Characterization of a Short Isoform of Human Tgs1 Hypermethylase Associating with Small Nucleolar Ribonucleoprotein Core Proteins and Produced by Limited Proteolytic Processing*

Received for publication, May 22, 2007, and in revised form, November 16, 2007 Published, JBC Papers in Press, November 26, 2007

Cyrille Girard, Céline Verheggen, Henry Neel, Anne Cammas, Stephan Vagner, Johann Soret, Edouard Bertrand, and Rémy Bordonné

From the Institut de Génétique Moléculaire de Montpellier, UMR 5535, IFR 122, CNRS and Université de Montpellier, 1919 Route de Mende, 34293 Montpellier Cedex 5, France and INSERM US63, Université Toulouse III Paul Sabatier, 31052 Toulouse, France

Tgs1 is the hypermethylase responsible for m3G cap formation of U small nuclear RNAs (U snRNAs) and small nucleolar RNAs (snoRNAs). In vertebrates, hypermethylation of snRNAs occurs in the cytoplasm, whereas this process takes place in the nucleus for snoRNAs. Accordingly, the hypermethylase is found in both compartments with a diffuse localization in the cytoplasm and a concentration in Cajal bodies in the nucleoplasm. In this study, we report that the Tgs1 hypermethylase exists as two species, a full-length cytoplasmic isoform and a shorter nuclear isoform of 65–70 kDa. The short isoform exhibits methyltransferase activity and associates with components of box C/D and H/ACA snoRNPs, pointing to a role of this isoform in hypermethylation of snoRNAs. We also show that production of the short Tgs1 isoform is inhibited by MG132, suggesting that it results from proteasomal limited processing of the full-length Tgs1 protein. Together, our results suggest that proteasome maturation constitutes a mechanism regulating Tgs1 function by generating Tgs1 species with different substrate specificities, subcellular localizations, and functions.

Small ribonucleoproteins (RNPs) are complexes required for processing RNA precursors into mature RNA species. Based on their intracellular location and function, these RNPs can be classified in two groups, the nucleoplasmic small nuclear RNPs (snRNPs) that play a role in the maturation of pre-mRNAs and the small nucleolar RNPs (snoRNPs) that reside in the cell nucleolus and are required for maturation of pre-rRNA (1). Each spliceosomal snRNP is composed of an snRNA (U1, U2, U5) complexed with a set of common (or core) proteins, also called Sm proteins (B–B’ in mammals, D1, D2, D3, E, F, and G) (2). In contrast, snoRNAs associate with distinct sets of evolutionarily conserved proteins: box C/D snoRNAs are bound by fibrillarin, Nop56p, Nop58p, and 15.5p, whereas box H/ACA snoRNAs associate with dyskerin (Cb5 in yeast), Nhp2p, Gar1p, and Nop10p (3). Formation of a complete core complex is required for accumulation/stabilization of snoRNAs as well as for correct nucleolar localization of snoRNPs (3).

The snRNP and snoRNP particles differ not only in their composition but also in their biogenesis, since production of mature spliceosomal snRNPs requires a cytoplasmic phase, whereas formation of snoRNPs is exclusively a nuclear process (4, 5). Despite these differences, snRNPs and snoRNPs share common features. For example, yeast snRNA and snoRNA precursors are stabilized by the La proteins, and they utilize common processing factors, such as the RNase III orthologues Rnt1 and the exosome for 3’ end maturation (6–11). Moreover, assembly of vertebrate snRNPs and snoRNPs depends on the SMN protein, the product of the survival of motor neuron gene (12). Finally, both in yeast and in mammals, snRNAs and a number of snoRNAs are substrates for the Tgs1 hypermethylase involved in the conversion of the m’G cap to a 2,2,7-trimethylguanosine structure (13–15).

The fact that a unique enzyme is responsible for m3G cap formation of snRNAs and snoRNAs, which are associated with different sets of common proteins, raises the question of how Tgs1 recognizes these various RNPs. By analogy with the snRNPs, where Tgs1 associates with basic domains found in the C-terminal extensions of Sm proteins, it has been proposed that the enzyme also interacts with basic rich domains of the common proteins of the snoRNPs (13). Consistent with this, common snoRNP proteins Nop58p, Nop56p, Cbf5p, and Gar1p possess highly charged domains with numerous K and R residues (16), and it has been shown that the C-terminal basic rich extensions of Nop58p and Cbf5p strongly interact with Tgs1 in yeast (13).

We previously reported that, in HeLa cells, the human Tgs1 hypermethylase locates diffusely in the cytoplasm (17), in agreement with former data showing that hypermethylation of vertebrate snRNAs is a cytoplasmic event (2). It also locates in the nucleus, where it concentrates in Cajal bodies (17, 18). Given that snoRNAs transit through Cajal bodies during their biogenesis, it is possible that hypermethylation of snoRNAs...
occurs in these structures, but this has never been directly shown. Moreover, the manner via which nuclear Tgs1 discriminates between snRNAs and snoRNAs remains to be determined, given that both RNA species transit through Cajal bodies during their biogenesis (19). In the process of investigating the function of Tgs1 in vivo using an RNA interference approach, we discovered that two isoforms of Tgs1 exist in HeLa cells and that these isoforms are differentially distributed; the full-length isoform with an apparent molecular mass of 110 kDa is located in the cytoplasm, whereas the shorter 65–70-kDa isoform is located in the nuclear compartment. We also show that the short isoform lacks the N-terminal region of Tgs1, retains methyltransferase activity, and binds preferentially to common proteins of C/D and H/ACA snoRNPs, consistent with a role in snRNA hypermethylation. Finally, we report that inhibitors of the proteasome block production of the small Tgs1 isoform, suggesting a role for the proteosomal machinery in its maturation.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructions**—The plasmid expressing the Tgs1-TAP fusion protein was constructed using the pZOMNE-N vector placed into the Gateway system according to the manufacturer’s instructions (Invitrogen). The bicistronic plasmid pCRTgs1 was constructed by tripartite ligation of a Tgs1-encoding fragment (nucleotides 1–903) generated by PCR amplification and cloned at corresponding restriction sites into the Gateway system. The pBST-dyskerin construct carrying the Nhp2 (20) under the T7 promoter were constructed using the Gateway system according to the manufacturer’s instructions. The pBST-dyskerin construct carrying the Nhp2 under the T7 promoter were constructed using the Gateway system according to the manufacturer’s instructions. The pBST-dyskerin construct carrying the Nhp2 under the T7 promoter were constructed using the Gateway system according to the manufacturer’s instructions.

Plasmids carrying genes (FibNter, Nop56Cter, 15.5, and Nhp2) under the T7 promoter were constructed using the Gateway system. The pBST-dyskerin construct carrying the dyskerin gene in the pKS(−) background was a gift from B. Jady (Toulouse).

To obtain the pGST-Tgs1Nter plasmid (encoding amino acids 1–475 of Tgs1), a BamHI-HindIII fragment (blunt-ended by Klenow treatment at the HindIII site) was isolated and cloned into the pGST-4T3 vector (Amersham Biosciences) digested with BamHI and SmaI. To obtain the pGST-Tgs1Cter construct (encoding from amino acid 477 to the end of Tgs1), a BamH1-HindIII fragment (blunt-ended by Klenow treatment at the HindIII site) was isolated and cloned at corresponding restriction sites into the pBluescript vector (Stratagene).

Plasmids carrying genes (FibNter, Nop56Cter, 15.5, and Nhp2) under the T7 promoter were constructed using the Gateway system. The pBST-dyskerin construct carrying the dyskerin gene in the pKS(−) background was a gift from B. Jady (Toulouse).

**Production of Anti-Tgs1 Antibodies**—The Tgs1 polyclonal antibodies were generated in rabbits (Eurogentec, Belgium) by injecting two peptides corresponding to amino acids 656–670 and 839–852 of the Tgs1 sequence (Fig. 3C). Tgs1-specific antibodies were affinity-purified using glutathione S-transferase (Amersham Biosciences) beads cross-linked to a GST-Tg1 fusion protein carrying the C-terminal region (577–end) of Tgs1 as described (21). Briefly, 5 ml of rabbit serum adjusted to pH 8.2 were applied to the column overnight at 4 °C. After washing the column, antibodies were eluted with 0.2 M glycine buffer (pH 2.5). The protein concentration was determined at 280 nm using the BCA protein assay kit (Pierce). Fractions containing the antibodies were dialyzed against PBS containing 50% glycerol and stored at −20 °C.

**Cell Culture, Transfection, RNA Interference, and Luciferase Assays**—HeLa cells were grown at 37 °C in 5% CO2 and in Dulbecco’s modified Eagle’s medium containing 10% fetal calf serum. Cells were transfected by the calcium phosphate coprecipitation procedure as described previously (18). The precipitates were left overnight on the cells and removed by solubilization in an isotonic solution without calcium and phosphate, and cells were then split.

Target-specific siRNA duplexes were designed as described by the manufacturer (Eurogentec), and selected siRNA sequences corresponding to region 788–808 of the Tgs1 mRNA sequence were generated. Control siRNAs were generated against the sequence of the firefly luciferase. Cells were transfected with the siRNA duplexes using Oligofectamine reagent (Invitrogen) according to the manufacturer’s protocol.

For luciferase assays, HeLa cells were seeded on 12-well plates at a density of 150,000 cells/well. The next day, plasmids (1 μg/well) were transfected using FuGene 6 (Roche Applied Science) following the supplier’s instructions. Cells were incubated at 37 °C for 24 h before analysis. The luciferase activities were measured using the Dual Glo Luciferase kit (Promega) and a microplate lumionometer (EG and G, Berthold).

**Extract Preparation, Subcellular Fractionation, and Western Blot Analysis**—For preparation of whole cell extracts, HeLa cells were harvested in HNTG buffer (20 mM HEPES, pH 7.9, 150 mM NaCl, 1% Triton X-100, 10% glycerol, 1 mM MgCl2, 1 mM EGTA, 1 mM phenylmethylsulfonyl fluoride, protease inhibitor mixture) and lysed on ice for 20 min. After centrifugation at 15,000 × g (10 min, 4 °C), supernatants were carefully removed and used for Western analysis or immunoprecipitation experiments.

For proteasome inhibition, cells were treated with 20 μM MG132 (10.5 mM stock in Me2SO) (Biomol International) diluted in Dulbecco’s modified Eagle’s medium. Culture medium was then supplemented with 25 μM MG132 for 10–12 h for a total incubation time of 18–20 h.

For cell fractionation, HeLa cells were harvested by trypsinization and washed twice with ice-cold PBS. Cells were resuspended at about 5 × 109/ml in HMKE buffer (20 mM Heps (pH 7.2), 10 mM KCl, 5 mM MgCl2, 1 mM EDTA, 250 mM sucrose) containing protease inhibitor mixture (complete EDTA-free; Roche Applied Science), phenylmethylsulfonyl fluoride (1 mM), and digitonin (200 μg/ml). Cells were left on ice for 10 min and then centrifuged at 500 × g 10 min at 4 °C to separate cytosol from membranes, nuclei, organelles, and cytoskeleton. The supernatant (cytosol) was carefully removed, and the pellet was washed in HMKE buffer without digitonin. To perform nuclear protein extraction, the pellet was solubilized in extraction buffer (0.1 M Tris-HCl, pH 9, 0.1 M NaCl, 5 mM KCl, 1 mM CaCl2, 0.5 mM MgCl2, 0.5% Nonidet P-40, and protease inhibitor mixture) for 20 min on ice. Samples were
Characterization of a Short Tgs1 Isoform

then centrifuged at 15,000 \( \times g \) for 10 min at 4 °C, and the supernatant was carefully removed (nuclear fraction).

For Western blot analysis, the protein content of the fractions was determined using the BCA protein assay kit (Pierce) and equal amounts of protein from each lysate were analyzed. Samples were boiled in SDS-sample buffer and analyzed on an SDS-8% polyacrylamide gel. Proteins were blotted on a Protran nitrocellulose membrane and incubated with purified anti-Tgs1 antibodies (diluted 1:400), followed by anti-rabbit secondary antibody. Detection was carried out by enhanced chemiluminescence (Pierce). Anti-TATA-binding protein antibodies reacting against the nuclear TATA-binding protein were used at a 1:5000 dilution and were purchased from Eurogentec. Anti-PAP antibodies (P1291; Sigma) were used at a 1:2000 dilution.

Production of Recombinant Proteins and Methyltransferase and in Vitro Binding Assays—Plasmids encoding pGST fusion proteins were transformed in Escherichia coli and induced by the addition of 0.5 mM isopropyl 1-thio-\( \beta \)-d-galactopyranoside, and recombinant proteins were purified as described (22). A methyltransferase assay was performed as described (14) in 20-\( \mu l \) reaction mixtures containing 50 mM Tris-HCl (pH 8.0), 5 mM dithiothreitol, 3 \( \mu M \) [\( ^{3}H \)-CH\(_{3}\)]AdoMet, 5 mM m\(^{3}\)GTP, and 2 \( \mu g \) GST-sTgs1 fusion protein. GST (2 \( \mu g \)) was used as a negative control. After incubation for 60 min at 37 °C, the reaction mixture was spotted onto DEAE-cellulose filters, and unincorporated [\( ^{3}H \)-CH\(_{3}\)]AdoMet was removed by washing the filters four times with 20 \( \mu l \) ammonium bicarbonate and twice with ethanol/ether (1:1). The filters were then dried and counted in scintillation fluid.

In vitro transcription–translation reactions with bacterial S30 or reticulocyte lysates were performed in the presence of \([{}^{35}S]\)methionine according to the manufacturer’s instructions (Promega). For in vitro protein/protein interaction studies, 1 \( \mu g \) of purified GST fusion proteins (or GST alone as a negative control) were incubated with 10 \( \mu l \) of glutathione-agarose and 20 \( \mu l \) of in vitro translated \([{}^{35}S]\)methionine-labeled proteins in 200 \( \mu l \) of final binding buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM MgCl\(_{2}\), 0.1% Nonidet P-40). After a 30-min incubation at 25 °C followed by a 30-min incubation at 4 °C with constant rotation, beads were pelleted and washed five times with 1 \( ml \) of binding buffer. Bound proteins were fractionated on SDS-polyacrylamide, and the gel was then dried and analyzed by fluorography.

Reverse Transcription–PCR, Immunoprecipitation Experiments, and Northern Blot Analysis—Total RNA was extracted from HeLa cells using TRI-Reagent (Sigma), and cDNA was synthesized with the First Strand cDNA kit (Amersham Biosciences) using the Tgs1-specific reverse primer covering the 3’-stop codon (5’-CCTCGCAGCTGCTGATAGTTAG-3’). PCR amplification was performed using the forward primer (5’-ATGTGCTGCGAGAAGTGGAGC3’) and the reverse primer. PCR products were separated by agarose gel electrophoresis and visualized by ethidium bromide staining.

For the immunoprecipitations, antibodies were coupled to 40 \( \mu l \) of protein A-Sepharose (0.1 g/ml; Amersham Biosciences) in 500 \( \mu l \) of buffer HNTG (20 mM HEPES, pH 7.9, 150 mM NaCl, 1% Triton X-100, 10% glycerol, 1 mM MgCl\(_{2}\), 1 mM EGTA, 1 mM phenylmethylsulfonyl fluoride, protease inhibitor mixture) for 1 h at 4 °C. The beads were washed four times with 1 ml of buffer, added to 500 \( \mu g \) to 1 mg of extract, adjusted up to 750 \( \mu l \) with HNTG buffer, and incubated for 2 h at 4 °C with constant stirring. The beads were then washed four times with 1 ml of buffer and resuspended in loading buffer. Immunoprecipitation of the Tgs1-TAP fusion proteins was performed using 20 \( \mu l \) of IgG-agarose beads (Sigma). For Northern blot analysis, total RNA was separated by agarose-formaldehyde gel electrophoresis and analyzed following standard protocols (23).

Immunocytochemistry and Immunofluorescence—Immunofluorescence was performed on cells grown on coverslips, washed twice in PBS, and fixed in 4% (w/v) paraformaldehyde in PBS at room temperature, followed by permeabilization with 0.1% Triton X-100 in PBS for 5 min at room temperature. Antibodies against fibrillarin were described previously (18). 800 colin was detected with a polyclonal rabbit anti-coilin (1:400 dilution). Anti-Tgs1 antibodies were used at a 1:200 dilution. Incubations were for 1 h at 25 °C in PBS containing 2% bovine serum albumin, and slides were then washed three times for 3 min in PBS. Secondary antibodies were diluted according to the supplier’s instructions. Coverslips were mounted on glass slides using mounting medium (Vectashield), and samples were observed using a Leica fluorescence microscope. Images were acquired with a Coolsnap camera (Photometrics) controlled by the Metamorph software (Universal Imaging).

RESULTS

Two Tgs1 Isoforms Are Expressed in HeLa Cells—We previously reported that polyclonal rabbit antibodies raised against a mixture of two peptides located in the C-terminal domain (amino acids 656–670 and 839–852) of the human Tgs1 hypermethylase allow the detection of the endogenous Tgs1 protein by immunofluorescence and Western blot analysis (17, 18). As shown in Fig. 1A, the full-length Tgs1 protein (fTgs1) migrates approximately as a 110-kDa polypeptide rather than at 96 kDa as predicted by its sequence, and this might be due to post-translational modifications or to its amino acid composition. For reasons of clarity, we will use this 110 kDa size to refer to the full-length Tgs1 polypeptide throughout. In addition to the fTgs1 band, we also detected additional faster migrating bands with estimated sizes of ~65–70 and 55 kDa. We previously assumed that these latter bands corresponded to nonspecific binders of the anti-Tgs1 antibodies. However, during our studies on Tgs1 depletion by RNA interference, a closer examination of the origin of these bands by Western analyses of total extracts revealed that the levels of fTgs1 (110 kDa) and sTgs1 (65–70 kDa) proteins are concomitantly decreased by a unique siRNA duplex covering region 788–808 of the Tgs1 sequence (Fig. 1B). In contrast, the level of the lower unknown bands around 55 kDa was not affected using this RNA interference duplex, suggesting that these bands result from cross-reactivity of the Tgs1 antibodies with other human proteins. In this regard, one of the immunizing peptides covers amino acids 656–670 located in the Tgs1 region encompassing conserved methyltransferase motifs. The unknown proteins at 55 kDa might therefore correspond to nonspecific binders containing methyltransferase motifs. That these proteins are unrelated to Tgs1 is also consistent with the fact that they were not gener-
...onation when a Tgs1-TAP fusion was expressed in HeLa cells (Fig. 3A), and they were not immunoprecipitated with anti-fibrillarin antibodies (Fig. 5A) (see below).

We next investigated the subcellular localization of the Tgs1 isoforms, since it was previously shown that the Tgs1 hypermethylase localizes both in the cytoplasm and in the nucleus in Cajal bodies (17, 18). As shown in Fig. 1C, fractionation of HeLa cell extracts into nuclear and cytoplasmic compartments indicates that the full-length Tgs1 protein is found mainly in the cytoplasm, whereas the short sTgs1 isoform is located in the nucleus. In order to test whether Tgs1 isoforms are detected when lysates are made under denaturing conditions, we prepared extracts by direct lysis of cells in Laemmli-SDS buffer. As shown in Fig. 1D, the same pattern of bands was observed when the blot was probed with anti-Tgs1 antibodies, suggesting that the short isoform is not produced by proteolytic degradation during the preparation of the extracts.

**Characterization of a Short Tgs1 Isoform**

We next investigated the mechanisms responsible for the production of the sTgs1 protein and tested whether this isoform could be generated through alternative splicing of the Tgs1 pre-mRNA. This hypothesis was first tested by performing Northern blot analysis of total RNA isolated from different tissues. As shown in Fig. 2A, only one transcript of ~3.2 kb was detected in the different tissues, and no RNA corresponding to an alternatively spliced version capable to encode the short Tgs1 isoform could be detected. A unique transcript was also detected by Northern analysis on total RNA prepared from HeLa cells (Fig. 2B).

The existence of alternative Tgs1 transcripts was also tested by reverse transcription-PCR analysis of poly(A)⁺ RNA from HeLa cells. As shown in Fig. 2C, this approach repeatedly amplified a PCR product with an expected size of ~2.5 kb, corresponding to the full-length isoform when primers covering the 5' -ATG and the 3'-stop codon were used. These results suggest that the sTgs1 isoform is generated via a posttranslational mechanism rather than by alternative splicing.

**The sTgs1 Isoform Contains the C-terminal Region and Exhibits Methyltransferase Activity**

The fact that two isoforms of Tgs1 were detected with the same anti-Tgs1 antibodies raised against the C terminus of the protein suggested that the shorter one arises from deletion of the N-terminal region. This was tested using a Tgs1-TAP fusion construct carrying the tag at the C-terminal end of Tgs1. After transient expression of the Tgs1-TAP construct into HeLa cells, total extract was prepared and analyzed by Western blotting. As shown in Fig. 3A (left), two Tgs1-TAP isoforms with the expected sizes were detected by the anti-Tgs1 antibody in addition to the endogenous forms of Tgs1. The same polypeptides were also detected when the blot was incubated with PAP antibodies.
Characterization of a Short Tgs1 Isoform

FIGURE 3. The small Tgs1 isoform contains the C-terminal domain. A, constructs carrying the empty TAP vector or a Tgs1-TAP fusion were transiently transfected in HeLa cells. Whole cell lysates were prepared and subjected to Western blot analysis with anti-Tgs1 antibodies (left) or anti-PAP antibodies specifically recognizing protein A of the TAP tag (right). M, protein mass standards in kDa. B, lysates prepared from HeLa cells expressing the Tgs1-TAP fusion protein were purified on IgG-agarose beads followed by Western blot analysis with anti-Tgs1 antibodies. C, schematic representation of the Tgs1 isoforms. Internal ATGs in frame with the initiating ATG of the full-length Tgs1 protein are found at positions 94, 288, 301, and 484. The hatched box delineates the region containing the methyltransferase motifs. The two regions used for the production of Tgs1 polyclonal antibodies are indicated by α. Dotted lines indicate that the end of the short isoform has not been precisely determined.

The small Tgs1 isoform contains the C-terminal domain. A, constructs carrying the empty TAP vector or a Tgs1-TAP fusion were transiently transfected in HeLa cells. Whole cell lysates were prepared and subjected to Western blot analysis with anti-Tgs1 antibodies (left) or anti-PAP antibodies specifically recognizing protein A of the TAP tag (right). M, protein mass standards in kDa. B, lysates prepared from HeLa cells expressing the Tgs1-TAP fusion protein were purified on IgG-agarose beads followed by Western blot analysis with anti-Tgs1 antibodies. C, schematic representation of the Tgs1 isoforms. Internal ATGs in frame with the initiating ATG of the full-length Tgs1 protein are found at positions 94, 288, 301, and 484. The hatched box delineates the region containing the methyltransferase motifs. The two regions used for the production of Tgs1 polyclonal antibodies are indicated by α. Dotted lines indicate that the end of the short isoform has not been precisely determined.

FIGURE 4. The short Tgs1 isoform exhibits methyltransferase activity. A, a recombinant GST-sTgs1 was expressed in E. coli and purified over glutathione beads. Purified protein was resolved by SDS-PAGE and stained with Coomassie Blue (left). Protein mass standards (in kDa) are indicated at the left. The proteins were electrophoretically transferred to a nitrocellulose membrane and incubated with anti-Tgs1 antibodies (right). The double arrows point to the GST-sTgs1 protein band. B, 2 μg of GST-sTgs1 or 2 μg of GST were incubated for 1 h at 37 °C in 20 μl of reaction buffer containing 50 mM Tris-HCl (pH 8.0), 5 mM dithiothreitol, and 3 μM [3H-CH3]AdoMet as well as 5 mM GTP or 5 mM m7GTP as substrates. The reaction was then spotted on DEAE filters, and radioactivity was counted by scintillation.

recognizing only the protein A domain of the TAP tag, thus showing that they represent the Tgs1-TAP isoforms (Fig. 3A, right). As discussed above, no smaller forms of Tgs1-TAP were revealed, again demonstrating that the unknown bands of 55 kDa detected with anti-Tgs1 antibodies are not related to Tgs1 and correspond to cross-reacting bands.

To definitively establish that both flTgs1-TAP and sTgs1-TAP represent fusions carrying the TAP tag, extracts were prepared from HeLa cells carrying the pTgs1-TAP construct and incubated with IgG-agarose beads, and the pellets were analyzed by Western blotting using anti-Tgs1 antibodies. As shown in Fig. 3B, both flTgs1-TAP and sTgs1-TAP were recovered on IgG beads, confirming that they contain the TAP tag. A schematic representation of the full-length Tgs1 protein together with the short isoform is shown in Fig. 3C. Due to low recovery, we could not precisely map the N-terminal end of the short Tgs1 isoform, but based on the size obtained by Western analysis, we estimate that it begins between residues 320 to 400 of the Tgs1 protein.

In order to test whether the short Tgs1 isoform possesses methyltransferase activity, we constructed a vector encoding for a GST-sTgs1(380–end) fusion protein carrying the Tgs1 region from amino acid 380 to the end. The GST-sTgs1(380–end) construct was expressed in E. coli, and the fusion protein was purified on glutathione beads. As shown in Fig. 4A, when the purified protein is subjected to SDS-PAGE, a protein band with an apparent molecular mass around 87 kDa is obtained, which is in good agreement with the expected molecular mass of GST-sTgs1 (79 kDa). This protein band is confirmed to be the GST-sTgs1 fusion protein by Western blotting with anti-Tgs1 antibodies (Fig. 4A, right). The faster migrating bands do not efficiently stain with the antibody and most likely represent degradation products or premature translation termination polypeptides. To test whether the recombinant GST-sTgs1 protein has methyltransferase activity, we used a filter assay allowing us to measure methyl transfer from [3H-CH3]AdoMet to the m7GTP but not to GTP as reported for a recombinant Schizosaccharomyces pombe Tgs1 protein (14). As shown in Fig. 4B, methyl transfer to m7GTP but not to GTP was repeatedly observed with GST-sTgs1, whereas no activity was observed with the GST protein alone. These results clearly show that the human sTgs1 isoform retains catalytic activity.

The sTgs1 Isoform Preferentially Associates with both snoRNP C/D and H/ACA Core Proteins—The Tgs1 protein is required for hypermethylation of different classes of RNPs (snRNPs and box C/D and H/ACA snoRNPs) complexed with different common proteins. Based on our subcellular fractionation experiments showing that the short isoform is predominantly located in the nuclear compartment (Fig. 1C), we investigated whether it associates with fibrillarin, a protein common to the box C/D snoRNPs. This was first tested by immunoprecipitation of total cell lysates with fibrillarin antibodies. As shown in Fig. 5A, these experiments allowed the recovery of the small Tgs1 isoform, whereas only background levels of full-length Tgs1 protein were detected in the pellet.

Interactions of core snoRNP proteins with Tgs1 were also tested by in vitro pulldown assays using GST-Tgs1Nter and...
Characterization of a Short Tgs1 Isoform

GST-Tgs1Cter fusion proteins carrying the N-terminal domain (amino acids 1–475) of Tgs1 and the C-terminal domain (amino acids 477–end) of Tgs1, respectively. We previously showed that the SmB protein strongly binds to the N-terminal domain of Tgs1 but not to the C-terminal region (17), suggesting that this latter domain can mediate interactions with snoRNPs common proteins. As shown in Fig. 5, B and C, only background levels of labeled 35S-polypeptides were bound by the N-terminal domain or the GST protein alone, whereas the GST-Tgs1Cter fusion protein significantly associated with [35S]methionine-labeled basic rich domains of fibrillarin and Nop56 as well as with [35S]methionine-labeled dyskerin but not with labeled 15.5 and Nhp2 proteins. These results demonstrate that the C-terminal domain (residues 477-end) of Tgs1 is able to bind to snoRNPs core proteins.

To investigate further the possible association of the sTgs1 isoform with snoRNPs core proteins, HeLa nuclear extracts were prepared after transient expression of a plasmid carrying the GST-sTgs1(380–end) fusion and incubated with glutathione beads. Associated proteins as well as aliquots of the extracts were analyzed by SDS-PAGE and Western blotting. As shown in Fig. 5D, endogenous dyskerin and Nop56 proteins were recovered, whereas no proteins were immunoprecipitated from extract prepared after transfection of the GST vector alone. This experiment demonstrates that a GST-sTgs1(380–end) fusion protein corresponding to the short Tgs1 isoform associates with common snoRNPs components. We also examined whether snRNAs rather than snRNAs can be pelleted by the sTgs1 isoform. As shown in Fig. 5E, the short isoform co-precipitated U3 snRNA, but not U1 snRNA, in extracts prepared from cells transfected with a plasmid encoding the GST-sTgs1(380–end) fusion, confirming that sTgs1 preferentially associates with snoRNPs.

The Short Tgs1 Isoform Is Not Generated from an Internal Translation Initiation Site—Truncated isoforms of a given polypeptide can be generated from internal translation start sites. This can occur by mechanisms such as leaky ribosome scanning and ribosomal shunting or by the use of an internal ribosomal entry site (20, 24). The use of an alternative translation start site for production of the short Tgs1 isoform was first examined using in vitro transcription/translation experiments. Indeed, based on the size of the short Tgs1 isoform, several internal Met residues could serve as translation start sites, especially a methionine at position 301, which is surrounded by a perfect Kozak sequence (PuXATG). As shown in Fig. 6A (lane 1), in vitro transcription/translation using reticulocyte lysates and a vector carrying the full-length Tgs1 cDNA gave rise to two bands, the first corresponding to initiation at methionine 1 of the Tgs1 sequence and the second to initiation at methionine 94. However, no band corresponding to the short isoform was obtained by in vitro translation. This was further demonstrated by Western blot analysis, since no

FIGURE 5. The small Tgs1 isoform binds to snoRNP components. A, immunoprecipitation of fibrillarin with sTgs1. HeLa cell extracts were mixed with anti-fibrillarin antibodies fixed on protein A-Sepharose and treated as described under "Experimental Procedures." The pellet was separated on SDS-PAGE, and immunoblotting was performed using anti-Tgs1 antibodies. Input (50 μg of total extract) represents 10% of the amount used in the immunoprecipitation experiments; control (Ctrl) corresponds to extracts mixed with protein A-Sepharose beads alone, and pellet represents the fraction bound to the resin. B, in vitro translated 35S-labeled fibrillarin- Nter, Nop56Cter, and 15.5 proteins were incubated with the indicated GST-recombinant fusion proteins or GST alone. Bound proteins were analyzed by SDS-PAGE and autoradiography. Input was as in A. C, binding assay of recombinant GST fusion proteins carrying the Nter or Cter domains of Tgs1. The recombinant fusion proteins were incubated with in vitro translated 35S-labeled dyskerin and Nhp2 proteins. Input and bound proteins were treated as described in C. D, ex vivo binding assays. A plasmid encoding the GST fusion protein carrying the region of Tgs1 encompassing amino acids 380–end as well as a plasmid encoding GST-Nter, Nop56Cter, and 15.5 proteins were incubated with the indicated GST-recombinant fusion proteins or GST alone. Bound proteins were analyzed by SDS-PAGE and autoradiography. Input was as in A. Ctrl corresponds to extracts mixed with protein A-Sepharose beads. E, Northern blot analysis. Extract prepared from HeLa cells expressing GST-sTgs1(380–end) was incubated with glutathione beads, RNA was recovered, and the pellet was analyzed by Northern blot hybridization. Probed RNA species are shown at the left. Input represents 10% of the lysate used in the experiment. Ctrl (Ctrl) corresponds to lysate incubated with CL6B-Sepharose beads.
Characterization of a Short Tgs1 Isoform

FIGURE 6. The production of the short Tgs1 isoform does not require an internal translation initiation site. A, in vitro coupled transcription/translation was performed with a TNT reticulocyte lysate using a construct carrying the full-length Tgs1 cDNA under the T7 promoter. The produced polypeptides were then separated on SDS-PAGE, transferred to nitrocellulose, and autoradiographed (lane 1). After exposure, the blot was probed with anti-Tgs1 antibodies to visualize exogenous Tgs1 isoforms (lane 2). A HeLa whole cell extract was analyzed similarly and used as control (lane 3). B, analysis of IRES activity in the Tgs1 sequence. The pCRHL construct contains a stable hairpin in the intercistronic region separating the LucR and LucF genes (negative control). Construct pCRMyc carries the IRES region found in Myc (positive control), whereas pCRTgs1 carries the N-terminal region of Tgs1 (nucleotides 1–903). The plasmids were transfected into HeLa cells, and luciferase activities were measured. The corresponding values represent the ratio between the activities of the two luciferase enzymes obtained with each construct.

FIGURE 7. Proteasome inhibitors block production of the small Tgs1 isoform. A, HeLa cells were cultured in the absence (−) or presence (+) of MG132 for 18 h, and whole cell extracts were prepared and separated by SDS-PAGE, transferred to nitrocellulose membrane, and hybridized with anti-Tgs1 antibodies (top) or anti-GAPDH antibodies (bottom) to demonstrate similar loading. B, HeLa cells were treated for 18 h as in A, except that the proteasome inhibitor epoxomycine was used during cell culture. C, Western blot analysis of Tgs1 in HeLa cells treated with MG132 for the indicated times. Cell lysates were resolved by SDS-PAGE, and proteins were transferred to a nitrocellulose membrane and hybridized with anti-Tgs1 antibodies. Untreated cells are labeled ctrl. D, HeLa cells were transfected with a construct carrying the Tgs1-TAP fusion in the absence (−) or presence (+) of MG132. Whole cell extracts were prepared, and Western blot analysis was performed using an anti-PAP antibody.

A prominent band was revealed upon incubation with anti-Tgs1 antibodies (Fig. 6A, lane 2).

We next tested whether synthesis of the short Tgs1 isoform might occur through an internal ribosome entry site (IRES), which has also been shown to modulate the relative expression of protein isoforms (25, 26). To this end, the sequence (residues 1–903) coding for the N-terminal Tgs1 region was cloned into a bicistronic vector expressing two luciferases, LucR and LucF, encoded by the same mRNA under the control of the cytomegalovirus promoter (25), and the plasmid obtained was named pCRTgs1. The pCRHL plasmid, used as negative control, carries a hairpin between both luciferases, whereas the pCRMyc plasmid was used as positive control and contains the Myc IRES (27). The plasmids were transfected into HeLa cells, and the ratio between the activities of the two luciferase enzymes observed in the cell extracts was calculated and compared with the data obtained with the negative and positive controls (Fig. 6B). Although the Myc IRES (pCRMyc) gave rise to a relative activity of 25 and the negative control to an activity of 1.7, expression of the pCRTgs1 construct gave rise to an activity of 5, indicating that the N-terminal region of Tgs1 does not exhibit efficient IRES activity. Altogether, these results show that use of an internal translation start site does not account for the production of the small Tgs1 isoform.

Proteasome Inhibitors Block Production of the Short Tgs1 Isoform—We next tested whether the short Tgs1 isoform is generated by proteolytic cleavage. Given that protein substrates are often targeted for site-specific cleavage or degradation by the proteasome (28), we investigated the effects of proteasome inhibitors on production of the Tgs1 isoforms. To this end, HeLa cells were treated with the proteasome inhibitors MG132 or epoxomycine, and extracts were analyzed by immunoblotting. As shown in Fig. 7, the sTgs1 isoform was no longer observed upon treatment with MG132 (A) or epoxomycine (B), whereas the full-length Tgs1 isoform was still detected. A time course analysis (Fig. 7C) showed that formation of the short isoform was already inhibited after 10 h of exposure to MG132, and the level continues to decrease up to 20 h, indicating that the disappearance of the short form is not an indirect consequence of the treatment with the drug.

In order to confirm the results obtained for the endogenous Tgs1 isoforms, we tested whether treatment with proteasomal inhibitors also affects the production of the small Tgs1-TAP isoform. To this end, HeLa cells were transfected with the Tgs1-TAP construct and treated with MG132. Immunoblotting of extracts with antibodies against the TAP tag revealed that the small Tgs1-TAP fusion protein was not produced upon MG132 treatment, whereas the two Tgs1 isoforms were detected in nontreated cells (Fig. 7D). These results are consistent with an involvement of the proteasome in the production of the small Tgs1 isoform.

The Short Tgs1 Isoform Accumulates in Cajal Bodies—Subcellular fractionation experiments reported in this study demonstrate that the short Tgs1 isoform locates in the nuclear compartment. Moreover, our immunochemistry analyses using Tgs1 antibodies previously showed that the hypermethylase localizes in the cytoplasm and nuclear Cajal bodies (17, 18). In order to test whether Cajal body staining can be observed with the sTgs1 isoform, HeLa cells were transfected with a plasmid encoding the GFP-sTgs1(380–end) fusion protein. As shown in Fig. 8A (bottom), the GFP-sTgs1(380–end) fusion protein
and testis, whereas only a 55-kDa protein was ubiquitously found in other tissues. The same study also reported that the cytoplasmic fraction derived from HeLa cells contains a 55-kDa isoform, whereas the larger 90-kDa form is nuclear (29). Although it is likely that these isoforms correspond to those that we report here, there is a discordance regarding the subcellular localization; we find that the short isoform is nuclear and the full-length is cytoplasmic. Whether these discrepancies are due to differences in the protocols used for fractionation of the cellular compartments remains to be determined. Nevertheless, the localizations of the full-length and short Tgs1 isoforms that we report are corroborated by the subcellular localizations detected using the GFP-Tgs1 and GFP-sTgs1(380–end) fusion proteins (Fig. 8A). This is also consistent with the nuclear localization of DTLd, the homolog of Tgs1 in Drosophila (30). Indeed, the DTLd protein, which is essential for development, is produced from a 2.6-kb transcript generating a 60-kDa protein that is located in the nucleus and in nuclear spots when expressed as a GFP fusion in HeLa cells. Interestingly, Western blot analysis of Drosophila S2 cell extracts using anti-DTLd antibodies allows the visualization of an additional high molecular mass 110-kDa polypeptide in addition to the 60-kDa isoform (30). The relationship between these proteins was not further analyzed, and as such, additional studies are required to test whether they are functionally equivalent to the isoforms that we detected in HeLa cells.

The short Tgs1 isoform encompasses all of the conserved methyltransferase motifs, and we could show that it possesses catalytic activity (Fig. 4B) by using a methyltransferase assay (14). Comparisons of Tgs1 family homologies indicate that hypermethylases from lower eukaryotes are mainly composed of the conserved MTase domain, whereas proteins from higher eukaryotes have long N-terminal domains with no obvious similarities to other known proteins (32). Concerning the former class, it has been shown that Tgs1 proteins from Saccharomyces cerevisiae and Trypanosoma brucei are essential for m4G cap formation of both snRNAs and snoRNAs (13, 33), indicating that the absence of N-terminal regions does not restrict the hypermethylation capabilities of the enzymes. Biochemical studies also showed that S. pombe Tgs1 is a bona fide guanine-specific methyltransferase and that the chemical steps of m4G synthesis do not require input from RNA or protein cofactors in vitro (14).

The nuclear localization of the sTgs1 isoform is in agreement with our experiments showing that it associates with core proteins of box C/D and H/ACA snRNPs (Fig. 5). Our results suggest a role for the short isoform in the biogenesis of snoRNAs biogenesis and specifically in the snoRNA cap hypermethylation process. This correlates with previous studies showing that snoRNPs do not transit through the cytoplasm during their synthesis (12, 31). Our finding that a GFP fusion protein carrying the N-terminal domain (amino acids 1–475) of human Tgs1 localizes to the cytoplasm (data not shown) is also consistent with a role for the full-length protein in hypermethylation of the cap structure of snRNAs, a cytoplasmic process (2, 4, 34). It is also consistent with the fact that human Tgs1 associates with SmB via its N-terminal domain (17). Accordingly, stepwise reconstitution of the U1 snRNP core particle indicates that the
Characterization of a Short Tgs1 Isoform

core U1 RNP is hypermethylated in vitro by a HeLa cytosolic extract, whereas a subcore U1 RNP lacking the SmB/B'-SmD3 proteins is not (35). Altogether, these observations strongly suggest that the N-terminal regions of Tgs1 proteins from higher eukaryotes are involved in protein/protein interactions required for subcellular localization of the enzyme and/or for association with specific RNP components.

Interestingly, our study reveals that the short isoform containing the C-terminal domain probably originates from the full-length protein by limited proteasomal processing. The use of such a mechanism was unexpected, since the proteasomal system appears to be involved in complete destruction of proteins rather than maturation of precursor proteins. However, recent studies indicate that the proteasome also plays nonproteolytic functions in multiple aspects of gene expression, such as recruitment of coactivators to promoters, initiation of transcription, and elongation (36, 37). To our knowledge, only the p50 and p52 subunits of NF-κB are generated by ATP-dependent processing of p105 and p100 in vivo and in vitro (38–40). During this process, the precursor proteins are cleaved, the N-terminal domain yielding the mature subunit while the C-terminal domain is degraded. In the case of Tgs1, it appears that the short isoform is produced from processing of the full-length protein, with the C-terminal domain yielding the sTgs1 isoform while the N-terminal domain is degraded. The structural determinants required for sTgs1 formation remain unknown, and due to its low abundance, we were unable to precisely map its amino-terminal residue. Further studies will be required to characterize the precise mechanism used for Tgs1 processing and to determine which motifs are involved in the production of the short Tgs1 isoform.

Our subcellular localization studies show that the short Tgs1 isoform is found in Cajal bodies and that fibrillarin distribution becomes altered. Its accumulation in Cajal bodies following proteasome inhibition (Fig. 8A) is consistent with a recent study reporting that MG132 treatment modulates the relative distributions of various nucleolar proteins, such as the snoRNPs core proteins fibrillarin and Nop58 (41). It also alters their nuclear distribution, since Nop58 becomes localized in Cajal bodies (41). In addition to its effect on the distribution of early processing factors, proteasome inhibition also results in a decrease in rRNA gene transcription as well as rRNA production. Thus, the proteasomal system has been shown to be involved in multiple steps of ribosome biogenesis (41).

Whether limited processing of Tgs1 occurs in the nucleus or in the cytoplasm remains unknown. Proteasomes are found in both compartments, and it has been shown that they diffuse rapidly within the cytoplasm and the nucleus, they can be transported unidirectionally over the nuclear membrane, and they can also enter the nucleus upon reassembly during cell division (42). Our findings that fibrillarin associates with the sTgs1 isoform and that proteasome inhibition results in a colocalization of Tgs1 and fibrillarin in Cajal bodies (Fig. 8A) suggest that maturation occurs in the nuclear compartment, directly in Cajal bodies. In this model, the full-length Tgs1 protein should be imported in the nucleus and locate in Cajal bodies in order to be processed. Accordingly, we have shown that a subset of the GFP-Tgs1 fusion protein localizes into the nuclear compartment in Cajal bodies (Fig. 8A) (17), a result consistent with our subcellular fractionation experiment, indicating that the full-length Tgs1 is predominantly but not exclusively found in the cytoplasm (Fig. 1C). It is noteworthy that one can envisage that the presence of the GFP at the N terminus of Tgs1 decreases the rate of its processing, allowing thereby the detection of the fusion protein in Cajal bodies. The proposed model also agrees with the fact that Tgs1 is still detected in these structures after proteasome inhibition when the production of the sTgs1 isoform is blocked (Fig. 8B). In conclusion, proteasome maturation could represent a mechanism regulating Tgs1 function by allowing production of Tgs1 species with different substrate specificities, subcellular localizations, and functions.

Acknowledgments—We thank L. Linares, I. Jariel, O. Coux, and M. Pichaczky for discussions, and we are grateful to N. Taylor for critically reading the manuscript. We also thank M. Caizergues-Ferrer and B. Jady for gifts of plasmids and antibodies as well as F. Maloum, O. Zosanit, and Y. Campion for help in pulldown assays and protein purification.

REFERENCES

1. Bertrand, E., and Bordonne, R. (2004) Prog. Mol. Subcell. Biol. 35, 79–97
2. Will, C. L., and Lührmann, R. (2001) Curr. Opin. Cell Biol. 13, 290–301
3. Kiss, T. (2004) J. Cell Sci. 117, 5949–5951
4. Mattaj, I. W. (1986) Cell 66, 905–911
5. Terns, M. P., and Dahlberg, J. E. (1994) Science 264, 959–961
6. Chanfreau, G., LeGrain, P., and Jacquier, A. (1998) J. Mol. Biol. 284, 975–988
7. Kufel, J., Allmang, C., Chanfreau, G., Petfalski, E., Lafontaine, D. L., and Tollervey, D. (2000) Mol. Cell. Biol. 20, 5415–5424
8. Marais, R. J., and Intine, R. V. (2002) Gene Expr. 10, 41–57
9. Seielt, R. L., Zheng, B., Asura, A., and Rymond, B. C. (1999) Nucleic Acids Res. 27, 587–595
10. Wolin, S. L., and Cedervall, T. (2002) Annu. Rev. Biochem. 71, 375–403
11. Xue, D., Rubinsson, D. A., Pannone, B. K., Yoo, C. J., and Wolin, S. L. (2000) EMBO J. 19, 1650–1660
12. Terns, M. P., and Terns, R. M. (2001) Curr. Biol. 11, R862–R864
13. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002) Mol. Cell. 9, 891–901
14. Hausmann, S., and Shuman, S. (2005) J. Biol. Chem. 280, 4021–4024
15. Lemm, I., Girard, C., Kuhn, A. N., Watkins, N. I., Schneider, M., Bordonne, R., and Lührmann, R. (2006) Mol. Biol. Cell 17, 3221–3231
16. Gautier, T., Berges, T., Tollervey, D., and Hurt, E. (1997) Mol. Cell. Biol. 17, 7078–7089
17. Mouaikel, J., Narayanam, U., Verheggen, C., Matera, A. G., Bertrand, E., Tazi, J., and Bordonne, R. (2003) EMBO Rep. 4, 616–622
18. Verheggen, C., Lafontaine, D. L., Samarsky, D., Mouaikel, J., Blanchard, J. M., Bordonne, R., and Bertrand, E. (2002) EMBO J. 21, 2736–2745
19. Cioce, M., and Lamond, A. I. (2005) Annu. Rev. Cell Dev. Biol. 21, 105–131
20. Vagner, S., Galy, B., and Pyronnet, S. (2001) EMBO Rep. 2, 893–898
21. Bharti, A. K., Olson, M. O., Kufe, D. W., and Rubin, E. H. (1996) J. Biol. Chem. 271, 1993–1997
22. Smith, D. B., and Johnson, K. S. (1988) Gene (Amst.) 67, 31–40
23. Ausubel, F. M., Brent, R., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1997) Current Protocols in Molecular Biology, pp. 491–498, John Wiley & Sons, Inc., New York
24. Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A. C., and Vagner, S. (2003) Biol. Cell 95, 169–178
25. Bonnal, S., Schaeffer, C., Creancier, L., Clamens, S., Moine, H., Prats, A. C., and Vagner, S. (2003) J. Biol. Chem. 278, 39330–39336
26. Bonnal, S., Pileur, F., Orsini, C., Parker, F., Pujol, F., Prats, A. C., and Vagner, S. (2005) J. Biol. Chem. 280, 4144–4153
27. Nanbru, C., Lafon, L., Audigier, S., Gensac, M. C., Vagner, S., Huez, G., and
Characterization of a Short Tgs1 Isoform

JANUARY 25, 2008•VOLUME 283•NUMBER 4•JOURNAL OF BIOLOGICAL CHEMISTRY 2069