Architecture of higher plant chloroplast ribosomes involves additional protein domains over that found in the Escherichia coli ribosome, although the rRNAs in these two kinds of ribosomes are very similar in length and sequence (Subramanian, A. R. (1993) Trends Biochem. Sci. 18, 177-180). Here, we show that two chloroplast-specific protein domains (a novel chloroplast ribosomal protein of the 30 S subunit, called Psrpl-1 or S22, and a divergent protein of the 50 S subunit with long terminal extensions and low homology to its E. coli counterpart, L21) are both incorporated in E. coli ribosomes and polysomes when their gene constructs are expressed in E. coli. Also, the 67-residue NH₂-terminal extension in chloroplast L21 by itself is incorporated. Thus, our results indicate preexisting binding sites for novel chloroplast-specific ribosomal proteins/domains on eu-bacterial ribosomes. Additionally, we observed cleavage of the chloroplast-targeting transit peptide (present in the expressed Psrpl-1 precursor), indicating protease(s) of the required specificity in E. coli cells. The expression of chloroplast L21 with its NH₂-terminal extension was inhibitory to E. coli growth, suggesting a drastic effect of the latter on some property of L21. Expression of Psrpl-1 was neutral, consistent with a function only in chloroplast translation. Based on analysis of the assembly of Psrpl-1 and various L21 fragments in E. coli ribosomes, a general model for studying ribosomal protein-ribosome interactions is suggested.

Chloroplast ribosomes have many similarities to eu-bacterial ribosomes (1, 2), but the most striking difference between them is in their protein moiety (3-6), which, in higher plants, is supplemented by 5-10 additional proteins (6). Five of these have been characterized to date, variously at protein (7-9), cDNA (7, 8, 10), and gene (11) levels. In four cases, no significant homology to any known ribosomal protein (RP) or other sequenced proteins is found (7, 8, 10), suggesting their evolution specifically for the chloroplast translational system. In the fifth case, homology to a polypeptide present in several thermophilic bacteria (12) was observed (9). All of these novel chloroplast RPs are encoded in the nuclear genome (2).

Over 35 further chloroplast RP sequences, each homologous to a corresponding Escherichia coli RP, have also been reported (reviewed in Refs. 13-16). Many of these RPs contain NH₂- and COOH-terminal extensions (NTE, CTE) of considerable length beyond the homologous sequence regions (2, 6). Nothing is known about the origin or role in translation of chloroplast-specific RPs and extensions, but it is possible to envision that they help generate the structural and functional uniqueness in chloroplast ribosomes. In contrast, the RNA of higher plant chloroplast ribosomes is very similar in chain length and primary/secondary structure to eu-bacterial rRNA (17, 18).

To understand the unique features of chloroplast ribosomes at the biochemical level, we have examined the behavior of two such proteins (and their derivatives) at the assembly stage of E. coli ribosomes. Psrpl-1 is an acidic, relatively high M, protein present as a stoichiometric component in two-dimensional gels of sucrose gradient-purified spinach chloroplast 30 S ribosomal subunits (7). It has been isolated and sequenced, and its cDNA clones (7, 8) and a genomic clone have been characterized (11). The Psrpl-1 (designated S22 in Ref. 11) sequence showed no detectable homology to any of the known eu-bacterial, archae-bacterial, mitochondrial, or cytosolic RPs (7, 8). The second protein we chose was chloroplast L21, which is double the size of its E. coli counterpart and contains a 67-residue NTE and 30-residue CTE (19, 20). Its cDNA clones and a genomic clone have been characterized (19-21). The central 103-residue region of chloroplast L21 is homologous to the entire E. coli L21. The extensions (present only in higher plant chloroplast L21; see Ref. 19) have no homology to any reported protein sequences in data bases.

We have taken the previously cloned (7, 19) coding regions of spinach chloroplast Psrpl-1 and L21 and have expressed various constructs made from them in E. coli. The presence of the expressed proteins in ribosomal particles was assayed by sucrose gradient fractionation followed by Western blotting/immunostaining with antibodies specific to chloroplast RPs. Thus, incorporation of these “foreign” proteins by the E. coli ribosome during assembly and utilization in protein biosynthesis, i.e. participation in polysome formation, could be determined. The results are presented and discussed in this paper.

MATERIALS AND METHODS

Spinach Chloroplast RP cDNAs—cDNAs of spinach (Spinacia oleracea, cv. Matador) chloroplast Psrpl-1 and L21 (7, 19), cloned in gpt11 as previously described (23), were used for this study.

DNA Manipulations—All DNA manipulations were done according to standard procedures (24, 25). DNA fragments were isolated from agarose gels after electrophoresis using GeneClean II (BIO 101, Inc., Vista, CA). E. coli strains XL1 and JM110 (Stratagene), and WR6 (26) were used for transformation and expression.

Psrpl-1 Constructs for Its Expression—The Psrpl-1 cDNA insert in gpt11 recombinant phage was subcloned into pT7/T3-19U (Pharmacia Biotech Inc.) in the EcoRI site and used as the starting material. The Asn1-EcoRI fragment of the insert, containing almost the whole cDNA sequence (without 26 base pairs of 5'-untranslated region and the first

(Received for publication, November 3, 1993, and in revised form, January 24, 1994)

Mikhail G. Bubunenko and Alap R. Subramanian

From the Max-Planck-Institut fur Molekulare Genetik, Ihnestrasse 73, 14195 Berlin-Dahlem, Germany
RESULTS AND DISCUSSION

Expression of Psrp-1 Protein and Its Deletion Derivatives in E. coli—Three derivatives of Psrp-1 and the complete cDNA encoding the mature protein and transit peptide were expressed from the thermosensitive promoter of pJLA502 (Fig. 1A). Since some of the constructs carried the amber (TAG) stop codon, all were expressed in the E. coli strain W66, which lacks the amber suppressor TRP1 (16). In no case was a distinct Coomassie Blue staining band observed; therefore, all further analyses were performed by immunostaining. Evidently, all four constructs were expressed in E. coli (Fig. 1B) albeit at a level not detected by dye staining.

With pMB1 plasmid, which contains the whole cytoplasmic precursor of Psrp-1, four specifically immunoreactive major bands were observed (Fig. 1B). The highest M, band corresponded to the size of Psrp-1 precursor (M, = 33,752) (7), whereas the lowest of the four bands migrated with about the same mobility of the mature protein (M, = 26,791). Therefore, the latter must have arisen in E. coli by an approximately
specific removal of the transit peptide. Processing of the transit peptides of certain chloroplast-targeted proteins by *E. coli* membrane signal peptidase has been previously reported (32). However, unlike these cases where the processed protein is secreted, PsrP-1 occurs in the cytoplasm (see below). Therefore, its transit peptide is cleaved off by an *E. coli* cytoplasmic protease. The other two major bands (of intermediate size) found in this experiment probably represent processing intermediates.

Protein expression from pMBS14, which lacks the 56-residue acidic region in PsrP-1 (7), was significantly lower; the processing intermediates (if present) were hardly visible (Fig. 1B). With the pMBS5 and pMBS13 encoding the COOH-terminal and NH₂-terminal regions, respectively, there is evidence that the higher molecular mass bands (Fig. 1B) are derived from translational read-through at the UGA stop codon (data not shown). With pMBS5 expression product that includes the acidic region, an apparent anomalous electrophoretic migration was noted; the mobility corresponded to a significantly higher Mr than the 14,270 of the expressed polypeptide (Fig. 1B). Such anomalous mobility is reminiscent of the acidic RNA-binding domain of *E. coli* S1 (33).

Incorporation of PsrP-1 in *E. coli* Ribosomes—The PsrP-1 protein and its derivatives were examined for binding to *E. coli* ribosomes by isolating ribosomes and polysomes from sucrose gradients (centrifuged in high-salt buffer conditions to avoid non-specific binding). With both the PsrP-1 precursor and its processed forms including the mature protein, a relatively efficient binding was observed, i.e. 84% of the expressed protein was present in the ribosomal particles (Fig. 2A). The precursor was present in polysomes (fractions 6–9) as well, while the mature protein was present in monosomes (MBS1, fractions 10–11) and subunits, a typical separation profile being shown in Fig. 4A. The presence of PsrP-1 in fractions containing 50 S subunits may arise from the overlapping of the 70 and 30 S peaks and the dissociation (from ribosomes) in the gradient. The NH₂-terminal (pMBS13) and COOH-terminal (pMBS5) parts of PsrP-1 bound less strongly (Fig. 2, C and D), i.e. 55 and 47% of the expressed protein, respectively, as would be expected for a protein whose binding sites are organized by both parts of the molecule interacting with the ribosome.

The interaction of the polypeptide with the deleted central region is of special interest (pMBS14, Fig. 2B). The expressed protein (>90%) was found mainly in 70 S ribosomes and 30 S subunits (fractions 10 and 14). Its preferential affinity for the small ribosomal subunit is in line with the mode of binding of intact PsrP-1. The strong binding of this deletion construct argues that the deleted central region (with its unusual sequence characteristics in the context of the entire PsrP-1 polypeptide chain discussed in Ref. 7) is involved in some other function than interaction with ribosomes.

Effect of PsrP-1 Expression on *E. coli* Growth—Expression of PsrP-1 or its derivatives did not reveal any effect on *E. coli* cell growth (data not shown), even though in all cases except that of pMBS14 the chloroplast protein is incorporated in *E. coli* polysomes, i.e. ribosomes actively participating in protein synthe-
Analysis of expression of all the L21 constructs led to findings that can be summarized as follows. All constructs containing the whole transit peptide (pMBS23, -27, and -32) were not expressed; no protein products could be detected (Fig. 3B). Removal of the 5’-terminal half of the transit peptide was sufficient to allow expression (e.g. pMBS43). The expression could be modulated to some extent by its complete removal (pMBS44) as well as by further manipulations within the L21 coding region, e.g. deletion of the 3’ coding and noncoding regions and/or insertion of an internal stop codon (Fig. 3, A and B). The 5’ half of the transit peptide coding region has an inhibitory effect on L21 expression. The 3’ noncoding sequence of L21 mRNA could form (by computer simulation) a stable stem-loop secondary structure with the coding region of the transit peptide, but removal of the 3’ region (e.g. pMBS32) did not relieve inhibition. While the real spatial structure of L21 mRNA is not known, we can only speculate that the 5’ region of the transit peptide might be involved in an alternative stable secondary structure that blocks translation of this mRNA by *E. coli* ribosomes.

A striking point observed during this study was the apparent high accessibility of L21 protein derivatives to proteolytic degradation (as inferred by the appearance of smearing and lower Mr bands), especially those containing the transit peptide or NTE (pMBS18, -19, -43, -44, and -46). Polypeptides encoded by pMBS26 and pMBS2 (two of the three smallest fragments of L21, we made) and whose expression was not detected were probably rapidly degraded. In contrast, the third smallest fragment we studied (pMBS49), the NTE of L21 by itself, was relatively stable. It represents a rather interesting case and will be discussed later.

Evidently, more experimental work is needed to clarify the problem of the poor expression of Psrp-1, L21, as well as other (34, 35) nuclear encoded chloroplast RPs in *E. coli*. In the case of L21, it seems likely that the low level of expression is post-transcriptionally mediated and that its main determinants are low translational efficiency of L21 mRNA and high proteolytic accessibility of L21 protein and its derivatives. The problem of low expression is intriguing when one considers the overexpression of chloroplast RPs that are encoded in the organelle DNA, e.g. maize chloroplast S18 and L23 (34). The assembly of both nuclear encoded and organelle-encoded RPs into chloroplast ribosomes proceeds in the organelle where, in principle, they would be equally accessible to organelle proteases. The difference in expression in *E. coli* of these two groups of genes most likely reflects a basic difference in the translational efficiency of their mRNAs by prokaryotic ribosomes. It may arise from some yet unknown difference in the structural organization between eukaryotic *cis-a-cis* organelle (prokaryotic-like) mRNAs. The usage of rare codons in the nuclear and chloroplast RP mRNAs is similar (34). The chloroplast L21 protein and its mRNA could be a particularly useful tool for studying this problem because its gene in a lower land plant (liverwort, *Marchantia polymorpha* (22)) is chloroplastic, in contrast to the case in spinach and probably other flowering plants where it is nuclear (19, 20). Thus, a direct comparison of two forms of a gene from the two genome compartments, with respect to their expression in *E. coli*, is possible.

**Incorporation of Chloroplast L21 Protein in E. coli Ribosomes as a Model to Study RP-Ribosome Interaction**—Spinach chloroplast L21 protein has only 30% sequence identity to its *E. coli* counterpart, but the conserved amino acid residues are distributed in clusters over the entire homologous region (19). Thus, it can be regarded as an extensively mutagenized *E. coli* L21 with

---

2 M. G. Bubunenko and A. R. Subramanian, unpublished observations.

---
Fig. 3. Gene constructs of chloroplast ribosomal protein L21 and expression in E. coli. A, schematic diagram of 14 constructs. Top, L21 cDNA and the coding regions for its distinct peptide tracts. Clones pMBS23 and pMBS19 contained this DNA. The transcription/translation signals in pTrc99A vector (cf. pJLA502, Fig. 1A) is shown below it. Among the depicted constructs are the following: pMBS44, for expressing the mature L21; pMBS30, for the homologous region including CTE; pMBS27 and pMBS32, for the NTE including the transit peptide; and pMBS49, to express essentially the NTE. The remaining constructs were for expressing various other regions of the L21. Open bar, unexpressed part of L21 coding region; dotted segment, (pMBS33), deleted coding region. See text for further details. MCS, multiple cloning site. B, Western immunoblots showing expressed chloroplast protein L21 and its various derivatives. Expression was induced with IPTG. Cells were harvested 5 h after induction, lysed in SDS, and electrophoresed in 15% acrylamide gel. The numbers above electrophoretic lanes indicate pMBS clone numbers. Controls of pool 43 and cell lysate with pTrc99A are shown at left. Arrow shows the position of mature chloroplast L21. The expressed protein level was too low to be detected in several cases (pMBS23, -26, -27, -32, and -52). pMBS19 expression was also low. In the experiment shown, a three times greater amount of sample (relative to others) has been loaded.
structurally and functionally important residues left unchanged. From this point of view and based on the earlier finding that chloroplast RPs are incorporated into *E. coli* ribosome (34–36), we used L21 protein and its deletion derivatives to develop a model system for studying the minimal structural requirements for RP-ribosome interaction. The absence of immunocross-reaction between chloroplast L21 and *E. coli* ribosome under physiological conditions. The whole protein molecule is involved in this in vivo physiological conditions.

Several L21 constructs representing the main parts of the protein were examined for incorporation in *E. coli* ribosome *in vivo*. The analysis showed that L21 derivatives covering the homologous region (pMBS30) or together with the NTE (pMBS18 and -44) were specifically incorporated in both ribosomes and polysomes (Fig. 4). Approximately 94, 64, and 84% of the expressed protein was found in ribosomal particles with pMBS30, pMBS18, and pMBS44, respectively. This roughly corresponds to 40–80% of all of the ribosomes in the cell extracts, based on $A_{260}$ nm quantitation of ribosomes and a semi-quantitative immunoassay of the protein. The COOH-terminal part of L21 (pMBS33) is incorporated in the *E. coli* ribosome (70% of the expressed protein) but not into polysomes (Fig. 5A). However, the NH₂-terminal protein fragment (pMBS42) is mainly found in the supernatant fraction of the gradient (only 20% in ribosomes). It showed some tailing into the subunit region of the gradient, indicating a weak affinity for ribosomes (Fig. 5B).

A particularly interesting finding was the incorporation into the *E. coli* ribosome of the chloroplast-specific NTE in L21. This 67-residue peptide has no homology to any known RPs (or to any sequence in data bases). From its length and amino acid sequence, the L21-NTE has the potential to form a new domain in the structure of chloroplast L21. Expression of L21-NTE (pMBS49) led to the accumulation of a relatively stable product in both cell extracts and in ribosomes (Fig. 5D). Its stability to cellular proteases (in contrast to the fate of other fragments of the same size, e.g. pMBS26 and -52) (Fig. 3B) indicates a compact structure, and thus it could be a distinct protein domain. Its incorporation into *E. coli* ribosome (35% of the expressed protein) can be taken as evidence for a relatively weak but specific ribosomal interaction. This conclusion is further supported by the data with pMBS48 that additionally include half the transit peptide (30% of the expressed protein incorporated in ribosomes) (Fig. 5C).

Additional proteolytic fragments (over those originally found in cell extracts) could be detected on ribosomes with pMBS44 (Fig. 4B), which carries the whole NTE. This could indicate that this extension is exposed on the surface of the ribosome and is readily accessible to proteases in a manner shown earlier with L13 (34). Multiple degradation products were also observed on ribosomes with products of pMBS19 and -43 (data not shown).

The results demonstrate that as low as 30% sequence identity in chloroplast L21 is sufficient for it to compete with *E. coli* L21 and assemble itself in *E. coli* ribosome under physiological conditions. The whole protein molecule is involved in this interaction (as expected for a structural RP), the COOH-terminal part apparently being rather important. Since most RPs are involved, with increasing evidence (37, 38), in creation and maintenance of rRNA conformation and ribosomal functional centers, they probably are bound in the ribosome via multiple contact sites. Hence, the principles of their interaction with ribosomes (and functional roles) would be difficult to study by simple site-directed mutagenesis. We suggest that a more
Fig. 5. Investigation of the regions of chloroplast L21 responsible for interacting with E. coli ribosome. A, COOH-terminal part of L21 (pMBS33); B, NH2-terminal part of L21 (pMBS42); C, NH2-terminal part of L21, including a segment of transit peptide (pMBS48); D, the NH2-terminal extension of L21 (pMBS49). The numbers above the lunes are of the sucrose gradient fractions; the number of the 70 S peak is underlined. CL, cell lysate; Mr, prestained protein markers.

Fig. 6. Inhibitory effect of chloroplast L21 expression on E. coli cell growth. E. coli XL1 cells were transformed with the different L21 constructs shown, and colonies were streaked on LB plates with or without IPTG and incubated at 37 °C. Plasmid pMBS44, which inhibited E. coli growth, contains the complete mature L21 with both extensions. Other constructs were the following: pMBS30, homologous region of L21; pMBS33, COOH-terminal part of L21; pMBS42, COOH-terminal part of L21. They showed no noticeable effect on cell growth.

A promising approach would be to use heterologous proteins and their engineered fragments (as demonstrated in this work). First, the minimal requirements for binding is roughly estimated; then, the fine resolution of binding sites is achieved by further dissection of the fragments in combination with site-directed mutagenesis of conserved regions. This approach would be especially valuable in conjunction with E. coli strains lacking one of the RPs; more than 15 such lacking mutants are available (39).

Effect of Chloroplast L21 Expression on E. coli Growth: Evolution of Chloroplast RP Extensions—What are the consequences of assembling a chloroplast RP in E. coli ribosome either as an extra component or as substitute for an E. coli homologue? Previously, the functional complementation of rpsL mutation (S12, streptomycin resistance) in E. coli with the chloroplast rps12 gene was experimentally shown (36). Since S12 is one of the most conserved RPs, this result was not surprising. It would be more interesting if a highly divergent protein like L21 could similarly complement, but no L21 mutants of E. coli have been reported (39). However, the observed ability of chloroplast L21 protein to interact with E. coli ribosomes in vivo offered at least the possibility to examine the physiological consequence of this interaction.

Five L21 constructs (pMBS30, -33, -42, -44, and -49) were chosen for this experiment. They represent the central homologous core of the protein with or without the extensions, or mainly the NH2- and COOH-terminal extensions. The results of the experiment (Fig. 6) showed that the expression of the full-length mature L21 protein carrying the NTE (pMBS44) severely inhibited E. coli growth, whereas the expression of any of the other constructs was almost neutral. This suggests that the presence of the NTE on the L21 molecule alters some basic character of the protein and makes its presence on the E. coli ribosome toxic to the cell.

The function of the L21 protein is not understood at present. It is not known to be a part of any of the functional centers of the ribosome, but it interacts as an assembly protein (40) with two regions of 23 S rRNA (41). It participates in building the structure of the functional 50 S subunit (42). Evidently, the presence of the 67-residue extension on the NH2 terminus of the higher plant chloroplast L21 protein does not disturb the latter’s binding to E. coli ribosomes. A similar observation was
made earlier with chloroplast L13 carrying a 43-residue NTE (35). This property of chloroplast RP s carrying additional long sequences that interact with E. coli ribosomes could give some insight into their evolution. The role of the long NTEs could be made earlier with chloroplast L13 carrying a 43-residue NTE (35). This property of chloroplast RP s carrying additional long sequences that interact with E. coli ribosomes could give some insight into their evolution. The role of the long NTEs could be to help create the specific chloroplast ribosomal structure in higher plants, quite similar (1), yet different (2), from that of its prokaryotic ancestor(s). They could have emerged as new structural modules, added to preexisting prokaryotic ancestral structural cores, and coevolved with the latter to adapt the core to the evolving chloroplast. The cases of chloroplast RP s L21 and especially S18 (encoded in the chloroplast DNA), with its variably repeated heptapeptide motif (43), may illustrate this process. Interestingly, chloroplast L21 of M. polymorpha, which is encoded the organelle DNA, does not contain an NTE (as first pointed out in Ref. 7), indicating that most NTEs may have appeared after the evolution of the common ancestor of land plants.

Screening of Ribosomes from Different Organisms for Presence of Psrp-1—A protein specifically cross-reacting with anti-serum to Psrp-1 was found in the chloroplast ribosomes of spinach, maize, and C. reinhardtii (Fig. 7). No such protein was detected in cytoplasmic ribosomes (maize). E. coli ribosomes/cell extracts, or in cell extracts of a cyanobacterium, Synechocystis PCC 6803 (data not shown). The oxygenic cyanobacterial group is thought to belong to the ancient endosymbiotic progenitors of chloroplasts (44). Thus, it appears that Psrp-1 is a ubiquitous chloroplast-specific RP present in mono- and dicot plants and algae. L21 homologues were detected in chloroplast ribosomes of maize and Chlamydomonas but with significantly low intensity of cross-reaction (Fig. 7).

The absence of Psrp-1 cross-reacting material in plant cytosolic and Synechocystis ribosomes makes it difficult to speculate on the evolutionary origin of this novel protein. As pointed out (2), it could be a ribosomal component that was present in the ancestral prokaryotic chloroplast progenitor (but now maintained only in chloroplasts) or a protein recruited into ribosomes at the early stage of photosynthetic eukaryotic evolution. A discernible identity of parts of the Psrp-1 sequence to certain bacterial open reading frames in data bases has been suggested (i.e. upstream open reading frames of div gene in Bacillus subtilis and pheA in E. coli; downstream of rpoN in Azotobacter vinelandii, Klebsiella pneumoniae, Pseudomonas putida, and Rhizobium melliloti.) The recently initiated work on sequence analysis of Synechocystis RP genes (29) and completion of the sequencing projects on eukaryotic cytosolic RP s of yeasts and mammals (e.g. Ref. 45) should shed further light on this problem.

Acknowledgment—We thank Klaus von Knoblauch for skillful technical assistance.

References

Novel and Divergent Chloroplast Ribosomal Proteins

41. Oswald, M., Greuer, B., and Brimacombe, R. (1990) Nucleic Acid Res. 18, 6775–6790