The Plastid Ribosomal Proteins (2): Identification of all the Proteins in the 50S Subunit of an Organelle Ribosome (Chloroplast)

Running title: Plastid ribosomal 50S subunit proteins

Kenichi Yamaguchi and Alap R. Subramanian*

From the Department of Biochemistry, The University of Arizona, Tucson, USA and the Max-Planck-Institut für molekulare Genetik, Berlin-Dahlem, Germany

*Corresponding Author:
Dr. A. R. Subramanian, Department of Biochemistry, BioSciences West, The University of Arizona, Tucson, AZ 85721-0088
Fax: +1 520 621 9288
E-mail: subraman@u.arizona.edu

The protein and nucleic acid sequences reported in this paper have been deposited in SWISS-PROT and GenBank databases, respectively. The accession numbers (total of over 40) are listed in Table I and in text.
ABBREVIATIONS

The abbreviations used are: RP, ribosomal protein; PRP, plastid ribosomal protein; PSRP, plastid-specific ribosomal protein; RRF, ribosome recycling factor; TP50, total protein from 50S subunit; TP70, total protein from 70S ribosome; 2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; HPLC, high-performance liquid chromatography; MS, mass spectrometry; LC/MS, reversed-phase HPLC coupled to electrospray ionization mass spectrometry; ESI MS, electrospray ionization mass spectrometry; EST, expressed sequence tag; AA, amino acid(s); NTE, N-terminal extension; CTE, C-terminal extension.
SUMMARY

We have completed identification of all the ribosomal proteins (RP's) in spinach plastid (chloroplast) ribosomal 50S subunit via a proteomic approach using 2-dimensional electrophoresis, electroblotting/protein sequencing, HPLC purification, PCR-based screening of cDNA library/nucleotide sequencing and mass spectrometry (LC/MS and ESI MS). Spinach plastid 50S subunit comprises 33 proteins, of which 31 are orthologues of *E. coli* RP's and 2 are plastid-specific RP's (PSRP-5 and PSRP-6) having no homologues in other types of ribosomes. Orthologues of *E. coli* L25 and L30 are absent in spinach plastid ribosome. Twenty-five of the plastid 50S RP's (PRP's) are encoded in the nuclear genome and synthesized on cytosolic ribosomes while 8 of the PRP's are encoded in the plastid organelle genome and synthesized on plastid ribosomes. Sites for transit peptide cleavages in the cytosolic RP precursors and formyl Met processing in the plastid-synthesized RP's were established. Post-translational modifications were observed in several mature plastid RP's, including multiple forms of L10, L18, L31 and PSRP-5, and N-terminal/internal modifications in L2, L11 and L16. Comparison of the RP's in gradient-purified 70S ribosome with those in the 30S and 50S subunits revealed an additional protein, in approximately stoichiometric amount, specific to the 70S ribosome. It was identified to be plastid ribosome recycling factor (P-RRF). Combining with our recent study of the proteins in plastid 30S subunit (ms. submitted), we show that spinach plastid ribosome comprises 59 proteins (33 in 50S subunit, 25 in 30S subunit and RRF in 70S), of which 53 are *E. coli* orthologues and 6 are plastid-specific proteins (PSRP-1 to -6). We propose the hypothesis that PSRP's were evolved to perform functions unique to plastid translation and its regulation, including protein targeting/translocation to thylakoid membrane via plastid 50S subunit.
INTRODUCTION

The plastid (chloroplast) ribosome is a plant-specific, organelle ribosome that produces proteins encoded by the plastid genome. Plastid ribosomes are responsible for the synthesis of huge amounts of biomass, since the large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (a most abundant protein in the biosphere) is synthesized in plastids. Plastid ribosomes are very similar to the eubacterial 70S-type ribosome, in composition and general mode of function (1-4). The rRNA's and most of the characterized ribosomal proteins (RP's) in plastid ribosomes also bear close resemblance to the corresponding components so far identified in cyanobacteria, a correlation supporting the importance of endosymbiotic theory in plastid evolution (5).

The *E. coli* ribosome, the most well-studied of the eubacterial ribosomes (6), is composed of 21 RP's in the 30S subunit and 33 RP's in the 50S subunit. Two more possible *E. coli* RP's have been suggested: protein Y, the product of *E. coli yfia* gene (7), bearing a distant sequence homology to a chloroplast-specific RP (PSRP-1), and a protein designated S22 (8). Post-translational modifications are found in many *E. coli* RP's, although a modification in L16 (Arg-81) remains yet to be characterized (see Results for plastid L16). We have recently identified all the RP's in spinach plastid 30S ribosomal subunit, including all its PSRP's and many post-translational modifications (Yamaguchi, K., von Knoblauch, K., and Subramanian, A. R., ms submitted for publication). The number of RP's in plastid 50S subunits has so far only been estimated (35~39, reviewed in ref. 2), but has not been determined.

Although the constituents of plastid translational machinery in general are similar to those of *E. coli*, the genes are distributed in two genome compartments: the plastid and the nucleus. The rRNA and tRNA genes are located in the plastid genome, whereas the genes for processing/modification enzymes, aminoacyl tRNA synthetases and 60% of the RP's are located in the nuclear genome (1-4). The plastid translation system also differs from the eubacterial system in other significant ways, e.g. chloroplast mRNA is often edited (9), about 60% of chloroplast mRNAs lack canonical ribosome binding sites (RBS)
found in *E. coli* mRNAs (10), mRNA levels in chloroplasts remain relatively unchanged through dark/light transitions whereas protein synthesis rates increase dramatically upon illumination (11, 12), nuclear-coded factors mediate light-regulated translation (13), and nuclear mutants occur with defects in chloroplast polysome assembly (14). Gene expression in chloroplasts depends overwhelmingly on nuclear gene products which mediate both transcription/post-transcriptional processing (15, 16) and translation. We speculated (ms. submitted to JBC) that some of the nuclear factors exert their roles through evolutionary alterations in plastid ribosome. Since few evolutionary changes are observed in the plastid rRNA, but many in plastid RP (e.g. PSRP's), a key to understanding light-dependent translational regulation might involve chloroplast RP's.

We have applied a proteomic approach (2D-PAGE, HPLC separation, protein sequencing, PCR-based screening/DNA sequencing, and mass spectrometry: LC/MS and ESI MS) to the plastid ribosomal 50S subunit to establish a complete identification of all its protein components. Proteins in both 50S subunits and 70S ribosomes were identified, yielding an unexpected result that plastid ribosomal recycling factor (P-RRF) is present in the approximate stoichiometry of one in the 70S ribosome. Transit peptide cleavage sites in all 25 cytosolically synthesized plastid RP precursors, post-translational modifications in many of the mature PRP's, the absence of the orthologues of two *E. coli* RP's, and possible function of a 70S ribosome-bound P-RRF are discussed.

**EXPERIMENTAL PROCEDURES**

*Preparation of Spinach Chloroplast 70S Ribosome, 50S Subunit, TP50 and TP70 -*  
Spinach (*Spinacia oleracea*, cv. Alwaro) ribosomes were prepared as described previously (17). Purification of 70S ribosomes and the 50S subunits were done by zonal centrifugation (5000 A260 units of ribosome/zonal rotor for 70S purification and 3000 A260 units of purified 70S/ zonal rotor for 50S isolation) as described previously (18). The approximate conversion factor used for the estimation of protein amount was: 1 A260 unit of 70S or 50S = 20 µg protein. TP50 and TP70 (total protein of 50S subunit and 70S
ribosome, respectively) were extracted basically according to Hardy et al. (19). Sample preparations for 2D-PAGE analysis, HPLC resolution (1 mg TP50/run), and LC/MS analysis were as described earlier (30S ms. submitted to JBC).

**Reversed-phase HPLC** - The HPLC system was assembled with a WellChrom Maxi-Star K-1000 HPLC pump (KNAUER, Germany), an injection valve, No.7125 (Rheodyne, USA), and an UV detector, UV-VIS S-3702 (Soma, Japan). Separation of TP50 was performed basically according to Kamp et al. (20) using Vydac C18 column (4.6 x 250 mm) and a gradient with 0.1% trifluoroacetic acid (TFA) and 0.1% TFA in isopropanol at a flow rate of 0.5 ml/min.

**Protein Electrophoresis and Electroblotting** - SDS-PAGE was done by the method of Laemmli et al. (21). Tricine SDS-PAGE was performed according to the method of Schägger and von Jagow (22). 2D-PAGE (basic and acidic) were done as described previously (23, 24). Electroblotting was done according to Walsh et al. (25) using tank-blotting chamber and performed in 25 mM Tris-HCl, pH 8.4, 0.5 mM dithioerythritol, 0.02% SDS at 500 mA for 16 h at 4 °C. PVDF-membrane (Sequi-Blot, BIO-RAD) was used for transfer membrane. After electroblotting, the resulting blot was rinsed three times in water for 5 sec and stained with 0.1% Amido Black 10B (Sigma) in 50% methanol for 5 min, destained with 50% methanol until the background disappeared, rinsed 3 times in water, then dried on a Whatman 3MM filter paper, and stored in the dark at -20 °C until used.

**Protein sequencing** - Protein sequencing was carried out at the Laboratory for Protein Sequencing and Analyze, the University of Arizona, using an Applied Biosystem 477A Protein/Peptide sequencer interfaced with a 120A HPLC analyzer to determine phenylthiohydantoin (PTH) amino acids. After the conversion step, 50 µl of the PTH-derivative (out of 135 µl or 37%) is injected into an ABI PTH-Narrowbore C18 column (2.1 x 250 mm) for detection of PTH-AA (remaining sample going to fraction collector).

**PCR Screening for cDNA’s of PrpL19, PrpL34, Psrp-5 and Psrp-6** - A lambda gt11 spinach cDNA library prepared in our laboratory previously (26) was screened by
thermal gradient PCR using a Mastercycler gradient PCR apparatus (Eppendorf Scientific, Inc., USA). Thermal gradient PCR was performed for 3 min at 94 °C, 35 cycles of 1 min at 94 °C, 1 min at 43-60 °C, 1 min 30 sec at 72 °C with 0.25U of Taq DNA polymerase (Gibco-BRL) in a 20 µl reaction volume containing 1 µl of spinach cDNA library (~10⁸ plaque forming units), 5 µM gene specific primer or 10 µM degenerate primer, 5 µM lambda arm primer (PF or PR), 20 µM each dNTP, 1.5 mM MgCl₂, and 50 mM KCl in 20 mM Tris-HCl (pH 8.4).

Plaque Screening of cDNA Library/Cloning of Spinach PrpL5 and PrpL34 - The lambda gt11 spinach cDNA library (26) was screened using random-primed, [³²P]-labeled Arabidopsis EST clone (E10B7T7) and a 5'-PRPL34 cDNA portion (PF/PL34R1), respectively, as probes for PrpL5 and PrpL34. Clone E10B7T7 was from Arabidopsis Biological Resource Center (Ohio State University). Radio-labeling was carried out as described by the supplier of Random Primed DNA Labeling Kit (Boehringer Mannheim). 150,000 plaque-forming units were plated on four 132 mm plates, and plaques were lifted onto Nylon filters. Prehybridization was performed in 500 mM sodium phosphate, pH 7.0, at 50 °C (for PrpL5) or 65 °C (for PrpL34) for 2 h and hybridization in 500 mM sodium phosphate, pH 7.0, 7% SDS at 50 °C for 4 h (for PrpL5) or at 65 °C for 16 h (for PrpL34). The filters were washed twice in 100 mM sodium phosphate, pH 7.0, 1% SDS at 37 °C for 10 min followed by a 10-min wash in 40 mM sodium phosphate, pH 7.0, 1% SDS at 37 °C and autoradiographed. Plaques giving positive signals were purified two further rounds of screening (27).

Oligonucleotide primers and DNA sequencing - The oligonucleotide primers used in this study were: PF, 5’-CGG-GAT-CCG-GTG-GCG-ACG-ACT-CCT-GGA-GCC-CG-3’; PR, 5’-CGG-GAT-CCC-AAC-TGG-TAA-TGG-TAG-CGA-CCG-GC-3’; PL5F1, 5’-TGG-CAC-TGA-TTA-CTG-GGC-AAA-GGC-3’; PL5R1, 5’-GTG-TTT-TGT-TAC-ACG-GAA-TGC-3’; PL5R2, 5’-TAC-CTT-CTC-TGA-CCT-TAA-ACC-CTG-3’; PL19F1, 5’-AAR-GAR-ATH-AAR-GTI-GTI-GCI-CAY-MG-3’; PL19F2, 5’-CAA-TGA-CTT-GAA-TTT-CCC-TG-3’; PL19F3, 5’-GCC-ATT-GAA-GAA-GCA-ATT-
AG-3'; PL19F4, 5'-GGT-GAT-ATT-CTG-CTG-AAC-ATC-AG-3'; PL19R1, 5'-CTT-
AGA-TAG-TAT-AGC-CTT-GCC-3'; PL19R2, 5'-CTG-ATT-TGC-ACA-ATG-TCT-
CC-3'; PL19R3, 5'-GCT-ATC-CTC-CGC-TTC-CGA-CC-3'; PL19R4, 5'-CTA-ATT-
GCT-TCT-TCA-ATG-GC-3'; PL34F1, 5'-GGI-AAR-GCI-GCI-YTR-IYI-ISI-YII-AAR-MG-3'; PL34F2, 5'-GTC-ATT-GGC-TCG-GAC-ACA-TG-3'; PL34R1, 5'-
GCT-CAT-TCG-CAG-ACG-AAA-ACC-3'; PL34R2, 5'-AGA-GCA-ATG-GAG-
TGA-CCC-GG-3'; PPSRP-5F1, 5'-GGA-ATT-CTA-GAT-ATC-GTC-GAC-GAG-
AGA-TGG-CAC-TCC-TTT-C-3'; PPSRP-5F2, 5'-GAA-GCT-AAC-ATC-TCA-GTT-
CAG-3'; PPSRP-5R1, 5'-GGA-ATT-CGT-CGA-CGC-GTT-TTT-GAG-AAA-AAG-
ATT-TAC-.ACT-G-3'; PPSRP-5R2, 5'-GCC-TGT-TCC-TTC-GGA-GTC-TG-3';
PPSRP-6F1, 5'-GGA-ATT-CTA-GAT-ATC-GTC-GAC-CCT-TCC-AKA-GCA-
AAA-TAG-AAA-AAA-AAG-AGG-3'; PPSRP-6R1, 5'-CAT-RTG-RTG-IGC-IGT-
ICC-YTT-YTT-YTG-3'; PPSRP-6R2, 5'-TCA-TCA-AAC-AGT-TCA-TAT-GC-3';
PTAG1, 5'-GGA-ATT-CTA-GAT-ATC-GTC-GAC-GAG-AGA-3'; PT7, 5'-TAA-TAC-GAC-TCA-CTA-TAG-GG-3'; PT3, 5'-AAT-
TAA-CCC-TCA-CTA-AAG-GG-3'.

PF (forward primer) and PR (reverse primer) are complementary to the cloning site of
lambda gt11. Degenerate oligonucleotide primers, PL19F1, PL34F1, and PPSRP-6R1
were designed from PRPL19 peptide 1 (sequence region: KEIKVVSHR), N-terminal
sequence of PRPL34 (sequence region: GKAALXLTKR), and N-terminal sequence of
PSRP-6 (sequence region: QKKGTAHHM), respectively (see also Table I). Gene-
specific PCR primers for PrpL19 and PrpL34 (PL19R1 and PL34R1) were designed from
the nucleotide sequences of PCR products PL19F1/PR and PL34F1/PR, respectively,
which were amplified using primer sets (PL19F1 and PR) and (PL34F1 and PR). Gene-
specific PCR primer for PsrP-6 (PPSRP-6F1) was designed from the nucleotide sequence
of PF/PPSRP-6R1 and tagged with PTAG1 sequence. Gene-specific PCR primers for
PsrP-5 (PPSRP-5F1 and PPSRP-5R1) was designed from the nucleotide sequence of
spinach chloroplast L40 (28), and tagged with PTAG1 and PTAG2 sequences,
respectively. Other sequencing primers shown above were designed from obtained DNA sequences during primer walking. PCR products were analyzed by agarose gel electrophoresis using 1% agarose gel and visualized by ethidium bromide staining. PrpL5 insert DNA in the phage clone (L5F2-1) was amplified by PCR using primer sets PF and PR, and sequenced using primers PF, PL5F1, PL5R1, and PL5R2. The nucleotide sequence of PrpL19 was obtained by sequencing PCR products PF/PL19R1 and PL19F1/PR using sequencing primers: PL19F2, PL19F3, PL19F4, PL19R2, PL19R3, and PL19R4. For PrpL34, insert DNA in the phage clone (L34D2-1-1) was amplified by PCR using primer sets PF and PR, then cleaved by *EcoRI* digestion and subcloned into the plasmid vector pBluescript SK* (Stratagene). The insert DNA in PrpL34 plasmid clone was sequenced using primers, PT3, PL34F2, PT7, and PL34R2. The nucleotide sequence of PsrP-5 was obtained by sequencing PPSRP-5F1/PPSRP-5R1 using sequencing primers: PTAG1, PPSRP-5F2, PTAG2, and PPSRP-5R2. The nucleotide sequence of PsrP-6 was obtained by sequencing PPSRP-6F1/PR using sequencing primers: PTAG1 and PPSRP-6R2. Nucleotide sequences were determined at the DNA Sequencing Facility, University of Arizona, using an Applied Biosytems model 377 sequencer.

**Mass Spectrometry** - Mass spectrometry was done at the Mass Spectrometry Facility, Chemistry Department, University of Arizona. Electrospray ionization mass spectrometry (ESI-MS) was carried out on a Finnigan LCQ system. Liquid chromatography - mass spectrometry (LC/MS) was done on the same LCQ system interfaced with a Michrom HPLC system (Magic 2002) using a Microbore C18 column (1 x 150 mm). The solvent system was: 0.1% TFA in 2% acetonitrile (solvent #1) and 0.1% TFA in 90% acetonitrile (solvent #2). 50 pmol purified PRP in 10 µl 4% acetic acid was subjected to ESI-MS. For LC/MS, 20 µl TP50 (100 pmol) in solvent A was injected to the Microbore C18 column.

**Internal peptide preparation** - ‘In-gel’ digestion of PRP19 was performed basically according to Hellman *et al.* (29). Five spots were excised from Coomassie Blue stained
2D-gels of TP50 (200 pmol/gel) and equilibrated with mixing for 40 min at 30 °C in 2 ml of 50 mM Tris-HCl, pH 8.5 in 50% acetonitrile. The gel pieces were completely dried in a Speed-Vac, then rehydrated with 150 µl of 50 mM Tris-HCl, pH 8.5, 0.02% polyoxyethylenesorbitan monolaurate (Tween 20), 10% acetonitrile, containing 0.4 µg (ca. 1/40 enzyme/substrate ratio by weight) endoproteinase Lys-C (Sigma), and incubated at 30 °C for 16 h. The enzyme reaction was stopped by adding 1/10 reaction volume of 10% TFA. Gel pieces were transferred into 500 µl of 0.1% TFA in 60% acetonitrile and peptides were extracted by shaking at 30 °C for 80 min. The extract was dried in a Speed-Vac, dissolved in 50 µl of 5% acetic acid, and peptides were purified by reversed-phase HPLC using a Vydac C8 (4.6 x 50 mm) column in TFA-acetonitrile solvent system. The purified peptides (PRPL19 peptides 1 and peptide 2, see Table I) were dried, dissolved in 25 µl of 30% acetic acid and subjected to a protein sequencing.

In order to obtain internal peptides of RRF (the 70S ribosome-specific protein, identified later as plastid ribosome recycling factor), it was purified from pool 37 from a previous work (30). Pool 37 showed two components, identified ~60% RRF and ~40% PSRP-2, by 2D-PAGE analysis. RRF was purified using a Vydac C4 (4.6 x 150 mm) column in TFA-acetonitrile solvent system. The purified protein (ca. 20 µg) was digested using 0.2 µg endoproteinase Asp-N (Sigma) in 50 mM Tris-HCl, pH 8.0, 2 M urea, at 37 °C for 16 h. The digest was dried in a Speed-Vac, then subjected to Tricine SDS-PAGE. Peptides separated on the gel were electroblotted onto a PVDF-membrane and stained as described above. A 3.5 kDa peptide band was excised from the blot and sequenced (RRF peptide 1 shown in Table I).

Small subunit of RuBisCo- Spinach SSU (small subunit of ribulose 1, 5-bisphosphate carboxylase/oxygenase) was present in a pool from previous work, pool 40 (30). It was purified by reversed-phase HPLC using a Vydac C8 (4.6 x 50 mm) column in TFA-acetonitrile solvent system, and also by 2D-PAGE/electroblotting. The N-terminal sequence of the purified protein (12 cycles), Methyl-MKVWPTQNMKRY, confirmed its identity as spinach SSU.
Computer analyses - The program BLAST from the National Center for Biotechnology Information (NCBI) was used for sequence searches. Protein sequence searches were performed using ‘blastp’ program versus nr (non-redundant Database of GenBank CDS translations (PDB, Swiss-Prot, PIR, and PRF). EST searches were done using ‘tblastn’ program versus dbEST (non-redundant Database of GenBank, EMBL, DDBJ EST Divisions). ORF’s (open reading frame) from cDNA sequences were analyzed using the ‘map’ program from GCG software (31). Isoelectric points and sequence masses were calculated using ‘peptidesort’ program (31). Sequence alignments and comparisons were performed using ‘pileup’ and ‘gap’ programs (31). LC/MS data was analyzed using the ThermoQuest Finnigan Xcalibur data system. Masses were deconvoluted from resulted m/z using ‘BIOMASS Deconvolution’ program and charge states were convoluted using ‘BIOMASS Calculation’ program.

RESULTS

Proteins of Chloroplast Ribosomal 50S Subunit - In order to identify all the PRP’s in pure spinach chloroplast 50S ribosomal subunit, spinach chloroplast 70S ribosomes were first purified on a zonal gradient and then run through a second dissociating zonal gradient to obtain 30S and 50S subunits. Efficient dissociation of chloroplast ribosomes required the development of a phosphate-containing dissociation buffer (see Ref. 18 for details and typical gradient profiles). TP50 was extracted from 50S subunits and the proteins were separated by 2D-PAGE and transferred onto PVDF membrane. Fig. 1A shows such a membrane, stained with Amido Black. The individual spots have a slightly different staining pattern with Amido Black as compared to staining with Coomassie Blue, e.g. acidic proteins like L12 are stained poorly by Amido Black (compare Fig. 1A with Fig.8A). Each of the spots was excised and was subjected to N-terminal sequencing. Since several of the protein spots were composite-looking, indicating overlapping separations (Fig. 1A), the experiment was repeated, with the spots excised into 2 or 3 segments for
the N-terminal analysis. The sequence data from all these experiments are summarized in Table I.

Most of the N-terminal sequences were sufficient to allow protein identification from similarity using BLAST search and, in several cases, from the identity to reported spinach PRP sequences in databases. In a few instances, however, the N-terminal data was insufficient to provide a positive identification. In those cases, additional experiments were done to obtain internal peptide sequences and/or to isolate cDNA’s from a spinach cDNA library for nucleotide sequencing. Plastid (chloroplast) ribosomal proteins (PRP’s) are designated in Table I as L1, L2,…to L36, as per their sequence similarity to the corresponding *E. coli* ribosomal proteins. The 2D-PAGE pattern is diagramatically represented in Fig. 1B, depicting the positions of all the identified proteins. The designations alpha, beta, .., PSRP-5, PSRP-6, etc. are discussed later.

To confirm whether all of spinach chloroplast 50S subunit PRP’s appeared in the 2D-PAGE pattern shown in Fig 1, two additional experiments were performed. The 2-D system used for Fig. 1 resolves mainly proteins of pI 4.5 or greater (23), as acidic proteins of lower pI do not migrate into the 1st dimension gel. We therefore ran TP50 in another 2-D system suitable for the resolution of acidic proteins (pI 5.5 or lower, ref. 24). A few spots appeared in this 2D-PAGE, but N-terminal sequencing did not reveal any new protein sequence (data not shown). A second problem arises if an N-blocked protein comigrates with an unblocked PRP. A single spot would then appear, giving a single N-terminal sequence under Edman degradation, suggesting the spot contained one protein. To overcome this problem, we resolved the 50S subunit PRP’s (TP50) on a reversed-phase HPLC column, where hydrophobic interaction (rather than the net charge or peptide chain length) is the key to resolution. HPLC could thus separate proteins that are often unresolved by 2D-PAGE. Fig. 2 shows the separation of plastid TP50 obtained by this procedure. Proteins in the individual peaks were identified by 2D-PAGE analysis. The HPLC experiment allowed clean separations of several proteins that were not well resolved by 2D-PAGE, e.g. L4/L21, L3/L13, L20/PSRP-5 forms (compare Figs 1 and 2).
Indeed, plastid L36 and the beta/gamma forms of PSRP-5 did not at all show up on 2D-PAGE, but gave distinct peaks in the HPLC run, permitting both N-terminal sequence and mass (by LC/MS) determination. Because these three small proteins are very basic, they might have migrated out of the 2D-gel. Interestingly, a few proteins on the other hand, L35, L33, L32, L27, and PSRP-5 (alpha-gamma), were each eluted in two different peaks (Fig. 2). Their identities were inferred from 2D-PAGE analysis, N-terminal sequencing, as well as mass determinations (ESI MS). This observation could suggest that some plastid RP’s may possibly exist as two distinct conformers on the plastid ribosome.

The average yield of PTH-AA recovered in the first three cycles of Edman degradation was calculated for each of the sequence runs and is summarized in Fig. 3. Most proteins showed in yields ranging from 20 to 50 pmol, indicating that this amount approximates the average stoichiometry (1 copy/50S subunit). Plastid L12 gave a yield of almost 200 pmol, four times the highest yield for the other proteins recovered in good yield. The result supports the existence of 4 copies of L12/plastid 50S ribosomal subunit, as has been previously deduced (17). Several of the plastid 50S subunit proteins, e.g. L10, L18, L31, and PSRP-5 exist in multiple forms (named alpha to gamma, Figs. 1 and 2 and Table I). The summed N-terminal yield of the multiple forms in each of these cases corresponded to the approximate stoichiometric amount for the other PRP’s (Fig. 3).

Very low yields of N-terminal amino acids were observed in three instances: L16, L22, and L34. As discussed later, 90% of PRPL16 is likely blocked by trimethylation; most of PRPL22 may be N-blocked by an unknown group; and PRPL34, one of the most basic proteins of the 50S subunit (like L36), partially runs out of the 2D-gel.

The N-terminal sequence data in Table I allowed the identification of the orthologues of 28 *E. coli* RP's in spinach plastid ribosome, beside two plastid specific proteins, PSRP-5 and PSRP-6. *E. coli* ribosomal 50S subunit contains [making 3 subtractions for L7, alpha-N-acetylated form of L12; L8, a complex of L7/L12 and L10, and L26=S20] 33 canonical RP’s: L1 to L36 (32). Hence our data so far has not allowed positive identifications for five possible *E. coli* orthologues. These are: L5, L19, L25, L30 and L34,
of which three (L5, L19, and L34) were identified with additional sequence data obtained (using the protein sequence information) by screening a spinach cDNA library, described earlier (26).

The N-terminal sequence of the protein, later identified as PRPL5 (25 AA), did not show significant homology to any database proteins (BLAST search using ‘blastp’ program), but the same data when used to screen the EST data bank with ‘tblastn’ program gave two matches with Arabidopsis EST clones. One of these, E10B7T7, was obtained from the Arabidopsis Biological Resource Center (Ohio State University) and used as probe to screen our spinach lambda gt11 cDNA library (26). Several positive clones were isolated and the longest cDNA clone (L5F2-1) was sequenced (data deposited in GenBank, accession # AF250923). Clone L5F2-1 encodes 207 AA residues, that is 13 residues short of the complete sequence of mature PRPL5. The 25-residue N-terminal sequence (Table 1), however, provided this missing sequence as well as a 12-residue overlap with the DNA-derived sequence. The sequence comparisons of spinach PRPL5 with its homologues in a higher plant (A. thaliana), an alga (Porphyra purpurea), a photosynthetic bacterium (Synechocystis PCC6803), and E. coli are shown in Fig. 4.

Spinach PRPL5 contains a 16-amino acid long NTE, and a positively charged 26-amino acid long CTE, as compared to E. coli L5. The plastid protein has amino acid identity of 72.5% to the A. thaliana L5, and 56.4%, 60.0%, and 48.6%, respectively to the L5 proteins of P. purpurea, Synechocystis PCC6803, and E. coli. The amino acid sequences of the NTE and CTE of plastid L5 yielded no significant matches when searched against protein databases.

The N-terminal and internal sequences of plastid L19 and the N-terminal sequence of plastid L34 did not show significant similarity to any proteins in databases. We therefore screened the spinach lambda gt11 cDNA library (26) using inosine-containing degenerate oligonucleotide primers designed from PRPL19 peptide 1 and the N-terminal sequence of PRPL34. Thermal gradient PCR allowed us to find the optimal amplification conditions using primer sets of degenerate primers and lambda-arm primers (PF or PR). Amplified
PCR products (PL19F1/PR and PL34F1/PR) were sequenced, and further PCR amplifications were done using sets of PF and gene specific primers (based on the obtained DNA sequence). The nucleotide sequences of PRPL19 cDNA was thus obtained entirely by sequencing of PCR products (submitted to GenBank, accession # AF 250384).

As compared to *E. coli* L19, spinach plastid L19 contains a negatively charged, 47-amino acid long NTE. However, it showed a remarkably low similarity to the corresponding two L19-like sequences in the *Arabidopsis* database (Ath1 and Ath2, Fig. 4), and gave no significant matches against all the sequences in protein databases. Overall, spinach PRPL19 showed amino acid identities of only 50.4% and 47.7% to the two *Arabidopsis* sequences (genes F3I6.17 and AT4g11630), and 38.5%, 42.7%, and 37.3%, respectively, to the L19 protein sequences from *P. purpurea*, *Synechocystis* PCC6803, and *E. coli*.

For PRPL34, a PCR product encoding the 5'-region of its cDNA (PF/PL34R) was first obtained and was used as probe to screen the spinach cDNA library. Several clones were isolated and the one containing the longest cDNA was subcloned into a plasmid vector and was sequenced. The nucleotide sequence is submitted to GenBank (accession # AF238221). As compared to *E. coli* L34, plastid L34 contains a 7-residue long NTE and 10-residue long CTE. These extensions are absent in both cyanobacterial and algal L34 proteins. Core sequences between PRPL34 and the L34 proteins from *P. purpurea*, *Synechocystis* PCC6803, and *E. coli* are relatively well conserved (Fig. 4), the percent identities being, 45.5%, 41.9% and 50.0%, respectively.

Including the three proteins identified via DNA work, a total of 31 orthologues of *E. coli* RP's are present in spinach plastid ribosome. There are thus two *E. coli* RP's (L25 and L30, see Discussion), for which we could obtain no evidence of occurrence in spinach plastid ribosome.

*Plastid-specific Ribosomal Proteins in Chloroplast 50S Subunit* - One of the additional N-terminal sequences in Table I (PSRP-5) corresponded to the reported sequence of
spinach plastid L40 (28), which is the homologue of an earlier reported protein from pea, named PsCL18 (33). Another N-terminal sequence we determined (PSRP-6, Table I) showed similarity to a second reported protein from pea, PsCL25 (33); spinach homologue of PsCL25 has not been reported. We have screened our spinach cDNA library (26) for the cDNA’s corresponding to these two proteins, and the clones obtained were sequenced. The proteins were designated PSRP-5 (previous L40) and PSRP-6 (homologue of pea PsCL25), in accordance with our proposed nomenclature (see Discussion; Nucleotide sequences of PSRP-5 and PSRP-6 cDNA’s deposited in GenBank, accession # AF261940 and AF245292, respectively). Our PSRP-5 data showed 6 amino acids differences, all at the C-terminal portion, from the reported L40 data (28), leaving it uncertain at this point whether the differences reflect spinach cultivar differences (cv. Alwaro versus G'leant d'hiver) or sequencing errors. Three post-translationally modified forms of PSRP-5 were identified from our protein sequencing results (Table I), the modifications were in the N-terminal portion of the protein (see Discussion). Both PSRP-5 and PSRP-6 are unique to the plastid ribosome, homologous sequences being not found in the RP's of E. coli, archaebacteria, yeast (cytosolic or mitochondrial) or mammals.

Mass Spectrometry of 50S PRP’s - Spinach TP50 was analyzed by LC/MS. Fig. 5A shows the relative abundance of summed mass/charge ratio (m/z) of the PRP’s in the m/z range 400 to 2,000 versus HPLC elution time. An example of a mass spectrum at the elution interval of 41.5-42.0 min (summed scans for 30 sec) is shown in Fig. 5B. Protein mass was calculated by deconvolution of the m/z series. For example, from deconvolution of the m/z series in Fig. 5B, a major mass of 13,811.0 Da and a minor mass 13850.2 Da were derived (Fig. 5C). As the observed mass 13,811.0 Da is very close to the sequence mass of PRPL12, i.e. 13,814.54 Da, this peak was identified as of plastid L12. In cases where sequence mass were not available, ESI MS of HPLC pools, containing proteins identified by 2D-PAGE (shown in Fig. 2), allowed mass identification. Every 30 sec interval was analyzed as stated above, and the resultant mass values are summarized in Table II. The combined LC/MS analysis of TP50 and ESI MS analysis of HPLC pools...
allowed the identification of most individual 50S PRP masses, except for PRP's L2, L20, the multiple forms of L10, L18, L31, and PSRP-6. An extreme case was PRPL2, which was not detected either by LC/MS of TP50 or ESI MS of HPLC pools; it probably represents a very poorly ionizing polypeptide.

Edman degradation of PRPL16 - The plastid located prpL16 gene codes for the N-terminal sequence formyl Met-Leu-Ser- (34), while our N-terminal sequence analysis of PRPL16 yielded Xaa-Leu-Ser- (Xaa being a modified, unidentified AA). The N-terminus of *E. coli* L16 is α-N-monomethylated (35), and so we suspected the same modification in plastid L16. To test this hypothesis, we purified spinach SSU (small subunit of ribulose 1, 5-bisphosphate carboxylase/oxygenase), which has monomethylated methionine at the N-terminus (36). About 200 pmol each of SSU HPLC (SSU purified by HPLC), SSU 2D/blot (SSU purified using 2D-PAGE/electroblotting) and PRPL16 2D/blot were subjected to N-terminal analysis (Fig. 6). The SSU HPLC yielded PTH-monomethyl methionine (NMM, retention time 26.37 min, appearing between PTH-tryptophan and PTH-phenylalanine), whereas SSU 2D/blot showed an unusual PTH-derivative of retention time 10.30 min, just after the N'N'-dimethyl-N'-phenylthiourea peak (DPTU, marked a; retention time 10.25 min). Alkylation of monomethyl methionine with acrylamide under mildly alkaline conditions (during 2D-PAGE and/or electroblotting), just as the similar modification of cysteine during the same procedure (37), could produce the latter (uncharacterized) compound. PRPL16 2D/blot showed the same unusual PTH-derivative with the same retention time in cycle 1. In cycle 2, PRPL16 2D/blot gave the expected residue, PTH-Leu, but the yield (2.7 pmol) was only about one-tenth the expected amount (the average yield of N-terminal PTH-AA of 50S PRP’s was 27.5 pmol, see Fig. 3), whereas the two SSU preparations gave the normal yield of PTH-Lys (see Fig. 6). We therefore infer that post-translationally about 10% of PRPL16 is alpha-N-monomethylated while 90% of it is blocked to Edman degradation, probably by trimethylation (see Discussion).

Identification of a 70S Ribosome-specific Protein - During the course of this work we observed that a prominent protein spot, always present in TP70 gels, was absent in either
TP30 or TP50 gels (Fig. 7). To investigate further, a TP70 blot was prepared, and this spot and several other spots (from both 30S and 50S proteins) were excised and subjected to N-terminal sequence analysis. The analysis yielded a single N-terminal sequence, specific to a protein found only on plastid 70S ribosome, the spot indicated by arrow in Fig. 7C. The N-terminal sequence of 21 residues (Table I), however, did not show homology to any of the reported proteins in database. Therefore, the 70S ribosome-specific protein was purified from a previously fractionated spinach chloroplast ribosomal protein pool, #37 (ref. 31), and an internal peptide was obtained by endoproteinase Asp-N digestion. At the time we sequenced the internal peptide, the nucleotide sequence of spinach chloroplast ribosome recycling factor (RRF) was reported (38). Our N-terminal and internal sequences matched 100% with the corresponding sequences of RRF (N-terminal positions 1 - 21 and internal positions 116 - 139). Thus, the 70S ribosome-specific protein we identified in this work is plastid RRF. In contrast to the plastid situation, *E. coli* RRF is mostly found in post-ribosomal supernatant (see Discussion).

We have previously identified all the proteins in spinach chloroplast 30S subunit, a total of 25 proteins - 21 *E. coli* orthologues and 4 PSRP's (Yamaguchi, von Knoblauch, and Subramanian, submitted). Here we have identified all the proteins in spinach chloroplast 50S subunit, a total of 33 proteins - 31 *E. coli* orthologues and 2 PSRP's. In addition, a protein found only on plastid 70S ribosome was characterized and this 70S-specific protein was identified as plastid RRF. In order to round off the picture, a 2D-PAGE of TP70 was done (Fig. 8A) and all the spots were cut out and analyzed. The protein identifications confirmed the results from the 30S and 50S subunit experiments and, as diagramatically represented in Fig. 8B, revealed no further proteins. It should be mentioned that PRPL36 and two of the forms of PSRP-5 (beta and gamma) did not show up in the 2D-PAGE patterns, but were identified with the help of HPLC (Fig. 2). As we noted in the 30S identification paper, the diffusely staining minor spots visible in the upper (high molecular) part of the 2D-PAGE (Figs 1A and 8A) represent mainly minor
aggregates of a few PRP's exhibiting tendency to polymerize (see legend to Fig. 1), and small amounts of nonribosomal proteins, that probably have functional associations with ribosomes.

DISCUSSION

In this paper we present sufficient protein and nucleic acid sequence data to establish the identification of all the protein components in the 50S subunit of a land plant chloroplast ribosome. Similar results for chloroplast 30S subunit has just been submitted for publication. Taken together these two publications present the first complete identification of the protein components of an organelle (plastid) ribosome, complementing the previously reported such complete identifications for *E. coli* (6), and the cytosolic ribosomes of yeast (39) and mammals (40). Recent reports indicate that a complete protein identification for a mitochondrial ribosome may soon be forthcoming (41-43).

In the companion 30S subunit paper we presented a nomenclature for plastid RP's and their genes that accords with both the current usage for mitochondrial RP's and the CPGN (Commission on Plant Gene Nomenclature) rules for plant gene names. In brief, chloroplast RP's are designated PRP's (Plastid Ribosomal Proteins), with gene names written in italics, but having the first letter in capital for the RP genes that are located in the nuclear genome and the first letter lowercase for the RP genes that are located in the plastid genome. Thus the plastid homologue of *E. coli* L1 is designated PRPL1 and its nucleus-located gene designated, *PrpL1*, whereas the plastid homologue of *E. coli* L2 is designated PRPL2, and its plastid-located gene is designated *prpL2*. As described in the companion paper, chloroplast ribosomes contain proteins that do not have homologues in *E. coli*, archaebacteria, or in yeast/mammalian cytosolic ribosomes. These proteins are designated PSRP-1 etc (Plastid Specific RP), and their genes, which are all located in the nuclear genome, are designated *Psrp-1* etc.
Proteins of Plastid Ribosomal 50S Subunit - Spinach chloroplast 50S ribosomal subunit contains 33 proteins, the same number as in *E. coli* 50S ribosomal subunit. However, only 31 of these proteins are orthologues of corresponding *E. coli* RP's, while the remaining two are plastid specific ribosomal proteins (PSRP's: Fig. 1B, Fig. 3 and Table I). The 31 orthologues of *E. coli* 50S RP's are designated (PRP): L1 - L6, L9 -L24, L27 - L29, L31-L36 (*E. coli* designations, L7, L8 and L26 do not stand for distinct RP's). Thus two *E. coli* 50S RP's, L25 and L30, do not have orthologues (or homologues) in the plastid 50S subunit. Since plastid 30S subunit contains the orthologues of all 21 *E. coli* 30S RP's (ms. submitted for publication), spinach plastid ribosome maintains all *E. coli* RP's except L25 and L30.

The orthologues of the two PSRP's we identified in spinach plastid 50S subunit were first reported in pea, derived from cDNA sequences named PsCL18 and PsCL25 (33). Based on the present comprehensive protein study, we can state that no additional PSRP's are present in spinach chloroplast 50S subunit (Figs. 1-3, Table I). The companion comprehensive study on plastid 30S subunit (ms. submitted), has revealed 4 PSRP's, (designated PSRP-1 to PSRP-4) in the 30S subunit. The plastid ribosome thus maintains 6 plastid-specific proteins, 4 of them in the 30S subunit and 2 in the 50S subunit. The two PSRP's in the 50S subunit, are designated PSRP-5 and PSRP-6.

Sequence similarity between a barley plastid ribosomal protein (BPRL28) and PSRP-6 has been recently noted (44). Interestingly, despite being specific to plastids, the sequence identities among PSRP-5 and PSRP-6 homologues are relatively low, e.g. only 53.8% and 54.0%, respectively, between the spinach and pea proteins (Fig. 4). There was no significant sequence similarity between PSRP-5 or PSRP-6 and any of the bacterial proteins in databases, suggesting their appearance in plants after the endosymbiotic event in plastid evolution. In contrast, all four 30S subunit PSRP's show some sequence homology to eubacterial proteins: PSRP-1, PSRP-2, and PSRP-3 to cyanobacterial proteins (ms submitted to JBC) and PSRP-4 to a protein only reported from *Thermus thermophilus* (ms in preparation), suggesting their evolution from ancestral eubacterial
genes. The data thus indicate a strong diverging trend for the 50S PSRP sequences in the plastid evolution.

Post-translational Processing: Plastid 50S RP’s Encoded in the Organelle Genome -
The genes that encode 8 of the spinach plastid 50S subunit PRP's are maintained in the plastid genome, and thus they are synthesized on the plastid ribosome with initiating formyl-Met tRNA. The post-translational processing of the N-terminal formyl-Met undergone by these PRP's are shown in Table III. Five of the proteins (L2, L20, L22, L32 and L33) have the entire formyl-Met group excised, while three (L14, L16 and L36) have only the formyl group removed, leaving methionine at the N-terminus. The N-terminal alanine of L2 is monomethylated, as reported earlier (ref. 30, 1st description of post-translational modification in a plastid RP). The modifications in plastid L16 are discussed later (see below). We had previously suspected N-terminal modification in plastid L36 (45), but the present study did not reveal any modification. The MS mass results for L14, L32, and L33 accorded with the calculated sequence molar masses (Table II), indicating no additional post-translational modifications in these PRP's.

Post-translational Processing: Plastid 50S RP’s Encoded in the Nuclear Genome -
Twenty-five 50S subunit PRP’s - orthologues of 23 E. coli RP's and the two PSRP's - are encoded in the nuclear genome and are thus synthesized on the cytosolic ribosomes as precursors; all plastid RP precursor molecules contain an extra, routing peptide sequence (transit peptide), which is cleaved off upon entry of the mature PRP into the plastid. Table III lists the immediate post-cleavage N-terminal sequences for all of these 25 PRP's. Their pre-cleavage peptide sequences are also listed in Table III, where they are known for 13 of the PRP's. A consensus, flanking peptide sequence that could specify the cleavage site for the transit peptidase cleavage enzyme was not discernible in this data.

In contrast to the situation for plastid 30S subunit, only a minority of 50S subunit RP genes located in the plastid DNA (8 out of 33; 24%), the majority being in the nuclear genome. For the 30S subunit, its RP genes are about equally distributed between the plastid genome and the nuclear genome. This may reflect a co-evolutionary linkage, as
pointed out in a previous review (3), arising from the main function of the 30S subunit, i.e. formation of 30S initiation complex, involving plastid-transcribed mRNA. A correlation between the PRP’s that are synthesized on the plastid ribosome and the early steps of ribosome assembly (plastid rRNA is transcribed in the plastid) has been suggested, but it is applicable for the 30S subunit and not for the 50S subunit.

Post-Translational Modifications in mature 50S PRP’s - The post-translational modifications of individual plastid 50S PRP’s, as derived from this study, are discussed below.

PRPL1 showed a minor fragment (spot ‘e’ in Fig. 1) with N-terminal truncation of 23-amino acids, i.e. starting with the N-terminal sequence, TLPSPTKPKKGKAAL, position nos. 24-38 of PRPL1 sequence (GenBank accession # X76932; ref. 46; M. Kavousi and A. R. Subramanian, unpublished results). Since this fragment was identified as present only in TP50 preparations but not in TP70 (compare Figs.1 and 8), its significance is at present unclear.

Spinach PRPL2 gene sequence has been reported [accession # X00797, ref. 47], but the C-terminal portion of the deduced amino acid sequence lacks homology to bacterial and reported chloroplast L2 sequences in databases. We have therefore re-sequenced the \( rpl2 \) gene from spinach chloroplast DNA, and our data (submitted to GenBank, accession # AJ244023) confirms a frame shift arising from a minor sequencing mistake in the earlier submission. The corrected spinach PRPL2 amino acid sequence shows a conserved C-terminal portion with the expected homology. Mature spinach plastid L2 is post-translationally modified, with \( \alpha-N \)-monomethyl alnine at the N-terminus (30).

In previous work using a different cultivar of spinach, the spot of PRPL4 had appeared in a different, slightly more basic 2D-PAGE position (Fig. 1 in ref. 2), while all the other 50S and 30S PRP’s showed the same positions as in this study. Thus PRPL4 may be one of the few plastid RP’s that harbor detectable strain-specific differences. A mass difference of 58-71 is seen between the sequence mass of PRPL4 (GenBank accession # X93160, cv. Melody; ref. 48) and the observed MS mass in this study (cv.
Alwaro, Table II). Possibly, it may again reflect a cultivar difference, or post-translational modification.

PRPL10 exists in three forms of differing size (designated alpha, beta and gamma). The relative amounts were: alpha form, 80%, beta, 12% and the minor gamma form, about 8%, as estimated from the yields of N-terminal PTH-AA’s. The major form, PRPL10 alpha, mass 20,305 Da, was identified by the LC/MS analysis (Table II), but the other two forms (both approximately 16.5 kDa, as estimated by SDS-PAGE) were not detected in the MS analysis, probably due to low ionization and smaller amounts. As all the 3 forms gave the same N-terminal sequence, their differences arise from post-translational modifications in the internal and/or C-terminal regions of the molecule.

PRPL11 has been previously shown to be epsilon-N-trimethylated at positions Lys-9 and Lys-45 (ref. 49; J. Schmidt and A. R. Subramanian, unpublished). The E. coli L11 protein is trimethylated at the corresponding positions in its sequence context (50). LC/MS data showed that plastid L11 has a mass increment of 81.5 Da over its sequence mass (Table II). This increment is close to the 84.2 Da mass for two trimethyl modifications. In the case of E. coli L11, the N-terminal amino acid Ala is alpha-N-trimethylated (50). The N-terminal Ala of PRPL11 is mainly unmodified, but there is indirect evidence for a partial modification of this residue, resulting in the presence of a minor form (this paper and ref. 49).

PRPL12 was obtained in a significantly higher yield (Fig. 3), consistent with the presence of 4 copies/50S subunit, as has been previously deduced for plastid 70S ribosome (17). In E. coli, this protein exists in two forms: L12, with free N-terminus and L7, alpha-N-acetylated form, the sum of the two forms constituting 4 copies/50S ribosomal subunit (51), and the ratio of the two forms altering during the bacterial growth cycle (52). The N-terminus of plastid L12 is essentially unmodified (Fig. 3), and thus the roles played by N-acetylation (and its variation with growth) in E. coli are apparently abolished in plastid metabolism. Interestingly, a minor mass of 13,850.2 Da was observed in the LC/MS of TP50, at the same elution interval as PRPL12 (mass of 13,811 Da), as
shown in Fig. 5C. The mass increment of 39.2 Da might indicate the presence of a minor modification (including a minor acetylated form, +42.04 Da).

PRPL16 is partially $\alpha$-N-monomethylated (10%), while the bulk of it is -N-blocked by an unknown modification (surmised as trimethylation, see Results). In E. coli the L16 protein is $\alpha$-N-monomethylated (35) and its Arg-81 is modified by an as yet uncharacterized group (35). The observed mass of E. coli L16 by a recent MALDI-TOF MS analysis (53) is 44.9 Da heavier than the sequence mass. Subtracting the mass of $\alpha$-N-monomethylation (14 Da) from 44.9 Da, the uncharacterized group in Arg-81 of E. coli L16 would be of mass 30.9 Da. Our LC/MS analysis of 50S PRP’s (Table II) indicated that plastid L16 is modified with a mass increment of 74.1 Da. Plastid L16 has a conserved Arg residue (Arg-82), in a relatively conserved sequence context, that would correspond to E. coli Arg-81. It is conceivable that Arg-82 of PRPL16 maintains the same modification as in E. coli Arg-81 (+30.9). Hypothetically, if most of plastid L16 N-terminus is blocked by trimethylation (+42.08 Da), the total mass increment would be 72.98 (30.9+42.08), close to the observed 74.1 Da increment in the mass of PRPL16 (Table II).

PRPL18 exists in two forms (alpha and beta) of different size but with the same N-terminal sequence; alpha form, 35% and beta form, 65%, as estimated from the yield of PTH-AA. The modification in PRPL18 was not characterized in this study.

PRPL19 is a highly diverged protein with a negatively charged, 46 amino acid residues long NTE compared to E. coli L19 (Fig. 4). The presence of a phosphorylated protein in plastid ribosomal 50S subunit has been reported (54, 55): the protein spot in the reports corresponds to PRPL19 (see Figs. 1 and 8), as correlated in ref. 2. As the observed protein mass of PRPL19 is very close to the sequence mass (Table II), MS data would suggest no post-translational modification. However, phosphorylation is a reversible process and the plastid ribosomes used in the present study could be de-phophorylated. The sequence -RRLSSLRASTSKS- in the C-terminal portion of Xenopus 40S subunit ribosomal protein S6 (56), is phosphorylated (S in bold). The consensus recognition motif
for S6 kinase II is reported to be -RXXS- (57), and for casein kinase II is -S/TXXE/D-
(58). The motif -RXXS- is present in both spinach (83-RRLS-86) and Arabidopsis
(RRVS) plastid L19 sequences, whereas (1-SEAE-4 and 26-SEAE-30) are present only
in spinach L19. Thus, this study cannot rule out (or rule in) plastid L19 phosphorylation.

Bubunenko et al. has reported that plastid L23 protein in the chloroplast ribosomes of
a certain group of plants (Caryophyllidae, spinach and relatives) is replaced by a
homologue of the cytosolic L23 protein, an unusual evolutionary event (59). Table II
shows the observed mass of plastid L23 by ESI MS analysis (13,553.5 Da), a value close
to the sequence mass (13,553.7 Da), calculated from the nucleotide of sequence of plastid
L23 cDNA (GenBank accession # X90414; C. Jayabaskaran and A. R. Subramaniam,
unpublished). The mature plastid L23 thus appears post-translationally unmodified. The
N-terminal sequence of PRPL23 shown in Table I corresponds to the reported cDNA; it
does not correspond to the deduced N-terminal sequence of a prokaryotic-type rpL23
gene occurring in spinach plastid DNA. Moreover, an N-terminal sequence corresponding
to the plastid rpL23 gene was not observed in this comprehensive study. These results
thus confirm the previous conclusion (59) that the plastid-encoded rpL23 gene in spinach
is a pseudogene. cDNA's derived from two distinct but closely related genes for spinach
cytoplasmic L23, the presumed progenitor of the nuclear gene for spinach plastid L23, have
been isolated and sequenced (C. Jayabaskaran and A. R. Subramaniam, unpublished,
GenBank accession # X92367 and X92350).

PRPL31 exists in three forms (alpha to gamma) with the same N-terminal sequence but
differing in charge: alpha form 55%, beta form 31%, and gamma 14% by estimation from
PTH-AA yields. E. coli L31 may apparently exist in two forms, a truncated form missing
the C-terminal sequence, RFNIPGSK, and the full length form as deduced from the rpL31
(rpmE) gene sequence (60, 61). The modifications in plastid L31 forms are
uncharacterized.

PRPL34 (61 AA) is the smallest and the most basic (pI=12.99) of the 25 nucleus-
coded RP's of spinach plastid 50S subunit. Its transit peptide is 91 AA long (Fig. 4), and
thus the longest of all spinach chloroplast RP transit peptides. PRPL35 (73AA, pI=12.15) also has a relatively long transit peptide of 86 AA (62). An average size plastid ribosomal protein, PRPL19 (156 AA, pI=10.51) has a 77 AA long transit peptide. The largest plastid ribosomal protein PRPS1 (370 AA, pI=4.83) has only a 41 AA long transit peptide (63). In general, a correlation can be made that shorter the mature protein, the longer the transit peptide in the RP precursor. Longer transit peptides might thus be a requirement for the proper import and processing of small basic ribosomal proteins.

Among the two plastid-specific proteins of the 50S subunit, PSRP-5 exist in 3 forms (alpha to gamma), differing apparently in the cleavage points in the precursor form. The reported spinach L40 sequence (28), corresponds to PSRP-5 alpha form in protein sequence length. The cDNA-derived protein sequences of L40 and PSRP-5 differ in 6 amino acids at the C-terminal portion (Fig. 4). These differences could be due to cultivar difference (Alwaro versus Géant d’hiver) or possibly sequencing errors. The protein mass of all the 3 forms of PSRP-5 were observed upon LC/MS analysis of TP50 (Table II), suggesting that all the forms exist on plastid 50S subunit. Mass data also indicate that two of the forms (beta and gamma) are unlikely to be post-translationally modified. In contrast, Edman degradation of PSRP-5 alpha gave only 6% of the expected average yield of PTH-AA, as compared to the yields of 28% and 30%, respectively, for PSRP-5 beta and PSRP-5 gamma (Fig. 3). We surmise that the major portion of PSRP-5 alpha is blocked at the N-terminus. As the LC/MS mass data for PSRP-5 alpha showed 41 Da increment over the calculated sequence mass (Table II), the blocking group could be acetyl (42.04 Da). The relative occurrence of the 3 forms of PSRP-5 appear to be alpha, 42%, beta, 28%, and gamma, 30%. The second plastid specific protein on the 50S subunit, i.e. PSRP-6, appears N-terminally unmodified (Table I), but since it was not observed in the MS analysis, we have no data on modifications elsewhere in the molecule.

The case of ribosomal proteins L25 and L30 - Plastid 50S subunit is missing the orthologues of E. coli L25 and L30 (the PSRP-5 and PSRP-6 proteins have no sequence similarity to L25/L30). In E. coli, L25, L18 and L5 are 5S rRNA binding proteins, the
binding of L18 stimulating the specific binding of L5 (64) and the association of 5S rRNA with 23S rRNA requiring all three proteins (65). Is the function of \textit{E. coli} L25 performed by another plastid 50S protein? It has been reported that there are striking similarities in the properties of \textit{E. coli} L25 and spinach plastid L22, in protecting a domain of \textit{E. coli} 5S rRNA comprising nucleotides 70-109 (66). Spinach plastid L22 consists of a long NTE and a CTE, and a central core homologous to \textit{E. coli} L22 (34); but apparently it is this core of plastid L22 that binds to 5S rRNA (67), even though it is known that \textit{E. coli} L22 does not bind 5S rRNA. A homologue of \textit{E. coli} L25 has been identified in \textit{Anabaena} (photosynthetic cyanobacterium) RP's (68) and evidence for the occurrence of L25 in \textit{Synechocystis} PCC 6803 genome sequence has been presented (68). Thus, it appears that L25 protein might have been lost during chloroplast evolution, with plastid L22 protein taking over some of its functions.

\textit{E. coli} L30 is a 23S rRNA binding, late assembly protein, that can be mutated out without loss of cell viability (69). Binding of L30 protein to nucleotides 931-938 of 23S rRNA has been established by cross-linking (70). Interestingly, the 3 nucleotides, 931-933, on the \textit{E. coli} 23S rRNA are replaced by a variable loop of 5 - 20 nucleotides in chloroplast 23S rRNA (71, 72; see Fig. 10). Protein sequences showing some similarity to \textit{E. coli} L30 could be found in several plant ESTs, but the deduced sequences lacked chloroplast transit peptide sequence. Those EST's thus probably represent cytosolic RP's, mitochondrial RP's which do not always require a separate transit peptide sequence (some MRP's are imported into mitochondria without transit peptide), or nonribosomal proteins.

Sequences having significant homology to \textit{E. coli} L25 and L30 are not always present in the many reported eubacterial genomes. Thus, while both L25 and L30 genes are identified in \textit{Haemophilus influenzae}, \textit{Chlamydia trachomatis}, \textit{Rickettsia prowazekii}, \textit{Neisseria meningitidis} strains MC58 and Z2491, the L25 gene is reported missing in \textit{Aquifex aeolicus}, \textit{Bacillus subtilis}, \textit{Borrelia burgdorferi}, and \textit{Thermotoga maritima}, and the L30 gene is missing in \textit{Synechocystis} PCC6803. Both L25 and L30 genes are missing in
*Helicobacter pylori* strains 26695 and J99, *Mycoplasma pneumoniae, Mycobacterium tuberculosis, Mycoplasma genitalium,* and *Treponema pallidum.* L30 gene homologues are present in the reported complete genome sequences from archaeabacteria, but L25 gene orthologues are absent. Thus, ribosomal proteins L25 and L30 (and their genes) appear to be evolutionary mavericks.

**Comparison of the 50S ribosomal subunits of chloroplast and *E. coli*** - Almost all of the plastid 50S RP's are larger than their *E. coli* counterparts (Fig. 9, histogram). The mass increases are essentially due to NTE's and/or CTE's added to the *E. coli* homologous core portions of plastid ribosomal proteins. Significant mass increases are present in PRP's L1, L4, L5, L13, L15, L19, L21, L22, L24, L27, L29, and L31. Interestingly, all these proteins except L22 are encoded in the nucleus. The summed mass of plastid specific NTE's, CTE's and the two PSRP's equal to 92.5 kDa. The protein mass of chloroplast 50s subunit is 529.6 kDa (as compared to the 437.1 kDa protein mass in *E. coli* 50S subunit); thus the increase corresponds to 21.2%. For chloroplast 30S subunit also there is a similar increase in protein mass: addition of 81.5 kDa, corresponding to a 19% increase over *E. coli* 30S protein mass (ms. submitted).

The rRNA of land plant plastid 50S subunit consists of 23S, 4.5S, and 5S rRNAs. The 4.5 S rRNA essentially represents the corresponding 3' end portion of bacterial 23S rRNA (73). The sum of tobacco 23S rRNA (2,804 nt) and 4.5 S rRNA (103 nt) is just 3 nucleotides larger than *E. coli* 23S rRNA (2,904 nt). The 5S rRNA is a conserved molecule (*E. coli* and tobacco 5S rRNA's are 120 and 121 nt, respectively), and spinach chloroplast 5S rRNA was shown to incorporate in vitro with RP's and 23S rRNA from *Bacillus stearothermophilus* to form functionally active 50S subunits (74).

Interestingly, there are several small but distinctly variable regions of nucleotide sequence, identifiable in both tobacco (a dicot plant) and maize (a monocot plant) 23S rRNA molecules (72). One is the variable sequence discussed earlier regarding L30 binding site. The variable regions are mostly localized in domains I, II, II and VI in 23S rRNA structure (Fig. 10). Interestingly, although these variable regions seem to be randomly
distributed in the secondary structure of 23S rRNA, they have a unique localization in the 3-D arrangement of the rRNA in 50S subunit (75). As seen in Fig. 10, except for the added loop at helix 38.2, most of the other changes are clustered around the bottom of the 50S subunit structure, near the exit for the nascent polypeptide.

The functions of the large ribosomal subunit include peptide bond formation at the peptidyl transferase center and co-translational protein targeting for membrane and lumenal proteins via interactions with signal recognition particle (SRP) and its receptor. SRP's in *E. coli* and eukaryotic cytosol are ribonucleoprotein complexes that co-translationally target proteins (to the ER and bacterial inner membrane, respectively). Plants have evolved an additional specific membrane in chloroplasts (i.e. the thylakoid where photosynthesis occurs), several components of which are synthesized on the plastid ribosome. Recently a chloroplast SRP has been reported as a novel SRP (cpSRP), as it lacks the RNA moiety found in bacterial and eukaryotic SRP's (76). It is conceivable that the plastid 23S rRNA variations (Fig. 10) and the two PSRP's in the 50S subunit have been evolved, in combination with an RNA-less cpSRP, for protein targeting/translocation functions at the 50S subunit-thylakoid membrane interface.

*Plastid Ribosome Recycling Factor (RRF)* -We identified plastid ribosome recycling factor (P-RRF) as a protein strongly associated with the plastid 70S ribosome (Fig. 8). Its stoichiometry was approximately one, similar to that of most PRP's in the zonal sucrose gradient-purified 70S ribosome preparations (Table I). The RRF protein was absent in either 30S or 50S subunits obtained from the 70S ribosome preparation (Fig. 7). In contrast, the bulk of the RRF content in *E. coli* is present in post-ribosomal supernatant and only a small amount, 0.08-0.2 mol/ribosome, is present in the ribosomal pellet; ribosomal pellet from midlog cells contained the lower amount and that from stationary cells contained the higher amount; the amount in gradient-purified ribosomes was negligible [ experimental results with radio-labeled *E. coli*, A. R. Subramanian (1979), unpublished]. Thus, there is a remarkable difference in the apparent ribosomal affinity of RRF, between the plastid and *E. coli* translation systems.
In *E. coli*, RRF is suggested to catalyze the fourth step of protein synthesis, i.e. the disassembly of the post-termination complex of ribosome, mRNA and deacylated tRNA (77). In a more recent report, the dissociation of 50S subunit from the 70S post-termination complex was proposed to be the step that is catalyzed by RRF, requiring EF-G dependent GTP hydrolysis (78). The finding that plastid RRF is associated, in approximately stoichiometric amount, with 70S ribosomes is apparently inconsistent with the catalytic role suggested for *E. coli* RRF. In an experiment where plastid 70S ribosomes were treated with 500 mM ammonium chloride in 50% ethanol (at 0 °C for 10 min), half the RRF amount was still found on the ribosome, whereas most of plastid L12 was released (data not shown). This observation further supports a strong binding between RRF and 70S ribosomes in plastids.

In *E. coli*, IF-3 is reported to be a ribosome dissociation factor, that dissociates run-off 70S ribosomes to 30S and 50S subunits, the free 30S subunit with bound IF-3 initiating a new round of translation (79, 80). In chloroplasts, the IF-3 activity of the plastid IF-3 molecule is 100-fold affected by the presence of its plastid-specific N/C terminal extensions: it has been suggested that plastid IF-3 is capable of being activated (through its NTE/CTE) by a nuclear regulatory factor under appropriate environmental signals, e.g. light (81). We speculate that in chloroplasts the plastid RRF might have a role in holding the 70S ribosome together (as an inactive form), prior to the activation of IF-3 by a light-dependent regulatory factor and the start of a new round of translation. Plastid RRF might thus function as a true ribosome anti-dissociation factor.

In conclusion, among the 8 spinach plastid 50S subunit RP genes located in the organelle genome, only *rpL16* gene contains an intron (*rpL2* gene is intron-containing in most plants, except spinach and relatives). To obtain a preliminary idea of the intron-exon structure in the nuclear genes that encode for plastid 50S RP’s, we cloned four PRP genes in our laboratory. One of them; *PrpL12* occurs as a cluster of 3 intron-less genes in *A. thaliana* (ref. 82), whereas the other three, *PrpL1, PrpL13*, and *PrpL35*, are present as
single copy genes, containing 6, 3, and 2 introns, respectively (ref. 46; M. Kavousi and A. R. Subramanian, unpublished). The complete picture of plastid ribosomal protein genes from a plant should soon be available (e.g. with the completion of the *Arabidopsis* genome sequencing project), with our protein work facilitating the PRP gene identifications.

Acknowledgment

We would like to thank Drs. K. Giese, B. R. Srinivasa, U. Markmann, Y. Ogihara, S. H. Phua, P.M. Smooker, C. H. Johnson, M. Bubunenko, C. Jayabaskaran and M. Kavousi (post-doctoral fellows/visiting scientists in A. R. Subramanian's Max-Planck lab, Berlin) for their early contributions, Dr. W. Schröder (Free University, Berlin) for his work on PRPL11 methylation, and Mr. K. von Knoblauch for devoted technical/computing assistance. Max-Planck-Gesellschaft, Germany, supported this research at the University of Arizona through a Sponsored Project Grant (protein synthesis and regulation).

REFERENCES


FIG. 1. **2D-PAGE pattern of spinach plastid 50S ribosomal subunit proteins and protein spot identification.**

* A. Amido Black stained electroblot (PVDF membrane) of TP50 (200 pmol) separated by 2D-PAGE, as described in ref. 23. The 1st dimension was at pH 5.0 in 8 M urea and the 2nd dimension was at pH 6.7 in 0.2% sodium dodecylsulfate (SDS). Amido Black stains basic proteins well but acidic proteins like L12 poorly (compare with Fig. 8).

* B. Schematic diagram of the spots in *A* with protein identification (see Discussion for RP nomenclature); prefix PRP (plastid ribosomal protein) is omitted in *B*. Some minor spots in the high molecular mass portion of the electropherogram were identified as: a, L2 dimer; b, L5 dimer; c, L20 dimer; d, L16 dimer; e, L1 fragment (see Discussion); f, minor form of L5 (seen only in some gels); g, L28 dimer; h, L34 dimer; i, not identified (yielded no N-terminal sequence and no HPLC/MS data). PRP's L36, and PSRP-5 alpha and gamma forms were not visible on the 2D-PAGE; they were resolved by HPLC (Fig. 2) and identified. See Results for the details of identification strategy.

FIG. 2. **Resolution of spinach plastid TP50 by reversed-phase HPLC and identification of each eluted PRP by 2D-PAGE.**

One mg of TP50 was resolved on a Vydac C18 column (4.6 x 250 mm) using a step linear gradient of solvent #1 (0.1% TFA) and solvent #2 (0.1% TFA in isopropanol). Program: 90% #1/10% #2 from 0-10 min, 75% #1/25% #2 at 70 min, 54% #1/46% #2 at 250 min, followed by a washing step, 20% #1/80% #2 at 270 min, at constant flow rate of 0.5 ml/min, fraction size, 375 µl. 2D-gel patterns for 3 pools (L15/L17, L3, L12) and their co-electrophoresis (left, PRP pool; right, PRP pool + TP50) are shown beneath the HPLC-profile. Note: PRP's L35, L33, L32, L27, and PSRP-5 alpha-gamma forms were each
eluted in two different peaks. ESI MS analysis of the pools of the distinct peaks showed the same protein mass (see Results).

FIG. 3. Yields of N-terminal amino acids from 50S PRP sequencer runs.
200 pmol of TP50 or TP70 were subjected to 2D-PAGE, electroblotted onto PVDF-membrane, and individual spots were subjected to N-terminal protein sequencing. Yield was calculated from the average PTH-AA recovery for the first three sequence cycles for each spot. The average apparent yield for all the 50S PRP’s (L12 was counted as 4 copies, see text) is 27.5 pmol/protein. The actual recovery of PTH-AA is 74 pmol (37%), because only 50 µl out of 135 µl is injected for PTH-AA identification. PRP’s L10, L18, L31 and PSRP-5 exist in multiple forms and are indicated in inset. Plastid RRF (ribosome recycling factor) is found to be a 70S ribosome-specific protein (this work, see text).

FIG. 4. Alignments of five spinach plastid 50S subunit protein sequences with homologous sequences from eubacteria, alga and land plants. The spinach plastid sequence (derived from cDNA, see Results) is at top. Arrows (black) indicate cleavage sites for removing the transit peptide; gray arrows indicate sites which are either alternative transit peptide cleavage site or processing points after PRP is imported into plastid. Tilde and dot indicate blanks and gaps, respectively. Underlined sequences are protein sequences determined experimentally (Table I). The chain length of mature protein (that of precursor in parenthesis), and percents identity (I) and similarity (S) are shown after C-terminus. PSRP-5 data: Sol1, PSRP-5 sequenced in this work; Sol2, sequence reported as spinach L40 (28): Psa, pea sequence reported as PsCL18 (33). PSRP-6 data: Sol, PSRP-6 sequenced in this work; Psa, pea sequence reported as PsCL25 (33). Abbreviations: Sol, Spinacia oleracea; Ath, Arabidopsis thaliana; Ppu, Porphyra purpurea (alga); Syn, Synechocystis PCC6803 (cyanobacterium); Eco, Escherichia coli;
Psa, *Pisum sativum* (pea). New accession numbers from this work are: Sol L5 (AF250923), Sol L19 (AF250384), Sol L34 (AF238221), Sol PSRP-5 (AF261940), and Sol PSRP-6 (AF245292).

**FIG. 5.** Liquid Chromatography-Mass Spectrometry (LC/MS) analysis of spinach chloroplast 50S subunit proteins.

*A.* Relative abundance (m/z) of PRP’s from the range of m/z 400 to m/z 2000 versus HPLC elution time. TP50 (100 pmol) was loaded on a Microbore C18 (1 x 150 mm) column and resolved by a step linear gradient of solvent #1: 0.1% TFA in 2% acetonitrile and solvent #2: 0.1% TFA in 90% acetonitrile (95% #1/5% #2 at 0 min, 65% #1/35% #2 at 22 min, 64% #1/36% #2 at 23 min, 35% #1/65% #2 at 55 min, followed by a washing step, 5% #1/95% #2 at 56 min) at constant flow rate of 50 µl/min. The ion series of a protein discussed in *B* is indicated by an arrow.

*B.* An example of mass spectrum for elution interval 41.5 - 42.0 min. Peaks represent individual charged ions. Convoluted charge and m/z values are indicated above each peaks.

*C.* Deconvoluted mass spectrum of the m/z series in *B*. It indicates a major protein (PRPL12) of mass 13811.0 Da, and a minor form, 13850.2 Da (see text).

**FIG. 6.** Evidence for alpha-N-monomethyl methionine in plastid L16.

Samples of L16 and SSU (small subunit of ribulose 1, 5-bisphosphate carboxylase/oxygenase), 200 pmol each, were subjected to N-terminal analysis. Results from the first two cycles are shown. NMM, PTH-alpha-N-monomethyl methionine; NMM-deriv., uncharacterized PTH-alpha -N-monomethyl methionine derivative; a, *N*,*N*-dimethyl-*N*'-phenylthiourea; b, *N*,*N*-diphenylthiourea. Minor amounts of PTH-Ala (L27) in cycle 1 and PTH-His (L27) in cycle 2 of PRPL16 sequencer run were from L27 contamination (see Fig. 1). SSU HPLC and SSU 2D/blot, and L16 2D/blot stand for SSU purified by HPLC,
SSU purified by 2D-PAGE/blotting, and L16 purified by 2D-PAGE/blotting, respectively (L16 could not be purified by HPLC, see Fig. 2).

**FIG. 7.** A protein is identified as present in plastid 70S ribosome but absent in either 30S or 50S subunits. A portion of Coomassie Blue stained 2D-PAGE patterns of TP30, TP50 and TP70 are shown. The arrow indicates a protein spot (marked RRF) in TP70, and its absence in the 2D gels of TP30 and TP50. PRP's S1 and L1 are shown for orientation. The N-terminal and an internal peptide sequences of the 70S-specific protein were determined that permitted its identification.

**FIG. 8.** Plastid ribosomal proteins of spinach chloroplast 70S ribosome separated by 2D-PAGE.

* A. Coomassie Blue stained electropherogram of spinach chloroplast 70S RP’s. TP70 (200 pmol) was subjected to 2D-PAGE as described (23), see legend to Fig. 1 for details.

* B. Schematic diagram of the spots in A with protein identification. 30S PRP’s are shown in light gray, 50S PRP’s are shown in dark gray and the 70S ribosome-specific protein (RRF) is shown in black. Plastid-specific ribosomal proteins (PSRP’s) are written in bold letters. Proteins L36, and the beta and gamma forms of PSRP-5 were not seen in the 2D-PAGE; they were isolated and identified by HPLC resolution (Fig. 2).

**FIG. 9.** Changes in the molecular mass of individual chloroplast 50S subunit proteins as compared to that of corresponding RP's of *E. coli*.

Chloroplast (plastid) ribosomal protein masses are taken from Table II. The *E. coli* 50S RP masses are taken from (53), as determined by mass spectrometry (MALDI-TOF).
FIG. 10. The secondary structure differences in Domains I, II, III and VI of *E. coli* and tobacco chloroplast 23S rRNA and locations in the 3-D arrangement of rRNA in the 50S subunit.

Helix numbers shown in the secondary structure schematics and 3-D structure are taken from ref. 75. Only those regions in domains I-III and VI are shown where there are some significant differences (domains IV and V are highly conserved). Domain VI includes 4.5S rRNA. The 3-D arrangement of rRNA in the 50S subunit is taken from Mueller *et al.* (75). Cp, central protuberance; StL7/12 stalk; Exit, exit site of nascent peptide, indicated by dotted circle (view from solvent side), or arrow (view from L7/L12 stalk side). De9 and De98, positions of deleted helices 9 and 98. Ad38.2, Ad58.2 and Ad100, added loops in chloroplast 23S rRNA. The Ad 38.2 loop (nucleotides 931-933) interrupts the L30 binding site in 23S rRNA (see text).
Table I. N-terminal and internal peptide sequences of spinach plastid 50S ribosomal proteins and a 70S ribosome-specific protein.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>N-terminal sequence</th>
<th>Swiss-Prot accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>50S subunit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>AAVAEAEVA DMEEEXGEXK XVA</td>
<td>P82178</td>
</tr>
<tr>
<td>L2</td>
<td>*AIGHYVGSTS STR</td>
<td>P06509</td>
</tr>
<tr>
<td>L3</td>
<td>SVDAGIGVNG TKLGMSSPE EDGTVVPVT IGFKEGNIVT QVKTE</td>
<td>P82191</td>
</tr>
<tr>
<td>L4</td>
<td>ELIPLPLINF SGEKVGET</td>
<td>O49937</td>
</tr>
<tr>
<td>L5</td>
<td>AAGTAFFDVK AAEITINL RK RYQEX VEEY</td>
<td>P82192</td>
</tr>
<tr>
<td>L6</td>
<td>KEGIRGROGI TVENVIAEM GDQLKVGKP</td>
<td>P82193</td>
</tr>
<tr>
<td>L9</td>
<td>Q/KVWKIRKI ILKEDIDPG KGQQLDVRA GLRNLFFLLP GKAAV</td>
<td>P82180</td>
</tr>
<tr>
<td>L10α</td>
<td>AISRTKKEET VETVQHLEX YLLA</td>
<td>P82139</td>
</tr>
<tr>
<td>L10β</td>
<td>AISRTK</td>
<td>P82181</td>
</tr>
<tr>
<td>L10γ</td>
<td>AISRTK</td>
<td>P82182</td>
</tr>
<tr>
<td>L11</td>
<td>APKPGKAKKV IGVIKLAE LA GKATPAPFVG</td>
<td>P31164</td>
</tr>
<tr>
<td>L12</td>
<td>AVEAPEKEIQ LGTQLGLT</td>
<td>P02398</td>
</tr>
<tr>
<td>L13</td>
<td>RKSTS ATRKC TEE</td>
<td>P12629</td>
</tr>
<tr>
<td>L14</td>
<td>MIQPGTHLNV</td>
<td>P09596</td>
</tr>
<tr>
<td>L15</td>
<td>SASSSNVSPS IGSGSETRF</td>
<td>P22798</td>
</tr>
<tr>
<td>L16</td>
<td>†MLSPKRFTRFR KCVR</td>
<td>P17353</td>
</tr>
<tr>
<td>L17</td>
<td>MKHRKRIHAL SXPADQXRL LR</td>
<td>P82194</td>
</tr>
<tr>
<td>L18α</td>
<td>XANTREDRT ARHVRIRKV EGTPERXXL</td>
<td>P82195</td>
</tr>
<tr>
<td>L18β</td>
<td>XANTREDRT ARHVRIRKV EGTPER</td>
<td>P82195</td>
</tr>
<tr>
<td>L19</td>
<td>SEAEDDSAP AESVAVAE ELPVE</td>
<td>P82413</td>
</tr>
<tr>
<td>L19 peptide 1</td>
<td>SPNIKIEKV SHRV</td>
<td>P82413</td>
</tr>
<tr>
<td>L19 peptide 2</td>
<td>LPLSTFK</td>
<td>P82413</td>
</tr>
<tr>
<td>L20</td>
<td>TRVKERGIAVR XRKEIRFFA SXFRGA</td>
<td>P28803</td>
</tr>
<tr>
<td>L21</td>
<td>AKRRRFQIEIP EELKAEFEF QRPF</td>
<td>P24613</td>
</tr>
<tr>
<td>L22</td>
<td>GFFKKK</td>
<td>P09594</td>
</tr>
<tr>
<td>L23</td>
<td>ATVEAVAPTE EAATSQ</td>
<td>submitted</td>
</tr>
<tr>
<td>L24</td>
<td>RIKWRRK</td>
<td>P27683</td>
</tr>
<tr>
<td>L27</td>
<td>AKKGAGGSTK NGRDSKGQRL GVKYIGDQVA KPGAIIRRQR GTFHPGKVN</td>
<td>P82190</td>
</tr>
<tr>
<td>L28</td>
<td>RRIXFPTGKK SNKARVXH</td>
<td>P82245</td>
</tr>
<tr>
<td>L29</td>
<td>VKKEDELKEL RYKTEELN EILQLK</td>
<td>P82248</td>
</tr>
<tr>
<td>L31α</td>
<td>RRSDIHFEFR EDARKVIXGE LVM</td>
<td>P82249</td>
</tr>
<tr>
<td>L31β</td>
<td>RRSDIHFEFR ED</td>
<td>P82246</td>
</tr>
<tr>
<td>L31γ</td>
<td>RRSDIHFEFR ED</td>
<td>P82247</td>
</tr>
<tr>
<td>L32</td>
<td>AVPKKRP SLY KIRKINIVK KKGYSALK</td>
<td>P28804</td>
</tr>
<tr>
<td>L33</td>
<td>AKGKDFRKV SKLCEGVRK SVNKGSRGS VS RYITQSNRHN TPCRELARKF</td>
<td>P28805</td>
</tr>
<tr>
<td>L34</td>
<td>GKAALXKTR</td>
<td>P82244</td>
</tr>
<tr>
<td>L35</td>
<td>AKGYQMVTNK ASAKRFRVTG KG1IVRXAG KQLHAXKNT KXNNXLKKL</td>
<td>P23326</td>
</tr>
<tr>
<td>L36</td>
<td>MKIRASVRXI CEKCRILRX R G</td>
<td>P12230</td>
</tr>
<tr>
<td>PSRP-5α</td>
<td>VSAAAE</td>
<td>submitted</td>
</tr>
<tr>
<td>PSRP-5β</td>
<td>ISVQNLPLES</td>
<td>submitted</td>
</tr>
<tr>
<td>PSRP-5γ</td>
<td>NLPLESKLQL</td>
<td>submitted</td>
</tr>
<tr>
<td>PSRP-6</td>
<td>SRRPQKKGTA HHMKTRP</td>
<td>P82411</td>
</tr>
</tbody>
</table>

| **70S ribosome-specific** |                     |                             |
| RRF           | ATMEEEVAAEK SLEITNTKQR M | P82231                     |
| RRF peptide 1| DRRKELSXV SKLAEGRKVA VRNI | P82231                     |

X, amino not identified. *, N-terminal alanine of L2 is α-N-monomethylated (30). †, 10% N-terminal methionine of L16 is α-N-monomethylated; 90% is blocked to Edman reaction. Protein sequences determined in the current study are represented in bold (other data are published from our laboratory.) Prefix PRP is omitted.
Table II. Mass spectrometry of spinach chloroplast 50S ribosomal proteins

<table>
<thead>
<tr>
<th>PRP name</th>
<th>Sequence mass (Da)</th>
<th>Elution time (min)</th>
<th>Observed mass (Da)</th>
<th>Mass difference (Da)</th>
<th>Observed mass (Da)</th>
<th>Mass difference (Da)</th>
<th>SDS-PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>30443.87</td>
<td>33.5</td>
<td>30442.7&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-1.2</td>
<td>30447.0</td>
<td>3.1</td>
<td>31.5</td>
</tr>
<tr>
<td>L2</td>
<td>29661.04&lt;sup&gt;g&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>30.5</td>
</tr>
<tr>
<td>L3</td>
<td>27163.50</td>
<td>38.0</td>
<td>27221.8&lt;sup&gt;g&lt;/sup&gt;</td>
<td>58.3</td>
<td>27235.0</td>
<td>71.5</td>
<td>31.0</td>
</tr>
<tr>
<td>L4</td>
<td>24209.96</td>
<td>37.5</td>
<td>24207.8&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-2.2</td>
<td>ND</td>
<td>ND</td>
<td>24.5</td>
</tr>
<tr>
<td>L5</td>
<td>29.0</td>
<td>20224.9&lt;sup&gt;g&lt;/sup&gt;</td>
<td>20244.0</td>
<td>21.5</td>
<td>204.4</td>
<td>ND</td>
<td>24.5</td>
</tr>
<tr>
<td>L6</td>
<td>31.5</td>
<td>17665.8&lt;sup&gt;g&lt;/sup&gt;</td>
<td>17674.8</td>
<td>16.0</td>
<td>ND</td>
<td>ND</td>
<td>21.5</td>
</tr>
<tr>
<td>L10α</td>
<td>35.5</td>
<td>20305.8&lt;sup&gt;g&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>16.5</td>
</tr>
<tr>
<td>L10β</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>16.5</td>
</tr>
<tr>
<td>L11</td>
<td>16667.62</td>
<td>34.5</td>
<td>16749.1&lt;sup&gt;g&lt;/sup&gt;</td>
<td>81.5</td>
<td>ND</td>
<td>ND</td>
<td>16.5</td>
</tr>
<tr>
<td>L12</td>
<td>13814.54</td>
<td>41.5</td>
<td>13811.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-3.5</td>
<td>13813.0</td>
<td>-1.5</td>
<td>17.5</td>
</tr>
<tr>
<td>L13</td>
<td>23361.83</td>
<td>31.5</td>
<td>23360.2&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-1.6</td>
<td>23365.0</td>
<td>3.2</td>
<td>24.5</td>
</tr>
<tr>
<td>L14</td>
<td>13464.58</td>
<td>37.0</td>
<td>13463.6&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-1.0</td>
<td>ND</td>
<td>ND</td>
<td>13.5</td>
</tr>
<tr>
<td>L15</td>
<td>26.0</td>
<td>22327.9&lt;sup&gt;g&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>22.0</td>
</tr>
<tr>
<td>L16</td>
<td>15327.85</td>
<td>25.5</td>
<td>15402.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>74.1</td>
<td>15467.0</td>
<td>139.2</td>
<td>15.0</td>
</tr>
<tr>
<td>L17</td>
<td>17566.17</td>
<td>26.5</td>
<td>17564.6&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-1.6</td>
<td>17566.0</td>
<td>-0.2</td>
<td>20.0</td>
</tr>
<tr>
<td>L18α</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>11.5&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>L19</td>
<td>15557.18</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>16.5</td>
</tr>
<tr>
<td>L20</td>
<td>22762.86</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>30.0</td>
</tr>
<tr>
<td>L21</td>
<td>23114.09&lt;sup&gt;g&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>24.5</td>
</tr>
<tr>
<td>L22</td>
<td>13553.66</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>13553.5&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-0.2</td>
</tr>
<tr>
<td>L24</td>
<td>16513.14</td>
<td>26.0</td>
<td>16406.5&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-106.6</td>
<td>16425.0</td>
<td>-88.1</td>
<td>16.5</td>
</tr>
<tr>
<td>L27</td>
<td>23.5</td>
<td>13646.5&lt;sup&gt;g&lt;/sup&gt;</td>
<td>13650.2</td>
<td>16.0</td>
<td>13650.2</td>
<td>-1.5</td>
<td>16.0</td>
</tr>
<tr>
<td>L28</td>
<td>25.0</td>
<td>8991.8&lt;sup&gt;g&lt;/sup&gt;</td>
<td>9071.5</td>
<td>13.0</td>
<td>9071.5</td>
<td>ND</td>
<td>13.0</td>
</tr>
<tr>
<td>L29</td>
<td>27.0</td>
<td>13433.6&lt;sup&gt;g&lt;/sup&gt;</td>
<td>13502.1</td>
<td>15.0</td>
<td>13502.1</td>
<td>ND</td>
<td>15.0</td>
</tr>
<tr>
<td>L31α</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>11.0&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>L31β</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td>L31γ</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td>L32</td>
<td>6503.67</td>
<td>24.5</td>
<td>6503.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-0.7</td>
<td>6505.0</td>
<td>1.3</td>
<td>11.5</td>
</tr>
<tr>
<td>L33</td>
<td>7516.81</td>
<td>19.5</td>
<td>7512.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-4.8</td>
<td>7515.0</td>
<td>-1.8</td>
<td>11.0</td>
</tr>
<tr>
<td>L34</td>
<td>6767.95</td>
<td>18.0</td>
<td>6767.1&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-0.8</td>
<td>6939.3</td>
<td>171.4</td>
<td>NA</td>
</tr>
<tr>
<td>L35</td>
<td>8433.02</td>
<td>21.0</td>
<td>8432.7&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-0.3</td>
<td>8449.9</td>
<td>16.9</td>
<td>12.5</td>
</tr>
<tr>
<td>L36</td>
<td>4401.35</td>
<td>19.0</td>
<td>4399.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-2.4</td>
<td>4492.6</td>
<td>91.3</td>
<td>NA</td>
</tr>
<tr>
<td>PSRP-5α</td>
<td>9254.97</td>
<td>23.0</td>
<td>9296.0&lt;sup&gt;g&lt;/sup&gt;</td>
<td>41.0</td>
<td>ND</td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td>PSRP-5β</td>
<td>7065.79</td>
<td>23.0</td>
<td>7065.0</td>
<td>-0.8</td>
<td>ND</td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td>PSRP-5γ</td>
<td>6638.29</td>
<td>22.5</td>
<td>6636.5</td>
<td>-1.8</td>
<td>ND</td>
<td>ND</td>
<td>NA</td>
</tr>
<tr>
<td>PSRP-6</td>
<td>7387.18&lt;sup&gt;g&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>12.0</td>
</tr>
</tbody>
</table>

<sup>a</sup> Calculated from translated DNA sequence using ‘peptidesort’ program (31). <sup>b</sup> Elution time from Fig. 5A. <sup>c</sup> Mass obtained using deconvolution program. <sup>d</sup> Calculated as observed mass minus sequence mass. <sup>e</sup> Mass obtained by ESI MS of HPLC pool containing identified PRP(s) in Fig. 2. <sup>f</sup> Taken from ref. 2. <sup>g</sup> Mass values used in Fig. 9. NA, not available. ND, not determined.
### Table III

Post-translational processing/modification of spinach 50S PRP’s and 70S ribosome-specific P-RRF

#### Plastid encoded (8 proteins)

<table>
<thead>
<tr>
<th>PRP name</th>
<th>N-terminal sequence</th>
<th>Post-translational processing/modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>Met NMA-Ile-His-Leu-Tyr-Lys</td>
<td>1) fMet removed; 2) α-N-Ala monomethylated</td>
</tr>
<tr>
<td>L14</td>
<td>Met-Ile-Gln-Pro-Gln-Thr-His</td>
<td>Formyl group removed</td>
</tr>
<tr>
<td>L16</td>
<td>NMM-Leu-Ser-Pro-lys-Arg-Thr</td>
<td>1) Formyl group removed; 2) 10% α-N-Met monomethylated; 3) unknown internal modification</td>
</tr>
<tr>
<td>L20</td>
<td>Met Thr-Arg-Val-Lys-Arg-Gly</td>
<td>fMet removed</td>
</tr>
<tr>
<td>L22</td>
<td>Met Gly-Phe-Phe-Lys-Lys-Lys</td>
<td>fMet removed</td>
</tr>
<tr>
<td>L32</td>
<td>Met Ala-Val-Pro-lys-Lys-Arg</td>
<td>fMet removed</td>
</tr>
<tr>
<td>L33</td>
<td>Met-Lys-Ile-Arg-Ala-Ser-Val</td>
<td>Formyl group removed</td>
</tr>
</tbody>
</table>

#### Nuclear encoded (25 50S proteins + P-RRF)

<table>
<thead>
<tr>
<th>PRP name</th>
<th>Transit peptide cleavage site</th>
<th>Additional post-translational processing/modification during/after import into chloroplast</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>---SPKSVFRVL ↓ AAVAAEAeva-------</td>
<td>No modification</td>
</tr>
<tr>
<td>L3</td>
<td>---XXXXXXXXXX ↓ SVDAGIGVMG-------</td>
<td>No modification</td>
</tr>
<tr>
<td>L4</td>
<td>---QTSSLISIKS ↓ ELIPLPILNF-------</td>
<td>No modification</td>
</tr>
<tr>
<td>L5</td>
<td>---XXXXXXXXXX ↓ AAGTAVFSDK-------</td>
<td>No modification</td>
</tr>
<tr>
<td>L6</td>
<td>---XXXXXXXXXX ↓ KESRIGQPI-------</td>
<td>Epsilon-trimethylation of Lys9 and Lys45 (+81.5 Da)</td>
</tr>
<tr>
<td>L9</td>
<td>---XXXXXXXXXX ↓ QKKVKKIRKI-------</td>
<td></td>
</tr>
<tr>
<td>L10</td>
<td>---XXXXXXXXXX ↓ AISRTKKEET-------</td>
<td></td>
</tr>
<tr>
<td>L11</td>
<td>---SHRRLSIVAM ↓ APKPGKAKKV-------</td>
<td></td>
</tr>
<tr>
<td>L12</td>
<td>---TRTTLRPIA ↓ AVEAPEKEQ-------</td>
<td></td>
</tr>
<tr>
<td>L13</td>
<td>---TVAFAVVSA ↓ RKSTSSASTRC-------</td>
<td></td>
</tr>
<tr>
<td>L15</td>
<td>---XXXXXXXXXX ↓ SASSSNVSPS-------</td>
<td></td>
</tr>
<tr>
<td>L17</td>
<td>---XXXXXXXXXX ↓ MKHGRKIHR-------</td>
<td></td>
</tr>
<tr>
<td>L18</td>
<td>---XXXXXXXXXX ↓ KAHTRREDRT-------</td>
<td></td>
</tr>
<tr>
<td>L19</td>
<td>---FPARNSFVVR ↓ SAEEDSSDAP-------</td>
<td></td>
</tr>
<tr>
<td>L21</td>
<td>---LSRLSLLPVA ↓ AKRRRFQEI-------</td>
<td></td>
</tr>
<tr>
<td>L22</td>
<td>---SFGRDLMAQ ↓ LTEAVAPTTE-------</td>
<td></td>
</tr>
<tr>
<td>L23</td>
<td>---EGHSPCIVM ↓ RIKWERKDC-------</td>
<td></td>
</tr>
<tr>
<td>L24</td>
<td>---XXXXXXXXXX ↓ AHKKGAGSTK-------</td>
<td></td>
</tr>
<tr>
<td>L25</td>
<td>---XXXXXXXXXX ↓ RPIXPFTGKK-------</td>
<td></td>
</tr>
<tr>
<td>L26</td>
<td>---XXXXXXXXXX ↓ VKKEDELKEL-------</td>
<td></td>
</tr>
<tr>
<td>L31</td>
<td>---XXXXXXXXXX ↓ RKSDIHPFPR-------</td>
<td></td>
</tr>
<tr>
<td>L34</td>
<td>---DRCRFVVRSA ↓ GKAALCLTKR-------</td>
<td></td>
</tr>
<tr>
<td>L35</td>
<td>---TSSPSTTFVA ↓ AKGYKMKTHK-------</td>
<td></td>
</tr>
<tr>
<td>PSRP-5</td>
<td>---AQKRGTVVM ↓ VSSAEEETAG-------</td>
<td></td>
</tr>
<tr>
<td>PSRP-6</td>
<td>---CGGGSITIIEC ↓ SSRPQKKGTA-------</td>
<td></td>
</tr>
<tr>
<td>P-RRF</td>
<td>---QNRAGTFCRA ↓ ATMEEVEAEK-------</td>
<td></td>
</tr>
</tbody>
</table>

XX, data not available (cDNA/genomic DNA not sequenced).
Figure 1
Yamaguchi & Subramanian
Figure 2
Yamaguchi & Subramanian
Figure 3
Yamaguchi & Subramanian
Figure 4
Yamaguchi & Subramanian
Figure 5

Yamaguchi & Subramanian
Figure 6
Yamaguchi & Subramanian
Figure 7
Yamaguchi & Subramanian
Figure 8
Yamaguchi & Subramanian
Figure 9
Yamaguchi & Subramanian
*E. coli* 23S + 5S rRNA 3-D arrangement

**Fig. 10**
Yamaguchi & Subramanian
The Plastid Ribosomal Proteins (2): Identification of all the Proteins in the 50S Subunit of an Organelle Ribosome (Chloroplast)
Kenichi Yamaguchi and Alap R Subramanian

J. Biol. Chem. published online June 28, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M005012200

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

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2000/06/28/jbc.M005012200.citation.full.html#ref-list-1