Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes

KATHLEEN SZICK*, MARK SPRINGER†, AND JULIA BAILEY-SERRES*‡§

*Interdepartmental Graduate Program in Genetics and Departments of †Biology and ‡Botany and Plant Sciences, University of California, Riverside, CA 92521-0124

Communicated by M. T. Clegg, University of California, Riverside, CA, December 17, 1997 (received for review June 2, 1997)

ABSTRACT The P-protein complex of eukaryotic ribosomes forms a lateral stalk structure in the active site of the large ribosomal subunit and is thought to assist in elongation phase of translation by stimulating GTPase activity of elongation factor-2 and removal of deacylated tRNA. The complex in animals, fungi, and protozoans is composed of the acidic phosphoproteins P0 (35 kDa), P1 (11–12 kDa), and P2 (11–12 kDa). Previously we demonstrated by protein purification and microsequencing that ribosomes of maize (Zea mays L.) contain P0, one type of P1, two types of P2, and a distinct P1/P2 type protein designated P3. Here we implemented distance matrices, maximum parsimony, and neighbor-joining analyses to assess the evolutionary relationships between the 12 kDa P-proteins of maize and representative eukaryotic species. The analyses identify P3, found to date only in mono- and dicotyledonous plants, as an evolutionarily distinct P-protein. Plants possess three distinct groups of 12 kDa P-proteins (P1, P2, and P3), whereas animals, fungi, and protozoans possess only two distinct groups (P1 and P2). These findings demonstrate that the P-protein complex has evolved into a highly divergent complex with respect to protein composition despite its critical position within the active site of the ribosome.

Translation is a complex, multi-step process that involves ribosomes; initiation, elongation, and release factors; aminoacyl-tRNAs; mRNA; and mRNA-binding proteins (1–3). Ribosomes, which catalyze polypeptide synthesis, consist of 3–4 rRNA molecules and up to 90 proteins assembled into large and small subunits. Both prokaryotic and eukaryotic ribosomes have been investigated, with the eukaryotic emphasis on ribosomes of rat and yeast (4, 5), and little detailed analysis of the ribosomes of plants (reviewed in ref. 6). The overall structure and function of the eukaryotic ribosome is considered to be conserved. The small ribosomal subunits of animals, fungi, and plants are very similar in molecular mass, whereas the mass of the large ribosomal subunits is quite variable. Large ribosomal subunits of plants have a lower molecular mass than that of rat or yeast (6, 7), which is due in part to nucleotide sequence differences in the 23S-like rRNA component (6, 7), but may also result from heterogeneity in ribosomal protein (r-protein) composition.

The majority of r-proteins are basic (pI > 8.5). There are, however, a group of acidic r-proteins with isoelectric points in the pH 3–5 range, a subset of which form a distinct and universally conserved lateral-stalk structure on the large ribosomal subunit (8, 9). The stalk structure is present in the active site of the ribosome where interactions between mRNA, tRNA, and translation factors occur during the late initiation, elongation, and termination phases of translation (8).

Abbreviations: r-protein, ribosomal protein; P-protein, acidic ribosomal phosphoprotein.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. U62748 (rpp2a-2), U62749 (rpp2a-3), and U62750 (rpp2a-4)].

*To whom reprint requests should be sent at the ‡ address. e-mail: serres@mail.ucr.edu.
In a recent report (26), we demonstrated by protein purification and cDNA sequence analysis that ribosomes of maize (Zea mays L.) roots possess one form of P1, two types of P2, and a distinct P1/P2-like protein, which was designated P3 because of its ambiguous classification as P1 or P2. Here we examine the evolutionary relationship of P1 and P2 with respect to the P3. We provide phylogenetic evidence, using both distance- and parsimony-based methods, that P3 is a highly divergent, evolutionarily distinct P-protein present in plants and is apparently absent in other eukaryotes. These results demonstrate surprising evolutionary divergence of the P-protein complex, an integral component of the active site of ribosomes.

**EXPERIMENTAL PROCEDURES**

**Sequence Divergence and Phylogenetic Analyses.** P-protein homology searches were performed via the BLAST (27) search option available through Genetics Computer Group (GCG) (28). Four maize P1/P2 type proteins were aligned with 4 archaeabacterial L12 sequences and 29 additional eukaryotic P1 and P2 sequences (Table 1). Expressed sequence tags were used for rice P2a and P3 and for Arabidopsis P3. Protein alignments were generated using CLUSTAL W (49) and were adjusted manually upon visual inspection (alignment available upon request). Pairwise uncorrected distances were calculated using PAUP test ver. 4.0.455 (50). The aligned protein sequences were subjected to maximum parsimony and neighbor-joining analyses (p-distances) using PAUP. Neighbor-joining analysis using Kimura-corrected amino acid distances (51) were carried out using PHYLIP Version 3.5c (52). Parsimony analysis was conducted using the heuristic search algorithm with 50 random input orders and tree bisection-reconnection branch swapping. The reliability of the phyllogeny identified by neighbor joining and maximum parsimony was estimated using bootstrapping (53) with 500 replicates and 1 input order per replicate.

**Isolation of Genomic DNA and Southern Blot Analysis.** Genomic DNA was isolated from 5- to 6-day-old Z. mays L. (maize inbred B73 provided by Pioneer Hi-Bred, Johnson, IA) seedling roots using the CTAB extraction procedure (54). Twenty-five micrograms of genomic DNA were digested separately with BamHI, EcoRI, or HindIII (GIBCO/BRL), fractionated on a 0.8% agarose gel and transferred onto nylon membranes (MagnaGraph; Micron Separations, Westboro, MA) (55). Membranes were hybridized with [α-32P]dATP-labeled cDNAs (26) overnight at 42°C in 6× SSC, 5× Denhardt’s solution, 0.5% SDS (wt/vol), 100 μg/ml denatured calf thymus DNA, and 50% formamide (vol/vol) (55), washed twice in 2× SSC/0.1% SDS (wt/vol), once in 0.2× SSC/0.1% SDS (wt/vol) for 20 min each at 65°C, and exposed to autoradiographic film (Hyperfilm; Amersham) for 4 days at −80°C with an intensifying screen.

**RESULTS**

**Phylogenetic Analyses Identify an Evolutionarily Distinct P-Protein in Higher Plants.** Rat, yeast, and maize ribosomes possess P1 and P2, the eukaryotic L12 homologs. Rat ribosomes contain one form of P1 and one form of P2 (16), yeast possess two forms of P1 and P2 (37–39), whereas maize root ribosomes possesses one form of P1, two forms of P2 (P2a and P2b), and a distinct P-protein, P3 (26). Fig. 1 shows a representative alignment of the deduced amino acid sequences of archaeabacterial L12 and eukaryotic L12 homologs used in our analyses. Sequence conservation is highest in the hinge and C-terminal regions, but is maintained throughout the N-terminal region. Of particular interest is the P3 protein of maize ribosomes (26). P3 exhibits characteristics of eukaryotic P1 and P2 proteins; it has a predicted pI of 4.3, a calculated molecular mass of 12.2 kDa, and contains the conserved C-terminal and hinge regions. The N-terminal region of P3 is distinct yet aligns within the C-terminal region of the eukaryotic P1, P2, and archaeabacterial L12 proteins. BLAST (27) searches of the GenBank and Swiss-Prot databases failed to identify P3 homologs except in plant species, all of which were expressed sequence tags [full-length cDNAs: O. sativa (rice), D15754 and A. thaliana, Z18207; partial cDNAs: Brassica campestris (Chinese cabbage), L35823 and Ricinus communis (castor bean), T24312].

To examine the evolutionary relatedness of the plant 12 kDa P-proteins, pairwise uncorrected distances were calculated.
Table 2. Divergence between plant 12-kDa P-proteins

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Maize</th>
<th>Rice</th>
<th>Arabidopsis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize P1</td>
<td>0.619</td>
<td>0.677</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Maize P2a</td>
<td>0.657</td>
<td>0.667</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Maize P2b</td>
<td>0.677</td>
<td>0.667</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Maize P3</td>
<td>0.677</td>
<td>0.667</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Rice P2a</td>
<td>0.661</td>
<td>0.677</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Rice P2b</td>
<td>0.677</td>
<td>0.667</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Rice P3</td>
<td>0.661</td>
<td>0.677</td>
<td>0.714 0.794</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>0.661</td>
<td>0.677</td>
<td>0.714 0.794</td>
</tr>
</tbody>
</table>

*Uncorrected distances given in upper matrix, mean character differences in lower matrix.

Analysis of a 12-kDa P-Protein Gene Family Indicates That Nonconservative Amino Acid Replacements in the N-Terminal Region Are Permitted. Plant r-proteins are frequently encoded by multigene families (57). The level of complexity of the genes that encode maize P1, P2, and P3 (rpp1, rpp2a, rpp2b, and rpp3, respectively) was estimated by Southern blot analysis (Fig. 3) and copy number reconstruction analysis (data not shown). cDNAs rpp1 and rpp2b hybridized to single fragments, whereas rpp2a hybridized to three to four fragments and rpp3 hybridized to three to five fragments of maize genomic DNA. These

![Fig. 1. Alignment of a subset of archaeobacterial and eukaryotic 12 kDa P-proteins. The amino acid sequences represent archaeobacteria [Sulfolobus solfataricus (Sso), Halobacterium marismortui (Hma)], yeast [Saccharomyces cerevisiae (Sce)], rat (Rattus rattus (Rra)), and plants [Zea mays (Zma), Oriza sativa (Osa), Arabidopsis thaliana (Ath)]. Sequences were aligned using CLUSTAL W (49) and were adjusted manually upon visual inspection. Gaps were introduced to ensure maximum homology. Amino acids of conserved physicochemical similarity are shaded based on the following criteria: (i) conserved amino acids must occur in three of the four P-protein groups (i.e., L12, P1, P2, and P3) and (ii) conserved amino acids must occur in at least 75% of the sequences. The N termini, acidic hinge regions, and highly conserved C termini of the 12-kDa P-proteins are indicated.

![Fig. 2. Amino acid phylogenetic analysis of the 12 kDa P-proteins. A phylogenetic tree was generated by the neighbor-joining method using Kimura-corrected distances in PHYLIP version 3.5c based on the amino acid sequence of 33 eukaryotic 12 kDa P-proteins and 4 archaeobacterial L12 proteins. Branch lengths are proportional to the amino acid distances along each branch. Bootstrap values from 500 replicates are indicated. Parsimony bootstrap values for clades supported above the 50% level are indicated below branches, whereas neighbor-joining bootstrap values based on Kimura-corrected distances or uncorrected distances are indicated above the branch (corrected distances are in boldface italics).
The P2a gene family was analyzed further to gain insight into the amino acid sequence variation of a P-protein. cDNAs that encode four P2a gene family members were characterized and designated rpp2a-1, rpp2a-2, rpp2a-3, and rpp2a-4. Fig. 4 presents an alignment of the deduced amino acid sequences of four maize P2a isoforms. The carboxyl terminus of the deduced peptides is thoroughly conserved and amino acid variation is limited to the amino-terminal and hinge regions. The amino acid substitution of Asp in P2a-1 for Glu in P2a-2, P2a-3, and P2a-4 at position 21 is conservative. Nonconservative substitutions include Ala for Thr in P2a-2 at position 88 and Ala for Ser in P2a-3 at position 16, in P2a-4 at position 25, and in P2a-2 at position 79.

**DISCUSSION**

The P-protein complex of eukaryotic ribosomes is composed of acidic phosphoproteins and assists in the late initiation and elongation phases of translation via interactions with tRNA, mRNA, and translation factors (8, 9). The complex is composed of r-proteins P0, P1, and P2 in a number of eukaryotic species. We recently demonstrated (26) that ribosomes of maize roots contain P0, one type of P1, two types of P2, and a distinct P1/P2-like protein, designated P3. Analysis of P-proteins in numerous eukaryotic species including yeast, rats, trypanosomes, and others (Table 1) has failed to identify a P3 type protein in these organisms. Completion of sequencing of trypanosomes, and others (Table 1) has failed to identify a P3 type protein in these organisms. Nonetheless, conclusions about universality of P3 within the plant kingdom await the investigation of the presence of the P3 gene in plants such as conifers, cycads, ferns, and mosses. To the best of our knowledge, P3 is the first plant ribosomal protein identified that has no counterpart in yeast or rat ribosomes.

To gain insight into the evolutionary relationship of the P-protein family of r-proteins, we applied molecular phylogenetic techniques that can be used to infer the evolutionary history of genes. Evolutionary relatedness between P1 and P2 type proteins was reported (24, 25). Alignment of P1 and P2 with the colinear archaeobacterial homolog, L12, indicated that P1 and P2 form distinct, monophyletic groups (25). It was suggested that eukaryotic P1 and P2 arose from duplication and subsequent divergence of an ancestral form of L12 that occurred very early in the eukaryotic lineage (21, 25). The analyses presented here confirm the established P1 and P2 groupings, and separate P3 of plants into a distinct monophyletic group. Taxonomic representation in two of the three groups (P1 and P2) is extensive and includes proteins from protozoans, fungi, plants, and animals. In contrast, the P3 group includes only proteins of plant species. The present phylogenetic analyses clearly identify three distinct monophyletic groups of 12 kDa P-proteins in plants, whereas in animals, fungi, and most protozoans there are only two distinct monophyletic groups. We propose that in plants, the P-protein family is composed of three evolutionarily distinct subfamilies, P1, P2, and P3.

The finding of evolutionarily distinct components of the P-protein complex is not limited to plants. An unusual P1 protein in the ciliate *Tetrahymena* was an additional exception to the previous finding of P1 and P2 subgroups in eukaryotes. Hansen et al. (40) characterized the *Tetrahymena* P1 gene and found that the divergence between the *Tetrahymena* P1 and other P1 type proteins is as least as great as the divergence between *Trypanosoma* P2 and other P2 type proteins. This is inconsistent with the phylogenetic tree generated with rRNA sequences, where the branching of trypanosomes from the eukaryotic lineage predates that of the ciliates (58). Further evidence of the lack of conservation of the eukaryotic P-proteins was revealed when mutant strains of yeast that lack functional P1 and P2 genes were transformed with *Dictyostelium* P1 and P2 genes. The fungus-like protozoan P1 and P2 were synthesized, but were unable to assemble into yeast ribosomes and derepress the slow growth mutant phenotype (59).

A model for the evolution of the P-proteins was proposed by Shimmin et al. (21). They predicted that the common primordial ancestor of eukaryotes had a single L12 gene. Duplication of L12, possibly to ensure elevated stoichiometry in the ribosome, and its further divergence, led to the P1 and P2 genes present in contemporary eukaryotes. Our identification of the evolutionarily distinct P3 protein in plants requires an expansion of the Shimmin model by a minimum of one additional gene duplication event. The duplication event that produced P3 might have occurred early within the eukaryotic lineage, prior to the separation of protozoans, fungi, plants, and

![Fig. 4. Maize P2a deduced peptide sequence. Deduced peptide sequences of the maize P2a family members were aligned using the PILEUP alignment program (28). Amino acids that differ from P2a-1 are indicated. Gaps were introduced to ensure maximum homology.](image-url)
FIG. 5. Alternative models depicting the evolution of the plant P3 protein. Model 1. Duplication and divergence of the ancestral L12-like gene occurred very early in the eukaryotic lineage resulting in P1, P2, and P3 type proteins in ancestral eukaryotes. Contemporary plants have retained the P3 gene, whereas the specific loss of P3 from other contemporary eukaryotes is necessary to explain the absence of P3 in these species. Model 2. Duplication and divergence of the ancestral L12-like gene occurred very early in the eukaryotic lineage to produce P1 and P2 type P-proteins in ancestral eukaryotes. Further duplication and divergence within the plant lineage produced the P3 seen in modern plants.

An alternative hypothesis is that the duplication event that gave rise to P3 occurred strictly within the plant lineage (Fig. 5, Model 2). Although forcing topological constraints such as P3 with the plant P1 group or P3 with the plant P2 group both require additional steps (nine and four, respectively) on the parsimony tree, Templeton (60) (approximately 0.2371 and 0.6185, respectively) and Kishino and Hasegawa (61) (P = 0.2371 and P = 0.6185, respectively) and Kishino and Hasegawa (61) (P = 0.1929 and P = 0.6013, respectively) tests indicate that this is not a significant increase in tree length (P values for the P1/P3 constraint are mean values from 21 equally most parsimonious trees and P values for the P2/P3 constraint are mean values from 15 equally most parsimonious trees). Hence, we cannot reject this hypothesis. If this hypothesis is true, it would appear that the P3 group has an ancient origin within the plant lineage. The presence of P3 in both monocotyledonous (Oryza and Zea) and dicotyledonous plant species (Arabidopsis, Brassica, and Ricinus) indicates that the duplication event that gave rise to P3 is at least as old as the divergence of monocots and dicots, approximately 200 million years (62). Furthermore, the mean sequence divergence separating P2 and P3 of plants (1.442; divergences are Kimura-corrected amino acid distances) is approximately eight-fold greater than the mean within-group divergence separating rice and maize (0.178) and three-fold greater than the mean within-group divergence separating monocots and dicots (0.516). Assuming a molecular clock (63) calibrated with the estimated time of rice/maize divergence (approximately 70 million years ago) (64) and monocot/dicot divergence coupled with the Kimura-corrected amino acid distance between these groups, plant P2 and P3 diverged approximately 560 million years ago. This predicted time of divergence between plant P2 and P3 precedes the predicted time of origin of terrestrial plants from an aquatic algal ancestor (approximately 425 million years ago) (65). Whichever hypothesis holds true, the duplication event in a progenitor P-protein gene that gave rise to P3 may have been followed by a period of rapid accumulation of amino acid replacements. These replacements could have been fixed as a result of positive Darwinian selection favoring the amino acid changes that serve to adapt this duplicated gene to a unique or distinct function. Evidence for the tolerance of substitutions in amino acid sequence following gene duplication of a 12 kDa P-protein was demonstrated by examination of the deduced polypeptide sequence of maize P2a isoforms (Fig. 4). Variation among the deduced P2a polypeptides was limited to the N terminus and included several nonconservative substitutions. No amino acid substitutions were observed within the C terminus, which is implicated in interactions with elongation factor-2 (66).

Different forms of 12 kDa P-proteins may have relevant functional distinctions. Two isoforms of P1 and P2 are present in both Baker’s and fission yeasts (37–39), whereas two isoforms of P2 (P2a and P2b) are also present in maize (26). The presence of isoforms may suggest functional redundancy, although data from various groups have suggested the contrary. Yeast mutant strains in which one or two of the genes encoding P1 and P2 were disrupted exhibit decreased cell growth rates (67). Double disruptant yeast mutants show various levels of decreased cell growth rates, but the most notable decrease was observed when both P1 or P2 genes were disrupted (68). Examination of cell doubling time in triple mutants indicate further heterologous roles of individual isoforms of P1 and P2 proteins (69).

Our data provide evidence that the lateral stalk structure of the large ribosomal subunit has evolved into a distinct complex in animals, yeast, protozoans, and plants despite its position within the active site of the ribosome, where structural and functional conservation would be expected. The significance of evolutionary distinctions in the stalk structure of the ribosome remains to be elucidated. Perhaps extreme environmental conditions led to specific selective pressures that imposed different structural and functional constraints on the P-proteins of plants. At this time we do not have any information on the individual functions of P1, P2a, P2b, and the P3 plant proteins. Our goals include the examination and quantification of the 12 kDa P-proteins in maize in response to environmental stress conditions and during development. The unique aspects of the plant P-protein complex, such as the presence of the P3 protein, modulation in protein composition or phosphorylation, may provide plants with an additional mechanism for the regulation of protein synthesis.

We thank Dr. S. Manjunath, Dr. Sanjay Jayachandran, Alan Williams, and Virginia Oberholzer for critical review of this manuscript; Sasipriya Vangala for many helpful comments and discussions; and Dr. David Swafford for permission to use Faup Version 4.0d55. K.S. thanks the University of California, Riverside Interdepartmental Graduate Program in Genetics for funding. M.S.S. thanks the Alfred P. Sloan Foundation. This research was supported by Grant MCB-9220192 from the National Science Foundation to J.B.S.