Original Articles

Cloning and Expression of *PARP-3 (Adprt3)* and *U3-55k*, Two Genes Closely Linked on Mouse Chromosome 9

(poly(ADP-ribose) polymerase / PARP / *Adprt* / *ADPRTL* / PARP-1 / PARP-2 / hU3-55k / Rrp9p / U3 snoRNA / U3 snoRNP / WD-repeat protein / bi-directional promoter / tissue-specific gene expression / orthologue)

P. URBÁNEK, J. PAČES, J. KRÁLOVÁ, M. DVOŘÁK, V. PAČES

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Abstract. Post-translational modification of nuclear proteins by poly(ADP-ribose) polymerase 1 (PARP-1) is involved in the regulation of DNA repair, cell death, and maintenance of genomic stability. Recently, several PARP-1 homologues have been identified constituting a family of poly(ADP-ribosyl)ating proteins. We cloned and sequenced the cDNAs of the mouse *PARP-3* (*Adprt3*) gene encoding poly(ADP-ribose) polymerase 3 and of the closely linked *U3-55k* gene coding for the U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. The two genes are located in a head-to-head orientation on mouse chromosome 9 and are linked by an approximately 1.5-kb putative bi-directional promoter region. This gene arrangement is conserved between mouse and human orthologues. Three alternative non-coding 5'-end exons were found in the mouse PARP-3 mRNA. The expression patterns of *PARP-3*, *U3-55k*, *PARP-2*, and *PARP-1* genes were determined using Northern blot with mRNA from various adult mouse tissues and organs. *PARP-3* expression was found to be regulated in a tissue-specific manner. The highest expression of *PARP-3* was detected in the skeletal muscle, high to moderate levels were found in the lung, liver, kidney, ovary, spleen and heart, while thymus, small intestine and colon contained lower levels of the *PARP-3* transcripts. Notably, *PARP-3* expression was barely detectable in the whole brain and testis mRNA. In contrast to *PARP-3*, the other three genes showed ubiquitous expression with less variable mRNA levels. Interestingly, the mouse and human *PARP-2* gene has recently been shown to be connected via a bi-directional promoter with the gene for the RNase P RNA subunit (Amé et al., *J. Biol. Chem.* 276: 11092-11099, 2001). As both the U3-55k protein and the RNase P RNA are involved in the processing of precursor RNAs of the protein-synthesizing machinery (pre-rRNA and pre-tRNA, respectively), it is tempting to hypothesize that expression of some members of the two groups of genes (i.e. *PARP* vs. protein-synthesizing machinery RNA-processing genes) may be coordinately regulated under certain physiological or pathological conditions and/or in some cell types.

Mammalian poly(ADP-ribose) polymerase 1 (PARP-1) is an abundant nuclear enzyme consisting of three main domains: the N-terminal DNA-binding domain, the C-terminal PARP catalytic domain and the automodification domain, located in the central part of the protein. PARP-1 binds to DNA-strand breaks in response to DNA damage induced by ionizing radiation or alkylating agents. Binding to DNA activates its catalytic domain, leading to the synthesis of branched ADP-ribose homopolymers on various nuclear proteins, including the automodification domain of the PARP-1 molecule itself (for reviews see de Murcia and Ménissier-de Murcia, 1994; Lindahl et al., 1995; D'Amours et al., 1999; Herceg and Wang, 2001). Numerous *in vitro* and *in vivo* studies, including analysis of *PARP-1* knock-out mice and cells, have documented functions of PARP-1 in DNA repair, replication, transcription, cell proliferation and death, as well as in maintaining the genomic integrity and suppressing tumorigenesis (for reviews see de Murcia and Ménissier-de Murcia, 1994; Lindahl et al., 1995; Jeggo, 1998; Le Rhun et al., 1998; D'Amours et al., 1999; Jacobson and Jacobson, 1999; Shall and de Murcia,

Received June 28, 2002. Accepted July 17, 2002.

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, grant No. A5052802, and in part by the Center for Integrated Genomics and by the Grant Agency of the Czech Republic, grant No. 204/00/0554.

Corresponding author: Pavel Urbánek, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 37 Prague, Czech Republic. Tel.: +420-2-20183404; fax: +420-2-24310955; e-mail: urbanek@img.cas.cz.

Abbreviations: bp – base pair(s), EST – expressed sequence tag, kb – kilobase(s), nt – nucleotide(s), 3'UTR – 3'-end untranslated region, gb:/emb:/dbj: – accession number in the GenBank/EMBL/DDBJ sequence databases, respectively, ref: – accession number in the RefSeq database of the National Center for Biotechnology Information (NCBI).

2000; Bürkle, 2001; Herceg and Wang, 2001; Smith, 2001; Tong et al., 2001).

Although *PARP-1*$^{-/-}$ mice showed genomic instability and were highly sensitive to genotoxic stress, they were viable and fertile (Wang et al., 1995; Ménissier-de Murcia et al., 1997; Wang et al., 1997; Masutani et al., 1999). This relatively mild mutant phenotype, as well as the detection of ADP-ribose polymer synthesis in the PARP-1$^{-/-}$ cells (Shieh et al., 1998) and the discovery of two PARP genes in higher plants encoding two structurally different PARP homologues (Lepiniec et al., 1995; Babiychuk et al., 1998; Mahajan and Zuo, 1998), suggested the presence of additional *PARP* genes in mouse that might functionally substitute for the lack of PARP-1 activity. Over the past three years several mammalian genes encoding proteins homologous to the PARP-1 catalytic domain were identified, constituting a family of poly(ADP-ribosyl)ating proteins (reviewed in Smith, 2001; see also Kaminker et al., 2001; Kuimov et al., 2001; Lyons et al., 2001; Ma et al., 2001; Monz et al., 2001). The least studied member of the *PARP* family is *PARP-3* (*ADPRTL3*), characterized so far only in humans (Johansson, 1999).

Here we report the cDNA sequence of the mouse *PARP-3* (*Adprt3*) gene and of the closely linked *U3-55k* gene that we identify as the mouse orthologue of the human gene encoding the hU3-55k protein (Pluk et al., 1998). hU3-55k is a WD-repeat-containing 55-kilodalton protein associated with the U3 small nucleolar ribonucleoprotein complex (U3 snoRNP) that is involved in the processing of the primary pre-ribosomal RNA transcript (Lübben et al., 1993; Pluk et al., 1998; for review on snoRNPs see Maxwell and Fournier, 1995). Its orthologues in *Saccharomyces cerevisiae* (designated Rrp9p) and *Xenopus laevis* were characterized recently (Lukowiak et al., 2000; Venema et al., 2000). The main structural feature common to these three proteins is the six (man, frog) or five (yeast) WD-repeat sequence motifs (Pluk et al., 1998; Lukowiak et al., 2000; Venema et al., 2000; for review on WD-repeat proteins see Smith et al., 1999), which are probably required for association with the U3 snoRNA (Lukowiak et al., 2000).

As the mouse *PARP-3* and *U3-55k* genes appear to be linked via an approximately 1.5-kb bi-directional promoter region, it was of interest to determine their expression patterns in various mouse tissues and organs. Here we show that *PARP-3* gene expression is tissue-specifically regulated, in contrast to the neighbouring *U3-55k* gene and also to the homologous *PARP-1* and *PARP-2* genes.

Material and Methods

Oligonucleotide primers

The sequences are written in the 5' to 3' direction, **(s)** indicates sense and **(as)** anti-sense relative to the corresponding cDNA (see below): **1 (as)**, cacccgagtgtggcatgat; **2 (as)**, agcccactggtgaggatggc; **3 (as)**, ccacattggtgccgtgccacag; **4 (s)**, ctgtggcacggcaccaatgtgg; **5 (as)**, gggcaagcagcttagaggtgaat; **6 (s)**, ctcatatacaaggagagccagtg; **7 (as)**, ggtgtcattaggtgatctaagct; **8 (s)**, caatgttctggttagggatggac; **9 (as)**, gcaccattcacaacaaccaactc; **10 (s)**, gggtagcggcaccttgaacag; **11 (as)**, aagggagagtgtcaggagccg; **12 (s)**, atgatggctctgtggccttgtg; **13 (as)**, ggaggctttaatacaaaagaggac; **14 (s)**, gaagggtcgctagatgggatc; **15 (s)**, ccagaactcaggactgttagc; **16 (s)**, tttctgctgcctcggggaacac; **17 (as)**, cctcggcagtggaccggaag; **18 (as)**, gagccctcagtctgcacagag; **19 (as)**, gctggctttgcccgcttacga; **20 (s)**, tctggaaggcgagtgctaaatg; **21 (as)**, gttctaggccttggcgctctg; **22 (s)**, ggaaaccgacacgttagcggag; **23 (as)**, ccagtacagtaataagcgtcgct.

Total RNA and poly(A)$^+$ RNA isolation

Total RNA was isolated from mouse tissues and organs using the TRIzol Reagent (Life Technologies, Inc. – Gibco BRL, Gaithersburg, MD). Whole organs of three months old mice (kidney, brain, heart, spleen, thymus, lung, ovary, testis) or their parts (liver, skeletal muscle, small intestine, colon) were used. The skeletal muscles were from hindlegs. Ovaries were dissected free of oviduct, fat pad and bursa. Testes were taken free of fat pad, epididymis and vas deferens. Poly(A)$^+$ RNA was isolated using the PolyATtract mRNA Isolation System III (Promega Corporation, Madison, WI).

cDNA and genomic DNA cloning and sequencing

Mouse PARP-3 cDNA was isolated by a combination of the rapid amplification of 5' cDNA-ends (5'-RACE) (Frohman et al., 1988) and reverse transcription – polymerase chain reaction (RT-PCR). For the 5'-RACE, three oligonucleotides (Nos. 1, 2 and 3) were designed according to the mouse expressed sequence tag (EST) (accession number gb:AI019500; Marra et al., 1999) highly similar to the human PARP-3 cDNA sequence (gb:AF083068) (Johansson, 1999). Oligonucleotide 1 was used to prime reverse transcription of mouse skeletal muscle and heart total RNA with the SuperScript II RNase H$^-$ Reverse Transcriptase (Life Technologies, Inc. – Gibco BRL, Gaithersburg, MD). Oligonucleotides 2 and 3 were used for the first and the second round of amplification, respectively. The first round consisted of initial annealing at 50°C for 2 min and extension at 72°C for 20 min, followed by 25 cycles: 40 s at 94°C, 1 min at 60°C, and 3 min at 72°C. The second round consisted of 25 cycles: 30 s at 95°C, 15 s at 66°C, and 3 min at 68°C. Both PCR amplifications were performed with the Advantage 2 Polymerase Mix (CLONTECH Laboratories, Inc., Palo Alto, CA). PCR products were fractionated by agarose gel electrophoresis, cloned with the pGEM-T Vector System (Promega Corporation, Madison, WI), and sequenced. Two cDNA forms with different 5'-end sequences were isolated initially, suggesting the existence of two alternative 5'-end exons (designated 1a and 1c) in the mouse PARP-3 mRNA. This result was confirmed by the

alignment of the cDNAs with the genomic sequence of the promoter region (Fig. 1). A third alternative 5'-end exon (designated 1b) was inferred from the alignment of the mouse EST sequence (gb:AW318917; Marra et al., 1999) and the genomic sequence (Fig. 1). The existence of the three alternative exons was further verified by RT-PCR, cloning and sequencing of the three cDNA fragments spanning over the exon 1 – exon 2 junction. The first strand cDNA for RT-PCR was prepared from skeletal muscle total RNA using the cDNA Cycle Kit (Invitrogen Corporation, Carlsbad, CA) and oligo(dT) primers. Fifty ng of the cDNA were then used as a template for PCR amplification with Taq DNA Polymerase (Sigma Chemical Company, St. Louis, MO) and primer pairs 14/17, 15/17 or 16/17 (35 cycles: 30 s at 95°C, 30 s at 60°C, and 2 min at 72°C). Primers 14, 15 and 16 are specific for the alternative exons 1a, 1b and 1c, respectively, while primer 17 is located in exon 2 common to all the three splice variants (Fig. 1).

RT-PCR was performed under the same conditions in order to clone the 3'-terminal part of the PARP-3 cDNA and the full-length U3-55k cDNA. Oligonucleotides 4 and 5 that were derived from mouse ESTs (gb:AI019500 and gb:AI183257; Marra et al., 1999) homologous to the human PARP-3 cDNA sequence were used to amplify the 3'-terminal part of the PARP-3 cDNA coding region. The 3'-end untranslated region (3'UTR) of mouse PARP-3 was amplified with the oligonucleotides 6, 7, 8 and 9 designed according to the mouse ESTs located in the 3'-end part of the PARP-3 transcript (gb:AI183257, gb:AA197532, gb:AA177323; Marra et al., 1999). Similarly, the coding region of the U3-55k cDNA was amplified with the oligonucleotides 10 and 11, and the 3'UTR of this gene with the oligonucleotides 12 and 13. These U3-55k-specific oligonucleotides were designed according to the mouse EST sequences (gb:AA409391, gb:AI447235; Ko et al., 1998; Marra et al., 1999) found by homology to the human U3-55k cDNA (emb:AJ001340) (Pluk et al., 1998). The 1454-bp genomic fragment spanning the promoter region of the mouse *PARP-3* and *U3-55k* genes was PCR-amplified with the oligonucleotides 18 and 19, derived from the second *PARP-3* exon and from the first *U3-55k* exon, respectively. All the PCR-amplified cDNA and genomic DNA fragments were fractionated by agarose gel electrophoresis, cloned and sequenced. All the nucleotide sequences originally determined in the cloned fragments were verified by direct sequencing of PCR fragments independently amplified from the cDNA (hindleg skeletal muscle) or genomic DNA (spleen) of the C57BL/6J inbred mice.

Northern blot analysis

Poly(A)$^+$ RNAs isolated from various tissues of C57BL/6J x 129/Sv F1 hybrid mice were fractionated by electrophoresis in 1% agarose-formaldehyde gel (3 μg of RNA per lane) and blotted onto the GeneScreen Hybridization Transfer Membrane (NEN Life Science Products, Inc., Boston, MA) according to the standard protocols (Sambrook et al., 1989). For each tissue, a pool of equal aliquots of RNA isolated independently from three animals of the same age and sex was used. Northern blot was hybridized overnight with the ^{32}P-labelled mouse PARP-3 cDNA probe in the ULTRAhyb Ultrasensitive Hybridization Buffer (Ambion, Inc., Austin, TX) and washed under high stringency conditions according to the manufacturer's protocol. Autoradiography was performed at –70°C using Kodak BioMax MS film and intensifying screen (Eastman Kodak Company, Rochester, NY). The membrane was then successively rehybridized with the mouse cDNA probes of U3-55k, PARP-2, and PARP-1. After each round of hybridization the probe was stripped from the membrane by boiling for 3–5 min in the solution containing 15 mM sodium chloride, 1.5 mM sodium citrate, pH 7.0 (i.e. 0.1x SSC), and 1% sodium dodecyl sulphate (SDS), and the completeness of the probe removal was checked by autoradiography.

Probes used for Northern blot hybridization

The following cloned mouse cDNA fragments were used: PARP-3 (nt 335–1316, i.e. starting from the underlined *Sma*I half-site in the sequence 5'-gggatacaggtccac-3' to the oligonucleotide 3 inclusive, gb:AF368233 – this work); U3-55k (nt 577–1461, i.e. starting at the underlined *Xho*I half-site in the sequence 5'-tcgagccaagaaggg-3' to the oligonucleotide 11 inclusive, gb:AF368232 – this work); PARP-2 (nt 43–1009, i.e. oligonucleotide 20 through oligonucleotide 21, emb:AJ007780) (Amé et al., 1999); PARP-1 (nt 22–1022, i.e. oligonucleotide 22 through oligonucleotide 23, emb:X14206) (Huppi et al., 1989). PARP-2 and PARP-1 cDNA fragments were prepared by RT-PCR amplification using the indicated oligonucleotides, cloned and verified by sequencing as described above for PARP-3 and U3-55k cDNAs. For hybridization, the cDNA fragments were labelled by nick translation with [α-^{32}P]dCTP according to Sambrook et al. (1989).

Database homology searches and sequence comparisons

The BLAST programmes (Altschul et al., 1990; Altschul et al., 1997) at the National Center for Biotechnology Information (NCBI) (Wheeler et al., 2002) (http://www.ncbi.nlm.nih.gov/BLAST/) were used to search for the PARP and U3-55k homologous sequences in the NCBI databases (nr, est, human and mouse genomic). The sequences referred to in this work and accessed through the NCBI Entrez system (Wheeler et al., 2002) originated from the following databases: GenBank (Benson et al., 2002), EMBL (Stoesser et al., 2002), DDBJ (Tateno et al., 2002) and NCBI RefSeq database (Pruitt and Maglott, 2001) (see also the