Identification of a Chloroplast Ribonucleoprotein Complex Containing *Trans*-splicing Factors, Intron RNA, and Novel Components*[S]

Jessica Jacobs‡, Christina Marx‡§, Vera Kock‡§, Olga Reifschneider‡, Benjamin Fränzel¶, Christoph Krisp¶, Dirk Wolters¶, and Ulrich Kück‡‖

Maturation of chloroplast *psaA* pre-mRNA from the green alga *Chlamydomonas reinhardtii* requires the *trans*-splicing of two split group II introns. Several nuclear-encoded *trans*-splicing factors are required for the correct processing of *psaA* mRNA. Among these is the recently identified Raa4 protein, which is involved in splicing of the tripartite intron 1 of the *psaA* precursor mRNA. Part of this tripartite group II intron is the chloroplast encoded *tscA* RNA, which is specifically bound by Raa4. Using Raa4 as bait in a combined tandem affinity purification and mass spectrometry approach, we identified core components of a multisubunit ribonucleoprotein complex, including three previously identified *trans*-splicing factors (Raa1, Raa3, and Rat2). We further detected *tscA* RNA in the purified protein complex, which seems to be specific for splicing of the tripartite group II intron. A yeast-two hybrid screen and co-immunoprecipitation identified chloroplast-localized Raa4-binding protein 1 (Rab1), which specifically binds *tscA* RNA from the tripartite *psaA* group II intron. The yeast-two hybrid system provides evidence in support of direct interactions between Rab1 and four *trans*-splicing factors. Our findings contribute to our knowledge of chloroplast multisubunit ribonucleoprotein complexes and are discussed in support of the generally accepted view that group II introns are the ancestors of the eukaryotic spliceosomal introns. ***Molecular & Cellular Proteomics 12: 10.1074/mcp.M112.026583, 1912–1925, 2013.***

Intron-containing genes from prokaryotic or organellar genomes carry either group I or group II introns, each of which has distinct features. The splicing mechanism of group II introns and the secondary structures of their presumed active sites were used as early arguments for the hypothesis that this class of introns represents the ancestors of eukaryotic spliceosomal introns (1, 2). It was further assumed that group II introns invaded the eukaryotic nucleus and subsequently proliferated at various genomic sites, leading to the degeneration of the catalytic intron structure into small nuclear RNAs (snRNAs).[1] This assumption was supported by the observation of naturally occurring variants of group II introns that are split into two or more pieces (3), reminiscent of eukaryotic spliceosomal RNA (1). Group II intron RNAs are characterized by six conserved domains, and tertiary interactions among these domains generate the compact native and catalytic complex. Some of these group II intron domains have been shown to act in *trans* on the splicing of other introns that lack the corresponding domain (4). *In vivo*, various RNA-binding proteins promote the formation of catalytically active intron RNA. In contrast to the nuclear spliceosome, which acts generally on a broad range of nuclear-encoded pre-mRNAs, proteins involved in organellar intron splicing seem to more efficiently stabilize the active three-dimensional RNA structure *in vivo*.

Several splicing factors in higher plants, such as the chloroplast RNA-splicing and ribosome maturation (CRM) domain protein CRS1, as well as the pentatricopeptide repeat proteins OTP51 and PPR4, have been reported to be involved in the splicing of single transcripts (5, 6). Nonetheless, there are splicing factors that carry out functions on a broad range of transcripts, including CRS2 and its associated proteins CAF1 and CAF2, and WTF1, a splicing factor containing a plant organelle RNA-recognition domain (5, 6). Sedimentation and co-fractionation experiments in, for example, maize have demonstrated that these proteins are part of large multiprotein and ribonucleoprotein complexes with their cognate RNAs (5, 7). In addition, these complexes resemble the nuclear spliceosome in which the snRNAs associate with more than 200 proteins (8).

From the ‡Department for General and Molecular Botany, Ruhr-University Bochum, D-44780 Bochum, Germany; ¶Department of Analytical Chemistry, Ruhr-University Bochum, D-44780 Bochum, Germany

Received December 18, 2012, and in revised form, February 27, 2013

Published, MCP Papers in Press, April 4, 2013, DOI 10.1074/mcp.M112.026583

[1] The abbreviations used are: CRM, chloroplast RNA-splicing and ribosome maturation; EMSA, electrophoretic mobility shift assay; MudPIT, multidimensional protein identification technology; Ni-NTA, nickel-nitrilotriacetic acid; qRT-PCR, quantitative real-time PCR; SD, synthetic dropout; snRNA, small nuclear RNA; TAP, tandem affinity purification.

The unicellular green alga *Chlamydomonas reinhardtii* is also known to contain high molecular weight complexes containing splicing factors (9, 10). In this alga, the chloroplast-encoded *psaA* gene, which encodes a major subunit of photosystem I, is split into three independently transcribed exons. Splicing of the *psaA* pre-mRNAs requires the assembly of two group IIB introns (11). For *psaA* intron 1, the catalytically active intron structure is fragmented into three chloroplast-encoded intron sequences, including the core *tscA* RNA (12). The *tscA* RNA is required in order to form the active intron structure, because it complements the tripartite intron by contributing domains D2 and D3, as well as parts of D1 and D4 (Fig. 1*A*). Several photosynthetic mutants have been identified that are deficient in the splicing of either intron 1 or intron 2, or both, or in the processing of *tscA* RNA. At least 14 nuclear loci are involved in *trans*-splicing, with six splicing factors identified to date (13). Two of them, Raa3 and Raa4, are directly involved in the correct splicing of intron 1, and Raa1, Rat1, and Rat2 are essential for the processing of *tscA* RNA from a polycistronic precursor, a prerequisite for intron 1 splicing. Besides its function in processing *tscA* RNA, Raa1 plays a role in splicing the second *psaA* intron. A further protein involved in splicing the second *psaA* intron is Raa2. Except for Rat1, which is significantly homologous to the NAD^+-binding domain of poly (ADP-ribose) polymerases, and Raa2, which shows similarities to pseudouridine synthases, all other *psaA*-splicing factors display only slight sequence homologies to other known proteins (14, 15).

So far little is known about the protein–protein interactions and the overall composition of the organellar ribonucleoprotein complexes involved in *psaA trans*-splicing. In this study, we identified basic subunits of a multipartite complex that contains four functional *trans*-splicing factors. To identify the components of this complex, we used a multifaceted approach combining tandem affinity purification (TAP), mass spectrometry, and yeast two-hybrid screening. We applied different environmental conditions (light, dark, anaerobiosis) to define true and essential subunits of the basic splicing complex that are present under various conditions. Further, we detected a novel intron RNA binding protein that interacts with at least four splicing factors. The protein–RNA complex described here points toward a chloroplast multisubunit splicing complex specific for a tripartite group II intron that is reminiscent of the nuclear spliceosome.

EXPERIMENTAL PROCEDURES

Strains, Conditions, and Transformation—*C. reinhardtii* strains and growth conditions are listed in supplemental Table S1. For TAP, *C. reinhardtii* cultures were grown in tris-acetate phosphate medium in the light. For the induction of anaerobic conditions, a concentrated and shaded *C. reinhardtii* culture was flushed with argon as described elsewhere (16). Hydrogenase activity was measured as described elsewhere (16). For dark adaptation, cells were dark incubated for 2 h. The nuclear transformation of algal cells was carried out according to the glass-bead method (17) with 5 μg circular or hydrolyzed DNA.

Molecular Biological Techniques—Procedures for standard molecular techniques were performed as reported elsewhere (14, 18). *Escherichia coli* strain XL1-blue MRF′ served as the host for general plasmid construction and maintenance (19). *S. cerevisiae* strain PJ69–4A (20) was used for homologous recombination as described by Colot *et al.* (21). The transformation of yeast cells was done by means of electroporation according to the method of Becker and Lundblad (22) in a Multiporator (Eppendorf, Hamburg, Germany) at 1.5 kV. Transformants were selected for tryptophan or leucine prototrophy. DNA extraction was performed using the E.N.Z.A. Plasmid Miniprep Kit I (Peqlab Biotechnologie, Erlangen, Germany) after treatment with glass beads.

C. reinhardtii total RNA was prepared as described elsewhere (11). PCR and RT-PCR experiments were performed as described elsewhere (18). One-step RT-PCR was performed with the OneStep RT-PCR Kit from Qiagen (Hilden, Germany) according to the manufacturer's instructions. Recombinant plasmids and oligonucleotides used for PCR experiments, protein synthesis, or generation of transgenic algal strains are listed in supplemental Table S2 and supplemental Table S3, respectively. If necessary, suitable restriction sites for cloning were added to oligonucleotides.

Quantitative RT-PCR—TAP eluates containing nucleic acids were purified via phenol/chloroform/isoamyl alcohol (25:24:1) extraction and precipitation at −20 °C. Genomic DNA was removed by means of DNase I treatment for 25 min at 25 °C. 1 μl of each 44-μl sample was subjected to One-Step qRT-PCR (KAPA Sybr Fast ABI Prism, Peqlab, Erlangen, Germany) using gene-specific oligonucleotides (supplemental Table S3). As a control for successful DNaseI treatment, each reverse transcription was carried out twice, once with and once without reverse transcriptase. qRT-PCR was performed in an ABI 5700 (Applied Biosystems, Foster City, CA) with a One-Step qRT-PCR Kit containing SybrGreen and ROX (KAPA Sybr Fast ABI Prism, Peqlab) in a volume of 20 μl. Each reaction was carried out in triplicate with an oligonucleotide primer at a concentration of 10 μM. Primers were selected to have melting temperatures of 56 °C to 61 °C and to yield amplicons of 147 to 185 bp. PCR conditions were as follows: 42 °C for 5 min, 95 °C for 1 min, and 40 cycles of 95 °C for 5 s and 60 °C for 20 s, followed by a melting curve analysis. Amplicon size was verified using gel electrophoresis. Primer pair efficiencies and expression ratios were calculated as described elsewhere (23). Each qRT-PCR experiment was done with two biologically independent samples.

Construction of Plasmids—To construct the Raa4 two-hybrid plasmids, cDNA fragments coding for amino acids 48–610 and 609–1143 were amplified (primers: for_Y2H1, rev_Y2H2; for_Y2H3, rev_Y2H4) and ligated in pDrive or pBIIKS+. After restriction with EcoRI and BamHI, the resulting fragment was cloned into EcoRI and BamHI sites of pGADT7 resulting in plasmids pGADT7_Raa4-A and pGADT7_Raa4-B. The full-length version of Raa4 (pGADT7_Raa4-FL) was obtained after digestion of pBIIKS+_ Raa4-B with *Srf*I and BamHI and ligation of the resulting fragment in *Srf*I and BamHI restricted pGADT7_Raa4-A.

Rab1 yeast-two hybrid vectors were generated as follows: DNA fragments coding for amino acids 51–725 and 668–1216 were amplified from cDNA (primers: OVK48, OVK49; OVK50, OVK51) and cloned into EcoRI and BamHI restriction sites of pGADT7 resulting in pGADT7_Rab1-A and pGADT7_Rab1-B.

For the generation of two-hybrid vectors carrying Raa3 subfragments, *RAA3* fragments coding for amino acids 674–1298, 70–675, 1296–1783, and 674–1783 were amplified from cDNA (primers: for_Raa3-1, rev_Raa3-2; for_Raa3–3, rev_Raa3–3; for_Raa3-1, rev_Raa3-1; for_Raa3-2, rev_Raa3-2) and inserted into EcoRI and SalI restriction sites of vector pGBKT7 resulting in plasmids pGBKT7_Raa3-A, pGBKT7_Raa3-B, pGBKT7_Raa3-C, and pGBKT7_

Raa3-D. For the Raa3 full-length construct (pGBKT7_Raa3-FL), cDNA was amplified (primers: for_Raa3-3, rev_Raa3-2) and cloned in EcoRI and BamHI sites of pGBKT7.

In order to construct *RAT2* two-hybrid plasmids, cDNA was amplified in three fragments (primers: for_pGADT7_Rat2, rev_Rat2_F1; for_Rat2_F2, rev_Rat2_F2; for_Rat2_F3, rev_pGADT7_Rat2) that overlapped each other and the pGADT7 cloning site. Full-length cDNA of *RAT2* (pGADT7_Rat2-FL) was obtained via homologous recombination in *S. cerevisiae* PJ69–4A. Two fragments, *RAT2*-A and *RAT2*-B, coding for amino acids 1–682 and 683–1376 were amplified from pGADT7_Rat2-FL (primers: for_pGADT7_Rat2, rev_Rat2-Y2H-F1; for_Rat2-Y2H-F2, rev_pGADT7_Rat2) and introduced in pGADT7 via homologous recombination. All *RAT2* fragments (Rat2-FL, Rat2-A, and Rat2-B) were cloned in pGBKT7 via NdeI and the compatible restriction sites XhoI and SalI resulting in pGBKT7_Rat2-FL, pGBKT7_Rat2-A, and pGBKT7_Rat2-B.

For the generation of yeast two-hybrid vectors containing the *RAA1-A* fragment, cDNA of *RAA1* was amplified in two fragments (primers: for_pGADT7_Raa1, rev_pGADT7_Raa1; for_Raa1-F3–3, rev_Raa1-F3–3). The two fragments showed regions overlapping each other and the pGADT7 cloning site and were introduced to pGADT7 by means of homologous recombination. *RAA1-A* was inserted into the NdeI and BamHI restriction site of pGBKT7.

For the construction of His_6::Raa4M, an 884-bp fragment of *RAA4* cDNA was amplified via PCR (primers: for_Raa4-M2 and rev_Raa4-M). After ligation into pTOPO and hydrolysis of the resulting plasmid with BamHI and HindIII, the 870-bp fragment was ligated into pQE30 cut with BamHI and HindIII, resulting in plasmid pQE30_Raa4-M2.

For cloning of the Rab1::One-STrEP-tag fusion construct, *RAB1* cDNA was amplified (primers: OVK_29, OVK_30) and cloned in pASG-IBA3 via StarGate® combinatorial cloning according to the manufacturer's instructions (IBA GmbH, Göttingen, Germany).

For the generation of Rab1cTP::cGFP, a genomic fragment was amplified using primers Rab1_cTP_for and Rab1_cTPlong_rev and cloned in pDrive. The resulting plasmid was restricted with NheI and cloned in NheI cut pCr1g resulting in pCr1g_Rab1cTP.

Generation of TAP-tagged RAA4—The c*TAP* gene was amplified from pUC57 using oligonucleotides Taptag1 and Taptag1 and cloned into plasmid pCrg1 (18) via BglII restriction sites resulting in plasmid pCM10. For the generation of an Raa4::TAP tag fusion construct, *RAA4* was amplified in two fragments (primers Raa4-A1, Raa4-A2 and Raa4-B1, Raa4-B2) from BAC subclone 2539_1A (14). Fragment *RAA4*-B was cloned via XbaI restriction sites in pCM10 resulting in plasmid pCM12. Fragment *RAA4*-A was cut with *Pme*I and cloned in *Pme*II restricted pCM12 resulting in plasmid pCM13. For deletion of the median *Pme*I restricition site, pCM13 was restricted with *Mau*BI and SrfI. The resulting 1.2-kb fragment was replaced with the corresponding fragment from plasmid 2539_1A resulting in plasmid pCM15, which comprises the genomic sequence of *RAA4* fused to the *TAP* tag gene under control of the artificial *RBCS2*/*HSP70* tandem promoter. For the construction of an RbcS1::TAP tag fusion construct, *RBCS1* was amplified from genomic DNA (primers: RbcS1_NheI_1, RbcS1_NheI_2) and cloned in pDrive. *RBCS1* was then introduced into plasmid pCM10 using restriction site NheI resulting in plasmid pCM18.

Laser Scanning Confocal Fluorescence Microscopy—The fluorescence emissions of transformed *C. reinhardtii* cells were analyzed via laser scanning confocal fluorescence microscopy using a Zeiss LSM 510 META microscopy system (Carl Zeiss, Jena, Germany) based on an Axiovert inverted microscope. cGFP and plastids were excited with the 488-nm line of an argon-ion laser. The fluorescence emission was selected by band pass filter BP505–530 and long pass filter LP560, respectively, using beam splitters HFT UV/488/543/633 and NFT545 as described elsewhere (18).

Heterologous Synthesis of RAA4 and RAB1 in E. coli—For the heterologous synthesis of *RAA4* and *RAB1*, *E. coli* BL21(DE3) was transformed with the respective plasmids (pQE30_Raa4-M, pASG-IBA3_Rab1). Protein production was performed in 0.5 l LB medium containing 100 μg ml^{-1} ampicilline. Fusion proteins were isolated from inclusion bodies according to the procedure described by Steinle *et al.* (24) as described in Ref. 14. The purification of refolded recombinant proteins was performed according to the manufacturer's instructions (GE Healthcare, Freiburg, Germany; Qiagen, Hilden, Germany).

Electrophoretic Mobility Shift Assays—For RNA mobility shift assays, uniformly ^{32}P-UTP-labeled run-off transcripts served as substrate RNAs and were generated by the *in vitro* transcription of plasmids as given in supplemental Table S2. *In vitro* transcription and EMSAs were performed as previously reported (14, 18, 25, 26). Unlabeled competitor RNAs and nonspecific competitor RNA derived from plasmid pBSIIKS$^+$ (Stratagene, La Jolla, CA) were synthesized as described elsewhere (18, 26). Recombinant His-tagged cNAPL protein or GST-tagged Raa4 were used as controls and were purified as described elsewhere (14, 18).

Sequence Analysis—Sequences were retrieved from the *C. reinhardtii* Joint Genome Institute database, v5.3 (27). Basic Local Alignment Search Tool (BLAST) searches were performed using NCBI's BLAST Server. Isoelectric point and sequence masses were calculated by the program Clone Manager 9 Professional Edition (Scientific & Educational Software, Cary, NC). Secondary structure analysis was performed using version IV of the GOR secondary structure prediction method (28). Protein motifs were predicted with Motif Scan (29). For the identification of RNA binding residues in proteins, the programs BindN (30) and RNABindR (31) were used. Putative targeting signal sequences were identified with PredAlgo (32), ChloroP 1.1 (33), TargetP 1.1 (34), and SignalP V4.0 (35).

Yeast Two-hybrid Analysis—For construction of the cDNA library, *C. reinhardtii* cultures were grown to mid-log phase in tris-acetate-phosphate medium in the light, tris-acetate-phosphate medium in the dark, high-salt medium in the light, and tris-acetate-phosphate medium in the light, and then they were shifted to dark conditions. Yeast two-hybrid cDNA library generation, screening, and mating assays were performed with the Matchmaker™ Library Construction and Screening Kit according to the manufacturer's instructions (Clontech Laboratories, Inc., Mountain View, CA). Alternatively, co-transformation of *S. cerevisiae* PJ69–4a with the two-hybrid plasmids was performed.

Pull-down Assay—*In vitro* binding assays were performed as described by Bals *et al.* (36). 5 μg of the indicated proteins were incubated in 100 μl of 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM Imidazol for 30 min at room temperature. Proteins were then applied to 30 μl Ni-NTA resin (Qiagen, Hilden, Germany) washed with 5 ml 50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM Imidazol and eluted in 50 μl 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM Imidazol. Rab1-Strep incubated with Ni-NTA resin and His-tagged chloroplast recognition particle cpSRP served as controls (36). For pull-down analysis with crude protein extracts, *E. coli* BL21(DE3) was transformed with the respective plasmids (pQE30_Raa4-M, pASG-IBA3_Rab1). Protein production was performed in 100 ml LB medium containing 100 μg ml^{-1} ampicillin. For the preparation of crude extracts, cells were pelleted; resuspended in 1 ml 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM Imidazol; and sonicated six times for 30 s each at 30% to 40% power with a Branson sonifier 250 (Branson Ultrasonics Corp., Danbury, CT). Extracts were then centrifuged at 13,000 rpm for 15 min at 4 °C, and supernatants were used for analysis. Supernatants (250 μl) were mixed and incubated on a rotating wheel for 30 min at room temperature. 80 μl Ni-NTA resin was added and incubated with the protein extract on a rotating wheel for 1 h at room temperature.

Ni-NTA resin was applied to Wizard Minicolumns (Promega, Madison, WI), washed four times with 10 ml 50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM Imidazol. Elution was performed with 80 μl 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM Imidazol.

TAP of Raa4 Interacting Proteins—For the preparation of crude extracts, pelleted *C. reinhardtii* cells were resuspended in lysis buffer (100 mM Tris, 150 mM NaCl, pH 8.0) containing protease inhibitors (Protease Inhibitor Mixture VI, Calbiochem, Bad Soden, Germany) and sonified (4*30 s, 30% to 40% power). Cell debris was sedimented via centrifugation, and the supernatant was applied to TAP according to the work of Bayram *et al.* (37) as described elsewhere (38).

Multidimensional Protein Identification Technology—The digestion of precipitated proteins was performed in 25 mM ammonium bicarbonate buffer pH 7.8 with sequencing grade trypsin (Promega, Madison, WI) (1:50 w/w) overnight at 37 °C. Extracted peptides were diluted in 0.1% trifluoroacetic acid and analyzed using multidimensional protein identification technology (MudPIT) as described elsewhere (39). Peptides were detected with an LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific). Proteome Discoverer software, version 1.2, was used to interpret acquired MS/MS data, searching the spectra against the *C. reinhardtii* database creinhardtii_223_peptide, which comprises 19,529 entries (27). In addition, our database included 69 entries derived from the *C. reinhardtii* chloroplast genome (40). The mass accuracy for precursors was set at 10 ppm, and for fragments at 0.8 Da. Oxidation of methionine was set as a possible peptide modification. Accepted proteins had at least two unique peptides with a false discovery rate of less than 1% using a decoy database for reversed database alignment. Detected proteins were considered as specific Raa4 binding partners if they were identified in all Raa4::TAP purification replicates but not in the negative controls (two TAP experiments with strain RST-1, expressing *RBCS::TAP* tag fusion gene, and two TAP experiments with the untagged wild type arg$^-$cw15).

RESULTS

Purification of Raa4 Complexes Using a Codon-optimized TAP Tag Reveals Novel Protein–Protein Interactions between Chloroplast Splicing Factors—To elucidate the protein interaction network that mediates *psaA* splicing, we selected the recently identified protein Raa4, which is involved in *trans*-splicing of the first *psaA* intron (14), as a target protein in protein–protein interaction studies. TAP enables the detection of protein–protein interactions under native conditions, the determination of direct interactions, and the identification of whole protein complexes with low background contamination (41). For the purification of native Raa4 complexes, a vector was generated in which a codon-optimized TAP tag consisting of protein A, a tobacco etch virus protease cleavage site, and a calmodulin-binding peptide was C-terminally fused to Raa4. The nucleotide sequence of the optimized gene exhibits 76% homology to the original *TAP* tag gene and has an increased guanine-cytosine content of 59% (guanine-cytosine content of the original *TAP* tag: 46%) (41). *C. reinhardtii trans*-splicing mutant *raa4* was transformed with the Raa4::TAP tag fusion construct, and transformants were selected under high light conditions. Several photosynthetically active transformants were obtained, indicating functionality of the Raa4::TAP tag fusion protein and the respective protein complexes. We analyzed genomic integration of the *RAA4::TAP* tag fusion gene using PCR and the expression of the fusion gene using RT-PCR (supplemental Figs. S1*A* and S1*B*). For further investigation, we selected a single transformant (R4T-1) that exhibited strong expression of the fusion gene (supplemental Fig. S1*C*). Immunoblotting finally detected the TAP-tagged Raa4 in crude extracts of R4T-1 (supplemental Fig. S1*D*). As a negative control, we generated an *RBCS::TAP* tag fusion construct that was used for transformation of strain *uvm4*, a *C. reinhardtii* strain that efficiently expresses transgenes (42). Several transformants were obtained and analyzed regarding the genomic integration and expression of the recombinant gene (supplemental Fig. S2). Algal transformant RST-1 was finally selected for further experiments.

To identify subunits that represent true and essential components of the basic *psaA* splicing complex, we used *C. reinhardtii* cultures adapted to three environmental conditions (light, dark, and anaerobiosis) for TAP. In addition to light/dark conditions that would probably only modulate the levels of protein subunits in the predicted spliceosome, anaerobic conditions were chosen for the propagation of cells. Splicing of *psaA* pre-RNAs has to occur under anaerobic conditions because PsaA is a major component of Photosystem I, which under anaerobic conditions is required in order to maintain electron transfer to hydrogenase Hyd1, a key enzyme of *C. reinhardtii* anaerobic metabolism (16). *C. reinhardtii* cultures were anaerobically induced by flashing a highly concentrated and shaded culture with argon. The success of the anaerobic adaptation was monitored by an *in vitro* hydrogenase activity assay, as hydrogen production only occurs under anaerobic conditions and is therefore a good measure for anaerobiosis (supplemental Fig. S3). For dark adaptation, cells were grown in the light and then shifted to darkness for 2 h.

TAP crude extracts obtained from the above-described *C. reinhardtii* cultures were applied to IgG beads. After tobacco etch virus protease cleavage, the eluted proteins were further purified using calmodulin beads and then analyzed using MudPIT, a non-gel-based approach for the identification of proteins from complex mixtures. TAP-MudPIT experiments were performed in biological replicates. In total we investigated four algal cultures grown in light, three dark-adapted, and three anaerobically induced cultures for TAP. In all cases, peptides specific for the Raa4::TAP tag were unambiguously identified (Table I). In order to discriminate between specific Raa4-binding partners and co-purifying contaminants, TAP was performed twice using either non-tagged extracts from wild-type strain arg$^-$cw15 or strain RST-1. Proteins that were recovered in these purifications were considered as false positives. Proteins that were identified in at least three out of four (light) or two out of three (dark, anaerobiosis) replicates were considered as high-confidence interactors. By comparing the three obtained datasets, we identified 32 proteins that were present under all three environmental conditions (supplemental Fig. S4). These proteins were then analyzed regarding their subcellular localization, because true Raa4 interac-

TABLE I

Proteins identified in TAP experiments. Ten independent TAP purifications (P1–P10) with protein extracts derived from strain R4T-1 (strain expressing RAA4::TAP tag fusion gene) were performed. For P1–P4, cells were grown under continuous light; for P5–P7, cells were dark incubated for 2 h; and for P8–P10, cells were anaerobically induced by flashing with argon. Proteins that were detected in at least three out of four (light) or two out of three (dark, anaerobiosis) replicates but not in the negative controls (two TAP experiments with strain RST-1, expressing RBCS::TAP tag fusion gene, and two TAP experiments with the untagged wild-type arg$^-$cw15) were considered as high-confidence interactors. Twenty-two proteins are present under all three environmental conditions and are considered to be chloroplast targeted. The name of the respective gene product or the functional annotation is given in column "Protein." Proteins were analyzed with web-based tools (PredAlgo, TargetP, and ChloroP) for the prediction of subcellular localization. "C" and "M" indicate the occurrence of predicted chloroplast and mitochondrial signal peptide sequences

						Light				Dark			Anaerobiosis		
						Σ≠ peptides[e]				Σ≠ peptides[e]			Σ≠ peptides[e]		
Accession	Protein	kDa[a]	PA[b]	TP[c]	CP[d]	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Trans-splicing factors															
Cre02.g125300.t1.1	Raa4 (bait)	116	-	C	C	12	29	34	20	36	32	24	36	19	14
Cre09.g388372.t1.2	Rat2, OPR domain	144	C	C	C	11	24	25	10	24	17	14	26	20	9
Cre09.g394150.t1.3	Raa1, OPR domain	210	C	C	C	13	18	26	10	16	18	10	20	8	-
Cre12.g531050.t1.3	Raa3	180	C	C	C	8	11	25	10	14	14	6	9	9	3
Not annotated															
Cre02.g111800.t1.2	No functional annotations	45	C	M	C	4	7	7	2	4	2	-	5	6	-
Cre10.g440000.t1.3	No functional annotations, OPR domain	269	C	C	C	27	44	47	16	26	34	18	46	25	14
Cre12.g524250.t1.2	No functional annotations	63	C	M	-	-	10	7	7	6	8	3	5	7	-
Cre17.g698750.t1.2	No functional annotations, OPR domain	92	M	M	C	6	12	13	4	8	6	2	14	9	-
Cre17.g712300.t1.3	No functional annotations	100	C	M	-	12	13	22	2	8	6	3	20	4	-
Cre17.g724450.t1.2	No functional annotations	71	M	C	C	5	16	15	10	13	12	8	16	10	3
g11582.t1	No functional annotations	139	M	M	C	5	16	22	10	9	17	13	22	14	3
g9821.t1	No functional annotations	83	C	M	C	3	4	2	-	3	3	-	10	2	-
g33.t1	No functional annotations	109	C	C	C	6	12	16	12	10	12	6	13	8	7
Gene expression															
Cre08.g373800.t1.3	U1 snRNP-specific protein C	113	M	M	C	9	13	18	8	11	16	9	18	15	10
Cre12.g556050.t1.2	Plastid ribosomal protein L9	22	C	M	-	4	6	6	2	4	3	-	4	2	-
Cre13.g592150.t1.2	U5 RNA helicase, DEAD/DEAH box helicase	103	C	C	-	8	2	10	2	-	4	6	11	3	-
g15179.t1	CRS1/YhbY (CRM) domain, RNA binding	109	M	M	C	2	2	7	-	2	-	2	4	3	-
DNA repair, protein folding, metabolism															
Cre02.g147850.t1.2	ATP-dependent subunit of mitochondrial HslUV protease	70	M	M	C	4	9	8	7	5	3	11	11	4	-
Cre07.g314650.t1.2	Chloroplast RecA recombination protein	45	C	C	C	8	3	5	3	4	2	2	8	2	-
Cre12.g514200.t1.2	GMC oxidoreductase	64	C	M	C	2	6	7	5	3	3	4	3	3	-
Cre12.g551500.t1.3	DnaJ-like protein	47	M	C	C	2	2	2	-	4	3	-	5	3	-
Cre14.g610500.t1.2	Short-chain dehydrogenase/reductase	40	C	M	C	3	5	7	3	-	3	4	9	2	-
g13120.t1	ATP-dependent CLP protease	109	C	-	-	11	7	12	8	3	7	4	12	7	-

[a] kDa, predicted molecular weight in kDa.

[b] PA, PredAlgo.

[c] TP, TargetP.

[d] CP, ChloroP.

[e] Σ≠ peptides, number of unique peptides identified in each R4T-1 TAP experiment.

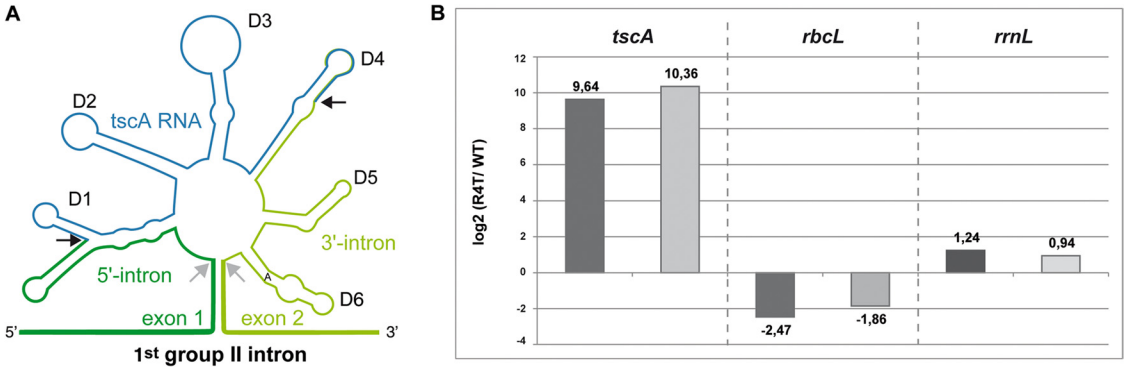

FIG. 1. ***tscA* RNA is co-purified with Raa4.** *A,* secondary structure model of the first *psaA* intron flanked by exon 1 and exon 2. The splicing sites are indicated by gray arrows. The tripartite group II intron consists of three separate RNA molecules (*tscA*, 5′-intron exon 1, and 3′-intron exon 2). The six conserved domains (D1–D6) of group II introns are indicated. The peripheral structures of split D1 and D4 domains were predicted via *in silico* methods (3). Fragmentation sites are indicated by black arrows. *B*, quantitative real-time PCR (qRT-PCR) analysis to detect *tscA* RNA in affinity purified Raa4 complexes. RNA was isolated from Raa4 TAP eluate and from TAP eluate derived from the untagged wild-type strain arg⁻cw15. All RNA samples were treated with DNase I to remove possible DNA contaminants. One-step qRT-PCR was performed with two biological replicates (dark gray bars/light gray bars) to determine the *tscA*, *rbcL*, and *rrnL* RNA content in RNA samples. The data are shown as ratios for strain R4T-1 *versus* wild type in log2 scale.

tion partners have to be localized to the chloroplast. Interestingly, most proteins were not identified in *C. reinhardtii* chloroplast proteome datasets (43). None of the known *trans*-splicing factors were identified in these datasets, suggesting that these factors are present in low abundance. Therefore, we analyzed Raa4 interaction partners with the web-based *in silico* tools TargetP, ChloroP, and the recently developed PredAlgo in order to better predict the subcellular localization of these putative Raa4 interaction factors (Table I) (32–34). In contrast to PredAlgo, which was trained on *C. reinhardtii* proteins, TargetP and ChloroP were first developed for higher plants proteins. Thus, they mispredict the localization of many *C. reinhardtii* proteins and predict chloroplast localized proteins as mitochondrial proteins (32, 43). Proteins were considered to be chloroplast localized if they had a chloroplast prediction in PredAlgo or a chloroplast/mitochondrion prediction in TargetP and ChloroP. According to this *in silico* analysis, 22 proteins can be assigned to the chloroplast proteome.

Remarkably, the most abundant proteins, which were identified with a large number of peptides, included the *trans*-splicing factors Raa1, Raa3, and Rat2 (Table I) with up to 26, 25, and 26 different peptides, respectively. We detected once an unknown protein in the purified protein complex, which as described below was identified in a yeast-two hybrid screening as an *R*aa4 *b*inding protein (Rab1). A second group of proteins identified here includes proteins that are not currently annotated or characterized. We analyzed the protein sequences with a motif scan but detected no conserved motifs or domains. However, two of these proteins were identified recently as octatricopeptide repeat (OPR) proteins. Transcript Cre10.g440000.t1.3 (in *C. reinhardtii* Joint Genome Institute database v5.3) encodes for a putative protein of 269 kDa and exibits 22 OPRs; transcript Cre17.g698750.t1.2 encodes for a protein with a predicted molecular weight of 92 kDa and 10 OPRs (44). A third group is composed of proteins that do not exhibit functions directly related to *psaA* splicing but have a general function in RNA metabolism. This includes predicted proteins with similarities to spliceosomal proteins (Cre08.g373800.t1.3, g11889.t1) or a putative protein harboring a CRM domain. A further group of Raa4 associated proteins includes proteins with similarities to proteins involved in DNA repair, protein folding, or metabolism.

We also tested the TAP eluates for the presence of intron RNA. qRT-PCR was performed with two replicates to detect *tscA* RNA in the affinity purified Raa4 splicing complex. As shown in Fig. 1*B*, the purified Raa4-splicing complex contained significant amounts of *tscA* RNA relative to the untagged wild-type strain. As a negative control, we analyzed the presence of *rbcL* and *rrnL* RNA, to exclude the possibility that unspecific RNAs are enriched in the purified complex. As shown in Fig. 1*B*, we were unable to detect increased amounts of *rbcL* and *rrnL* transcripts in the Raa4 complex.

Taken together, these observations demonstrate that the two-step affinity purification was applied successfully and that it is possible to isolate native splicing complexes with TAP-tagged Raa4 as bait. Moreover, the co-purification of Raa4 with Raa1, Raa3, and Rat2 demonstrates interactions among various splicing factors and shows that *psaA trans*-splicing factors are organized in a heteromeric ribonucleoprotein complex.

A Deep Yeast Two-hybrid Screen identifies Raa4-interacting Proteins—To identify direct interaction partners of Raa4, an extensive yeast two-hybrid screen with 1.4×10^9 clones was performed. Raa4 lacking the putative transit peptide was fused to the DNA-binding domain of GAL4 and used as bait to screen a cDNA library from *C. reinhardtii.* To generate the cDNA library, *C. reinhardtii* cultures were grown under various

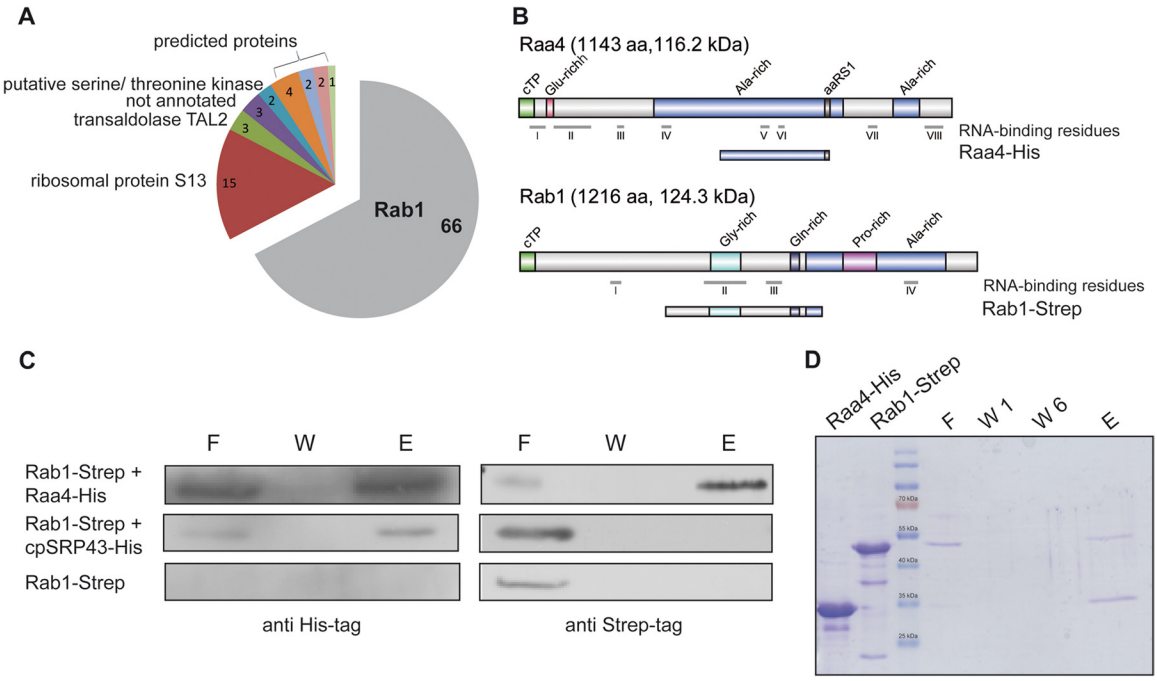

FIG. 2. **Raa4 interacts with Rab1.** *A*, yeast two-hybrid screening with Raa4 as bait. A total of 98 clones were identified in a yeast two-hybrid screening using Raa4 as bait. 66 clones carried cDNA fragments corresponding to the *RAB1* gene, which encodes for an unknown protein. *B*, primary structure of Raa4-His (amino acids 533–819) and Rab1-Strep (amino acids 393–800). Depicted are chloroplast transit peptides (cTP) and glutamine-, alanine-, glycine-, and proline-rich regions (Glu-rich, Ala-rich, Gly-rich, and Pro-rich, respectively). Raa4 shows loose homology to an aminoacyl tRNA synthetase class I signature (aaRS1). RNA-binding residues were predicted using BindN and RNABindR. *C*, analysis of the *in vitro* binding between Raa4-His and Rab1-Strep. 5 μg of recombinant Raa4-His was incubated with 5 μg of recombinant Rab1-Strep. Proteins were repurified by Ni-NTA resin, and eluted proteins were detected via SDS-PAGE analyses and immunoblotting using antibodies against His- and Strep-tag. An unrelated His-tagged protein (cpSRP43) served as a negative control. 5 μg Rab1-Strep was incubated with Ni-NTA resin to rule out the possibility of unspecific interactions between Rab1 and Ni-NTA resin. F, flow through; W, wash step; E, eluate. *D*, *in vitro* pull-down assays with Raa4-His and Rab1-Strep. 5 μg of recombinant Raa4-His was incubated with 5 μg of recombinant Rab1-Strep. Proteins were repurified by Ni-NTA resin. Aliquots of purified proteins (Raa4-His and Rab1-Strep), flow through (F), wash steps 1 and 2 (W1 and W2), and eluate (E) were analyzed via SDS-PAGE.

environmental conditions to ensure a wide range of cDNA representation in the total cDNA population. Clones were selected on selective synthetic dropout (SD) medium and were further characterized through qualitative and quantitative lacZ tests. Ninety-eight clones representing nine different cDNA sequences were detected with this approach. Interestingly, 66 out of 98 clones contained cDNA fragments encoding truncated forms of an uncharacterized protein (transcript Cre03.g157050.t1.3) that was designated Rab1 (*Raa*4-*b*inding *1*; Fig. 2*A*). For further analysis, plasmid pGADT7_Rab1-B (carrying one of the *RAB1* cDNA fragments) was reintroduced into yeast and mated against strains carrying genes encoding Raa4 subfragments (Raa4-FL, Raa4-A, Raa4-B) fused to the GAL4 DNA-binding domain. To exclude the possibility of transactivation, all yeast strains were mated against control strains carrying plasmids with either the empty GAL4 DNA-binding domain or a GAL4 activation domain. Growth tests on selective SD medium clearly confirmed the sole interaction of Raa4 and Rab1 subfragments.

We further investigated the interaction between Raa4 and Rab1 using *in vitro* pull-down assays. Because of the low-level expression of full-length open reading frames in *E. coli*, subfragments of the corresponding proteins were used for all *in vitro* experiments. A truncated version of *RAA4* was synthesized in *E. coli* as a 30.3-kDa His-tag fusion protein (Raa4-His) and purified via Ni-NTA affinity chromatography, and a 44.5-kDa truncated Rab1 variant was synthesized as a One-STrEP-tag fusion protein (Rab1-Strep) and purified via Strep-Tactin affinity chromatography (Fig. 2*B*). Purified proteins or, alternatively, crude *E. coli* extracts containing both proteins were incubated with each other and then repurified using Ni-NTA resin. SDS-PAGE and immunoblot analysis of the eluates demonstrated that Rab1-Strep co-purified with Raa4-His, indicating a strong interaction between the proteins (Figs. 2*C* and 2*D*, supplemental Fig. S5). Rab1-Strep incubated with Ni-NTA resin and His-tagged chloroplast recognition particle cpSRP served as controls to rule out the possibility of unspecific interactions.

Rab1 Localizes to the Chloroplast and Binds to tscA RNA—Cloning and sequencing of the complete *RAB1* cDNA revealed an exon/intron organization that is in agreement with the annotated gene structure, and these data underlie the

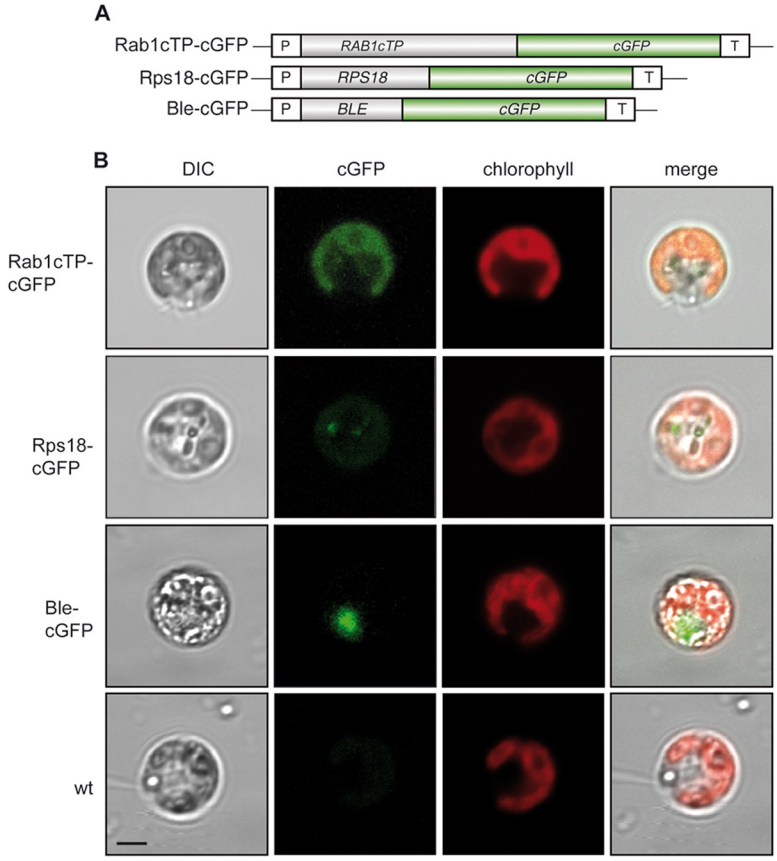

FIG. 3. **Laser scanning confocal fluorescence microscopy (LSCFM) to demonstrate chloroplast localization of Rab1 protein.** *A*, schematic drawing of fusion protein constructs used for cGFP assays. *B*, LSCFM of *C. reinhardtii* arg⁻cw15 transformants. Transformants were analyzed by means of differential interference contrast microscopy (DIC) or confocal fluorescence microscopy. DIC, cGFP fluorescence (green), and chlorophyll autofluorescence (red) were merged as indicated. Scale bar represents 5 μm. *BLE*, phleomycin resistance gene of *Streptoalloteichus hindustanus*; *cGFP*, synthetic GFP; P, *HSP70A/RBCS2* promoter; *RAB1cTP*, 5′-region containing chloroplast signal sequence of *RAB1*; *RPS18*, cytoplasmic ribosomal protein S18 of *C.reinhardtii*; T, 3′-UTR of *LHCB1* or of *RBCS2* gene.

gene structure depicted in supplemental Fig. S6. *RAB1* gives rise to a predicted protein of 1216 amino acids with a molecular weight of 124.3 kDa (Fig. 2*B*). Further, secondary structure analysis indicated that Rab1 is an α-helical (43.67%) protein with 7.64% acidic amino acids and 8.71% basic amino acids with a predicted pI of 6.46. Pattern and profile searches revealed no distinct protein domains or functional motifs, with the exception of glycine-rich (residues 508–587), glutamine-rich (residues 721–742), proline-rich (residues 860–945), and alanine-rich (residues 761–1135) profiles.

Although *in silico* tools for the prediction of subcellular localization did not localize Rab1 to a specific subcellular compartment, two putative chloroplast cleavage sites occur within the first 33 amino acids of the N-terminus between positions 29 and 30 (Val-Glu-Ala29 ↓ Arg30) or 33 and 34 (Val-Arg-Ala33 ↓ Val34). Of note is that in *C. reinhardtii* the amino acid sequence Val-X-Ala is a well-conserved motif at position -3 to -1 relative to the cleavage site of transit peptides that target proteins to the chloroplast stroma (45). Additionally, the length of the putative transit peptide (29 or 33 amino acids) is close to the mean length (29 amino acids) of other *C. reinhardtii* chloroplast transit peptides (45).

To verify the subcellular localization of Rab1 *in vivo*, a *RAB1cPT*::c*GFP* fusion construct under the control of the *HSP70A/RBCS2* tandem promoter (46) was expressed in *C. reinhardtii* (Fig. 3). Laser scanning confocal fluorescence microscopy revealed the co-localization of the chimeric Rab1cTP::cGFP fusion protein with the chloroplast autofluorescence, indicating that Rab1 is localized in the chloroplast. We used the ribosomal protein Rps18 fused to cGFP as a cytoplasmic control (18). *C. reinhardtii* strains transformed with this fusion construct exhibited GFP fluorescence that clearly localized outside the chlorophyll autofluorescence of the chloroplast. A Ble::cGFP fusion protein served as a control for nuclear localization.

Several putative RNA-binding residues spread over the entire protein sequence are predicted by BindN (30) and RNABindR (31). Electromobility shift assays were conducted to evaluate the RNA-binding properties of Rab1; subfragment Rab1-Strep was incubated with radioactively labeled RNA comprising domains D2 (155 nt), D3 (192 nt), D2+D3 (337 nt), or D5+D6 (102 nt) of *psaA* intron 1 (Fig. 1*A*). The RNA–protein complexes were separated on native polyacrylamide gels and analyzed. Histidine-tagged cNAPL (40.8 kDa), a chloroplast RNA-binding protein (18), or GST-tagged Raa4 fusion protein (66.9 kDa) (14) and a functionally unrelated One-STrEP-tag fusion protein (43.9 kDa) served as positive and negative controls, respectively. Bandshifts were observed when Rab1-Strep was incubated with domain D2 or D3 of *tscA*, whereas no binding to domains D5 and D6 was observed (Fig. 4*A*). Competition experiments showed that binding of Rab1-Strep to *tscA*D2+D3 is specific, as incubation with unlabeled non-

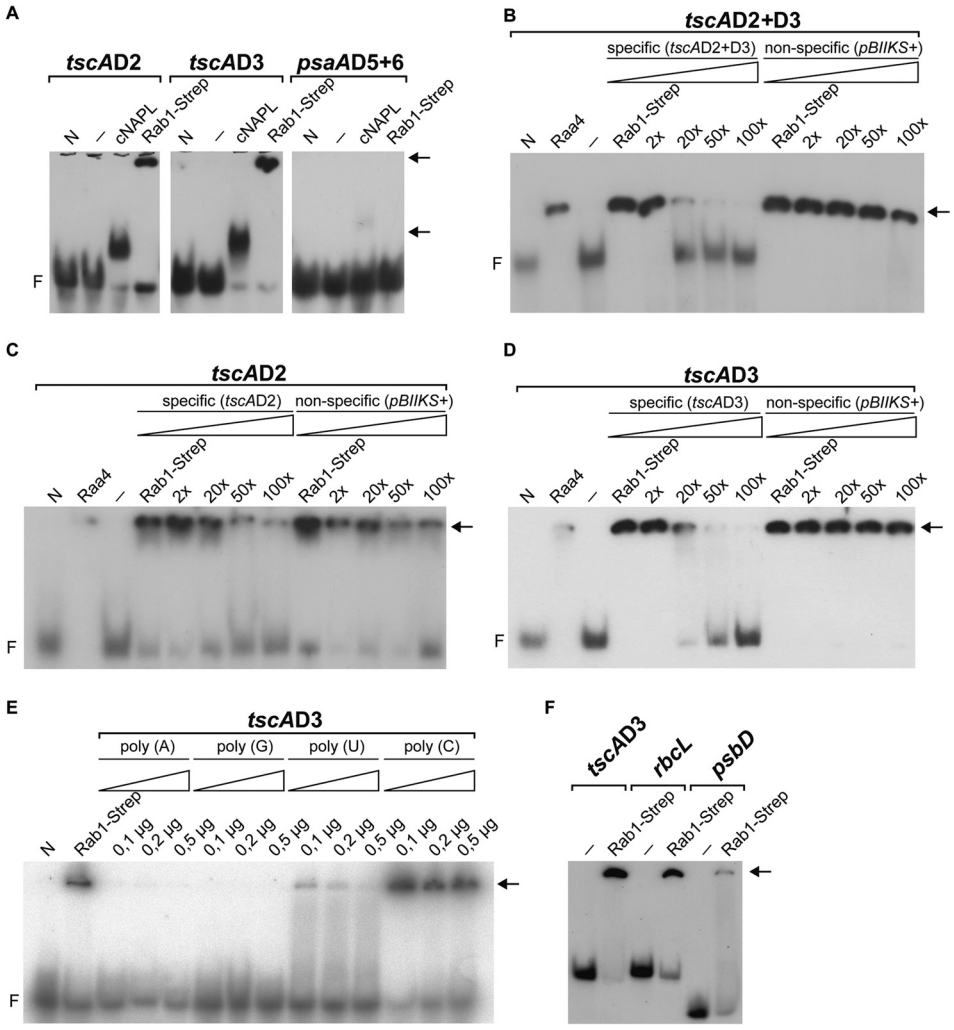

FIG. 4. ***In vitro* binding of Rab1 to representative intron domains of *tscA* RNA.** *A*, 5 µg Rab1-Strep were incubated with 30 fmol labeled *tscA*D2, *tscA*D3, and *psaA*D5+6 transcript and separated on a 5% native polyacrylamide gel. An unrelated One-STrEP-tag protein (N) was used as a negative control, and His-tagged cNAPL (18) was used as a positive control. Lanes marked with "–" represent labeled transcript without the addition of protein. Arrows indicate shifted bands, and F indicates free RNA. *B–D*, competition assays of 5 µg Rab1-Strep and 15 fmol radioactive probes of internally labeled intron RNA (*tscA*D2+3, *tscA*D2, *tscA*D3) and excess of cold specific (*tscA*D2+3, *tscA*D2, *tscA*D3) and nonspecific competitor RNA (pBSIIKS+). Lanes beneath triangles are the 2-, 10-, 50-, and 100-fold molar excess of the competitor. An unrelated One-STrEP-tag protein (N) served as a negative control, and GST-tagged Raa4 (14) was used as a positive control. Lanes marked with "–" represent labeled transcript without the addition of protein. Free RNA, negative control, and shifted bands are indicated as in *A*. *E*, competition experiments with Rab1-Strep (2 µg). EMSA of domain D3 of *tscA* RNA (15 fmol) with Rab1 in the presence of unlabeled RNA homopolymer competitors as indicated. *F*, chloroplast transcripts (15 fmol) corresponding to fragments of *tscA*D3, *rbcL*, and *psbD* were incubated with Rab1-Strep (2 µg) as indicated.

specific competitor RNAs, derived from *in vitro* transcription with plasmid pBIIKS+ (121 nt), had no effect on the formation of the *tscA*-Rab1complex (Fig. 4*B*). To analyze differences in binding specificities, competition analyses with RNA comprising either domain D2 or domain D3 were performed. The addition of a 50-fold molar excess of unlabeled specific competitor transcript led to a substantial loss of the Rab1-*tscA*D3 complex (Fig. 4*D*). Incubation with the same amount of specific competitor transcript had only a minor effect on the formation of the Rab1-*tscA*D2 complex (Fig. 4*D*). To study sequence preferences, we compared the ability of A, G, U, or C RNA homopolymers to compete for binding of Rab1-Strep to the labeled D3 domain. Even low amounts of poly(C) abolished Rab1 binding, but the same or excess amounts of poly(G) or poly(A) had no significant effect on complex formation (Fig 4*E*). The use of poly(U) reduced the formation of the Rab1-*tscA*D3 complex. The competition assays suggest that Rab1-Strep preferentially interacts with C-rich sequences and binds to a lesser extent to U-rich sequences. We further investigated the binding of Rab1-Strep to transcripts of two chloroplast genes. As shown in Fig. 4*F*, Rab1 binds to chloroplast *rbcL* RNA and to a lesser extent to *psbD* RNA.

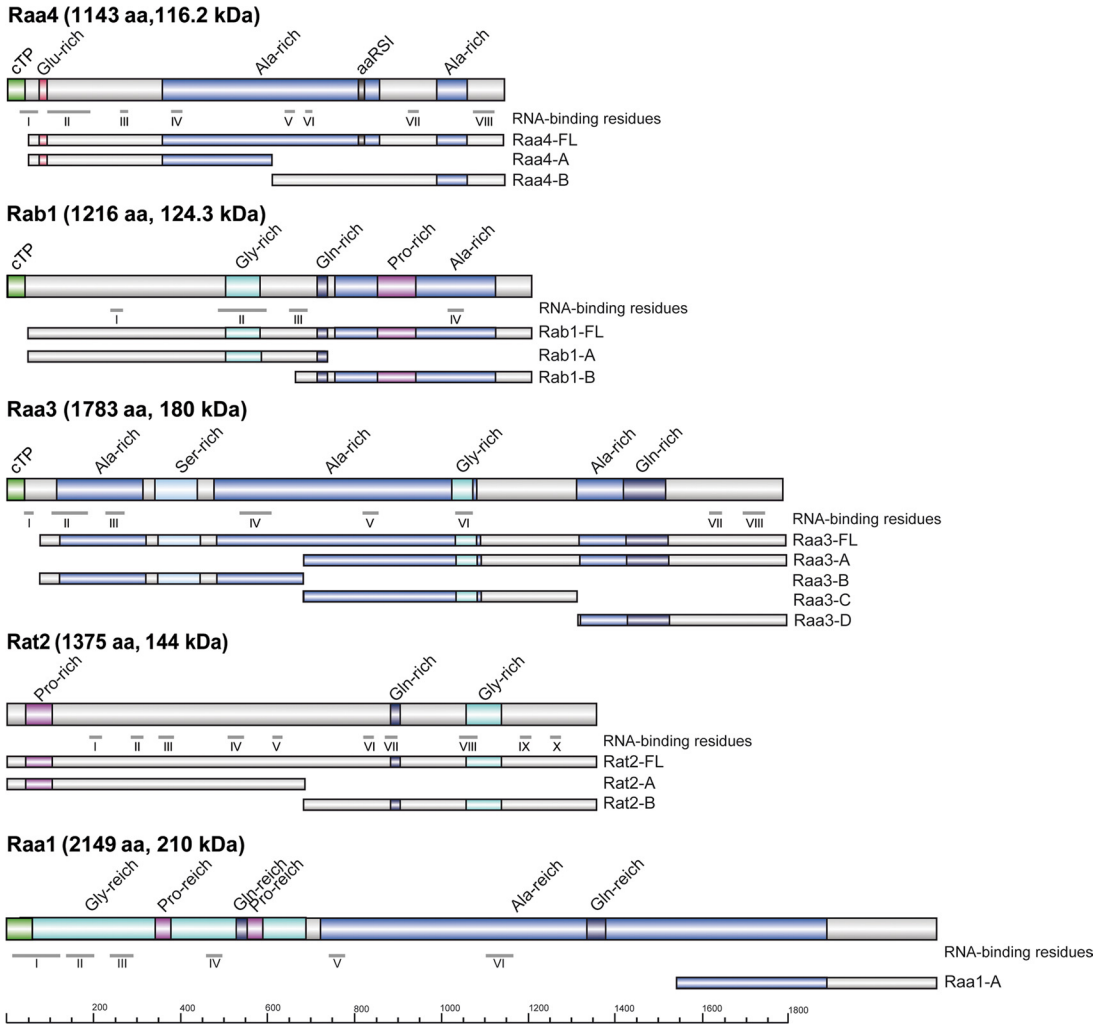

FIG. 5. **Protein subfragments used in two-hybrid assays.** Depicted are primary structures of Raa4, Rab1, Raa3, Rat2, and Raa1. RNA-binding residues were predicted using BindN and RNABindR. Raa4 shows loose homology to an aminoacyl tRNA synthetase class I signature (aaRS1). cTP, chloroplast transit peptide; Gly-rich, Glu-rich, Pro-rich, Ala-rich, and Ser-rich, glycine-, glutamine-, proline-, alanine-, and serine-rich domains, respectively.

psaA Trans-splicing Factors Show Protein–Protein Interactions in a Heteromeric Protein Complex—Yeast two-hybrid assays were carried out to test direct interactions between *trans*-splicing factors and the *tscA*-binding protein Rab1. Therefore, full-length versions and derivatives of Raa4, Rab1, Raa1, Raa3, and Rat2 were fused to either the GAL4 activation domain or the GAL4 DNA-binding domain (Fig. 5).

Yeast strains carrying the above-mentioned fragments were mated, and the diploids were selected for on selective SD medium lacking tryptophan and leucine (diploids will carry both plasmids and will therefore survive in the absence of those nutrients) (supplemental Fig. S7). Diploids carrying Rab1 and Raa4, Rab1 and Raa3, Rab1 and Rat2, or Rab1 and Raa1 fusion proteins also exhibited growth on selective SD medium lacking tryptophan, leucine, adenine, and histidine. Growth on this medium indicates interaction between the two fusion proteins (Fig. 6*A*). In addition, growth of diploids carrying Rat2 and Rab1, Rat2 and Raa4, Rat2 and Raa3, or Rat2 and Raa1 was detected. Interestingly, strains expressing both Rat2 fusion proteins were also able to grow. Furthermore, growth of diploids carrying Raa4 and Raa1 fragments was observed. In contrast, no growth of strains carrying Raa4 and Raa3 fusion proteins was detected.

Taken together, these results confirm the interaction between Raa4 and Rab1. Furthermore, direct binding between Rab1, Rat2, or Raa1 and all tested *psaA trans*-splicing factors as well as interaction of Rat2 with itself was observed (Fig. 6*B*). Thus, these proteins, together with the 19 chloroplast components identified in the mass spectrometry analysis, most probably form a multisubunit complex.

DISCUSSION

More than 200 components make up the eukaryotic nuclear spliceosome including the five snRNAs U1, U2, U4, U5, and

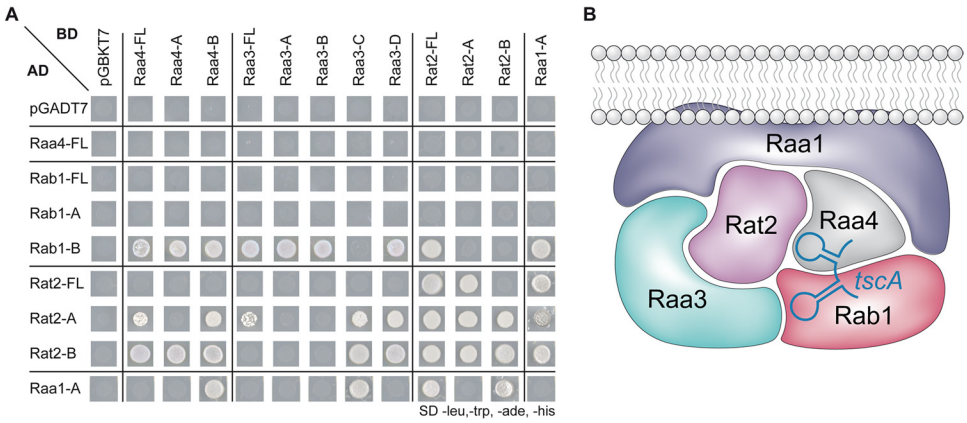

FIG. 6. **Protein–protein and protein–RNA interactions between *trans*-splicing factors.** *A*, strains carrying different subfragments of Raa4, Rab1, Raa3, Rat2, and Raa1 fused to the GAL4 activation (AD) or DNA-binding domain (BD) were mated and spotted onto SD medium lacking leucine, tryptophane, adenine, and histidine. Growth on SD medium lacking leucine, tryptophane, adenine, and histidine reflects the interaction between proteins fused to the GAL4 activation and DNA-binding domains. To exclude the possibility of transactivation, all yeast strains were mated against control strains carrying either the empty GAL4 DNA-binding domain (pGADT7) or the GAL4 activation domain plasmid (pGBKT7). *B*, schematic drawing representing direct protein–protein interactions (based on yeast two-hybrid data) between factors that are involved in the *trans*-splicing of *psaA* intron 1. Direct binding of Raa4 and Rab1 to *tscA* intron RNA was demonstrated by EMSAs. Raa1 is also involved in the splicing of *psaA* intron 2 and was detected in a membrane-associated complex (10).

U6. Protein–protein and protein–RNA interactions play an important role in this multi-megadalton machinery, as it consists predominantly of proteins (47). Large ribonucleoprotein particles also exert critical functions in chloroplast splicing and have been described in several organisms, including *Z. mays* and *Arabidopsis thaliana* (15, 48). *Trans*-splicing of the fragmented *psaA* gene of *C. reinhardtii* is likewise mediated by high molecular weight ribonucleoprotein complexes (3, 15). In this investigation, we have examined the intricate protein network involved in *psaA* splicing via an experimental approach that combines diverse methods to identify novel subunits and to study protein–protein and protein–RNA interactions.

TAP has proven useful for the purification of protein complexes and the analysis of protein interactions in organisms such as *Saccharomyces cerevisiae*, *A. thaliana*, and *Aspergillus nidulans* (25–27). Because the expression of foreign genes in *C. reinhardtii* is often poor as a result of inappropriate codon usage (28), we generated a codon-optimized variant of the TAP tag and placed the *RAA4::TAP tag* fusion gene under control of the *HSP70A/RBCS2* tandem promoter (18). For TAP we used cells adapted to three different environmental conditions (light, dark, and anaerobiosis) to define the basic components of the *psaA* splicing complex.

Using this approach, we co-purified several yet uncharacterized proteins with Raa4, including two proteins that exhibit multiple OPR motifs. OPRs are found in several proteins that have functions in the post-transcriptional regulation of chloroplast gene expression, as, for instance, in the chloroplast translation factors Tbc2 or Tab1 from *C. reinhardtii*, but also in the *psaA trans*-splicing factors Raa1 and Rat2 (44, 49). They show a degenerate consensus sequence comprising the amino acids PPPEW, and it is suggested that they fold into arrayed α-helices. This places them into the helical repeat superfamily that includes, for example, tetratricopeptide repeat and pentatricopeptide repeat proteins that have diverse functions in RNA metabolism (5, 50). Furthermore, we identified a protein exhibiting a CRM domain. The CRM domain is an RNA-binding module that appears to be restricted to archea, bacteria, and plants. In plants it occurs in a protein family having multiple copies of these domains. In *Z. mays* and *A. thaliana*, several members of the CRM family with a function in splicing have been described. However, further experiments are necessary to validate the interaction of Raa4 with these proteins and to analyze their involvement in *psaA*-splicing.

A remarkable result of our investigation was the detection of the functionally characterized splicing factors Raa1, Raa3, and Rat2, which all are involved in splicing of the first group II intron of the *psaA* precursor RNA (9, 10, 51) (Table I). Raa1 is also involved in the splicing of second group II intron of *psaA* pre-mRNA (52). Raa1 and Rat2 play a role in *tscA* processing, whereas Raa3 functions in intron 1 splicing. We therefore propose that the co-purified proteins are components of a complex that couples the protein machineries for *tscA* maturation and *psaA* intron 1 splicing. Sedimentation and co-fractionation experiments have demonstrated that *trans*-splicing factors are organized in high molecular weight ribonucleoprotein complexes, the exact composition of which is not clearly defined. Raa3, for instance, was identified in a soluble, stroma-localized ribonucleoprotein complex of about 1700 kDa that is associated with *tscA* RNA and *psaA* exon 1 precursor transcripts (9). The membrane-associated Raa1 protein, however, was identified in a complex of 670 kDa together with unidentified RNAs (10). Considering the molecular weights of these complexes, it is likely that unknown

components might exist. Further, a 400 to 500 kDa membrane complex that contains Raa1 and Raa2 has been described (53). Therefore, it is also possible that the splicing of intron 2 involves a different protein complex including Raa1, Raa2, and other proteins. This is consistent with our MudPIT results, from which we were unable to detect Raa2 in the purified complex. Splicing complexes may be dynamic, with their compositions changing as a result of the addition and loss of proteins. This possibility is supported by the observation that Raa1 and Raa3 co-purify with Raa4, although sedimentation analyses have shown that Raa3 is found in the chloroplast stroma, whereas Raa1 is associated with membranes (9, 10).

The presence of Raa1 suggests that this *trans*-splicing factor—among others—could form the core of a spliceosome-like complex that is capable of recruiting single intron-specific splicing factors. In other systems, organellar splicing factors have already been described that carry out functions on a broad range of transcripts (7, 54). Thus, unlike the nuclear spliceosome, the chloroplast splicing complex is an evolutionarily "young" complex that probably contains a core of only a few subunits common for different organelle transcripts. During splicing of single organellar introns, the "core complex" recruits several specific splicing factors that assemble to form a functional splicing complex.

To gain a deeper understanding of the direct protein–protein interactions involved in *psaA* splicing, we applied a yeast two-hybrid screen using Raa4 as bait. The extensive yeast two-hybrid screen used in this work led to the discovery of the Raa4 interaction partner Rab1. *In vitro* pull-down assays indicated strong and direct binding of Rab1 to Raa4. As a *psaA trans*-splicing factor, Rab1 is expected to localize within the chloroplast, which we confirmed using a *RAB1cTP*::*cGFP* fusion construct. Rab1 shows no significant sequence homology to other proteins, but it exhibits three low-complexity regions and several scattered, putative RNA-binding residues. Although low-complexity regions are quite abundant in proteins, recent systematic approaches have indicated that they occur frequently in proteins associated with the regulation of gene expression. These regions are believed to function in protein–protein interactions (55); however, recent observations have shown that they also participate in RNA recognition (56). Interestingly, low-complexity regions also occur in other *psaA*-splicing factors such as Raa4, which harbors alanine- and glutamine-rich regions (14). Another similarity to Raa4 is the absence of typical RNA-binding domains such as the K homology domain, the RNA recognition motif, the CRM domain, and pentatricopeptide repeats (57, 58). Nevertheless, we demonstrated the direct binding of Rab1 to *psaA* intron RNA via electromobility shift assays. Experiments with further RNAs revealed that Rab1 also interacts with chloroplast transcripts, and competition experiments with RNA homopolymers show preferential binding with poly(C) and poly(U). We are aware that a truncated variant of Rab1 has been used and thus might show an altered RNA binding property relative to the full-size protein. *In vivo*, specific protein–RNA binding might require the cooperation of other proteins such as Raa4.

Future experiments will test whether Rab1 can also be considered a *trans*-splicing factor. Artificial miRNA experiments to down-regulate *RAB1* expression, however, have already indicated that the *RAB1* transcript level is rather low; we failed to get any strain in more than 800 transformants that showed a down-regulated *RAB1* gene.[2] Therefore, one has to await knock-out libraries for *C. reinhardtii* in order to get a functional analysis of the *RAB1* gene.

For the yeast two-hybrid analyses, we used full-length variants as well as derivatives of splicing factors to increase the detection and sensitivity of interactions. This was done because fusion proteins frequently cannot fold correctly in yeast and thus are incapable of interacting. Moreover, full-length proteins can be locked in a "closed" formation that masks binding domains (59, 60). This effect might explain why, for example, subfragment Rab1-B interacts with several fusion proteins, whereas its full-length variant Rab1-FL fails to bind the respective proteins. Another aspect that has to be considered with respect to yeast two-hybrid experiments is the possibility of false-positive interactions. These nonspecific interactions are of diverse origins. In many cases, the source of these interactions is the high expression level of fusion proteins or their localization in a subcellular compartment that does not correspond to the proteins' natural environment (61). However, we were able to verify the direct interaction of Raa4 and Rab1 with pull-down assays. Moreover, TAP with Raa4 as bait points toward direct or indirect interactions between the tested splicing factors. Our analyses demonstrated that Rab1, Rat2, or Raa1 interacts directly with all tested *psaA trans*-splicing factors. Moreover, our data indicate that Rat2 interacts with itself and possibly forms a homodimer or a homomultimer. Future work will have to determine whether these interactions occur simultaneously or successively.

Pairs of direct protein–protein interactions demonstrated by yeast-two hybrid assays have been described as well for other chloroplast group II intron splicing factors. A direct interaction with CAF1 and CAF2 was described for CRS2 from *Z. mays* (62). It has been assumed that the CAF proteins create an RNA-binding platform that recruits CRS2. We propose that some *C. reinhardtii* splicing factors also function as protein adaptors or scaffolding proteins to provide a platform for the binding of factors directly involved in the splicing of *psaA* RNA. Scaffold proteins have a central role in signaling pathways because they act as protein-binding platforms for signaling components such as kinases. Scaffolding proteins also participate in the nuclear spliceosome. For example, the large protein Prp8 interacts with the pre-mRNA, with the U5 and U6 snRNAs, and with several other spliceosomal proteins

[2] V. Kock and O. Reifschneider, personal communication.

and is thought to be the master regulator of the spliceosome (63).

A significant finding of our investigation is the detection of *tscA*-RNA in the affinity purified protein complex. This result indicates further that the splicing complex represents a ribonucleoprotein complex involved in *trans*-splicing. This is consistent with the identical binding preferences of Rab1 and Raa4, with both interacting with *tscA* domains D2 and D3 (14). The specific binding of splicing factors to their target RNAs has been described in several cases (62, 64). It is assumed that these proteins participate in intron folding by stabilizing functionally active structures and folding intermediates (65). It is thus possible that Rab1 has a similar stabilizing effect on *psaA* intron 1.

Future work will focus on identifying additional participating components to allow us to define in detail the ribonucleoprotein complexes involved in group II intron *trans*-splicing. Thus, because these complexes functionally resemble the nuclear spliceosome, the elucidation of their precise composition and structure is of particular interest from an evolutionary standpoint, as group II introns are proposed to be the ancestors of nuclear spliceosomal introns (2, 66).

Acknowledgments—We thank Katja Schmitt and Ingeborg Godehardt for their excellent technical assistance. We also thank Dr. Beatrix Dünschede and Prof. Dr. Danja Schünemann for support with *in vitro* pull-down experiments, and Dr. Anja Hemschemeier and Prof. Dr. Thomas Happe for support with the anaerobic adaption of *C. reinhardtii* cultures.

* This work was funded by the Deutsche Forschungsgemeinschaft (SFB480 B3, KU 517/13-1, JA 2296/1-1).

S This article contains supplemental material.

‖ To whom correspondence should be addressed: Tel.: +49 234 3226212, Fax: +49 234 3214184, E-mail: ulrich.kueck@ruhr-uni-bochum.de.

§ Authors contributed equally to this work.

REFERENCES

1. Sharp, P. A. (1991) "Five easy pieces". *Science* **254,** 663
2. Lambowitz, A. M., and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb. Perspect. Biol.* **3,** a003616
3. Glanz, S., and Kück, U. (2009) *Trans*-splicing of organelle introns—a detour to continuous RNAs. *BioEssays* **31,** 921–934
4. Jarrell, K. A., Peebles, C. L., Dietrich, R. C., Romiti, S. L., and Perlman, P. S. (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products. *J. Biol. Chem.* **263,** 3432–3439
5. Stern, D. B., Goldschmidt-Clermont, M., and Hanson, M. R. (2010) Chloroplast RNA metabolism. *Annu. Rev. Plant Biol.* **61,** 125–155
6. Barkan, A. (2011) Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant Physiol.* **155,** 1520–1532
7. Kroeger, T. S., Watkins, K. P., Friso, G., van Wijk, K. J., and Barkan, A. (2009) A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 4537–4542
8. Sperling, J., Azubel, M., and Sperling, R. (2008) Structure and function of the pre-mRNA splicing machine. *Structure* **16,** 1605–1615
9. 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
10. Perron, K., Goldschmidt-Clermont, M., and Rochaix, J. D. (2004) A multiprotein complex involved in chloroplast group II intron splicing. *RNA* **10,** 704–711
11. Kück, U., Choquet, Y., Schneider, M., Dron, M., and Bennoun, P. (1987) Structural and transcriptional analysis of two homologous genes for the P_{700} chlorophyll α-apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo trans*-splicing. *EMBO J.* **6,** 2185–2195
12. 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
13. 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
14. Glanz, S., Jacobs, J., Kock, V., Mishra, A., and Kück, U. (2012) Raa4 is a *trans*-splicing factor that specifically binds chloroplast *tscA* intron RNA. *Plant J.* **69,** 421–431
15. Jacobs, J., Glanz, S., Bunse-Grassmann, A., Kruse, O., and Kück, U. (2010) RNA *trans*-splicing: identification of components of a putative chloroplast spliceosome. *Eur. J. Cell. Biol.* **89,** 932–939
16. Hemschemeier, A., Melis, A., and Happe, T. (2009) Analytical approaches to photobiological hydrogen production in unicellular green algae. *Photosynth. Res.* **102,** 523–540
17. Kindle, K. L. (1990) High-frequency nuclear transformation of *Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. U.S.A.* **87,** 1228–1232
18. Glanz, S., Bunse, A., Wimbert, A., Balczun, C., and Kück, U. (2006) A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in *Chlamydomonas reinhardtii. Nucleic Acids Res.* **34,** 5337–5351
19. Jerpseth, B., Greener, A., Short, J. M., Viola, J., and Kretz, P. L. (1992) XL1-Blue MRF′ *E. coli* cells: McrA$^-$, McrCB$^-$, McrF$^-$, Mrr$^-$, HsdR$^-$ derivative of XL1-Blue cells. *Strateg. Mol. Biol.* **5,** 81–83
20. James, P., Halladay, J., and Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144,** 1425–1436
21. Colot, H. V., Park, G., Turner, G. E., Ringelberg, C., Crew, C. M., Litvinkova, L., Weiss, R. L., Borkovich, K. A., and Dunlap, J. C. (2006) A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 10352–10357
22. Becker, D. M. and Lundblad, V. 2001. Introduction of DNA into Yeast Cells. Current Protocols in Molecular Biology. 27:13.7.1–13.7.10.
23. Nowrousian, M., Ringelberg, C., Dunlap, J. C., Loros, J. J., and Kúck, U. (2005) Cross-species microarray hybridization to identify developmentally regulated genes in the filamentous fungus *Sordaria macrospora. Mol. Genet. Genomics* **273,** 137–149
24. Steinle, A., Li, P., Morris, D. L., Groh, V., Lanier, L. L., Strong, R. K., and Spies, T. (2001) Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics* **53,** 279–287
25. Balczun, C., Bunse, A., Schwarz, C., Piotrowski, M., and Kück, U. (2006) Chloroplast heat shock protein Cpn60 from *Chlamydomonas reinhardtii* exhibits a novel function as a group II intron-specific RNA-binding protein. *FEBS Lett.* **580,** 4527–4532
26. Bunse, A. A., Nickelsen, J., and Kück, U. (2001) Intron-specific RNA binding proteins in the chloroplast of the green alga *Chlamydomonas reinhardtii. Biochim. Biophys. Acta* **1519,** 46–54
27. Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Maréchal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riano-Pachon, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S.,

Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martinez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318,** 245–250

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. Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C. V., Hau, J., Martin, O., Kuznetsov, D., and Falquet, L. (2007) MyHits: improvements to an interactive resource for analyzing protein sequences. *Nucleic Acids Res.* **35,** 433–437
30. Wang, L., and Brown, S. J. (2006) BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. *Nucleic Acids Res.* **34,** W243–W248
31. Terribilini, M., Sander, J. D., Lee, J. H., Zaback, P., Jernigan, R. L., Honavar, V., and Dobbs, D. (2007) RNABindR: a server for analyzing and predicting RNA-binding sites in proteins. *Nucleic Acids Res.* **35,** W578–W584
32. Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugiere, S., Hippler, M., Ferro, M., Bruley, C., Peltier, G., Vallon, O., and Cournac, L. (2012) PredAlgo, a new subcellular localization prediction tool dedicated to green algae. *Mol. Biol. Evol.* **29,** 3625–3639
33. Emanuelsson, 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
34. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2,** 953–971
35. Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8,** 785–786
36. Bals, T., Dünschede, B., Funke, S., and Schünemann, D. (2010) Interplay between the cpSRP pathway components, the substrate LHCP and the translocase Alb3: an *in vivo* and *in vitro* study. *FEBS Lett.* **584,** 4138–4144
37. Bayram, O., Krappmann, S., Ni, M., Bok, J. W., Helmstaedt, K., Valerius, O., Braus-Stromeyer, S., Kwon, N. J., Keller, N. P., Yu, J. H., and Braus, G. H. (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. *Science* **320,** 1504–1506
38. Bloemendal, S., Bernhards, Y., Bartho, K., Dettmann, A., Voigt, O., Teichert, I., Seiler, S., Wolters, D. A., Poggeler, S., and Kück, U. (2012) A homolog of the human STRIPAK complex controls sexual development in fungi. *Mol. Microbiol.* **84,** 310–323
39. Wolters, D. A., Washburn, M. P., and Yates, J. R., 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **73**, 5683–5690
40. Maul, J. E., Lilly, J. W., Cui, L., dePamphilis, C. W., Miller, W., Harris, E. H., and Stern, D. B. (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* **14,** 2659–2679
41. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17,** 1030–1032
42. Neupert, J., Karcher, D., and Bock, R. (2009) Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes. *Plant J.* **57,** 1140–1150
43. Bienvenut, W. V., Espagne, C., Martinez, A., Majeran, W., Valot, B., Zivy, M., Vallon, O., Adam, Z., Meinnel, T., and Giglione, C. (2011) Dynamics of post-translational modifications and protein stability in the stroma of *Chlamydomonas reinhardtii* chloroplasts. *Proteomics* **11,** 1734–1750
44. Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J. D. (2012) Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. *Plant J.* **27,** 652–661
45. Franzen, L. G., Rochaix, J. D., and von Heijne, G. (1990) Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* share features with both mitochondrial and higher plant chloroplast presequences. *FEBS Lett.* **260,** 165–168
46. Schroda, M., Blocker, D., and Beck, C. F. (2000) The *HSP70A* promoter as a tool for the improved expression of transgenes in *Chlamydomonas.* *Plant J.* **21,** 121–131
47. Will, C. L., and Lührmann, R. (2006) Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **3**, a003707
48. de Longevialle, A. F., Small, I. D., and Lurin, C. (2010) Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. *Mol. Plant* **3,** 691–705
49. Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y., and Wollman, F. A. (2011) Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of atpA transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Plant J.* **67,** 1055–1066
50. Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y. S., Bond, C. S., and Small, I. (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* **8**, e1002910
51. Balczun, C., Bunse, A., Hahn, D., Bennoun, P., Nickelsen, J., and Kück, 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
52. Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J. D., and Goldschmidt-Clermont, M. (2006) A novel multifunctional factor involved in *trans*-splicing of chloroplast introns in *Chlamydomonas. Nucleic Acids Res.* **34,** 262–274
53. Perron, K., Goldschmidt-Clermont, M., and Rochaix, J. D. (1999) A factor related to pseudouridine synthases is required for chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii. EMBO J.* **18,** 6481–6490
54. Asakura, Y., and Barkan, A. (2007) A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *Plant Cell* **19,** 3864–3875
55. Coletta, A., Pinney, J. W., Solis, D. Y. W., Marsh, J., Pettifer, S. R., and Attwood, T. K. (2010) Low-complexity regions within protein sequences have position-dependent roles. *BMC Syst. Biol.* **4,** 43
56. Kala, S., and Salavati, R. (2010) OB-fold domain of KREPA4 mediates high-affinity interaction with guide RNA and possesses annealing activity. *RNA* **16,** 1951–1967
57. Jacobs, J., and Kück, U. (2011) Function of chloroplast RNA-binding proteins. *Cell. Mol. Life Sci.* **68,** 735–748
58. Schmitz-Linneweber, C., and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13,** 663–670
59. Boxem, M., Maliga, Z., Klitgord, N., Li, N., Lemmens, I., Mana, M., de Lichtervelde, L., Mul, J. D., van de Peut, D., Devos, M., Simonis, N., Yildirim, M. A., Cokol, M., Kao, H. L., de Smet, A. S., Wang, H., Schlaitz, A. L., Hao, T., Milstein, S., Fan, C., Tipsword, M., Drew, K., Galli, M., Rhrissorrakrai, K., Drechsel, D., Koller, D., Roth, F. P., Iakoucheva, L. M., Dunker, A. K., Bonneau, R., Gunsalus, K. C., Hill, D. E., Piano, F., Tavernier, J., van den Heuvel, S., Hyman, A. A., and Vidal, M. (2008) A protein domain-based interactome network for *C. elegans* early embryogenesis. *Cell* **134,** 534–545
60. Koch, M. R., and Pillus, L. (2009) The glucanosyltransferase Gas1 functions in transcriptional silencing. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 11224–11229
61. Bruckner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int. J. Mol. Sci.* **10,** 2763–2788
62. Ostheimer, G. J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. (2003) Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* **22,** 3919–3929
63. Grainger, R. J., and Beggs, J. D. (2005) Prp8 protein: at the heart of the spliceosome. *RNA* **11,** 533–557
64. Singh, R. N., Saldanha, R. J., D'Souza, L. M., and Lambowitz, A. M. (2002) Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J. Mol. Biol.* **318,** 287–303
65. Pyle, A. M., Fedorova, O., and Waldsich, C. (2007) Folding of group II introns: a model system for large, multidomain RNAs? *Trends Biochem. Sci.* **32,** 138–145
66. Cech, T. R. (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **44,** 207–210