Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 14136–14141, November 1998
Biochemistry

A functional homolog of a yeast tRNA splicing enzyme is conserved in higher eukaryotes and in *Escherichia coli*

SHERRY L. SPINELLI*, HARMIT S. MALIK†, SANDRA A. CONSAUL*, AND ERIC M. PHIZICKY*‡

*Department of Biochemistry and Biophysics, School of Medicine and Dentistry; and †Department of Biology, University of Rochester, Rochester, NY 14642

Communicated by Fred Sherman, University of Rochester School of Medicine and Dentistry, Rochester, NY, October 6, 1998 (received for review September 1, 1998)

ABSTRACT **tRNA splicing in the yeast *Saccharomyces cerevisiae* requires an endonuclease to excise the intron, tRNA ligase to join the tRNA half-molecules, and 2′-phosphotransferase to transfer the splice junction 2′-phosphate from ligated tRNA to NAD, producing ADP ribose 1′′–2′′ cyclic phosphate (Appr>p). We show here that functional 2′-phosphotransferases are found throughout eukaryotes, occurring in two widely divergent yeasts (*Candida albicans* and *Schizosaccharomyces pombe*), a plant (*Arabidopsis thaliana*), and mammals (*Mus musculus*); this finding is consistent with a role for the enzyme, acting in concert with ligase, to splice tRNA or other RNA molecules. Surprisingly, functional 2′-phosphotransferase is found also in the bacterium *Escherichia coli*, which does not have any known introns of this class, and does not appear to have a ligase that generates junctions with a 2′-phosphate. Analysis of the database shows that likely members of the 2′-phosphotransferase family are found also in one other bacterium (*Pseudomonas aeruginosa*) and two archaeal species (*Archaeoglobus fulgidus* and *Pyrococcus horikoshii*). Phylogenetic analysis reveals no evidence for recent horizontal transfer of the 2′-phosphotransferase into Eubacteria, suggesting that the 2′-phosphotransferase has been present there since close to the time that the three kingdoms diverged. Although 2′-phosphotransferase is not present in all Eubacteria, and a gene disruption experiment demonstrates that the protein is not essential in *E. coli*, the continued presence of 2′-phosphotransferase in Eubacteria over large evolutionary times argues for an important role for the protein.**

tRNA splicing is ubiquitous, occurring in all three major kingdoms. In Eubacteria, tRNA introns are all self-splicing group I or group II introns, which effect splicing by two RNA-catalyzed transesterification reactions mediated by the well structured intron (1–3). By contrast, tRNA splicing in Eukarya and Archaea is enzyme catalyzed (4, 5). In these kingdoms, tRNA introns occur as small insertions, which are invariably located one base 3′ of the anticodon in eukaryotes (6), and in slightly variable locations in archaea (7). Splicing is initiated in both kingdoms by an endonuclease that excises the intron to yield half-molecules with ends containing a 2′-3′ cyclic phosphate and a 5′-OH (8, 9). Although details of pre-tRNA substrate recognition differ in different organisms (10–12), the endonucleases from both kingdoms are closely related phylogenetically (13–15).

In eukaryotes, two different mechanisms have been observed for the joining steps of tRNA splicing. In the yeast *Saccharomyces cerevisiae*, which is the best studied system, tRNA ligase joins the half-molecules to generate a splice junction bearing a 2′-phosphate (16), and 2′-phosphotransferase transfers the 2′-phosphate to NAD to form ADP ribose 1′′–2′′ cyclic phosphate (17, 18). Both yeast enzymes are known to be required for tRNA splicing, because conditional ligase (*rlg1*) mutants accumulate tRNA half-molecules (19), and conditional phosphotransferase (*tpt1*) mutants accumulate 2′-phosphorylated ligated tRNAs (20) as the cells stop growing in nonpermissive conditions. The enzymes in this pathway are conserved in eukaryotes: a similar ligase activity has been observed in extracts from wheat germ (21–24), *Chlamydomonas* (25), and HeLa cells (26), and a highly similar ligase gene has been isolated from the yeast *Candida albicans* (27). Moreover, 2′-phosphotransferase activity similar to that in yeast has been detected in both HeLa cell extracts (28) and microinjected *Xenopus* oocytes (18). These findings support a universal ligase/phosphotransferase pathway for the last steps of tRNA splicing in eukaryotes. However, vertebrates also appear to have a second and completely different ligase, which directly joins the 5′-OH and the cyclic phosphate ends of the half-molecules to generate a junction with a normal 3′-5′ phosphodiester bond but no extra 2′-phosphate (29, 30). This vertebrate ligase is active in microinjected *Xenopus* oocytes (31); furthermore, a similar ligase also may be present in the archaeal species *Haloferax volcanii* (32).

To further explore the conservation of the yeast-like ligase/phosphotransferase pathway in different organisms, we have sought genes whose products could substitute for the yeast enzyme in *tpt1* mutants that lacked 2′-phosphotransferase. We noted earlier that the amino acid sequence of the *TPT1* gene was somewhat conserved in an ORF from *Schizosaccharomyces pombe*, the C-terminal end of expressed sequence tags (ESTs) from mouse and rice, and an ORF in *Escherichia coli* (33). We report here that the *S. pombe*, mouse, and *E. coli TPT1* homologs are functional members of a 2′-phosphotransferase gene family, which also includes functional homologs in the distantly related yeast *C. albicans* and in a higher plant, *Arabidopsis thaliana*. Further searching of the database shows that members of this *TPT1* gene family are present in another bacterial species and in two archaeal species. The widespread occurrence of the 2′-phosphotransferase in eukaryotes confirms the conservation of the ligase/phosphotransferase pathway in eukaryotes. However, the presence of a functional eukaryotic tRNA splicing enzyme in bacteria, which are not known to splice tRNA by this pathway, is a puzzle. Phylogenetic analysis suggests that the eubacterial gene did not arise by recent horizontal transfer from either Eukarya or Archaea and may be over 3 billion years old. This result suggests that the bacterial gene is important, even if there is no obvious splicing of this class in bacteria.

MATERIALS AND METHODS

Strains Libraries, Plasmids, and DNAs. SC974 (*MATα* ura3–52, *leu2-3, 112*, *ade2-101*, *his3*-Δ200 trp1-Δ901

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.

© 1998 by The National Academy of Sciences 0027-8424/98/9514136-6$2.00/0
PNAS is available online at www.pnas.org.

Abbreviations: CIP, calf intestinal phosphatase; EST, expressed sequence tag.

‡To whom reprint requests should be addressed. email: Eric_Phizicky@urmc.rochester.edu.

tpt1-Δ1::LEU2 lys2$^-$ *cyh2*^{r} [pEMP1135–*CEN TRP1 CYH2 TPT1*]) was derived from SC814 (*MATα ura3-52*, *leu2–3, 112*, *ade2-101*, *his3-Δ200 trp1-Δ901 tpt1-Δ1*::*LEU2*, *lys2*$^-$ [pGMC1 (URA3, *CEN IV*, *TPT1*]) by selection for a *cyh2*^{r} derivative on yeast extract/peptone/dextrose plates containing 10 μg/ml cycloheximide, followed by transformation of the strain with pEMP1135 and selection against the *URA3* plasmid with 5-fluoroorotic acid. SC839 was derived from SC814 by transformation of pEMP 1062 [*CEN IV TRP1. LYS2, TPT1*] (20), followed by selection against the *URA3* plasmid. pEMP1135 was made by cloning the *Eco*RI–*Sac*I *TPT1* fragment into a *CEN TRP1* plasmid (34), followed by ligation of a *CYH2* gene into the *Pst*I site. The *URA3* cDNA expression libraries from *A. thaliana* (35) and *S. pombe* (36) and the genomic library from *C. albicans* (37) have been described.

Mouse cDNAs W65960 and AA245980 were obtained from Research Genetics (Huntsville, AL) and sequenced. The cDNA of AA245980, which encodes almost the entire ORF, was ligated into pET24B (Novagen) for expression in *E. coli*. PCR amplification introduced an ATG at the 5′-end, which was ligated into pBM150 (38) to make pEMP1275 (*CEN URA3* P_{GAL10} −*mTPT*), for expression in yeast under P_{GAL10} control.

The *E. coli kptA* gene was isolated from a K12 strain (EMP804–*leuB600*, Δ*lacx74*, *hsdR*, *strA*, *galE*, *galK*) by PCR amplification with primers Ecol3A (5′-CGGAATTCGCGCATATGTGGAAAAGGCTC-3′) and Ecol3 (5′-GAATTCCCCTTACCATCTAAGAAAC-3′), followed by digestion with *Eco*RI and ligation into pBM150 to generate pSLS18–2 (pBM150-*CEN IV URA3* P_{GAL10}−*kptA*)], in which the GTG start was changed to an ATG start. Amplification with primers Ecol3A and Ecol4 (5′-GCGGAATTCCCACTCAAGAAACTCAAC-3′) followed by cleavage with *Eco*RI and *Nde*I and ligation into pET24a allowed expression of the protein in *E. coli* as a His-6 fusion.

Selection for Functional Complementation of *S. cerevisiae tpt1* Mutants. Strain SC974 was transformed with either the *S. pombe* (700,000 transformants) or the *A. thaliana URA3* cDNA library (200,000 transformants), and transformants were outgrown in liquid medium lacking uracil for 24 hr, plated directly to yeast extract/peptone/dextrose plates containing 7.5 μg/ml cycloheximide, and screened for those Cyh^r colonies that were Ura$^+$, Trp$^-$, Leu$^+$, and 5-fluoroorotic acid. Three *S. pombe* and 35 *A. thaliana* transformants passed all the screens, and two of the candidates from each library were sequenced and found to contain the same gene. The *C. albicans TPT1* gene was obtained by transformation of SC839 with the genomic library, followed by screening of Ura$^+$ transformants (47,000) for those that grew on medium containing α-aminoadipic acid (which selects against the *LYS2* gene on the plasmid) and were also Trp$^-$, Lys$^-$, Leu$^+$, and sensitive to 5-fluoroorotic acid. Sixteen transformants passed all the screens, and the three DNAs that were examined contained the same *TPT1* homologous gene.

Assay of 2′-Phosphotransferase Activity. [α^{32}P]ATP ligated $tRNA^{PHE}$ substrate was prepared by *in vitro* transcription and splicing, and phosphotransferase activity was assayed as previously described (33), except that reactions with the *E. coli* protein were incubated at 37°C. Products were resolved on polyethyleneimine cellulose thin layer plates that were developed in buffer containing 2M sodium formate, pH 3.5.

Proteins. Phosphotransferase protein was prepared from *E. coli* cells expressing *TPT1*, m*TPT*, or *kptA*, by Blue Sepharose (Pharmacia) column chromatography of extracts, as described (20). Cyclic phosphodiesterase fractions were from a Blue Sepharose column fractionation of yeast extracts (39). Calf intestinal phosphatase was from Boehringer Mannheim.

Construction of the *kptA* Knockout. The *kptA* ORF (with 962 and 871 base pairs of 5′ and 3′-flanking sequence, respectively) was amplified from *E. coli* DNA using primers ECF-5 (5′-CTACCCGGGGTCGACCTGAATGATGACAAACGC-3′) and ECF-3 (5′-GTTCCCGGGGAGCTCGTGTTATTGATTCGCGTC-3′), digested with *Sal*I and *Sac*I, and ligated into pUC18. Digestion with *Eco*RV and *Sna*BI removed 267 base pairs from the *kptA* ORF, into which a *Sma*I fragment containing the Cm^r gene was placed. The *kptA* knockout strain was generated as described earlier, following ligation of the *Sal*I–*Sac*I fragment containing *kptA*::Cm^r into pCVD442 (*oriR6K mobR4P* Amp^r *sacB*) (40).

Alignment and Phylogenies. Phosphotransferases were identified in the database by using each 2′-phosphotransferase as the template for BLASTP searches vs. the nonredundant database, as well as TBLASTN searches vs. the EST database (41). Proteins were aligned by using the multiple alignment feature of CLUSTAL W (42) and presented by using MACBOXSHADE. The subsequent alignment was then used to generate a Neighbor-Joining Tree (43), and bootstrap trials were carried out by using CLUSTAL W. The alignment also was used to do an "exhaustive search" for the most parsimonious tree by using PAUP 3.1 (44).

RESULTS

Genes from several different eukaryotes and one bacterial species functionally complement a yeast *tpt1* mutant strain. Because the 2′-phosphotransferase encoded by the yeast *TPT1* gene is essential, we used plasmid shuffle methods to obtain functional *TPT1* genes from other organisms. To this end, we constructed a strain with relevant genotype *tpt1-Δ1*::*LEU2* (p *CEN TPT1 TRP1 CYH2*). Transformation of a *URA3* expression library into this strain, followed by selection against the resident *TPT1* plasmid (with cycloheximide) and appropriate screening, efficiently yields the functional homologs. Using this plasmid shuffle approach, we obtained the *TPT1* genes from *S. pombe*, *C. albicans*, and *A. thaliana*. The *S. pombe* gene proved to be the same gene that we had previously identified by homology (33), and the corresponding ORFs from the other two organisms also had similar sequences. In addition, extracts containing the Tpt1 homologs and NAD had substantial 2′-phosphotransferase activity, as measured by the transfer of phosphate from ligated tRNA to material that comigrated with Appr>p (data not shown). Thus, there is a distinct family of *TPT1* genes conserved in higher plants and distantly related yeast species.

Both the mouse (*mTPT1*) and the *E. coli* (*kptA*) candidate *TPT1* homologs that were present in the database also encode 2′-phosphotransferase activity. To show this, we cloned the corresponding genes directly to test their function (see *Materials and Methods*). After expression of a His-6 fusion of each protein in *E. coli*, we obtained crude extracts (data not shown) and partially purified preparations with activity that was indistinguishable from that of the yeast protein (Fig. 1). Removal of the phosphate from ligated tRNA requires both KptA (or mTpt1) protein and NAD, and yields product that comigrates with Appr>p in both this TLC system (lanes b, h, and n) and in a system containing 1 M LiCl. As expected for Appr>p, it is resistant to calf intestinal phosphatase (lanes c, i, and o) and is hydrolyzed to material that comigrates with Appr-1′′p when incubated with yeast cyclic phosphodiesterase (lanes d, j, and p), which is highly specific for this cyclic phosphate bond (39). As noted earlier (39), this preparation of cyclic phosphodiesterase also has a contaminating nuclease activity, which produces AMP from either ligated tRNA or dephosphorylated tRNA (compare lanes g and m with lanes d and j).

As expected, it is the 2′-phosphate from ligated tRNA that is transferred by KptA protein (see Fig. 2). Whereas Appr>p is formed from 2′-phosphorylated ligated tRNA in the presence of KptA and NAD (lanes d–h), no Appr>p is detected with either intron-containing pre-tRNA (lanes i–k) or ligated tRNA whose 2′-phosphate has been previously removed (lanes

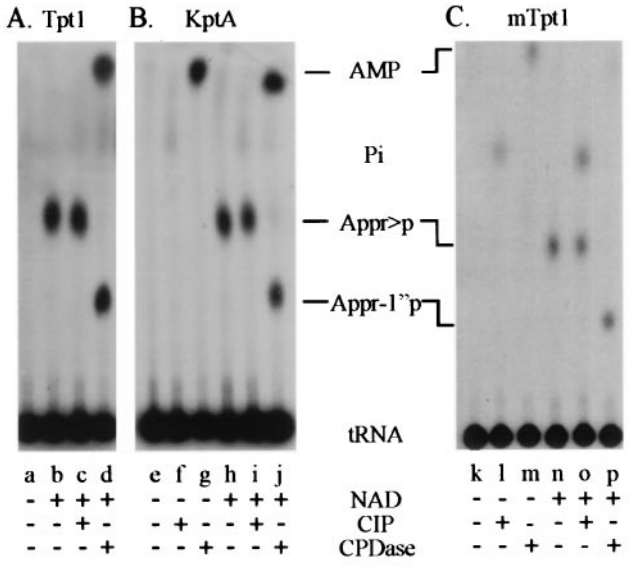

FIG. 1. The *E. coli* and mouse Tpt1 homologs have phosphotransferase activity. Partially purified phosphotransferase from yeast (Tpt1; *A*), *E. coli* (KptA; *B*), and mouse (mTpt1; *C*) were incubated with ligated $tRNA^{PHE}$ substrate and 1 mM NAD, as indicated. Portions of the mixtures were subsequently incubated with calf intestinal phosphatase (CIP), or yeast cyclic phosphodiesterase (CPDase), as shown, and samples were applied to TLC plates to resolve products.

l–n). Thus, KptA protein, like yeast phosphotransferase (17), is transferring the 2′-phosphate of ligated tRNA and not some other labeled residue. A similar result was obtained with the mouse protein (data not shown).

Both the mouse and the *E. coli* ORFs complement a yeast *tpt1* mutant when expressed in yeast under control of the

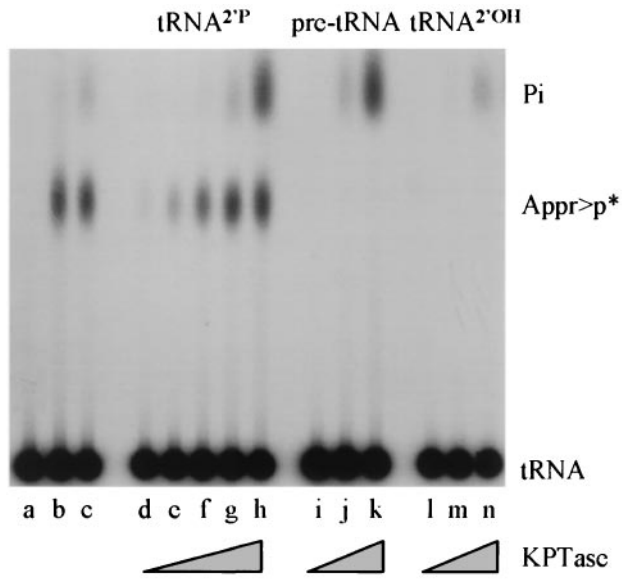

FIG. 2. *E. coli* KptA protein transfers the 2′-phosphate from the substrate ligated tRNA. Different ^{32}P-labeled derivatives of the substrate were tested as the source of the transferred phosphate, with 1 mM NAD and 5-fold serial dilutions of KptA protein, starting from 125 units of activity. After incubation, 0.1 unit of calf intestinal phosphatase was added to each tube, followed by a further 15-min incubation, to convert contaminating nuclease products (but not Appr>p) to Pi, and samples were applied to TLC plates. Lanes d–h, ligated $tRNA^{PHE}$ with a labeled 2′-phosphate (the standard substrate); i–k, labeled pre-tRNA as substrate; and l–n, mature tRNA (previously dephosphorylated with yeast Tpt1 protein). Lanes a–c, controls with ligated tRNA substrate and NAD: a, buffer control, followed by CIP buffer; b, Tpt1p, followed by CIP buffer; c, Tpt1p, followed by CIP.

P_{GAL10} promoter, which activates gene expression in galactose and represses it in glucose. This was demonstrated by the plasmid shuffle complementation test described above, in medium containing galactose. Complementaton with the *E. coli kptA* gene was somewhat weaker than the corresponding complementation with the yeast or mouse genes, since glucose repression of *kptA*, but not of *TPT1* or *mTPT1*, was strong enough to be lethal. We note that complementation by the *E. coli* gene was not due to fortuitous recombination, either with the remaining portion of the chromosomal *tpt1* gene or with the plasmid-borne copy of the *TPT1* gene that is present before cycloheximide selection to remove it. This fact was established by reisolation of the yeast plasmid bearing the *E. coli* gene and sequencing the ORF. Thus the *E. coli* KptA protein is a fully functional 2′-phosphotransferase. The observation that an *E. coli* gene can complement a yeast mutant defective in a tRNA-splicing catalytic activity is at once puzzling and fascinating because of the lack of this class of tRNA splicing in bacteria.

Tpt1 Is a Member of a Widespread Family of Phosphotransferases of Ancient Origin. An alignment of the various functional *TPT1* ORFs, together with other similar sequences found in the database, is presented in Fig. 3. It is evident from the alignment that the known functional 2′-phosphotransferases share the same distinct blocks of conserved sequence, suggesting that the basic architecture and active site residues of 2′-phosphotransferase have been preserved in all of these proteins. The other (untested) homologs, including another eubacterial ORF (*P. aeruginosa*), and three Archaeal ORFs (two in *A. fulgidus* and one in *P. horikoshii*), also share similar conserved blocks, suggesting that they too are functional 2′-phosphotransferases. The finding of *TPT1* homologs in Archaea extends the family to a third kingdom and may indicate a role of the protein in archaeal tRNA splicing. However, the role of the eubacterial protein is difficult to reconcile with splicing because of the lack of introns of this class.

To explore the origin of the eubacterial phosphotransferase, a phylogenetic analysis of the protein was done. The results are summarized in Fig. 4, as an unrooted tree. Using a Neighbor-Joining analysis, the various phosphotransferases fall into three distinct clades that represent the three major kingdoms: Archaea, Eubacteria, and Eukarya. A tree with the same topology (branching arrangement) also is obtained by using maximum parismony methods for phylogeny reconstruction. Because the tree conforms to the expected phylogeny of organisms and has high confidence levels (given by the bootstrap analysis), it strongly argues against the possibility of a trans-kingdom transfer accounting for the presence of the Eubacterial or Archaeal clade. Rather, the results from the phylogeny (although limited by the number of homologs) suggest that *TPT1* homologs have been present in Eubacteria since their separation from Archaea between 3 and 4 billion years ago (45). This phylogenetic analysis cannot rule out the possibility of horizontal transfer within Eubacteria. However, *kptA* appears to have been present in *E. coli* for a substantial period of time. A recent comprehensive analysis of the *E. coli* genome, based on GC content and codon usage, did not earmark this gene as "recently horizontally transferred" (100 million years) (46).

A small discrepancy in the unrooted tree compared with the organismal phylogeny is the position of the *S. pombe* homolog closer to the plant and animal phosphotransferases than to the other fungal homologs. Although this might be explained by an ancient horizontal transfer within the Eukaryotic kingdom, it is more likely to be a case of a paralogous (related through gene duplication) rather than orthologous (related through descent) origin. For example, the ancestral eukaryotic lineage of phosphotransferases could have undergone an early gene duplication, with a different copy propagated in *S. pombe* (and

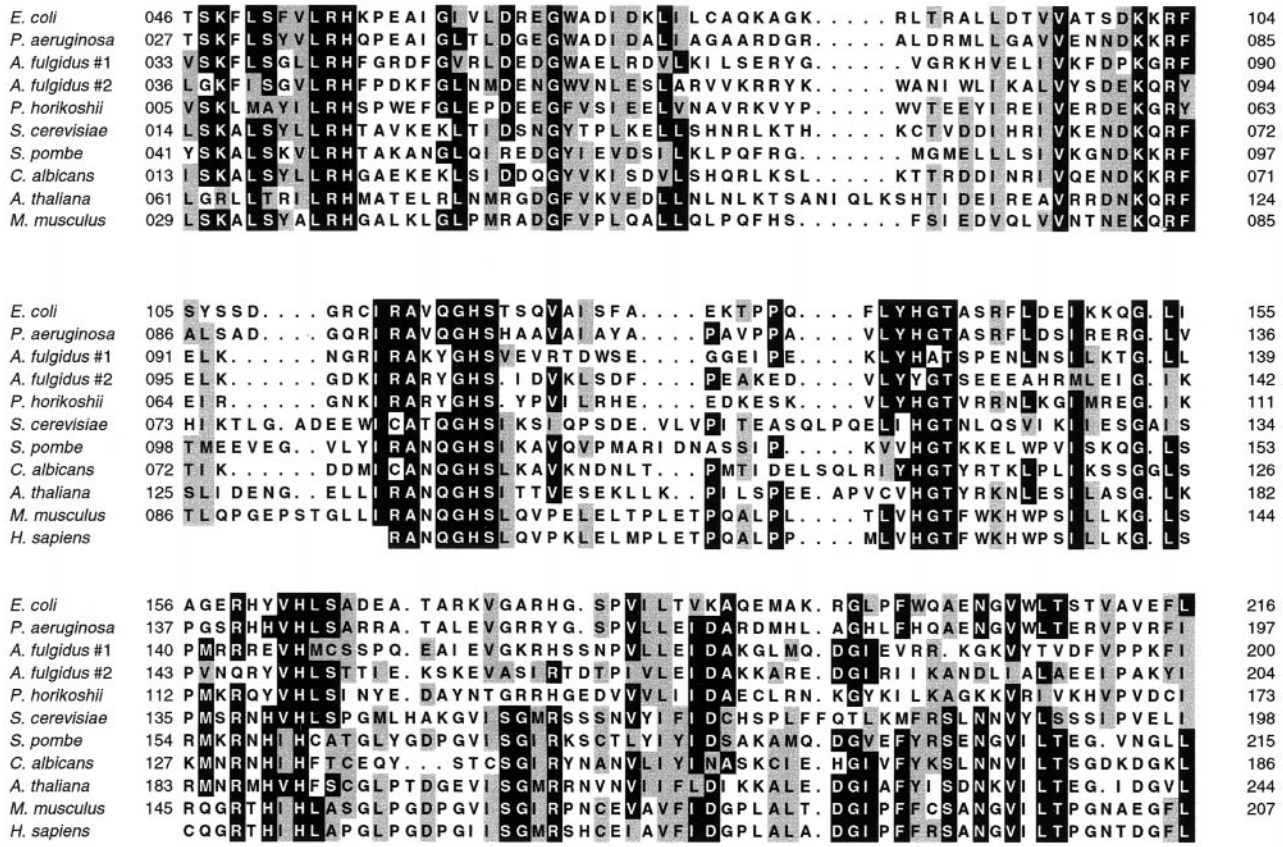

FIG. 3. Alignment of the phosphotransferase homologs. Phosphotransferases identified by biochemical and database searches are presented. The alignment is shaded to a 50% consensus by using MACBOXSHADE, with dark and light shading, indicating identical and similar residues, respectively. The human sequence is only a partial EST, and the beginning of the mouse sequence is absent from its EST. Numbers indicate the residue number for each homolog. *E. coli*, sp|P39380|YJII (ORF218); *P. aeruginosa*, gnl|PAGP|Contig581; *S. pombe*, EMB|Z99259|SPAC2C4; *A. fulgidus* #1,gb|AE001076; *A. fulgidus* #2, gb|AE000995; *P. horikoshii*, dbj|AB009470; *A. thaliana*, gb|AC002387; *M. musculus*, AA245980; A rice EST is not shown.

plants and animals) than the one in other fungi. Direct evidence for propagation of two paralogous lineages is found in the two *A. fulgidus* phosphotransferase homologs.

The retention of the bacterial phosphotransferase reading frame for what appears to be over 3 billion years argues strongly that the gene or its product has an important selective advantage in bacteria, even if not for tRNA splicing. However, two lines of evidence argue that it is not essential for life. First, there is a notable absence of a recognizable 2′-phosphotransferase in several other completely sequenced Eubacterial species, including *Bacillus subtilis*, *Haemophilus influenzae*, *Helicobacter pylori*, and *Mycoplasma genitalium*. Second, a gene disruption experiment in *E. coli* demonstrates that it is not essential there. We constructed a strain containing a tandem duplication of the *kptA* gene and flanking DNA (one copy of which was deleted and replaced by a Cm^r marker) and selected for excisive recombination to eliminate one of the *kptA* loci (40). Approximately 50% of the selected colonies had the *kptA*-Δ::Cm^r copy of the *kptA* gene (which was confirmed by PCR analysis), demonstrating that the *kptA* gene is not essential for growth. Although phosphotransferase is clearly not essential in all Eubacterial species [by contrast to yeast (33)], we presume that these proteins have another, as yet unrecognized, function to account for their retention over such a long evolutionary time.

DISCUSSION

Data presented in this paper demonstrate that the yeast *TPT1* gene, whose product is known to be involved in the phosphotransferase step of tRNA splicing, is part of a widespread family of conserved genes found in eukaryotes, bacteria, and archaea. This conservation is functional for two reasons: First, otherwise lethal *S. cerevisiae tpt1* mutants can be complemented by expression of the corresponding genes from yeasts (*C. albicans* and *S. pombe*), plants (*A. thaliana*), mammals (mouse), and bacteria (*E. coli*). Second, extracts from each of the complemented strains (as well as from the partially purified mouse and *E. coli* proteins) have 2′-phosphotransferase activity. The strong conservation of several distinct blocks of sequence among these proteins suggests a common structure and active site, which appears also to be present in gene products from another bacterial species (*P. aeruginosa*) and two archaeal species (*P. horikoshii* and *A. fulgidus*). Phylogenetic analysis suggests that the phosphotransferase family is ancient, having arisen before or shortly after the separation of the three kingdoms.

The conservation of 2′-phosphotransferase throughout eukaryotes was anticipated, for three reasons. First, the ligase that generates its substrate, an RNA splice junction with a 2′-phosphate, is widely conserved in plants, yeasts, and humans (21, 25, 27, 28). Second, this phosphotransferase is the only activity that can efficiently remove splice junction 2′-phosphates from tRNA in yeast and HeLa extracts (28, 47). Third, the known role of both ligase and phosphotransferase in yeast tRNA splicing (19, 20), coupled with the widespread occurrence of tRNA introns in eukaryotes, argues that this metabolic pathway would be widely conserved in eukaryotes (although some vertebrates appear to have another ligation pathway, which bypasses the need for a phosphotransferase).

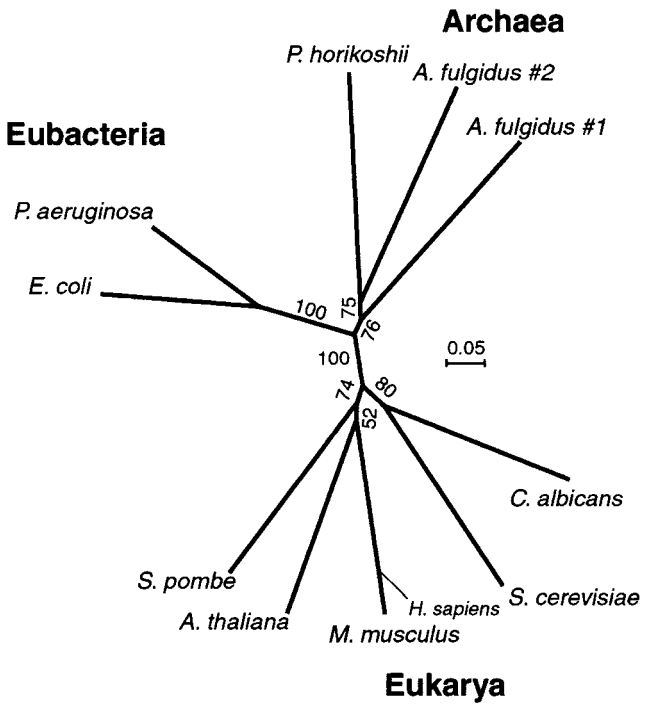

FIG. 4. Phylogeny of the phosphotransferase homologs. An unrooted phylogram, based on the Neighbor-Joining method, of the various phosphotransferases is presented, using the alignment in Fig. 3. Numbers next to each branch represent bootstrap values as a percentage of 1,000 trials. The branch lengths represent actual amino acid divergences based on the scale shown (0.05 denoting 5% amino acid divergence).

The finding that the joining step of yeast *HAC1* mRNA splicing (a transcriptional regulator of the unfolded protein response) is catalyzed by tRNA ligase (48) extends this argument; it suggests that tRNA ligase (and perhaps phosphotransferase) may participate in a separate class of mRNA-splicing reactions in eukaryotes.

The existence of a 2′-phosphotransferase in Archaea could, as in yeast, imply a role for the protein in RNA splicing. This interpretation would be consistent with the existence of similar archaeal and eukaryotic tRNA introns and homologous endonucleases to excise them (12–15). However, such a role is not clear for two reasons: First, no ligase that generates a 2′-phosphate has yet been found in these organisms. Indeed, the one reported archaeal ligase may be more similar to the vertebrate-type ligase, which does not generate a 2′-phosphate (32). Second, it is clear from the database that several fully sequenced archaeal species (e.g., *Methanococcus jannaschii*) have no obvious 2′-phosphotransferase, although they have both an endonuclease and intron-containing tRNAs. As discussed below, there may be other roles of this family of proteins, perhaps in catalyzing related activities.

The existence of the eubacterial phosphotransferase is an enigma, in view of the lack of known ligases or introns of the appropriate class in this kingdom. The best interpretation is that the eubacterial protein has some important other function that is not routinely required; this would account for both the conservation of the reading frame over evolutionary times and for the fact that the eubacterial gene is not essential. The finding of an enzyme, like 2′-phosphotransferase, that is present in both *S. cerevisiae* and *E. coli*, but essential only in *S. cerevisiae* (33), is rare but not unique. One example of this type is the enzyme ATP(CTP) tRNA nucleotide transferase. This enzyme is essential for the addition of CCA to the ends of tRNA in yeast (49) but not in *E. coli*, which has CCA encoded on the ends of its tRNA genes (50). Its function in *E. coli*, which also is likely conserved in yeast, is to repair degraded CCA ends (50, 51)].

It seems likely that the function of the bacterial phosphotransferase is intimately related to the binding pocket and active site of the protein. To account for its activity in yeast cells, the bacterial protein must still have a high degree of substrate specificity, substrate selectivity, and chemical reactivity; otherwise it would not be able to find ligated tRNA substrates *in vivo* and efficiently dephosphorylate them. The retention of so many features of the protein over such large evolutionary periods is most simply explained by the conservation of a highly related reaction. Indeed, our results indicate that the *E. coli* protein and the yeast protein are highly similar in both their recognition of substrate (M. A. Steiger, R. Kierzek, D. Turner, and E.M.P., unpublished observations) and their mechanism (S.L.S., R. Kierzek, D. Turner, and E.M.P., unpublished observations).

We can think of four possible roles of the *TPT1* gene product in Eubacteria: First, there may be 2′-phosphorylated RNAs generated in these organisms, despite the absence of an obvious yeast-like ligase. Second, the Eubacterial phosphotransferase may transfer a phosphate from other molecules to NAD to form Appr>p as part of a metabolic or a regulatory circuit. Either Appr>p or the dephosphorylated molecule could be the important product. In this connection, we note that NADP, which has a 2′-phosphate, is not a phosphate donor for the *E. coli* KptA protein (M. A. Steiger, R. Kierzek, D. Turner, and E.M.P., unpublished observations). Third, there may be a related but superficially different chemical reaction catalyzed by the bacterial protein. Although evolutionarily related proteins generally catalyze similar chemical reactions, and sometimes acquire new substrates, the reactions can appear to be quite different from one another (52). Fourth, there may be a completely different and unexpected role of this protein. This would be reminiscent of the dual roles of T4 RNA ligase in RNA ligation and tail fiber attachment (53), the conserved aconitase activity of the iron response element binding protein (54), the role of various glycolytic enzymes such as lactate dehydrogenase in the lens crystallins (55), and the tRNA synthetase activity of group I-splicing cofactors (56, 57). Detailed knowledge of the mechanism of the *E. coli* 2′-phosphotransferase reaction, and of its substrate recognition determinants, will greatly help us in deducing the nature of any possible related functions of the *E. coli* protein. These functions also may be retained in other eukaryotes and archaea.

We thank L. Guarente, F. Sherman, and J. Normanly for the *S. pombe*, *C. albicans*, and *A. thaliana* libraries, and are grateful to L. Wright and R. Silver for strains and for advice on *E. coli* genomic knockouts. We also thank T. H. Eickbush, E. J. Grayhack, and M. A. Steiger for discussions about this work and for critical reading of the manuscript. This work was supported by a National Institutes of Health Grant GM52347 (to E.M.P.). H.S.M. was supported by a National Science Foundation Grant MCB-9601198 (awarded to T. H. Eickbush).

1. Xu, M. Q., Kathe, S. D., Goodrich-Blair, H., Nierzwicki-Bauer, S. A. & Shub, D. A. (1990) *Science* **250,** 1566–1570.
2. Reinhold-Hurek, B. & Shub, D. A. (1992) *Nature (London)* **357,** 173–176.
3. Ferat, J. L. & Michel, F. (1993) *Nature (London)* **364,** 358–361.
4. Belfort, M. & Weiner, A. (1997) *Cell* **89,** 1003–1006.
5. Abelson, J., Trotta, C. R. & Li, H. (1998) *J. Biol. Chem.* **273,** 12685–12688.
6. Hopper, A. K. & Martin, N. C. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 2, pp. 99–141.
7. Lykke-Andersen, J., Aagaard, C., Semionenkov, M. & Garrett, R. A. (1997) *Trends Biochem. Sci.* **22,** 326–331.

8. Peebles, C. L., Gegenheimer, P. & Abelson, J. (1983) *Cell* **32,** 525–536.
9. Thompson, L. D. & Daniels, C. J. (1988) *J. Biol. Chem.* **263,** 17951–17959.
10. Reyes, V. M. & Abelson, J. (1988) *Cell* **55,** 719–730.
11. Thompson, L. D. & Daniels, C. J. (1990) *J. Biol. Chem.* **265,** 18104–18111.
12. Fabbri, S., Fruscoloni, P., Bufardeci, E., Di Nicola Negri, E., Baldi, M. I., Attardi, D. G., Mattoccia, E. & Tocchini-Valentini, G. P. (1998) *Science* **280,** 284–286.
13. Kleman-Leyer, K., Armbruster, D. W. & Daniels, C. J. (1997) *Cell* **89,** 839–847.
14. Trotta, C. R., Miao, F., Arn, E. A., Stevens, S. W., Ho, C. K., Rauhut, R. & Abelson, J. N. (1997) *Cell* **89,** 849–858.
15. Lykke-Andersen, J. & Garrett, R. A. (1997) *EMBO J.* **16,** 6290–6300.
16. Greer, C. L., Peebles, C. L., Gegenheimer, P. & Abelson, J. (1983) *Cell* **32,** 537–546.
17. McCraith, S. M. & Phizicky, E. M. (1991) *J. Biol. Chem.* **266,** 11986–11992.
18. Culver, G. M., McCraith, S. M., Zillmann, M., Kierzek, R., Michaud, N., LaReau, R. D., Turner, D. H. & Phizicky, E. M. (1993) *Science* **261,** 206–208.
19. Phizicky, E. M., Consaul, S. A., Nehrke, K. W. & Abelson, J. (1992) *J. Biol. Chem.* **267,** 4577–4582.
20. Spinelli, S. L., Consaul, S. A. & Phizicky, E. M. (1997) *RNA* **3,** 1388–1400.
21. Konarska, M., Filipowicz, W., Domdey, H. & Gross, H. J. (1981) *Nature (London)* **293,** 112–116.
22. Schwartz, R. C., Greer, C. L., Gegenheimer, P. & Abelson, J. (1983) *J. Biol. Chem.* **258,** 8374–8383.
23. Pick, L., Furneaux, H. & Hurwitz, J. (1986) *J. Biol. Chem.* **261,** 6694–6704.
24. Pick, L. & Hurwitz, J. (1986) *J. Biol. Chem.* **261,** 6684–6693.
25. Kikuchi, Y., Tyc, K., Filipowicz, W., Sanger, H. L. & Gross, H. J. (1982) *Nucleic Acids Res.* **10,** 7521–7529.
26. Zillmann, M., Gorovsky, M. A. & Phizicky, E. M. (1991) *Mol. Cell. Biol.* **11,** 5410–5416.
27. Baymiller, J., Jennings, S., Kienzle, B., Gorman, J. A., Kelly, R. & McCullough, J. E. (1994) *Gene* **142,** 129–134.
28. Zillman, M., Gorovsky, M. A. & Phizicky, E. M. (1992) *J. Biol. Chem.* **267,** 10289–10294.
29. Laski, F. A., Fire, A. Z., RajBhandary, U. L. & Sharp, P. A. (1983) *J. Biol. Chem.* **258,** 11974–11980.
30. Filipowicz, W. & Shatkin, A. J. (1983) *Cell* **32,** 547–557.
31. Nishikura, K. & De Robertis, E. M. (1981) *J. Mol. Biol.* **145,** 405–420.
32. Gomes, I. & Gupta, R. (1997) *Biochem. Biophys. Res. Commun.* **237,** 588–594.
33. Culver, G. M., McCraith, S. M., Consaul, S. A., Stanford, D. R. & Phizicky, E. M. (1997) *J. Biol. Chem.* **272,** 13203–13210.
34. Gietz, R. D. & Sugino, A. (1988) *Gene* **74,** 527–534.
35. Minet, M., Dufour, M. E. & Lacroute, F. (1992) *Plant J.* **2,** 417–422.
36. McNabb, D. S., Tseng, K. A. & Guarente, L. (1997) *Mol. Cell. Biol.* **17,** 7008–7018.
37. Janbon, G., Rustchenko, E. P., Klug, S., Scherer, S. & Sherman, F. (1997) *Yeast* **13,** 985–990.
38. Johnston, M. & Davis, R. W. (1984) *Mol. Cell. Biol.* **4,** 1440–1448.
39. Culver, G. M., Consaul, S. A., Tycowski, K. T., Filipowicz, W. & Phizicky, E. M. (1994) *J. Biol. Chem.* **269,** 24928–24934.
40. Donnenberg, M. S. & Kaper, J. B. (1991) *Infect. Immun.* **59,** 4310–4317.
41. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
42. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
43. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
44. Swofford, D. L. (1993) PAUP, Phylogenetic Analysis Program Using Parsimony (Illinois Natural History Survey, Champaign, IL), Version 3.1.
45. Feng, D.-F., Cho, G. & Doolittle, R. F. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13028–13033.
46. Lawrence, J. G. & Ochman, H. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9413–9417.
47. McCraith, S. M. & Phizicky, E. M. (1990) *Mol. Cell. Biol.* **10,** 1049–1055.
48. Sidrauski, C. & Walter, P. (1997) *Cell* **90,** 1031–1039.
49. Aebi, M., Kirchner, G., Chen, J. Y., Vijayraghavan, U., Jacobson, A., Martin, N. C. & Abelson, J. (1990) *J. Biol. Chem.* **265,** 16216–16220.
50. Zhu, L. & Deutscher, M. P. (1987) *EMBO J.* **6,** 2473–2477.
51. Wolfe, C. L., Hopper, A. K. & Martin, N. C. (1996) *J. Biol. Chem.* **271,** 4679–4686.
52. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D. & Kozarich, J. W. (1993) *Trends Biochem. Sci.* **18,** 372–376.
53. Snopek, T. J., Wood, W. B., Conley, M. P., Chen, P. & Cozzarelli, N. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3355–3359.
54. Kaptain, S., Downey, W. E., Tang, C., Philpott, C., Haile, D., Orloff, D. G., Harford, J. B., Rouault, T. A. & Klausner, R. D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10109–10113.
55. Wistow, G. J., Mulders, J. W. & de Jong, W. W. (1987) *Nature (London)* **326,** 622–624.
56. Lambowitz, A. M. & Perlman, P. S. (1990) *Trends Biochem. Sci.* **15,** 440–444.
57. Mohr, G., Caprara, M. G., Guo, Q. & Lambowitz, A. M. (1994) *Nature (London)* **370,** 147–150.