Fibrillarin-associated Box C/D Small Nucleolar RNAs in Trypanosoma brucei

SEQUENCE CONSERVATION AND IMPLICATIONS FOR 2'-O-RIBOSE METHYLATION OF rRNA*

Received for publication, February 11, 2000
Published, JBC Papers in Press, March 9, 2000 DOI 10.1047/jbc.M001180200

David A. Dunbar‡, Steven Wormsley§, Todd M. Lowe¶, and Susan J. Baserga†**

From the Departments of ‡Therapeutic Radiology, §Cell Biology, and ¶Genetics, Yale School of Medicine, New Haven, Connecticut 06520-8040 and the †Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

We report the identification of 17 box C/D fibrillarin-associated small nucleolar RNAs (snoRNAs) from the ancient eukaryote, Trypanosoma brucei. To systematically isolate and characterize these snoRNAs, the T. brucei cDNA for the box C/D snoRNA common protein, fibrillarin, was cloned and polyclonal antibodies to the recombinant fibrillarin protein were generated in rabbits. Immunoprecipitations from T. brucei extracts with the anti-fibrillarin antibodies indicated that this trypanosomalid has at least 30 fibrillarin-associated snoRNAs. We have sequenced seventeen of them and designated them TBR for T. brucei RNA 1–17. All of them bear conserved box C, D, C’, and D’ elements, a hallmark of fibrillarin-associated snoRNAs in eukaryotes. Fourteen of them are novel T. brucei snoRNAs. Fifteen bear potential guide regions to mature rRNAs suggesting that they are involved in 2'-O-ribose methylation. Indeed, eight ribose methylations have been mapped in the rRNA at sites predicted by the snoRNA sequences. Comparative genomics indicates that six of the seventeen are the first trypanosome homologs of known yeast and vertebrate methylation guide snoRNAs. Our results indicate that T. brucei has many fibrillarin-associated box C/D snoRNAs with roles in 2'-O-ribose methylation of rRNA and that the mechanism for targeting the nucleotide to be methylated at the fifth nucleotide upstream of box D or D’ originated in early eukaryotes.

In all eukaryotes the rRNA genes are transcribed in the nucleolus as large 35–45 S precursor transcripts. In yeast and metazoans the rRNA precursor is processed into the mature 18 S, 5.8 S, and 28 S (25 S in yeast) rRNAs of the ribosome (1, 2). In trypanosomes, however, the large subunit rRNA (28 S) is further processed into six rRNAs, called 28 Sα, 28 Sβ, sr1, sr2, sr4, and sr6 (3–5). These processing steps are not peculiar to trypanosomes because Euglena gracilis, which shares a common ancestor with trypanosomatids, also contains a multiply fragmented 28 S rRNA (6, 7). It has been hypothesized that the origin of contiguous high molecular weight rRNAs started from an ancient ribosome that consisted of primarily fragmented rRNAs (4, 8, 9).

Small nucleolar RNAs (snoRNAs)1 are required for both processing of the pre-rRNA precursor and in the extensive nucleotide modification (2'-O-ribose methylation and pseudouridine formation) that occurs on the rRNAs (reviewed in Refs. 2, 10, 11–22). There are two major classes of snoRNAs that are named for specific conserved nucleotide sequences: box C/D and box HI/ACA snoRNAs (23). In yeast, the box C/D snoRNAs are characterized by their association with the nuclear proteins fibrillarin, Nop5/Nop58, and Nop56 (and perhaps other proteins, 24, 25–29) in small nucleolar ribonucleoproteins. Of the 41 box C/D snoRNAs predicted to function in 2'-O-ribose methylation of rRNA in Saccharomyces cerevisiae, 37 have been experimentally confirmed (30, 31). Only four other known box C/D snoRNAs in yeast perform other functions (U3) or have no assigned function (snR190, snR4, and snR45). Many of the vertebrate box C/D snoRNAs are also required for site-specific rRNA 2'-O-ribose methylation (31–35). However, the vertebrate U3, U22, and U8 box C/D snoRNAs are required for pre-rRNA cleavage events (36–40). The U14 snoRNA is unusual in that it is the only box C/D snoRNA shown so far to function in both 18 S rRNA maturation and 18 S rRNA nucleotide modification (35, 41–43).

The box C/D snoRNAs involved in nucleotide modification, also called methylation guide RNAs, are characterized by the presence of conserved box C and D sequences near their 5’- and 3’-ends, respectively. The methylation guide RNAs also often bear internal box D’- and box C’-sequences (44, 45). Most methylation guide RNAs have a single region of complementarity to mature rRNA regions upstream from their box D or D’ sequences; however, some of them contain two complementary sequences and are referred to as double methylation guide RNAs. Results from several laboratories suggest that the fifth nucleotide upstream of the box D or D’ sequence, within the complementary sequence, specifies the ribose of the nucleotide to be methylated in the target rRNA (31, 32, 35). Investigation of functional constraints on the guide RNA-rRNA duplex indicates that both the length and composition of the complementary sequence influence the extent of the methylation reaction (46).

Our current understanding of the trypanosomatid snoRNAs involved in ribosome biogenesis is just beginning to emerge.

1 The abbreviations used are: snoRNA, small nucleolar RNA; GAR, glycine and arginine rich; TBR, T. brucei RNA 1–17; PCR, polymerase chain reaction; nt, nucleotide(s); TMG, trimethylguanosine.

2 D. Lafontaine and T. Tollervey, personal communication.

* This work was supported by a grant from the Council for Tobacco Research (to S. J. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF168719-AF168736.

** To whom correspondence should be addressed: Yale School of Medicine, P. O. Box 208040, 333 Cedar St., New Haven, CT 06520-8040. Tel.: 203-785-4618; Fax: 203-785-6309; E-mail: susan.baserga@yale.edu.

This paper is available on line at http://www.jbc.org
Previous studies revealed the existence of a U3 homolog in Trypanosoma brucei (47–50). The first methylation guide box C/D snoRNA in trypanosomes was described by Levitan et al. (51) who determined that it has the potential to guide methylation of a 5.8 S rRNA nucleotide. Two additional box C/D snoRNAs from Leishmania tarentolae, Trypanosoma cruzi, and T. brucei were identified by Roberts et al. (52). However, it is clear from the work of Hartshorne and Agabian (48) that T. brucei has many more fibrillarin-associated snoRNAs. The identity of the other fibrillarin-associated snoRNAs in T. brucei remains unknown.

A comparison of the box C/D snoRNA sequences from such widely divergent species as trypanosomes, yeast, and vertebrates is likely to yield important information about the evolution of ribosome biogenesis, particularly with reference to pre-rRNA cleavage and 2'-O-ribose methylation. We reasoned that the box C/D snoRNAs in T. brucei could be isolated and sequenced by enrichment via immunoprecipitation with anti-fibrillarin antibodies; however, antibodies specific to the T. brucei fibrillarin are not available and antibodies that cross-react are scarce. To identify the box C/D snoRNAs in trypanosomes in a systematic way, we cloned the T. brucei fibrillarin cDNA and determined that it bears both the glycine- and arginine-rich (GAR) and methyltransferase domains present in fibrillarins characterized in other eukaryotes. We expressed the T. brucei fibrillarin in Escherichia coli and purified it and then used the recombinant protein to generate polyclonal antibodies in rabbits. Immunoprecipitation experiments on T. brucei extracts using the anti-fibrillarin antibodies indicated that trypanosomes contain at least thirty fibrillarin-associated snoRNAs. Seventeen of them were sequenced and designated TBR for T. brucei RNA 1–17. Their sequences and mapping of the methylation sites in the rRNA revealed that, like in yeast and metazoans, many have the potential to be guide RNAs for 2'-O-ribose methylation of rRNA, suggesting that the box C/D snoRNAs exist in early eukaryotes. Interestingly, six of them appear to be homologs of methylation guide snoRNAs found in yeast and vertebrates. This implies that the mechanism of specifying the methylation site in rRNA has been conserved from an ancient eukaryote.

**EXPERIMENTAL PROCEDURES**

**Growth of T. brucei**—The procyclic form of Trypanosoma brucei rhodesiense YTAT1.1 strain obtained from Elisabetta Ullu (Yale University School of Medicine) was used in this study. Cells were grown at 28 °C in SM medium supplemented with 20% fetal calf serum (53). The procyclic form of Trypanosoma brucei—Growth of T. brucei—The procyclic form of Trypanosoma brucei was grown at 37 °C in SM medium supplemented with 20% fetal calf serum (53). The procyclic form of Trypanosoma brucei was generated by a heat shock of 40 °C for 1 h and twice in 1 N HCl.

**Reverse transcription on total RNA**—The cDNA clone representing this expressed sequence tag was made in the following way. Seven of the cloned T. brucei sequences and identified a candidate partial clone in Leishmania tarentolae, Trypanosoma cruzi, and T. brucei. The first methylation guide box was used to screen a cDNA library (from strain YTaT1.1) cloned into Lambda Zap (Stratagene). The library was screened by performing PCR in the presence of complementary to the 3'-ends of snoRNA sequences (see Table I). RNA was isolated from an anti-fibrillarin immunoprecipitation performed on 50 ml of T. brucei cells and used for individual primer extension reactions to generate full-length cDNAs. These cDNAs were gel-eluted and poly(A)-tailed in 25 μl of buffer containing 0.1 μM potassium cacodylate (pH 7.2), 20 mM MgCl₂, 0.2 mM diithiothreitol, 0.2 mM dATP, and 15 μM of each terminal deoxynucleotidyl transferase. Incubation was performed by Immunization Services at the Yale School of Medicine. The purified, and plasmids recovered according to the manufacturer's instructions. Automated DNA sequencing of one positive clone was carried out on an Applied Biosystems 373 Stretch sequencer by primer walking of both strands.

**Expression and Purification of the T. brucei Fibrillarin Protein in E. coli**—The polymerase chain reaction was used to amplify the full-length fibrillarin cDNA with the appropriate restriction sites for cloning in frame into the E. coli expression vector PET28a (Novagen). This vector places a 6x histidine tag on the amino terminus of the protein. The correct oligonucleotide contains a BamH1 site (Tbfib.3, 5'-GAGCGGTATCTGCGAGTTGGTTGGCG) and also a 5'-BamHI site (Tbfib.3, 5'-CCCGCGGATTCTGCGAGTTGGTTGGCG). The 3'-oligonucleotide is complementary to the last 15 nucleotides and has an AscI site (Tbfib.4, 5'-GTCAGGCTGAGCATATTATTTGTTGACTGG). These oligonucleotides were used in the polymerase chain reaction with the fibrillarin plasmid as a template with the following cycling steps: 94 °C for 15 s, and 72 °C (15 s). The product was purified by ethanol precipitation, digested with BamHI and AscI, and the resulting band was gel-purified. This fragment was ligated into BamHI, and XhoI cut PET28a.

The fibrillarin cDNA cloned into pET28a was transformed into BL21 (DE3) cells expressing the E. coli expression and protein purification. Partial solubility of the fibrillarin protein was obtained when the cells were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside and grown overnight at 30 °C. The resulting soluble protein was purified by metal chelation chromatography and eluted from the column with 1 M imidazole. The fibrillarin protein, insoluble upon elution, was dialyzed overnight into phosphate-buffered saline prior to injection into rabbits. Injections were performed by Immunization Services at the Yale School of Medicine. The antibodies were checked for reactivity with fibrillarin expressed in E. coli by Western blot analysis.

**Immunoprecipitations and RNA Analysis**—For anti-fibrillarin immunoprecipitation experiments, 3 mg of protein A-Sepharose CL-4B (Amersham Pharmacia Biotech) was mixed with 50 μl of either rabbit anti-fibrillarin or preimmune serum in 0.5 ml of NET-2 (40 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.05% Nonidet P-40) overnight at 4 °C on a rotating nutator. The bound antibodies were washed four times with cold NET-2 and used as described below. For anti-trimethylguanosine (TMG) experiments, 15 μl of anti-TMG antibody and 15 μl of rabbit anti-mouse IgG were mixed with protein A-Sepharose and then incubated and washed in NET-2 as described above.

T. brucei cells were grown to a density of 2 × 10⁷ cells/ml. Cells from a 50 ml culture were pelleted and resuspended in 10 ml of wash buffer (20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 3 mM MgCl₂). Cells were repelleted and resuspended in 0.6 ml of either NET-2, NET-5, NET-6, or NET-7 containing protease inhibitors (2 μg of aprotinin/ml, 1 μg of leupeptin/ml, and 1 μg of pepstatin A/ml). NET-5, 6, and 7 are the same as NET-2 except that they contain either 500, 600, or 700 mM NaCl, respectively. Cells were lysed by vigorous vortexing with one-half volumes (65–65 μM NaCl) for five minutes. The lysate was cleared by centrifugation at 13,000 × g for 10 min at 4 °C. The cleared lysate containing either 150, 500, 600, or 700 mM NaCl was added to the protein A-Sepharose beads treated with either rabbit anti-fibrillarin serum, preimmune serum, or anti-TMG antibodies and incubated for 1 h at 4 °C on a rotating nutator. The beads were washed six times with either NET-2, -5, -6, or -7 buffer. RNA was recovered by addition of NET-2 (25 μl) and extraction with phenol/chloroform/isooamyl alcohol and ethanol precipitation. Immunoprecipitated RNAs were labeled with 5'-³²P-pyridine-3',5'-bisphosphate and T4 RNA ligase according to Ref. 55 and separated on a 10% sequencing gel.

**Sequence and Cloning of Trypanosome snoRNAs**—For direct RNA sequencing, immunoprecipitations and 3'-end labeling of snoRNAs were performed as described above. Individual labeled snoRNAs were gel-isolated and eluted in 0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA (pH 8.0), 0.1% SDS overnight at room temperature on a rotating nutator. Approximately 30 nt of the 3'-end of labeled RNAs was obtained by enzymatic sequencing using RNAses T1, U2, Phym, and Bacillus cereus (Amersham Pharmacia Biotech).

Full-length cDNAs were obtained using Superscript II reverse transcriptase (Life Technologies, Inc.) with ³²P-labeled deoxyoligonucleotides complementary to the 3'-ends of snoRNA sequences (see Table I). RNA was isolated from an anti-fibrillarin immunoprecipitation performed on 50 ml of T. brucei cells and used for individual primer extension reactions to generate full-length cDNAs. These cDNAs were gel-eluted and poly(A)-tailed in 25 μl of buffer containing 0.1 μM potassium cacodylate (pH 7.2), 20 mM MgCl₂, 0.2 mM dithiothreitol, 0.2 mM dATP, and 15 μM of each terminal deoxynucleotidyl transferase. Incubation was performed by Immunization Services at the Yale School of Medicine. The poly(A)-tailed cDNAs were extracted with phenol/chloroform/isoamyl alcohol and ethanol-precipitated. The poly(A)-tailed cDNAs were amplified using a (dT) oligomer and the 3'-oligomer by PCR with these cycling steps: 94, 55, and 72 °C (30 s each, 30 cycles). The PCR products were cloned using the CLONTECH TA-cloning kit and sequenced using an Applied Biosystems 373 Stretch sequencer.
The sequences of the 3'-ends were verified independently. Adenosine nucleotides were added to the 3'-ends of total snoRNAs isolated from an anti-fibrillarin immunoprecipitation experiment with 5 units of poly(A) polymerase (Amersham Pharmacia Biotech) in 50 μl containing 40 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 2.5 mM MnCl₂, 250 mM NaCl, 250 μM ATP, 50 μg/ml bovine serum albumin for 30 min at 37 °C. The 3'-poly(A)-tailed RNAs were extracted with phenol/chloroform/isoamyl alcohol and ethanol-precipitated. Full-length cDNAs complementary to the poly(A)-tailed RNAs were made with a (dT) oligomer and Superscript II reverse transcriptase. Aliquots of the primer extension reaction were then used to amplify individual snoRNAs by PCR using a (dT) oligomer and an oligomer containing the 5'-sequence specific for a particular snoRNA using the cycling conditions described above (Table I). The PCR products were cloned and sequenced as described above.

Northern Blot Analysis of snoRNAs—To isolate T. brucei total RNA, a 50-ml culture of cells grown to a density of 2–3 × 10⁷ cells/ml was pelleted and resuspended in wash buffer. Cells were repelleted and resuspended in 250 μl of solution D (4 M guanidinium thiocyanate, 26.4 mM sodium acetate, pH 7.0, 0.5% sarcosyl, 0.72% β-mercaptoethanol). The lysed cells were extracted with an equal volume of acid phenol and 50 μl of chloroform and precipitated with three volumes of ethanol. RNA from an anti-TMG cap immunoprecipitation and an anti-fibrillarin immunoprecipitation were resolved with total RNA on a 10% sequencing gel and transferred to a Zeta-probe membrane. Plasmids containing the TBR17 snoRNA clone and a plasmid containing the T. brucei U4 small RNA were used in PCR reactions with clone-specific primers and [α-32P]dCTP (300 Ci/mmol) to make labeled probes. Blots were hybridized to 1 × 10⁶ cpm of labeled probes in 3 x saline/sodium phosphate/EDTA, 10 x Denhardt’s solution, 7% SDS at 65 °C. Blots were washed twice in 3 x SSC 0.1% SDS at room temperature for 15 min each wash and once in 3 x SSC 0.1% SDS at 65 °C for 10 min and exposed to x-ray film.

Mapping of 2'-O-Ribose Methylations by Primer Extension—The ribose methylation mapping protocol was modified from Ref. 56. Oligonucleotides complementary to rRNA 3' to sites of predicted 2'-O-ribose methylations were synthesized. 7 μg of T. brucei total RNA was annealed to 15 ng of the 32P-labeled oligonucleotides for 10 min at 70 °C. Primer extension reactions were carried out in 20 μl of a buffer containing 50 mM Tris-HCl, pH 8.5, 8 mM MgCl₂, 30 mM KCl, and 1 mM dithiothreitol in the presence of 0.5 units/μl avian myeloblastosis virus reverse transcriptase (Roche Molecular Biochemicals) with either 1 mM dNTPs/μl or 0.004 mM dNTPs/μl for 1 h at 42 °C. The extension products were resolved on an 8% sequencing gel next to an RNA sequencing ladder. For generation of RNA sequencing ladders, 5 μg of T. brucei total RNA

---

**TABLE I**

| Oligonucleotides used to clone the T. brucei snoRNAs |
|-----------------------------------------------------|
| The 5’-oligos were used to identify snoRNA 3’-ends, and the 3’-oligos were used to identify snoRNA 5’-ends. See text for descriptions of both the 5’- and 3’-snoRNA cloning strategies. |

| TBR1 5’ | 5’GGCCATAGGATAAAAAACGCTC3’ |
| TBR1 3’ | 5’CGCTCTGGATGGA3’ |
| TBR2 5’ | 5’CCGTTGAAGAGTTGCGACGGG3’ |
| TBR2 3’ | 5’CATGAGATATTCTGCAACGCT3’ |
| TBR3 5’ | 5’AACTGACAAAAAACCCTTCA3’ |
| TBR3 3’ | 5’TCGAAGTTCGCTTGATGGC3’ |
| TBR4 5’ | 5’CGGTTGATTAGCGAGCTCTCCACCTAA3’ |
| TBR4 3’ | 5’TCTCATTTACAAATTTGATATAA3’ |
| TBR5 5’ | 5’GAAGTGTGAGACACCTAGGCCC3’ |
| TBR5 3’ | 5’ATCATGAGACATGAAATGTT3’ |
| TBR6 5’ | 5’GGCCGAGTAGACATCACAAGT3’ |
| TBR6 3’ | 5’TTGAGATTTTCTTGTG3’ |
| TBR7 5’ | 5’GAAATTGGACAGACAAAAACCT3’ |
| TBR7 3’ | 5’TCTGAGGGGCAACAATTTCC3’ |
| TBR8 5’ | 5’GGACAGTGAHTAATATGTC3’ |
| TBR8 3’ | 5’TCAGTGCAAGCTGTTTATGGGTGC3’ |
| TBR9 5’ | 5’CGCTGAGAAGTTGATGATG3’ |
| TBR9 3’ | 5’CTGAGAAAGATGCTTGTTTGCTG3’ |
| TBR10 5’ | 5’CCAGAGGGATTGTCATATTAT3’ |
| TBR10 3’ | 5’CATTCTAGTTTCTCCCTC3’ |
| TBR11 5’ | 5’TATCATGATATGAGACACCGCTGCTG3’ |
| TBR11 3’ | 5’TCTCATTTTACTAATTTGATATAA3’ |
| TBR12 5’ | 5’CTGAGATAGTGCAACCTCTTAC3’ |
| TBR12 3’ | 5’GTGAGTAGTACAGGATGAAATC3’ |
| TBR13 5’ | 5’CAAATGATGTCAATACTCAGGCGATTT3’ |
| TBR13 3’ | 5’CATGAGATTTGTCACCTCCAT3’ |
| TBR14 5’ | 5’CAGATGATTTCTCAGCTTGATCTG3’ |
| TBR14 3’ | 5’TCAAGTTGCGGTAATGTC3’ |
| TBR15 5’ | 5’CCCATGTAGTGTAGTGCAACATTC3’ |
| TBR15 3’ | 5’TCAAGATCTTCTGGTGG3’ |
| TBR16 5’ | 5’AGCATGATGATCATAGCTTGCAAT3’ |
| TBR16 3’ | 5’TGGAGCTTGTGTTTTG3’ |
| TBR17 5’ | 5’CCCGGCAAGGCTGAGCCTGAGGACACACCC3’ |
| TBR17 3’ | 5’GGTACAGTTATAAGCTGATGGTGTGCTTAA3’ |
was hybridized to 15 ng of labeled oligonucleotide for 10 min at 70 °C. Dideoxy RNA sequencing was performed in a 20-μl reaction using 10 units/μl Superscript II reverse transcriptase, 1 mM dNTPs/μl and 1 mM of the appropriate ddNTP at 42 °C for 1 h.

RESULTS

Cloning and Expression of T. brucei Fibrillarin—A radiolabeled short fragment (600 nt) of the putative T. brucei fibrillarin gene was used to screen a T. brucei cDNA library to obtain the full-length fibrillarin clone. The nucleotide and amino acid sequence are shown in Fig. 1A. The T. brucei fibrillarin is 300 amino acids long and has a predicted molecular mass of 31.7 kDa. The T. brucei fibrillarin is therefore among the shortest of the fibrillarins, the shortest being the Tetrahymena fibrillarin at 292 amino acids (57). Like most of the other fibrillarins, the T. brucei fibrillarin has a GAR domain at its amino terminus (58) and a potential S-adenosylmethionine binding/methyltransferase domain (amino acids 141–157; 59, 60). Comparison of the T. brucei fibrillarin to the nine other known eukaryotic fibrillarins (Fig. 1, B and C) indicates a high degree of conservation. It is most identical to the fibrillarin of a related trypanosomatid, L. major, with 84% sequence identity; it is least similar to the Giardia lamblia fibrillarin with 47% identity. Fig. 1B indicates that the highest sequence conservation among fibrillarins occurs carboxyl-terminal to the GAR domain.

Previous studies indicate that the human fibrillarin can re-
store growth to a *S. cerevisiae* strain with a null fibrillarin allele (Δnop1), though the strain is temperature-sensitive (61). The human and *S. cerevisiae* fibrillarins are 60% identical/68% similar. We assessed whether the *T. brucei* fibrillarin, which is 53% identical/64% similar to the budding yeast fibrillarin, could restore growth at the nonpermissive temperature to strains with five temperature-sensitive fibrillarin alleles (nop1.2, nop1.3, nop1.4, nop1.5, and nop1.7; 62). The *T. brucei* fibrillarin cDNA was cloned into the yeast expression vector, p415GPD (63), and transformed into the temperature-sensitive strains. Subsequent restreaking of the colonies and growth at both 22 and 37 °C indicated that the *T. brucei* fibrillarin does not complement these temperature-sensitive fibrillarin alleles (data not shown). This suggests that the *T. brucei* fibrillarin is too dissimilar to function in *S. cerevisiae*. It is also possible that the *T. brucei* fibrillarin cannot attain the correct cell compartment in yeast because its nuclear import/nucleolar targeting signals are different.

The *T. brucei* fibrillarin was cloned into pET28a for expression in *E. coli* as a histidine-tagged fusion protein. When the cells were induced and grown at 37 °C, the fibrillarin protein is insoluble, and upon solubilization in 6 M guanidine, HCl does not bind to the metal chelation column. Growth at 30 °C allowed partial solubility, and the protein was purified under nondenaturing conditions. However, when fibrillarin was eluted from the metal chelation column it became insoluble and could be visualized in column fractions as a cloudy precipitate. Gel electrophoresis of this precipitate revealed a protein of the expected mobility (32 kDa). The cloudy precipitate was injected into rabbits for the production of antibodies.

**Anti-fibrillarin Antibodies Immunoprecipitate the U3 and Many Other snoRNAs**—Sera from two rabbits were tested for reactivity with the fibrillarin produced in *E. coli* by Western blots, and one was chosen for further study. Immunoprecipitations were performed on *T. brucei* whole cell extracts with anti-fibrillarin polyclonal antibodies and compared with immunoprecipitations performed with preimmune serum at different salt concentrations (0.5–0.7 M). RNAs from immunoprecipitations were labeled at their 3'-ends and resolved on a 10% denaturing polyacrylamide gel. Anti-fibrillarin antibodies immunoprecipitated at least thirty specific bands at all three sodium chloride concentrations (Fig. 2, lanes 2, 4, and 6). None of these RNAs were immunoprecipitated with preimmune serum, indicating that they result from co-immunoprecipitation with anti-fibrillarin antibodies (compare lanes 1 and 2, 3 and 4, and 5 and 6).

Several strategies were used to determine the snoRNA sequences. The full-length sequences of RNAs TBR1–10 (Fig. 2, lanes 2, 4, and 6) were determined by direct RNA sequencing and by using a 5'-end cloning strategy. Because some of the nucleotides obtained by direct enzymatic sequencing were ambiguous, the RNA 3'-end sequence was confirmed using a 3'-end cloning strategy. The sequences of TBR11, -13, -15, and -16 were obtained when attempting to clone other snoRNAs with deoxyoligonucleotides based on the RNA 3'-end sequence. Similarly, the 3'-end of TBR12 was obtained when attempting to verify the 3'-end sequence of a different snoRNA. The 5'-end sequence of the *T. brucei* homolog of the *L. collosoma* snoRNA-2 was obtained using the 5'-end cloning strategy with a labeled deoxyoligonucleotide complementary to the 3'-end of the *Leptomonas collosoma* snoRNA-2 (51). The 3'-end sequence of this snoRNA was verified using the 3'-end cloning strategy and named TBR14. The presence of each snoRNA in total RNA and in fibrillarin immunoprecipitates was verified by Northern blot analysis (data not shown).

Sequencing revealed that all sixteen snoRNAs have box C, D, C', and D' sequences, characteristic of snoRNAs associated with fibrillarin (Fig. 3). Box C and D sequences are located at the 5'- and 3'-ends of the RNAs, respectively, as are all but one (snR13) of the 2'-O-ribose methylation guide RNAs described previously in yeast and vertebrates. Fourteen of the snoRNAs that we have sequenced are newly identified in *T. brucei*. TBR5 and TBR7 were previously identified and proposed to be box C/D snoRNAs (52, 64) but had not been shown to be associated with the fibrillarin protein. Direct RNA sequencing suggested the presence of at least two TBR10 species, A and B, and revealed the previously identified U3 snoRNA and a U3 3'-end breakdown product that migrates just above 5 S rRNA (Fig. 2).

Using a snoRNA search algorithm and model scoring program, the snoRNAs were tested for complementarity to the *T. brucei* rRNAs, U2, U3, U4, U6, spliced leader RNA, spliced leader associated RNA, and 7SL RNA. This program has been used to identify 22 novel methylation guide snoRNAs in yeast; the algorithm and scoring scheme are described elsewhere (30). Complementary sequences that gave the highest scores for each individual snoRNA are shown in Fig. 3. Because they have complementarity to rRNA upstream of box D or D', 15 of them are potential methylation guide snoRNAs, 4 with sequence complementary to 18 S rRNA (TBR3, TBR7, TBR8, and TBR12), 1 with sequence complementary to both 18 S and 5.8 S rRNA (TBR14), and 8 with sequence complementary to 28 S rRNA. Of the large subunit rRNA methylation guide RNAs, two are complementary to 28 Sα (TBR6, TBR9) and eight are complementary to 28 Sβ (TBR1, TBR2, TBR4, TBR10, TBR11, TBR13, TBR15, and TBR16). Given confirmed methylation sites at the predicted targets, all the potential guide regions yielded scores that were in the same range as the previously characterized yeast snoRNA guide regions. None of the

---

**FIG. 2.** Anti-fibrillarin antibodies immunoprecipitate U3 and many other snoRNAs. Immunoprecipitations were performed on *T. brucei* whole cell extracts with rabbit preimmune serum (lanes 1, 3, and 5) or anti-fibrillarin rabbit serum (lanes 2, 4, and 6) in the indicated salt concentrations. The RNAs isolated from the immunoprecipitations were labeled with 5'-[32P]cytidine 3',5'-bisphosphate and T4 RNA ligase and analyzed on a 10% denaturing polyacrylamide gel. The sizes of the labeled pBR322-MspI markers are indicated.
snoRNAs had a significant degree of complementarity to RNAs other than rRNA. One, TBR5, does not have any significant complementarity to any known RNA.

We mapped some of the 2'-O-ribose methylation sites in trypanosome rRNA to demonstrate that they do indeed occur at sites targeted by the snoRNAs that we have identified (Fig. 4).
Methyllations occur at 28S mA740 (target of TBR9), 28S mA2713 (target of TBR10), 18S mc18 (target of TBR14), 18S mU652 (target of TBR12), 5.8S mG75 (target of TBR14), 28S mC3573 (target of TBR16), and 28S mG3578 (target of TBR11). In all cases, the methylation site occurs complementary to the nucleotide in the snoRNA that is 5 nucleotides upstream of box D or D’, suggesting that trypanosome snoRNAs conform to the “box D + 5 rule” as in other eukaryotes.

Several T. brucei snoRNAs Are Possible Functional Homologs to Yeast and Vertebrate snoRNAs—Trypanosomes are among the earliest diverged eukaryotes, yet our search program has identified several potential T. brucei functional homologs to yeast and vertebrate snoRNAs based on the presence of 5’- and 3’-box C, D, C’, and D’ sequences and rRNA complementary sequences (Table II). We define functional homologs as box C/D snoRNAs with similar (some are identical) rRNA complementary sequences that have the potential to methylate the ribose of the same nucleotide. These include functional homologs to the yeast snR73/vertebrate U35 (TBR1), yeast snR56/vertebrate U25 (TBR8), yeast/snR48/vertebrate U18 (TBR9), yeast snR35 (TBR11), yeast/snR48/vertebrate U60 (TBR11), yeast snR77 (TBR12), and yeast snR13/vertebrate U15 (TBR13). Thus, trypanosomes are the earliest diverged eukaryote with snoRNAs homologous to yeast and humans. This implies conservation of the mechanism to target 2’-O-ribose RNA methylation throughout evolution.

TBR17 Is a Box C/D Fibrillarin-associated snoRNA—Roberts et al. (52, 64) have previously identified a T. brucei snoRNA of 270 nt, which is somewhat larger in size than the other identified trypanosome snoRNAs. We will refer to this snoRNA as TBR17. We have investigated whether it is one of the box C/D fibrillarin-associated snoRNAs. First, as it appeared that Roberts et al. (52) had only a partial T. brucei TBR17 sequence, we cloned the full-length TBR17 snoRNA using the 5’- and 3’-poly(A)-tailing PCR-based technique. The full-length sequence of TBR17 was then used as a probe to examine whether it is associated with fibrillarin.

Immunoprecipitations were performed on trypanosome whole cell extracts using anti-TMG antibodies and anti-fibrillarin antibodies and compared with immunoprecipitations performed with preimmune serum. The Northern blot was hybridized with probes to U3, U4, and TBR17. As expected, the anti-TMG antibodies immunoprecipitate the U3 and U4 RNAs, because they both bear a TMG cap at their 5’-ends (Fig. 5A, lane 2). However, TBR17 is not immunoprecipitated with anti-TMG antibodies, indicating that this RNA does not contain a 5’-TMG cap (Fig. 5A, lane 2). This is consistent with the fact that the majority of the T. brucei box C/D snoRNAs do not possess a 5’-TMG cap (this work, data not shown, and Ref. 48). The anti-fibrillarin antibodies specifically immunoprecipitate both the U3 and TBR17 RNAs (Fig. 5A, compare lanes 3 and 4). These results demonstrate that like U3, TBR17 is a fibrillarin-associated RNA.

The TBR17 sequence that we obtained was 99% identical to the published sequence and contains both box C and D elements (52). However, the 3’-end of TBR17 that we obtained has seven additional nucleotides when compared with the published sequence, including a consensus box D element located between two nucleotides from the 3’-end of the molecule (Fig. 5B). The L. tarentolae TBR17 homolog has a box D element located at its 3’-end as well (52). We also observe a potential box C element located about 95 nucleotides from the 3’-end of the molecule (Fig. 5B). The snoRNA search algorithm could not find potential guide regions to the T. brucei snoRNAs or any of the other RNAs tested. Taken together, these results suggest that TBR17 is a genuine box C/D fibrillarin-associated snoRNA.

**DISCUSSION**

To explore the nature of methylation guide snoRNAs in an ancient eukaryotic organism, we characterized the box C/D snoRNAs of T. brucei by direct sequencing of RNAs immunoprecipitated with anti-fibrillarin antibodies. The T. brucei fibrillarin cDNA was cloned, and the predicted protein sequence was found to bear the GAR and methyltransferase domains and to be conserved across species. The fibrillarin protein was expressed in and purified from E. coli and was used to raise polyclonal rabbit antibodies. Immunoprecipitation of T. brucei extracts with these antibodies indicated that there are at least thirty fibrillarin-associated snoRNAs in this trypanosomatid. We have identified and characterized seventeen of them and have named them TBR1–17. Sixteen have conserved homologs to yeast and vertebrate snoRNAs (Table III). This pattern of box C and D nucleotide usage is generally similar to that observed in yeast and vertebrates (30, 70). One exception is that TBR1 has an atypical box D element located between two nucleotides from the 3’-end of the molecule (Fig. 5B). The snoRNA search algorithm could not find potential guide regions to the T. brucei snoRNAs or any of the other RNAs tested. Taken together, these results suggest that TBR17 is a genuine box C/D fibrillarin-associated snoRNA.
methylation in all other eukaryotes studied. Our computer search algorithm predicts that fifteen of the seventeen snoRNAs are potential guide RNAs for 2'-O-ribose methylation of rRNA (Fig. 3). Because the related trypanosome *Crithidia fasciculata* contains 95–100 sites of 2'-O-ribose methylation (71), it is likely that the rRNA of *T. brucei* is also extensively methylated. Previously, three rRNA methylation sites were either mapped in *T. brucei* RNA or can be inferred from mapping of rRNA in a related trypanosomatid: Gm75 in 5.8 S rRNA, Gm1867 in 18 S rRNA (erroneously specified as 1868), and Gm3968 in large subunit rRNA (51, 52). Levitan et al. (51) sequenced the *L. collosoma* snoRNA that is likely to guide methylation of the 5.8 S rRNA at a position that would be complementary to the fifth nucleotide upstream of the box D sequence. We mapped 7 additional methylation sites in *T. brucei* rRNA: 28 S mA740, 28 S mA2713, 18 S mC18, 18 S mU652, 5.8 S mG75, 28 S mC3573, and 28 S mG3578. We have identified snoRNAs that can mediate the methylation reactions of eight of the mapped sites. Furthermore, 6 of the 15 guide snoRNAs are potential functional homologs to snoRNAs from yeast and/or vertebrates. In each case, the nucleotide whose ribose undergoes methylation is the fifth nucleotide upstream of either box D or box D', within the snoRNA-rRNA complementarity (31, 72). This strongly suggests that trypanosome

| trypanosome snoRNA | yeast snoRNA | human snoRNA |
|--------------------|--------------|--------------|
| LSU 3722 AGACCGUGUGAG 28S | LSU 2956 UAGACCGUGUGA 25S | LSU 4496 AGACCGUGUGAGA 28S |
| 3' - AGUAAAUCUGCCAGCAUUC TBR1 box D' | 3' AGUCACAGUGCCAGCACU snR73 box D' | 3' AGCCCGUCUGGUACACUCU U35 box D' |
| SSU 1867 GCAGGUGUGG 18S | SSU 1425 ACAGGUCUGUG 18S | SSU 1490 ACAGGUCUGUGA 18S |
| 3' - AGUCACAGUGCCAGCACU TBR8 box D | 3' AGUCAGUGCCAGCACU snR56 box D' | 3' AGUCGUGGUCAGACU U25 box D' |
| LSU 740 GAAACCGGAACCA 28S | LSU 647 AAACCGGAACCA 25S | LSU 1306 GAAACCGGAACCA 28S |
| 3' - AGACCUUUGGCGGGUUUC U18 box D' | 3' AGUAAUUGGCGGGUUUC snR48 box D' | 3' AGUCUGGACAGACUUU U60 box D' |
| LSU 3556 AGACGUGUCAG 28S | LSU 2790 UAGAGGUGCCAGAA 25S | LSU 4330 AGAGGUGCCAGAA 28S |
| 3' - AGUCUCUCUCCACAGUC TBR11 box D | 3' UGUAUUCUCAGCGCUU snR48 box D' | 3' AGUAUCGGAACACAGCU U60 box D' |
| LSU 3578 ACAGGGAUACUG 28S | LSU 2812 ACAGGGAUACUG 25S | 3' AGUCUGGACCACU snR38 box D |
| 3' - AGUCUACCCAUAAAGUC TBR11 box D | | |
| SSU 652 GGUAAUUGCC 18S | SSU 578 CCGUAAUUGCCG 18S | |
| 3' - AGUCUACACUAAGUGC TBR12 box D | 3' AGUCCGUCUAAAGGUGCA snR77 box D' | |
| LSU 2853 CAAAGCGUCCUG 28S | LSU 2279 GCGAAAGGC 25S | LSU 3754 CAAAGGCCUC 28S |
| 3' - AGUAAUGUAAAGGCGAGC TBR13 box D | 3' AGUCUGUAAAGGCGA snR13 box D' | 3' AGUCUGUAAAGGCGAGC U15 box D |

**TABLE II**

Guide regions of potential snoRNA homologs between trypanosomes, yeast, and humans showing rRNA complementary nucleotides

The asterisk indicates the rRNA methylated nucleotide according to the box D + 5 rule of yeast and vertebrate box C/D guide snoRNAs (37–39).
rRNA methylation does indeed follow the box D rule, in contrast to a previous proposal (52). Our results and those of Levitan et al. (51) suggest that methylation guide snoRNAs originated early in eukaryotic evolution and that the strategy for targeting methylation according to the box D rule evolved with them.

Studies in several eukaryotic organisms indicate that the biogenesis of the box C/D snoRNAs occurs by a number of strategies (17, 20). In vertebrates, the majority of the box C/D snoRNAs are processed from introns of pre-mRNA transcripts. The yeast box C/D snoRNAs are transcribed as monocistronic or polycistronic RNAs or as introns of pre-mRNAs (73). All known plant box C/D snoRNAs are also processed from polycistronic transcripts (74, 75). It will be interesting to determine the mode of biogenesis of the trypanosome box C/D snoRNAs.

Based on the work of Roberts et al. (52, 64), it is likely that at least some are clustered in the trypanosome genome. Are they also processed from polycistronic transcripts? Which RNA polymerase transcribes the newly identified trypanosome snoRNAs? So far it is known that the trypanosome U3 snoRNA, like the trypanosome U2, U4, and U6 small RNAs, is transcribed by RNA polymerase III (76, 77). Are the other trypanosome box C/D snoRNAs also transcribed by RNA polymerase III or by a different RNA polymerase?

Acknowledgments—We thank Elisabetta Ullu and Chris Tschudi and members of their laboratory for influencing our work in this direction and for trypanosome-related reagents and instruction. We acknowledge Chris Tschudi for suggesting that we search for an expressed sequence tag for the T. brucei fibrillarin. We thank Leo Otake and Christine Smith in Joan Steitz's laboratory for advice about RNA sequencing and cloning of snoRNA 5'- and 3'-ends. We are grateful to Ed Hurt and David Tollervey for providing the yeast strains bearing mutant fibrillarin alleles. We thank Sean Eddy for his support of the snoRNA analysis. We thank Sean Eddy, Elisabetta Ullu, and François Dragon for critical reading of the manuscript.

REFERENCES

1. Eichler, D. C., and Craig, N. (1995) Prog. Nucleic Acids Res. Mol. Biol. 49, 179–239
2. Venema, J., and Tollervey, D. (1995) Yeast 11, 1629–1650
3. White, T. C., Rudenko, G., and Borst, P. (1986) Nucleic Acids Res. 14, 5789–5801
4. Spencer, D. F., Collings, J. C., Schnare, M. N., and Gray, M. W. (1987) EMBO J. 6, 1063–1071
5. Campbell, D. A., Kubo, K., Clark, C. G., and Bothevroyd, J. C. (1987) J. Mol. Biol. 196, 113–124
6. Schnare, M. N., Cook, J. R., and Gray, M. W. (1990) J. Mol. Biol. 215, 85–91
7. Schnare, M. N., and Gray, M. W. (1990) J. Mol. Biol. 215, 85–91
8. Gray, M. W., Boer, P. H., Collings, J. C., Heinonen, T. Y. K., and Schnare, M. N. (1989) Highlights of Modern Biochemistry (Skoda, J., Paces, V., and Kostka, V., eds) pp. 521–530, VSP International Science Publishers, Ziest, The Netherlands
9. Clark, C. G. (1987) J. Mol. Evol. 25, 343–350
10. Maden, B., and Hughes, J. (1987) Chromosoma (Berl.) 105, 391–400
11. Maden, B. E. H. (1997) Nature 389, 129–131
Fibrillarin-associated Box C/D Small Nucleolar RNAs in *Trypanosoma brucei*:
SEQUENCE CONSERVATION AND IMPLICATIONS FOR 2′-O-RIBOSE
METHYLA TION OF rRNA

David A. Dunbar, Steven Wormsley, Todd M. Lowe and Susan J. Baserga

*J. Biol. Chem. 2000, 275:14767-14776.*
doi: 10.1074/jbc.M001180200 originally published online March 9, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M001180200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 74 references, 28 of which can be accessed free at
http://www.jbc.org/content/275/19/14767.full.html#ref-list-1