Sequence and proposed secondary structure of the *Tetrahymena thermophila* U3-snRNA

Henrik Ørum, Henrik Nielsen[1] and Jan Engberg
The Royal Danish School of Pharmacy, Department of Biology, 2 Universitetsparken, DK-2100 Copenhagen and [1]Department of Biochemistry B, The Panum Institute, University of Copenhagen, 3 Blegdamsvej, DK-2200 Copenhagen, Denmark

Received April 14, 1993; Accepted April 17, 1993 EMBL accession no. X71349

A nucleolar snRNA previously proposed to be the *T.thermophila* homologue of mammalian U3-snRNA (1) was isolated by preparative polyacrylamide get electrophoresis and sequenced by a combination of established methods (2). The sequence identified two very similar RNA species of identical size (256 nt.) and with unique 5′ and 3′ ends (Figure 1).

Several features of the sequence warrant the classification of the isolated snRNA as the *T.thermophila* homologue of the mammalian U3-snRNA. At the primary sequence level the snRNA matches the phylogenetically conserved U3-snRNA sequence elements termed box A, C and D (cf refs. 3–5). Moreover, the location of these boxes within the sequence and their relative spacing are similar to those found in U3-snRNAs of other species. However, no clear equivalent to the phylogenetically conserved box B (cons:AG$^{U}/_{C}$G$^{U}/_{A}$GAA; ref.3) can be identified in the *T.thermophila* sequence. At the secondary structure level, the 3′ two-thirds of the sequence can be folded into a compound helical structure that closely resembles that proposed for the U3-snRNAs of other organisms (3–5) whereas the 5′ third of the snRNA does not appear to be able to form a stable secondary structure.

Southern hybridization data indicate that there are four U3-snRNA genes per haploid genome equivalent in *T.thermophila* (data not shown). Two of these, U3-1 and -2, were isolated from a genomic library by colony hybridization and the remaining two, U3-3 and -4, were cloned by PCR amplification using primers directed against the 5′ and 3′ end of the coding region. The sequences of all four genes exhibit minor nucleotide differences all of which are consistent with the proposed secondary structure (Figure 1). The sequence of the U3-snRNA was colinear with the sequence of both the U3-1 gene and the U3-3 PCR clone. The sequence alterations associated with the U3-2 gene and U3-4 PCR clone could not be detected in the U3-snRNA preparation, suggesting that these genes were not expressed under the experimental growth conditions examined (exponential growth).

REFERENCES

1. Pedersen,N., Hellung-Larsen,P. and Engberg,J. (1985) *Nucleic Acids Res.* **13**, 4203–4223.
2. Ørum,H., Nielsen,H. and Engberg,J. (1991) *J. Mol. Biol.* **222**, 219–232.
3. Jeppesen,C., Stebbins-Boaz,B. and Gerbi,S. (1988) *Nucleic Acids Res.* **16**, 2127–2149.
4. Wise.J.A. and Weiner,A.M. (1980) *Cell* **22**, 109–118.
5. Hughes,J.M.X., Konings,D.A.M. and Cesareni,G. (1987) *EMBO J.* **6**, 2145–2155.

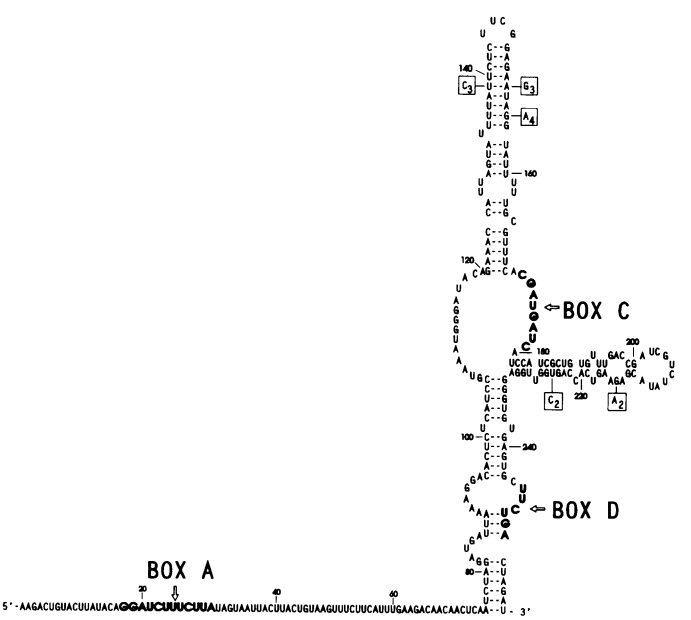

Figure 1. Sequence and proposed secondary structure of the *T.thermophila* U3-snRNA. The sequence variations found in the U3-2, -3 and -4 genes are shown in boxes with subscripts referring to the corresponding gene. The nucleotides constituting the putative boxes A, C and D are shown in bold letters. Primers used for PCR amplification of the U3-genes annealed to position 15–38 and 226–243.