Biochemical and Biophysical Research Communications **288,** 16–21 (2001)
doi:10.1006/bbrc.2001.5720, available online at http://www.idealibrary.com on IDEAL®

U86, a Novel snoRNA with an Unprecedented Gene Organization in Yeast

Daniela Filippini,* Fabiana Renzi,* Irene Bozzoni,* and Elisa Caffarelli†,[1]

**Istituto Pasteur Fondazione Cenci-Bolognetti, Dipartimento di Genetica e Biologia Molecolare, Università "La Sapienza," Rome, Italy; and †Centro di Studio per gli Acidi Nucleici of Consiglio Nazionale delle Ricerche, Rome, Italy*

Received September 3, 2001

The *Xenopus laevis* Nop56 gene (XNOP56), coding for a snoRNP-specific factor, belongs to the 5′-TOP gene family. XNOP56, as many 5′-TOP genes, contains an intron-encoded snoRNA. This previously unidentified RNA, named U86, was found as a highly conserved species in yeast and human. While in human it is also encoded in an intron of the hNop56 gene, in yeast it has an unprecedented gene organization: it is encoded inside an open-reading frame. Both in *X. laevis* and yeast, the synthesis of U86 snoRNA appears to be alternative to that of the cotranscribed mRNA. Despite the overall homology, the three U86 snoRNAs do not show strong conservation of the sequence upstream from the box D and none of them displays significant sequence complementarity to rRNA or snRNA sequences, suggesting a role different from that of methylation.

***Key Words: Xenopus laevis;* Nop56 gene; 5′-TOP genes; box C/D small nucleolar RNAs; intron; splicing; processing.**

The nucleolus of eukaryotic cells contains a multitude of small nucleolar RNAs (snoRNAs) which participate, at different levels, in ribosome biosynthesis. They interact with the nascent prerRNA and allow the correct folding, cleavage, and nucleotide modifications of the RNA precursor. One of the two major families, the box C/D snoRNAs, are mainly involved in site-specific 2′-O-ribose methylation, not only of rRNA (1), but also of snRNA substrates (2–4). In vertebrates, snoRNAs display a peculiar gene organization. Most of them are localized inside introns of protein coding genes. The biosynthesis of such RNAs can be splicing-dependent or splicing-independent. In the latter case the snoRNA is released by endonucleolytic cleavages of the intron. In yeast, which also contains a few intron-encoded specimens, most of the snoRNAs are encoded by independent genes (5). We have identified a novel snoRNA, U86, which is intron-encoded in vertebrates, while in yeast it displays an unprecedented gene organization, since it is encoded inside an open-reading frame.

Abbreviations used: *X. laevis, Xenopus laevis;* snoRNAs, small nucleolar RNAs; snoRNPs, small nucleolar ribonucleoprotein particles; ORF, open-reading frame; 5′-TOP, 5′-terminal oligopyrimidine tract; PCR, polymerase chain reaction.

[1] To whom correspondence should be addressed. Fax: 39-6-49912500. E-mail: elisa.caffarelli@uniroma1.it.

MATERIALS AND METHODS

Analysis of U86 snoRNA. For the detection of U86 in *X. laevis* oocytes RNase protection assay was performed. This was carried out using the RPA II kit (Ambion) and a 213-antisense RNA probe, targeting the Nop56 intron sequence including U86, produced by *in vitro* transcription in the presence of [α-^{32}P]UTP. The template for RNA probe was obtained by PCR amplification of the XNop56 genomic clone using oligonucleotides U1 (GATCCTGGGAGCGGAGAAGGCTCTGTTC) and U2 (TAATACGACTCACTATAGGGCCAATGGCAGCGATACAAC). For the detection of U86 in *X. laevis* embryos total RNA, corresponding to three embryos, was used. RNA was run on a 6% polyacrylamide-7M urea gel and blotted to Hybond membranes (Amersham). Hybridizations were carried out with the U86 oligonucleotide (GGACTATCACTCACAGAGCT). To detect hU86 snoRNA and yU86 snoRNA, 10 μg of total human fibroblast RNA and 5 μg of total *Saccharomyces cerevisiae* RNA were used; hybridizations were carried out with the following probes: the oligonucleotide bodyhU86 (CATAGCAGAGGCCTGTATAGGTCAG) and the oligonucleotide yU86 (GGTTGGATACGAGCTCTTGCGACACATTG) for human and yeast RNA, respectively. To detect yU86 snoRNA coding region on RIB1 mRNA, RNase protection assay was performed on total yeast RNA using a 120 nt antisense RNA probe targeting 100 nt, including 50 nt of the 3′ end of yU86 snoRNA and 50 nt of the downstream RIB1 sequence. The template for the RNA probe was obtained by PCR amplification on total yeast DNA using the oligonucleotides Rib1-U86A (ACCTCGATCAAGTCTATCGGTACAATGTGTCGCAA) and Rib1-U86B (TAATACGACTCACTATAGGGATGGCTAGATGTTCCTTG).

In vitro transcription, oocyte microinjection, and analysis of RNA processing products. The preU86 template was obtained by PCR amplification using oligonucleotides UT7 (TAATACGACTCACTATAGGGGTGCCCGTCTCATCTCCCAC) and U*Hin*dIII (AAGCTTTCATGGCGGCTCGGCCAAT) on the XNop56 genomic clone and then cloned into PvuII digested pBS vector. The preU16 template was constructed as already described (9). RNA substrates were synthesized *in vitro* as previously described (6). For oocyte microinjection [^{32}P]-labeled transcripts were dissolved in double-distilled water

0006-291X/01 $35.00

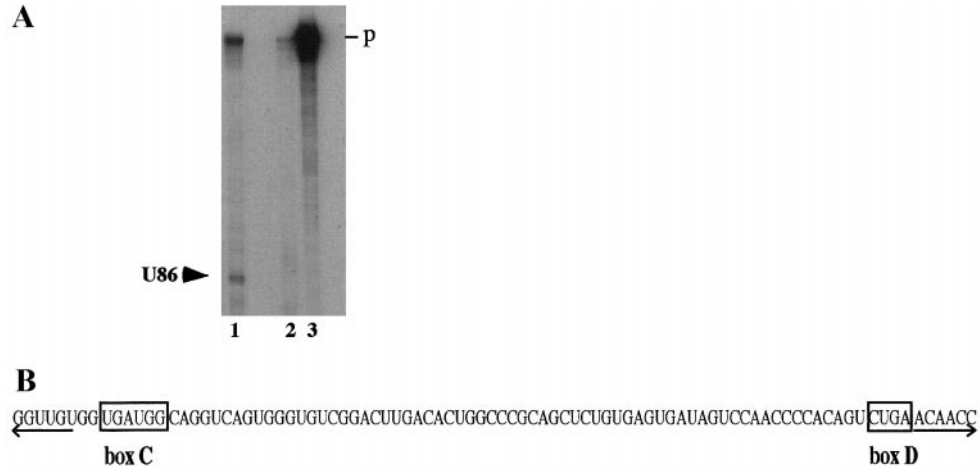

FIG. 1. Identification of U86 snoRNA in *X. laevis* oocytes. (A) RNase protection analysis was performed on 10 μg of total RNA extracted from *X. laevis* oocytes (lane 1) and, as a control, on 10 μg of total yeast RNA (lane 2). (Lane 3) Nonreacted probe (p). The probe used, 213 nucleotides long, was obtained by PCR amplification of the intron portion including U86 snoRNA (see Materials and Methods). (B) Nucleotide sequence of *X. laevis* U86 snoRNA. The conserved C and D elements are boxed, while the sequences forming the 5′–3′ terminal stem are underlined.

at a concentration of 0.2–0.4 pmol/μl and 9.2 nl were injected into germinal vesicles of stage VI oocytes. After injection of [^{32}P]-labeled U16- and U86-containing precursors (preU16 and preU86), the oocytes were incubated at 19°C for 10 min, 2 h, and 16 h. After each incubation time nuclei were manually isolated, total RNA was purified and analyzed on a 10% polyacrilamide-7M urea gel.

Rnase H assay. RNA molecules, eluted from gel slices, were purified and incubated in 25 μl of 500 μM ATP, 20 mM creatine phosphate (CP), 2.4 mM $MgCl_2$ together with 8 μM oligonucleotides and 60% HeLa nuclear extract. After 30 min of incubation at 30°C, the samples were purified and analyzed on 6% acrylamide-urea gel. The oligonucleotides used were: α (CAGCTACAGCCAATGGCAGC) complementary to an intronic region downstream from U86 snoRNA; β (CTGTAATCCAATTGGTCATG) complementary to an intronic region upstream from exon 2; γ (CGCTCCCAGGATCTGCACCG) complementary to a region of exon 1.

Accession numbers. The DDBJ/EMBL/GenBank Accession Nos. for the *X. laevis*, human, and *S. cerevisiae* U86 snoRNA are AJ311852, AJ311853, and AJ311854, respectively.

RESULTS

An Intron of XNop56 Gene Encodes for a Novel Box C/D snoRNA

Inspection of the 5′ end of XNop56 mRNA showed the typical features of the 5′-TOP genes (manuscript in preparation). Partial sequencing of Nop56 gene in *X. laevis* revealed the presence of an intronic sequence, 294 nucleotides long, containing some elements diagnostic for the occurrence of a small nucleolar RNA: (i) the conserved boxes C and D, shared by all the members of box C/D family, and (ii) six nucleotides, upstream and downstream from the boxes, which can pair in a terminal stem (7). In order to test whether this sequence is indeed expressed as an independent and stable RNA in *X. laevis* cells, we performed RNase protection assay and Northern blot analysis. The results of these experiments revealed the presence of a stable 90-nucleotide long snoRNA species both in *X. laevis* oocytes and embryos (Figs. 1A and 3B). This previously unidentified RNA was named U86 (DDBJ/EMBL/GenBank Accession No. AJ311852). Its size corresponds to that expected if the RNA termini map at the ends of the terminal stem, as it generally occurs in the snoRNAs identified so far (1). In order to show that U86 was produced by processing of the XNOP56 host-intron, we microinjected a [^{32}P]-labeled precursor RNA, containing the intron and portion of the flanking exons, into the nucleus of *X. laevis* oocytes as previously done for the study of U16 and U18 snoRNAs biosynthesis (8). Processing of U16 snoRNA was studied in great detail both *in vivo* and *in vitro* (8, 9). We were able to demonstrate that U16-containing premRNA can undergo two alternative pathways: splicing, with the production of mRNA, or processing with the release of the snoRNA. We also showed that the commitment towards the processing pathway rely on the assembly of specific snoRNP complexes, containing fibrillarin, on the premRNA. Only molecules lacking in such complexes would be able to assemble into functional spliceosomes (8). The choice between the two pathways is allowed by the presence of sub-optimal splicing consensus sequences (the 5′ splice site in U16-containing intron). In addition, we have characterized the endonucleolytic activity (XendoU) which, by cleavages upstream and downstream from U16 coding region, is involved in the first step of U16 snoRNA biosynthesis (10). U86 snoRNA, as U16, is contained in a very poorly spliceable intron, in that it contains a noncanonical pyrimidine tract at its 3′ end. Figure 2A (panel preU86) shows the pattern of U86 snoRNA excision after injection of U86-containing precursor into *X. laevis* oocytes. It is evident that while the splicing of

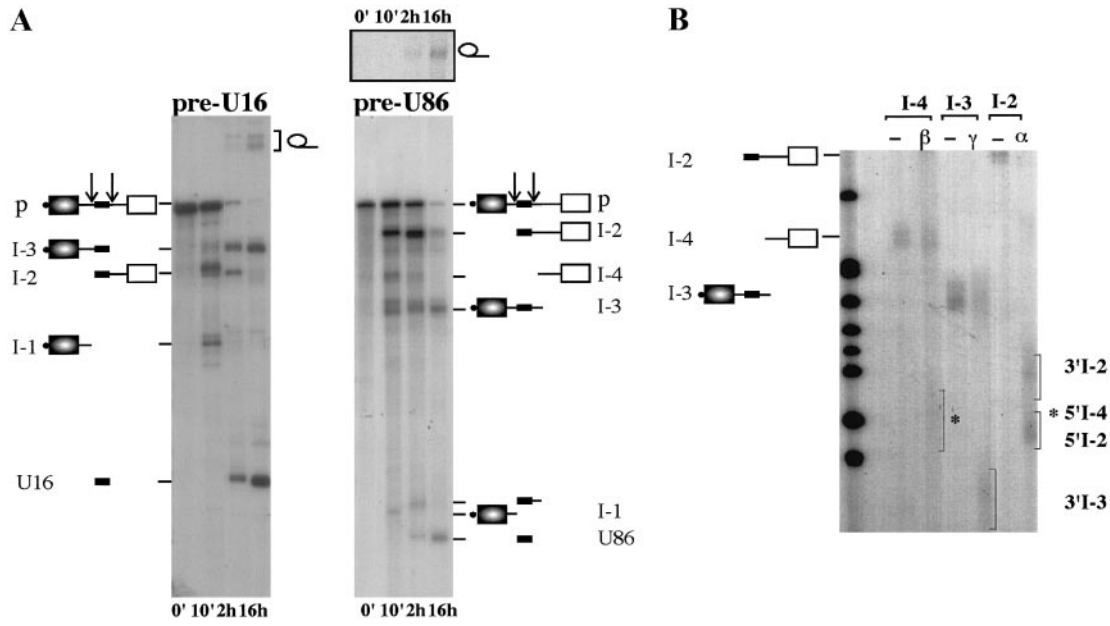

FIG. 2. *In vivo* processing analysis of U86-containing precursor. (A) *In vivo* processing of U16 (preU16) and U86 (preU86) precursors. [^{32}P]-labeled RNAs were injected into the nuclei of *X. laevis* oocytes and incubation was allowed to proceed for the times indicated above the lanes. The RNA precursors (p) and intermediates (I-1, I-2, I-3, and I-4) are diagrammed on the sides (black dots, CAP structure; boxes, exons; continuous line, intron; thicker line, snoRNA coding region; arrows indicate the sites of endocleavage). Mature snoRNA molecules are also indicated. An overexposure of the upper part of preU86 gel, containing the intron-lariat, is shown in the upper panel. (B) RNase H analysis. The different cleavage RNA molecules I-4, I-3, and I-2, obtained after injection in the experiment of (A), were loaded on a 6% acrylamide-urea gel before (lanes –) or after RNase H treatment with the oligonucleotides α, β, and γ (see Materials and Methods). On the left side the RNA molecules treated with RNase H are schematically represented (black dot, CAP structure; boxes, exons; continuous line, intron; thicker line, snoRNA coding region). 5′I-2 and 3′I-2 indicate the 5′ and 3′ cut off products of the I-2 molecule; 5′ I-4 and 3′I-3 indicate the 5′ half and the 3′ half of I-4 and I-3 molecules, respectively. The complementary cut off products of these fragments are short and run off the gel.

preU86 is quite inefficient, as demonstrated by the almost undetectable accumulation of splicing products (the intron released in the lariat form is visible only after an overexposure, see the upper panel of Fig. 2A), the processing reaction generates, besides U86 snoRNA, the intermediates products I-1, I-2, I-3, and I-4 which are visible at short incubation times (10 min). The half-life of these products is quite short *in vivo* because the unprotected ends are rapidly attacked by exonucleases. In particular I-1 and I-4 are rapidly degraded, whereas I-3 and I-2 are more stable. On the basis of their size, these molecules can be predicted to be truncated products generated by cleavages occurring upstream and downstream from the snoRNA region (8). The identity of the processing products was determined by RNase H assay (Fig. 2B). The results of the analysis are schematically represented on the side of the gel (Fig. 2A). The I-3 and I-4 molecules represent the 5′ and 3′ cut-off molecules originating from cleavage downstream from U86 sequence, while I-1 and I-2 molecules are the cut off products of cleavage upstream from U86. The higher stability of I-3 and I-2 molecules can then be explained by the fact that they contain the snoRNA sequence which, assembling the specific box C/D protein factors, protect RNA from degradation.

To identify potential RNA substrates for U86 snoRNA, search for complementarity of U86 to known methylated targets (1, 13, 14) was performed following the rules described for methylating snoRNAs (11); no significant match was found.

U86 Has Homologs in Human and Yeast

Human genome database search revealed the existence of a sequence 62% homologous to the *X. laevis* U86 snoRNA (Fig. 3A). This sequence is nested in one intron of the human Nop56 gene. Northern blot hybridization on total RNA from human fibroblasts (Fig. 3C) demonstrated the existence of U86 (hU86) (DDBJ/EMBL/GenBank Accession No. AJ311853) as an RNA species shorter than 90 nucleotides.

A parallel analysis in the yeast genomic database was carried out. Blastn research was run introducing XU86 nucleotide sequence to *S. cerevisiae* complete genomic sequence. Only one match (score, 117; P(N), 0.89) was found on the Crick strand of chromosome II (coordinates 159601–159519). This sequence is localized inside the RIB1 ORF (coordinates 159656–158619) and is 59% homologous to the *X. laevis* U86 snoRNA (Fig. 3A). RIB1 gene encodes for a GTP-cyclohydrolase III, a 38-kDa protein involved in the riboflavin biosynthetic pathway (15). Northern analysis of Fig. 3D on *S. cerevisiae* total RNA demonstrated the existence of such an RNA species with an apparent

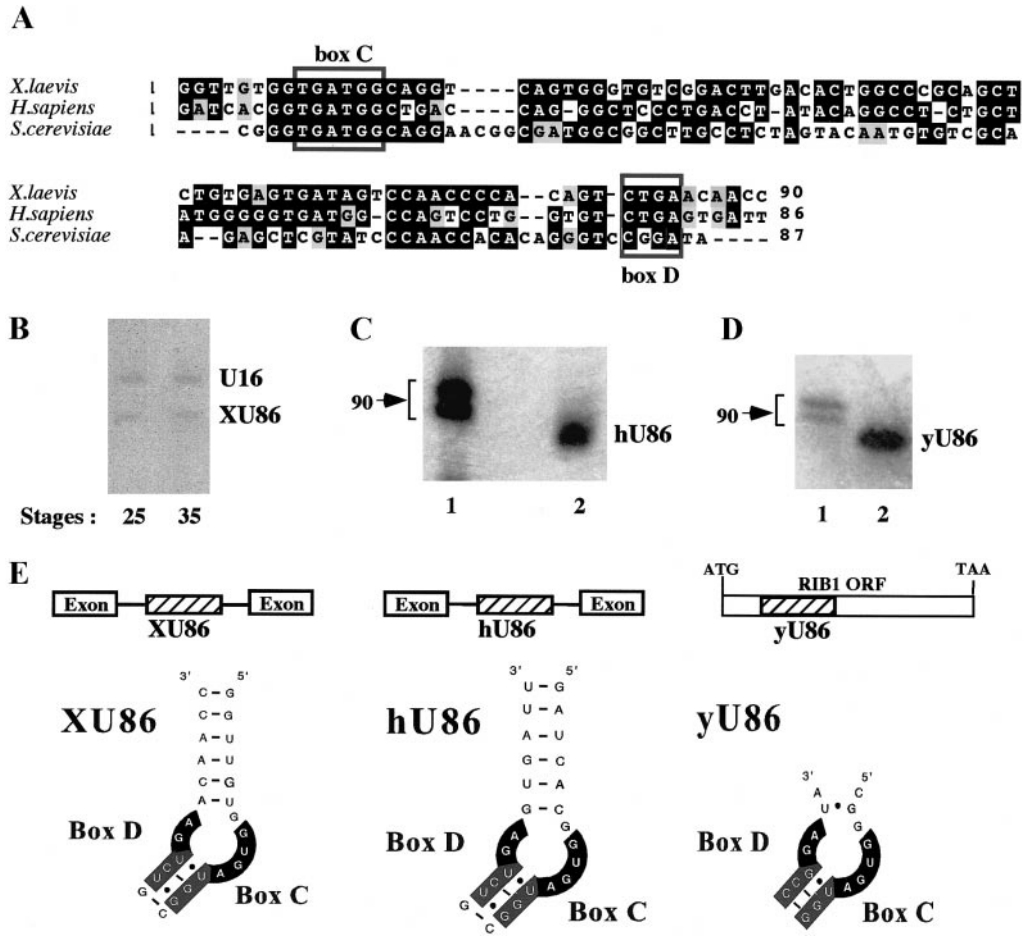

FIG. 3. Identification of human and yeast U86 snoRNAs. (A) Alignment and sequence comparison of *X. laevis*, human and yeast U86 snoRNAs. Identical nucleotides are indicated by white letters on black blocks and purine changes are indicated by grey shading. Dashes indicate spaces inserted to aid with the nucleotide sequence alignment. The boxes C and D are boxed. (B) Northern blot of total RNA from three stage 25 and 35 *X. laevis* embryos run on a 6% polyacrylamide-urea gel. A U86 and U16 mixed probe was used; the identified RNAs are indicated on the right. (C) Northern analysis of 10 μg of total human fibroblast RNA (lane 2) run in parallel with a [^{32}P]-labeled *Msp*I-digested pBR322 DNA (lane 1). (D) Northern analysis of 5 μg of total *S. cerevisiae* RNA (lane 2) run in parallel with a [^{32}P]-labeled *Msp*I-digested pBR322 DNA (lane 1). The probes utilized in (B), (C), and (D) are described under Materials and Methods. (E) Genomic organization of *X. laevis*, human and yeast U86 snoRNAs is schematically represented in the upper part; in the lower part of the box C/D motifs of *X. laevis*, human and yeast U86 snoRNAs are schematically drawn, according to the recently redesigned secondary structure (16).

size shorter than 90 nucleotides (yU86) (DDBJ/EMBL/GenBank Accession No. AJ311854). The demonstration that U86 snoRNA sequence is indeed embedded within the RIB1 mRNA derived from an RNase protection experiment performed on total *S. cerevisiae* RNA. The probe used is 120 nucleotides long and includes 50 nucleotides of the 3′ end of U86 snoRNA sequence and 50 nucleotides of the downstream RIB1 sequence (Fig. 4). A protected fragment of 100 nucleotides is expected when the probe recognizes the RIB1 mRNA, while a band of 50 nucleotides is obtained if the U86 coding region is expressed as a stable RNA. Figure 4 shows that both bands are obtained when total yeast RNA is used, confirming that two distinct RNA species originate from the same genomic sequence.

It is interesting to observe that, even if YU86 snoRNA displays a sub-optimal D box (CGGA instead than CUGA) it does not affect the conserved secondary structure of the box C/D motif (Fig. 2E) as recently redesigned (16, 17).

Similarly to *X. laevis* U86, also the human and yeast counterparts do not show any significant complementarity to rRNAs or snRNAs; in addition, despite their overall homology, the three snoRNAs do not display strong conservation of the sequence upstream from the D box, which is involved in RNA target recognition (Fig. 3A).

DISCUSSION

Many 5′-TOP genes share the property of harbouring intron-encoded snoRNAs. The analysis of a XNOP56 intron revealed the presence of sequences diagnostic for a box C/D small nucleolar RNA. Northern analysis

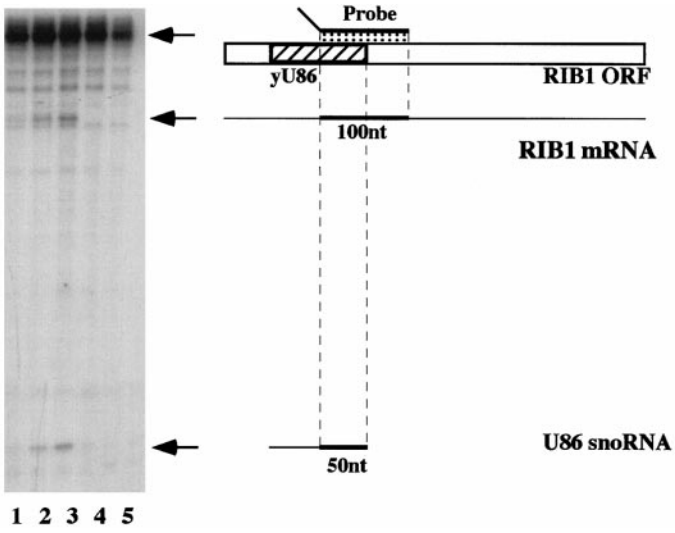

FIG. 4. U86 snoRNA and RIB1 mRNA originate from the same transcriptional unit. RNase protection analysis was performed on 1 μg (lane 1), 5 μg (lane 2), and 10 μg (lane 3) of total yeast RNA and, as a control, on 10 μg (lane 5) of the total RNA extracted from *X. laevis* oocytes. As further control the reaction was carried out in the same conditions but in the absence of yeast RNA (lane 4). The 120-nucleotide long RNA probe used is schematically represented on the right side. It was produced by *in vitro* transcription of a PCR product, obtained as described under Materials and Methods, and includes 50 nt of the 3′ end of yU86 snoRNA and 50 nt of the downstream RIB1 sequence. The specific products of the reaction are also represented on the side and indicated by arrows.

and RNase mapping proved the existence of such an RNA species in *X. laevis* total RNA; this RNA was named U86. As in the case of U16 snoRNA (8), also U86 is encoded in a poorly spliceable intron. Microinjection experiments in *X. laevis* oocytes showed that U86 is released from its host intron through a splicing-independent pathway consisting of endonucleolytic cleavages of the premRNA. In this case, as for U16, splicing and processing appear to be mutually exclusive, indicating that the production of U86 is alternative to that of XNop56 mRNA. This conclusion is further supported by experimental data showing that in different batches of *X. laevis* oocytes an inverse correlation between efficiency of splicing and snoRNA processing is observed (data not shown).

Database sequence analysis revealed the occurrence of sequences highly homologous to U86 in human and yeast genomes. In both cases the presence of stable RNA species of corresponding size was confirmed. The genomic organization of U86 snoRNA appears to be conserved in vertebrates. In fact, also in human it is encoded in one of the introns of the Nop56 gene. On the contrary, a very different organization was found in yeast. U86 snoRNA is embedded in the open-reading frame of the RIB1 gene (15). Thus, as in the case of *X. laevis*, the production of U86 would be alternative to that of the cotranscribed mRNA since the two RNA species (U86 snoRNA and RIB1 mRNA) should originate from the same transcriptional unit. This arrangement, which has no precedents, opens the interesting question of how yU86 snoRNA synthesis is coordinated with that of Rib1 mRNA. Interestingly, a 17-nucleotide long stem, previously shown to be indispensable for processing of the intron-encoded snoRNAs (18), is present in the ORF regions flanking U86.

Sequence comparison indicated that the three identified U86 snoRNAs do not share a strong conservation in the region upstream from the box D; in addition, these regions do not show significant complementary to either rRNAs or snRNAs (1–4). Recently a large number of new mouse snoRNAs, including box C/D species, with unidentified targets was reported (19). Altogether these data suggest that these new RNA species could have noncanonical targets or play a different role than methylation. Accumulation of U86 snoRNA and of the cotranscribed mRNA are mutually exclusive, therefore a regulatory role for the snoRNA can be suggested.

ACKNOWLEDGMENTS

We are grateful to Massimo Arceci for skillful technical help. This work was partially supported by grants from PRIN 40% of MURST, from CNR "Target Project on Biotechnology," MURST-CNR Progetto Strategico legge 449/97, and CNR progetto strategico: Tecnologie di base della post-genomica.

REFERENCES

1. Bachellerie, J. P., Michot, B., Nicoloso, M., Balakin, A., Ni, J., and Fournier, M. J. (1995) Antisense snoRNAs: A family of nucleolar RNAs with long complementarities to rRNA. *Trends. Biochem. Sci.* **20,** 261–264.
2. Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998) Modification of U6 spliceosomal RNA is guided by another small RNA. *Mol. Cell* **2,** 629–638.
3. Ganot, P., Jady, E. B., Bortolin, M. L., Darzacq, X., and Kiss, T. (1999) Nucleolar factors direct the 2′-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.* **19,** 6906–6917.
4. Jady, B. E., and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.* **20,** 541–551.
5. Bachellerie, J. P., Cavaillé, J., and Liang-Hu, Q. (2000) The ribosome: Structure, function, antibiotics, and cellular interactions, ASM Press, Washington, D.C.
6. Filippini, D., Bozzoni, I., and Caffarelli, E. (2000) p62, a novel *Xenopus laevis* component of box C/D snoRNPs. *RNA* **6,** 391–401.
7. Ling, X., Watkins, N., and Maxwell, S. (1997) Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. *RNA* **3,** 17–26.
8. Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P., and Bozzoni, I. (1996) Processing of the intron-encoded U16 and U18 snoRNAs: The conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J.* **15,** 1121–1131.
9. Caffarelli, E., Arese, M., Santoro, B., Fragapane, P., and Bozzoni, I. (1994) *In vitro* study of processing of the intron-encoded U16

small nucleolar RNA in *Xenopus laevis. Mol. Cell. Biol.* **14,** 2966–2974.

10. Caffarelli, E., Maggi, L., Fatica, A., Jiricny, J., and Bozzoni, I. (1997) A novel Mn^{++}-dependent ribonuclease that functions in U16 snoRNA processing in *X. laevis. Biochem. Biophys. Res. Commun.* **233,** 514–517.
11. Kiss-Làszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* **85,** 1077–1088.
12. Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1996) A mammalian gene whose introns instead of exons generate stable RNA products. *Nature* **3,** 464–466.
13. Ganot, P., Jady, E. B., Bortolin, M. L., Darzacq, X., and Kiss, T. (1999) Nucleolar factors direct the 2′-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.* **19,** 6906–6917.
14. Jady, B. E., and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.* **20,** 541–551.
15. Oltmanns, O., and Bacher, A. (1972) Biosynthesis of riboflavine in *Saccharomyces cerevisiae:* The role of genes rib1 and rib7. *J. Bacteriol.* **110,** 818–822.
16. Watkins, N. J., Segault, W., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Luhrmann, R. (2000) A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* **103,** 457–466.
17. Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5 kD protein bound to U4 snRNA fragment. *Mol. Cell* **6,** 1331–1342.
18. Villa, T., Ceradini, F., and Bozzoni, I. (2000) Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *S. cerevisiae. Mol. Cell. Biol.* **20,** 1311–1320.
19. Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J. P., and Brosius, J. (2001) RNomics: An experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* **20,** 2943–2953.