BIOLOGIJA. 2007. Vol. 53. No. 4. P. 1–4

Identification of *Tetrahymena pyriformis* ribonuclease P subunits: RNA and protein Pop1

Violeta Jonušienė,

Sofija Sasnauskienė,

Benediktas Juodka

Department of Biochemistry and Biophysics, Vilnius University, M. K. Čiurlionio 21, LT-03101 Vilnius, Lithuania

Ribonuclease P (RNase P) removes 5'-leader sequences from precursor tRNAs and occurs in all three phylogenetic domains. The enzyme is universally composed of both protein and RNA moieties. RNase P requires both RNA and protein subunits for functioning *in vivo*. Little is known about the RNase P from ciliated protozoan. In this study, we have identified a *Tetrahymena pyriformis* sequence showing a homoly to Pop1, the largest protein subunit of RNase P. cDNA containing a complete coding sequence for the protein was cloned. We report the cloning and characterization of the gene encoding RNase P RNA and the present secondary structure model of the predicted RNA.

Key words: RNase P, pre-tRNA processing, Pop1

INTRODUCTION

During the production of a mature molecule, precursor tRNAs (pre-tRNAs) must undergo a complex series of processing steps. Ribonuclease P (RNase P) cleaves pre-tRNA molecules to produce mature 5'-termini in both prokaryotes and eukaryotes. The bacterial enzyme is composed of a large RNA subunit (typically 350–400 nt) and a single small protein (~10% by mass). The bacterial RNA subunit alone is able to recognize and cleave pre-tRNA substrates *in vitro* under high Mg^{2+} concentrations, while both RNA and protein subunits are required for activity *in vivo* [1]. The archeal and eukaryal nuclear RNase P enzymes have a more complex subunit composition, retaining a structurally related RNA (typically 350–400 nt), but with an increased number of protein subunits. At least four proteins have been found in archeal enzymes (about 45% by mass), nine proteins in yeast and ten proteins in human (about 70% by mass) [2, 3]. The lack of catalytic activity for eukaryotic RNase P RNAs *in vitro* suggests a more critical role for protein subunits in eukaryotes than that provided by protein subunits in bacterial holoenzymes [4]. However, the RNase P RNR subunit from humans and the lower eukaryote *Giardia lamblia* were shown to catalyze the processing of four different pre-tRNAs in the absence of protein. The level of RNase P catalytic activity in the human RNR subunit is one millionth of what is found in bacterial RNR, if the assay for biochemical catalysis is carried out at pH 6 [5]. New discoveries on the catalytic activity of eukaryotic RNase P have provided the further impetus for studying this enzyme in eukaryotes. The eukaryotic RNase P has evolved into several different enzymes expressing their activity in the nucleus and organelles. Studies of RNase P in organelles suggest a diversity of their composition. The well-characterized mitochondrial RNase P in *Saccharomyces cerevisiae* clearly contains an RNA subunit and one 105 kDa protein subunit [6]. In contrast, *Aspergillus nidulans* mitochondrial RNase P contains seven polypeptides [7]. An interesting case has been detected in spinach chloroplasts where RNase P has the properties of a solely protein enzyme [8].

Here, we present the cloning two components, protein Pop1 and RNR subunit, of the nuclear RNase P complex from the ciliated protozoan *Tetrahymena pyriformis*.

MATERIALS AND METHODS

Cell culture. *Tetrahymena pyriformis* (strain GL) cells were grown at 26 °C with aeration in a medium containing 0.6% of meat peptone, 0.6% of yeast extract, 1 mM Mg_2SO_4, 50 μM ferrum citrate (pH 6.4).

Oligonucleotides. Oligonucleotides (Table) were obtained from AB Fermentas or MWG Biotech.

Isolation of DNA. 1×10^8 *T. pyriformis* cells were harvested by centrifugation, washed with cold 10 mM Tris-HCl (pH 7.4 at 4 °C) and suspended in 10 ml of the same buffer. The cell suspension was mixed with 1 volume of the lysis buffer (10 mM Tris HCl, pH 7.4 at 25 °C, 0.1 M Na_2EDTA, 1% SDS) and 0.5 volume of phenol, incubated for 10 min at 60 °C, then extracted with chloroform / isoamyl alcohol twice. The aqueous phase was recovered, dialyzed against SSC, and subsequently treated with 100 μg/ml of RNase A and 100 μg/ml of proteinase K [9].

Gene amplification. Adaptor PCR: genomic DNA samples were digested in parallel with one of the following restriction enzymes: *Bam*HI, *Bgl*II or *Xba*I. Adaptors were annealed by mixing 2.5 μg of the oligonucleotides Bam or Xba with 2.5 μg Uni in 100 μl H_2O, heating for 5 min at 85 °C and slowly cooling to room temperature. Digested genomic DNA (100 ng) was then ligated to 50 ng of an appropriate adaptor by incubation at 16 °C overnight [10]. Ligation mixtures were used for amplification with the sequence-specific and Uni primers.

E-mail: violeta.jonusienė@gf.vu.lt

Table. **Structure and abbreviations of DNA oligonucleotides used in this study**

Abbreviation	Sequence	Destination
RNP3	5′-AGTAGTCTGAATTGGGTT-3′	RNase P RNP gene
RNP5	5′-CTGGGAAGGTCTGAGA-3′	RNase P RNP gene
CRI	5′-GGTGGGAARKTCGGTG-3′	RNase P RNP gene
CRV	5′-DTARTCGGWATCGVGTT-3′	RNase P RNP gene
RNP3′	5′-CGGTCAAAGTTTGCGTTGAG-3′	3′-end of RNase P RNP gene
RNP5′	5′-TCGGTCAAGAGTAGTCTGAATTG-3′	5′-end of RNase P RNP gene
POP1	5′-CTNCCNAGRCAYATGAGRAGAAG-3′	RNase P POP1 gene
POP2	5′-TGCCADATRTGDGTRGRNARCCA-3′	RNase P POP1 gene
POP3	5′-AATGTCCCACAATAGAATACAGAATACC-3′	RNase P POP1 gene
POP4	5′-GGTATTCTGTATCTATTGTGG-3′	RNase P POP1 gene
Uni	5′- CTCGTAGACTGCGTACC-3′	adaptor
Bam	5′-GATCGGTACGCAGTC-3′	adaptor
Xba	5′-CTAGGGTACGCAGTC-3′	adaptor

The nucleotide of RNase P RNA subunit has been deposited in the GenbankTM / EBI Data Bank under the accession number EF014343 and nucleotide sequence of RNase P Pop1 gene as EU136732.

RESULTS AND DISCUSSION

Isolation of the gene encoding a protein homologous to Pop1

Previously, three polypeptides of about 100, 45 and 35 kDa in sizes have been found to co-purify with the activity of ribonuclease P (RNase P) from *T. pyriformis*, suggesting that these protein subunits could play a role in RNase P activity. Alternatively, at least one of these polypeptides could be a subunit of RNase P [11]. It has been discovered that two 45 and 35 kDa proteins, which co-fractionated with RNase P activity, are α and γ subunits of the eukaryotic translation elongation factor eEF1B. The co-purification is based on a nonspecific interaction of eEF1B γ with the RNA subunit of RNase P [12].

The function of the individual protein subunits of eukaryotic RNase P are largely unknown, although it has been established that several protein subunits can bind RNA. We have predicted that the 100 kDa polypeptide we purified can be a homologue of the yeast Pop1p (100 kDa), the largest protein subunit of RNase P. Pop1p is important for assembly and functioning of RNase P and is conserved from yeast to human [13, 14]. Pop1p and Pop4p bind directly to RNase P RNA [15]. Alignment of the Pop1p sequences from yeast, worm and human revealed three highly conserved regions [16]. The COR1 and COR2 regions have been designated as the "Pop1 domain", the signature sequences of Pop1 family [14]. We synthesized the degenerated primers corresponding to the conserved regions COR1 and COR2 and used them for cDNA amplification. Three 200–350 bp products were amplified, cloned and sequenced. We have identified a clone that has an open reading frame theoretically coding for the polypeptide of 70 amino acids and exhibits an identity with Pop1. This fragment was then used for cloning a full-length Pop1 cDNA. The complete open reading frame has the potential to encode the basic protein (pI = 9.22), of 718 amino acids, with a predicted molecular weight of 84.96 kDa.

The sequence of the putative protein was aligned with those of Pop1 from *Tetrahymena termophyl*a, *Paramecium tetraurelia*, *Mus musculus*, *Anopheles gambiae*, using the Clustal W algorithm (http://us.expasy.org). The alignment confirmed the presence of highly conserved regions COR1, –2, and –3 (Fig. 1). The conserved basic and aromatic amino acids in COR1 and COR2 could potentially contribute to RNA binding by Pop1. Conserved positions in COR3 are more dispersed. COR3 has been named the POPLD (NUC188) domain. This domain is found in POP1-like nucleolar proteins. Sequence conservation of Pop1 family proteins outside the conserved regions is low, with an only 4% sequence identity [14]. The presence of the dispersed conserved sequence blocks suggests that, despite the low overall conservation, this protein is homologous to Pop1. A comparative analysis of the cDNA sequence and the genomic DNA fragment has revealed that the Pop1gene is approximately 3051 base in length and has five exons and four introns (Fig. 2).

Structure of *Tetrahymena pyriformis* ribonuclease P RNA

The RNase P RNAs in Bacteria, Archaea and Eukarya are all homologues, of common ancestry, but they have undergone a substantial evolutionary diversification. A decrease of the RNR catalytic activity and an increase in protein content of eukaryotic enzymes appears to accompany complexity reduction in the eukaryotic RNA structure. Secondary structure models for eukaryal RNase P RNA have been developed through phylogenetic comparisons of diverse eukaryotic RNAs, and a core structure shared with bacterial RNAs is evident. Although sequence conservation among the eukaryal RNA is low, it includes distinct sequence elements, conserved regions CR-I, CR-V and more weakly conserved CR-IV, found in all archaeal and bacterial RNAs [14, 17]. We used the known sequences of CR-I and CR-V regions to design oligonucleotide primers for RNA amplification from *T. pyriformis* genomic DNA. Amplified products about 200–400 bp in length were cloned into the pTZ57 / T vector for the subsequent sequence analysis. The candidates were analysed to check whether the other conserved features of RNA were present. The fragment containing the full-length RNA gene was amplified with primers RNP3' and RNP5', homologous to 5' and 3' ends of *T. thermophila* RNase P RNA. We identified the RNase P RNA

```
Conserved region COR1
POP1_Tpyriformis        RGKENILFIIQLAYQMVKKHLRRRAMSHNRYRIPSRIRNRMNVQELDVIEKTQDP--IRC 88
POP1_Ttermophyla        QLNSKKNGSNKLAYQMVKKHLRRRAMSHNRYRIPSRIRNRMNVQELDVIEKTQDP--IKC 93
POP1_Paramecium         TLKSKFLNSTAQPFQVIHKHMRRRAMSHNRYRIPSSLRKHMSGPGLEERDLKQKT--PKC 82
POP1_Mus                AVTQK--SSNSLVFQTLPRHMRRRAMSHNVKRLPRRLQEMAKKEAEKAAHQKKEHSKNKC 178
POP1_Anopheles          TMNSG-SNQRKLMHQSLPCHMRRRAMSYNVRRLPRRFRQVHTAQFNKSGVSEKKK--RPS 97

Conserved region COR2
POP1_Tpyriformis        RKHLRKPKLLMNSYIRRSKSD-KWMETHLWHAKRMRMIKYFQYKIARTPNEKGDRACYRF 147
POP1_Ttermophyla        RKHLRKPKLLMNSYIRRSRSD-KWMETHLWHAKRMRMIKYFQYKIARTPNEKGDRACFRF 152
POP1_Paramecium         RKHIRKRRDLI--FLNRSNKI-LWMETHLYQVKRMKMINYCGFKVPEKANEKNFRAVHRY 139
POP1_Mus                HKARRCHINRTLEFNRRQQKN-IWLETHIWHAKRFHMVKKWGYCLGERPTAKSHRACYRA 237
POP1_Anopheles          RRYRRKPTNLLREYERRKRAF-VWLETHVWHAKRFHMTSRWGYKLPLAPCSKGYRSSYRA 156

POP1_Tpyriformis        ANHDCTVYDRSYYEFLFIQGTN--LLEFLQVHLLQKDYKTVLVNQKKNFKISDPIDVYKN 205
POP1_Ttermophyla        ANHDCTVYDRSYYEFIFIQGSK--ILEFMASHLQQKDYKILIQNQKKNFKISDNLDVYKN 210
POP1_Paramecium         FDHGSMIFDASYYQIVEGS---------IENELVVDDHTIIHPQEVPDTDFVVHNNKYN- 189
POP1_Mus                MTNLCLLQDLSYYCCLELKGKEEELLKALSQMCSIDAGLTFAAVHCLSGKRQGSLMLYRA 297
POP1_Anopheles          SSKHCLVHDLSYEGCVEVSGEETQLKEGFRRLCSERVGLTLAAKAYTAGDRAGYVWLYRD 216

Conserved region COR3
POP1_Tpyriformis        EEVEVEDTEQQQNQEQMQIEEIKLANDLVSSDINMVIIFR----NSKQHKGYGQGVTLVV 507
POP1_Ttermophyla        --IQTEEVEAQSNN-SIQLEEIKFAKDLVESEIHLMIIFN----NSKQHKAFGQGVTLVV 509
POP1_Paramecium         ----EEDLQKLQQNTTLKPKDIQIK---------MLIIPP----SNQIKKQIDIGYKLII 423
POP1_Mus                NKLLDQDLNRMRSELLVPGSQLDLGARESKIPILLIQQPGKV--TGEDRLGWGSGWDVLI 652
POP1_Anopheles          ----THELSVMRSVETLVAGELCRSETALQPLPVLLLQNPGSQDAAYKRLGYGAGWELIV 507

POP1_Tpyriformis        NQGRGVLVWRLLSFIACKAIGLSEYNNIIEESGQKIFPNDYPLTNAYSILQKEQLEDKVT 567
POP1_Ttermophyla        NQGRGILVWRLLSFIACKAIGLSEYNNIIEESGQKVFPNDYPLTQAYSTLNKE------- 562
POP1_Paramecium         NLGGGLIVWRLFNQLKLKPVGLRDIRQLCLEN--NVLFNEYENEDLIEKYLKSTN----- 476
POP1_Mus                PKGWGMAFWIPFIYRGARVGGLKEATVHSQYRRSPNIPGDFPDCRAGVLFAEDQAKDLLE 712
POP1_Anopheles          PAGYGLAVWHSLVMWGARPVGQLQLDMLELETG--IDRSGVPDTVLGQEEASRRHAEALS 565

POP1_Tpyriformis        KHFMKPKGKRVNYARINSPFPLKPNFDYILSKYLE-------------------------- 602
POP1_Ttermophyla        ---------------LINSPFPFKPNFDYILSKYLE-------------------------- 583
POP1_Paramecium         -----------KINYLALQFPYPFSFDFSLYPSWC-------------------------- 500
POP1_Mus                KYRRRPPAKRPNYVKLGTLAPFCCPWEQLTRDWESRVHAQEAAIIASMPGAQETDPRRLG 772
POP1_Anopheles      KYFHRPSNKRVNYTKLAIASPFLCPWTQLVQEWNKASDGPLPFFVLRD---------- 613
```

Fig. 1. Alignment of the conserved regions (COR1-3) of Pop1 from *Tetrahymena pyriformis, Tetrahymena termophyla, Paramecium tetraurelia, Mus musculus, Anopheles gambiae*. The sequences were aligned with Clustal W algorithm. Similar residues have a light grey background; identical residues have a dark grey background. The underlined region represents the conserved regions COR1, COR2 and COR3

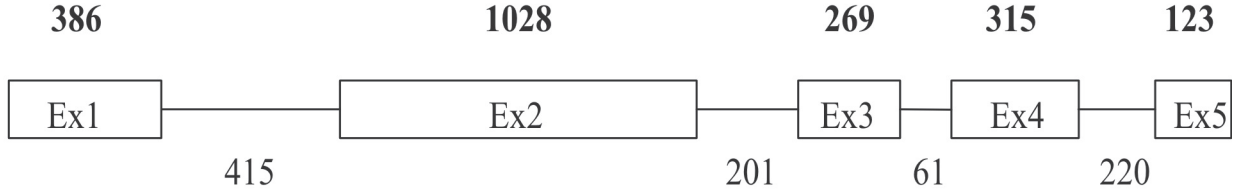

Fig. 2. Schematic map of Pop1 gene from *Tetrahymena pyriformis* that contains five exons and four introns

gene from *T. pyriformis* 316 bp long in size. The secondary structure model of the RNA is shown in Fig. 3. RNA secondary structure predictions were carried out by MFOLD [18]. *T. pyriformis* RNase RNA has a subset of eukaryal structures – stems eP8, eP9, eP19, but lacks eP15 stem. The eP15 and eP19 stems, unlike eP8 and eP9, are not found in eukaryal RNAs [17]. In *T. pyriformis*, like in most eukaryotes, the P3 stem has a large internal loop, and there is a limited sequence conservation. The eukaryotic P3 stem is required for a correct assembly and functioning of the holoenzyme and appears to bind the protein subunit Pop1 [15].

Received 3 September 2007
Accepted 18 October 2007

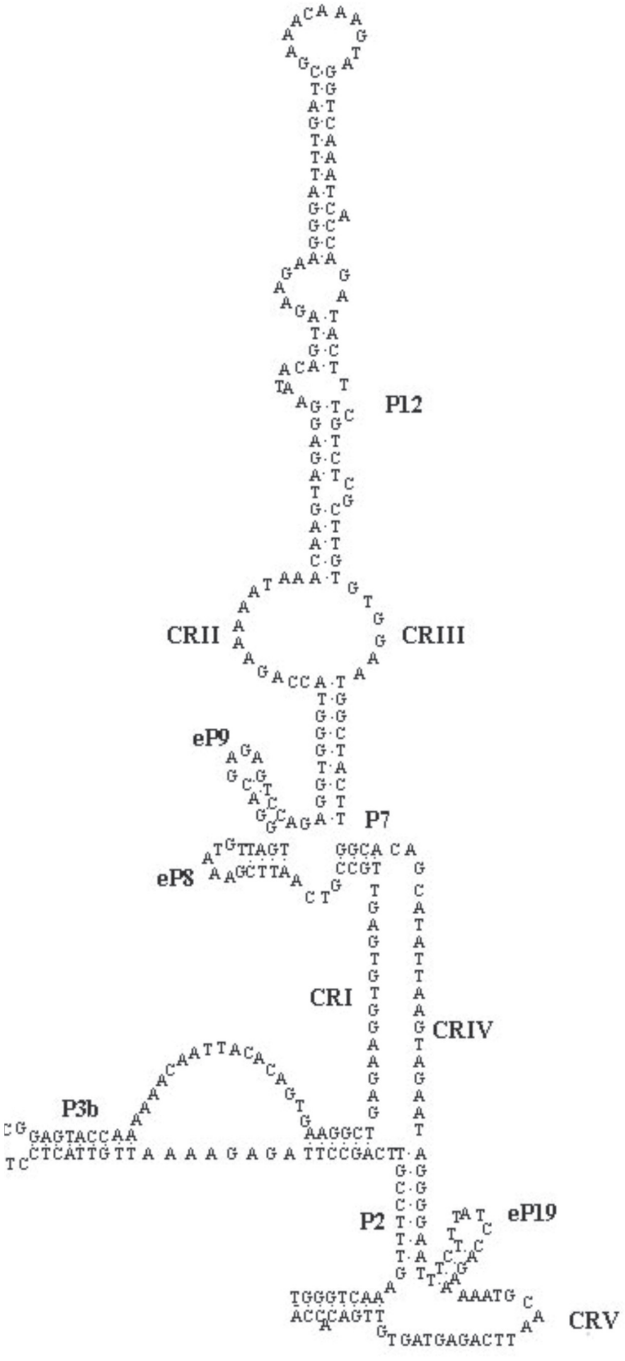

Fig. 3. Secondary structure model of RNase P RNA subunit from *Tetrahymena pyriformis*

References

1. Guerrier-Takada C, Altman S. Science 1984; 223: 285–6.
2. Jarrous N, Altman S. Methods in Enzimology 2001; 342: 93–100.
3. Walker SC, Engelke DR. Crit Rev Biochem Mol Biol 2006; 41: 77–102.
4. Xiao S, Houser-Scott F, Engelke D. J Cell Physiol 2001; 187: 11–21.
5. Kikovska E, Svard SG, Kirsebom LA. Proc Natl Acad Sci U S A 2007; 104(7): 2062–7.
6. Martin NC. Organellar tRNAs: Biosynthesis and Function. Söll D, RajBhandary U, editors. tRNA: Structure, biosynthesis and function, ASM Press, Washington, D.C 1995.
7. Lee YC, Lee BJ, Hwang DS, Kang HS. Eur J Biochem. 1996; 235(1–2): 289–96.
8. Wang MJ, Davis NW, Gegenheimer P. EMBO J 1998; 7: 1567–74.
9. Junker V, Teichmann T, Hekele A, Fingerhut Ch, Beier H. Nucleic Acid Res 1997; 25: 4194–200.
10. Frank D, Adami C, Ehringer MA, Pitule C, Pace NR. RNA 2000; 6: 1895–904.
11. Vainauskas S, Stribinskis V, Padegimas L, Juodka B. Biochimie 1998; 80: 595–604.
12. Jonušienė V, Sasnauskiene S, Juodka B. Cellular & Molecular Biology Letters 2005; 10: 689–96.
13. Lygerou Z, Mitchell P, Petfalski E, Seraphin B, Tollervey D. Genes Devel 1994; 8: 1423–33.
14. Xiao S, Hsieh J, Nugent RL, Coughlin DJ, Fierke CA, Engelke DR. RNA 2006; 12: 1023–37.
15. Ziehler WA, Morris J, Scott FH, Millikin C, Engelke DR. RNA 2001; 7: 565–75.
16. Lygerou Z, Pluk H, van Venrooij WJ and Seraphin B. EMBO J 1996; 15: 5936–48.
17. Piccinelli P, Rosenblad MA, Samuelsson T. Nucleic Acids Res 2005; 33: 4485–95.
18. Zuker M. Science 1989; 244: 48–52.

Violeta Jonušienė, Sofija Sasnauskienė, Benediktas Juodka

RIBONUKLEAZĖS P IŠ *TETRAHYMENA PYRIFORMIS* RNR IR POP1 BALTYMO SUBVIENETŲ NUSTATYMAS

Santrauka

Ribonukleazė P – ribonukleoproteinas, atsakingas už tRNR pirmtakų 5’ papildomos sekos pašalinimą tiek prokariotuose, tiek eukariotuose, yra vienas įdomiausių tRNR brendimo fermentų. Fermentą sudarantys RNR bei vienas ar keli baltyminiai subvienetai yra būtini RNazės P funkcionavimui *in vivo*. RNazė P buvo išgryninta Biochemijos ir biofizikos katedroje. Visuose chromatografinio gryninimo etapuose RNazės P aktyvumu pasižyminčiose frakcijose buvo aptikti trys pagrindiniai polipeptidai, iš kurių du yra transliacijos elongacijos veiksnio eEF1B α ir γ subvienetai, savo ypatybių dėka besigryninantys kartu su ribonukleaze P. Buvo tiriama, ar trečiasis baltymas galėtų būti homologiškas Pop1, didžiausiam RNazės P baltyminiam subvienetui. Buvo klonuotas genas, koduojantis baltymą, kurio sekoje esantys konservatyvūs elementai leido įvardyti jį kaip eukariotų RNazės P Pop1 baltymo homologą. Nustatyta, kad geną sudaro 5 egzonai ir 4 intronai. Nustatytas *T. pyriformis* RNazės P sandas – RNR subvienetas ir pasiūlytas jo antrinės struktūros modelis.