A novel multifunctional factor involved in trans-splicing of chloroplast introns in Chlamydomonas

Livia Merendino, Karl Perron, Michèle Rahire, Isabelle Howald, Jean-David Rochaix and Michel Goldschmidt-Clermont*

Department of Molecular Biology and Department of Plant Biology, University of Geneva 30, quai E. Ansermet, CH-1211 Geneva 4, Switzerland

Received October 31, 2005; Revised and Accepted December 16, 2005

ABSTRACT

In the chloroplast of Chlamydomonas reinhardtii, psaA mRNA is spliced in trans from three separate precursors which assemble to form two group II introns. A fourth transcript, tscA, completes the structure of the first intron. Of the fourteen nucleus-encoded factors involved in psaA splicing, only two are required for splicing of both introns. We cloned and characterized the first of these more general factors, Raa1. Consistently with its role in psaA splicing, Raa1 is imported in the chloroplast where it is found in a membrane fraction and is part of a large ribonucleoprotein complex. One mutant, raa1-L137H, is defective for splicing of both introns, but another allelic mutant, raa1-314B, still expresses the 3′ part of the Raa1 gene and is deficient only in splicing of intron 2. This observation and a deletion analysis indicate the presence of two domains in Raa1. The C-terminal domain is necessary and sufficient for processing of tscA RNA and splicing of the first intron, while the central domain is essential for splicing of the second intron. The combination of these two functional domains in Raa1 suggests that this new factor may coordinate trans-splicing of the two introns to improve the efficiency of psaA maturation.

INTRODUCTION

Group II introns are typically characterized by small elements of conserved primary sequence and a common secondary structure involving a central core and six radiating domains with helical stems. They also share many conserved tertiary structure interactions among the domains, as well as between the domains and the exons [reviewed in (1,2)]. Group II introns are prevalent in the organelle genomes of chloroplasts and mitochondria, and are also found in bacteria and archaea [reviewed in (3)]. These introns are thought to be evolutionarily related to the introns in nuclear pre-mRNAs, where catalysis is mediated by the ribonucleoprotein complexes of the spliceosome, and proceeds by a similar mechanism involving two trans-esterification reactions and a lariat intermediate (4).

In the green alga Chlamydomonas reinhardtii, the psaA gene is composed of three exons which are at distant positions in the chloroplast genome and are transcribed separately (5,6). The precursor transcripts are spliced in trans to generate the mature psaA mRNA (Figure 1D). Flanking the exons, the sequences which constitute the split introns have characteristic features of group II introns. The second intron is assembled from the precursors of exon 2 and of exon 3. However the first intron is actually composed of three transcripts: the precursors of exon 1 and of exon 2, and a small non-coding RNA, transcribed from a separate locus, tscA (7). This intron in three pieces can be seen as an intermediate between the usual introns of group II, containing their own catalytic sequences, and the introns of nuclear pre-mRNAs, where part of the structure and catalytic activity have been proposed to reside in the trans-acting RNAs of the snRNPs.

While some group II introns are well-studied ribozymes that can self-splice in vitro, others have been refractory to self-splicing in the test tube. Both types require accessory protein factors for efficient splicing in vivo. Because the catalytic activity is carried by the RNA, these factors may function as RNA chaperones, may assist proper folding of the intron, or may stabilize it in the active conformation (8).

*To whom correspondence should be addressed. Tel: +41 22 379 6188; Fax: +41 22 379 6868; Email: michel.goldschmidt-clermont@molbio.unige.ch

Present addresses:
Livia Merendino, Laboratoire Plastes et Différenciation Cellulaire, UMR 5575, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France
Karl Perron, Department of Genetics and Microbiology, CMU, 9 avenue de Champel, CH-1211 Geneva, Switzerland

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org
In *Chlamydomonas*, a set of at least fourteen nuclear loci are essential for trans-splicing the two split group II introns of *psaA* (9). Three of these factors have been characterized more in detail, Raa3, Raa2 and Rat1. Raa3 is necessary for trans-splicing of the first intron [class C, (10)]. This novel protein is part of a large complex which also contains the precursor of exon 1 and the *tscA* RNA. Raa2 is essential for trans-splicing of the second intron [class A, (11)]. Although Raa2 is related to pseudouridine synthases, this enzymatic activity is not required for its function in trans-splicing. Rat1 is required for processing of *tscA* RNA from a larger precursor and for splicing of the first intron (12). Rat1 shows sequence similarity to the NAD+ binding domain of poly(ADP-ribose) polymerase, but this domain can be mutated without affecting the function of Rat1 in splicing. These factors are highly specific since they are each required for splicing of only one of the introns. However, two nuclear loci encode factors that are required for splicing of both introns, and may thus have a more general role in splicing of group II introns [class B, (9)]. Here we report the characterization of the first of these factors, Raa1, which is required for trans-splicing of the two introns of *psaA*. We map two distinct domains of Raa1, one involved in processing of *tscA* and splicing of intron 1, the other in splicing of intron 2. Our results show that Raa1 is a multifunctional protein required, as a component of a large membrane-bound and RNA-containing complex, for the splicing of these group II introns.

**MATERIALS AND METHODS**

**Strains and media**

Procedures for growing *C. reinhardtii* and media were described (13) (TAP: Tris Acetate Phosphate; HSM: High Salt Minimal). The strain *arg2 cwd* also has a plastome mutation conferring spectinomycin resistance and was obtained from Dr R. Loppes (University of Liège, Belgium). Mutant L137H was obtained by mutagenesis with 5-fluoro-deoxyuridine and UV, as described previously (6,9).

**Oligonucleotides**

See Supplementary Table I.

**Generation of the 314B mutant strain**

*arg2 cwd* cells were transformed with plasmid pARG7.8 (14) and selected on TAP medium without arginine in the dark. Screening of ca. $4 \times 10^5$ colonies yielded six strains with the fluorescence induction kinetics of mutants deficient in photosystem I or the b6f complex and a *psaA* RNA profile characteristic of trans-splicing mutants (9). Mutant 314B was chosen for further study because analysis of fourteen progenies from a back-cross to the wild-type showed co-segregation of the pARG7.8 insertion with the trans-splicing mutant phenotype, suggesting that the mutation was tagged.

**Transformation of Chlamydomonas**

For transformation of *raa1-314B*, $3 \times 10^7$ cells were treated with 1 μg of the appropriate plasmid DNA by the glass bead method (15) and spread on a single agar plate (HSM or TAP plus 10 μg/ml paromomycin).

---

**Figure 1.** Characterization of *raa1-314B* and *raa1-L137H*. (A) Immunoblot analysis of total cell extracts from the wild-type (WT), the mutants *raa1-314B* and *raa1-L137H*, and the mutants rescued with the genomic DNA (cosmid #2), or the midigene construct. Top, PsaA; bottom, small subunit of Rubisco. (B) Immunoblot analysis with anti-Raa1 antibodies (affinity purified) of total cell extracts from the WT and the mutants *raa1-L137H* and *raa1-314B*. The arrow labels a non-specific band. (C) RNA blot analysis with probes (indicated on the left) specific for *psaA* exons 1, 2, 3, the 5' part of intron 2 (5'i2), *tscA* RNA or *psbA* transcript as a control. Symbols correspond to the transcripts on the splicing scheme (panel D), except the white square which labels a non-specific hybridization signal. (D) Trans-splicing of *psaA* mRNA. The precursors of exon1 (black square, 0.4 kb), exon 2 (black circle, 3.6 kb) and exon 3 (open circle, 2.4 kb) are spliced in trans to produce intermediates (exon1-exon2, asterisk, 2.6 kb or exon2-exon3, diamond, 3.8 kb) and mature *psaA* mRNA (black triangle, 2.8 kb). Splicing of the first split intron requires *tscA* RNA (black star, 0.4 kb) which is processed from longer precursors (open stars) containing chlN.
The raal-L137H strain was transformed by electroporation [modified from (16)]. A total of $10^6$ cells of the mutant were transformed with 1 µg of the appropriate plasmid DNA in presence of 50 µg of salmon sperm DNA and spread on a single plate as above. The electric pulse was applied in the absence of a shunt resistor with 700 V and 25 µF.

RNA blot hybridization

RNA blot analysis of psaA transcripts and the hybridization probes were described previously (11). The tscA probe was a 720 bp PCR fragment (7).

Cloning of the Raa1 gene

For cloning of Raa1, a genomic library of BamHI-digested DNA fragments from raal-314B was constructed in bacteriophage lambda EMBL4 (17). The library was screened by hybridization with radiolabelled pBR329, the plasmid vector of pARG7.8 (14). Restriction mapping of one of the positive phage (λP23) and DNA blot hybridizations with the pARG7.8 probe, or with total genomic DNA from C.reinhardtii as a probe, allowed the identification of a 1.8 kb HindIII fragment that contained neither parts of the pARG7.8 plasmid, nor repetitive sequences. This fragment was used to screen a cosmid library of wild-type C.reinhardtii DNA (18). Three positive cosmids were obtained and used for transformation of the raal-314B mutant and selection on minimal medium (HSM). All three were capable of rescuing phototrophic growth. Cosmid #2 was used for restriction mapping and subcloning. A minimal region was identified spanning the ClaI to ScaI restriction mapping and subcloning, and was found to contain the wild-type phenotype. This cosmid (30 kb insert) was used for restriction mapping and subcloning, and was found to contain the Raa1 gene.

Complementation analysis in vegetative diploids

Vegetative diploids were obtained by fusion of either raal-L137H or raal-314B with raal-314B. To allow selection of diploids, paromomycin-resistant derivatives of raal-L137H and raal-314B strains were obtained by transformation with the aphVIII gene [pSI103, (20)]. The raal-314B strain has a chloroplast mutation conferring resistance to spectinomycin. Removal of the cell wall was achieved by incubating loopfuls of each strain with autolysin in dim light for 1 h at 25°C. Vegetative fusion was induced by spreading the cells on selective plates in fusion solution [83 mM PEG6000, 20 mM CaCl2, 20 mM glycine and NaOH (pH 8)] or a control solution [20 mM CaCl2, 20 mM glycine and NaOH (pH 8)]. Diploids were selected in dim light on TAP plates containing paromomycin (10 µg/ml) and spectinomycin (120 µg/ml). The genotype of raal-L137H / raal-314B diploids was verified by PCR amplification of the aphVIII gene (characteristic of raal-L137H) with oligos paro2-5’ and paro2-3’ and of the pARG 7.8 plasmid (characteristic of raal-314B) with oligos pBRBam2 and ArgSau2. The genotype of raal-314B/ raal-314B diploids was verified by amplification of aphVIII and of Raa2 [which is present in raal-314B but partly deleted in raal-A18, (11)] with oligos A18/rbcS and A18/Nci1.

Immunoblot analysis

Protein samples were analyzed on 5% polyacrylamide gels reinforced with 0.15% linear poly(acrylamide-co-acryl acid) (Aldrich). Proteins were transferred to nitrocellulose for 16 h at 25 V. Raal rabbit antibodies were obtained by immunization with a mixture of three histidine-tagged polypeptides corresponding to residues 1009–1349, 1510–1829 and 1830–2103 of Raa1. They were affinity purified using the three antigens immobilized on Ni NTA. The monoclonal HA-11 antibody was obtained from Eurogentec.

Raa1 genomic and midigene constructs

The subclone gAhI (116b) was obtained by inserting the 12 kb HindIII fragment from cosmid #2 into Bluescript KS (+) digested with HindIII. Subclone gAsc was obtained by inserting the 11 kb EcoRI-ScaI fragment from 116b into Bluescript KS (+) digested with EcoRI and SmaI. Subclone gAsa was derived from 116b by partial digestion with Sall and religation.

The cDNA clone C11 (Supplementary Figure 1) was isolated by screening a library constructed in bacteriophage lambda gt10 (21) with the 0.8 kb Sall–BstXI genomic fragment containing exons 11 and 12. The cDNA clone 405 (Supplementary Figure 1) was found by screening a library constructed in lambda NM1149 (H. Sommer, personal communication) with the 1.5 kb Sall–SpH1 fragment from C11 (exons 6–14). The cDNA clone 551 (Supplementary Figure 1) was obtained from a dedicated cDNA library. This library was prepared by priming synthesis of the first strand of cDNA with oligos i2 and m3 and cloning the double-stranded cDNA with EcoRI-NotI adapters (Amersham-Pharmacia) into lambda gt10. This library was screened with a 0.68 kb EcoRI–BglIII fragment from 405 (exons 4–6). All cDNA inserts were excised with EcoRI and cloned into Bluescript KS (+). The open reading frame (ORF) that the cDNA clones defined encodes a predicted polypeptide of 2103 amino acid. It is preceded by stop codons in every frame, implying that the entire coding sequence has been obtained.

The HA-tagged midigene, mWT (also called pl2), was constructed as follows. Subclone B04 was derived by cloning the 9 kb BscI (ClaI) fragment from gAhI into Bluescript KS (+) digested with BscI. A 3.4 kb Sall fragment from cDNA C11 was cloned into Sall-digested B04, to yield plasmid 269 (the Sall site used for this fusion is in exon 6). The C11 cDNA (recloned in pET15b, Novagen) was tagged with a His6-Myc-BstErI-Kanr sequence introduced in the SmaI site in exon 6 (5’-ctcatggeggtcaccacaccacacactggaagctatcggaggagggcgggttaac-Kanr-gtaacctggtc-3’). The Kan’ cassette was derived from pUC4K, Amersham-Pharmacia). This tagged cDNA sequence was then transferred as a Sall–Stul fragment into plasmid 269, to yield plasmid 349. The triple HA epitope fragment (22) was then inserted as a NruI–NaeI fragment into 349 digested with BstEII to replace the Kan’ cassette and obtain plasmid 490. Finally the 3.5 kb ScaI fragment of 490 (one ScaI site in the polylinker, the other in exon 3) was replaced with the 7 kb fragment from 116b to provide the HA-tagged midigene pl2.
The *aphVIII* gene, conferring paromomycin resistance (20) was amplified by PCR from plasmid pSL17 (S. Lemaire and J. D. Rochaix, unpublished data). The construct mΔBp was derived from pl2 by digestion with BglII (at the intron 6 and exon 7 junction of Raa1) and KpnI (in the polylinker) and ligation to the *aphVIII* gene (amplified with Paro Bgl 5′ and Paro 3′ and digested with BglII and KpnI). The construct mΔkp was derived from pl2 by digestion with KpnI (in exon 16 of Raa1 and in the poly linker) and ligation to *aphVIII* (amplified with Paro Kpn 5′ and Paro 3′ and digested with KpnI). The construct mΔAM was derived from pl2 by digestion with AgeI, blunting with mung bean nuclease, digestion with MscI (in exon 2 of Raa1) and religation. In a second step, the *aphVIII* gene was amplified with Paro Mlu 5′ and Paro Mlu 3′, digested with MluI and inserted in the MluI site, upstream of *Raa1*. The clone pl4 was similarly derived from pl2 by inserting the *aphVIII* gene in the MluI site. The construct mANs was derived from pl4 by digestion with NruI and Smal (in exons 2 and 4 of *Raa1*, respectively) and religation.

**RT–PCR analysis of *Raa1* transcripts**

Total RNAs were isolated as described (23) and then treated with RQ1 DNase (Promega). Reverse transcription was performed at 50°C with M-MLV reverse transcriptase (Promega). To analyze three different regions (exons 1 and 2, exons 12 and 13 and exons 17 and 18) of the Raa1 transcript, the corresponding first strand cDNA was synthesized with oligos AUG2, middle 4 or oligo dT, respectively. PCR amplification was performed with the following pairs of oligos: AUG5'-1 and AUG3'-1; m5'-3 and m3'-3; STOP5'-6 and STOP3'-6, respectively. The PCR products span two exons so that amplification of the cDNA can be distinguished from amplification of any contaminating genomic DNA which would include the intervening intron. As control, RT–PCR analysis was also performed for *Cbl* transcripts (24), using oligo dT to synthesize the first strand cDNA and oligos UP and DOWN for the PCR. The Arg7–Raa1 fusion transcript in *raa1-314B* was revealed by synthesizing the first strand cDNA with oligo middle 4 (in exon 14 of Raa1) and PCR amplification with oligos Arg7ex11 and m3'-3 (in exon 13 of Raa1).

**Sequence analysis of *Raa1***

The sequence of the *Raa1* cDNA is deposited in the EMBL database under accession no. AJ605114. The genomic sequence is now available (scaffold 33) in the draft of the *C.reinhardtii* genome, at the Joint Genome Institute of the US Department of Energy (http://genome.jgi-psf.org/chlr2/chlr2.home.html). Comparison of the cDNA and genomic sequences shows that the *Raa1* gene is composed of 18 exons with very different sizes (from 60 to 2473 bp). The polyA tail is preceded 11 bp upstream by the TGTAA consensus signal typical of *C.reinhardtii* (25).

Transit peptide sequences were predicted using the algorithms ChloroP (26) (http://www.cbs.dtu.dk/services/ChloroP/) and Predotar (27) (version 0.5; http://www.inra.fr/predotar/). Protein secondary structure was predicted with GORIV (28) (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)

**Cell fractionation**

Cells were lysed with a nebulizer and chloroplasts were purified on Percoll gradients as described previously (10) in the presence of protease inhibitor mix A (PIM A: 1 mM e-caproic acid, 0.5 mM benzamidine, 0.6 µg/ml leupeptin, 2 µM E64, 2 µM pepstatin and 0.5 mM AEBSF). Chloroplasts were resuspended in hypotonic buffer [10 mM HEPES (pH 7.8), 5 mM MgCl2, 1 mM DTT] supplemented with protease inhibitor mix B (PIM B: 1 mM e-caproic acid, 0.5 mM benzamidine, 0.6 µg/ml leupeptin, 2 mM ortho-phenanitrole and 1/200 Protease Inhibitor Cocktail P8849 from Sigma). For further fractionation, the chloroplasts (250 µl, 7 mg protein/ml) were lysed by sonication, the solution was adjusted to 20 mM HEPES (pH 7.8), 50 mM KCl and 10 mM MgCl2 (HepKM) and subjected to centrifugation at 100 000 g for 30 min at 4°C in a TLA45 rotor (Beckman). The supernatant was recovered for analysis, and the pellet was washed by resuspension in HepKMS (0.2 M sucrose, HepKM, PIM B) and centrifugation at 50 000 g for 15 min at 4°C in a TLA45 rotor. The supernatant was discarded and the pellet was resuspended in HepKMK plus PIM B.

For analysis of the protein complexes, cells were collected by centrifugation, resuspended at a density of 2.5 10⁸/ml in HepKM plus PIM A, and lysed by freezing and thawing. The lysate (1 ml aliquot) was subjected to centrifugation at 18 000 g for 20 min at 4°C in a TLA45 rotor (Beckman). The pellet containing the membranes was resuspended in HepKM, 0.2 M (NH₄)₂SO₄, PIM B. Where needed, 100 µl of RNaseA were added [10 mg/ml in 100 mM Tris–HCl (pH 7.5) and 10 mM Na acetate] or the same amount of buffer without RNase (mock treatment), followed by incubation at 16°C for 20 min. Membranes and insoluble material were removed by two rounds of centrifugation at 120 000 g for 20 min in a TLA45 rotor, and the supernatant was loaded on sucrose gradients (10 ml; 0.1–1.3 M sucrose, HepKM, PIM A). After centrifugation at 115 000 g for 17 h in the SW40 rotor (Beckman), 10 fractions (1.1 ml) were collected from the bottom by puncturing the tube. Sedimentation was calibrated with the High Molecular Weight Calibration Kit from Amersham-Pharmacia.

**RESULTS**

**Characterization of the *raa1-314B* mutant deficient in splicing of *psaA* intron 2 and cloning of the *Raa1* gene**

Insertional mutants were generated by transformation of a *Chlamydomonas* arginine auxotroph with a plasmid containing the wild-type Arg7 gene. Mutant *raa1-314B* was unable to grow phototrophically (data not shown) and had no detectable PsaA protein (Figure 1A, lane 2). RNA blot hybridization (Figure 1C) showed that the pattern of transcripts that accumulate in *raa1-314B* is characteristic of a block in trans-splicing of the second intron of *psaA* [class A, (6)]. Probes specific for each of the three exons hybridize to mature *psaA* mRNA (black triangle, see Figure 1D) in the wild-type (Figure 1C, lane 1) but not in the mutant (lane 2). In the *raa1-314B* mutant, the probe for exon 3 only detects an unspliced precursor of exon 3 (open circle), which is also present in the wild-type. In the mutant, probes for exon 1
or for exon 2 reveal a splicing intermediate (asterisk) with exons 1 and 2 spliced together, which migrates slightly faster than mature psaA mRNA. This intermediate, with the 5′ part of intron 2 still attached to exon 2 (5′/2 probe, lane 9), is not detected in the wild-type (lane 8). The precursor of exon 1 (black square) over accumulates in raal-314B, as was observed previously for all other psaA splicing mutants (6).

Because of these defects in splicing of psaA transcripts, we have called the mutant raal-314B (RNA maturation of psaA).

Using the Arg7 insertion as a molecular tag, we cloned the raal mutant locus. We then obtained wild-type RAA1 cosmids that could rescue the raal-314B mutant by transformation. These results also showed that the raal-314B mutation is recessive to the wild-type. Additionally, Southern blot analysis of raal-314B DNA indicated that the locus was rearranged (data not shown), confirming that the locus that we cloned was indeed RAA1. In the strains transformed with cosmid #2, PsaA protein is recovered (Figure 1A, lanes 7 and 8, 314B, as was observed previously for all other psaA splicing mutants (6).

To map the gene within the 34 kb insert in cosmid #2, we tested subclones for their ability to restore phototrophic growth of raal-314B by transformation and identified a minimal region of 7.4 kb. We obtained cDNA clones from this region which encode a predicted polypeptide of 2103 amino acid (See Materials and Methods). To ascertain that the appropriate cDNA had been cloned, a ‘midigene’ was constructed with the genomic sequence containing the 5′ part of the gene (including the promoter) fused to the cDNA from the 3′ part (Supplementary Figure 1D). This construct transformed the raal-314B mutant to prototrophic growth as efficiently as the parental mutant (Figure 1C, lanes 2 and 4, black triangle). The splicing intermediate containing the 5′ part of intron 2 is not detected any more (Figure 1C, lane 11, asterisk) and the levels of exon 1 precursor are much lower than in the parental mutant (Figure 1C, lanes 2 and 4, black square).

To test the ability of these two mutants as raal-314B and raal-L137H, we constructed vegetative diploids to test whether the two mutants could genetically complement (Supplementary Data). The observed lack of complementation, together with the fact that transformation with the Raal gene can rescue psaA splicing in both mutants, indicate that raal-314B and L137H are allelic. Furthermore full-length Raa1 protein was not detectable by immunoblotting in the two mutant strains, raal-314B and L137H (Figure 1B, lanes 2 and 3). We thus refer to the two mutants as raal-314B and raal-L137H.

The raal-L137H strain is also defective in maturation of tscA RNA

HN31, another mutant defective in trans-splicing of both introns of psaA (class B), is allelic to L137H (9). Interestingly, HN31 was previously shown to be defective in the maturation of tscA (29), the small non-coding RNA required for splicing of intron 1 (Figure 1D). Mature tscA RNA could not be detected in raal-L137H (Figure 1C, lane 10, black star). In the mutant, several larger transcripts were more prevalent (white stars). These precursor transcripts also contain sequences of chlN, which is downstream of tscA (29). However, normal levels of mature tscA RNA were recovered in the raal-L137H strains transformed with the Raal cosmids or the ‘midigene’ (Figure 1C, lanes 12 and 14). Mature tscA RNA was observed in raal-314B (lane 9), as expected for a strain competent in splicing of intron 1.

A truncated form of Raal is expressed as a fusion to Arg7 in raal-314B

Our data indicate that the mutants are allelic, but their splicing defects are different: in raal-L137H there is a block in maturation of the tscA RNA and in splicing of both psaA introns, while in raal-314B only splicing of intron 2 is defective. This could be explained if Raal is partly functional in raal-314B, and more severely impaired in raal-L137H. To further investigate the basis for the different phenotypes, we analysed the expression of Raal transcripts by RT–PCR (Figure 2).

We designed primers to amplify a 5′ region (exons 1 and 2), an intermediate region (exons 12 and 13) and a 3′ region (exons 17 and 18) of Raal mRNA (Figure 2A).

In the wild-type strain the Raal transcript could be detected with the three primer pairs (Figure 2B, lane 2). In raal-L137H (lane 3) the three parts of the transcript were also amplified,
indicating the presence of the Raa1 mRNA. However, the full-length Raa1 protein was not detected by immunoblot analysis (Figure 1B, lane 2).

In raa1-314B (Figure 2B, lane 4) the intermediate and 3' regions of the Raa1 transcript could be amplified. However, no amplification of the 5' region (exons 1 and 2) was detected. This qualitative analysis suggested that the 5' part of the gene was altered but that a truncated form of Raa1 might be expressed. To investigate this further, we sequenced a genomic fragment from the raa1-314B mutant (LP23, Supplementary Figure 1A) and found that the Arg7 gene and part of the pBR329 vector were joined to intron 9 of the Raa1 gene (Figure 2C). Indeed, the presence of an Arg7–Raa1 fusion transcript was revealed by RT–PCR with a 3' primer in exon 11 of Arg7 and a 5' primer in exon 13 of Raa1. Its sequence showed that the fused mRNA is spliced using a cryptic 5' donor site within the last exon of Arg7 (exon 13) and the 3' acceptor site for exon 10 of Raa1 (Figure 2C). The exons are fused in-frame, such that a polypeptide is predicted with all but the last 14 residues of the Arg7 polypeptide (i.e. arginino-succinate lyase) fused to the 644 C-terminal residues of Raa1. Arg7 is a nucleus-encoded plastid protein (30), so the fusion protein is also predicted to be targeted to the chloroplast. It thus appears that in raa1-314B the C-terminal domain of Raa1 is sufficient for processing of tscA RNA and splicing of intron 1, but not for splicing of intron 2. The predicted fusion protein of 116 kDa was not detected by immunoblotting probably because of low expression or high turnover, or because it was masked by non-specific bands.

Different domains of Raa1 are required for its functions in psaA trans-splicing

Since the C-terminal domain of Raa1 expressed in raa1-314B is sufficient for processing of tscA RNA and splicing of intron 1, we sought to determine which parts of the protein are required for these functions or for splicing of intron 2. We conducted an initial deletion analysis using genomic fragments with 3' truncations of Raa1 (Figure 3C, construct names beginning with g for genomic). We transformed the
Figure 3. Deletion analysis of Raa1. (A) Immunoblot analysis with anti-HA antibodies of total cell extracts from transformants of raa1-L137H (lanes 1–10) or raa1-314B (lanes 11–16) with the midigene constructs depicted in panel C. (B) RNA blot hybridization of total RNA from the wild-type (lanes 1 and 12), the parental mutants (lanes 2, 3, 13 and 14) and the transformants of raa1-L137H (lanes 4–11) or raa1-314B (lanes 15–20). Symbols are as in Figure 1. (C) Schematic representation of the Raa1 polypeptide predicted for the wild-type or for the deletion derivatives (open bars, see Supplementary Materials and Methods). Names shown on the left begin with ‘g’ for genomic constructs, with ‘m’ for the midigene tagged with the HA epitope (white triangle). The part of Raa1 that is fused to Arg7 in raa1-314B is also represented as an open bar. Black lines at the bottom represent the domains involved in splicing of intron 1 or intron 2. In the columns marked ‘phototrophy’, growth of the two mutants on minimal medium is indicated by ‘+’, lack of growth by ‘−’. Splicing of intron 1 (i1) or of intron 2 (i2) is indicated by ‘+’, lack of splicing by ‘−’. As deduced from the RNA blot analysis in panel B. (nd: not determined, na: not applicable). The position of the restriction sites used for the deletions are shown at the top (they truncate the predicted polypeptide sequence after the following amino acid residues: AgeI, 85; NruI, 300; MscI, 731; Smal, 944; SalI, 1136; BglII, 1173; ScaI, 1549; HindIII, 1785 and KpnI, 1853). Black triangles represent the 38 amino acid repeats (not drawn to scale).
**A**

MRRHPTCGIAATPOAQYVGAVFPRLPVAPVPRAPSGQLPPRPCRRAVLCAARRPAGTGTRASGD
AGGGAGEGEGATAAADSHGTSURTPTGLEGRASGVSAPVAGIGGAGSSDGDASAQSRG
PRRGRPGRPMQMTAVANPNTATPAPSSASAVQVQCAANSSPSGSGDEGSGLA
TRRRGPRRSTADASSGSGCISGGSTRAGTCASADYEAAASAAAPMPGMPLASGREGDCQP
QPQSCQPLQQQTEAGGSTATAPGGGSGGSGSGSTQDRGSGRAGGAGGFLRSRQRS
RAMGGVGAAPNATMPPGAPPGDSSLQAFQSMFSEPAAERSDSTEDEGAVUGRSGGAT
SRRQTPPPPDLQHPAAPQQCQCPPDPAAGAASAGSMATDDMLGLSDEDGASLAGLAF
GSPGCGASSSQDFIIIAALDLVGVSTAPPLSDDLSDDWGDTCAGGGCGWDASTGSA
GAMGALDGTVGVLGATGSITIGIWGDGASASAAVAGDASAALQTPSQAALALSDWGF
QDDEAPQPQOPQPQPQPQAPQAQPQALQTPQALQPQGQPQPQVPVLLKGAPAECSPR
PVPAGGGGGGAEATQSGSFRMDSDPHGGLDSDFDLVSSYDAEMSYPDADVCGLLPFEGVAAA
AAAGAGAPGGGGWSEGVLSDMAADALPDLSYSGGSGGSGGGGSGSGSGGSGGSGSGGSGGGSGSGG
AVAGAACRPCVASSSTPIAAPPFPATAPAAPAPAAPPAAKAAGAAPAATGAAKGQAQAAS
GSSGGSSGSSGSSDGSSRTAFAEALLGAGGGSHAGGRADPLVQRARVLQVESLGLAPE
EAAAAADAAAELPGVSLPLFLLPRVLPAPQAQLLEFAEPATARLAAAAAEEAAGGLAAP
PRQVEAPPASDEWALLGGAAPASAASIASAASAAASACTSARGFYGCGYETDTTGDG
SSWRREREREQGDEADFTERRLEAGPAYSAYLRAVVLGLGSGSGLGSGSGGRGGCG
VQAGAVQAALALLEQVEAAAGDWLQLCWLLEAAAQQGVALGFRAAAAFQKAIALVPQAPALP
RLAATTATTATGGLAAEAASYQLCDRARGLAAYGAVLGSVVSAASTHTAARAGGRGSTR
RGVPPITAGGRGGGAGSVAASGSDVAGSQCIGYMGALTICPCQLYGAVLVSASQL
AALTRRRAASLAATLAAAALRSAAVPTSDSARGGGGDGVRRVSDEIEMAAAAS
SAVQPNHASELPPQQQLQQQQQTGSPGFLSGAAATKASALTAAAGVWTAADLIAALAW
GVNAYAGSAAVAAASAAATADCATASQAARKQQRQORQQQORPPRPKHLQRLQQTQP
QPLPPPLPVGGGDWLRSVLAGSEYMATECGGGGGGCGCPPQWMLGWSREPRLYAGPSQ
WMLALLSSRAEAACLQSFQDETCGCRLLWIALAAMDYVERPLMLRVRAGQLQRARRADFTPDQV
TLHCALARLYGAPRPFECVALHAARAILPMPLMTQFALAAHLAAHASAPAWRPGFLVA
ASATGGAAAPVPSASSPAAVQMFAQALMALATRLVLMQRATTEEAAAVASPAGGAA
NAQQQQAPQTALAAAADVGSGLSMALWALLLRQQQCOQQOQPAADVQLMLPEWLAAWAA
AEAEPAVAATFATCVCQSLWALAELEAPGPLHGSAAALAAASGASAATTYADARA
GQQQQAAQAAGAVAILALLALVFVPQGQAATADLSTTIALADLQYFSPQWMFALTAEARB
RLCIAATATATCTGTTAATTATATNEDHGLIAYGLAHLGWPLSEAVQWELAAGGYRAMAGAS
1860
GEGIALALLWLGSASROWGTSASTGQREFDLFRESGSKWDSCPGPAVGLLYCAVAADMMPPQOE
1920
P1PWWQQQLLVKALRLLQVRPRTRALLPAALRTAGGCGLPLTRAERQQLQWAGAAAAAR
1980
AAGGGGGGIGLVLVAGEAYTAVPSGSLAGAVVWALQPCHRPPSPSLPGKGWAPQPPCPG
2040
SgaEDVQDPEEVGLASRVLPLCELCEEFPHPAALAAAVAEGRSWWAHDVAAELRARWGLV
2100
RWS

**B**

Raa1

| 1423 | LWLANSFARLGYAPSAQPENMLALLSREAEACLQSFQDTEG |
| 1461 | LCRILNLAAAMDYVPERMLRAVAGQLQQARADFTPDQ |
| 1499 | VTTYCALARLYGAPRPFECVALHAARAILPMPLMTQF |
| 1772 | LSTTALADLQYFSPQWMFALTAEARBRLGTAATA |
| 1826 | HGLIAAYGLAHLGWPLSEAVQWELAAGGYRAMAGAS |

**Raa1 consensus**

L--L-wALA-LGY-PS--W--AL-A-A-R-----T---
* * * * * * * * *
L----R-----RFSP-HL----L--L--LG-RPPPEW--A

**Tbc2**

- TYNALINAYAK-G--EAA--LY--M--G--FN- PPR

---AY--LG--Y-----YE-A---Y...KAL-LNPNN

helix A  helix B
two raa1 mutants and selected for phototrophy on minimal medium, an indication of whether psaA splicing is restored. However, in this approach, when phototrophy is not rescued, for instance because only one of the two introns is spliced, transformants are not recovered and the effect of the mutations on the excision of the individual introns cannot be analysed in detail. Therefore, we also constructed deletions of the Raa1 midigene carrying an aphIII cassette, allowing selection on paromomycin (20) and included the HA epitope tag (Figure 3C, construct names beginning with m for midigene). Many of the transformants which were initially positive in (Figure 3C, construct names beginning with m for midigene). Likewise only one line stably expressing Raa1 could be recovered for mraa1-L137H in the transformant and for RNA analysis (Figure 3B). Some lines were particularly prone to silencing: only low levels of HA-tagged Raa1 could be detected in the transformant raal-L137H mΔNS #4, and none in raal-L137H mΔNS #7 (Figure 3A, lanes 8 and 9). Likewise only one line stably expressing Raa1 could be recovered for mΔAM in raal-L137H (mΔAM #3, lane 10) or for mΔNS in raal-314B (mΔNS #1, lane 14), and none for mΔBg in raal-314B.

In raal-314B, constructs mΔKp, gΔHi and gΔSc rescued phototrophy (Figure 3C). Indeed, spliced psaA mRNA was detected in the transformants mΔKp #4 and #7 (Figure 3B, black triangle, lanes 16 and 17). These results define a domain upstream of Scal which is sufficient for splicing of 2. This part of Raa1 and the domain expressed as an Arg7–Raa1 fusion in raal-314B are thus capable of intragenic complementation. However in raal-314B, mΔBg and gΔSa did not rescue phototrophy (Figure 3C), indicating that a region between the BglII and Scal sites is essential for splicing of 2. In the raal-L137H mutant, the shortest 3’ deletion, mΔKp, failed to rescue phototrophy (Figure 3C). In these raal-L137H mΔKp transformants, no psaA mRNA was detected (Figure 3B, exon 3, black triangle, lanes 7 and 8), but the intermediate with exon 2 spliced to exon 3 accumulated (exon 3, diamond). This confirms that mΔKp constructs can still support intron 2 splicing and also implies that the deletions at the C-terminal end of Raa1 affect splicing of 1 (Figure 3C). In raal-L137H, mΔBg and gΔSa did not rescue phototrophy (Figure 3C), and in mΔBg transformants, the pattern was identical to the raal-L137H mutant (Figure 3B, lanes 5 and 6 versus lane 2). Thus the domain required for intron 2 splicing is affected in this deletion, confirming the analysis of the raal-314B transformants.

To analyse the N-proximal part of the protein, we constructed internal deletions that leave the putative transit peptide intact. The mΔAM deletion rescued phototrophy in both mutants (Figure 3C), and restored splicing of both introns to produce mature psaA mRNA (Figure 3B, exon 3, black triangle, lanes 11, 19 and 20). Thus a large domain in the N-terminal part of Raa1 is not essential for splicing of either intron. Transformation of raal-314B with mΔNS does not rescue phototrophy or splicing of intron 2: the mature psaA mRNA is not made (Figure 3B, exon 3, black triangle, lane 18) and the exon1–exon2 intermediate is still accumulated (Figure 3B, 5’12, asterisk, lane 18). Comparison of the mΔNS and mΔAM raal-314B transformants indicate that part of the MscI to Scal segment is necessary for intron 2 splicing. In the raal-L137H mΔNS transformant #4, the intermediate with exon 2 spliced to exon 1 (Figure 3B, 5’12, asterisk, lane 9 versus lane 2) indicates that splicing of intron 1 is restored, in spite of the low levels of Raa1 protein expressed in this strain (Figure 3A, lane 8). This confirms that the C-terminal domain of Raa1 is sufficient for splicing of intron 1. Only a very faint band is observed for raal-L137H mΔNS #7 (Figure 3B, 5’12, asterisk, lane 10) which does not detectably express Raa1 (Figure 3A, lane 9).

In this analysis, the domain of Raa1 required for splicing of intron 1 is not distinguishable from the domain required for tscA processing: in the mΔKp and mΔBg transformants of raal-L137H where intron 1 is not spliced, mature tscA fails to accumulate (Figure 3B, tscA, black star, lanes 5–8). Conversely intron 1 is spliced and mature tscA RNA is formed in the raal-314B mutant, or in the mΔNS and mΔAM transformants of raal-L137H (Figure 3B, tscA, black star, lanes 3, 9 and 11).

Taken together, these results delineate two functional domains of Raa1 (Figure 3C). A central region involved in splicing of intron 2 extends from Scal to somewhere between BglII and Scal, with a possible additional contribution from the N-terminal region upstream of AgeI. A C-terminal region involved in splicing of intron 1 and tscA maturation starts at the breakpoint of the Arg7–Raa1 fusion in raal-314B (Figure 3C, last open bar at the bottom) and ends somewhere between KpnI and the terminus. The two domains may partly overlap in a short region to the left of the Scal site.

**Raa1, a novel protein with internal repeats**

The Predotar and ChloroP algorithms predict that Raa1 is imported into the chloroplast with an N-terminal transit peptide of 54 residues (Figure 4). This localization was confirmed by cell fractionation experiments (see below). A striking feature of the predicted Raa1 polypeptide (Figure 4A) is the occurrence of numerous stretches of repeated amino acids, such as (Ala)3, (Gly)3, (Gln)3 or (Pro)3. Raa1 also contains repetitive sequences like (Pro–Gln)x or (Ser–Ala–Ala)y. Similar repetitive stretches also occur in other C.reinhardtii proteins, notably in another psaA splicing factor, Raa3 (10) and in other proteins involved in chloroplast gene expression (see Discussion). Raa1 harbours five copies of a 38 amino acid repeat (Figure 4B). The first three repeats are in tandem arrangement (residues 1423–1536), the other two are separate, further on the C-terminal side (1792–1809 and 1826–1863). They are related to the PPPEW repeats that are found in Tbc2, a
In order to determine the subcellular localization of the Raa1 protein, chloroplasts were prepared by Percoll gradient centrifugation. Raa1 was found in the chloroplast fraction, which also contains the other chloroplast proteins PsAA and Rubisco, but only trace levels of the cytosolic translation factor eIF4A (Figure 5A, lane 2). When isolated chloroplasts were further fractionated, Raa1 was enriched in the crude membrane pellet along with the integral thylakoid protein PsAA (lane 4), and depleted from the supernatant which contains the soluble protein Rubisco (lane 3). Raa1 thus fractionates as a chloroplast protein, as expected from the genetic data, although its presence in other organelles, such as mitochondria, is not ruled out. The membrane association of Raa1 was confirmed by sucrose gradient equilibrium centrifugation (data not shown).

**Raa1 is part of a large RNase sensitive complex**

The Raa1 protein can be released from the membrane fraction by treatment with moderate concentrations of salts such as 0.2 M ammonium sulphate (data not shown). An extract was prepared by washing membranes with salt and was fractionated by sedimentation on sucrose gradients (Figure 5B). Part of Raa1 was found in the lower part of the gradient (fractions 2–4), as a large complex which sediments just ahead of the 670 kDa molecular mass marker. Part of Raa1 was also found in a complex that sediments more slowly (fractions 6–8), where it is associated with Raa2 (32).

When the ammonium sulphate extract was incubated with RNase before sucrose gradient sedimentation, the position of the large complex shifted towards the top of the gradient. This shift was not observed when the extract was mock-incubated in the absence of RNase. The sensitivity to RNase shows the presence of RNA in the large Raa1 complex.

**DISCUSSION**

**Raa1 is a multifunctional splicing factor**

We have characterized a new mutant, raal-314B, defective in trans-splicing of only the second intron of psaA (class A), and a mutant that fails to splice both introns, raal-L137H (class B). The two mutants are clearly allelic since they fail to complement in genetic tests and both can be rescued by the wild-type Raa1 gene. Thus Raa1 has at least two distinct functions, in splicing of intron 1 and of intron 2. In raal-HN31, which is an allele of raal-L137H (9), there is no detectable mature tscA (29), the small RNA required for splicing intron 1 (7). The same defect is also observed in raal-L137H, but in contrast raal-314B splices intron 1 and accumulates mature tscA RNA. Thus the function of Raa1 in splicing of intron 1 may be direct (in this case Raa1 would have three distinct roles), or may be indirectly mediated by tscA. Two other factors which are required for intron 1 splicing and for processing of tscA were described recently, Rat1 and Rat2 (12). Like for Raa1, it is still unclear whether they are directly or indirectly involved in intron 1 splicing. The pattern of longer tscA transcripts that accumulate in raal-L137H and raal-HN31 and in the rat1 and rat2 mutants are similar (12,29). Thus a complex processing pathway requiring at least three factors (Raa1, Rat1 and Rat2)
is involved in the maturation of \textit{tscA} RNA from longer precursors. The partial defect in \textit{raa1-314B} can be explained by our observation that in \textit{raa1-314B} the 3' part of \textit{Raai} is expressed as a fusion to Arg7; this domain is thus sufficient to promote \textit{tscA} maturation and splicing of intron 1. Our deletion analysis of \textit{Raai} confirms that this 3'-terminal domain is required for \textit{tscA} maturation and splicing of intron 1, while a more central domain is required for splicing of intron 2. However we did not reveal separate domains required for intron 1 splicing and for \textit{tscA} RNA maturation. This is the first time that in \textit{Chlamydomonas} a factor involved in splicing of the two introns of \textit{psaA} (class B) is cloned and characterized. Interestingly, of the fourteen nuclear loci that are required for maturation of \textit{psaA} transcripts, only two belong to this class. The other factors described to date, Raa2 (11), Raa3 (10), Rat1 and Rat2 (12) were shown to be required for either one of the two intervening sequences (classes A and C). The presence of the two functional domains in the same factor suggests that Raai may coordinate splicing of the two \textit{psaA} introns. Mutants defective in splicing of either intron 1 or 2 show that the intervening sequences can be excised independently from each other and in either order (6). However, a mechanism that allows coordinate splicing of the two introns may improve the efficiency of maturation of \textit{psaA} transcripts. Chloroplast factors with dual targets were found in \textit{Arabidopsis} where PGR3 is required for expression of the \textit{petL} operon and of an \textit{ndh} gene (33) and in maize, where CRP1 functions in processing of \textit{petD} RNA, and in translation of \textit{petA} (34).

\textbf{Raa1 is part of a large RNA–protein complex associated with membranes}

We have cloned the \textit{Raal} gene, which encodes a large predicted polypeptide of \textasciitilde 206 kDa. We have shown that Raai is found in the chloroplast fraction, as expected from its predicted transit peptide and from its role in plastid gene expression, although its presence in other organelles cannot be ruled out. Surprisingly it is associated with membranes, and shares this property with Raa2 (formerly Maa2), a protein that is required for trans-splicing of the second intron of \textit{psaA} (11), and Rat1, a factor required for processing of \textit{tscA} and splicing of the first intron (12). This membrane association may be relevant because PsaA is a core subunit of photosystem I embedded in the thylakoid membrane. However Raa3, a factor required for splicing of the first intron of \textit{psaA}, is mostly a stromal protein (10), like proteins involved in splicing of maize chloroplast group II introns, CRS1 and CRS2 (35,36). Chloroplast polyribosomes and some RNA-binding proteins are associated with membranes, and this association may facilitate the assembly of the photosystems and other membrane complexes [reviewed in (37)]. Whether the membrane association of Raai, Raa2 and Rat1 reflects a coupling between RNA splicing, translation and protein assembly into the membrane complexes remains a matter of speculation.

When Raai is released from membranes with moderate concentrations of salt, it sediments in large complexes which can be resolved into two peaks in sucrose gradients. Raai is in part associated with Raa2, with which it can be co-immunoprecipitated (32). The complex containing Raai and Raa2 corresponds to the first peak. The faster sedimenting component of the second peak (\textasciitilde 700 kDa) is sensitive to treatment with RNase, implying that this Raai complex contains RNA. This finding fits with the role of Raai in splicing. However the precursor transcripts are broadly distributed in the gradients (data not shown), and we have not been able to conclusively determine whether they are part of the complex. The RNase-resistant fraction still exhibits splicing, indicating that Raai is in a multimeric complex, perhaps in association with other proteins. Raa3 is also part of a large ribonucleoprotein complex that contains exon 1 and the \textit{tscA} RNA (10). It will be interesting to investigate interactions of Raai with Rat1 in splicing of intron 1. Although Raa3 is mainly in the soluble fraction, and Raa1 and Rat1 in the membrane fraction, it is possible that they could also interact transiently.

\textbf{Raai is a novel splicing factor with repeats}

Because of the limited sequence homology with other known proteins, it is difficult to predict the mode of action of Raai. As already suggested for other proteins found to be essential for group II intron splicing, we expect that Raai may function as an RNA chaperone or a stabilizer of active RNA conformations (8). In maize, the CRS1 protein is specifically required for splicing of a group II intron in \textit{atpF}, it binds the intron and promotes its folding (38). CRS2, together with CAF1 or CAF2, is required for splicing nine other group II introns (35,39). CRS1, CAF1 and CAF2 contain conserved CRM domains which may be RNA-binding modules (40). In plants, genes for proteins similar to intron-encoded maturases are found in the chloroplast and in the nucleus (41,42). Other proteins which associate in the chloroplast with group II introns have been described, but their role in splicing has not yet been established (43,44).

One of the striking features of the predicted Raai protein is the presence of numerous stretches of repeated amino acids. Such repetitive stretches occur occasionally in other proteins, but in Raai they are particularly long and numerous, and involve diverse residues (Ala, Gln, Ser, Pro and Arg). Several other proteins that share this feature with Raai are also involved in chloroplast gene expression, such as Raa3 which is also a \textit{psaA} trans-splicing factor (10), Nac2, Mbb1 and Mca1 which are involved in RNA processing and stability (45–47), or Tbc2 which is implicated in translation (31). The functional significance of these repetitive stretches is unknown. They may play a structural role and may occur in parts of the protein where the precise sequence is not important so that it has drifted during evolution. There are five copies of a 38 amino acid repeat in Raai, related to the 38 amino acid PPPEW repeats of Tbc2 (Figure 4). Four of the repeats are retained in the Arg7–Raai fusion of \textit{raa1-314B} which is sufficient for \textit{tscA} processing and intron 1 splicing, and the last repeat is interrupted in the m\textit{ADkpn} deletion which disrupts these activities (Figure 3C). Two of the repeats are in the region of possible overlap with the domain required for intron 2 splicing, just upstream of the Scal site.

We speculate that although they have a different sequence, the 38 amino acid repeats in \textit{Chlamydomonas} might be a counterpart of the repeats in the PPR and TPR proteins, many of which are also implicated in RNA maturation. The PPR proteins form a large family in plants where a majority are
predicted to be located in the organelles and many are known to interact with RNA [reviewed in (48)]. The TPR-like repeats are found in Mbb1 and Nac2 of *C. reinhardtii* (45,46) and in HCF107 of *Arabidopsis* (49). Each TPR repeat unit fold back to form two anti-parallel α-helices which are stacked with the helices of the adjacent repeats to generate a superhelical structure (50). The related PPR repeats could form a similar structure (51), and one side of the superhelix could provide a surface for protein–RNA interactions, analogous to the protein–RNA interactions between the Puf repeats in Pumilio and their RNA target (52). Thus the repeat-containing chloroplast factors may have been recruited during evolution, by adaptation of a modular protein fold, to recognize new partners or substrates. It will be of interest to determine whether the repeats in Raa1 actually bind RNA, and specifically the psaA repeats in Raa1, which is required for *psaA* precursor transcripts or the *tscA* RNA.

In conclusion, in this report we describe a novel factor, Raa1, which is required for *trans*-splicing of both *psaA* introns in *Chlamydomonas*. Distinct domains in Raa1 are responsible for at least two functions, processing of *tscA* (the small RNA which is required for splicing of *psaA* intron 1) and splicing of *psaA* intron 2. Raa1 may thus play a role in coordinating the excision of the two *psaA* introns. Our analysis of Raa1 provides a basis to further investigate how different events of chloroplast RNA maturation are connected.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors wish to thank Nicolas Roggli for preparing the figures, the members of the Rochaix laboratory for helpful discussions, Hans Sommer for providing a cDNA library, Cris Kuhlemeier for a gift of eIF4A antibody and Frédé Barneche, Stéphane Bellafiore, Paula Duek, Angela Falciatore and Frédéric Iseni for helpful comments on the manuscript. L.M. was supported by a long-term HFSP fellowship. This work was supported by a grant of the Swiss National Fund for Scientific Research (3100-0667763.02). Funding to pay the Open Access publication charges for this article was provided by the Swiss National Fund for Scientific Research.

**Conflict of interest statement.** None declared.

**REFERENCES**

1. Qin,P.Z. and Pyle,A.M. (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.*, 8, 301–308.
2. Bone,H., Presting,G., Roach,J., Schuler,D. and Vogel,J. (2001) The ins and outs of group II introns. *Trends Genet.*, 17, 322–331.
3. Toro,N. (2003) Bacteria and Archaea Group II introns: additional mobile genetic elements in the environment. *Environ. Microbiol.*, 5, 143–151.
4. Villa,T., Pleiss,J.A. and Guthrie,C. (2002) Spliceosomal snRNAs: Mg(2+)/dependent chemistry at the catalytic core? *Cell*, 109, 149–152.
5. Kück,U., Choquet,Y., Schneider,M., Dron,M. and Bennoun,P. (1987) Structural and transcription analysis of the two homologous genes for the P700 chlorophyll a-apoprotein in *Chlamydomonas reinhardtii*: evidence for *in vivo* *trans*-splicing. *EMBO J.*, 6, 2185–2195.
6. Choquet,Y., Goldschmidt-Clermont,M., Girard-Bascou,J., Kuck,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.
7. Goldschmidt-Clermont,M., Choquet,Y., Girard-Bascou,J., Michel,F., Schirmer-Rahire,M. and Rochaix,J.D. (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. *Cell*, 65, 135–143.
8. Schroeder,R., Grossberger,R., Pichler,A. and Waldschütz,C. (2002) RNA folding *in vivo*. *Curr. Opin. Struct. Biol.*, 12, 296–300.
9. 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.
10. Rivier,C., Goldschmidt-Clermont,M. and Rochaix,J.D. (2001) Identification of an RNA–protein complex involved in chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.*, 20, 1765–1773.
11. Perron,K., Goldschmidt-Clermont,M. and Rochaix,J.D. (1999) A factor related to pseudouridine syntheses is required for chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.*, 18, 6481–6490.
12. Balcun,C., Bunse,A., Hahn,D., Bennoun,P., Nickelsen,J. and Kuck,U. (2005) Two adjacent nuclear genes are required for functional complementation of a chloroplast *trans*-splicing mutant from *Chlamydomonas reinhardtii*. *Plant J.*, 43, 636–648.
13. Harris,E.H. (1989) *The Chlamydomonas sourcebook*. A comprehensive guide to biology and laboratory use. Academic Press, Inc., San Diego, CA.
14. Debuchy,R., Purton,S. and Rochaix,J.D. (1989) The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: an important tool for nuclear transformation and for correlating the genetic and molecular maps of the ARG7 locus. *EMBO J.*, 8, 2803–2809.
15. Kindle,K.L. (1990) High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA.*, 87, 1228–1232.
16. Shimogawara,K., Fujisawa,S., Grossmann,A. and Usuda,H. (1998) High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. *Genetics*, 148, 1821–1828.
17. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Purton,S. and Rochaix,J.D. (1994) Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. *Plant Mol. Biol.*, 24, 533–537.
19. Zhang,H., Herman,P.L. and Weeks,D.P. (1994) Gene isolation through genomic complementation using an indexed library of *Chlamydomonas reinhardtii* DNA. *Plant Mol. Biol.*, 24, 663–672.
20. Siwója,J., Fuhrmann,M. and Hegemann,P. (2001) A *Streptomyces rimosus* ahpVIII gene coding for a new type phototransferase provides stable antibiotic resistance to *Chlamydomonas reinhardtii*. *Gene*, 277, 221–229.
21. Goldschmidt-Clermont,M. and Rahire,M. (1986) Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *J. Mol. Biol.*, 191, 421–432.
22. Silljö,C.D., LaVoie,M., Tam,L.W., Tousey,S., Sanders,M., Wu,W., Borodovsky,M. and Lefebvre,P.A. (2001) The Vf1 Protein in *Chlamydomonas* localizes in a rotationally asymmetric pattern at the distal ends of the basal bodies. *J Cell Biol.*, 153, 63–74.
23. Merendino,L., Falciatore,A. and Rochaix,J.D. (2003) Expression and RNA-binding properties of the chloroplast ribosomal protein S1 from *Chlamydomonas reinhardtii*. *Plant Mol. Biol.*, 53, 371–382.
24. Schloss,J.A. (1990) *A Chlamydomonas gene encodes a G protein beta subunit-like polypeptide*. *Mol. Gen. Genet.*, 221, 443–452.
25. Silljö,C. (1998) Organization of the nuclear genome. In Rochaix,J.D., Goldschmidt-Clermont,M. and Merchant,S. (1998) *The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas*. Kluwer Academic Publishers, Dordrecht, The Netherlands, Vol. 7, pp. 25–40.
26. Emmanuelsen,O., Nielsen,H. and von Heijne,G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.*, 8, 978–984.
27. Small,L., Peeters,N., Legeai,F. and Lunel,C. (2004) Pethor; a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics*, 4, 1581–1590.
28. Garnier,J., Gibrat,J.F. and Robson,B. (1996) GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol.*, 266, 540–553.
29. Hahn, D., Nickelsen, J., Hackert, A. and Kück, U. (1998) A single nuclear locus is involved in both chloroplast RNA trans-splicing and 3′ end processing. *Plant J.*, 15, 575–581.

30. Baginsky, S., Siddiqui, A. and Gruissem, W. (2004) Proteome analysis of tobacco bright yellow-2 (BY-2) cell culture plastids as a model for undifferentiated heterotrophic plastids. *J. Proteome Res.*, 3, 1128–1137.

31. Auchincloss, A.H., Perron, K., Girard-Bascou, J. and Rochaix, J.D. (2002) Characterization of the chloroplast group II intron reverse transcriptase/maturases are encoded by nuclear genes in *Chlamydomonas reinhardtii*. *J. Cell. Biol.*, 157, 953–962.

32. Perron, K., Goldschmidt-Clermont, M. and Rochaix, J.D. (2004) A multiprotein complex involved in chloroplast group II intron splicing. *RNA*, 10, 704–710.

33. Yamazaki, H., Tasaka, M. and Shikanai, T. (2004) PPR motifs of the nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta.*, 1519, 46–54.

34. Fisk, D.G., Walker, M.B. and Barkan, A. (1999) Molecular cloning of the nuclear-encoded factor CRS1: a novel group II intron splicing factor. *Plant J.*, 19, 3366–3376.

35. Matthews, B.W. (2005) Structural analysis of the group II intron splicing factor CRS2 yields insights into its protein and RNA interaction surfaces. *EMBO J.*, 24, 2514–2525.

36. Jenkins, B.D. and Barkan, A. (2001) Recruitment of a peptidyl-tRNA hydrolyase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.*, 20, 872–879.

37. Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, 82, 583–601.

38. Ostrowski, O., Cooke, A.M., Watkins, K.P. and Barkan, A. (2005) CRS1: a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. *Plant Cell*, 17, 241–255.

39. Liere, K. and Link, G. (1995) RNA-binding activity of the matK protein encoded by the chloroplast trnK intron from mustard (*Sinapis alba L*). *Nucleic Acids Res.*, 23, 917–921.

40. Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbuhl, F. and Rochaix, J.D. (2000) The Nac2 gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in psbD mRNA stability. *EMBO J.*, 19, 3366–3376.

41. Takanami, T., Takezaki, M. and Shikanai, T. (2004) PPR motifs of the nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta.*, 1571, 3366–3376.

42. Ostrowski, O., Cooke, A.M., Watkins, K.P. and Barkan, A. (2005) CRS1: a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA*, 7, 1227–1238.

43. Yamazaki, H., Tasaka, M. and Shikanai, T. (2004) PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in *Arabidopsis*. *Plant J.*, 38, 152–163.

44. Fisk, D.G., Walker, M.B. and Barkan, A. (1999) Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.*, 18, 2621–2630.

45. Jenkins, B.D. and Barkan, A. (2001) Recruitment of a peptidyl-tRNA hydrolyase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.*, 20, 872–879.

46. Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, 82, 583–601.

47. Ostrowski, O., Cooke, A.M., Watkins, K.P. and Barkan, A. (2005) CRS1: a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. *Plant Cell*, 17, 241–255.

48. Lin,N., Andres,C., Aubourg,S., Bellaoui,M., Bitton,F., Bruyere,C., Caboche,M., Debast,C., Gualberto,J., Hoffmann,B. et al. (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, 16, 2089–2103.

49. Felder,S., Meierhoff,K., Sane,A.P., Meurer,J., Driemeier,C., Plucken,H., Klaff,P., Stein,B., Bechtold,N. and Westhoff,P. (2001) The nucleus-encoded HCF107 gene of *Arabidopsis* provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. *Plant Cell*, 13, 2127–2141.

50. Das, A.K., Cohen, P.W. and Barford, D. (1998) The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein–protein interactions. *EMBO J.*, 17, 1192–1199.

51. Small, I.D. and Peeters, N. (2000) The PPR motif: a TPR-related motif prevalent in plant organelar proteins. *Trends Biochem. Sci.*, 25, 46–47.

52. Wang, X., McLachlan, J., Zamore, P.D. and Hall, T.M. (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell*, 110, 501–512.