



**HAL**  
open science

## Proteomic Characterization of the *Chlamydomonas reinhardtii* Chloroplast Ribosome

Kenichi Yamaguchi, María Verónica Beligni, Susana Prieto, Paul A Haynes,  
W. Hayes Mcdonald, John R. Yates Iii, Stephen Mayfield

► **To cite this version:**

Kenichi Yamaguchi, María Verónica Beligni, Susana Prieto, Paul A Haynes, W. Hayes Mcdonald, et al.. Proteomic Characterization of the *Chlamydomonas reinhardtii* Chloroplast Ribosome. *Journal of Biological Chemistry*, 2003, 278 (36), pp.33774-33785. 10.1074/jbc.M301934200 . hal-02120760

**HAL Id: hal-02120760**

**<https://hal.science/hal-02120760>**

Submitted on 6 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Proteomic Characterization of the *Chlamydomonas reinhardtii* Chloroplast Ribosome

IDENTIFICATION OF PROTEINS UNIQUE TO THE 70 S RIBOSOME\*<sup>§</sup>

Received for publication, February 24, 2003, and in revised form, June 9, 2003  
Published, JBC Papers in Press, June 24, 2003, DOI 10.1074/jbc.M301934200

Kenichi Yamaguchi<sup>‡§</sup>, María Verónica Beligni<sup>‡</sup>, Susana Prieto<sup>‡¶</sup>, Paul A. Haynes<sup>||</sup>,  
W. Hayes McDonald<sup>\*\*</sup>, John R. Yates III<sup>\*\*</sup>, and Stephen P. Mayfield<sup>‡ ††</sup>

From the <sup>‡</sup>Department of Cell Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, the <sup>||</sup>Torrey Mesa Research Institute of Syngenta, San Diego, California 92121, and the <sup>\*\*</sup>Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037

We have conducted a proteomic analysis of the 70 S ribosome from the *Chlamydomonas reinhardtii* chloroplast. Twenty-seven orthologs of *Escherichia coli* large subunit proteins were identified in the 50 S subunit, as well as an ortholog of the spinach plastid-specific ribosomal protein-6. Several of the large subunit proteins of *C. reinhardtii* have short extension or insertion sequences, but overall the large subunit proteins are very similar to those of spinach chloroplast and *E. coli*. Two proteins of 38 and 41 kDa, designated RAP38 and RAP41, were identified from the 70 S ribosome that were not found in either of the ribosomal subunits. Phylogenetic analysis identified RAP38 and RAP41 as paralogues of spinach CSP41, a chloroplast RNA-binding protein with endoribonuclease activity. Overall, the chloroplast ribosome of *C. reinhardtii* is similar to those of spinach chloroplast and *E. coli*, but the *C. reinhardtii* ribosome has proteins associated with the 70 S complex that are related to non-ribosomal proteins in other species. In addition, the 30 S subunit contains unusually large orthologs of *E. coli* S2, S3, and S5 and a novel S1-type protein (Yamaguchi, K. *et al.*, (2002) *Plant Cell* 14, 2957–2974). These additional proteins and domains likely confer functions used to regulate chloroplast translation in *C. reinhardtii*.

In the chloroplast, where proteins of the photosynthetic apparatus and the carbon-fixing enzymes are synthesized, gene expression is primarily regulated during translation (1). Chloroplast translation has been thought to be similar to transla-

tion in bacterial systems, mainly because of similarities in ribosomal RNA and the sensitivity of chloroplast ribosomes to bacterial antibiotics. These similarities support the endosymbiotic theory that chloroplasts originated from a photosynthetic prokaryote, cyanobacteria (2, 3). It is now recognized that chloroplast gene expression and chloroplast translation are unique and quite different from bacterial systems (1, 4–7). The chloroplast ribosome contains plastid-specific ribosomal proteins (PSRPs)<sup>1</sup> in addition to bacterial orthologs (5–7). It has been proposed that PSRPs may take part in the unique light-dependent aspects of chloroplast translation (8). Bacterial gene expression is strongly influenced by the rate of transcription, and translation and transcription are often closely coupled. In the chloroplast, transcription is often globally regulated, and mRNA accumulation can be unrelated to the rate of translation of a protein (reviewed in Refs. 1 and 9–13). Translation of many chloroplast mRNAs is activated in response to light illumination, with little change in the corresponding mRNA levels (14–19). A majority of the work on chloroplast translation has been carried out in the unicellular green alga, *Chlamydomonas reinhardtii* (reviewed in Refs. 13, 20, and 21), because it is amenable to both genetic and biochemical analysis (22, 23). Identification of all of the proteins required for chloroplast translation in *C. reinhardtii* would facilitate our understanding of the mechanisms of chloroplast translation, photosystem biogenesis, plastid differentiation, and ultimately plant development and function.

Ribosomal proteins (RPs) of the small and large subunits from the *C. reinhardtii* chloroplast ribosome have been characterized by two-dimensional PAGE (24–26). Schmidt *et al.* (26) designated these proteins S-1 to S-31 and L-1 to L-33, according to their estimated size on two-dimensional PAGE, and identified their sites of synthesis by *in vivo* pulse-labeling in the presence of inhibitors of cytoplasm (anisomycin) and chloroplast (lincomycin) translation. Immunological characterization of chloroplast RPs from *C. reinhardtii* was carried out with antisera available for 15 bacterial RPs (27, 28) and suggested that chloroplasts contained orthologs for each of these bacterial proteins. However, to date the majority of pro-

\* This work was supported by Grant GM54659 from the National Institutes of Health, Contract DE-FG03-93ER70116 from the Department of Energy, and a grant from Syngenta Corp. (all to S. P. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> The on-line version of this article (available at <http://www.jbc.org>) contains a supplementary figure.

The nucleic acid sequences reported in this paper have been submitted to the GenBank™/EMBL Data Bank with accession numbers AY177617 (Rap38) and AY177616 (Rap41).

<sup>¶</sup> Supported by a Skaggs postdoctoral fellowship. Present address: Division of Biochemistry, Faculty of Fisheries, Nagasaki University, Bunkyo-machi, Nagasaki 852-8521, Japan.

<sup>||</sup> Present address: Centre de Recherche de Biochimie Macromoléculaire (CRBM), CNRS UPR 1086, 34293 Montpellier Cedex 5, France.

<sup>††</sup> To whom correspondence should be addressed: Dept. of Cell Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-784-9848; Fax: 858-784-9840; E-mail: mayfield@scripps.edu.

<sup>1</sup> The abbreviations used are: PSRP, plastid-specific ribosomal protein; PRP, plastid ribosomal protein; RP, ribosomal protein; ORF, open reading frame; LC-MS/MS, reversed-phase HPLC coupled with tandem mass spectrometry; LC/LC-MS/MS, two-dimensional liquid chromatography coupled with tandem mass spectrometry; pRRF, plastid ribosome recycling factor; CRP, cytoplasmic ribosomal protein; MudPIT, multi-dimensional protein identification technology; EST, expressed sequence tag; RAP, ribosome-associated protein; UTR, untranslated region; NTE, N-terminal extension.

teins of the *C. reinhardtii* chloroplast ribosome have not been identified.

Recently, sequence databases for the nuclear and chloroplast genomes of *Chlamydomonas* have become available (29–31).<sup>2</sup> Although similarity searches of open reading frames (ORFs) within these databases can help us speculate on the composition of chloroplast RPs, these databases will not allow us to identify actual RPs, especially proteins unique to this alga. Actually, ORFs can be precisely annotated only after protein sequencing and protein characterization. Precise identification of all the RPs in the *C. reinhardtii* chloroplast ribosome based on a precise proteomic characterization would give a much clearer picture of how algal RPs are similar, and different, from those of bacteria and higher plants.

In our recent study, we characterized the small subunit of the *C. reinhardtii* chloroplast ribosome via a proteomic approach. By using LC-MS/MS and LC/LC-MS/MS, we identified a novel S1 domain-containing protein (named PSRP-7) and unusually large bacterial orthologs of S2, S3, and S5 (7). Structural predictions, based on the crystal structure of the *Thermus thermophilus* 30 S subunit, suggest that the additional domains of S2, S3, and S5 are located adjacent to each other on the solvent side of the ribosome near the binding site for the S1 protein. We proposed that these additional domains interact with the S1 protein and PSRP-7 to function in mRNA recognition and translation initiation in *C. reinhardtii* chloroplast.

In this paper, we report the proteomic characterization of the chloroplast ribosomal large subunit and the complete 70 S ribosome. All of the proteins identified in the large subunit have bacterial orthologs and higher plant homologs. Overall, the large subunit proteins are very similar to those of *Escherichia coli* and spinach chloroplast in terms of size, isoelectric point, and amino acid sequence. We have classified all of the *C. reinhardtii* large subunit proteins in accordance with plastid ribosomal protein (PRP) nomenclature (5–7). In addition, proteomic and immunological analyses revealed that the 70 S ribosome of *C. reinhardtii* chloroplast contains two additional proteins of 38 and 41 kDa (RAP38 and RAP41), and these two proteins are not components of either the 30 S or 50 S subunits. Nucleotide sequencing and phylogenetic analysis revealed that RAP38 and RAP41 are paralogs of an ancestral nucleotide binding protein related to a bacterial epimerase/dehydratase and to a higher plant (spinach) chloroplast RNA-binding protein CSP41. This analysis has shown that the chloroplast translation machine of *C. reinhardtii* is compositionally and structurally similar to those of higher plants and eubacteria, but that *C. reinhardtii* ribosomes contain additional domains and proteins previously unidentified in other 70 S ribosomes.

#### EXPERIMENTAL PROCEDURES

**Nomenclature**—We describe *C. reinhardtii* chloroplast ribosomes, the large subunits, and the small subunits as 70 S, 50 S, and 30 S, respectively. This is to keep the nomenclature consistent with their higher plant and bacterial counterparts, although the sedimentation values of ribosomes and subunits from this algal chloroplast have been reported to be somewhat higher (32, 33). Likewise, we use the PRP nomenclature system adapted for plant PRPs, based on sequence similarity to bacterial RPs (5, 6). That is, chloroplast/plastid orthologs of *E. coli* L1–L36 are to be designated PRP L1–L36. In accordance with the approved Commission on Plant Gene Nomenclature designation for plant genes, gene names are written in italics, with nuclear genes having capital first letters, and organelle genes having lowercase first letters (34). For example, the gene for nuclear-encoded PRP L1 is *PrpL1*, whereas the gene for plastid-encoded PRP L2 is *prpL2*. We propose to use the cytoplasmic ribosomal protein (CRP) nomenclature for the plant and algal homologs of rat CRPs (35) to distinguish them

from PRP and mitochondrial RP. Prefixes PRP or CRP may be omitted when it is obvious what specified RP is being discussed.

**Preparation of *C. reinhardtii* Chloroplast Ribosomes and Subunits**—*C. reinhardtii* strain cc3395 (cw15/Arg7) was used for this study. The chloroplast ribosomes and subunits were prepared as described in our previous report (7), which is a modified procedure of nitrogen bomb extraction (36) and successive sucrose gradient centrifugation (37). Details on centrifugal conditions are described in the figure legends. RPs were extracted from purified ribosomes or subunits by an acetic acid extraction method as described (6).

**Proteomic Analyses**—SDS-PAGE was performed as described by Laemmli (38), using 1.5-mm-thick, 12% acrylamide gel. The molecular weight markers used were the BENCHMARK™ prestained protein ladder (Life Technologies, Inc.). For LC-MS/MS analysis, 10 pmol of total protein from the 50 S subunit (TP50) and 70 S ribosome (TP70) were separated by SDS-PAGE and then stained with Coomassie brilliant blue R-250 as described (7). The gels were sectioned into pieces as shown in Fig. 1E, and each gel piece was further fragmented into 1-mm<sup>2</sup> pieces and transferred into one well of a 96-well plate. The plates were transferred to a Massprep digestion robot (Micromass, Beverly, MA) for destaining (39), reduction/alkylation (iodoacetamide), and in-gel digestion with trypsin or endoproteinase Lys-C (40). After digestion, tryptic peptides were extracted from the gel pieces on the Massprep robot. The extracted peptides were then subjected to LC-MS/MS equipped with a microbore HPLC system (Surveyor, ThermoFinnigan, San Jose, CA), a Surveyor autosampler, and a ThermoFinnigan LCQ-Deca ion trap mass spectrometer (ThermoFinnigan) as described previously (7). Spectra were scanned over the range 400–1400 mass units. Automated peak recognition, dynamic exclusion, and product ion scanning of the top two most intense ions were performed using the Xcalibur software as described previously (41).

Direct analysis of total large subunit proteins was performed using ~10 µg of TP50. The TP50 was reduced and alkylated by iodoacetamide and then digested with Endoproteinase Lys-C (Boehringer Mannheim) and Porozyme Trypsin Beads (Perseptive Biosystems) as described in our previous report (7). The tryptic peptide mixture was analyzed by multidimensional protein identification technology (MudPIT) as described previously (42–44).

MS/MS data obtained were analyzed using SEQUEST, a computer program that allows the correlation of experimental data with theoretical spectra generated from known protein sequences (45, 46). In this work, the general criteria for a preliminary positive peptide identification for a doubly-charged peptide were a correlation factor >2.5, a  $\delta$  cross-correlation factor >0.1 (indicating a significant difference between the best match reported and the next best match), a minimum of one tryptic peptide terminus, and a high preliminary scoring. For triply-charged peptides, the correlation factor threshold was set at 3.5. All matched peptides were confirmed by visual examination of the spectra. All spectra were searched against a FASTA-format database generated from *Chlamydomonas* ESTs and ORFs in the *C. reinhardtii* plastid genome (SWISS-PROT) (29).

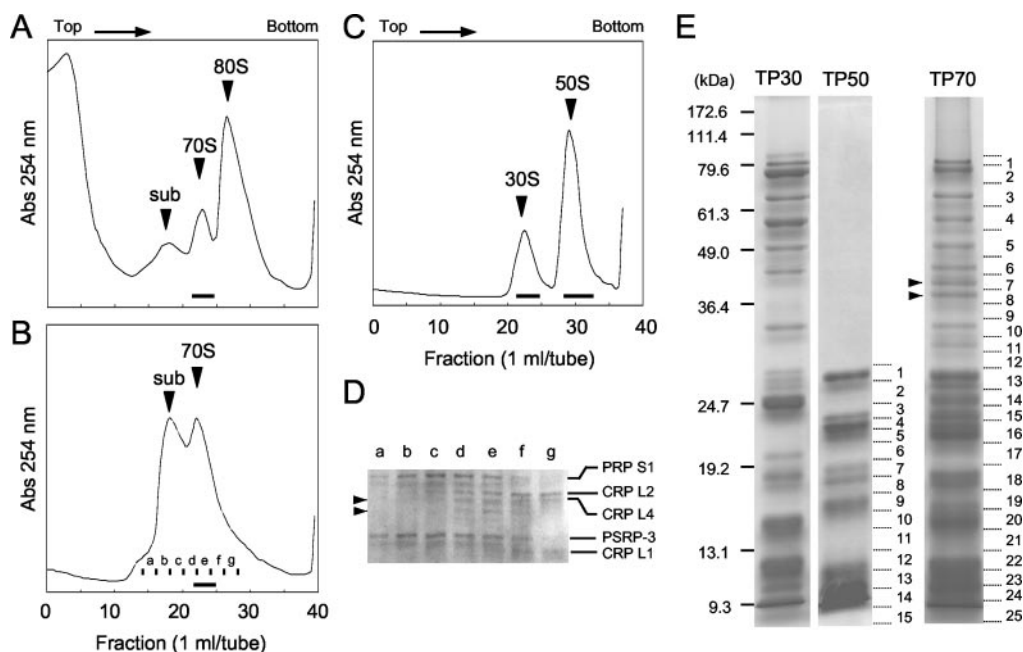
**Western Blot Analysis**—Samples were run in SDS-PAGE under reducing conditions and transferred to nitrocellulose. Membranes were blocked in TBST (10 mM Tris-HCl, pH 8.0, 0.9% NaCl, and 0.1% Tween 20) containing 5% skim milk and 0.03% Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and then incubated overnight in the same solution containing primary antibodies. Anti-*Chlamydomonas reinhardtii* RAP38 polyclonal antibodies were raised in rabbits (The Scripps Research Institute, Protein and Nucleic Acid Core Facility) and used at 1:2000 dilutions. Anti-spinach chloroplast RP S1 and L2 polyclonal antisera were provided by Dr. Subramanian and used at 1:2000 dilutions. Western blots were developed using nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (Sigma).

**DNA Sequencing**—Nucleotide sequencing was carried out with capillary electrophoresis technology using an ABI 3700 DNA analyzer (Nucleic Acid Core Facility, The Scripps Research Institute). The reactions were performed using thermal cycle sequencing conditions with fluorescently labeled terminators.

**Computational Analyses**—EST contig (assembled EST) was obtained from ChlamyDatabase using a WU-BLAST search.<sup>3</sup> BLAST (National Center for Biotechnology Information) was used for general sequence searches. Homology comparison was done using BLAST 2 SEQUENCES (National Center for Biotechnology Information). Multiple sequence alignments were performed using CLUSTAL X (47) and refined manually for representation. Isoelectric points and sequence

<sup>2</sup> Internet address: [genome.jgi-psf.org/chlre1/chlre1.home.html](http://genome.jgi-psf.org/chlre1/chlre1.home.html).

<sup>3</sup> Internet address: [www.biology.duke.edu/chlamy\\_genome/blast/blast\\_form.html](http://www.biology.duke.edu/chlamy_genome/blast/blast_form.html).



**FIG. 1. Isolation of *C. reinhardtii* chloroplast ribosomes and subunits and SDS-PAGE profiles of their total proteins.** A, sucrose gradient of the S-40 fraction from the total extract of *C. reinhardtii* cells in  $T_{25}K_{25}M_{10}D_5$  buffer (25 mM Tris-HCl, pH 8.0, 25 mM KCl, 10 mM  $MgCl_2$ , and 5 mM dithiothreitol). Peaks corresponding to 30 S, 40 S, 50 S, and 60 S ribosomal subunits (*sub*), chloroplast ribosomes (70S), and cytoplasmic ribosomes (80S) are indicated by arrowheads. Underlined 70 S fractions were pooled and subjected to a second sucrose gradient purification. B, second sucrose gradient of 70 S ribosomes in  $T_{25}K_{25}M_{10}D_5$  buffer. Some of the 70 S ribosomes are dissociated in the 30 S and 50 S subunits (*sub*). The underlined subunit fractions were pooled and subjected to a third sucrose gradient to obtain 30 S and 50 S subunits. The underlined 70 S fractions were collected for TP70 preparation. C, separation of chloroplast 30 S and 50 S subunits in dissociation buffer ( $T_{25}K_{100}M_5D_5$ ). The underlined fractions were collected for TP30 and TP50 preparations, respectively. D, SDS-PAGE analysis of proteins in fractions a to g in the second sucrose gradient of 70 S ribosomes from B. Two proteins indicated by arrowheads are seen in fraction e corresponding to the 70 S ribosomal fraction. PRPs S1, PSRP-3, and CRPs L1, L2, and L4 are indicated as landmarks. E, total proteins extracted from 10 pmol of each chloroplast small subunit (TP30), large subunit (TP50), and ribosome (TP70) were resolved on a SDS-PAGE gel. Proteins were stained with Coomassie Brilliant Blue R-250. Two bands of 41 kDa (section 7) and 38 kDa (section 8), present in TP70 but absent in TP30 and TP50, are indicated by the arrows. The TP50 and TP70 lanes were sectioned into 15 and 25 pieces, respectively, for in-gel digestion, as indicated by the dotted lines.

masses were calculated by the ProtParam in the ExPasy proteomic tools.<sup>4</sup> Prediction of cleavage sites for chloroplast transit peptides were obtained using the ChloroP program (48). Phylogenetic analysis was performed by a neighbor-joining method using PAUP\* version 4.0 Beta, and the reliability of the created tree was estimated by bootstrapping (49).

## RESULTS

**Isolation of the *C. reinhardtii* Chloroplast Ribosomes and Subunits**—Chloroplast ribosomes were purified from the post-mitochondrial S-40 fraction of total cell extracts by successive sucrose gradient centrifugation, according to the method of Chua *et al.* (37), to obtain biologically active ribosomes. The chloroplast 70 S ribosomes were first separated from the cytoplasmic 80 S ribosomes in a 10–40% sucrose gradient (Fig. 1A) and then purified on a second sucrose gradient (Fig. 1B). About half of the 70 S ribosomes loaded were dissociated into subunits, because of the fragile nature of the 70 S ribosome of the *C. reinhardtii* chloroplast (37, 50, 51). Purified 70 S ribosomes were collected and dissociated into 30 S and 50 S subunits in a 10–30% sucrose gradient containing dissociation buffer (Fig. 1C). The purity of the 70 S ribosomes, and 30 S and 50 S subunits, was assessed by the RP pattern observed on SDS-PAGE (Fig. 1, D and E) and by rRNA analysis on agarose gels (data not shown).

**In-gel Digestion and Liquid Chromatography-Tandem Mass Spectrometry**—Total proteins (TP30, TP50, and TP70) extracted from the 30 S subunits, 50 S subunits, and 70 S ribosomes, respectively, were resolved by SDS-PAGE and sectioned

into gel pieces (Fig. 1E). Proteins in each section were digested with trypsin or Lys-C, and the generated peptide fragments were subjected to LC-MS/MS analyses (see “Materials and Methods” for details), as done previously for the 30 S subunit proteins (7).

An example of the LC-MS/MS analysis of the peptides derived from the chloroplast 70 S ribosome is shown in Fig. 2, demonstrating identification of the L1 protein. In the first step of LC-MS, the trypsin fragments from section 13 of the TP70 gel yielded the mass chromatogram shown in Fig. 2A. A peptide eluting at 26.21 min yielded the MS spectrum shown in Fig. 2B. In the following step of MS/MS, the abundant precursor ion of  $m/z$  737.64 generated collision-induced dissociation spectra shown in Fig. 2C. Subsequent SEQUEST analysis of the collision-induced dissociation spectrum identified the peptide sequence to be DAGADVVGDDLIEK, the precursor ion mass ( $M + H$ )<sup>+</sup> to be 1474.56 (*i.e.*  $m/z$  737.64 is doubly charged ion), and the sequence belonged to an EST (AV620102), as listed in Table I. A BLAST search indicated AV620102 to be a homolog of chloroplast and bacterial RP L1. An EST contig (ACE number 20011023.1212.1) in the ChlamyEST database was identified by WU-BLAST search probing with AV620102 and confirmed as encoding a full-length chloroplast precursor protein homologous to the bacterial L1 protein. In the same way, other proteins from the TP70 and TP50 fractions were identified by LC-MS/MS analyses. As a complementary analysis, we also applied MudPIT for 50 S protein identification by the same procedure that we used for 30 S protein identification (7). MudPIT identifies protein components directly from an enzymatic digest of the entire large protein complex (here, it is TP50), using LC/LC-MS/MS without resolving proteins by poly-

<sup>4</sup> Internet address: [www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html).

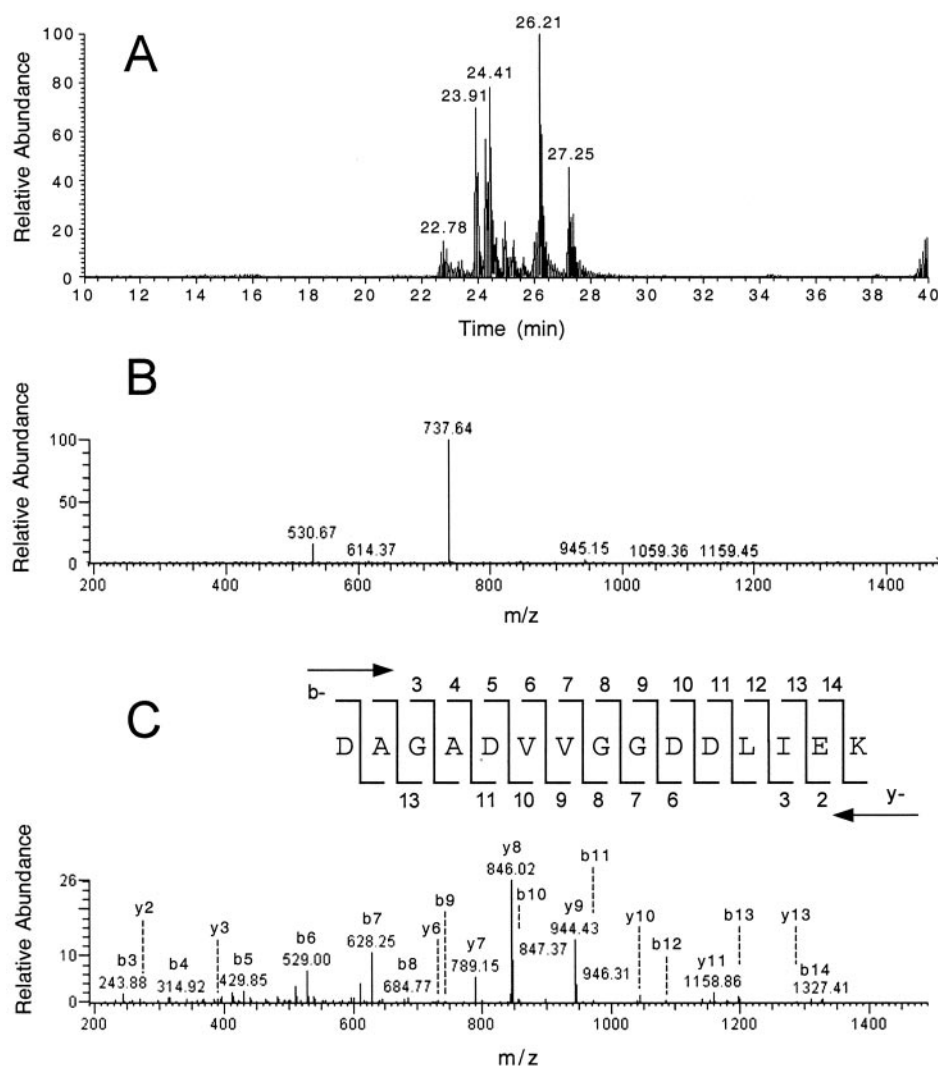


FIG. 2. LC-MS/MS analysis of an in-gel tryptic digest obtained from the *C. reinhardtii* chloroplast 70 S subunit proteins. A, total ion chromatogram obtained from a TP70 gel (section 13 in E). The represented chromatogram shows the relative abundance of the detected ions during retention times of 10–40 min, at which peptides were eluted from an LC column. B, MS spectrum of the precursor ions detected at 26.21 min in A. C, the MS/MS spectrum derived from the most abundant ion shown in B. Fragmentation of the doubly charged precursor ion at  $m/z$  737.64 yielded detectable singly charged b-ion species (b3–14) and y-ion species (y2, y3, y6–11, and y13), for which the sequence is indicated.

acrylamide gel electrophoresis (42–44). Table I summarizes all of the peptides identified by LC-MS/MS of TP50 and TP70 and LC/LC-MS/MS of TP50. Table I also shows their precursor ion masses, accession numbers of the identified ESTs or chloroplast ORFs, the protein identified and peptide origin, the SDS-PAGE gel section numbers, and the enzyme used or MudPIT.

**The Large Subunit Proteins**—From the chloroplast 50 S subunit, 28 proteins were identified including orthologs of bacterial L1, L2, L3, L4, L5, L6, L9, L10, L11, L12, L13, L14, L15, L16, L17, L18, L19, L20, L21, L22, L23, L24, L27, L28, L31, L32, and L35 and a homolog of spinach PSRP-6. An overview of the characteristics of the *C. reinhardtii* 50 S subunit proteins is shown in Table II comparing the *C. reinhardtii* 50 S PRPs with 50 S RPs from *Arabidopsis* chloroplast, *Synechocystis* sp. PCC 6803, and *E. coli*. The accession numbers and gene allocations for *C. reinhardtii* 50 S PRPs and percent similarity of 50 S RPs found in *C. reinhardtii* chloroplast, *Arabidopsis* chloroplast, *Synechocystis*, and *E. coli* are shown. The mature protein sizes and isoelectric points (pI) were predicted after removal of the predicted transit peptide or the *N*-formyl methionine and are summarized in Table III. Apart from proteomic analyses, we independently searched for homologs of the spinach 50 S PRPs and *E. coli* 50 S RPs from the *Chlamydomonas* databases and identified 31 potential 50 S PRP genes, suggesting the potential occurrence of L33, L34, and L36 proteins that were not identified by our proteomic analysis (see Tables II and III). These small and highly basic proteins likely escaped mass

spectrometric detection ( $m/z$  400–1400) because of digestion by trypsin or Lys-C generating peptides too small to be unequivocally sequenced using this method. No homologs of bacterial L29, L25, L30, and spinach PSRP-5 were identified in our proteomic analysis or from computation analysis of the complete chloroplast genome sequence and the currently available EST databases. This may be attributable to either incomplete EST databases or to the absence of these genes and proteins in this alga. L25 and L30 are likely absent from the chloroplast ribosome as discussed previously (6).

**The Small Subunit Proteins**—We previously identified 21 proteins from the small subunit of the *C. reinhardtii* chloroplast ribosome (7). In this study, we identified the same corresponding proteins from the 70 S ribosome (data not shown), and SEQUEST reanalysis of MudPIT data of TP30 using the complete chloroplast genome sequence identified one additional protein, S8 (Table I). Thus, we have identified 22 small subunit proteins: orthologs of *E. coli* S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S12, S13, S14, S15, S16, S17, S18, S19, S20, and S21; a homolog of spinach PSRP-3; and a novel S1 domain-containing protein (PSRP-7).

**Proteins Unique to the 70 S Ribosome: Identification of 38- and 41-kDa Ribosome-associating Proteins**—In spinach, a plastid ribosome recycling factor (pRRF) was identified in the chloroplast 70 S ribosome in stoichiometric amounts with other proteins of the 30 S and 50 S subunits. The pRRF was not present in either of the isolated 30 S or 50 S subunits (6).

TABLE I Identification of C. reinhardtii PRPs by tandem mass spectrometry

Table with 13 columns: Peptide sequence, TP50/TP30 MudPIT (MP), TP70-gel section No., Precursor Ion Mass (MH+), Accession number, Protein ID, Peptide position, TP50/TP30 MudPIT (MP), TP70-gel section No., Precursor Ion Mass (MH+), Accession number, Protein ID, Peptide position, TP50/TP30 MudPIT (MP), TP70-gel section No., Precursor Ion Mass (MH+), Accession number, Protein ID, Peptide position, TP50/TP30 MudPIT (MP), TP70-gel section No., Precursor Ion Mass (MH+), Accession number, Protein ID, Peptide position, TP50/TP30 MudPIT (MP), TP70-gel section No.

TABLE I—continued

Peptide sequence	Precursor Ion Mass (M+H) <sup>+</sup>	Accession number	Protein ID	Peptide position	TP50-gal section No. or TP50/TP30 MudPIT (MP)	TP70-gal section No. or TP70 MudPIT (MP)
<i>50S proteins</i>						
R.VYVDAVEGQTLGR.L	1508.67	AV388425	PRP L13	76-88	8T	
R.LASLAAIYR.G	1091.34	AV388425	PRP L13	89-98	6-11T	19L
K.HLPTTFSDMAGY.V	1587.68	AV392722	PRP L13	101-114	MP	
K.HLPTTFSDMAGYVIVNAQ.V	2437.08	AV392722	PRP L13	101-122	MP	
R.FGSIWTFESNELQAR.L	1735.69	AV392722	PRP L13	143-157	MP	
W.TFISNELQAR.L	1308.52	AV392722	PRP L13	147-157	MP	
K.KYKGTCHCHAGQPDITK.E	2224.44	BE211863	PRP L14	186-204	10L	
-MKPLSYLVADNSGARE	1850.88	CAA33226	PRP L14	4-17	MP	
K.P.SYLVADNSGARE	1477.62	CAA33226	PRP L14	6-17	MP	
R.SYLVADNSGARE	1267.39	CAA33226	PRP L14	4-17	MP	
KESALNPMVYK.D	1428.67	CAA33226	PRP L14	31-44	MP	
R.FDNNAAVINKK.E	984.95	CAA33226	PRP L14	45-53	MP	
R.LHLSPEQGR.R	1220.37	CA33226	PRP L14	147-157,MP	MP	
R.LHLSPEQGR.R	1207.33	AV390528	PRP L15	79-89	MP	
R.GRYAAGGGTCTGFNGRQK.A	2090.11	AV390528	PRP L15	34-44	MP	
R.GYAAAGGGTCTGFNGR.G	1474.24	AV390528	PRP L15	51-70	4L,5L	
R.GYAAAGGGTCTGFNGR.Q	1473.56	AV390528	PRP L15	53-67	MP	
R.GPSPVRRGEGGQITPLR.R	1906.11	AV390528	PRP L15	53-68	3T,5T	
R.PGEGGQITPLR.R	1322.46	AV390528	PRP L15	73-80	4T,MP	
R.PGEGGQITPLR.VVNLDDLEKH	2336.65	AV390528	PRP L15	79-90	4T	
R.PGEGGQITPLR.VVNLDDLEKH	2318.30	AV390528	PRP L15	97-119	5T	
R.PGEGGQITPLR.VVNLDDLEKH	1472.83	AV390528	PRP L15	105-119	MP	
K.HMAGEDVTLAAKEL.V	1719.93	AV390528	PRP L15	120-135	MP	
K.HMAGEDVTLAAKEL.V	1244.43	AV390528	PRP L15	138-149	2L,4L	
K.SISGRDASLPIK.V	2097.42	AV390528	PRP L15	138-158	4L	
R.DASLPIK.V	743.57	AV390528	PRP L15	143-149	MP	
K.SLNHATAFESSEK.A	1462.60	AV398154	PRP L16	159-172	4L,5L,MP	
K.VFDFEALQGEPC.W	1595.43	CA228835	PRP L16	27-40	MP	
K.VFDFEALQGEPC.W	2239.70	CA228835	PRP L16	27-45	MP	
K.PGDFVLAQGEPCWITSR.Q	1496.71	CA228835	PRP L16	86-101	12T,MP	
K.PDYVVAVVHPGK.I	1368.58	CA228835	PRP L16	90-101	12L	
K.ILYEMQVSETIAR.Q	1610.87	CA228835	PRP L16	102-115	12T,MP	
R.SYDVTIR.H	917.09	AV393223	PRP L17	79-86	MP	
K.VYDHMITLAK.D	1191.43	AV393223	PRP L17	103-112	MP	
R.QAMAFVYDQVGFENAFPER.Y	2543.44	AV393223	PRP L17	121-142	MP	
R.DVGNVFNAPER.Y	1475.53	AV393223	PRP L17	130-142	MP	
R.YAERPGGY.T	912.48	BE056672	PRP L17	143-150	MP	
K.SNEHYVQDDSK.G	1404.18	BE056672	PRP L17	161-173	MP	
K.GHTLYASATYMK.D	1647.78	AV396824	PRP L18	59-72	14L,15T,MP	
K.DKLESIGNANNAATAVGK.V	1231.45	AV396824	PRP L18	73-84	14L	
K.ELESIGNANNAATAVGK.V	2117.36	AV396824	PRP L18	85-105	14L	
K.ELESIGNANNAATAVGK.V	1632.77	AV396824	PRP L18	88-104	6L,13L,14T	
K.TNLPVAVGDNK.V	1760.94	AV396824	PRP L18	88-105	14L	
K.VGLAVQEGNGK.T	1326.54	BE227743	PRP L19	41-53	6L,13T,14L,MP	
R.TQIDGVIAAHG.T	1072.21	BE227743	PRP L19	54-64	MP	
R.TQIDGVIAAHG.T	1299.61	BE227743	PRP L19	67-78	MP	
R.TQIDGVIAAHGDSRR.T	1812.88	BE227743	PRP L19	67-84	MP	
R.VGGVGVVELVMPVHPVQIEVYR.P	2767.69	BE227743	PRP L19	91-115	MP	
L.VMPVHPVQIEVYR.P	1639.07	BE227743	PRP L19	100-115	MP	
K.VGLNYSERWNYLKT.I	1632.66	AF396929	PRP L20	72-84	MP	
K.VHPSPLTAPATL.A	1438.78	AV391113	PRP L21	65-78	MP	
K.LTLYRPLVAVVAVVEELRQK.V	2722.86	AV391113	PRP L21	133-144	MP	
K.PYLAVVAVVAVVEELRQK.V	1973.34	AV391113	PRP L21	128-144	MP	
K.SAATDAPAPVAVVAVVEELRQK.V	1654.07	AV398591	PRP L22	45-64	MP	
S.AATDAPAPVAVVAVVEELRQK.V	1766.99	AV398591	PRP L22	46-64	MP	
D.APAPVAVVAVVEELRQK.V	1408.84	AV398591	PRP L22	50-64	12L	
A.PAPVAVVAVVEELRQK.V	1337.56	AV398591	PRP L22	51-64	MP	
R.SYEEALVLCVMPYR.A	1098.54	AV398591	PRP L22	54-64	MP	
K.CVLSAANAK.H	1925.93	AV398591	PRP L22	81-105	MP	
K.VVSECTFADGGVPLK.R	1037.19	AV398591	PRP L22	113-122	MP	
K.VVSECTFADGGVPLK.R	1663.87	AV644217	PRP L22	132-146	MP	
K.VVSECTFADGGVPLK.R	2298.99	AF396929	PRP L23	40-59	MP	
R.AITLKE	658.52	AF396929	PRP L23	77-92	MP	
K.IGGGRKWEHLELPGGK.A	1792.05	AV398258	PRP L24	46-62	13L	
LGGGRKWEHLELPGGK.A	1678.89	AV398258	PRP L24	47-62	13T,MP	
R.KWEHLELPGGK.A	1351.55	AV398258	PRP L24	51-62	13L,MP	
K.KGDVYQIAGK.D	1130.33	AV398258	PRP L24	72-92	13L,MP	
K.KGDVYQIAGK.D	1002.16	AV398258	PRP L24	73-92	MP	
K.VVGVVQQGK.H	1157.31	AV398419	PRP L24	100-110	6L,13T,MP	
K.TLEFPHHSVMVLYS	1701.84	AV398419	PRP L24	125-138	MP	
K.TLEFPHHSVMVLYS	1917.23	AV398419	PRP L24	125-140	12L	
K.KVYKPEHYVYADHPK.A	1915.25	AV397116	PRP L27	100-115	12L	
K.KVYKPEHYVYADHPK.A	1787.97	AV397116	PRP L27	101-115	12L,13L	

This table was prepared in accordance with our previous paper (7). Peptide position indicates peptide location on precursor protein or iso-protein (see also supplemental figure).

TABLE II  
Percent similarities of *Chlamydomonas* 50S PRPs with homologs found in *Arabidopsis*, *Synechocystis*, and *E. coli*

Highest % similarity value is indicated by bold.

Protein <sup>a</sup>	<i>C. reinhardtii</i>		<i>Arabidopsis</i>			<i>Synechocystis</i>	<i>E. coli</i>
	Acc. no. <sup>b</sup>	Gene allocation	Acc. no.	Gene allocation	% similarity	% similarity	% similarity
50 S PRP							
L1	20011023.1212.1	Nucleus	At3g63490	Nucleus	<b>70</b>	63	63
L2	AF396929 ( <i>rpl2</i> ) <sup>c</sup>	Plastid	NP_051123	Plastid	<b>71</b>	<b>77</b>	70
L3	20011023.5271.1	Nucleus	At2g43030	Nucleus	<b>71</b>	70	63
L4	20011023.4457.1	Nucleus	At1g07320	Nucleus	<b>64</b>	<b>64</b>	54
L5	AF396929 ( <i>rpl5</i> ) <sup>c</sup>	Plastid	At4g01310	Nucleus	66	<b>73</b>	70
L6	20011023.5316.1	Nucleus	At1g05190	Nucleus	<b>69</b>	66	62
L9	20011023.3892.1	Nucleus	At3g44890	Nucleus	56	<b>62</b>	52
L10	20021010.7676.2	Nucleus	At5g13510	Nucleus	<b>52</b>	50	44
L11	20011023.6183.1	Nucleus	At1g32990	Nucleus	76	<b>82</b>	73
L12	20011023.1714.1	Nucleus	At3g27830	Nucleus	58	<b>69</b>	59
L13	20011023.1311.1	Nucleus	At1g78630	Nucleus	<b>77</b>	70	66
L14	AF396929 ( <i>rpl14</i> ) <sup>c</sup>	Plastid	NP_051094	Plastid	86	<b>87</b>	75
L15	20011023.1916.1	Nucleus	At3g25920	Nucleus	<b>66</b>	60	53
L16	AF396929 ( <i>rpl16</i> ) <sup>c</sup>	Plastid	NP_051095	Plastid	85	<b>87</b>	73
L17	20011023.6156.1	Nucleus	At3g54210	Nucleus	<b>74</b>	<b>74</b>	57
L18	20011023.1722.1	Nucleus	At1g48350	Nucleus	66	<b>74</b>	62
L19	20011023.3532.2	Nucleus	At4g17560	Nucleus	60	<b>68</b>	57
L20	AF396929 ( <i>rpl20</i> ) <sup>c</sup>	Plastid	NP_051082	Plastid	71	<b>82</b>	75
L21	20011023.3009.1	Nucleus	At1g35680	Nucleus	48	<b>67</b>	58
L22	20011023.2183.1	Nucleus	NP_051097	Plastid	71	<b>75</b>	61
L23	AF396929 ( <i>rpl23</i> ) <sup>c</sup>	Plastid	NP_051122	Plastid	58	<b>65</b>	54
L24	20011023.3290.3	Nucleus	At5g54600	Nucleus	61	<b>74</b>	56
L27	20011023.6913.1	Nucleus	At5g40950	Nucleus	61	<b>75</b>	69
L28	20011023.1369.1	Nucleus	At2g33450	Nucleus	65	<b>75</b>	57
(L29)			At5g65220	Nucleus			
L31	20011023.6817.1	Nucleus	At1g75350	Nucleus	62	<b>69</b>	55
L32	20011023.267.1	Nucleus	NP_051107	Plastid	<b>76</b>	62	39
(L33)	20011023.6737.1	Nucleus	NP_051080	Plastid	<b>64</b>	<b>64</b>	60
(L34)	20011023.570.1	Nucleus	At1g29070	Nucleus	61	<b>86</b>	80
L35	20011023.3217.1	Nucleus	At2g24090	Nucleus	47	<b>56</b>	47
(L36)	AF396929 ( <i>rpl36</i> ) <sup>c</sup>	Plastid	NP_051092	Plastid	<b>93</b>	91	81
(PSRP-5)			At3g56910	Nucleus			
PSRP-6	2001 1023.3600.1	Nucleus	At5g17870	Nucleus	<b>50</b>	ND <sup>d</sup>	ND <sup>d</sup>
70 S-specific proteins							
RAP38	AY177617	Nucleus	At1g09340 <sup>e</sup>	Nucleus	<b>84</b>	67	ND <sup>d</sup>
RAP41	AY177616	Nucleus	At3g63140 <sup>e</sup>	Nucleus	<b>62</b>	ND <sup>d</sup>	ND <sup>d</sup>

<sup>a</sup> Unidentified PRPs in this proteomic study were searched using *Arabidopsis* PRPs as probes and are indicated between parentheses.

<sup>b</sup> Accession numbers of EST contigs were obtained from the ChlamyEST database.

<sup>c</sup> Gene in the chloroplast genome (29).

<sup>d</sup> Homologs are not detected in database searches.

<sup>e</sup> Probably non-ribosomal protein.

Therefore, the protein composition of the 70 S ribosome is not necessarily equal to the sum of the 30 S proteins and the 50 S proteins. For this reason, we also analyzed proteins in the *C. reinhardtii* 70 S subunit. We had previously identified two proteins in the TP70 pattern resolved by SDS-PAGE that were not seen in the TP30 or TP50 patterns (7). Here we confirmed that the 38- and 41-kDa proteins cosedimented with the plastid 70 S ribosomes but not with the plastid 30 S or 50 S subunits nor with cytoplasmic 80 S ribosomes (Fig. 1D). Because the 38- and 41-kDa proteins were associated with the 70 S ribosome, we have designated these proteins RAP38 and RAP41 (Ribosome Associated Protein of 38 kDa and 41 kDa). As shown in Fig. 3, Western blot analysis of proteins from whole cells and from isolated 30 S, 50 S, and 70 S ribosomes revealed that RAP38 is associated with purified 70 S subunits but not with either of the purified dissociated subunits (30 S or 50 S). Anti-S1 and anti-L2 antibodies were used as controls to show that the ribosomal preparations contained the RP expected for each fraction. The SDS-PAGE profile of TP70 (Fig. 1, D and E) also shows that the RAP38 and RAP41 are present in similar amounts as other RPs (e.g. PRP S1 or PSRP-3), judged by the intensity of the Coomassie staining of these proteins. Thus, RAP38 and RAP41 appear to be present in stoichiometric amounts on the ribosome.

No homolog of spinach pRRF was identified from the *C. rein-*

*hardtii* 70 S ribosome. However, pRRF is likely present as a non-RP in *C. reinhardtii*, because the corresponding EST can be identified by computational search of the *C. reinhardtii* EST database. Homologs of 19 CRPs S6, S8, S17, S25, L3, L7, L11, L13a, L17, L1a, L23a, L24, L26, L27, L27a, L32, L36, P0, and P1 were also identified from our TP70 fractions by the sensitive LC-MS/MS analysis. Some of these proteins were also identified in the TP30 (7) and TP50 fractions (this study). These proteins are contaminations of cytoplasmic 80 S ribosomes and their subunits because all 19 proteins were identified in purified 80 S ribosomes.<sup>5</sup> None of these 19 CRPs show putative chloroplast targeting sequences.

*RAP38 and RAP41 Are Homologs of a Cyanobacterial Ancestor Related to Bacterial Epimerase/Dehydratase*—As summarized in Table I, five tryptic peptides of RAP38 and a Lys-C peptide of RAP41 were identified from section 8 and section 7 of the TP70 gel. The peptides belonged to ESTs AV620219 and AV621571 encoding RAP38 and to EST AV635725 encoding RAP41. We obtained the longest EST clone for each protein (AV620219 for RAP38; AV635725 for RAP41) from the Kazusa DNA Research Institute and sequenced the entire EST. The sequence data of RAP38 and RAP41 were submitted to the

<sup>5</sup> K. Yamaguchi, P. A. Haynes, and S. P. Mayfield, unpublished data.



TABLE III  
 Characteristics of *C. reinhardtii* 50 S PRPs and 70 S-specific RAPs

Protein	Predicted chain length and Isoelectric point				Molecular mass (kDa)	
	Precursor or pro-protein <sup>a</sup>	Transit peptide <sup>b</sup> or Met removal <sup>c</sup>	Mature protein	pI <sup>d</sup>	Sequence mass	SDS-PAGE <sup>e</sup> section/band
50 S PRP						
L1	297	36	261	9.4	28.5	26-29
L2	278	Met (-)	277	11.1	30.8	26-29
L3	259	23	236	10.0	25.4	23-24
L4	243	22 <sup>f</sup>	221	9.8	23.2	21-22
L5	206	Met (-)	205	9.9	23.1	17-19
L6	207	10	197	9.8	21.0	17-19
L9	204	21	183	9.6	20.2	14-16
L10	>230	29 <sup>g</sup>	>201	9.5	>22.2	21-23
L11	176	36	140	9.8	14.7	11-13
L12	162	32	130	4.9	13.4	11-13
L13	225	30	195	9.9	21.3	19-21
L14	122	Met (+)	122	10.3	13.4	10-11
L15	246	25	221	10.2	23.3	22-23
L16	136	Met (+)	136	11.5	15.5	ND <sup>g</sup>
L17	173	49	124	10.5	14.2	11-13
L18	145	26	119	9.7	12.9	10-11
L19	153	31	122	11.2	13.2	11-13
L20	112	Met (-)	111	11.8	13.4	ND <sup>g</sup>
L21	179	51	128	9.7	14.2	ND <sup>g</sup>
L22	175	43	132	9.6	14.5	11-13
L23	95	Met (+)	95	10.0	11.2	ND <sup>g</sup>
L24	170	33	137	10.0	14.9	11-13
L27	161	29 <sup>g</sup>	132	10.4	14.1	13-14
L28	195	33 <sup>i</sup>	162	9.8	18.4	16-17
L31	136	31	105	9.2	11.6	<10
L32	98	36	62	9.7	6.9	10-11
L33 <sup>f</sup>	86	11	75	10.2	8.6	NI <sup>h</sup>
L34 <sup>f</sup>	124	64	60	11.9	6.5	NI <sup>h</sup>
L35	114	42 <sup>i</sup>	72	10.5	7.9	<10
L36 <sup>f</sup>	37	Met (+)	37	10.8	4.3	NI <sup>h</sup>
PSRP-6	66	13	53	10.3	5.9	<10
70 S-specific proteins						
RAP38	401	41	360	9.2	40.2	38
RAP41	439	35	404	9.3	44.3	41

<sup>a</sup> Cytosolic precursor and plastid pro-protein sequences deduced from nucleic acid sequences.

<sup>b</sup> Predicted by ChloroP program (48).

<sup>c</sup> Predicted by penultimate amino acid residues (66).

<sup>d</sup> From predicted mature protein sequence.

<sup>e</sup> Estimated from major peptides found in the TP50/70-gel section (see Fig. 1E and Table I).

<sup>f</sup> Unidentified proteins by LC-MS/MS analyses.

<sup>g</sup> ND, not determined due to MudPIT identification.

<sup>h</sup> NI, not identified on this proteomic analysis.

<sup>i</sup> Predicted by mass spectrometric results (see supplemental figure).

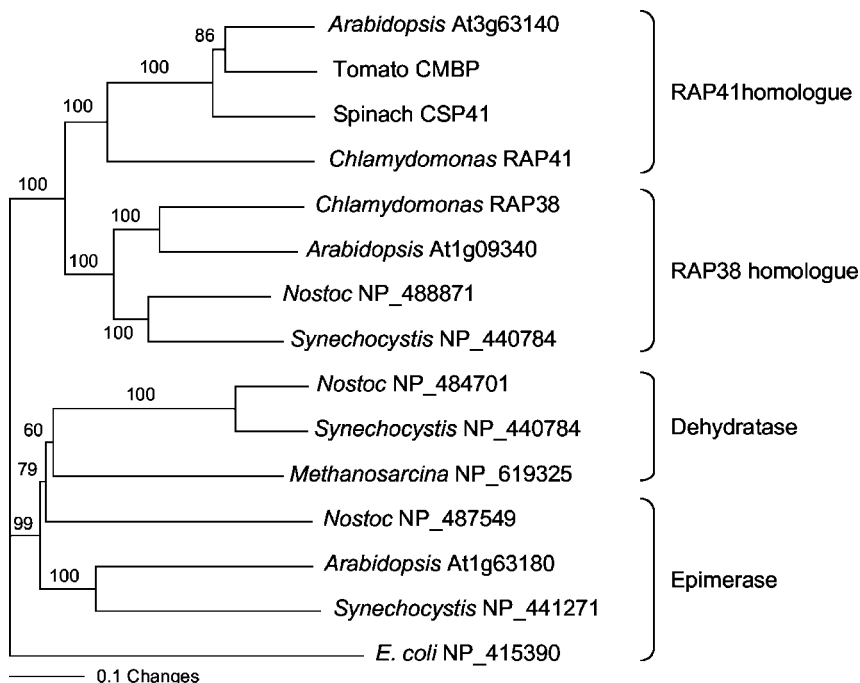
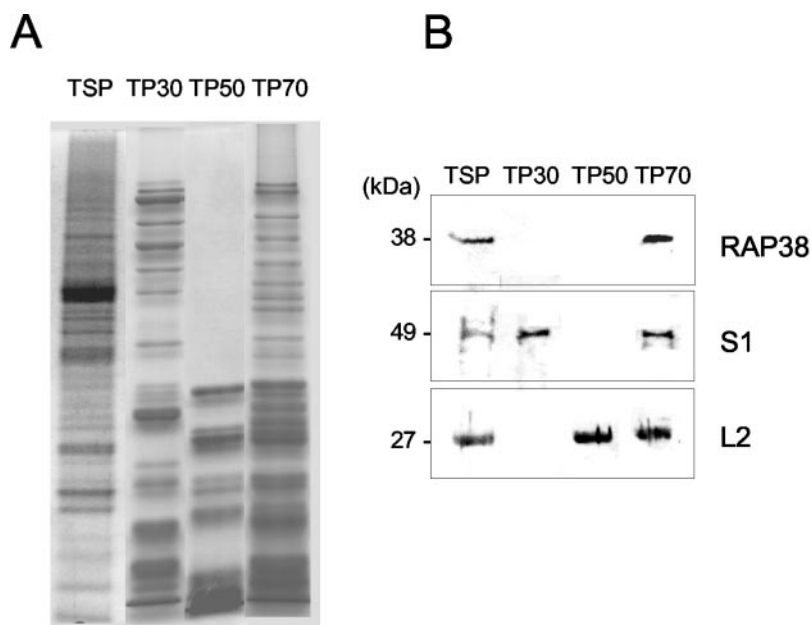
<sup>j</sup> N-terminal sequence reported (60).

GenBank database under accession numbers AY177617 (gene name: *Rap38*) and AY177616 (gene name: *Rap41*), respectively. RAP38 and RAP41 cDNAs encoded polypeptides of 401 and 439 amino acids, respectively. Chloroplast transit peptides of RAP38 and RAP41 were estimated to be 41 and 35 amino acids using the ChloroP program (48); thus, the mature protein chain lengths are 360 and 404 amino acids, respectively. Molecular masses of the mature RAP38 and RAP41 calculated from the sequences are 40.2 and 44.3 kDa, which are in reasonable agreement with molecular mass estimated by SDS-PAGE (Table III). Pairwise comparison of RAP38 and RAP41 showed 35% identity and 52% similarity to one another, indicating that these are related proteins. A BLAST search revealed that RAP38 and RAP41 have significant sequence similarity with a higher plant homolog of spinach, CSP41, a chloroplast RNA-binding protein/endoribonuclease (52), and hypothetical proteins of cyanobacteria (see Table II for percent sequence similarities). RAP38 and RAP41 also showed a distant but still significant sequence similarity with several sugar-nucleotide epimerases/dehydratases of archaea, eubacteria, and eukarya.

To examine phylogenetic and evolutionary relationships between RAP proteins and their homologs, we performed an analysis using a neighbor-joining method, and the reliability of the

created tree was estimated by bootstrapping (Fig. 4). The distance tree was constructed by rooting protein members of epimerase/dehydratase family as the outgroup, because proteins belonging to this enzyme family are well conserved between eubacteria and eukaryotes, including one of the earliest lineage in *Methanosarcina* (an archaeobacterium). The tree shows three distinct clades: 1) RAP41 homologs, 2) RAP38 homologs, and 3) the epimerase/dehydratase family. No RAP homologs are identifiable in the complete genome sequences of *E. coli* or *Methanosarcina*. In contrast, a hypothetical protein encoded by a single gene (NP\_440784) can be identified as a RAP38 homolog in the complete genome sequence of the cyanobacterium, *Synechocystis* sp. PCC6803. Similarly, another cyanobacterium, *Nostoc* sp. PCC 7120, has a RAP38-related gene (NP\_488871). The *Arabidopsis* genome contains both a RAP38 homolog (At1g09340) and a RAP41 homolog (At3g63140). RAP41 and its homolog were identified only in algae and plants and were not found in photosynthetic bacteria. Thus, phylogenetic analysis suggests that RAP38 and RAP41 are paralogs of a cyanobacterial ancestor diverged from bacterial epimerase/dehydratase, and that RAP41 is closely related to spinach CSP41. This phylogenetic analysis is consistent with a previously reported analysis that was predicted by a

**FIG. 3. RAP38 is present in purified 70 S subunits.** Ribosomes were purified on sucrose gradients as described. Samples were collected and analyzed by SDS-PAGE (A) and Western blot (2  $\mu\text{g}/\text{lane}$ ) with anti-RAP38 (upper panel), anti-S1 (middle panel), and anti-L2 (lower panel) polyclonal antisera (B). TSP, total soluble proteins from *C. reinhardtii*. TP30, TP50, and TP70, total proteins from the 30 S, 50 S, and 70 S subunits, respectively.



**FIG. 4. Distance tree of *C. reinhardtii* RAPs and their homologs.** The phylogram was rooted with proteins belonging to the epimerase/dehydratase family. The tree was constructed by the neighbor-joining method. Numbers between internal nodes indicate bootstrap values as percentages of 1000 replications. See "Experimental Procedures" for details on the analysis. *Tomato CMBP*, chloroplast mRNA-binding protein (64); *Spinach CSP41*, chloroplast stem-loop RNA-binding protein/endoribonuclease (52); accession numbers are indicated for other proteins.

conserved motif analysis, ancient divergence from a common ancestor of epimerases/dehydratase and an mRNA-binding protein with ribonuclease activity (53). Fig. 5 shows the amino acid sequence alignment of RAP38, RAP41, and their homologs, along with the motif structures as determined in the three-dimensional analysis of epimerases and dehydrogenases (53). The N-terminal  $\beta\text{A}-\alpha\text{B}-\beta\text{B}$  region of the sugar-nucleotide epimerase/dehydratase involved in binding the adenine of NAD (54) is significantly conserved with RAPs and their homologs.

**Unexpected PRP Gene Allocation and Evolution in *C. reinhardtii***—In algae and plants, the PRP genes are partially distributed in the plastid genome and partially in the nuclear genome. In higher plants, the PRP genes are distributed between the two genomes, with approximately one-third of  $\sim 60$  PRP genes found in the chloroplast and the remaining two-thirds in the nucleus (5, 6, 55). The transfer of PRP genes to the nuclear genome from the plastid genome occurred during plant/alga evolution. Unexpectedly, our proteomic analyses of the *C. rein-*

*hardtii* chloroplast ribosome revealed that the number of transferred PRP genes in the nucleus of this alga is greater than that in higher plants. We reported previously that the *C. reinhardtii* PRP genes for S15 and S16 are nuclear genes, whereas these genes are plastid genes in higher plants (7). Likewise, *C. reinhardtii* PRP genes for L22, L32, and L33 are allocated in the nucleus, whereas these genes are maintained in plastids in *Arabidopsis* and other higher plants (Table II). In *C. reinhardtii*, the *prpL21* gene is not present in plastid genomes, whereas our proteomic analysis identified a nuclear *PrpL21* gene (Tables I and II). In higher plants, the *prpL21* gene has been lost from plastid genomes after the divergence from bryophytes, and a nuclear gene of mitochondrial origin has replaced the chloroplast gene (56). This is not the case for the *PrpL21* nuclear gene of *C. reinhardtii*, in which phylogenetic analysis suggests that the gene originated from an ancestral plastid gene sharing significant homology with *rpL21* of cyanobacteria and other algal plastid genomes (data not shown).

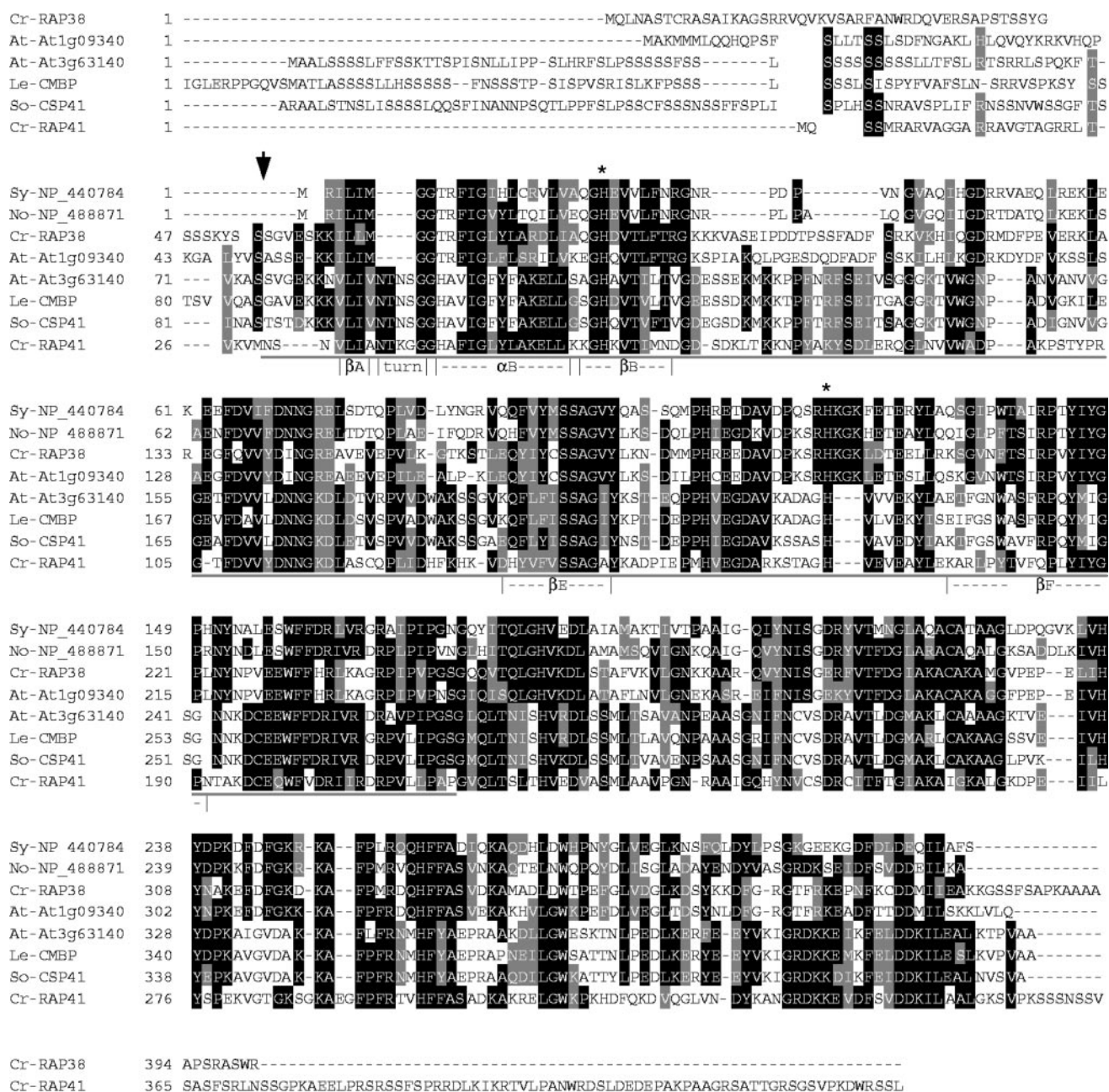


FIG. 5. Amino acid sequence alignment of RAP38 and RAP41 precursors with related proteins from higher plants and photosynthetic bacteria. The arrow indicates the chemically determined cleavage site of the chloroplast transit peptide of spinach CSP41 (52). The catalytic domain of spinach CSP41 (amino acids 1–191 of mature protein), by analyzing C-terminal deletions (59), is underlined. Asterisks indicate two histidine residues, which may take part in the catalytic site. Identical residues are boxed in black, and similar residues are boxed in gray. Cr, *Chlamydomonas reinhardtii*; At, *Arabidopsis thaliana*; Sy, *Synechocystis* sp. PCC 6803; So, *Spinacia oleracea* (spinach); No, *Nostoc* sp. PCC 7120; Le, *Lycopersicon esculentum* (tomato).

## DISCUSSION

We previously characterized the *C. reinhardtii* 30 S ribosomal subunit via a proteomic analysis and identified 21 proteins: 19 *E. coli* orthologs, a spinach PSRP-3 homolog, and a novel protein, PSRP-7, that contains two S1 domains (7). We identified unusually large orthologs of bacterial S2 (57 kDa), S3 (76 kDa), and S5 (84 kDa) that are significantly larger than their counterparts in higher plants and bacteria. PSRP-7 is also an unusually large RP. The total mass of *C. reinhardtii* 30 S proteins was estimated to be ~600 kDa, which is 1.7-fold and 1.4-fold greater than those of *E. coli* 30 S (350 kDa) and spinach 30 S (430 kDa) proteins. To fully understand the chloroplast translation machinery, we have now characterized the components of the 50 S subunit and the 70 S ribosome of the *C. reinhardtii* chloroplast. From the 50 S

subunit, 28 proteins were identified: 27 *E. coli* orthologs and a homolog of spinach PSRP-6. As shown in Fig. 6, the molecular masses of *C. reinhardtii* 50 S PRPs range from 4.3 kDa (L36) to 30.8 kDa (L2), and most of the 50 S PRPs are similar in size and sequence to their counterparts in spinach and *E. coli* (Table III and supplemental figure). The *C. reinhardtii* PRP L28 is twice as large as its spinach and *E. coli* counterparts because of the presence of a C-terminal extension. Other large subunit RPs show short insertion sequences or extension at their N-/C-termini (see supplemental figure), but these short extensions do not affect the net mass of the large subunit. The total mass of the *C. reinhardtii* 50 S PRPs is estimated to be ~490 kDa, which is similar to those of *E. coli* 50 S (440 kDa) and spinach 50 S (530 kDa) subunits.

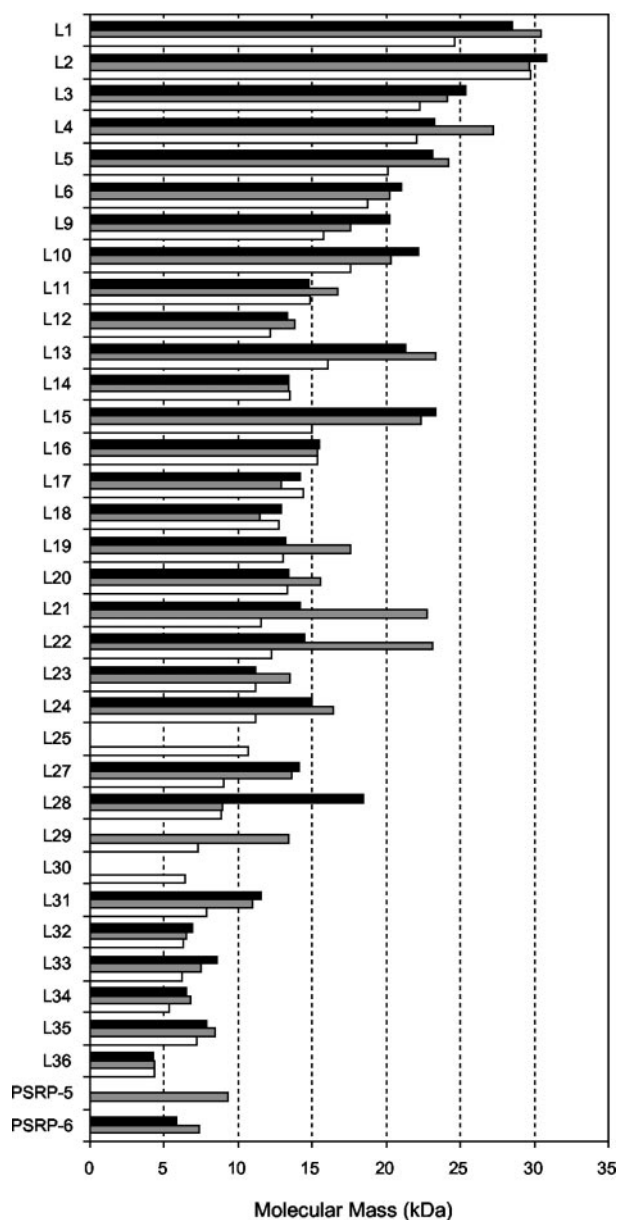


FIG. 6. Comparison of the molecular masses of 50 S RPs. Molecular masses of *C. reinhardtii* 50 S PRPs (black columns) were compared with those of spinach 50 S PRPs (gray columns) and *E. coli* 50 S RPs (white columns). Molecular masses shown in this histogram are mature spinach 50 S PRPs (6), mature *E. coli* 50 S RPs (65), and predicted mature *C. reinhardtii* 50 S PRPs from Table III.

Previous studies have examined immunological similarity of the chloroplast RPs of *C. reinhardtii* with RPs of spinach chloroplasts, the cyanobacterium *Anabaena*, and *E. coli* (27, 28). Using 15 selected antisera against *C. reinhardtii* S-7, S-11, S-19, S-20, S-21, L-1, L-6, L-13, L-17, L-18, L-21, L-26, L-27, L-29, and L-30 (see Ref. 26 for nomenclature), 9 cross-reacting *E. coli* proteins were identified: S1 (anti-S-7 and -S-11), S3 (anti-S-7), L2 (anti-L-1), L5 (anti-L-13), L9 (anti-L-30), L12 (anti-L-26, -L-29, and -L-30), L16 (anti-L-17), L23 (anti-L-29), and L27 (anti-L-18). Among them, L-18 has been identified as the homolog of *E. coli* L27 by N-terminal protein sequencing (57). In general, the greatest immunological similarity is found between *Anabaena* and *Chlamydomonas*, and the similarity is greatest for chloroplast-synthesized proteins of the large ribosomal subunit. As shown in Table II, a majority of the *C. reinhardtii* chloroplast RPs have greater sequence similarity with cyanobacterium (*Synechocystis*) than with higher plants

(*Arabidopsis*) or eubacterium (*E. coli*). Table II also shows that plastid-encoded proteins are more conserved between different organisms than nuclear-encoded proteins. Our results (this study and Ref. 7) have confirmed previous observations on the conservation of proteins within 70 S ribosomes and have expanded on these studies to include a complete molecular description of the entire set of RPs from *C. reinhardtii* chloroplasts.

Four of the *C. reinhardtii* PRPs are more than twice as large as the spinach and *E. coli* counterparts because of either N-terminal extensions (NTEs), insertion sequences, or C-terminal extensions (C-terminal extension): S2 (NTE), S3 (insertion sequence), PSRP-3 (NTE), and L28 (C-terminal extension). Recently, sensitive profile search techniques, such as PSI-BLAST, have identified a novel conserved domain, TRAM, predicted to be an RNA-binding domain common to tRNA uracil methylation and adenine thiolation enzymes (58). A TRAM domain is also identifiable from eukaryotic translation initiation factor-2 $\beta$  in *Thermoplasma*. Two TRAM domains are found at the N terminus of PRP S2 (ORF570) in *C. reinhardtii*. Accordingly, we speculate that the NTEs, C-terminal extensions, and insertion sequence domains in *C. reinhardtii* PRPs S2, S3, PSRP-3, and L28 may take part in regulation of translation, perhaps by directly binding chloroplast mRNAs with TRAM domains or other uncharacterized RNA-binding domains.

Two nuclear-encoded proteins have been identified from the 70 S ribosome that were not found in either of the subunits. These two ribosome-associated proteins were designated RAP38 and RAP41. Western blot analysis of total proteins and proteins from isolated ribosomal subunits using RAP38 antisera revealed that RAP38 is associated with 70 S ribosomes only and not with the individual subunits. These data give rise to at least two different hypotheses: that RAP38 and RAP41 have a role in nontranslating 70 S ribosomes, or that RAP38 and RAP41 participate in translation but associate with the 70 S ribosome after the subunits have initiated translation. The role of these proteins is being studied using genetic knock-outs and in *in vitro* and *in vivo* experiments. In the proteomic characterization of the spinach 70 S chloroplast ribosome (6), RAP38- and RAP41-like proteins were not identified, although a pRRF was identified as a 70 S-specific protein. Identification of RAP38 and RAP41 from the 70 S subunit is entirely unexpected, because these two proteins share sequence similarity with spinach CSP41, a chloroplast RNA-binding protein with endoribonuclease activity. How does the *C. reinhardtii* chloroplast ribosome contain RAP38 and RAP41, whereas the spinach chloroplast ribosome does not? In *C. reinhardtii*, RAP38 and RAP41 have C-terminal extensions of 16 and 83 amino acids, respectively, compared with spinach CSP41. These additional domains may allow RAP38 and RAP41 to bind the *C. reinhardtii* 70 S ribosome with high affinity. Several additional domains in the 30 S and 50 S proteins unique to the *C. reinhardtii* (e.g. NTE of S2, C-terminal extension of L28) may also interact with the RAP proteins to allow binding to the ribosome.

The *Arabidopsis* genome contains both RAP38 and RAP41 homologs, whereas the cyanobacterial lineage has only a RAP38-like protein. Phylogenetic analysis suggests that RAP38 and RAP41 are paralogs of a cyanobacterial protein, but this protein has not yet been biochemically characterized. Sequence and motif homologies between CSP41, RAP38, and RAP41 suggest that similar RNA-binding and nuclease activities can be expected for these proteins as well. Histidine residues are known to be important for catalytic function in ribonucleases, and three histidine residues (His-34, His-130, and His-140) found in the catalytic domain (amino acid 1–191 re-

gion; Ref. 59) of CSP41 are conserved between CSP41 homologs and RAP proteins, whereas these histidines are not conserved in epimerases/dehydratases. In spinach, CSP41, CSP29, and CSP55 are thought to form a stem-loop RNA-protein complex that is required for processing of the *petD* mRNA at the 3' end (reviewed in Ref. 11). Similarly, 5' UTR processing of some chloroplast mRNAs is necessary for translation in algae and higher plants. We have shown previously that processing of the *C. reinhardtii psbA* mRNA 5' UTR depends upon factors mediating ribosome association (60), whereas failure to process the *C. reinhardtii petD* mRNA 5' UTR is also associated with the loss of translation (61, 62). In barley chloroplast, *rbcL* mRNA with an unprocessed 5' UTR is not associated with polyribosomes (63). We propose that RAP38 and RAP41 are functional counterparts of CSP41 and may be involved in mRNA processing, directly or indirectly with other factors and/or RPs.

The proteomic, bioinformatics, and biochemical analyses presented here have shown that the chloroplast translation machine of *C. reinhardtii* is compositionally and structurally similar to those of higher plants and eubacteria, but that *C. reinhardtii* ribosomes contain additional domains and proteins previously unidentified in other 70 S ribosomes. Many of these additional domains reside on the 30 S subunit of the ribosome, the subunit responsible for mRNA discrimination during translation initiation. The part of the ribosome responsible for peptidyl bond formation between amino acids, the 50 S subunit, is more conserved. These results are not unexpected, because peptide bond formation would be expected to be a conserved enzymatic function, whereas mRNA discrimination would be anticipated to be organism dependent. Defining specific roles for individual proteins and domains in chloroplast translational regulation should now be greatly facilitated by having the complete set of RPs defined.

**Acknowledgments**—We thank Dr. Alap R. Subramanian for generous gifts of antisera against spinach chloroplast RPs S1 and L2, Dr. Benjamin F. Cravatt for reviewing the manuscript, and Kazusa DNA Research Institute for providing EST clones.

## REFERENCES

- Mayfield, S. P., Yohn, C. B., Cohen, A., and Danon, A. (1995) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 147–166
- Bogorad, L. (1975) *Science* **188**, 891–898
- Gray, M. W. (1992) *Int. Rev. Cytol.* **141**, 233–357
- Harris, E. H., Boynton, J. E., and Gillham, N. W. (1994) *Microbiol. Rev.* **58**, 700–754
- Yamaguchi, K., von Knoblauch, K., and Subramanian, A. R. (2000) *J. Biol. Chem.* **275**, 28455–28465
- Yamaguchi, K., and Subramanian, A. R. (2000) *J. Biol. Chem.* **275**, 28466–28482
- Yamaguchi, K., Prieto, S., Beligni, M. V., Haynes, P. A., McDonald, W. H., Yates, J. R., III, and Mayfield, S. P. (2002) *Plant Cell* **14**, 2957–2974
- Yamaguchi, K., and Subramanian, A. R. (2003) *Eur. J. Biochem.* **270**, 190–205
- Rochaix, J. D. (1996) *Plant Mol. Biol.* **32**, 327–341
- Sugita, M., and Sugiura, M. (1996) *Plant Mol. Biol.* **32**, 315–326
- Monde, R. A., Schuster, G., and Stern, D. B. (2000) *Biochimie (Paris)* **82**, 573–582
- Zerges, W. (2000) *Biochimie (Paris)* **82**, 583–601
- Somanchi, A., and Mayfield, S. P. (2001) in *Advances in Photosynthesis and Respiration*, Vol. 11 (Aro, E.-M., and Andersson, B., eds) pp. 137–151, Kluwer Academic Publishers, Dordrecht, the Netherlands
- Fromm, H., Devic, M., Fluhr, R., and Edelman, M. (1985) *EMBO J.* **4**, 291–295
- Inamine, G., Nash, B., Weissbach, H., and Brot, N. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 5690–5694
- Laing, W., Kruezn, K., and Apel, K. (1988) *Planta* **176**, 269–276
- Malnoë, P., Mayfield, S. P., and Rochaix, J. D. (1988) *J. Cell Biol.* **106**, 609–616
- Berry, J. O., Breiding, D. E., and Klessig, D. F. (1990) *Plant Cell* **2**, 795–803
- Keller, M., Chan, R. L., Tessier, L. H., Weil, J. H., and Imbault, P. (1991) *Plant Mol. Biol.* **17**, 73–82
- Bruick, R. K., and Mayfield, S. P. (1999) *Trends Plant Sci.* **4**, 190–195
- Barnes, D., and Mayfield, S. P. (2003) *Antioxid. Redox Signal.* **5**, 89–94
- Harris, E. H. (2001) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 363–406
- Dent, R. M., Han, M., and Niyogi, K. K. (2001) *Trends Plant Sci.* **6**, 364–371
- Hanson, M. R., Davidson, J. N., Mets, L. J., and Bogorad, L. (1974) *Mol. Gen. Genet.* **132**, 105–118
- Brugger, M., and Boschetti, A. (1975) *Eur. J. Biochem.* **58**, 603–610
- Schmidt, R. J., Richardson, C. B., Gillham, N. W., and Boynton, J. E. (1983) *J. Cell Biol.* **96**, 1451–1463
- Schmidt, R. J., Myers, A. M., Gillham, N. W., and Boynton, J. E. (1984) *Mol. Biol. Evol.* **1**, 317–334
- Randolph-Anderson, B. L., Gillham, N. W., and Boynton, J. E. (1989) *J. Mol. Evol.* **29**, 68–88
- Maul, J. E., Lilly, J. W., Cui, L., de Pamphilis, C. W., Miller, W., Harris, E. H., and Stern, D. B. (2002) *Plant Cell* **14**, 2659–2679
- Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H., and Tabata, S. (1999) *DNA Res.* **6**, 369–373
- Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y., and Tabata, S. (2000) *DNA Res.* **7**, 305–307
- Bourque, D. P., Boynton, J. E., and Gillham, N. W. (1971) *J. Cell Sci.* **8**, 153–183
- Margulies, M. M., and Tiffany, H. L. (1979) *Biochim. Biophys. Acta* **563**, 171–180
- Price, C. A., and Reardon, E. M. (2001) *Nucleic Acids Res.* **29**, 118–119
- Wool, I. G., Chan, Y. L., and Gluck, A. (1995) *Cell Biol.* **73**, 933–947
- Cohen, A., Yohn, C. B., Bruick, R. K., and Mayfield, S. P. (1998) *Methods Enzymol.* **297**, 192–209
- Chua, N. H., Blobel, G., and Siekevitz, P. (1973) *J. Cell Biol.* **57**, 798–814
- Laemmli, U. K. (1970) *Nature* **229**, 680–685
- Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S., and Mische, S. M. (1999) *Electrophoresis* **20**, 601–605
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68**, 850–858
- Haynes, P. A., Fripp, N., and Aebersold, R. (1998) *Electrophoresis* **19**, 939–945
- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., III. (1999) *Nat. Biotechnol.* **17**, 676–682
- Washburn, M. P., Wolters, D., and Yates, J. R., III. (2001) *Nat. Biotechnol.* **19**, 242–247
- Wolters, D. A., Washburn, M. P., and Yates, J. R., III. (2001) *Anal. Chem.* **73**, 5683–5690
- Eng, J., McCormack, A. L., and Yates, J. R., III. (1994) *J. Am. Soc. Mass Spectrom.* **5**, 976–989
- Yates, J. R., III, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) *Anal. Chem.* **67**, 1426–1436
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* **25**, 4876–4882
- Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999) *Protein Sci.* **8**, 978–984
- Swofford, D. L. (2002). *PAUP\*, Phylogenetic Analysis Using Parsimony (\*and Other Methods)*, Version 4.0b10, Sinaur Associates, Sunderland, MA
- Sager, R., and Hamilton, M. G. (1967) *Science* **157**, 709–711
- Chua, N. H., Blobel, G., and Siekevitz, P. (1973) *Biochem. Soc. Symp.* **38**, 163–174
- Yang, J., Schuster, G., and Stern, D. B. (1996) *Plant Cell* **8**, 1409–1420
- Baker, M. E., Grundy, W. N., and Elkan, C. P. (1998) *Biochem. Biophys. Res. Commun.* **248**, 250–254
- Murzin, A. (1994) *Struct. Biol.* **1**, 146–147
- Subramanian, A. R. (1993) *Trends Biochem. Sci.* **18**, 177–181
- Gallois, J. L., Achard, P., Green, G., and Mache, R. (2001) *Gene* **274**, 179–185
- Liu, X. Q., Gillham, N. W., and Boynton, J. E. (1988) *Mol. Gen. Genet.* **214**, 588–591
- Anantharaman, V., Koonin, E. V., and Aravind, L. (2001) *FEMS Microbiol. Lett.* **197**, 215–221
- Yang, J., and Stern, D. B. (1997) *J. Biol. Chem.* **272**, 12874–12880
- Bruick, R. K., and Mayfield, S. P. (1998) *J. Cell Biol.* **143**, 1145–1153
- Sakamoto, W., Chen, X., Kindle, K. L., and Stern, D. B. (1994) *Plant J.* **6**, 503–512
- Sakamoto, W., Sturm, N. R., Kindle, K. L., and Stern, D. B. (1994) *Mol. Cell Biol.* **14**, 6180–6186
- Reinbothe, S., Reinbothe, C., Heintzen, C., Seidenbecher, C., and Parthier, B. (1993) *EMBO J.* **12**, 1505–1512
- Vian, A., Henry-Vian, C., and Davies, E. (1999) *Plant Physiol.* **121**, 517–524
- Arnold, R. J., and Reilly, J. P. (1999) *Anal. Biochem.* **269**, 105–112
- Giglione, C., and Meinel, T. (2001) *Trends Plant Sci.* **6**, 566–572

**Proteomic Characterization of the *Chlamydomonas reinhardtii* Chloroplast Ribosome: IDENTIFICATION OF PROTEINS UNIQUE TO THE 70 S RIBOSOME**

Kenichi Yamaguchi, María Verónica Beligni, Susana Prieto, Paul A. Haynes, W. Hayes McDonald, John R. Yates III and Stephen P. Mayfield

*J. Biol. Chem.* 2003, 278:33774-33785.

doi: 10.1074/jbc.M301934200 originally published online June 24, 2003

---

Access the most updated version of this article at doi: [10.1074/jbc.M301934200](https://doi.org/10.1074/jbc.M301934200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2003/07/17/M301934200.DC1>

This article cites 65 references, 18 of which can be accessed free at

<http://www.jbc.org/content/278/36/33774.full.html#ref-list-1>