Volume 307, number 3, 337–342 FEBS 11376 August 1992

A novel class of nucleolar RNAs from *Tetrahymena*

Henrik Nielsen[a], Henrik Ørum[b] and Jan Engberg[b]

[a]*Department of Biochemistry B, Panum Institute, University of Copenhagen, Blegdamsvej 3, DK 2200 Copenhagen N, Denmark* and [b]*Department of Biological Sciences, The Royal Danish School of Pharmacy, Universitetsparken 2, DK-2100 Copenhagen, Denmark*

Received 12 June 1992

We describe a family of at least four nucleolar RNAs (snoRNAs) from the ciliate, *Tetrahymena*. The snoRNAs are 120–140 nucleotides long, moderately AU-rich and contain no modified nucleotides. Their 5′ ends are blocked by a cap of unknown nature. The snoRNAs can be folded into similar secondary structures consisting of two hairpins separated by a single-stranded AU-rich spacer. The sequences and secondary structures show no extensive sequence or secondary structure resemblance to any other small RNAs in the public databases.

Nucleolar RNA; RNA sequencing; *Tetrahymena thermophila*

1. INTRODUCTION

Nucleoli are the sites of ribosome biogenesis in the eukaryotic cell. Although this was established more than 30 years ago, some fundamental problems remain unsolved, e.g. the site of transcription within the nucleolus and the mechanism of pre-rRNA processing [1,2]. One way to approach these important questions is to study nucleolar structure and function in a variety of organisms. We have chosen the ciliate, *Tetrahymena*, because this organism offers the advantage of having extrachromosomal, amplified rDNA (reviewed in [3]).

Tetrahymena is a dinucleate ciliate. The germinal micronucleus is essentially transcriptionally inactive whereas the polyploid macronucleus is transcriptionally active (for an introduction to the molecular biology of *Tetrahymena*, see [4]). In exponentially growing ciliate cultures, 500–1,000 nucleoli are found in the macronucleus, equally distributed over the inner surface of the nuclear double membrane. The nucleoli show no morphological change during the amitotic divisions of the macronucleus [5]. The nucleoli from *Tetrahymena* are easy to purify [6,7] and the molecular components have been described in considerable detail.

Highly purified nucleoli contain only one kind of DNA, the rDNA [7]. The macronuclear rDNA is derived from a single micronuclear gene copy during sexual reorganization, and is therefore homogenous in sequence [3]. About 10,000 extrachromosomal, palindromic rDNA molecules are found in the macronucleus, each containing two transcriptional units. We have reported the complete sequence of these 21-kb long rDNA molecules from *T. thermophila* [8]. More than 50 protein bands are seen in SDS-PAGE of proteins from highly purified nucleoli [7]. Very few nucleolus-specific proteins have been studied in detail, but nuclear proteins such as histones, topoisomerases and telomere binding proteins have been studied through their association with rDNA (references in [4]). Several nucleolar RNAs were identified in a paper by Pedersen et al. [9] on snRNAs from *Tetrahymena*. Apart from RNAs transcribed from rDNA, three types of RNA were found to be highly enriched in crude nucleoli. These RNAs were termed T3, T6 and T7, according to their order of mobility in denaturing polyacrylamide gels. We later identified T7 as U3 and characterized T6 (H. Nielsen et al., manuscripts in preparation). T3 was originally described as a cluster of bands of roughly similar electrophoretic mobilities corresponding to several different RNAs. In this paper, we describe the sequence and a possible secondary structure model of four of the T3 RNAs. We find by several criteria that these molecular species are structurally related and thus constitute a new family of nucleolar RNAs. As a consequence, we have re-named these RNAs snoRNA01–04 (*s*mall *no*ucleolar RNA).

Abbreviations: snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; rRNA, ribosomal RNA; tRNA, transfer RNA; rDNA, ribosomal RNA genes; m_3G, trimethyl guanosine.

Correspondence address: H. Nielsen, Department of Biochemistry B, Panum Institute, University of Copenhagen, Blegdamsvej 3, DK 2200 Copenhagen N, Denmark. Fax: (45) 31 35 60 42.

2. MATERIALS AND METHODS

2.1. *Cell material*

The strain used in the present study was *T. thermophila* B1868VII. The ciliates were grown in complex medium and harvested in the exponential growth phase at approximately 200,000 cells/ml. RNA from whole cells and nuclei was isolated according to Pedersen et al. [9] and individual RNAs purified from denaturing-PAGE as previously [10]. For preparation of crude nucleoli, cells from a 1.5 l culture

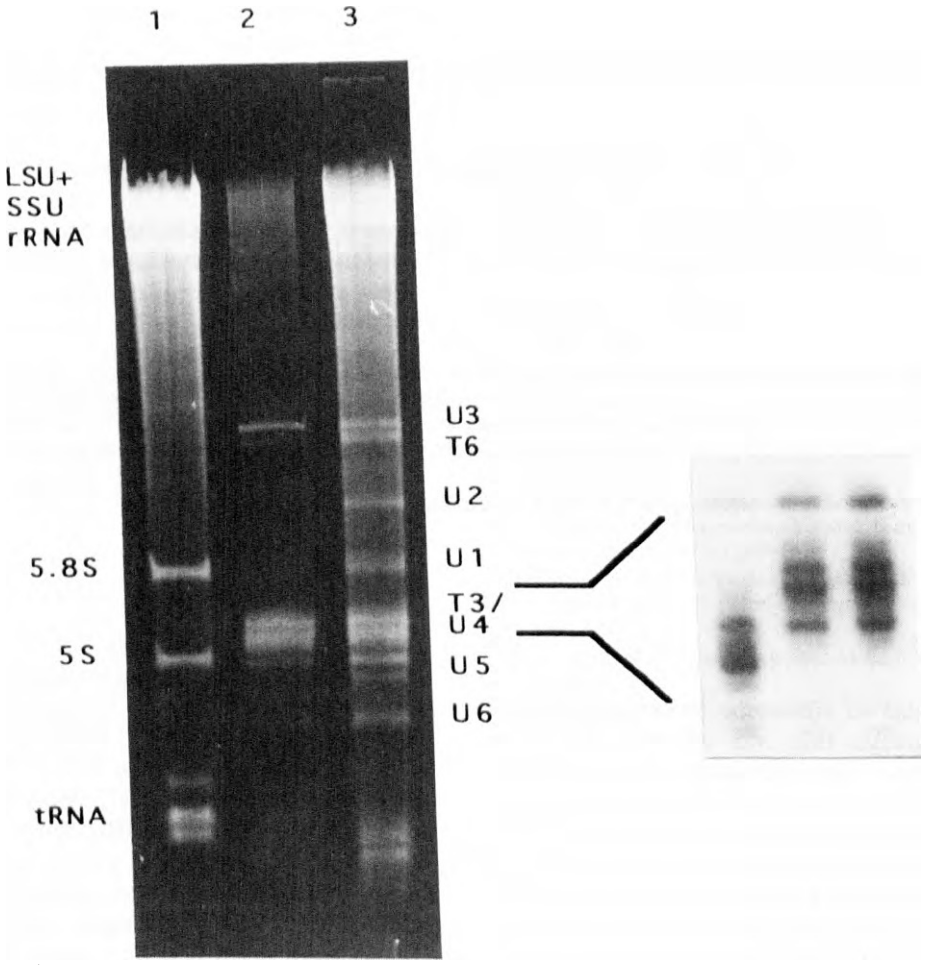

Fig. 1. Gel-electrophoretic analysis of nuclear RNA. (A) Samples of whole cell RNA (100 μg, lane 1), nucleolar RNA (10 μg, lane 2) and nuclear RNA (100 μg, lane 3) were run on a 5% denaturing (8 M urea) polyacrylamide gel and ethidium bromide stained. The positions of some previously identified cytoplasmic and nuclear RNAs are indicated. (B) The T3/U4 cluster of bands was cut out of the gel in three size fractions, end labelled, and run on an 8% denaturing (8 M urea) sequencing gel. Individual bands were cut out of the gel, the RNA eluted and subjected to sequencing analysis.

were washed in Mita-buffer (0.1 M sucrose, 10 mM NaCl, 3 mM $CaCl_2$, 10 mM Tris-HCl, pH 7.5) and lysed in the same buffer containing 0.3% NP40. The nuclei were harvested from the lysate by centrifugation at 1,000 × *g* in a Sorvall SS34 rotor and washed twice with 20 ml of Mita-buffer containing 0.3% NP40. The resulting nuclear pellet was resuspended in 50 ml of a buffer consisting of 140 mM NaCl, 1 mM $MgCl_2$, 1 mM β-mercaptoethanol, 10 mM Tris-HCl, pH 7.4, and 10% glycerol. The suspension was homogenized in a Dounce homogenizer (10 strokes, loose pistel) and the nuclei spun down at 1,000 × *g*. Finally, the nucleoli were harvested from the supernatant by centrifugation at 12,000 × *g* for 10 min. The nucleoli were resuspended in 1 ml of Mita-buffer. The yield and purity of the nucleoli were assayed by a gel-electrophoretic analysis of the rDNA content in the preparation. We estimate a yield of about 20%, and that more than 90% of the DNA was rDNA, corresponding to a ca. 50-fold purification at the DNA level. Nucleolar RNA was isolated after treatment of the crude nucleoli with proteinase K and RNAse-free DNAse I, as described in [9].

2.2. *Sequencing methods*

Sequencing of individual RNA species was as previously described [10]. Briefly, the purified RNA species was labelled at the 3′ end using T4 RNA ligase and [^{32}P]pCp. The labelled RNA was further gel-purified and sequenced by the chemical sequencing method. From the obtained sequence, an oligonucleotide with a sequence complementary to the 3′ end of the RNA was synthesized. This oligonucleotide was used in dideoxy sequencing using AMV reverse transcriptase and total nuclear RNA in order to obtain the remaining sequence. Then, an oligonucleotide with a sequence corresponding to the 5′ end of the RNA was synthesized and used in dideoxy sequencing of the first strand cDNA made from in vitro polyadenylated nuclear RNA. In this way, we were able to verify the result from the chemical sequencing and obtain a reliable sequence from the very 3′ end of the RNA. The sequences have been deposited in the EMBL database under accession nos. X66429–66432.

2.3. *Nucleolar localization by reverse transcription*

The nucleolar localization of the individual RNA species was demonstrated by a reverse transcription experiment. A primer cocktail consisting of 2 pmol of each of the relevant oligonucleotides labelled with ^{32}P at their 5′ ends [12] was mixed with nuclear RNA (50 μg) or nucleolar RNA (5 μg) in AMV reverse transcriptase buffer (50 mM Tris-HCl, pH 8.2, 6 mM $MgCl_2$, 10 mM DTT). The mixture was heated to 90°C for 3 min and transferred to 55°C for 45 min for primer annealing (this temperature is 5–15°C below the calculated T_m of the oligonucleotides involved). RNAse inhibitor (10 U), dNTPs (final concentration 0.2 mM of each) and AMV reverse transcriptase (10 U, Boehringer, Mannheim) was added and incubation continued at 50°C

Fig. 2. Sequences and potential secondary structures of four *T. thermophila* snoRNAs. The indicated sizes of the RNAs include an unidentified nucleotide at the 5' end (N, except in snoRNA04 for which the gene sequence is known), but not the cap designated 'X' in the figure. The foldings were assisted by the use of the GCG version of Zukers RNA folding programme [11].

for an additional 45 min. The nucleic acids were precipitated by the addition of 200 μl 2 M ammonium acetate, pH 4.5, 1 μl of *E. coli* tRNA (20 mg/ml) and 2.5 vols. of 96% ethanol. The precipitate was resuspended in loading buffer (8 M urea, 0.25% Bromophenol blue, 0.25% Xylene cyanol FF), heat denatured and analysed on 5% denaturing polyacrylamide gels.

3. RESULTS

A gel-electrophoretic analysis of RNA extracted from crude nucleoli is displayed in Fig. 1A, lane 2. By comparison to the nuclear RNA in lane 3 we find, as in our previous work [9], that U3 and a cluster of bands, termed T3, are enriched in nucleoli to the same extent, whereas T6 is enriched to a lesser extent. Compared to our earlier studies, we now find additional bands in the T3 cluster of a size comparable to 5 S rRNA, as well as several smaller RNAs of a size similar to the tRNAs. This is probably not due to cytoplasmic contamination, since no 5.8 S rRNA is observed. The differences compared to our earlier studies could be due to the fact that different protocols for nucleolar isolation were applied or that our previous observations were based on in vivo-labelled RNA in contrast to the EtBr staining pattern in Fig. 1. The T3 cluster was cut out of the gel in three slices representing different size fractions, and the RNA was eluted and 3′ end-labelled with [^{32}P]pCp using T4 RNA ligase. The result of this labelling is shown in Fig. 1B. At least eight different bands can be discerned in the T3 cluster. These bands all correspond to individual RNAs as evidenced by partial RNA sequence analysis (not shown). Four of the RNAs were selected for detailed sequencing studies and named snoRNA01–04.

The sequencing strategy is outlined in section 2, and the sequences are displayed in Fig. 2. The sequence at the very 5′ end is uncertain because of the presence of a cap of unknown nature (see later). The snoRNAs are 134, 130, 135 and 127 nucleotides in length, respectively. They contain no modified nucleotides [9] and are moderately AU-rich (60–65%). In this respect, it should be kept in mind that the overall genomic AT-content of *Tetrahymena* is 75% and that the AU-richness of the snoRNAs therefore could be a reflection of this general tendency rather than a specific feature. All four sequences can be folded into similar secondary structures as shown in Fig. 2. The structure consist of two hairpins separated by a long single-stranded spacer. The loops of all hairpins are rich in AUs compared to the whole molecule. The centrally located 35–40 nucleotides are

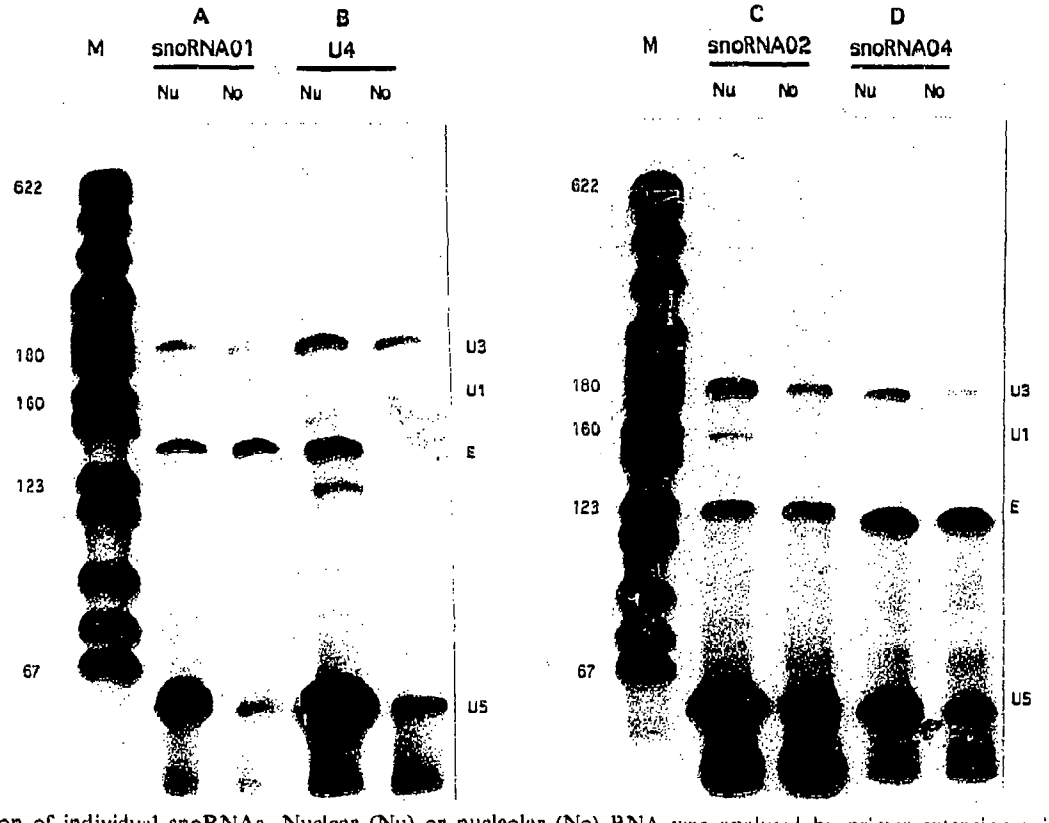

Fig. 3. Localization of individual snoRNAs. Nuclear (Nu) or nucleolar (No) RNA was analysed by primer extension using a mixture of oligonucleotides, and the extension products analysed on a 5% denaturing (8 M urea) polyacrylamide gel. The oligonucleotides complementary to the internal control RNAs were complementary to U3 pos. 170–189 (unpublished), U1 pos. 139–161 and U5 pos. 35–59 [10], respectively (the last number is also the length of the extension product in the present experiment). The experimental oligonucleotides (E) were complementary to snoRNA01 pos. 112–133, snoRNA02 pos. 101–121, snoRNA04 pos. 97–120, and U4 pos. 114–133 [11]. The oligonucleotides were used separately in control experiments (not shown).

significantly more AU-rich than the rest of the molecule. These nucleotides make up most of the spacer between the two hairpins and contain in all cases a stretch of 8–9 purines in the middle. We are currently trying to analyse the secondary structure of the snoRNAs by chemical and enzymatic probing of the RNA and by phylogenetic studies. Dideoxy sequencing of snoRNA04 from *T. pyriformis* revealed 9 substitutions in 96 nucleotides analysed, all of which were in accordance with the proposed structure (not shown).

In order to demonstrate the nucleolar localization of the snoRNAs separately, we devised a reverse transcription experiment. Nuclear and nucleolar RNAs were primed with a primer cocktail containing oligonucleotides complementary to nucleolar (U3) and nucleoplasmic (U1 and U5) controls plus the snoRNA in question. The outcomes of four such experiments are shown in Fig. 3. In Fig. 3A, the intensity of the bands representing extension products on U1 (in longer exposures of the autoradiogram) and U5 templates is clearly diminished in the nucleolar lane (No) compared to the nuclear lane (Nu), whereas U3 and snoRNA01 are present in slightly diminished or constant amounts, respectively. We therefore conclude that snoRNA01 is highly enriched in nucleoli. Similar results were obtained with snoRNA02 (Fig. 3C), snoRNA03 (not shown) and snoRNA04 (Fig. 3D). In Fig. 3B shows the analysis of a RNA which was originally thought to be a snoRNA because of its gel-electrophoretic mobility. This RNA is clearly not enriched in nucleoli. Sequence analysis has demonstrated this RNA to be snRNA U4 [10].

4. DISCUSSION

Nucleolar RNA families have been described from mammals and yeast. The mammalian family comprise U3, U8, U13, U14, X and Y RNAs, all of which have an m_3G cap structure and contain two short nucleotide sequences, Boxes C and D. These RNAs can be precipitated from extracts by anti-fibrillarin antibodies [13,14]. The yeast snoRNA family comprise at least 10 members and include the U3 and U14 homologues [15]. They carry an m_3G cap, and some, but not all, have a recognizable Box C sequence. All of the yeast snoRNAs can be precipitated by anti-fibrillarin antibodies (op. cit.). The *Tetrahymena* snoRNAs described in this paper show no extensive sequence or secondary structure resemblance to the above mentioned nucleolar RNAs or to any other small RNA sequences in the public databases. We suggest that the four RNAs are related and thus constitute a new family of nucleolar RNAs, although their common characteristics are somewhat different from those of the above described families: (i) all of the RNAs are highly enriched in nucleoli and are probably exclusively nucleolar in localization. They do not appear to be hydrogen bonded to rRNA ([9] and unpublished observations), as is the case for most of the yeast snoRNAs [15]; (ii) they have approximately the same size, similar nucleotide composition and absence of modified nucleotides [9]; (iii) they can be folded into similar secondary structures; (iv) they have blocked 5′ ends. They could not be labelled by kinasing with polynucleotide kinase and [γ-^{32}P]ATP with or without previous treatment of the RNA with alkaline phosphatase. The cap is probably neither an m_3G cap nor a γ-methyl cap since enzymatic (pyrophosphatase) and chemical (periodate/aniline [16]) decapping was unsuccessful ([9] and unpublished observations); (v) none of the snoRNAs have a recognizable Box C. Such a sequence in found in U3 from *Tetrahymena* (H. Nielsen et al., manuscript in preparation) which can be immunoprecipitated by heterologous anti-fibrillarin antibodies (Ron Pearlman, personal communication). Work is in progress to find out if this is also true for the snoRNAs; and finally (vi) Southern blot analyses indicate that the snoRNAs are encoded by single copy genes (data not shown). The gene encoding snoRNA04 has been cloned and shown to resemble other polymerase III genes of the of the U6/7SK type [17]. In conclusion, the snoRNAs described in this paper constitute a relatively homogenous family, so far comprising four members. Work is in progress to further characterize their structure and to assess their function in the nucleolus. In this respect, it is important that an in vitro pre-rRNA processing system [18], as well as a gene replacement protocol for rDNA [19], is available for *Tetrahymena*.

Acknowledgements: This work was supported by The Danish Medical Research Council and The Carlsberg Foundation. We would like to thank Mr. Franz Frenzel for skillful technical assistance and Ms. Susanne Søgaard Nielsen for help in preparation of Fig. 1.

REFERENCES

[1] Scheer, U. and Benavente, R. (1990) BioEssays 12, 14–21.
[2] Gerbi, S.A., Savino, R., Stebbins-Boaz, B., Jeppesen, C. and Rivera-Leon, R. (1990) in: The Ribosome (Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D. and Warner, J.R. eds.), pp. 452–469, Am. Soc. Microbiol., Washington, DC.
[3] Engberg, J. (1985) Eur. J. Cell Biol. 36, 133–151.
[4] Gall, J.G. (1986) The Molecular Biology of Ciliated Protozoa, Academic Press, Orlando, FL.
[5] Flickinger, J. (1965) J. Cell Biol. 27, 519–529.
[6] Higashinakagawa, T., Sezaki, M. and Kondo, M. (1979) Dev. Biol. 69, 601–611.
[7] Higashinakagawa, T., Narushima-Iio, M., Saiga, H., Kondo, S. and Mita, M. (1992) Chromosoma (in press).
[8] Engberg, J. and Nielsen, H. (1990) Nucleic Acids Res. 18, 6915–6919.
[9] Pedersen, N., Hellung-Larsen, P. and Engberg, J. (1985) Nucleic Acids Res. 13, 4203–4224.
[10] Ørum, H., Nielsen, H. and Engberg, J. (1991) J. Mol. Biol. 222, 219–232.
[11] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[12] Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning, Cold Spring Harbor Press, NY.
[13] Tyc, K. and Steitz, J.A. (1989) EMBO J. 8, 3113–3119.

[14] Baserga, S.J., Yang, X.W. and Steitz, J.A. (1991) EMBO J. 10, 2645–2651.
[15] Tollervey, D. and Hurt, E.C. (1990) Mol. Biol. Rep. 14, 103–106.
[16] Furuichi, Y. and Shatkin, A.J. (1989) Methods Enzymol. 180, 164–176.
[17] Ørum, H., Nielsen, H. and Engberg, J. (1992) J. Mol. Biol. (in press).
[18] Kister, K.-P., Kaffenberger, W. and Eckert, W.A. (1988) Eur. J. Cell Biol. 46, 233–243.
[19] Sweeney, R. and Yao, M.-C. (1989) EMBO J. 8, 933–938.