes 1 and 2). As expected for a protein translocated to the microsomal lumen, all of the Lac was recovered in the microsomal pellet after potassium acetate extraction (Fig. 3*C*, lanes 5 and 6). Na_2CO_3 extraction released all of the Lac to the membrane-free supernatant (Fig. 3*C*, lanes 3 and 4). The failure to detect Lac in the membrane pellet demonstrates that the microsomal vesicles were fully disrupted and that nonmembrane proteins were not trapped and recovered in the membrane pellet.

The full-length pea p-LHCPII synthesized *in vitro* by the reticulocyte lysate (p-LHCP) was not processed by canine microsomes (Fig. 3*B*, lanes 1 and 2). p-LHCP was equally distributed between the potassium acetate-extracted microsomal membrane pellet and supernatant (Fig. 3*B*, lanes 5 and 6). Na_2CO_3 extraction removed p-LHCP from the microsomal membranes (Fig. 3*B*, lanes 3 and 4). Densitometric analysis of the fluorographs indicates that only 28% of p-LHCP was recovered in the microsomal pellet after Na_2CO_3 extraction. Since studies with Lac failed to find evidence for nonspecific trapping of proteins within the Na_2CO_3-extracted membrane pellet, the Na_2CO_3-resistant association of p-LHCP with microsomes is not an artifact produced by incomplete vesicle disruption or protein trapping. It appears that, as reported by others (6), LHCPII spontaneously forms nonspecific alkali-resistant associations with microsomal membranes. The propensity of LHCPII to form nonspecific alkali-resistant associations with microsomes does not explain the large amount of i-NLH1 (70%) relative to the small amount of p-NLH1 (8%) and pea p-LHCP (28%) recovered with Na_2CO_3-extracted membranes. The differential recovery is most likely due to the fact that i-NLH1 has been inserted and anchored within the membrane while p-NLH1 and pea p-LHCPII form nonspecific hydrophobic associations.

DISCUSSION

Euglena chloroplasts are surrounded by three membranes rather than two as found in higher plants and green algae (28). The origin of this third chloroplast membrane, the perichloroplast membrane, is controversial. It has been postulated that the perichloroplast membrane evolved from the plasma membrane of the original photosynthetic symbiont, that it evolved from a phagocytic vacuole membrane, or that it is part of the ER (28). Regardless of the origin of the perichloroplast membrane, *Euglena* cytoplasmically synthesized chloroplast-localized proteins must contain targeting information for transit through a membrane closely related to ER membranes in addition to information for transit through chloroplast membranes. Studies of a recently characterized *Euglena* ER-targeted protein identified a typical eukaryotic signal peptide at the N terminus (17). This protein was cotranslationally processed and translocated into canine microsomes, demonstrating that ER targeting in *Euglena* is not different from targeting in other organisms (17).

A 33-kDa protein, p-NLH1, composed of the *Euglena* pLHCPII N-terminal extension and a truncated LHCPII, was cotranslationally processed *in vitro* by canine microsomes. Signal peptidase resides on the lumenal side of the microsomal membrane (14). Cotranslational removal of a 3-kDa peptide demonstrates that the *Euglena* pLHCPII N-terminal extension contains a functional signal peptide domain that can insert the protein into the membrane and initiate translocation into the lumen. The functional signal peptidase cleavage site, 35 amino acids from the N terminus, is at the position predicted by the method of von Heinje (15, 16).

The processed protein i-NLH1 was not protected from degradation by exogenous protease, indicating that it was not fully translocated into the lumen. All of the unprocessed p-NLH1, all of the soluble lumen-resident protein β-lactamase, and 72% of the pea p-LHCP containing the three membrane-insertion domains found in *Euglena* LHCPII were released from the microsomal membrane by Na_2CO_3 extraction. On the other hand, only 30% of the i-NLH1 was removed by Na_2CO_3 extraction. Taken together with the cotranslational conversion of p-NLH1 to i-NLH1, the Na_2CO_3-resistant association of i-NLH1 with microsomes suggests that i-NLH1 is inserted into the membrane and anchored within the lipid bilayer by a membrane anchor stop-transfer sequence.

Euglena pLHCPII is synthesized on membrane-bound polysomes (11) and transported to the ER and Golgi appa-

ratus prior to chloroplast localization (10, 13). Two other *Euglena* cytoplasmically synthesized chloroplast proteins, PBGD (20, 29) and SSU (13), are transported directly into the chloroplast. The N-terminal extensions of pPBGD and pSSU show no sequence identity to the pLHCPII extension. All three presequences are, however, extremely large (134–141 amino acids) and contain a signal peptide domain with predicted signal peptidase cleavage sites 32–35 amino acids from the N terminus. The signal peptide domain of the chloroplast presequences is larger than but structurally similar to most eukaryotic signal peptides (3, 16), including the signal peptide of a *Euglena* ER-targeted protein (17).

Attempts to directly demonstrate cotranslational processing and import of pPBGD by canine microsomes were unsuccessful (20). *E. coli* signal peptidase did, however, remove approximately 35 amino acids from pPBGD, suggesting the presence of a functional signal sequence (20). Although the complete 139-amino acid pPBGD presequence is removed in one step during post-translational import into isolated *Euglena* chloroplasts, import is abolished by removal of the first 30 amino acids, which compose the signal peptide domain of the presequence (20). This suggests that import of pPBGD into *Euglena* chloroplasts is initiated by a prokaryotic signal peptide domain which is nonfunctional in a eukaryotic, canine microsome, processing system.

In vitro chloroplast import studies with pPBGD (20) and our studies of microsomal processing of pLHCPII suggest a bipartite structure for the N-terminal extension found on *Euglena* cytoplasmically synthesized chloroplast proteins. A signal peptide domain which may include a signal peptidase cleavage site appears to initiate translocation through the ER-like perichloroplast membrane. The remainder of the presequence appears to function as a chloroplast targeting sequence initiating translocation through the two chloroplast membranes. This bipartite structure is analogous to but functionally the reverse of the bipartite presequence of thylakoid lumen proteins, where the N-terminal domain contains information for translocation through the chloroplast envelope membranes, while the C-terminal domain is a signal-peptide-like targeting signal for transport through the thylakoid membrane into the lumen (4).

In vitro processing studies demonstrated that processed pLHCPII was anchored within the microsomal membrane with one or more segments exposed on the outer, protease-accessible, cytoplasmic membrane face. A 19-amino acid hydrophobic domain is present in the chloroplast targeting portion of the bipartite pLHCPII N-terminal extension and probably functions as a membrane anchor stop-transfer sequence. This stop-transfer membrane anchor domain found in the pLHCPII presequence is not present in the pPBGD and pSSU presequences. The presence of a stop-transfer membrane anchor sequence may explain the initial sorting of *Euglena* pLHCPII to the Golgi apparatus (10, 13), and pPBGD (20, 29) and pSSU (13) directly to the chloroplast. The pLHCPII stop-transfer sequence would prevent translocation of the chloroplast targeting presequence domain through the perichloroplast membrane. The chloroplast-targeting domain would remain on the cytoplasmic side of the perichloroplast membrane, unable to initiate translocation into the chloroplast. Membrane-anchored pLHCPII would enter the default secretory pathway and be transported to the Golgi apparatus and then through an uncharacterized pathway to the chloroplast. The absence of a stop-transfer sequence would permit pPBGD and pSSU to be transported through the perichloroplast membrane. The chloroplast-targeting presequence domain would now be able to initiate import into the chloroplast.

Euglenoids are not the only algal group containing chloroplasts surrounded by more than two membranes. The chloroplasts of dinoflagellates are surrounded by three membranes, while chloroplasts of diatoms, brown algae, chrysophytes, and a number of other algal types are surrounded by four membranes (28). The precursors to the diatom counterpart of LHCPII, the fucoxanthin chlorophyll *a*/*c*-binding proteins, do not contain a typical transit sequence (30, 31). The *Phaeodactylum* fucoxanthin chlorophyll *a*/*c*-binding protein presequences are shorter (31 amino acids) than the *Euglena* presequences but, as found for *Euglena*, they have the three-domain structure characteristic of a signal peptide (30, 31). The *Phaeodactylum* presequence initiates cotranslational import into canine microsomes and is cleaved (31). Cotranslational transport of chloroplast precursor proteins into an ER-like compartment is thus not unique to *Euglena* but appears to be a common mechanism for targeting proteins to chloroplasts in photosynthetic organisms other than higher plants and green algae. *Euglena* is, however, the only organism identified to date as having polyprotein precursors that are transported to the Golgi apparatus prior to chloroplast localization.

This work was supported by National Science Foundation Grants DCB-8817250 and MCB-9118721 and by funds from the University of Nebraska–Lincoln Research Council. U.S.M. was supported by a University of Nebraska–Lincoln Center of Biotechnology Graduate Research Fellowship.

1. Buetow, D. E., Chen, H., Erdos, G. & Li, L. S. H. (1988) *Photosynth. Res.* **18**, 61–97.
2. Jansson, S. & Gustafsson, P. (1990) *Plant Mol. Biol.* **14**, 287–296.
3. von Heijne, G. (1988) *Biochim. Biophys. Acta* **947**, 307–333.
4. Keegstra, K., Olsen, L. J. & Theg, S. M. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 471–501.
5. Auchincloss, A. H., Alexander, A. & Kohorn, B. D. (1992) *J. Biol. Chem.* **267**, 10439–10446.
6. Huang, L., Adam, Z. & Hoffman, N. E. (1992) *Plant Physiol.* **99**, 247–255.
7. Cline, K., Fulsom, D. R. & Viitanen, P. V. (1989) *J. Biol. Chem.* **264**, 14225–14232.
8. Rikin, A. & Schwartzbach, S. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5117–5121.
9. Muchhal, U. S. & Schwartzbach, S. D. (1992) *Plant Mol. Biol.* **18**, 287–299.
10. Schiff, J. A., Schwartzbach, S. D., Osafune, T. & Hase, E. (1991) *J. Photochem. Photobiol. B* **11**, 219–236.
11. Kishore, R. & Schwartzbach, S. D. (1992) *Plant Sci.* **85**, 79–89.
12. Apel, K. & Kloppstech, K. (1980) *Planta* **150**, 426–430.
13. Osafune, T., Schiff, J. A. & Hase, E. (1991) *Exp. Cell Res.* **193**, 320–330.
14. Shelness, G. S., Lin, L. & Nicchitta, C. V. (1993) *J. Biol. Chem.* **268**, 5201–5208.
15. von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105.
16. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4691.
17. Gumpel, N. J. & Smith, A. G. (1992) *Eur. J. Biochem.* **210**, 721–727.
18. Chan, R. L., Keller, M., Canaday, J., Weil, J. H. & Imbault, P. (1990) *EMBO J.* **9**, 333–338.
19. Sharif, A. L., Smith, A. G. & Abell, C. (1989) *Eur. J. Biochem.* **184**, 353–359.
20. Shashidhara, L. S., Lim, S. H., Shackleton, J. B., Robinson, C. & Smith, A. G. (1992) *J. Biol. Chem.* **267**, 12885–12891.
21. Kishore, R., Muchhal, U. S. & Schwartzbach, S. D. (1992) *Mol. Biol. Cell* **3**, 315 (abstr.).
22. Tessier, L. H., Keller, M., Chan, R. L., Fournier, R., Weil, J. H. & Imbault, P. (1991) *EMBO J.* **10**, 2621–2625.
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
24. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
25. Scheele, G. (1983) *Methods Enzymol.* **96**, 94–111.
26. Nothwehr, S. F., Hoeltzli, S. D., Allen, K. L., Lively, M. O. & Gordon, J. I. (1990) *J. Biol. Chem.* **265**, 21796–21803.
27. Fujiki, Y., Hubbard, A. L., Fowler, S. & Lazarow, P. B. (1982) *J. Cell Biol.* **93**, 97–102.
28. Gibbs, S. P. (1981) *Ann. N.Y. Acad. Sci.* **361**, 193–208.
29. Shashidhara, L. S. & Smith, A. G. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 63–67.
30. Grossman, A. R., Manodori, A. & Snyder, D. (1990) *Mol. Gen. Genet.* **224**, 91–100.
31. Bhaya, D. & Grossman, A. (1991) *Mol. Gen. Genet.* **229**, 400–404.