Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 2378–2383, March 1998
Evolution

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 the 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 are 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). In *Escherichia coli*, the lateral stalk is a pentameric complex that contains one molecule of r-protein L10 (17 kDa) and two heterodimers of r-proteins L7 (12 kDa) and L12 (12 kDa) (10). The L7/L12 dimers attach to L10 through their amino termini, and L10 interacts with the 23S rRNA of the large subunit within the rRNA GTPase domain (11). The stalk assists in elongation factor-G binding and is required for GTPase activity during translocation (8). Analyses by proton nuclear magnetic resonance and attachment of fluorescent probes to L7/L12 within the ribosome demonstrated that the stalk is a mobile component of the ribosome (12–14), which may facilitate movement of mRNA and the removal of deacylated tRNA during elongation (8, 9). L7 and L12 are encoded by a single gene, but the proteins differ in that the N-terminal Ser of L7 is posttranslationally aminoacetylated (15); hence, these proteins of prokaryotic ribosomes are referred to as L12.

In eukaryotes, the L10 homolog is P0 (34 kDa) and the L12 homologs are P1 (11–12 kDa) and P2 (11–12 kDa) (16). These acidic ribosomal proteins are called P-proteins because they are phosphorylated. P0, P1, and P2 have distinct amino-terminal regions, a stretch of acidic residues followed by a highly conserved carboxyl-terminal region. The organization of the eukaryotic P-proteins into the stalk structure is thought to be similar to that of prokaryotes, with the complex formed by two P1 proteins, two P2 proteins, and one P0 protein (17). Analogous to the prokaryotic stalk structure, the eukaryotic P-protein complex is required for elongation factor-2 binding and GTP hydrolysis (8, 9). P1 and P2 are unusual r-proteins. They do not assemble onto preribosomes in the nucleolus, but cycle between ribosomes and a cytosolic pool, with only the phosphorylated forms found associated with ribosomes (18). In addition, they are the only r-proteins that occur in multiple copies per ribosome (19).

L12 homologs fall into two categories based on structural similarities, one that includes archaebacterial L12 and eukaryotic 12 kDa proteins and another that includes eubacterial, mitochondrial, and chloroplastic L12 proteins. An interkingdom alignment of L12 homologs indicated that the archaebacterial L12 and the eukaryotic P1 and P2 proteins are colinear, whereas the eubacterial L12 protein has undergone rearrangements, possibly through gene fusion events (20, 21). Mitochondrial and chloroplastic L12 proteins are structural homologs of eubacterial L12 (22, 23) and are therefore distinct from the cytosolic 12 kDa P-proteins. Phylogenetic analyses (24, 25) support the hypothesis that a single ancestral L12 gene may have duplicated and subsequently diverged very early in the eukaryotic lineage to produce the P1 and P2 proteins found in contemporary eukaryotes (20, 21).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

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 archaebacterial 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. 4.0d55 (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 phylogenies identified by neighbor joining and maximum parsimony was estimated using bootstrapping (53) with 500 replicates and 1 input order per replicate.

Table 1. The 12-kDa P-protein sequences used in evolutionary analyses

Source/organism	Common name	Protein	Accession no.	Ref.
Halobacterium halobium	Archaebacterium	L12	X06736	29
Halobacterium marismortui	Archaebacterium	L12	X51430	30
Haloferax volcanii	Archaebacterium	L12	X58924	31
Sulfolobus solfataricus	Archaebacterium	L12	X59038	32
Artemia salina	Brine shrimp	P1	X02633	33
Chlamydomonas reinhartii	Green algae	P1	X66411	*
Dictyostelium discoideum	Slime mold	P1	X56193	34
Drosophila melanogaster	Fruit fly	P1	Y00504	35
Homo sapiens	Human	P1	M17886	36
Rattus rattus	Rat	P1	X15097	16
Saccharomyces cerevisiae	Baker's yeast	P1α	M26504	37
		P1β	M26507	38
Schizosaccharomyces pombe	Fission yeast	A1	M33137	39
		A3	M33139	39
Tetrahymena thermophila	Ciliate	P1	M59428	40
Trypanosoma cruzi	Trypanosome	P1	X65025	41
Zea mays	Maize	P1	U62752	26
Artemia salina	Brine shrimp	P2	X02632	33
Cladosporium herberum	Mold	P2	X77253	42
Dictyostelium discoideum	Slime mold	P2	X56192	43
Drosophila melanogaster	Fruit fly	P2	X05016	44
Homo sapiens	Human	P2	M17887	36
Leishmania infantum	Trypanosome	P2-1	X68015	45
		P2-2	X68016	45
Oryza sativa	Rice	P2a	D16065	*
Rattus rattus	Rat	P2	X15098	16
Saccharomyces cerevisiae	Baker's yeast	P2α	M26503	38
		P2β	M26505	38
Schizosaccharomyces pombe	Fission yeast	A2	M33138	39
		A4	M33142	39
Trypanosoma cruzi	Trypanosome	P2α	X65065	46
		P2β	X52323	47
Zea mays	Maize	P2a	U29383	48
		P2b	U62753	26
Arabidopsis thaliana	Arabidopsis	P3	Z18207	*
Oryza sativa	Rice	P3	D15754	*
Zea mays	Maize	P3	U62751	26

*Direct GenBank submission.

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 *Bam*HI, *Eco*RI, or *Hin*dIII (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-^{32}P]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.

rpp2a cDNA Isolation and Sequencing. A cDNA library [kindly provided by Bruce Veit (Massey University, New Zealand); ref. 56] from maize immature ear mRNA (of the inbred B73) in the λZAP vector (Stratagene) was screened with a maize P2a cDNA (GenBank accession no. T18290; ref. 48) using standard recombinant DNA procedures. Clones were sequenced by dideoxy chain termination using Sequenase Version 2.0 (United States Biochemicals) or the cycle-sequencing method using *Taq* polymerase (Promega–Fisher) according to manufacturers' protocols, using commercially available or custom primers. Deduced polypeptide sequences were determined using GCG (28).

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 archaebacterial 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 with the N termini of the eukaryotic P1, P2, and archaebacterial 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.

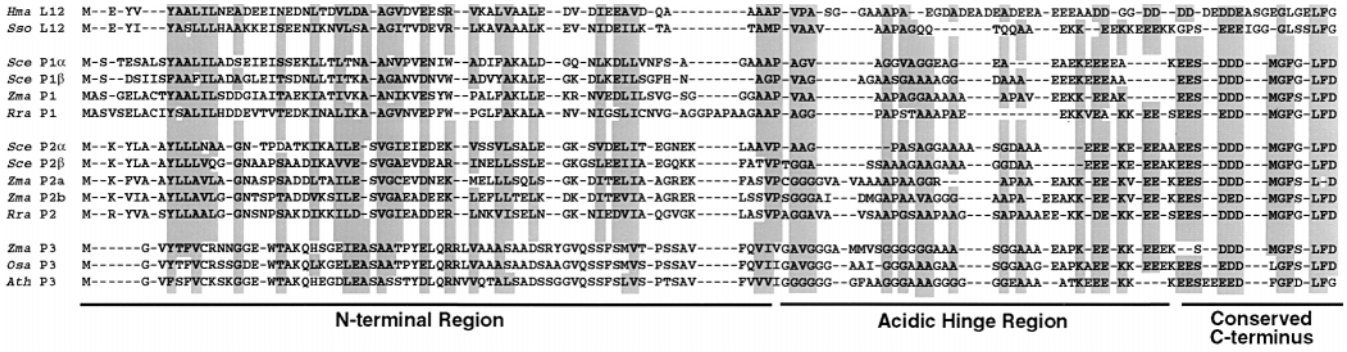

FIG. 1. Alignment of a subset of archaebacterial and eukaryotic 12 kDa P-proteins. The amino acid sequences represent archaebacteria [*Sulfolobus solfataricus* (*Sso*), *Halobacterium marismortui* (*Hma*)], yeast [*Saccharomyces cerevisiae* (*Sce*)], rat (*Rattus rattus* (*Rra*)], and plants [*Zea mays* (*Zma*), *Oryza 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.

Table 2 illustrates that P3 is distinct from P1 and P2. Within-group distances [maize P2a to rice P2a (0.145) and maize P3 to rice P3 (0.171)] are substantially less than between-group distances [maize P1 to maize P3 (0.677), maize P1 to maize P2a (0.619), and maize P2a to maize P3 (0.667)]. Even though taxonomic representation for the P1 and P2 groups is extensive, mean within-group distances calculated for the eukaryotic P1 (0.457), P2 (0.459), and P3 (0.306) proteins surveyed (Table 1) are less than distances between plant P-protein groups, further substantiating P3 as a distinct protein.

To study further the evolutionary relatedness of archaebacterial L12 and eukaryotic 12 kDa P-proteins we used maximum parsimony and neighbor-joining methods of phylogenetic analyses. Neighbor-joining analysis (using Kimura-corrected distances), based on 37 amino acid sequences, resulted in the tree shown in Fig. 2. Maximum parsimony analysis resulted in 18 1,041-step most parsimonious trees. Aside from the placement of the *Tetrahymena* P1, the eukaryotic 12 kDa P-proteins are separated into three distinct clusters corresponding to the P-protein "types," designated as P1, P2, and P3. Both the P1 and P2 groups include proteins from protozoans, fungi, plants, and animals, yet composition of the strongly supported P3 group is restricted to plant proteins (parsimony and uncorrected distance bootstrap % = 100; Kimura-corrected distance bootstrap % = 71). Both parsimony and neighbor-joining analyses suggest an association of the P2 and P3 proteins to the exclusion of the P1 proteins; however, parsimony bootstrap support for this arrangement is only 66%. The analyses group P1 and P2 isoforms of the yeasts *S. cerevisiae* (P1α and P1β) and *Sch. pombe* (A1 and A3; A2 and A4) within their respective clades. In addition, the groupings of P2 isoforms of the trypanosomes *L. infantum* (P2–1 and P2–2) and *T. cruzi* (P2α and P2β) and of maize *Z. mays* (P2a and P2b) suggest additional recent duplication events within these lineages.

Table 2. Divergence between plant 12-kDa P-proteins

	Maize				Rice		*Arabidopsis*
Proteins	P1	P2a	P2b	P3	P2a	P3	P3
Maize P1		0.619	0.657	0.677	0.614	0.714	0.794
Maize P2a	60		0.287	0.667	0.145	0.657	0.650
Maize P2b	65	31		0.736	0.333	0.692	0.676
Maize P3	67	70	78		0.661	0.171	0.402
Rice P2a	62	16	37	72		0.661	0.636
Rice P3	70	69	74	20	72		0.345
Arabidopsis P3	77	67	71	46	68	40	

*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

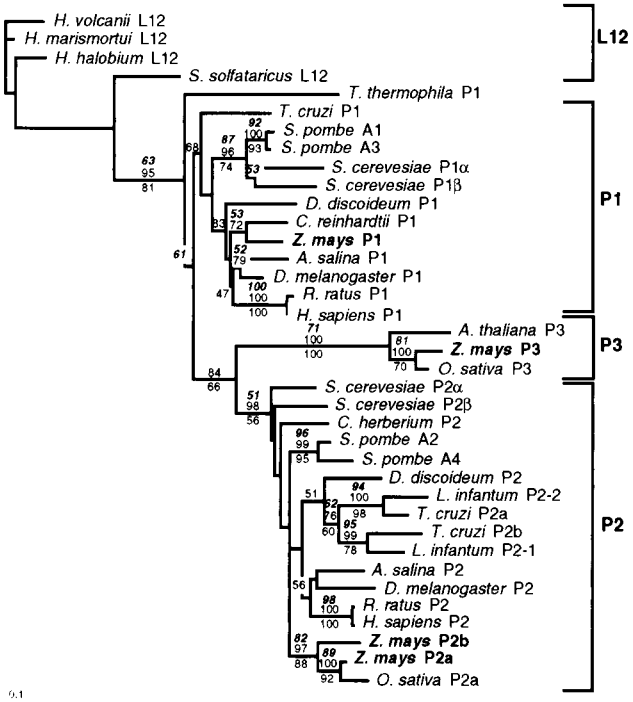

FIG. 2. Amino acid phylogenetic analysis of the 12 kDa P-proteins. A phylogenetic tree was generated by the neighbor-joining method using Kimura-corrected distances in PHYLIP version 3.5c based on the amino acid sequence of 33 eukaryotic 12 kDa P-proteins and 4 archaebacterial L12 proteins. Branch lengths 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).

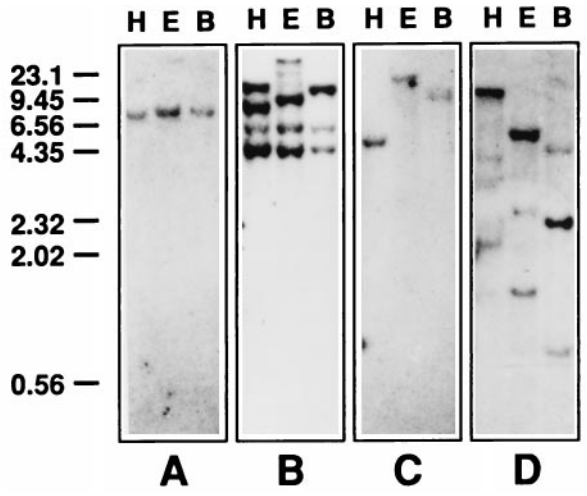

FIG. 3. Maize genomic DNA Southern blot analysis. Genomic DNA digested with *Bam*HI (B), *Eco*RI (E), or *Hin*dIII (H) was separated, blotted, and hybridized at high stringency with the [^{32}P]-labeled maize cDNAs encoding P1, P2a, P2b, and P3, respectively. (*A*) *rpp1*, (*B*) *rpp2a-1*, (*C*) *rpp2b*, (*D*) *rpp3*. Migration of DNA standards (in kilobases) are indicated.

data indicate that P1 and P2b are each encoded by a single gene, P3 is encoded by at least two genes, and P2a is encoded by at least four genes in maize.

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 the yeast (*S. cerevisiae*) genome did not reveal a gene that encodes a P3 homolog, indicating that P3 is absent from this organism. Our extensive DNA and protein database search identified P3 in monocotyledonous (*Oryza* and *Zea*) and dicotyledonous (*Arabidopsis, Brassica*, and *Ricinus*) plant species. 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 archaebacterial 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* is 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 *Trypanasoma* 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

```
          10        20        30        40        50        60        70        80        90        100       110       120
P2a-1  MKFVAAYLLAVLAGNASPSADDLTAILESVGCEVDNEKMELLLSQLSGKDITELIAAGREKFASVPC-GGGGVAVAAAAPAAGGRAPAAEAKKEEKVEEKEESDDDMGFSLFD
P2a-2  ....................E.............R.............................-..........S.........T.....................
P2a-3  .......S....E.....................................................G....M.....G.....................................
P2a-4  ..E...S.....................................................................G.....................................
```

N-terminal Region — Acidic Hinge Region — Conserved C-terminus

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.

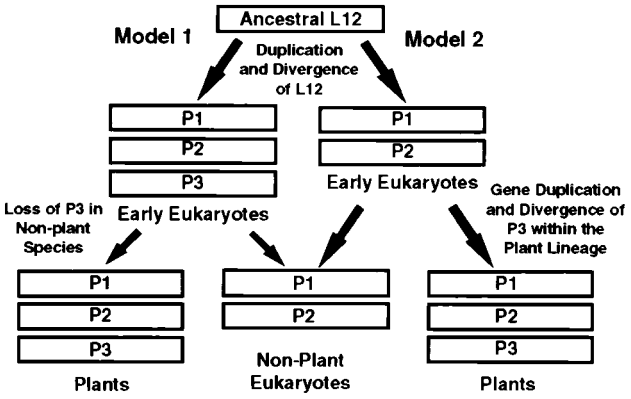

FIG. 5. Alternative models depicting the evolution of the plant P3 protein. Model 1. Duplication and divergence of the ancestral L12-like gene occurred very early in the eukaryotic lineage resulting in P1, P2, and P3 type proteins in ancestral eukaryotes. Contemporary plants have retained the P3 gene, 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.

animals (Fig. 5, Model 1). Several lines of evidence support such an ancient origin for the P3 protein: (*i*) sequence divergence separating P3 from both P1 and P2 is comparable to the divergence separating P1 and P2 (see Table 2); (*ii*) there is a lack of resolution of the base of the P-protein tree, that is, the exact branching order of the eukaryotic 12 kDa P-proteins is unclear; and (*iii*) a closer examination of the P1 and P2 subgroups reveals that within each subgroup there are distinct clusters that correspond to animals, plants, fungi, and protozoans, strengthening the notion that P-protein divergence occurred very early in the eukaryotic lineage. Fig. 2 supports this model. If such an ancient duplication event occurred, then the absence of P3 in other eukaryotes must be explained by the specific loss of P3 in these species.

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) ($P = 0.2371$ 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 moncots 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 moncot/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 yeast (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 Swofford for permission to use PAUP 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.

1. Nygård, O. & Nilsson, L. (1990) *Eur. J. Biochem.* **191,** 1–17.
2. Hershey, J. W. B. (1991) *Annu. Rev. Biochem.* **69,** 717–755.
3. Merrick, W. C. & Hershey, J. W. B. (1995) *Translational Control* (Cold Spring Harbor Lab. Press, Plainview, NY).
4. Wool, I. G., Chan, Y.-L. & Gluck, A. (1995) *Biochem. Cell. Biol.* **73,** 933–947.
5. Warner, J. R. (1989) *Microbiol. Rev.* **53,** 256–271.
6. Bailey-Serres, J. (1998) in *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants.* eds.

Bailey-Serres, J. & Gallie, D. R. (Am. Soc. Plant Physiologists, Rockville, MD), pp. 125–144.
7. Bielka, H. (1982) *The Eukaryotic Ribosome* (Springer, Berlin).
8. Möller, W. (1990) *The Ribosome: Structure, Function and Evolution* (Am. Soc. Microbiol, Washington, DC), pp. 380–389.
9. Liljas, A. (1991) *Int. Rev. Cytol.* **124,** 103–135.
10. Petersson, I., Hardy, S. J. S. & Liljas, A. (1976) *FEBS Lett.* **64,** 135–138.
11. Gudkov, A. T., Tumanova, L. G., Gondgaze, G. M. & Bushev, V. N. (1980) *FEBS Lett.* **109,** 34–38.
12. Cowgill, C., Nichols, B. G., Kenny, J. W., Butler, P., Bradbury, E. M. & Traut, R. R. (1984) *J. Biol. Chem.* **259,** 15257–15263.
13. Traut, R. R., Dey, D., Bochkariov, D. E., Oleinikov, A. V., Jokhadze, G. G., Hamman, B. & Jameson, D. (1995) *Biochem. Cell. Biol.* **73,** 949–958.
14. Hamman, B. D., Oleinikov, A. V., Jokhadze, G. G., Bochkariov, D. E., Traut, R. R. & Jameson, D. M. (1996) *J. Biol. Chem.* **271,** 7568–7573.
15. Terhorst, C., Möller, W., Laursen, R. & Wittmann-Liebold, B. (1973) *Eur. J. Biochem.* **34,** 138–152.
16. Wool, I. G., Chan, Y. L., Gluck, A. & Suzuki, K. (1991) *Biochimie* **73,** 861–870.
17. Uchiumi, T., Wahba, A. J. & Traut, R. R. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5580–5584.
18. Zinker, S. & Warner, J. R. (1976) *J. Biol. Chem.* **251,** 1799–1807.
19. Hardy, S. J. S. (1975) *Mol. Gen. Genet.* **140,** 253–274.
20. Ramirez, C., Shimmin, L. C., Newton, C. H., Matheson, A. T. & Dennis, P. P. (1989) *Can. J. Microbiol.* **35,** 234–244.
21. Shimmin, L. C., Ramirez, C., Matheson, A. T. & Dennis, P. P. (1989) *J. Mol. Evol.* **29,** 448–462.
22. Leijonmarck, M., Liljas, A. & Subramanian, A. R. (1984) *Biochem. Int.* **8,** 69–76.
23. Marty, L. & Fort, P. (1996) *J. Biol. Chem.* **271,** 11468–11476.
24. Goddemeier, M. L., Rensing, S. A. & Feix, G. (1996) *Plant Mol. Biol.* **30,** 655–658.
25. Liao, D. & Dennis, P. P. (1994) *J. Mol. Evol.* **38,** 405–419.
26. Bailey-Serres, J., Vangala, S., Szick, K. & Lee, C.-H. K. (1997) *Plant Physiol.* **114,** 1293–1305.
27. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
28. Genetics Computer Group (1994) *Program Manual for the Wisconsin Package, Version 8* (Genetics Computer Group, Madison, WI).
29. Itoh, T., Kumazaki, T., Sugiyama, M. & Otaka, E. (1988) *Biochim. Biophys. Acta* **949,** 110–118.
30. Arndt, E. & Weigel, C. (1990) *Nucleic Acids Res.* **18,** 1285.
31. Shimmin, L. C. & Dennis, P. P. (1996) *J. Bacteriol.* **178,** 4737–4741.
32. Ramirez, C., Shimmin, L. C., Leggatt, P. & Matheson, A. T. (1994) *J. Mol. Biol.* **244,** 242–249.
33. Maassen, J. A., Schop, E. N., Brands, J. H. G. M., Van Hemert, F. J., Lenstra, J. A. & Möller, W. (1985) *Eur. J. Biochem.* **149,** 609–616.
34. Prieto, J., Candel, E. & Coloma, A. (1991) *Nucleic Acids Res.* **19,** 1340.
35. Wigboldus, J. D. (1987) *Nucleic Acids Res.* **15,** 10064.
36. Rich, B. E. & Steitz, J. A. (1987) *Mol. Cell. Biol.* **7,** 4065–4074.
37. Mitsui, K., Nakagawa, T. & Tsurugi, K. (1988) *J. Biochem.* **104,** 908–911.
38. Remacha, M., Saenz-Robles, M. T., Vilella, M. D. & Ballesta, J. P. G. (1988) *J. Biol. Chem.* **263,** 9094–9101.
39. Beltrame, M. & Bianchi, M. E. (1990) *Mol. Cell. Biol.* **10,** 2341–2348.
40. Hansen, T. S., Andreason, P. H., Dreisig, H., Hojrup, P., Nielsen, H., Engberg, J. & Kristiansen, K. (1991) *Gene* **105,** 143–150.
41. Vazquez, M. P., Schijman, A. G. & Levin, M. J. (1992) *Nucleic Acids Res.* **20,** 2599.
42. Zhang, L., Muradia, G., Curran, I. H., Rode, H. & Vijay, H. M. (1995) *J. Immunol.* **154,** 710–717.
43. Prieto, J., Candel, E. & Coloma, A. (1991) *Nucleic Acids Res.* **19,** 1341.
44. Qian, S., Zhang, J. Y., Kay, M. A. & Jacobs-Lorena, M. (1987) *Nucleic Acids Res.* **15,** 987–1003.
45. Soto, M., Requena, J. M., Garcia, M., Gomez, L. C., Navarrete, I. & Alonso, C. (1993) *J. Biol. Chem.* **268,** 21835–21843.
46. Vazquez, M. P., Schijman, A. G., Panebra, A. & Levin, M. J. (1992) *Nucleic Acids Res.* **20,** 2893.
47. Schijman, A. G., Dusetti, N. J., Vazquez, M. P., Lafon, S., Leuy-Yeyati, P. & Levin, M. J. (1990) *Nucleic Acids Res.* **18,** 3399.
48. Vangala, S. & Bailey-Serres, J. (1995) *Plant Physiol.* **109,** 721.
49. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
50. Swofford, D. L. (1997) *Phylogenetic Analysis Using Parsimony (*PAUP) Test Version 4.0d55 (Ill. Natural History Survey, Champaign).
51. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, MA).
52. Felsenstein, J. (1993) *Phylogenetic Inference Package Version 3.5c* (Univ. of Washington, Seattle).
53. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
54. Boyce, T. M., Zwick, M. E. & Aquardo, C. F. (1990) *Genetics* **123,** 825–836.
55. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
56. Schmidt, R. J., Veit, B., Mandel, M. A., Mena, M., Hake, S. & Yanofsky, M. (1993) *Plant Cell* **5,** 729–737.
57. Browning, K. S. (1997) *Plant Mol. Biol.* **32,** 107–144.
58. Sogin, M. L., Elwood, H. J. & Gunderson, J. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1386–1387.
59. Bermejo, B., Prieto, J., Remacha, M., Coloma, A. & Ballesta, J. P. G. (1995) *Biochim. Biophys. Acta* **1263,** 45–52.
60. Templeton, A. R. (1983) *Evolution* **37,** 221–244.
61. Kishino, H. & Hasegawa, M. (1989) *J. Mol. Evol.* **29,** 170–179.
62. Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M. & Li, W. H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6201–6205.
63. Zuckeland, E. & Pauling, L. (1965) in *Evolving Genes and Proteins*, Bryson, V. & Vogel, H. J., eds. (Academic, New York).
64. Clark, L. G., Zhang, W. & Wendel, J. F. (1995) *Syst. Bot.* **20,** 436–460.
65. Graham, L. E. (1993) *The Origin of Land Plants* (Wiley, New York).
66. Köpke, A. K. E., Leggatt, P. A. & Matheson, A. T. (1992) *J. Biol. Chem.* **267,** 1382–1390.
67. Remacha, M., Santos, C. & Ballesta, J. P. G. (1990) *Mol. Cell. Biol.* **10,** 2182–2190.
68. Remacha, M., Santos, C., Bermejo, B., Naranda, T. & Ballesta, J. P. G. (1992) *J. Biol. Chem.* **267,** 12061–12067.
69. Remacha, M., Jimenez-Diaz, A., Bermejo, B., Rodriguez-Gabriel, M. A., Guarinos, E. & Ballesta, J. P. G. (1995) *Mol. Cell. Biol.* **15,** 4754–4762.