Cell, Vol. 65, 135–143, April 5, 1991, Copyright © 1991 by Cell Press

A Small Chloroplast RNA May Be Required for *Trans*-Splicing in Chlamydomonas reinhardtii

Michel Goldschmidt-Clermont,* Yves Choquet,*† Jacqueline Girard-Bascou,‡ François Michel,§ Michèle Schirmer-Rahire,* and Jean-David Rochaix*
*Departments of Plant Biology and Molecular Biology
University of Geneva
30 Quai E. Ansermet
1211 Genève 4
Switzerland
‡Institut de Biologie Physico-Chimique
13 Rue Pierre et Marie Curie
75005 Paris
France
§CNRS Centre de Génétique Moléculaire
91190 Gif-sur-Yvette
France

Summary

In C. reinhardtii, the mature psaA mRNA is assembled by a process involving *trans*-splicing of three separate transcripts encoded at three widely scattered loci of the chloroplast genome. At least one additional chloroplast locus (*tscA*) is required for *trans*-splicing of exons 1 and 2. We have mapped this gene by transformation of a deletion mutant with a particle gun. The 0.7 kb region of the chloroplast genome that is sufficient to rescue *tscA* function has been subjected to insertion mutagenesis, showing that it does not contain significant open reading frames. We suggest from these experiments that the product of the *tscA* gene may be a small chloroplast RNA that acts in *trans* in the first *trans*-splicing reaction of psaA. A model for the mode of action of this RNA is presented, in which the characteristic structure of group II introns is assembled from three separate transcripts.

Introduction

In Chlamydomonas reinhardtii, the chloroplast gene *psaA* has a very unusual structure: it consists of three exons that are widely scattered on the plastid genome (Figure 1; Kück et al., 1987). They encode one of the two related chlorophyll apoproteins of the photosystem I reaction center. The sequences adjacent to these exons have the conserved elements that are characteristic of group II introns (Michel and Dujon, 1983). However, in this case each of the two corresponding "introns" is split into separate parts.

Mutants that are defective in the assembly of the mature mRNA have been isolated and characterized (Choquet et al., 1988; Goldschmidt-Clermont et al., 1990; Herrin and Schmidt, 1988; Roitgrund and Mets, 1990). They accumulate abnormal transcripts that are thought to represent precursors and intermediates in the pathways of psaA mRNA maturation. Their analysis led us to propose a model where the three exons of psaA are transcribed separately to produce three individual precursors. The precursors are then assembled to produce the mature mRNA in two consecutive steps of *trans*-splicing, starting either with exons 1 and 2 or with exons 2 and 3. The mutants can be grouped in three classes depending on whether they are affected in one, or the other, or both of the *trans*-splicing reactions: mutants in class A fail to assemble exons 2 and 3 (but splice exons 1 and 2), mutants in class C cannot assemble exons 1 and 2 (but splice exons 2 and 3), and mutants in class B are defective in both steps. The genetic analysis of the mutants showed that at least 14 nuclear genes and one chloroplast locus (*tscA*) are required (Goldschmidt-Clermont et al., 1990; Roitgrund and Mets, 1990). These genes probably encode a number of factors that are involved in psaA mRNA maturation.

The chloroplast mutants belong to class C (defective in the splicing of exons 1 and 2) and have large deletions affecting the chloroplast DNA fragments R12 and R16, but leaving the psaA exons intact (Figure 1; Choquet et al., 1988; Roitgrund and Mets, 1990; Goldschmidt-Clermont et al., 1990). These mutants also have an additional phenotype when grown in the dark: they are defective in chlorophyll synthesis and are therefore yellow, unlike the wild type, which is normally green in the dark (Girard-Bascou, 1987; Roitgrund and Mets, 1990). Using particle gun transformation (Boynton et al., 1988), we have shown that the deletions affect two separate genes, *tscA* and *gidA*, whose functions can be restored using wild-type DNA fragments (Y. C., J. G.-B., M. G.-C., and J.-D. R., unpublished data). The defect in *trans*-splicing of the deletion mutants (*tscA*) can be rescued by transformation with the wild-type chloroplast fragment R12. We have used particle gun transformation to define further the *tscA* gene and to identify its product: a small chloroplast RNA.

Results

Intron 1 Accumulates As a Y-Shaped Structure

Whereas conventional splicing (in *cis*) of group II introns releases the intron as a lariat with a 2′-5′ phosphodiester bond at the branch point, *trans*-splicing is expected to produce the intron as a Y-shaped molecule. In the case of *psaA* in C. reinhardtii, the precursor of exon 2 also contains the sequence of the *psbD* gene, which is thus part of the split intron. In our model the splicing of exon 1 to exon 2 produces a Y-shaped molecule with the 5′ part of the first intron now attached to the 3′ part. Such a molecule was indeed detected, but its structure was not established (Choquet et al., 1988).

It is possible to cleave specifically the molecule close to the putative branch point using an oligonucleotide that hybridizes to the 3′ part of intron 1 and RNAase H, which only digests RNA that is part of an RNA–DNA hybrid (Figure 2, bottom left). This treatment releases a small RNA that is only detected using a probe specific for the 5′ part

† Present address: Institut de Biologie Physico-Chimique, 13 Rue Pierre et Marie Curie, 75005 Paris, France.

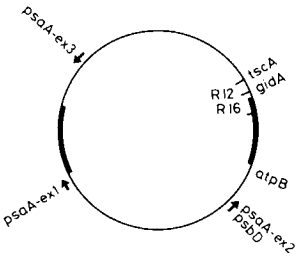

Figure 1. Schematic Map of the C. reinhardtii Chloroplast DNA

The circular genome (190 kb) has inverted repeats indicated by thick bars. The three exons of psaA with arrows indicating the direction of transcription are shown on the outside, together with other loci relevant to this work. The positions of the EcoRI restriction fragments R12 and R16 (Rochaix, 1978) are shown on the inside.

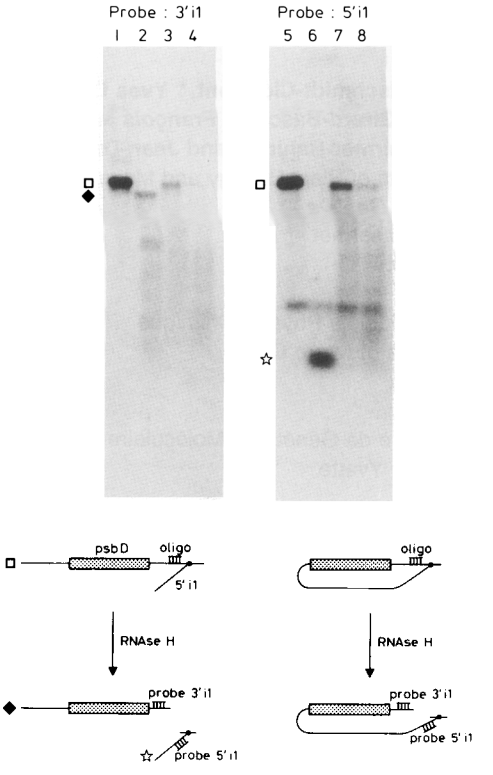

Figure 2. The First Intron Is Released As a Y-Shaped RNA

As shown schematically at the bottom left, total RNA was hybridized with the DNA oligonucleotide 31 (indicated as "oligo" in the diagrams) and then digested with RNAase H. This enzyme is specific for RNA that is hybridized to DNA and thus specifically cuts at the position of the oligonucleotide. The RNA was then subjected to agarose gel electrophoresis, blotted to nitrocellulose, and hybridized with a probe specific for the 3′ part of intron 1 (oligonucleotide 3′i1, lanes 1–4) or with a probe specific for the 5′ part of intron 1 (oligonucleotide 5′i1, lanes 5–8). Symbols refer to the Y-shaped intron (open square) and to the two fragments that are generated by RNAase H cleavage (closed diamond, star), as represented in the schematic diagram at the bottom of the figure on the left. The result that would have been obtained with a lariat intron is shown by the diagram at the bottom right. Lanes 1 and 5, untreated RNA; lanes 2 and 6, complete reaction; lanes 3 and 7, mock-treated control (no RNAase H); lanes 4 and 8, mock-treated control (no oligonucleotide).

of intron 1 (Figure 2, lane 6) and a larger product that only hybridizes to a probe specific for the 3′ part of intron 1 (lane 2). This is in contrast with the untreated molecule that hybridizes to both probes (lanes 1 and 5) and therefore probably has a Y-shaped structure. If the two parts of intron 1 had first been ligated end-to-end and then spliced in *cis*, a lariat would have been generated; RNAase H cleavage would then have given a single RNA still containing both parts of the intron (Figure 2, bottom right). Our results are thus clearly consistent with a mechanism involving splicing in *trans*. We cannot rule out, however, a mechanism involving ligation and *cis*-splicing to produce a lariat, followed by specific processing of the lariat to yield a Y-shaped intron. Nevertheless, in both alternatives the two separate precursors must interact in *trans* and undergo a splicing reaction with the release of a Y-shaped molecule. The interaction probably involves base pairing to form the characteristic secondary structure of group II introns. The difference between the two alternatives is then only whether such a structure is nicked or not.

The nature of the bond at the branch point remains unclear: although by analogy with other group II introns a 2′-5′ phosphodiester bond might be expected, our attempts to release the branch by treatment with debranching extract from HeLa cells were not successful.

Fine Mapping of the *tscA* Gene by Chloroplast Transformation

The chloroplast mutant H13 is deleted for a region of its chloroplast DNA comprising parts of fragments R12 and R16. It lacks two distinct functions, *trans*-splicing of psaA (*tscA*) and chlorophyll synthesis in the dark (*gidA*). Because psaA encodes an essential component of photosystem I, mutants of *tscA* are deficient in photosynthesis and only grow in the presence of a source of reduced carbon (acetate). Chlorophyll synthesis in the dark is dispensable, and *gidA* mutants simply have a yellow phenotype in the dark. Both of these functions can be rescued by transformation using a particle gun to deliver the appropriate DNA fragments to the chloroplast (Y. C., J. G.-B., M. G.-C., and J.-D. R., unpublished data). The phototrophic growth of H13 can be restored by bombardment with a plasmid-containing fragment R12, indicating that this fragment contains all the genetic information of *tscA* that is deleted in H13. Additional sequences from fragment R16 are necessary to rescue *gidA*. To locate precisely *tscA* within R12, we have tested the ability of a number of subclones and deletions of the fragment to restore photoautotrophy to H13 by particle gun transformation. A summary of the most relevant results is shown in Figure 3. We have never observed revertants of H13 in the untreated controls. The presence of the transforming DNA in the phototrophic colonies was verified by Southern analysis (data not shown).

Of the subclones we have analyzed, the smallest that can rescue the phototrophic growth of H13 (albeit at very low frequency) contains the internal KpnI–BglII fragment. We have also tested two series of nested deletions from the two ends of R12 (only those that define the borders of

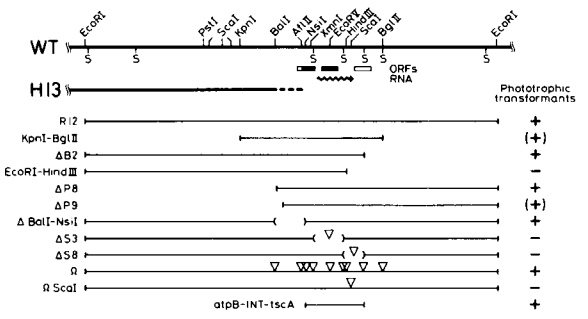

Figure 3. Mapping of the *tscA* Gene by Transformation

A *tscA* deletion mutant (H13) was bombarded with microprojectiles carrying plasmid DNA. Transformed colonies were selected for phototrophic growth on minimal medium. Successful transformation is indicated by a plus sign; lack of transformed colonies is indicated by a minus sign (see Table 1 in Experimental Procedures). The upper line shows a restriction map of the 4.95 kb EcoRI fragment R12 with sites relevant to this work. Bars represent the short *ORF*s discussed in the text (the bars are filled in black downstream of possible start codons; see Figure 6). The lines below them represent the DNA that is retained in the host H13 and in the derivatives of R12 used for transformation (subclones, deletions, and insertions). Triangles on one line indicate the positions of all the individual Ω insertions that were positive, and on a separate line the site of the ScaI insertion that was negative. The last line shows the 720 bp segment that was inserted at the *atpB* locus (see Figure 4).

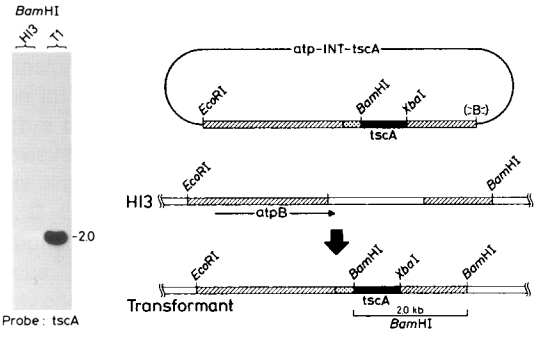

Figure 4. Integration of *tscA* at the *atpB* Locus

The right panel shows a map of the plasmid that was used for transformation with the 720 bp *tscA* fragment (atpB-INT-tscA), a map of the recipient DNA (H13), and of the result expected after integration (for a double cross-over or gene conversion event). In atpB-INT-tscA, the 3′ end of atpB is replaced by the 3′ end of psbD (stippled bar) and a linker with two cloning sites (BamHI and XbaI) used to insert the tscA fragment (black bar). The hatched bars show the regions of homology between the transforming plasmid and the recipient DNA. Note the absence of one of the BamHI sites in the plasmid, the absence of the other BamHI site in the recipient, and the presence of both in the transformant. The left panel shows a Southern blot of BamHI-digested DNA from the host (H13) and a transformant (T1). The expected 2.0 kb fragment hybridizes to the *tscA* probe.

the sequences required to rescue H13 are shown in Figure 3). While deletion ΔB2 (retaining the sequences to the left of residue 1059; see Figure 6) is positive in this assay, the EcoRI–HindIII subclone (extending only to position 848) does not give rise to phototrophic colonies. Thus, the righthand border of the required sequence must lie between the HindIII site (position 848) and the end of ΔB2 (position 1059). On the lefthand side, the last deletion that is clearly positive is ΔP8 (retaining sequences to the right of nucleotide 3), but the next deletion, ΔP9 (which only retains sequences to the right of nucleotide 73) transforms only at very low frequency. However, this also corresponds to the limit of the region that is deleted in H13, and the result could be due to the lack of homology between the transforming DNA and the recipient H13 DNA. If homologous recombination is required for efficient transformation or for marker rescue, then the lack of homology could interfere in the transformation assay. This is consistent with the finding that an internal deletion removing the sequences between the BalI and NsiI sites (residues 4 to 360) is positive: it retains a long region of homology to the recipient DNA, while removing more sequence to the right of residue 73 than ΔP9. Two other internal deletions, ΔS3 and ΔS8 (between the Sau3A sites at positions 454, 808, and 1060) are negative (the deleted DNA is replaced by the Ω marker, see below). Taken together, these results place the *tscA* sequences that are required to restore the function of *tscA* in H13 between residues 360 and 1052. This does not necessarily imply that all of the *tscA* gene is in the interval; part of the gene could be located to the left, in the region that is not deleted in H13. Recombination between the transforming DNA and the recipient would then restore the complete gene, possibly with an intron spanning the region of BalI–NsiI.

To test whether the *tscA* gene is entirely within the interval, the corresponding fragment (720 bp) was prepared using the polymerase chain reaction and inserted into a vector (atpB-INT) capable of integrating at a remote locus (Figure 4). The *tscA* fragment is thus inserted just downstream of the *atpB* gene, in a region that is dispensable for chloroplast function (Blowers et al., 1989; J.-D. R., unpublished data). When this construct is used to transform H13, phototrophic colonies are obtained. Substitution of the resident sequence with the *tscA* construct generates a new BamHI fragment that is observed as expected in the Southern analysis of the transformants (Figure 4). We conclude that all the essential sequences of the *tscA* gene are in the interval 347 to 1056.

The *tscA* Gene Encodes a Small RNA

Using the 720 bp *tscA* fragment as a probe, we have analyzed total RNA from wild-type and from the H13 deletion mutant by Northern blot hybridization. We thus identify a small transcript of 430 ± 20 bases in the wild-type but not in the deletion mutant (Figure 5a, lanes 6 and 7). When *tscA* is inserted at the *atpB* locus, a transcript of identical size is also observed as expected (Figure 5a, lanes 1 and 2). Using a synthetic oligonucleotide complementary to residues 827 to 847 near the HindIII site, the same transcript is again detected (data not shown); we conclude that transcription is from left to right. The 5′ end of the transcript was mapped by S1 nuclease protection using the NsiI–HindIII fragment labeled at the HindIII site as a probe (data not shown): the transcript starts at position 501 ± 3. This result was confirmed by extending a primer hybridized to tscA RNA at the ScaI site (residues 893 to 914) with reverse

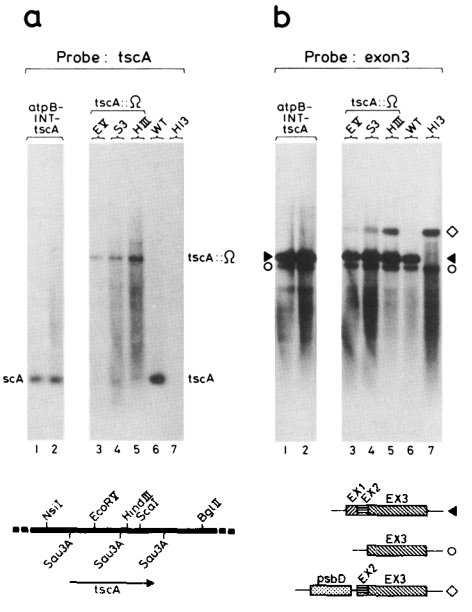

Figure 5. Total RNA was subjected to agarose gel electrophoresis, transferred to nitrocellulose, and hybridized with the 720 bp *tscA* fragment (a) or a probe from psaA exon 3 (b). Lanes 1 and 2, RNA from two independent atpB-INT-tscA transformants; lanes 3–5, RNA from transformants with Ω insertions (tscA::Ω) at the EcoRV (EV), Sau3A (S3), and HindIII (HIII) sites it the *tscA* locus; lane 6, RNA from the wild type; lane 7, RNA from H13, the host used in the transformation experiments (*tscA* deletion mutant). In (a), the positions of the 430 base tscA RNA and of the 2.4 kb *tscA*-Ω transcript are indicated. In (b) symbols show the positions of the mature mRNA (solid arrowhead), the precursor of exon 3 (open circle), and the intermediate with exons 2 and 3 and psbD (open diamond).

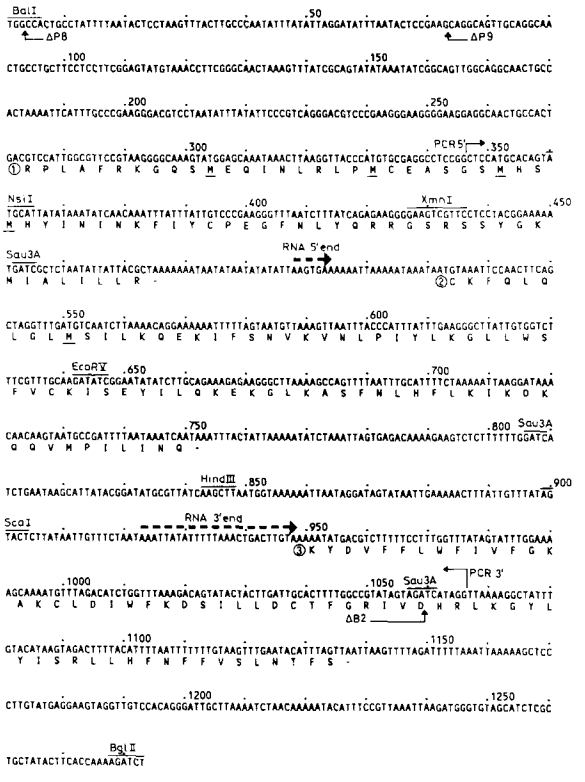

Figure 6. Nucleotide Sequence of the *tscA* Locus

The positions of the 5′ end of the tscA RNA and the estimated position of its 3′ end are shown by thick dashed arrows. Three small *ORF*s are indicated below the sequence (1, 2, and 3). The relevant restriction sites, the end points of deletions P8, P9, and B2, and the ends of the 720 bp *tscA* fragment generated by a polymerase chain reaction (PCR5′, PCR3′) are also shown.

transcriptase; a 410 base fragment was observed, as expected from the S1 protection experiment (data not shown). From the size of tscA RNA (430 ± 20 bases), we expect the 3′ end to map between residues 910 and 950, but S1 mapping experiments to confirm this inference were not successful. The comparison of the S1 protected fragment with the primer extension shows that the RNA is colinear with the DNA at least up to the ScaI site at position 901 (400 bases).

Does the *tscA* gene encode a small polypeptide? There are three small open reading frames (*ORF*s) within the interval defined by transformation (Figure 6). The first two (ORF1 and ORF2) could encode polypeptides of 57 and 67 amino acids starting from methionine codons. The third (ORF3) does not have a start codon (neither AUG nor GUG; Rochaix et al., 1989) but is open for 62 residues. Furthermore, ORF3 is only partly within the 720 bp fragment defined by transformation and is interrupted at residue 39 when *tscA* is moved to the *atpB* locus; ORF3 is thus probably not relevant to *tscA* function. ORF1 is also only partly within the 720 bp fragment, but the internal Met residue at position 351 could allow the translation of 41 residues. Only ORF2 is entirely contained within the tscA RNA as defined by S1 mapping and primer extension; ORF1 and probably most of ORF3 are outside this RNA (Figure 6).

To test the function of the *ORF*s, we disrupted them by insertional mutagenesis using the Ω fragment. This element carries translation stop codons in all three frames at both its ends and a marker than can be selected in Escherichia coli (Prentki and Krisch, 1984). Most Ω insertions in fragment R12 are still positive in the transformation assay and restore the phototrophic growth of H13 (see Figure 3). The positive insertions include those at the NsiI, XmnI, and Sau3A sites (ORF1, positions 360, 435, and 454; see Figure 5), the EcoRV site (ORF2, residue 644), and a Sau3A site (ORF3, position 1060), as well as insertions at a Sau3A site and the HindIII site (positions 808 and 847) and insertions outside the *tscA* interval. The insertion at a ScaI site (position 901) was negative in the transformation assay. It is noteworthy that this insertion is outside any of the ORFs. We suggest that the *tscA* gene probably does not encode a polypeptide, and that its functional product is apparently an RNA.

The question then arises of the effect of the Ω insertions on *tscA* function. As shown above, three insertions within the transcript (at the EcoRV, Sau3A, and HindIII sites) still give rise to photoautotrophic transformants of H13, while only one (at the ScaI site) does not. When RNA from the transformants with Ω at the three sites is analyzed by

Northern blot hybridization, the 430 base tscA RNA is not observed, but is replaced by a 2.5 kb transcript as expected if transcription proceeds across the Ω insert and terminates at the end of the *tscA* gene (see Figure 5a, lanes 3–5). An abnormal but stable RNA thus accumulates in these insertion mutants. Its function in *trans*-splicing is sufficient for photosynthetic growth, but is nevertheless significantly impaired compared with the wild type; in addition to mature psaA RNA, the splicing intermediate with exons 2 and 3 (and psbD) but lacking exon 1 also accumulates in the three transformants (Figure 5b, compare lanes 3–5 and 6 and 7). Thus, the splicing of exons 1 and 2 is hindered when the tscA RNA carries the Ω insertion at three different positions. Because we cannot recover colonies with the insertion at the ScaI site, we do not know whether the tscA transcript in that case is completely inactive or whether it fails to accumulate.

Discussion

Structure of the Transcript Released by Intron 1 *Trans*-Splicing

It is intriguing that exon 2 of psaA (a photosystem I gene) is cotranscribed with *psbD*, a gene encoding one of the core polypeptides of photosystem II. The split intron is thus apparently released as a Y-shaped molecule containing psbD with the 5′ part of intron 1 (sequences initially adjacent to exon 1) attached. It is not known whether this molecule can also function as a messenger RNA for *psbD*. There is as yet no evidence for coordinate regulation of *psbD* and *psaA*, but their expression can clearly be uncoupled: first, *psbD* mRNA can also accumulate independently of psaA exon 2, as is the case in mutants blocked in exon 1–exon 2 splicing (classes B and C). Second, in a mutant that destabilizes transcripts containing *psbD* sequences, exon 2 precursor is *trans*-spliced faster than it is degraded so that psaA mRNA is not affected (Kuchka et al., 1989). One possibility is that the promoter of the *psbD* gene may simply be used as an entry site for the transcription of exon 2.

The Product of the *tscA* Gene: A Small Chloroplast RNA?

The many genetic loci that are required for the assembly of psaA mRNA indicate that there are multiple factors involved in the *trans*-splicing of the separate exons. While some are necessary for both *trans*-splicing reactions (class B), others are specific to one of the steps. The assembly of exons 1 and 2 is particularly interesting in this respect, because it specifically requires at least seven nuclear genes, and also one chloroplast gene, *tscA* (class C). Mutations in this chloroplast gene thus reveal the requirement for a factor acting in *trans* in the *trans*-splicing of exons 1 and 2.

Using the *tscA* deletion mutant H13 as a host and particle gun transformation, we have defined a 720 bp interval that contains all the essential sequences of the *tscA* gene. When it is integrated at a different locus, this segment is sufficient to restore *tscA* function and *trans*-splicing. However, in its usual context, additional sequences to the left of the NsiI site (position 360) are required. This may reflect a requirement for homologous recombination in the transformation process, because the end point of the deletion in the recipient H13 DNA also maps to the region left of the NsiI site. There are additional complications, however, because in H13 the remaining chloroplast DNA has an unusual rearranged structure (Y. C., J. G.-B., M. G.-C., and J.-D. R., unpublished data). A small RNA (approximately 430 bases) is transcribed from the *tscA* locus. Three small *ORFs* are present in this region, but only one is within the *tscA* transcript. We propose that the ORFs are not relevant for *tscA* function for the following reasons:

First, ORF1 is not within the transcript we have detected. The BalI–NsiI deletion mutant is still functional, but retains only the C-terminal part of ORF1 with 8 residues downstream of the remaining ATG. Ω insertions that stop translation at the NsiI, XmnI, and Sau3A sites within ORF1 do not block *tscA* function.

Second, ORF2 is interrupted near its center in the Ω insertion at EcoRV, but with only a minor effect on *trans*-splicing. Insertions at other sites outside the ORFs have a stronger effect (HindIII, Sau3A: see Figure 5b; ScaI: discussed below).

Third, ORF3 does not contain any initiator codons (ATG or GTG), and is probably entirely outside the tscA transcript observed. Only about half of ORF3 is retained in the B2 deletion strain or in the 720 bp *tscA* fragment, both of which are functional for *trans*-splicing. The Ω insertion at the Sau3A site interrupts ORF3 near the middle, but does not prevent the function of *tscA*.

Last, the Ω insertion at the ScaI site, which is outside of the three ORFs, has a much more drastic effect (loss of photosynthetic competence) than the insertions within the ORFs.

We are thus led to propose that *tscA* does not encode a polypeptide, and that its product is a small chloroplast RNA.

The finding that three of the four Ω insertions within the tscA RNA still allow residual *trans*-splicing and photosynthetic growth is somewhat puzzling, because the tscA RNA now contains 2 kb of additional sequence. However, the transformants obtained with the three insertions that are still photosynthetically active display a significant slow down of psaA mRNA maturation (Figure 5b): they accumulate a splicing intermediate that is not detectable in the wild type. The Ω fragment has an inverted repeat of 147 bases at its ends, and the transcript of the insert should thus form a very stable hairpin loop, with the base of the hairpin at the insertion site. The flanking bases would therefore remain in close proximity, and the continuity of the tscA RNA and of its secondary structure would not be grossly altered. If this occurred in a nonessential region, then part of its activity could thus be retained.

Function of the tscA RNA

It will be of interest to determine how tscA RNA functions in *trans*-splicing. One possibility might be that the requirement is indirect, and that the tscA RNA is required for another process that would itself be necessary for *trans*-

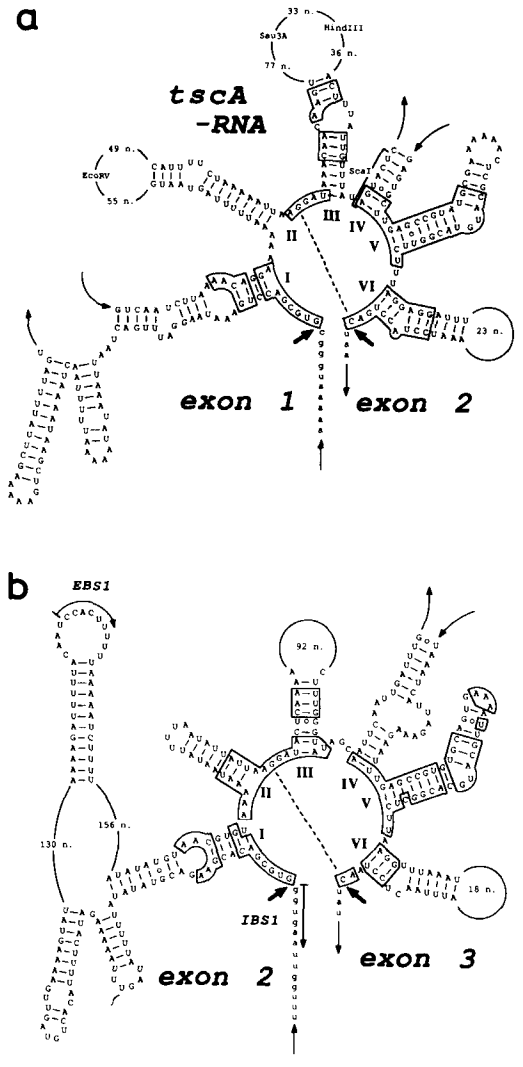

Figure 7. Secondary Structure Models of the Composite psaA Introns

(a) Intron 1. Thick arrows point to intron–exon junctions. Thin arrows indicate 5′ to 3′ strand polarity of the three separate transcripts. Roman numerals correspond to the six major domains of group II introns. Boxed residues are characteristic of subgroup II B1 introns (see Michel et al., 1989). The γ–γ′ tertiary interaction is symbolized by a dashed line (Jacquier and Michel, 1990). EcoRV, Sau3A, HindIII, and ScaI indicate the recognition sequences (6 bases, 4 bases for Sau3A) of these restriction enzymes in the corresponding DNA (see the text). The peripheral structure of domain I could not be unambiguously determined and is not represented.

(b) Intron 2. Blocked arrows indicate the span of the EBS1–IBS1 pairing (Jacquier and Michel, 1987). In this panel, the boxed residues are those that are shared by the two psaA composite introns. Other symbols as in (a).

splicing. Another possibility might be that *tscA* acts as a DNA site on which a tertiary structure is assembled to bring the coding sequences of exon 1 and exon 2 in close proximity; this would partly circumvent the kinetic problem of the bimolecular splicing reaction involving two separate transcripts.

It is perhaps more likely that tscA RNA is directly involved in exon 1–exon 2 splicing and functions as a part of the splicing machinery, maybe in association with some of the nuclear-encoded factors. This would be analogous to the role of snRNAs in the spliceosomes that process nuclear pre-mRNAs (Guthrie and Patterson, 1988). Its specific role in only one of the two *trans*-splicing steps would then imply that there be two subsets of splicing factors for the two "introns" of *psaA*.

More specifically we propose that tscA RNA may be involved in bridging the two parts of the first intron by base pairing to the two separate precursors as shown in the secondary structure model in Figure 7a. The tscA RNA could thus contribute to the formation of the characteristic structure of group II introns and participate in the assembly of the catalytic core. The proposed structure has many evolutionarily conserved residues typical of introns in subgroup II B1 (shown boxed in Figure 7a; Michel et al., 1989). Many secondary and tertiary interactions have been conserved during the evolution of group II introns; they play an essential role for self-splicing in vitro (reviewed by Michel et al., 1989). The helix at the base of domain I is composed of the sequences downstream of exon 1 (5′ part of intron 1) base paired to the 5′ part of the tscA RNA. Domains II and III are contributed entirely by the tscA RNA. There are five highly conserved residues between helices II and III, including the G residue that interacts with the last C residue of intron 5γ in the *oxi3* gene of yeast mitochondria (γ–γ′ pairing; Jacquier and Michel, 1990). The tscA RNA also pairs to the 3′ part of intron 1 to form the helix at the base of domain IV. Domain V, which is very well conserved, and domain VI, with the bulging A residue presumed to be the branch site, are formed by the sequences just upstream of exon 2. The bulging A is 7 bases from the end of the intron, a feature more typical of introns in subgroup IIA. The structure is thus tripartite, formed by the interaction in *trans* of the exon 1 transcript, of tscA RNA, and of the exon 2 transcript. In this model the Ω insertions that have only a partial effect on splicing (EcoRV, Sau3A, HindIII) are at the periphery, while the insertion that abolishes *trans*-splicing (ScaI) disrupts the stem of domain IV.

The revised secondary structure model shown in Figure 7b for the second composite intron also conforms quite well with the consensus. In this case, domains I–III are contributed by the 5′ part of intron 2, domains V and VI by the 3′ part of intron 2, and the two parts are paired to form the helix of domain IV. The well-conserved bulging A at the presumed branch site in the stem of domain VI is not found in its usual secondary structure configuration, but there are other examples of similar deviations (Michel et al., 1989). Thus, while intron 1 is split twice, in domains I and IV, intron 2 is split only once in domain IV. In the secondary structure models, tscA RNA is only required in the *trans*-splicing of intron 1, in agreement with the fact that the *tscA* deletion mutants belong to class C (capable of assembling exons 2 and 3 but not exons 1 and 2). The interruptions in domain IV of the two *psaA* composite introns are reminiscent of in vitro experiments with intron 5γ of the yeast mitochondrial *oxi3* gene (a self-splicing group II intron): Jarrell et al. (1988) have shown that this intron can be *trans*-spliced in vitro when it is interrupted in domain IV. In that case, the helix of domain IV is dispensable and

trans-splicing also occurs when domain IV is deleted. For psaA *trans*-splicing in vivo, the helix of domain IV may play an important role, because the Ω insertion at the ScaI site in *tscA* prevents intron 1 *trans*-splicing. In the secondary structure model for the *trans*-spliced *rps12* intron of Nicotiana tabacum and Marcantia polymorpha, the interruption is in domain III (Ozeki et al., 1987; Kohchi et al., 1988; Michel et al., 1989).

The similarity of the reaction mechanisms for the splicing of group II and of nuclear pre-mRNA introns suggests that the two may be evolutionarily related (Cech, 1986; Jacquier, 1990). The catalytic role of the self-splicing group II intron would have been gradually transferred to separate, *trans*-acting factors (snRNAs and polypeptides) of the spliceosome. In Trypanosome and Caenorhabditis elegans nuclear *trans*-splicing, a short RNA contains the 5′ exon that is added to many different mRNAs. In Trypanosomes, this SL RNA may be part of an snRNP and play a catalytic role in the splicing reaction, circumventing the requirement for the U1 snRNP (Bruzik et al., 1988). This could represent an intermediate stage in the evolution of nuclear pre-mRNA splicing. The role of tscA RNA in psaA *trans*-splicing in C. reinhardtii may also represent an intermediate stage: although still specific to a single group II intron, the tscA RNA is encoded separately.

One evolutionary perspective suggests that psaA represents an ancestral gene structure: the archaic genes may have been split not only with respect to their separate exons, but their introns may also have been separated into several *trans*-acting components. The opposite view would suggest that psaA was originally continuous, but that the tripartite gene arose through DNA rearrangements, perhaps involving transpositions mediated by the reverse transcriptase that may be encoded in some group II introns (Michel and Lang, 1985).

Experimental Procedures

Strains and Media

C. reinhardtii 137c (mt+) and the chloroplast mutants FUD3 (Girard-Bascou, 1987) and H13 (Goldschmidt-Clermont et al., 1990) were previously described. Cells were grown in Tris acetate phosphate (TAP) medium or high salt minimal (HSM) medium and handled as described by Harris (1989).

Oligonucleotides

DNA oligonucleotides used in this work are as follows.

Oligo HindIII: 5′-AGCTTGATAACGCATATCCGT, complementary to bases 827–847 (Figure 6).

Oligo ScaI: 5′-CAATTATAAGAGTACTATAAAC, complementary to bases 893–914.

tscA5: 5′-CAGACTGGATCCGCCTCCATGCACAGTATGC, BamHI linker and bases 347–363.

tscA3: 5′-CAGACTTCTAGACCTATGATCTACTATACGGC, complementary to bases 1047–1066 and XbaI linker.

31: 5′-ATCCTCCTAAAGAACCGTACATGC, 3′ part of intron 1, complementary to bases 130–153 (as numbered by Kück et al., 1987).

3′i1: 5′-TCGCATACGGCTCAAGCATC, 3′ part of intron 1, complementary to bases 101–121.

5′i1: 5′-TCCTTATTTCAGGTCGTCCAG, 5′ part of intron 1, complementary to bases 327–345.

Oligonucleotides 3′i1 and 5′i1 are numbered 312 and 51, respectively, in our collection.

RNA Analysis

RNA preparation by phenol–chloroform–isoamylalcohol extraction and Northern analysis (Goldschmidt-Clermont et al., 1990), S1 nuclease mapping (Erickson et al., 1984), and primer extension (Ausubel et al., 1990) were described previously.

RNAase H Digestions

RNA (6 μg) and oligonucleotide 31 (500 ng) were precipitated together with ethanol and dissolved in 9 μl of H_2O. The solution was then brought to 20 mM HEPES–KOH (pH 7.5), 50 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, and 30 μg/ml bovine serum albumin (BSA) by the addition of 1 μl of 10× concentrated buffer solution. The tube was heated to 65°C for 5 min, allowed to cool to room temperature in 10 min, and incubated at 32°C for 10 min. After the addition of 2 μ of RNAase H, the incubation was continued for 1 hr at 32°C. Control incubations lacking the oligonucleotide or the enzyme were treated in parallel. The RNA was then subjected to gel electrophoresis, Northern blotting, and hybridization with ^{32}P-labeled oligonucleotides as previously described (Choquet et al., 1988).

DNA Constructs

Procedures for the preparation of recombinant DNA plasmids are described by Sambrook et al. (1989). The chloroplast EcoRI fragment R12 was cloned in plasmid Bluescript KS (+) (Stratagene) (pR12-3) and used to derive the RI–HindIII subclone and the deletions B2, P8, and P9 (Y. C., J. G.-B., M. G.-C., and J.-D. R., unpublished data). The KpnI–BglII subclone was obtained in two steps by digestion of pR12-3 with KpnI and religation to give the Kpn–RI derivative, which was in turn digested with BglII and BamHI and religated. The BalI–NsiI deletion was obtained by digestion of pR12-3 with the two enzymes, treatment with Klenow DNA polymerase, and religation. Ω insertions (Prentki and Krisch, 1984) were constructed by linearizing pR12-3 by partial or complete digestion with the appropriate restriction enzyme, treatment with Klenow DNA polymerase to remove protruding ends (when necessary), and blunt-end ligation with SmaI-digested Ω fragment.

The 720 bp tscA fragment was prepared using the polymerase chain reaction with pR12-3 (cut with KpnI and BglII) and the two oligonucleotides tscA5 and tscA3. The reaction contained 0.5 μg of each oligonucleotide, 0.5 μg of DNA, and 2.5 μ of Ampli-Taq DNA polymerase (Perkin-Elmer Cetus) in 50 μl of 50 mM KCl, 10 mM Tris–HCl (pH 8.3), 2 mM $MgCl_2$, 0.2 mM each deoxynucleotide triphosphate, 100 μg/ml BSA, and was subjected to 5 incubation cycles (2.5 min at 94°C, 2 min at 45°C, and 4 min at 72°C). The resulting fragment was inserted using the XbaI and BamHI sites in the primers into the insertion vector atpB-INT (M. G.-C. and J.-D. R., unpublished data).

Transformation

For transformation by particle gun bombardment, cells were grown to a density of 2–4 × 10^6 cells/ml, concentrated by centrifugation, and approximately 10^7 cells in 0.3 ml were plated on TAP medium in 55 mm dishes and incubated 16–48 hr in dim light (200 lux) to obtain a dense confluent layer. The tungsten microprojectiles (approximately 1 μm diameter) were prepared and delivered using a particle gun as described (Zumbrunn et al., 1989). After transformation, the plates were incubated for 4–20 hr in dim light. The cells were then resuspended in 1 ml of HSM medium, plated on HSM medium (directly or with 2.5 ml of 0.5% agar in HSM medium), and incubated under strong light (approximately 300 lux) for 2–3 weeks. Table 1 lists the number of colonies that were obtained. Representative transformants were grown for DNA isolation and Southern hybridization (Rochaix et al., 1988).

Acknowledgments

We thank Prof. U. Kück for kindly providing unpublished sequence data on the 5′ part of intron 1, P. Malnoe, P. Prentki, and U. Schibler for stimulating discussions, and O. Jenni, F. Veuthey, and F. Bujard-Ebener for their help in preparing the figures. This work was supported by grants from the Swiss National Research Fund (31-26345.89) and from the Centre National de la Recherche Scientifique (UA 1187).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *"advertisement"* in accordance with 18 USC Section 1734 solely to indicate this fact.

Received November 5, 1990; revised January 23, 1991.

Table 1. Transformation of *tscA* Deletion Mutants

Plasmid	# of Colonies	Plasmid	# of Colonies
pR12-3	*11, 1, 8, 6, 12, 44* 10, 11, 40, 35, 55	KpnI–BglII	1, 0, 0, 0, 0, 0
Deletion B2	2, 0, 3, 11, 3	RI–HindIII	*0, 0, 0, 0,* 0, 0, 0, 0
Deletion P8	2, 1, 0, 0, 0, 0, 0	Deletion P9	0, 0, 1, 0
Deletion BalI–NsiI	10, 10, 2, 8	Deletion S3	0, 0, 0, 0, 0, 0
Deletion S8	0, 0, 0, 0	Ω BalI	49, 12
Ω AflII	52, 55, 47, 49	Ω NsiI	1, 0, 57, 112, 10, 34
Ω XmnI	131, 8, 65, 78	Ω S13	50, 22, 44, 12
Ω EcoRV	41, 27, 74, 185, 3, 5	Ω S21	47, 16, 9, 74
Ω HindIII	41, 70, 1, 10	Ω ScaI	0, 0, 0, 0, 0, 0, 0, 0
Ω S26	13, 1, 70, 112	Ω BglII	*10, 9*
No DNA	*0 in 10 plates* 0 in 18 plates		

For each DNA plasmid used for transformation of the *tscA* deletion mutant H13, the table lists the number of phototrophic colonies obtained on individual plates. In some initial experiments, FUD3 was used instead of H13 as a host (numbers in italics). The two mutants have very similar defects in their chloroplast DNA (Y. C., unpublished data; Goldschmidt-Clermont et al., 1990). Ω insertions at the Sau3A sites at positions 454, 808, and 1060 are called S13, S21, and S26, respectively. In the deletions S3 (454–808) and S8 (808–1060), the DNA between the Sau3A sites is removed and replaced by the Ω fragment.

References

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1990). Current Protocols in Molecular Biology (New York: J. Wiley & Sons).

Blowers, A. D., Bogorad, L., Shark, K. B., and Sanford, K. C. (1989). Studies on *Chlamydomonas* chloroplast transformation: foreign DNA can be stably maintained in the chromosome. Plant Cell *1*, 123–132.

Boynton, J. E., Gillham, N. W., Harris, E. H., Hosler, J. P., Johnson, A. M., Jones, A. R., Randolph-Anderson, B. L., Robertson, D., Klein, T. M., Shark, K. B., and Sanford, J. C. (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science *240*, 1534–1538.

Bruzik, J.-P., Van Doren, K., Hirsch, D., and Steitz, J. A. (1988). Trans splicing involves a novel form of small nuclear ribonucleoprotein particles. Nature *335*, 559–562.

Cech, T. R. (1986). The generality of self-splicing RNA: relationship to nuclear mRNA splicing. Cell *44*, 207–210.

Choquet, Y., Goldschmidt-Clermont, M., Girard-Bascou, J., Kück, U., Bennoun, P., and Rochaix, J.-D. (1988). Mutant phenotypes support a *trans*-splicing mechanism for the expression of the tripartite *psaA* gene in the C. reinhardtii chloroplast. Cell *52*, 903–913.

Erickson, J. M., Rahire, M., and Rochaix, J. D. (1984). *Chlamyodomonas reinhardtii* gene for the 32000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. EMBO J. *3*, 2753–2762.

Girard-Bascou, J. (1987). Mutations in four chloroplast loci of *Chlamydomonas reinhardtii* affecting the photosystem I reaction center. Curr. Genet. *12*, 483–488.

Goldschmidt-Clermont, M., Girard-Bascou, J., Choquet, Y., and Rochaix, J. D. (1990). *Trans*-splicing mutants of *Chlamydomonas reinhardtii*. Mol. Gen. Genet. *223*, 417–425.

Guthrie, C., and Patterson, B. (1988). Spliceosomal snRNAs. Annu. Rev. Genet. *22*, 387–419.

Harris, E. H. (1989). The *Chlamydomonas* Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. (San Diego: Academic Press).

Herrin, D. L., and Schmidt, G. W. (1988). *Trans*-splicing of transcripts for the chloroplast psaA1 gene. J. Biol. Chem. *263*, 14601–14604.

Jacquier, A. (1990). Self-splicing group II and nuclear pre-mRNA introns: how similar are they? Trends Biochem. *15*, 351–354.

Jacquier, A., and Michel, F. (1987). Multiple exon-binding sites in class II self-splicing introns. Cell *50*, 17–29.

Jacquier, A., and Michel, F. (1990). Base-pairing interactions involving the 5′ and 3′-terminal nucleotides of group II self-splicing introns. J. Mol. Biol. *213*, 437–447.

Jarrell, K. A., Dietrich, R. C., and Perlman, P. S. (1988). Group II intron domain 5 facilitates a *trans*-splicing reaction. Mol. Cell. Biol. *8*, 2361–2366.

Kohchi, T., Umesono, K., Ogura, Y., Komine, Y., Nakahgashi, K., Komano, T., Yamada, Y., Ozeki, H., and Ohyamy, K. (1988). A nicked group II intron and *trans*-splicing in liverwort, *Marchantia polymorpha*, chloroplasts. Nucl. Acids Res. *16*, 10025–10036.

Kuchka, M. R., Goldschmidt-Clermont, M., van Dillewijn, J., and Rochaix, J.-D. (1989). Mutation at the Chlamydomonas nuclear NAC2 locus specifically affects stability of the chloroplast psbD transcript encoding polypeptide D2 of PSII. Cell *58*, 869–876.

Kück, U., Choquet, Y., Schneider, M., Dron, M., and Bennoun, P. (1987). Structural and transcription analysis of two homologous genes for the P700 chlorophyll *a* apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo trans*-splicing. EMBO J. *6*, 2185–2195.

Michel, F., and Dujon, B. (1983). Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. EMBO J. *2*, 33–38.

Michel, F., and Lang, F. (1985). Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. Nature *316*, 641–643.

Michel, F., Umesono, K., and Ozeki, H. (1989). Comparative and functional analysis of group II catalytic introns—a review. Gene *82*, 5–30.

Ozeki, H., Ohyama, K., Inokuchi, H., Fukuzawa, H., Kohchi, T., Sano, T., Nakahigashi, K., and Umesono, K. (1987). Genetic system of chloroplasts. Cold Spring Harbor Symp. Quant. Biol. *52*, 791–804.

Prentki, P., and Krisch, H. (1984). *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene *29*, 303–313.

Rochaix, J. D. (1978). Restriction endonuclease map of the chloroplast DNA of *C. reinhardtii*. J. Mol. Biol. *126*, 597–617.

Rochaix, J. D., Mayfield, S., Goldschmidt-Clermont, M., and Erickson, J. (1988). Molecular biology of *Chlamydomonas*. In Plant Molecular Biology: A Practical Approach, C. H. Shaw, ed. (Oxford: IRL Press, pp. 253–275.

Rochaix, J. D., Kuchka, M., Mayfield, S., Schirmer-Rahire, M., Girard-Bascou, J., and Bennoun, P. (1989). Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast psbC gene product in *Chlamydomonas reinhardtii*. EMBO J. *8*, 1013–1021.

Roitgrund, C., and Mets, L. J. (1990). Localization of two novel chloroplast genome functions: trans-splicing of RNA and protochlorophyllide reduction. Curr. Genet. *17*, 147–153.

Sambrook, K., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Zumbrunn, G., Schneider, M., and Rochaix, J. D. (1989). A simple particle gun for DNA-mediated cell transformation. Technique *1*, 204–216.

GenBank Accession Number

The accession number for the sequence reported in this paper is M61198.