The Chlamydomonas reinhardtii chloroplast-localized poly(A)-binding protein RB47 is predicted to contain a non-conserved linker (NCL) sequence flanked by highly conserved N- and C-terminal sequences, based on the corresponding cDNA. RB47 was purified from chloroplasts in association with an endoribonuclease activity; however, protein sequencing failed to detect the NCL. Furthermore, while recombinant RB47 including the NCL did not display endoribonuclease activity in vitro, versions lacking the NCL displayed strong activity. Both full-length and shorter forms of RB47 could be detected in chloroplasts, with conversion to the shorter form occurring in chloroplasts isolated from cells grown in the light. This conversion could be replicated in vitro in chloroplast extracts in a light-dependent manner, where epitope tags and protein sequencing showed that the NCL was excised from a full-length recombinant substrate, together with splicing of the flanking sequences. The requirement for endogenous factors and light differentiates this protein splicing from autocatalytic inteins, and may allow the chloroplast to regulate the activation of RB47 endoribonuclease activity. We speculate that this protein splicing activity arose to post-translationally repair proteins that had been inactivated by deleterious insertions or extensions.

The chloroplast transcriptome is highly dependent on RNA processing, because pervasive activities has to be balanced with activities that discard nonfunctional or deleterious transcripts (1). Among the most prominent of these activities are endo- and exoribonucleases that exert RNA quality control and create mature RNA termini (2–4). We previously examined the mechanism of 3' end maturation for the Chlamydomonas reinhardtii chloroplast atpB mRNA, whose 3' terminus is defined by a stem-loop structure. The atpB 3'-extended precursor undergoes two-step maturation, in which an endonuclease cleaves at a specific site called the endonuclease cleavage site (ECS) (2) ~10 nucleotides (nt) distal to the stem-loop, followed by exonucleolytic trimming (5). Of particular interest was that the sequence downstream of the ECS was extremely rapidly degraded both in vivo and in vitro, even in ectopic contexts (6). We sought to purify this novel ribonuclease activity from Chlamydomonas chloroplasts, with the purification assay based on endoribonucleolytic cleavage near the ECS.

During the course of purification, we identified an endoribonuclease activity that appeared to cleave the RNA substrate near the ECS. As shown below, the endoribonuclease we isolated turned out to be RB47, a member of the polyadenylate-binding protein family (PABP). RB47 had been discovered previously in Chlamydomonas chloroplasts through affinity purification of proteins binding to the psbA mRNA 5'-UTR (7, 8). Based on a variety of correlative experiments, RB47 was proposed to regulate psbA translation in response to light (9–11) and redox poise (12). RB47, for which the isolated cDNA encoded a protein of 68 kDa (80 kDa by SDS-PAGE), could be converted during in vitro chloroplast import to the 47 kDa size found in chloroplasts (7), a difference far in excess of what can be accounted for by removal of the short N-terminal transit peptide (8). This size difference was proposed to be the result of a proteolytic cleavage, which removed ~30 kDa from the C terminus. Here we demonstrate that this conversion in fact results from splicing following the removal of an internal RB47 precursor sequence, the non-conserved linker. This splicing requires both light and chloroplast factors, unlike the spontaneous splicing of inteins (13), and additionally unmasks the otherwise-cryptic endonuclease activity of RB47. We suggest that protein splicing of this type is an underappreciated generator of protein diversity, both within and outside of organelles.

**Results**

*Purification and Identification of Endoribonuclease Activity—*An in vitro assay using total chloroplast soluble proteins had previously been established for an endoribonuclease activity cleaving at the atpB 3'-UTR ECS (5). To identify the protein(s) involved we used sequential chromatography steps (Fig. 1A) and assayed column fractions for endoribonuclease activity using the previously described RNA substrate (Fig. 1B). A final gel filtration column showed that the peak of activity eluted at ~50 kDa (Fig. 1C). The activity generated an endoribonuclease product consistent with cleavage near the ECS (band 1), as well as several smaller products (bands 2 and 3). A size marker similar in size to band 1 was generated by processing the same substrate with a spinach chloroplast extract, which trims exonucleolytically to the ECS (Sp), and a no protein control was included (No). Subsequent 5' and 3'-end labeling showed that band 1 in fact represents the downstream endonuclease cleavage product that is normally rapidly degraded in total soluble chloroplast protein extracts, suggesting that the
Chloroplast Protein Splicing

endoribonuclease had been separated from that activity and furthermore, that the purified protein was cleaving near, but not at the ECS (see below).

To identify candidate proteins for the activity peaking near 50 kDa, the peak gel filtration fractions were subjected to denaturing gel electrophoresis and bands in the 50-kDa size range, which were the dominant species in these fractions, were excised and submitted for analysis by mass spectroscopy (MS).

FIGURE 1. **Endoribonuclease activity and peptide mapping of RB47 purified from chloroplasts.** A, outline of the purification scheme beginning with soluble chloroplast lysate. B, diagram of the 32P-labeled *atpB* 3′-UTR RNA precursor substrate used for in vitro assays, and the products expected from cleavage at the ECS (arrowhead). C, gel filtration column fractions were assayed for endoribonuclease activity. Sp, spinach chloroplast lysate was used to generate a 206 nt size marker by exonuclease degradation of the precursor; No, buffer only. RNase cleavage products discussed in the text are labeled at right. The elution profile of molecular mass standards is indicated below the gel. D, diagram of *C. reinhardtii* RB47 indicating the location and approximate size of the TP, the four RNA recognition motifs (RRM1–4), the NCL and CCT. Amino acids are numbered below, and peptides identified by MS are shown in red. E, alignment of RNA sequence data from genomes.ucsc.edu/Cre454/ and the RB47 gene model from the most recent release of the *Chlamydomonas* genome (the gene location is chromosome_1: 5372774–5378221) as given on www.phytozome.net. The black line denotes the extent of contiguous sequences that encode the indicated domains of RB47.

As shown in Fig. 1D, only peptides from the RRM and CCT domains were found in our MS analysis, with the only significant stretches of predicted amino acid sequence not detected being the chloroplast transit peptide (TP), as expected, and the NCL. To exclude that the NCL was not removed at the level of RNA splicing, we examined the collective results of RNA deep sequencing available for *Chlamydomonas* (Fig. 1E). This analysis showed that while there is evidence for alternative splicing, namely partial inclusion of the intron preceding the NCL, the NCL itself is expressed in cDNA at an equivalent level to all other exonic sequences. Therefore, it is highly improbable that any significant amount of RB47 mRNA lacks sequences encoding the NCL. These results, together with previously published results showing conversion of full-length RB47 to a 47-kDa form under *in vitro* chloroplast import conditions (8) suggested that the NCL domain is removed upon translocation into the chloroplast, and that the CCT domain is then somehow joined to the RRM domains.

**Domain Analysis of Endonucleolytic Activity—Because no** PABP has been previously shown to exhibit endoribonuclease activity, we wished to ascertain if recombinant RB47 possessed this capacity. To do so, we tested several versions of recombinant RB47, as described in detail under “Experimental Procedures” and in Fig. 2A. These included a glutathione-S-transferase (GST) fusion to the full-length protein predicted from cDNA (GST-RB47), an internal deletion lacking most of the NCL (missing sequence underlined in Fig. 2A) and approximating the version we had sequenced (GST-RB47 ID), and a truncated version (with the truncation beginning with the start of the *underlined* sequence) lacking the NCL and CCT (GST-RB47-CTD). A GST-CCT fusion (with the CCT sequence in *italics* in Fig. 2A) that served as a source of antigen for production of RB47 antibodies was also tested. The recombinant proteins were purified along with a GST-only control (Fig. 2B), and the specificity of the CCT antibody was confirmed by immunoblotting (Fig. 2C).

We next tested the recombinant RB47 proteins described above and diagrammed in Fig. 3A for endoribonuclease activity (Fig. 3B), along with spinach chloroplast extract and no protein controls as described for Fig. 1C (*lanes 1 and 2*). The results showed that while neither GST alone nor GST fused to full-length RB47 had detectable activity (*lanes 4 and 5*), the GST-RB47 proteins with internal (ID) or C-terminal (CTD) deletions had robust activity (*lanes 6 and 7*). The reaction products recapitulated the pattern seen when assaying the gel filtration peak fractions (Fig. 1A). On the other hand, the CCT alone (lane 8) did not display activity. Both 5′- and 3′-end labeled substrates were used to identify the RNA cleavage products produced by recombinant RB47 CTD (Fig. 3C). This analysis showed that of
the two major products, band 1 represented the downstream (3') part and band 2 the upstream (5') part of the full-length substrate, with band 3 apparently containing neither end. Size markers were produced from uniformly-labeled substrate by digestion with RNase H. Comparison of these size markers with bands 1, 2, and 3 produced by RB47 CTD (Fig. 3) demonstrated that the two CTD cleavage sites are within the stem-loop structure (Fig. 3) with RNase H. Sequences used for the GST-RB47 ID are indicated at right. C, immunoblot of the recombinant proteins using an antibody against the RB47 CTD domain. Only proteins which contain the CCT (RB47, RB47 ID, RB47 antigen) are recognized.

In Vitro Processing of Full-length RB47 to Version Lacking the NCL—As mentioned above, in vitro-synthesized RB47 was reported (8) to undergo proteolytic cleavage upon import into pea chloroplasts, which removed an as-yet unidentified ~30-kDa segment of the protein. Our analysis of RB47 purified from chloroplasts suggested that along with short the N-terminal transit peptide, most or all of the internal ~30-kDa NCL was removed by an unusual processing event that allowed retention of an apparently contiguous stretch comprising the RRM and CCT domains. To facilitate additional studies, an in vitro processing assay was developed relying on homologous (Chlamydomonas) chloroplast extracts. Initially, chloroplasts were lysed in buffer containing detergent and the soluble fraction was incubated with recombinant GST-RB47 (i.e. the full-length protein) bound to a glutathione-Sepharose column. After incubation, the beads were thoroughly washed to avoid contamination by proteins from the chloroplast extract, and the recombinant substrate and products were released from beads by heating, then analyzed by gel electrophoresis.

Results from processing of GST-RB47 and controls were analyzed by Coomassie Blue staining (Fig. 4A), and immunoblotting with the antibody raised against the CCT domain (Fig. 4B). The stained gel revealed that the amount of input GST-RB47 (~100 kDa, lane 1) was greatly reduced following incubation with chloroplast lysate (lane 3), concomitant with the appearance of smaller bands consistent with protein cleavage activity. These bands were not present when the lysate was incubated with the beads in the absence of recombinant RB47 (lane 2). When treated with thrombin, which cleaves between GST and the RB47 sequences, still smaller species were seen (lane 4). To find which of these species contained the CCT, immunoblot analysis was used. We found that three of these bands seen in lane 3 (marked with asterisks) contained both the GST moiety, since they were bound to the glutathione-Sepharose column, and the CCT, since they were identified with the antibody.

The three GST-RB47-derived species include an SDS-resistant form (top band), which barely entered the gel, residual unprocessed GST-RB47 (middle band, which co-migrated with the input; lane 1), and a 75-kDa species, a size consistent with the removal of the linker domain. To verify that the 75-kDa species was not in some way derived by removal of the GST moiety, thrombin cleavage was used (lanes 4). Based on immunoblot analysis (Fig. 4B), the major species accumulating following incubation with the chloroplast extract was ~47 kDa, consistent with the size of RB47 sequenced from chloroplasts. The predominance of this band also suggested that the SDS-resistant complex primarily contained processed RB47, that became amenable to gel migration following treatment with thrombin. An additional control lane was included in Fig. 4B (lane 5) to rule out the possibility that the 47-kDa band seen in Fig. 4B, lane 4, might represent RB47 from the chloroplast lysate which had bound to the recombinant GST-RB47. This lane was loaded with whole lysate corresponding to the amount of lysate used for the processing reactions ultimately loaded into lanes 3 and 4. No chloroplast RB47 was detected in the amount of lysate used and therefore the bands detected in lanes 3 and 4 were derived from the recombinant GST-RB47.

Adherence to a GST column, presence of the thrombin cleavage site, and recognition by the antibody collectively supported the notion that the 47 kDa species had undergone internal processing. To ascertain the protein sequences directly, we subjected the precursor GST-RB47 and the processed product (arrowheads in Fig. 4A) to peptide mapping by trypsin/formic acid digestion and MS. As shown earlier, RB47 purified from chloroplasts had MS coverage in most regions excepting the transit peptide (TP) and NCL (Fig. 4C, top). On the other hand, unprocessed GST-RB47 had reasonable cover-
age of both those regions and the NCL, suggesting that the NCL is not inherently refractory to MS. Using the same method, however, we were not able to identify peptide sequences from the linker region in the processed GST-RB47 species that migrated at 47 kDa (Fig. 4C, bottom). The precise extents of sequenced peptides are detailed in Fig. 5. Taken together, we tentatively conclude that factors in the *Chlamydomonas* chloroplast extract had excised the the majority of the NCL while splicing the RRMs to the CCT.

**Fine Analysis of RB47 Splicing**—While the above evidence is strongly suggestive of protein splicing, we wished to demonstrate formally that the RRM and CCT domains had been joined together as a linear sequence of amino acids, rather than as a covalent cross-linking between amino acid side chains or by some other mechanism. The locations of the splice site(s) could not be precisely determined by peptide mapping alone, because intermittent portions of the NCL could be not be detected by sequencing recombinant unprocessed RB47 (Fig. 5, middle). To begin to narrow down the borders of sequence excision, histidine tags were inserted near the likely upstream (Sph) and downstream (Bsp) splice sites within the NCL, as estimated by processed protein migration and MS results (Fig. 6A). Our strategy was to use the presence or absence of the tags in spliced RB47 as an indication of on which side of the tag splicing had occurred. Because our original GST-RB47 fusion possessed a His tag at the extreme C-terminal end (Fig. 5, bottom two sequences), this tag had to be removed. This generated a version of GST-RB47 lacking the C-terminal tag (RB47-His), into which the two tags within the NCL could be inserted (see “Experimental Procedures”).

In this, and in all subsequent experiments, the growth conditions were altered from the continuous light used for cells from which chloroplast extracts were prepared for use in the experiments in Fig. 4. Instead, chloroplasts were isolated two
hours into the light phase of a light/dark growth cycle (16 h light/8 h dark). This change did not affect splicing per se, but did seem to affect the nature of the SDS-resistant complex seen in Fig. 4 (lane 3). Rather than a single, slowly migrating complex, which barely entered the gel, the SDS-resistant form manifested as a series of complexes ranging in migration from slightly larger than the unspliced form of RB47, to barely entering the gel (Fig. 6B, lanes 3, 4, 7, 8, 11, 12). Furthermore, these complexes seemed more resistant to digestion with thrombin than the complex shown in Fig. 4. The nature of the SDS-resistant complex has not been determined, but may be explained by a cross-linking between RB47 and a component(s) of the putative splicing complex (see below).

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

FIGURE 4. In vitro protein splicing reactions. A, Coomassie-stained SDS gel of in vitro protein processing reactions using GST-RB47 bound to glutathione beads and treated as indicated at the bottom. Protein size markers are at right. Bands indicated by (*) were recognized by the CCT antibody (see panel B), and the bands indicated by arrowheads in lanes 1 and 3 were excised from the gel and submitted for MS analysis. B, Immunoblot of the gel from panel A. Lanes 1–4 as in panel A; lane 5, chloroplast lysate equivalent to the amount used for the reactions in lanes 3 and 4. C, MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

The three His-tagged precursor proteins were subjected to processing in the chloroplast lysate as described above, and an immunoblot analysis with the anti-RB47 CCT antibody is shown in Fig. 6B. Lanes 1, 5, and 9 show the intact precursors, from which GST could be removed by thrombin digestion (lanes 2, 6, 10). When incubated in the chloroplast lysate a series of bands including an SDS-resistant species was observed in all cases (lanes 3 and 4). MS peptide mapping results for RB47 proteins purified from chloroplasts, recombinant RB47 from panel A, lane 1 indicated by arrowhead or GST-RB47 after in vitro protein processing from panel A, lane 3 indicated by arrowhead. Regions identified by MS peptides are in red.

We had anticipated detecting the excised NCL following splicing, if the His tags were found to be within the region removed during the reaction, as was indeed the case. We could rule out that the spliced NCL was trapped within an SDS-resistant complex, because anti-His tag immunoblots showed no signal whatsoever following splicing. We therefore considered it likely that the excised NCL, or at least those portions containing His tags, had been degraded. Since the chloroplast lysate used in the reactions consisted was a crude soluble protein preparation, it was likely to contain numerous proteases and exopeptidases. The inability to detect the extreme C-terminal peptides from cleavage (compare lanes 1 and 2 with 5/9 and 6/10, respectively). We found that the signal from the internally tagged proteins was completely lost after incubation with chloroplast lysate, with or without thrombin cleavage (lanes 7, 8, 11, 12). This suggested that the 47 kDa species was created by excising sequences within or upstream of the Sph His tag, and within or downstream of the Bsp His tag, and that the higher molecular weight bands in the splicing reactions all were species lacking the NCL, most likely aggregates with different degrees of resistance to SDS denaturation.
either chloroplast RB47 or recombinant RB47 exposed to chloroplast lysate (Figs. 4C and 5) is consistent with exopeptidase activity. To address this issue, and to obtain more highly purified preparations of splicing activity, additional fractionation steps were carried out as outlined in Fig. 7A. Briefly, chloroplasts were initially lysed without detergent and separated into soluble (S) and insoluble (membrane-associated) fractions. The membrane-associated fraction was further extracted with detergent and tested for activity. The activity was primarily found associated with the detergent extracted fraction, which was then applied to a series of three columns for further purification. In each case, the activity was found in the flow-through fraction. Splicing assays for the membrane-associated, detergent-extracted fraction and the final column flow-through (DE) are shown in Fig. 7B and C, using the Bsp internal His-tagged substrate.

Independently, we had found that RB47 splicing activity was resistant to heating to 85 °C for 20 min, and we used this treatment to obtain additional purity by precipitating other proteins that were insoluble under these conditions. As shown in Fig. 7B, splicing carried out by heat-treated proteins (H) was indistinguishable from that catalyzed by the membrane-extracted (ME) or DEAE Sepharose (DE) column-purified fractions, in terms of production of the 47-kDa product. On the other hand, when reaction products were probed with the anti-His tag antibody, a strong signal remained when the heat-treated proteins were used, unlike the wholesale loss of signal from the His tag when less purified fractions were used (Fig. 7C; ME, DE). Unfortunately, the signal remained trapped within SDS-resistant complexes under standard denaturation conditions, precluding meaningful size analysis. This supports our hypothesis that loss of the His tag signal in cruder fractions was due to the activity of a heat-labile protease/peptidase(s).

Based on the experiments in Figs. 6 and 7, the splice junctions are located N-terminal to the Sph tag and C-terminal to the Bsp tag. To identify the splice junctions precisely, we used the Bsp-His substrate for splicing with the most purified fraction, namely the heated lysate. After splicing and thrombin digestion, the 47-kDa product was excised from a Coomassie-stained gel and subjected to trypsin digestion and mass spectroscopy. The peptides identified were queried using a custom database comprising all the possible complete and partial tryptic peptides that could be generated by splicing between the two junction regions identified using the His tags. The resulting positives were then screened against a human peptide database to eliminate false positives. Two tryptic peptides remained after these steps, as shown in Fig. 8. Peptide 123 was P-QQK and peptide 243 PPNPMAVTS-QK, where the dash indicates the inferred splice junction. The initial proline of each tryptic pep-
tide must be preceded by a basic residue (K/R) since trypsin cleaves C-terminal to these residues (although rarely when K/R is followed by a proline). The (K/R) proline combination of both peptides must lie upstream of the SphI site in which the His tag, when inserted, is excised during splicing. There is only one combination, (R) proline, between that site and sequences present, based on MS, in the spliced product (Fig. 8, first black arrowhead). The QQK of peptide 123 is unambiguous and is found shortly downstream of the His tag that was present in the substrate used for splicing. Peptide 243 appears to be derived from splicing immediately upstream of the SphI site, with the downstream junction one amino acid displaced from the site defined by peptide 123. We conclude that the RRM and CCT domains of the RB47 Bsp-His substrate are joined as a linear peptide sequence in at least these two configurations (see “Discussion”).

RB47 Is Present in Unspliced Form in Vivo—All previous reports documenting RB47 purified from chloroplasts described the 47-kDa form, but no evidence for the longer, unspliced form predicted by cDNA (8, 11). Many of the analyzed fractions had been highly purified, or were assayed by UV-crosslinking rather than immunoblotting, which might have led to failure to detect the unspliced form. Another possibility was that splicing occurs rapidly and quantitatively during or immediately after import of RB47 from the cytosol, where it is synthesized, leading to a very low steady-state level of the unspliced form. To explore whether unspliced RB47 is present in chloroplasts, we examined isolated chloroplasts. Furthermore, as RB47 was posited to be a light-regulated translation factor (7, 9, 17), chloroplasts were isolated from cells grown for 4 days in either continuous light or darkness and prepared for immunoblotting either immediately (time 0), or following 30, 60, or 120 min of incubation in the dark at room temperature. Fig. 9, A and B (lanes 1) show that a substantial pool of unspliced RB47 (80 kDa) is present in the chloroplasts of both light- and dark-grown cells, along with several bands between 40 and 50 kDa. Over time, in chloroplasts from cells grown in the light, but not from those grown in the dark, the pool of unspliced RB47 was apparently converted to the 47 kDa form of RB47, which eventually became the slightly more predominant form. This suggested that RB47 splicing can occur within chloroplasts, and is strongly influenced by their prior exposure to light.

As noted above, two other minor species were identified with the RB47 antibody (marked by asterisks in Fig. 9, A and B). These either increased slightly in abundance (Fig. 9A) or changed little (Fig. 9B) as chloroplasts were incubated over a 2-h time course. To determine the source of these proteins, we compared immunoblot profiles of proteins from whole, light grown cells to those from isolated chloroplasts (Fig. 9C). The minor bands, migrating at ~51 kDa and 41 kDa, were readily found in whole cell preparations, but were absent from highly purified chloroplasts. This suggests that these two bands likely represent proteins found in some other organelle(s), possibly...
Chloroplast Protein Splicing

FIGURE 10. In vitro splicing of RB47 requires light. A, in vitro protein splicing reactions were performed in either the light or dark for the times indicated below, prior to immunoblotting with anti-RB47 CCT antibody. Unspliced (80 kDa) and spliced (47 kDa) proteins after digestion with thrombin are indicated by arrowheads at left. Protein size markers are at right. B, in vitro protein splicing reactions were performed in either the light (L), dark (D), red light (R), or blue light (B) prior to immunoblotting with anti-RB47 CCT antibody. Unspliced (80 kDa) and spliced (47 kDa) proteins after digestion with thrombin are indicated by arrowheads at left. Protein size markers are at right.

mitochondria (18, 19), which had co-purified at a low level with the chloroplast preparations shown in Fig. 9A.

Light Effects on in Vitro RB47 Splicing—In the isolated chloroplast experiment described above, chloroplasts from light-grown cells appeared to have higher splicing activity than those prepared from dark-grown cells. To see whether light also had an effect on splicing in chloroplast lysates, recombinant RB47 precursor was incubated with chloroplast lysate in either the light or in darkness Preliminary experiments demonstrated a low level of splicing in the dark (data not shown), similar to the degree observed in chloroplasts isolated from cells grown in the dark (Fig. 9B). However, this was not observed if the chloroplast lysate was first stored in the dark for 30 min prior to initiating the in vitro splicing reactions (Fig. 10A). In this assay, splicing had a remarkable and absolute dependence on light. This included light-dependent formation of the SDS-resistant complex discussed for Fig. 6. Complex formation requires both the combination of RB47 and lysate (Fig. 10A, compare lanes 1, 2 to lanes 4–6) and direct exposure to light during the splicing reaction (compare lanes 4–6 to lanes 10–12). Thus, light is required for the splicing of RB47 in vitro and perhaps the energy is temporarily stored by the putative splicing complex. We hypothesize that the low level of splicing observed in chloroplasts isolated from cells grown in the dark (Fig. 9B) may have been due to a brief exposure to light.

To determine whether photosynthetically active wavelengths were required, recombinant RB47 was incubated under splicing conditions using different light qualities (Fig. 10B). We found that red light supported both accumulation of the SDS-resistant complex and splicing, but that blue light supported neither. Thus, blue light appears unable to act as a protein splicing catalyst.

Discussion

Our results demonstrate that the post-translational removal of the NCL sequence from RB47 unmasks an endoribonuclease activity, explaining why our aim to purify an RNA 3′ processing factor yielded spliced, but not unspliced RB47. Given that RB47 is a member of the PABP protein family, this raises the possibility that the previously documented truncation of cytosolic PABPs by viral proteases (16) might also activate an otherwise suppressed endoribonuclease activity, as an additional mechanism of inhibiting cellular gene expression. In the case of RB47, the primary role of endoribonuclease activity may be maturation or degradation of chloroplast mRNA, for atpB and potentially other transcripts, possibly via the polyadenylation-stimulated pathway (20). As mentioned earlier, RB47 has been postulated to be a D1 (psbA) translation factor. It also binds in vitro near a 5′ maturation site for psbA mRNA. To our knowledge, its ability to catalyze 5′-end maturation has not been tested directly, but several mutants deficient in D1 translation also have abnormally long psbA mRNA 5′-ends (21).

We classify RB47 splicing as non-intein protein-splicing (NIPS) because unspliced RB47 could be purified as a recombinant protein and was also observed in isolated chloroplasts. Canonical inteins are capable of self-excision, with the concomitant splicing of the flanking protein sequences. Intein removal occurs spontaneously upon translation and does not require any coenzymes or source of energy (13). RB47 splicing was observed, however, after incubation with chloroplast factors in vitro, and requires light, presumably as an energy source for the excision and/or splicing reaction(s). It has been observed that enzyme-catalyzed protein splicing occurs during the processing of minor histocompatibility antigens (22) and gene-encoded cyclic peptides (23, 24). It has also been suggested that enzyme-catalyzed protein splicing might act as an important means of creating protein diversity (25). RB47 protein processing is unusual, both mechanistically in that an internal sequence is excised from a large polypeptide, and insofar as it activates an enzymatic function.

Although NIPS and intein splicing appear to occur through distinctly different mechanisms, similarities can also be hypothesized. Inteins are embedded within a larger protein, with the protein sequence N-terminal to the intein referred to as the N-extein and the sequence C-terminal to the intein referred to as the C-extein. All of the active site residues, and the structural information required to form the active site and bring the splice junction residues within close proximity, are contained entirely within the intein sequence. Inteins have a conserved serine or cysteine residue at their N terminus and an asparagine residue at their C terminus. The chemical reactions...
which excise the intein, and join together the two exteins, occur by a sequential, but almost simultaneous, exchange of chemical bonds, as compared with a distinct breakage and then formation of new bonds, which would result in side products that could be isolated as distinct fragments. Instead, the intein and the two exteins remain joined together by covalent bonds until the reaction is complete (13). However, mutations of the intein active site residues can result in intermediate side products that can be isolated. For example, mutation of the intein’s N-terminal active site residue can result in the release of the C-extein, with the N-extein remaining attached to the intein (26).

We have observed that RB47 forms light-dependent, SDS-resistant complexes during splicing reactions, which are detectable with the RB47 antibody and may also include other factor(s) from the lysate. It is tempting to speculate that these may be the result of an exchange of covalent bonds, not intramolecularly as with inteins, but intermolecularly between the the unknown splicing factor(s) and the RB47 substrate. Our identification of the splice site junction formed in peptide 243 indicates that a serine residue is the N-terminal junction residue, in common with inteins, and a glutamine residue is the C-terminal junction residue. This is different from inteins, which use asparagine, but both amino acids have a reactive amino group in common on their side chains, which may serve the same function in the splicing mechanism. Peptide 123 contains a proline-glutamine junction. It is difficult to imagine how a proline residue can serve as a reactive active site residue and we are therefore must consider the possibility that peptide 123 is a false positive.

The accumulation of unprocessed RB47 was not documented in earlier studies of its interaction with psbA mRNA. While differences between strains or growth conditions could underlie prior failure to detect the 80 kDa form using immuno-blotting, as discussed above we believe a more likely cause to be that all previous biochemical analyses of RB47 relied on purification protocols that would allow further splicing and/or removal of the unprocessed form during protein preparation (8, 10, 11, 21). Why both unprocessed and spliced forms are retained in vivo remains to be determined. It is possible that each form performs a different function, for example, the reported psbA message-specific translational activator function (11) may be performed by the unspliced RB47 while the spliced form has RNase activity, or the unspliced form is simply an inactive state that is stored until environmental or other conditions trigger its activation.

It seems unlikely that a NIPS machinery would have arisen, or be retained, solely to process RB47. Given that the insertion of inteins, transposons, and recombination events are, in combination, an almost universal phenomenon, a critical question becomes how cells mitigate such events when they are deleterious, for example when inactivating insertions arise within essential genes. In the case of *Chlamydomonas*, the chloroplast genome encodes several essential proteins involved in transcription, translation, or proteolysis, which possess internal intron DNA insertions or extensions (27–29), which would seemingly make the resulting proteins non-functional. A NIPS machinery may have evolved to post-translationally repair and restore functionality to such genes (Fig. 11). This model need not be limited to a splicing activity, since this would presumably happen only when the insertion is folded such that the future junction site residues are in close enough proximity to be joined together. One possible example of an alternative NIPS activity is the *C. reinhardtii* ClpP1 subunit of the chloroplast ClpP protease complex, which contains a large insertion that is removed by an unknown endoprotease(s) (29). This is reminiscent of the “excision, no splicing” activity that we predicted for the insertion repair function of NIPS machinery in Fig. 11.

We refer to a NIPS machinery rather than to a NIPS enzyme because it seems unlikely that a single enzyme would have the capability to both recognize a potentially diverse range of insertions/extensions, and also possess the ability to catalyze splicing. In addition, a protease activity may be part of the putative NIPS complex, to degrade the excised insertion sequence and permit another catalytic cycle. This may explain why the NCL-degrading activity we observed seemed to co-purify with the NIPS activity prior to the heat treatment step (Fig. 7).

It is important to note that for PABPs as a whole, the NCL is not considered to be a non-functional insertion. It is an interesting question as to whether PABPs may have initially been ribonucleases before acquiring an insertion which altered their function, in which case protein splicing of RB47 restores its original function, or whether the adventitious event which caused RB47 to be redirected to the chloroplast resulted in a...
new enzymatic function upon protein splicing. Taken together, our observations with RB47 are consistent with a new class of enzyme-catalyzed protein splicing that may generate enzymatic diversity, restore enzymatic activity to proteins that would otherwise be defective due to genomic DNA insertions or faulty RNA splicing, or possibly in the case of RB47 (and perhaps other PABPs), restore the original function to a protein, which has been inactivated due to an insertion and has subsequently evolved to perform a new role.

**Experimental Procedures**

*Cell Growth Conditions—C. reinhardtii* strain CC-406 (cw15/mt-), a cell wall-deficient mutant, was grown in Tris-acetate-phosphate (TAP) medium supplemented with 1% sorbitol under constant fluorescent light (100 μmol m⁻² s⁻¹), or under 16 h day/8 h night cycles as indicated, to a density of 1–3 × 10⁷ cells/ml.

**Chloroplast Isolation**—Chloroplasts were isolated using a combination of osmotic shock and detergent treatment. Cells were harvested by centrifugation at 1,000 g for 5 min and resuspended in 100 ml of TAP plus sorbitol for every 1–3 × 10⁷ cells. NaCl was added to a concentration of 250 mM, and the suspension was incubated at room temperature for 15 min before adding Triton X-100 to a final concentration of 0.005%. The suspension was immediately mixed by inversion, then centrifuged at 1,000 × g for 2 min. The cell pellet was gently spread across the surface of the centrifuge tube before adding 10 ml cell lysis buffer (20% Percoll, 20 mM KH₂PO₄ pH 7, 150 mM mannitol, 1 mM EDTA) for every 1–3 × 10⁷ cells. The intact chloroplasts were separated by resuspension and centrifugation at 1,000 g, 80 °C. Typically, 0.5–1 ml of pelleted chloroplasts was collected for every 1–3 × 10⁷ cells.

**Purification of RB47 from Chloroplasts**—Frozen chloroplasts were thawed and resuspended in an equal volume of Buffer E (20 mM HEPES pH 8, 60 mM KCl, 12 mM MgCl₂, and 17% glycerol), and lysed using a Dounce homogenizer. Insoluble material was removed by centrifugation at 10,000 × g for 10 min and the supernatant was applied to a DE-52 (Whatman) column equilibrated in Buffer E. The flow through fraction was applied to an SP Sepharose (GE Healthcare) column equilibrated in Buffer E, and protein was eluted from the column with a 0.1–2 M KCl step gradient in Buffer E. Fractions containing nuclease activity (0.3–0.5 M KCl) were consolidated and precipitated with 50% saturated [(NH₄)₂SO₄] and centrifuged at 10,000 × g for 10 min. The pellet was dissolved in Buffer E and applied to an oligo(dT) cellulose (Amersham Biosciences) column equilibrated in Buffer E. Nuclease activity was eluted from the column with 2 M KCl, precipitated with (NH₄)₂SO₄, and dissolved in Buffer E before being applied to a Superdex 200 gel filtration column (GE Healthcare) equilibrated in Buffer E.

**Plasmids for Recombinant Protein Expression**—The *E. coli* expression vector pGex 4T-1 (GE Healthcare) encodes an IPTG-inducible promoter followed by the GST open reading frame flanked by a thrombin cleavage site, followed by a multiple cloning site (MCS) and stop codons in all three reading frames. All RB47 sequences were inserted into this vector using the BamHI/EcoRI restriction sites in the MCS such that the protein would be in-frame with the GST and thrombin cleavage site sequences. A cDNA encoding the C-terminal conserved domain was generated from total *C. reinhardtii* RNA, using the oligonucleotides CACCGATCCCTGTACCCGCAGGTG and GAATTCCTAAAGCCTTGTCTCCTC for RT-PCR. The cDNA, (nt 1716–1896, aa 572–632) was digested with BamHI/EcoRI and ligated into pGex 4T-1 to create the plasmid GST-RB47 antigen. The complete RB47 sequence, optimized for protein expression in *E. coli*, was synthesized (GenScript) to create the plasmid GST-RB47. The resulting protein sequence included the chloroplast transit peptide at the N terminus and a His₅ tag at the C terminus. The synthesized RB47 sequence was also designed to include an AvrII restriction site at nt 1305 (aa 435) and a BspEI site at nt 1695 (aa 563). The GST-RB47 ID plasmid was created by digesting GST-RB47 with these enzymes and inserting a linker of two complementary oligonucleotides with appropriate AvrII and Bsp EI overhangs, pCTAGGCGCCATGCACTG and pCCGGAGCTGCATGGC to effectively delete nt 1317 (aa 440) to nt 1689 (aa 563). A similar strategy was used to create the GST-RB47 CTD plasmid except that two linkers were ligated together in opposite orientation and inserted into the AvrII site of GST-RB47. This resulted in the same deletion as in GST-ID, the addition of five amino acids (ALHGA), and a stop codon after aa 565. The same strategy was used to insert His₅ tags within the NCL of the recombinant RB47 protein. As a preliminary step, a construct was generated by PCR using the primers CACAGATCCCATGCACTG and TGGGAATTCTTACGCTTTGTTTTC for RT-PCR. The resulting enzyme sequences were expressed in *E. coli* and inserted into the pGex vector to create the plasmid GST-RB47 minus His (-His). The plasmids GST-RB47 Sp6-His and GST-RB47 Bsp6-His were created by inserting linkers into either the Sp6I (nt 1245) or the BspEI (nt 1692) sites in the GST-RB47-His plasmid. The complementary oligonucleotides pCCATCACCATCACCATCACTTAG and pCTAAGTGATGGTGATGGCATG were ligated into the Sp6I site and inserted the amino acid sequence HHHHHHLC. The complementary oligonucleotide pCCGGAGCTGCATGGCAGGCG was ligated into the BspEI site and inserted the amino acid sequence HHHHHHLC. The complementary oligonucleotides pCCGGGCTGTAGTGTGATGGCATG and pCTAGGCGCCATGCACTG were ligated into the GST-RB47 plasmid and inserted into the pGex vector to create the plasmid GST-RB47 minus His (-His). The plasmids GST-RB47 Sp6-His and GST-RB47 Bsp6-His were created by inserting linkers into either the Sp6I (nt 1245) or the BspEI (nt 1692) sites in the GST-RB47-His plasmid. The complementary oligonucleotides pCCATCACCATCACCATCACTTAG and pCTAAGTGATGGTGATGGCATG were ligated into the Sp6I site and inserted the amino acid sequence HHHHHHLC.
for 2 h at 20 °C and cells were collected by centrifugation. The cell pellet was resuspended in phosphate buffered saline pH 7.2 (PBS), 0.1% Triton X-100, and 1 mg lysozyme and frozen at −20 °C. Frozen cells were thawed and lysed by sonication before centrifugation at 10,000 × g for 10 min to remove insoluble material. The supernatant was applied to a glutathione-Sepharose column, washed with 10 column volumes of PBS plus 1% Triton X-100, and then with 10 column volumes of PBS. Bound protein was eluted with 20 mM glutathione in PBS, precipitated with 50% saturated [(NH₄)₂SO₄] and centrifuged at 10,000 × g for 10 min. The pellet was dissolved in Buffer E and frozen at −80 °C.

Antibody Production and Affinity Purification—Polyclonal rabbit antiserum was raised against the RB47 CCT domain by Lampire Biologicals using the purified, soluble GST-RB47 antigen. Specific antibodies were affinity purified by binding the purified GST and GST-antigen proteins to a PVDF membrane and excising the sections of membrane to which the specific proteins were bound. These sections were used to affinity purify the RB47 CCT-specific antibodies by first passing the antiserum over the GST-bound sections in order to remove any GST-specific antibodies and then passing it over the RB47 antigen-bound sections to bind any RB47 CCT-specific antibodies. The bound antibodies were eluted with 100 mM glycine (pH 2.5), neutralized with 1x Tris pH 8, precipitated with 50% saturated [(NH₄)₂SO₄] and dissolved in PBS. anti-His₆ antibodies were purchased from Sigma-Aldrich.

RNA Substrate Synthesis and in Vitro RNA Processing—RNA substrate synthesis and in vitro RNA processing were as described previously (5). The substrate used was the one described as atpB-WT. 3′-end-labeled precursor RNA was produced by synthesizing unlabeled RNA and then labeling the 3′-end with 32P-labeled cordycepin and yeast poly(A) polymerase. RNA size markers produced by RNase digestion were produced by hybridizing uniformly labeled precursor RNA with DNA oligonucleotides complementary to specific sequences within the precursor (M1, nt 206 GCTAATATGCATAT; M2, nt 96 TTATTTTTATGAGCAGC) in RNase H buffer and then digesting with RNase H.

Total Chloroplast Lysate for in Vitro Protein Processing—Frozen chloroplasts were thawed and resuspended in an equal volume of Buffer E plus 1% Triton X-100 and lysed using a Dounce homogenizer. Insoluble material was removed by centrifugation at 10,000 × g for 10 min and the supernatant was diluted with an equal volume of Buffer E and the centrifugation repeated. The supernatant was concentrated by placing it in dialysis tubing (MW cutoff 3500 Daltons), and covering the tubing with dry PEG 8000. The concentrated lysate was then frozen at −80 °C. For in vitro protein processing, GST-RB47 (50 μg) was bound to glutathione beads and incubated with chloroplast lysates (50 μl) at room temperature in Buffer E in a total volume of 200 μl. After incubation for 1–3 h, the samples were extensively washed in Buffer E and, where indicated, further digested with 2 NIH units thrombin (Sigma Aldrich) for 1 h prior to analysis by SDS-PAGE. In the light/dark experiments, the light intensity used was 100 μmol m−2 s−1. Red (part 129526) and blue (part 129580) gel filters were purchased from Norman Lights and placed over fluorescent fixtures.

Column Purification of Chloroplast Lysate—Thawed Chlamydomonas chloroplasts were resuspended in an equal volume of Buffer E and lysed using a Dounce homogenizer. Insoluble (membrane-associated) material was pelleted by centrifugation at 10,000 × g for 10 min. The pellet was then extracted with 1% TX-100 in Buffer E, diluted to 0.5% TX-100, centrifuged at 10,000 × g for 10 min and the supernatant was applied to a DE-52 column equilibrated in Buffer E. The flow-through fraction was applied to an SP Sepharose column equilibrated in Buffer E, and that flow-through fraction was applied to a DEAE Sepharose column (GE Healthcare). The flow-through from the DEAE column was concentrated by placing it in dialysis tubing, MW cutoff 3500 Dalton, and covering the tubing with dry PEG 8000. Further purification was achieved by heating the DEAE-obtained fraction to 85 °C for 20 min, slow cooling to room temperature and centrifugation at 10,000 × g for 10 min. The purified supernatant was stored at −80 °C.

Mass Spectroscopy—Protein bands were excised from the Coomassie Blue-stained gels and submitted to the Cornell University Life Sciences Core Laboratories Center. RB47 purified from chloroplasts was digested with trypsin and analyzed by MALDI-TOF MS. Recombinant RB47 was digested with trypsin and formic acid and analyzed by LC-MS. The splice junction(s) were identified by nano LC/MS/MS analysis of tryptic peptides from the spliced, recombinant RB47 protein. The peptides identified were queried using a custom database comprising all the possible complete and partial tryptic peptides that could be generated by splicing between the two junction regions identified using the His tags. The resulting positives were then screened against the Uniprot human peptide database to eliminate false positives; the remaining peptides are the two described under “Results.” The mass spectrometry proteomics data derived from the recombinant protein have been deposited to the ProteomeXchange Consortium via the PRIDE (14, 30) partner repository with the dataset identifier PXD004285.

Author Contributions—S. C. designed and conducted experiments, analyzed data, and contributed to the writing of the manuscript; D. S. analyzed data and contributed to experimental design and to the writing of the manuscript.

References
1. Germain, A., Hotto, A. M., Barkan, A., and Stern, D. B. (2013) RNA processing and decay in plastids. Wiley Interdiscip. Rev. RNA 4, 295–316
2. Stern, D. B., Goldschmidt-Clermont, M., and Hanson, M. R. (2010) Chloroplast RNA Metabolism. Annu. Rev. Plant Biol. 61, 125–155
3. Barkan, A. (2011) Expression of plastid genes: organelle-specific elaboration and role of U1 snRNA. EMBO J. 30, 2277–2285
4. Stoppel, R., and Meurer, J. (2012) The cutting crew - ribonucleases are key players in the control of plastid gene expression. J. Plant Physiol. 172, 1309–1313
5. Barkan, A., and Kindle, K. L. (1993) 3′ end maturation of the Chlamydomonas reinhardtii chloroplast atpB mRNA is a two-step process. Mol. Cell. Biol. 13, 2277–2285
6. Hicks, A., Drager, R. G., Higgs, D. C., and Stern, D. B. (2002) An mRNA 3′ processing site targets downstream sequences for rapid degradation in Chlamydomonas chloroplasts. J. Biol. Chem. 277, 3325–3333
7. Danon, A., and Mayfield, S. P. (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins. EMBO J. 10, 3993–4001
Chloroplast Protein Splicing

8. Yohn, C. B., Cohen, A., Danon, A., and Mayfield, S. P. (1998) A poly(A)-binding protein functions in the chloroplast as a message-specific translation factor. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2238–2243

9. Danon, A., and Mayfield, S. P. (1994) Light-regulated translation of chloroplast messenger RNAs through redox potential. *Science* **266**, 1717–1719

10. Yohn, C. B., Cohen, A., Danon, A., and Mayfield, S. P. (1996) Altered mRNA binding activity and decreased translational initiation in a nuclear mutant lacking translation of the chloroplast psbA mRNA. *Mol. Cell. Biol.* **16**, 3560–3566

11. Yohn, C. B., Cohen, A., Rosch, C., Kuchka, M. R., and Mayfield, S. P. (1998) Translation of the chloroplast psbA mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. *J. Cell Biol.* **142**, 435–442

12. Alargand, T., Peled-Zehavi, H., Katz, Y., and Danon, A. (2006) The chloroplast protein disulfide isomerase RB60 reacts with a regulatory disulfide of the RNA-binding protein RB47. *Plant Cell Physiol.* **47**, 540–548

13. Paulus, H. (2000) Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* **69**, 447–496

14. Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L. *et al.* (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226

15. Lemay, J. F., Lemieux, C., St-André, O., and Bachand, F. (2010) Crossing the borders: poly(A)-binding proteins working on both sides of the fence. *RNA Biol.* **7**, 291–295

16. Smith, R. W., and Gray, N. K. (2010) Poly(A)-binding protein (PABP): a common viral target. *Biochem. J.* **426**, 1–12

17. Danon, A., and Mayfield, S. P. (1994) ADP-dependent phosphorylation regulates RNA-binding in vitro: implications in light-modulated translation. *EMBO J.* **13**, 2227–2235

18. Klein, U., Chen, C., Gibbs, M., and Platt-Aloia, K. A. (1983) Cellular fractionation of *Chlamydomonas reinhardtii* with emphasis on the isolation of the chloroplast. *Plant Physiol.* **72**, 481–487

19. Mason, C. B., Bricker, T. M., and Moroney, J. V. (2006) A rapid method for chloroplast isolation from the green alga *Chlamydomonas reinhardtii*. *Nat. Protoc.* **1**, 2227–2230

20. Lange, H., Sement, F. M., Canaday, J., and Gagliardi, D. (2009) Polyadenylation-assisted RNA degradation processes in plants. *Trends Plant Sci.* **14**, 497–504

21. Bruick, R. K., and Mayfield, S. P. (1998) Processing of the psbA 5¢ untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. *J. Cell Biol.* **143**, 1145–1153

22. Warren, E. H., Vigneron, N. J., Gavin, M. A., Coulie, P. G., Stroobant, V., Dalet, A., Tykodi, S. S., Xuereb, S. M., Mitto, J. K., Riddell, S. R., and Van den Eynde, B. J. (2006) An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444–1447

23. Saska, I., Gillon, A. D., Hatsugai, N., Dietzgen, R. G., Hara-Nishimura, I., Anderson, M. A., and Craik, D. J. (2007) An asparaginyl endopeptidase mediates in vivo protein backbone cyclization. *J. Biol. Chem.* **282**, 29721–29728

24. Craik, D. J. (2006) Chemistry. Seamless proteins tie up their loose ends. *Science* **311**, 1563–1564

25. Bouabe, H. (2008) Polypeptide rearrangement hypothesis and its implication in genetic diversity. *J. Proteomics Bioinform.* **1**, 336–342

26. Mathys, S., Evans, T. C., Chute, I. C., Wu, H., Chong, S., Benner, J., Liu, X. Q., and Xu, M. Q. (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231**, 1–13

27. Fong, S. E., and Surzycki, S. J. (1992) Organization and structure of plastome psbF, psbL, petG, and ORF712 genes in *Chlamydomonas reinhardtii*. *Curr. Genet.* **21**, 527–530

28. Fong, S. E., and Surzycki, S. J. (1992) Chloroplast RNA polymerase genes of *Chlamydomonas reinhardtii* exhibit an unusual structure and arrange their coding regions: facile production of protein building blocks for protein ligation. *Curr. Genet.* **21**, 485–497

29. Derrien, B., Majeran, W., Wollman, F. A., and Vallon, O. (2009) Multistep protein backbone cyclization. *Biochem. J.* **426**, 1–12

30. Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, X. W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–456