

U87 RNA, a novel C/D box small nucleolar RNA from mammalian cells

Irina K. Gogolevskaya^a, Julia A. Makarova^a, Larisa N. Gause^b, Valentina A. Kulichkova^c,
Irina M. Konstantinova^c, Dmitri A. Kramerov^{a,*}

^aEngelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov Street, Moscow 119991, Russia

^bKol'tsov Institute of Developmental Biology, Russian Academy of Sciences, 26 Vavilov Street, Moscow 117808, Russia

^cInstitute of Cytology, Russian Academy of Sciences, 4 Tikhoretsky Avenue, St. Petersburg 194064, Russia

Received 11 February 2002; received in revised form 3 April 2002; accepted 7 May 2002

Received by E. Sverdlov

Abstract

A novel 72 nt small nucleolar RNA (snoRNA) called U87 was found in rat liver cells. This RNA possesses the features of C/D box snoRNA family: boxes C, D', C', D, and 11 nt antisense element complementary to 28S ribosomal RNA (rRNA). The vast majority of C/D box snoRNAs direct site-specific 2'-O-ribose methylation of rRNAs. U87 RNA is suggested to be involved in 2'-O-methylation of a G₃₄₆₈ residue in 28S rRNA. U87 RNA was detected in different mammalian species with slight length variability. Rat and mouse U87 RNA gene was characterized. Unlike the majority of C/D box snoRNAs U87 RNA lacks the terminal stem required for snoRNA processing. However, U87 gene is flanked by 7 bp inverted repeats potentially able to form a terminal stem in U87 RNA precursor. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Rodents; Alpha-ribonucleoprotein; Low-molecular weight RNA; Small nucleolar RNA gene; RNA processing; Ribosomal RNA methylation; Terminal stem

1. Introduction

Small nucleolar RNAs (snoRNAs) constitute a large class of RNAs found in eukaryotic cells (reviewed in Maxwell and Fournier, 1995; Smith and Steitz, 1997). Some of the snoRNAs are directly required for various cleavages of pre-ribosomal RNA (rRNA), but most of them are involved in the site-specific modifications of nucleotides of the pre-rRNA (Tollervey, 1996; Smith and Steitz, 1997). The vast majority of snoRNAs fall into two families (box C/D and H/ACA) distinguished by the common sequence boxes (Balakin et al., 1996). Members of the box C/D snoRNA family are involved in 2'-O-methylation of pre-rRNA precursor (Kiss-Laszlo et al., 1996) while the box H/ACA RNAs are required for pseudouridylation of pre-rRNA (Ganot et al., 1997).

The 2'-O-methylation guide snoRNAs possess the conserved C (consensus RUGAUGA) and D (CUGA)

boxes. Internal regions of box C/D snoRNAs carry imperfect copies of the C and D boxes, called C' and D' boxes (Kiss-Laszlo et al., 1998). The D and/or D' boxes are preceded by long (10–21 nt) recognition sequences that form double helices with rRNA sequences and, thus, position the D or D' box of the snoRNA exactly 5 nt from the 2'-O-methylated ribosomal nucleotide (Bachelierie and Cavaille, 1997).

Although C and D boxes are generally distant from each other in the primary snoRNA sequences, they are brought into proximity in the folded RNAs as a result of base pairing between the complementary sequences flanking the box elements. The resulting structure, consisting of box C and box D as well as one or two adjacent helices, has been referred to as the stem-box structure (Qu et al., 1995) or the C/D motif (Samarsky et al., 1998) and is required for the correct processing, the nucleolar localization (Narayanan et al., 1999) and the metabolic stability (Huang et al., 1992) of snoRNA. In addition, the motif can be a protein-binding signal, and the proteins with the affinity to this motif are believed to protect snoRNA from exonucleolytic degradation (Lafontaine and Tollervey, 1999; Lyman et al., 1999). Indeed, the C/D motif is required for binding of

Abbreviations: DEAE, diethylaminoethyl cellulose; EST, expressed sequence tag; IR, inverted repeats; PA, polyacrylamide; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; RNP, ribonucleoprotein; snoRNA, small nucleolar RNA

* Corresponding author. Tel.: +7-95-135-9997; fax: +7-95-135-1405.

E-mail address: kramerov@genome.eimb.relarn.ru (D.A. Kramerov).

an essential nucleolar protein, fibrillarin, common for this family of snoRNAs (Caffarelli et al., 1998).

Genes coding for snoRNAs have been found in diverse genomic contexts (reviewed by Maxwell and Fournier, 1995; Weinstein and Steitz, 1999). In vertebrates, the vast majority of C/D box snoRNAs are encoded within introns of a host gene, which quite often encodes a protein involved in translation or ribosome biogenesis (Liu and Maxwell, 1990), although several genes with introns containing snoRNA genes and non-coding exons have been found (Tycowski et al., 1996).

The starting point for this work was our investigation of alpha-RNA from alpha-ribonucleoprotein (RNP) particles. Alpha-RNP was found in mammalian cells as a RNP fraction tightly bound to chromatin (Konstantinova et al., 1984, 1987). Alpha-RNP carries some proteasome antigens and a set of unidentified low-molecular RNAs (Konstantinova et al., 1988). In the course of characterization of RNA isolated from alpha-RNP we surprisingly found a novel small RNA of the C/D box snoRNA family and named it U87. This RNA was shown to carry C/D box snoRNA sequence motifs and was localized in the nucleolus where it can mediate 2'-O-methylation of G₃₄₆₈ in 28S rRNA.

2. Materials and methods

2.1. Cell fractionation and isolation of RNA

Nuclear alpha-RNP were isolated from rat liver chromatin as an acid-soluble fraction (Konstantinova et al., 1977). RNA was isolated from alpha-RNP by proteinase K treatment and phenol extraction as described elsewhere (Konstantinova et al., 1987). Rat liver cells were separated into cytoplasmic, nucleolar, and nucleoplasmic fractions (Muramatsu et al., 1974). Total RNA from human HeLa cells or liver of animals was isolated by guanidine thiocyanate/phenol-chloroform extraction (Chomczynski and Sacchi, 1987).

2.2. Preparation of cDNA library

Alpha-RNA was polyadenylated using *Escherichia coli* poly(A) polymerase. The ³²P-labeled cDNA was synthesized with M-MLV reverse transcriptase and *Xba*I(T)₁₅ primer: 5'-GTCTGACTCTAGA(T)₁₅. After purification by electrophoresis in 4% NuSieve agarose gel and electroelution on a diethylaminoethyl cellulose (DEAE) membrane, the 3'-end of cDNA was tailed with poly(dG) using terminal deoxynucleotidyl transferase. cDNA was then amplified by 30 polymerase chain reaction (PCR) cycles with primers *Xba*I(T)₁₅ and *Eco*RI(C)₁₀: 5'-CGGAATTCGT(C)₁₀. The resulting double-stranded cDNA was digested with *Xba*I and *Eco*RI, purified by electrophoresis in 4% NuSieve agarose gel, and cloned into *Xba*I-*Eco*RI sites of pGEM7Z. The library was screened by hybridization with ³²P-labeled first strand of alpha-cDNA.

2.3. Cloning the full-length cDNA

Total liver RNA (150 µg) was fractionated on 6% denaturing polyacrylamide (PA) gel and the material from the gel zone containing U87RNA (according to size markers) was transferred onto a DEAE membrane by semi-dry electroblotting. Then RNA was eluted from the appropriate piece of the DEAE membrane in 1 M NaCl, 10 mM Tris-HCl, pH 7.5, and 1 mM ethylenediaminetetra-acetic acid at 60°C for 20 min. The isolated RNA was used for cDNA synthesis followed by cloning as described above for alpha-RNA. The cDNA library was screened by hybridization to U87-specific probe.

2.4. Preparation of a probe and hybridization

A 55 bp fragment of U87 cDNA was amplified by PCR. The longest of the originally isolated U87 cDNA clones was added to the reaction as a template. Oligonucleotides U87 DIR (5'-ATGTTTTTGCCGTTTACC) and U87 REV (5'-GCTCAGTCTTAAGATTCT) were used as primers. The amplified fragment was isolated from 5% NuSieve agarose gel after electrophoresis. The isolated DNA fragment (0.5%) was added to the PCR mixture containing 50 µCi alpha-[³²P]dATP. After 17 cycles of PCR the labeled PCR product was purified by ethanol precipitation. This probe was used in all hybridization experiments. Hybridization was carried out in 4 × 0.15 M NaCl/0.015 M Na₃-citrate as previously described (Serdobova and Kramerov, 1998).

2.5. RNA analyses

For Northern analyses 50–100 µg of total, nuclear, nucleolar, nucleoplasmic, or cytoplasmic RNA was separated on a 20 cm 6% denaturing PA gel and electroblotted onto a Hybond-N membrane. U87 RNA size was determined in a similar way except for using a sequencing gel electrophoresis. Primer-extension analysis (Kingston, 1992) of the 5'-termini of rat U87 RNA was performed using 5'-end labeled oligonucleotide U87 REV as a primer, and the sample was fractionated on a 6% sequencing PA gel.

2.6. U87 gene cloning

In order to determine the exact size of the DNA fragment containing U87 gene, 10 µg of rat or mouse genomic DNA was digested with *Hind*III and *Eco*RI, fractionated on a 0.8% agarose gel, blotted onto a nylon membrane, and hybridized to U87-specific probe. The genomic DNA fraction enriched in U87 gene was isolated using 0.8% agarose gel electrophoresis of 100 µg rat or mouse DNA digested with the same enzymes. The fraction (3.0 and 3.5 kb for rat and mouse, respectively) was electroeluted from 0.8% agarose gels using DEAE membrane and cloned in pSL1190 plasmid. Screening of 4–5 × 10⁴ clones with U87 probe yielded two and one positive rat and mouse clones, respectively.

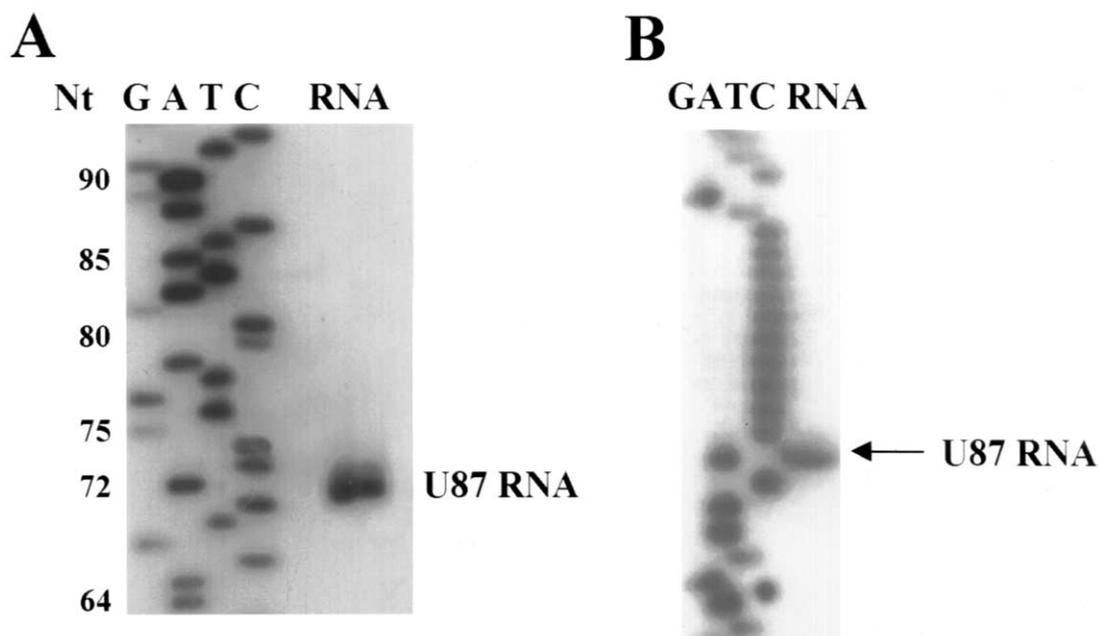


Fig. 1. Characterization of rat U87 snoRNA. (A) Measuring rat U87 RNA size. Rat liver nuclear RNA was fractionated on a 6% sequencing gel, blotted onto a membrane, and hybridized to the U87-specific probe ('RNA' lane). Sequencing lanes of a known DNA was used as a size marker. (B) Primer extension analysis of the 5'-terminus of U87 RNA. Terminally labeled oligonucleotide (U87 REV) specific for U87 RNA was annealed to rat liver RNA (<12S) and extended with reverse transcriptase ('RNA' lane). Lanes G, A, T, and C represent dideoxy sequencing reaction using the same oligonucleotide as a primer and plasmid carrying the presumably full-length U87 cDNA as a template.

2.7. Sequencing and computer analysis

Double-stranded plasmid templates were sequenced using the dideoxynucleotide method with Sequenase 2.0 (USB-Amersham). The BLAST search in GenBank was conducted using web-service of NCBI (NIH, USA).

3. Results and discussion

Characterization of a cDNA library of rat alpha-RNP allowed us to reveal a novel small RNA named U87 RNA in addition to several known small RNAs (e.g. ID, 4.5 S_I, 4.5 S_H, and U1A). Length difference of the corresponding nucleotide sequences in two clones (40 and 59 bp) suggested that an incomplete sequence of this RNA had been cloned. To overcome this problem the fraction enriched in U87 RNA was isolated from total rat liver RNA by polyacrylamide gel electrophoresis (PAGE) and subjected to cDNA synthesis and PCR amplification. The PCR products were cloned, the clones containing U87 RNA sequence were selected by hybridization and sequenced. Three such clones carried identical sequences 72 bp long (GenBank, acc. no. AF272707). The exact size of U87 RNA was determined by high-resolution PAGE followed by Northern-blot hybridization, and primer-extension analysis was performed for this RNA (Fig. 1). Results of these experiments confirmed that the cloned sequence of U87 RNA was full-sized. Additionally, 3 kb *Hind*III/*Eco*RI frag-

ments of rat genomic DNA containing U87 locus were cloned and sequenced (acc. no. AF396685). Fig. 2 shows the part of this fragment with the U87 RNA-coding region bolded. U87 RNA possesses C/D box snoRNA sequence elements (Fig. 2): the C, C', D' and D boxes with single substitution in box C' (this box usually represents a slightly altered version of the C box) and an 11 nt rRNA recognition motif perfectly complementary to the rat 28S rRNA (positions 3466–3477). Thus, U87 RNA carries the elements essential for nucleolar accumulation and function of 2'-O-methylation guide snoRNAs. It was shown that a G residue at position 3468 of 28S rat rRNA is actually 2'-O-methylated (Maden, 1988). This G residue could be involved in a presumptive Watson-Crick helix formed by the antisense element of U87 RNA and the complementary rRNA sequence (positions 3466–3477). In this duplex, the G₃₄₆₈ is located exactly 5 bp from the D' box (Fig. 2). The same mutual arrangement is observed for other methylated residues of rRNA and D' box of the corresponding snoRNAs (Bachelierie and Cavaille, 1997).

Intracellular distribution of the U87 RNA was assayed by Northern analysis of the subcellular fractions isolated from rat liver. The U87 RNA was abundant in the nucleolar fraction but was hardly detectable in the nucleoplasmic fraction (data not shown). Hence, the U87 RNA is a novel C/D box snoRNA that seems to guide 2'-O-methylation of G₃₄₆₈ in rat 28S rRNA.

Northern analysis revealed this RNA in all species tested (mouse, mole rat, jerboa, cat, hedgehog, and human), although the signal intensity was different (Fig. 3). These

the terminal stem in mature state are probably transferred into nucleolus as a longer IR-containing precursors. A specific pathway may exist for snoRNAs of this type. Interestingly, we originally found a stemless U87 snoRNA in alpha-RNP particles rather than in nucleolus. Probably, the lack of the terminal stem allows a small fraction of U87 RNA to be within alpha-RNP (out of nucleolus). Alternatively, a fraction of nucleolar alpha-RNP may exist.

Sequence analysis of U87 RNA indicates that this gene is not an independent transcription unit; rather it is located within an intron of another gene like most genes coding for C/D box snoRNAs (Weinstein and Steitz, 1999). Search through the GenBank expressed sequence tag (EST) database revealed a number of relatively short cDNA fragments highly similar to the cloned sequences (e.g. nos. AI232081, W35693, and AA316590). Each of the rat, mouse, or human EST was similar to two to four regions of the U87 gene locus. The two EST-homologous regions nearest to the U87 gene are underlined in Fig. 2. These regions seem to be the exons of an unknown gene and the U87 RNA gene is located within its intron. This is also supported by the presence of canonical splice sites at the presumptive exons/intron junction sites. EST found in this locus do not contain long open reading frames. Thus these exons are the part of either an untranslated region of mRNA or correspond to a protein non-coding gene. Identification of this host gene is a goal of our current research.

4. Note added in proof

Recently Huttenhofer et al. (2001) cloned incomplete cDNA derived from the mouse homologue of U87 RNA named MBII-276.

Acknowledgements

We are grateful to Drs A. Bannikova (Moscow State University) and A. Puzachenko (Institute of Nature Conservation and Reserves, Moscow) for providing animal tissues and Dr N. Vassetzky for helpful discussion. This study was supported by State Research Program 'Frontiers in Genetics' and Russian Foundation for Basic Research (grants 99-04-49487 and 02-04-49772).

References

- Bachelierie, J.P., Cavaille, J., 1997. Guiding ribose methylation of rRNA. *Trends Biochem. Sci.* 22, 257–261.
- Balakin, A.G., Smith, L., Fournier, M.J., 1996. The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* 86, 823–834.
- Caffarelli, E., Losito, M., Giorgi, C., Fatica, A., Bozzoni, I., 1998. In vivo identification of nuclear factors interacting with the conserved elements of box C/D small nucleolar RNAs. *Mol. Cell. Biol.* 18, 1023–1028.
- Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156–159.
- Darzacq, X., Kiss, T., 2000. Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure. *Mol. Cell. Biol.* 20, 4522–4531.
- Ganot, P., Bortolin, M.L., Kiss, T., 1997. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89, 799–809.
- Huang, G.M., Jarmolowski, A., Struck, J.C., Fournier, M.J., 1992. Accumulation of U14 small nuclear RNA in *Saccharomyces cerevisiae* requires box C, box D, and a 5', 3' terminal stem. *Mol. Cell. Biol.* 12, 4456–4463.
- Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachelierie, J.P., Brosius, J., 2001. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* 20, 2943–2953.
- Kingston, R.E., 1992. Primer extension. In: Ausubel, F.M., et al. (Eds.). *Short Protocols in Molecular Biology*. Wiley, New York, pp. 21–22.
- Kiss-Laszlo, Z., Henry, Y., Bachelierie, J.P., Caizergues-Ferrer, M., Kiss, T., 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* 85, 1077–1088.
- Kiss-Laszlo, Z., Henry, Y., Kiss, T., 1998. Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* 17, 797–807.
- Konstantinova, I.M., Kulichkova, V.A., Vorob'ev, V.I., Ratovitski, E.A., Shaposhnikov, J.D., 1977. Studies of poly(A +)-RNA in mouse hepatoma and cortisone-stimulated rat liver. *Mol. Cell. Biochem.* 18, 29–32.
- Konstantinova, I.M., Turoverova, L.V., Petukhova, O.A., Vorob'ev, V.I., 1984. Cortisone-induced small RNP tightly bound to chromatin. *FEBS Lett.* 177, 241–245.
- Konstantinova, I.M., Volkova, I.V., Kulichkova, V.A., Turoverova, L.V., Gauze, L.N., 1987. The small RNP acceptor of glucocorticoids bound to chromatin in liver cell nuclei. *Mol. Biol. (Mosk)* 21, 1360–1366.
- Konstantinova, I.M., Kozlov Yu, Y., Kulitchkova, V.A., Petukhova, O.A., 1988. Small cytoplasmic RNA associated with polyadenylated RNA is involved in the hormonal regulation of gene expression. *FEBS Lett.* 238, 320–324.
- Lafontaine, D.L., Tollervey, D., 1999. Nop58p is a common component of the box C + D snoRNPs that is required for snoRNA stability. *RNA* 5, 455–467.
- Lange, T.S., Ezrokhi, M., Borovjagin, A.V., Rivera-Leon, R., North, M.T., Gerbi, S.A., 1998. Nucleolar localization elements of *Xenopus laevis* U3 small nucleolar RNA. *Mol. Biol. Cell* 9, 2973–2985.
- Liu, J., Maxwell, E.S., 1990. Mouse U14 snRNA is encoded in an intron of the mouse cognate hsc70 heat shock gene. *Nucleic Acids Res.* 18, 6565–6571.
- Lyman, S.K., Gerace, L., Baserga, S.J., 1999. Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *RNA* 5, 1597–1604.
- Maden, B.E., 1988. Locations of methyl groups in 28 S rRNA of *Xenopus laevis* and man. Clustering in the conserved core of molecule. *J. Mol. Biol.* 201, 289–314.
- Maxwell, E.S., Fournier, M.J., 1995. The small nucleolar RNAs. *Annu. Rev. Biochem.* 64, 897–934.
- Muramatsu, M., Hayashi, Y., Onishi, T., Sakai, M., Takai, K., 1974. Rapid isolation of nucleoli from detergent purified nuclei of various tumor and tissue culture cells. *Exp. Cell. Res.* 88, 245–251.
- Narayanan, A., Speckmann, W., Terns, R., Terns, M.P., 1999. Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli. *Mol. Biol. Cell.* 10, 2131–2147.
- Qu, L.H., Henry, Y., Nicoloso, M., Michot, B., Azum, M.C., Renalier, M.H., Caizergues-Ferrer, M., Bachelierie, J.P., 1995. U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA. *Nucleic Acids Res.* 23, 2669–2676.
- Samarsky, D.A., Fournier, M.J., Singer, R.H., Bertrand, E., 1998. The

- snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J.* 17, 3747–3757.
- Serdobova, I.M., Kramerov, D.A., 1998. Short retroposons of the B2 superfamily: evolution and application for the study of rodent phylogeny. *J. Mol. Evol.* 46, 202–214.
- Smith, C.M., Steitz, J.A., 1997. Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell* 89, 669–672.
- Tollervey, D., 1996. Small nucleolar RNAs guide ribosomal RNA methylation. *Science* 273, 1056–1057.
- Tycowski, K.T., Shu, M.D., Steitz, J.A., 1996. A mammalian gene with introns instead of exons generating stable RNA products. *Nature* 379, 464–466.
- Weinstein, L.B., Steitz, J.A., 1999. Guided tours: from precursor snoRNA to functional snoRNP. *Curr. Opin. Cell. Biol.* 11, 378–384.