Testis-specific Expression of an Intronless Gene Encoding a Human Poly(A) Polymerase

Young Jae Lee1,2, Hana Kim1, Jae Hoon Chung2, and Younghoon Lee1,*

1 Department of Chemistry and Center for Molecular Design and Synthesis, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea;
2 Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea.

(Received February 12, 2001; Accepted March 26, 2001)

A mouse intronless gene, encoding a testis-specific poly(A) polymerase (mPAPT), was previously identified. mPAPT may play a role as a putative enzyme that is responsible for polyadenylation regulation during mouse spermatogenesis. In order to understand how PAPT genes are conserved in mammals, we isolated a human cDNA homolog encoding a human PAPT (hPAPT), which was specifically expressed in the testis. The structure of hPAPT was very similar to that of mPAPT. The about 100 residues at the C-terminal region of a nuclear poly(A) polymerase, PAP II, were missing in both PAPT proteins. An analysis of the genomic DNA showed that the hPAPT gene is an intronless gene that is similar to the mPAPT gene. Interestingly, the sequence homology between hPAPT and mPAPT was much lower than the homology between hPAP II and mPAP II. The phylogenetic analysis suggests that PAPTs arose through retrotransposition after the amphibian-amniote split during evolution.

Keywords: 3’-End Processing; Human Testis; Phylogenetic Analysis; Poly(A) Polymerase; Poly(A) Polymerization.

Introduction

Poly(A) polymerase (PAP) is a key protein in the mRNA 3’-end processing reaction. At least six components [cleavage- and polyadenylation-specificity factor (CPSF), cleavage-stimulation factor (CstF), two cleavage factors (CF I and CF II), PAP, and poly(A)-binding protein II (PABII)] are involved in the reaction (barabin and Keller, 1999; Colgan and Manley, 1997; Keller, 1995; Sachs and Wahle, 1993; Wahle, 1995). PAP adds the poly(A) tails of 200–250 residues at specific sites of pre-mRNAs that are introduced by CPSF, CstF, and CFs. PAP contains several functional domains: a catalytic core, an RNA-binding domain, two nuclear localization signals (NLS1 and NLS2), and a serine/threonine-rich domain (Martin and Keller, 1996; Raabe et al., 1994). Several forms of PAP have been identified from mammals and other organisms (Ballantyne et al., 1995; Gebauer and Richter, 1995; Lingner et al., 1991; Raabe et al., 1991; Wahle et al., 1991; Wilson, 1994; Zhao and Manley, 1996). Analyses of genomic DNA and cDNA clones encoding PAP in human, mouse, and bovine revealed the following: various forms of PAP are produced from the same PAP pre-mRNA by (1) alternative splicing or alternative polyadenylation in the mRNA level, and (2) multiple phosphorylation in the protein level (Zhao and Manley, 1996).

Recently, we identified a novel PAP gene from a mouse (Lee et al., 2000). The analysis of genomic and cDNA sequences of this PAP gene revealed that it is an intronless gene (Lee et al., 2000). Since the PAP gene showed testis-specific and developmental stage-specific expressions, we named it the mPAPT gene [mPAPT: mouse testis-specific poly(A) polymerase]. The same gene was also identified by an independent study by Kashiwabara et al. (2000), although it was termed the TPAP gene. The mPAPT of 641 residues lacks about 100-residues at the C-terminus compared to a nuclear poly(A) polymerase (mPAP II). The missing region comprises NLS2, one of two nuclear localization signals, and the consensus-phosphorylation sites of the serine/threonine-rich domain. The N-terminal part (residues 1 to 507) of mPAPT, which carries a catalytic core, an RNA binding domain, and an NLS1, shares a 91.3% amino acid identity with that of mPAP II. The C-terminal part of mPAPT (residues 508 to 641) is less conserved and has an amino...
acid identity of 65.4%. The highly dominant expression in testis and its developmental regulation (Kashiwabara et al., 2000; Lee et al., 2000) suggested that mPAPT may play specific roles in the male germ-cell development of mice. However, it is still unknown how PAPT genes are conserved in mammals. To answer this question, in this study, we isolated a cDNA encoding a human testis-specific PAP (hPAPT). Interestingly, the amino acid sequence comparison revealed that the homology of PAPTs between mice and humans is much lower than that of the PAP IIIs. Therefore, we also performed a phylogenetic analysis to assess the evolutionary divergence of PAP II and PAPT.

**Materials and Methods**

**Isolation of hPAPT cDNA clones** The human testis cDNA library (Clontech, USA) was screened with the random-primed mouse testis-specific poly(A) polymerase (mPAPT) cDNA with [α-32P]dCTP as a probe (Lee et al., 2000). Positive cDNA clones were isolated and subcloned into the EcoRI site of the pGEM-3Zf(+) vector (Promega, USA). The DNA inserts were sequenced using a Bigdye Terminator cycle sequencing kit (Perkin-Elmer).

**Nucleotide sequence accession number** The cDNA encoding a human testis-specific poly(A) polymerase (hPAPT) was deposited in GenBank (accession No. AF218840).

**PCR-amplification of DNA fragments containing the hPAPT gene** DNA fragments containing the hPAPT were amplified by PCR with the following primers: primer 1f (nt 1–24, 5'-GCA-GCTAGAACGGCCCTGCTCC-3') and primer 1r (nt 417–440, 5'-CCTTTTCGATGT AACTCAAACT-3'); primer 2f (nt 371–394, 5'-CTGTAATTGAAAACGTTGGAGGA-3') and primer 2r (nt 827–850, 5'-AAACACCTCGAGGAAACTTA-3'); primer 3f (nt 781–804, 5'-AGAGCCATCAAACCTGGGCAAAG-3') and primer 3r (nt 1237–1260, 5'-TTCTGTTGASTGCATCTGGCAGGAAG-3'); primer 4f (nt 1191–1214, 5'-CTTGGAAAGCTCAAGCGTTCCTTCTT-3') and primer 4r (nt 1647–1670, 5'-CTGTCGTCGGCCCTAATCTTGGGCACAG-3') and primer 5f (nt 1601–1624, 5'-CTCAGTCTACGTTCCAGGACAGA-3') and primer 5r (nt 2008–2031, 5'-CTCAGCTACCAACTGAGATTAC-3'). The PCR products were sequenced to examine the presence of introns, as described in Oh et al. (2000).

**Southern and Northern analyses** For Southern analysis, the human-genomic DNA was prepared from blood using the QIAGEN Genomic DNA Pre Kit. The human-genomic DNA of 10 µg were digested with EcoRI or HindIII, electrophoresed on an 0.8% agarose gel, and transferred to a Hybond N membrane (Amersham, UK). For Northern analysis, a human multi-tissue Northern-hybridization filter that contained approximately 2 µg of poly(A)+ RNAs of the brain, heart, kidney, stomach, testis, and placenta (purchased from OriGene, USA) was used. The probe was prepared by random-priming of the hPAPT cDNA (nucleotides 285 to 1690) with [α-32P]dCTP. Hybridization was carried out at 63°C with a QuickHyb solution (Stratagene, USA) according to the manufacturer’s instruction.

**Construction of phylogenetic tree** Amino acid sequences of PAPs of *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Xenopus*, chicken, mice, and humans were retrieved from public databases. A multiple alignment of amino acid sequences was generated using the ClustalW program (Thompson et al., 1994). The phylogenetic tree (Perriere and Gouy, 1996) was constructed based on the ClustalW alignment using *S. cerevisiae* PAP as an outgroup.

**Results and Discussion**

**Cloning of a cDNA encoding a testis-specific poly(A) polymerase and analysis of its genomic DNA** To isolate cDNAs encoding a human homolog of mouse testis-specific poly(A) polymerase (mPAPT), a human testis cDNA library was screened with the mPAPT cDNA as a probe. Among positive clones, one cDNA clone, λ (hPAPT4), was isolated as a clone carrying a putative human PAP (hPAPT) cDNA that contained the full, open-reading frame (Fig. 1). The hPAPT cDNA in λ (hPAPT4) has a short 3'-UTR. This is probably because the human testis cDNA library used in this study was constructed with an XhoI linker, and there was an XhoI site 8 bp downstream of the termination codon of cDNAs coding for hPAPT. The λ(hPAPT4) clone was chosen for a cDNA encoding hPAPT by three criteria. (1) It differed from the human PAP II cDNA in the nucleotide sequences. (2) It could not be generated from a common PAP pre-mRNA by alternative splicing or alternative polyadenylation. (3) The structure of its deduced protein sequence was very similar to that of mouse PAPT since about 100 residues at the C-terminal region were missing, as compared to that of human PAP II (hPAPT II). Like mPAPT, the hPAPT of 636 amino acids carries a catalytic core, an RNA binding region, and a putative nuclear localization signal (NLS1). It also lacks the second putative nuclear localization signal (NLS2) and a part of the serine/threonine-rich domain of human PAP II, due to the absence of about 100 residues at the C-terminus. There is a 86.1% identity in the deduced amino acid sequence of hPAPT with hPAP II (Fig. 1C). This degree of identity is very similar to that of the identity of mPAPT with mPAP II (85.7%). However, hPAPT showed an 86.6% identity with mPAPT. This relatively low identity is very unusual because hPAP II shares a 95.3% identity with mPAP II. The differences were mostly found in the C-terminal region (residues 508 to 636) of hPAPT. The 5'-UTR, including the translation-initiation region of hPAPT mRNA
Figure 1. Continuous
Fig. 1. Characterization of an hPAPT cDNA. A. The restriction map of the hPAPT cDNA in λ (hPAPT4). The translation start codon (ATG) and stop codon (TAG) of hPAPT are indicated. The PstI fragment used as a hybridization probe for Southern and Northern analyses is represented by the black bar. B. Nucleotide and deduced amino acid sequences of hPAPT. The deduced amino acid sequence is written below the nucleotide sequence. C. Multiple-sequence alignment of PAPs and PAPTs. Alignment was performed using the ClustalW program. The amino acids identical with those of bovine PAP II are represented with dots. The amino acid sequences of chicken PAP (gPAP), corresponding to residues 1–3 and 62–119 of bovine PAP II, are not yet determined. The GenBank accession numbers of the sequences being compared are: bPAPII, bovine poly(A) polymerase type II, X61585; hPAPII, human poly(A) polymerase type II, X76770; mPAPII, mouse poly(A) polymerase type II, L22658; chicken poly(A) polymerase, gPAP AF061758; mPAPT, mouse testis-specific poly(A) polymerase, AF039957; hPAPT, human testis-specific poly(A) polymerase; xPAP, Xenopus poly(A) polymerase type I, U19973; cPAP, C. elegans poly(A) polymerase, Z47356; yPAP, S. cerevisiae poly(A) polymerase, P29468. Some parts of the amino acid sequences are compensated by ESTs. D. Schematic diagram showing significant features of PAPII and PAPT. The conserved N-terminal, and less-conserved C-terminal regions, are shown as the open and hatched bars, respectively. The solid bars represent the nuclear localization signal sequences (NLS). The catalytic core and RNA binding regions are also shown. D, indicates the conserved aspartate.

showed CpG island features (Fig. 1B), which were also observed in mPAPT mRNA. The CpG island of the hPAPT gene may be involved in its testis-specific and/or developmental expression.

The copy number of the hPAPT gene in human genome was determined by Southern analysis using BamHI and HindIII (Fig. 2). A single and double bands were detected in the EcoRI-digested and HindIII-digested DNA lanes, respectively. Since the appearance of the double bands of the HindIII-digested DNA resulted from the presence of a
HindIII restriction site in the middle of the hPAPT-coding sequence (Fig. 1A), the hPAPT gene is a single-copy gene in the human genome.

When the DNA fragments, obtained by PCR-amplification with human-genomic DNA as a template, were sequenced and their sequences compared with the cDNA sequence, no intron was found. This result suggests that the hPAPT gene is also an intronless retrogene like the mPAPT gene (Lee et al., 2000). Some functional retrogenes, Zfa, G6pd-2, Pgk-2, and Pdha-2, compensate for lost functions of their isotype genes in spermatogenic cells (Ashworth et al., 1990; Hendriksen et al., 1997; McCarrey et al., 1992; Takakubo and Dahl, 1992). Retrogenes are presumed to originate from the retrotransposition of mRNAs containing a functional enhancer and/or promoter sequences (Chakrabarti et al., 1995; Vanin, 1985).

Expression of hPAPT  The expression pattern of the hPAPT gene in various human tissues was investigated by a Northern hybridization analysis (Fig. 3). The mRNAs that were isolated from the human brain, heart, kidney, stomach, testis, and placenta were hybridized with the hPAPT cDNA probe. The highly predominant bands of 2.8 and 4.5 kb as two major transcripts, and 8.0 kb as a minor transcript, were observed in the testis on a human multiple-tissue blot. Since the 4.5 kb transcript was, although barely, detected in other tissues, the 4.5 kb signal could include a cross-hybridization signal derived from hPAP II or hPAP II-related transcripts. We performed a Northern analysis, under the condition where mPAP II mRNA signals were undetected with the mPAPT cDNA probe (Lee et al., 2000). Since a sequence homology between hPAPT and hPAP II is similar to that of mPAPT and mPAP II, we expected that hPAP II or hPAP II-related transcripts would not be cross-hybridized with the human cDNA probe. The fact that no weak bands, besides the strong bands resulting from a single hPAPT gene, were observed in the Southern blot (Fig. 2) supports this expectation. Therefore, two major transcripts of 2.8 and 4.5 kb seem to be generated, but with different 3′-untranslated regions, from the same human PAPT pre-mRNA. However, the origin of the minor transcript of 8.0 kb still remains unclear because of its large size. In any case, the result of Fig. 3 confirmed that hPAPT is specifically expressed in the human testis as mPAPT in mouse testis. The testis-specific expression of both hPAPT and mPAPT further supports that PAPT plays important roles in the testis (Kashiwabara et al., 2000; Lee et al., 2000).

Phylogenetic analysis of PAPs  Poly(A) polymerase (PAP) is an essential protein for the mRNA 3′-end processing reaction in eukaryotes (Barabino and Keller, 1999; Colgan and Manley, 1997; Keller, 1995; Sachs and Wahle, 1993; Wahle, 1995) and the functional domains of the protein are highly conserved (Martin and Keller, 1996; Raabe et al., 1994). It is very interesting that the homology between hPAPT and mPAPT is much lower than the homology between hPAP II and mPAP II. Therefore, we examined how PAPs, including PAPTs, are phylogenetically related. Based on the alignments using full-length PAP amino acid sequences (Fig. 1C), we constructed a phylogenetic tree to see their evolutionary relationships (Fig. 4). The phylogenetic tree suggests that the retrotransposition of PAPT postdates the amphibian-amniote split. Since the first amniotes allegedly appeared about 350 million years ago (Wallace et al., 1996), the
Fig. 4. Evolutionary relationships of PAPs and PAPT s. The phylogenetic tree was constructed based on the ClustalW alig nment using S. cerevisiae PAP as an outgroup (Fig. 1C). The regions of the chicken PAP, corresponding to residues 1–3 and 62–119 of the bovine PAP II, are excluded in the ClustalW alignment since the amino acid sequences of the regions are not yet available. The estimated genetic distance between the sequences is proportional to the length of the horizontal lines connecting the sequences. The number at each branch node indicates the local bootstrap probability.

retrotransposition of PAPT appears to be a much more ancient event than that of usual retrogenes, which supposedly occurred during the past 30 million years (Vanin, 1985). The lower identity of hPAPT and mPAPT results from a relatively high-rate of divergence.

Acknowledgments This work was supported by a grant from the Molecular Medicine Research Group Program (00-J03-01-01-A-05) from the Ministry of Science and Technology, and in part by the Brain Korea 21 Project in 2000.

References


