Nucleic Acids Research, Vol. 20, No. 13 3443–3451

An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution

Véronique Ségault, Annie Mougin, Anne Grégoire, Josette Banroques[1] and Christiane Branlant*
Laboratoire d'Enzymologie et de Génie Génétique, Université de Nancy I, URA CNRS 457, BP 239, 54506 Vandoeuvre-Les-Nancy Cedex and [1]Laboratoire de Génétique Moléculaire, Ecole Normale Supérieure, 46 rue d'Ulm, 75230 Paris Cedex 05, France

Received February 25, 1992; Revised and Accepted May 28, 1992

ABSTRACT

The conformation of *Saccharomyces cerevisiae* U3 snRNA (snR17A RNA) in solution was studied using enzymatic and chemical probes. *In vitro* synthesized and authentic snR17A RNAs have a similar conformation in solution. The *S.cerevisiae* U3 snRNA is folded in two distinct domains. The 5'-domain has a low degree of compactness; it is constituted of two stem-loop structures separated by a single-stranded segment, which has recently been proposed to base-pair with the 5'-ETS of pre-ribosomal RNA. We demonstrate that, as previously proposed, the 5'-terminal region of U3 snRNA has a different structure in higher and lower eukaryotes and that this may be related to pre-rRNA 5'-ETS evolution. The *S.cerevisiae* U3 snRNA 3'-domain has a cruciform secondary structure and a compact conformation resulting from an higher order structure involving the single-stranded segments at the center of the cross and the bottom parts of helices. Compared to tRNA, where long range interactions take place between terminal loops, this represents another kind of tertiary folding of RNA molecules that will deserve further investigation, especially since the implicated single-strands have highly evolutionarily conserved primary structures that are involved in snRNP protein binding.

INTRODUCTION

Nuclei from all eukaryotic species contain a series of abundant and metabolically stable U snRNAs (1, 2). Whereas the implication of nucleoplasmic U1, U2, U4, U5 and U6 snRNAs in pre-messenger splicing has been experimentally demonstrated several years ago (3 for review), the expected role of the nucleolar U3 snRNA in pre-ribosomal RNA maturation has only been demonstrated recently (4).

In mammalian cells, rRNA is transcribed as a 47S precursor (5), which is subsequently cleaved in multiple steps to yield mature 18S, 5.8S and 28S rRNA species. The processing sites are not restricted to the borders of the mature rRNAs; primary cleavages occur upstream or downstream of these sites (6 for review). In particular, a first very rapid event consists in the cleavage of a 5'-terminal fragment, whose length is comprised between 400 and 800 nucleotides depending on the mammalian species considered (7–9).

Based on the observation that it was associated with the 28S-35S nucleolar RNA from deproteinized mammalian nucleolar extracts (1, 10) and with pre-ribosomal particles >60S from cell extracts (11), U3 snRNA was suggested to function in some aspects of the processing of the ribosomal RNA precursor and to base-pair with the pre-RNA. Nevertheless, no clear identification of the postulated base-pairing could be obtained, until recently (12–20).

The primary rRNA processing reaction has been successfully reproduced in extracts of mouse cultured cells (7). Abolition of the processing event when the cellular extract was depleted of U3 snRNA, by oligonucleotide-directed RNase H digestion, demonstrated an implication of U3 snRNA in the primary event (4). This was in agreement with the results of *in vivo* crosslinking between U3 snRNA and pre-rRNA by a psoralen derivative. In human (15) and rat (16) cell cultures, and in yeast cells (20), a crosslink between U3 snRNA and the pre-rRNA was observed in the vicinity of the first pre-rRNA cleavage site. This was also confirmed by a disruption of the U3 snRNA gene in the yeast *Saccharomyces cerevisiae* (21).

A clear understanding of U3 snRNA function in the primary pre-rRNA maturation events required information on the U3 snRNA structure and on its evolution. The first model for the U3 snRNA secondary structure was proposed on the basis of common potential helical structures of rat and *Dictyostelium discoïdeum* U3 snRNAs (22, 23). An improved model was then proposed on the basis of an experimental study on the human U3 snRNA secondary structure that was performed on both naked RNA and RNA in a crude cell extract (24). In this model (24), the 64 nucleotide 5'-terminal segment is folded into a single helical structure (helix 1), whereas the 3'-domain of the molecule (nt 75 to 217) is folded into a Y-shaped structure containing 3 helical structures (24). The biological significance of this Y-shaped structure is strongly supported by its phylogenetic conservation (19, 25, 26). In contrast, the 5'-terminal region of lower eukaryote and plant U3 snRNAs was proposed to have a secondary structure different from that of vertebrate U3

* To whom correspondence should be addressed

snRNAs: two short helical regions instead of a single, long one (26, 27). This statement was based on a prediction given by computer RNA-folding program and no experimental proof was provided. In addition, the *Saccharomyces cerevisiae* U3 snRNA sequence used for the comparison was wrong. The sequence used was that of the gene, and we showed later, that in this yeast species the U3 snRNA coding sequence is interrupted by an intron (28). Thus, the 14-nucleotide sequence previously considered to be at the 5'-end of the U3 snRNA is in fact an intronic sequence (28). The U3 snRNA 5'-domain may play a key role in pre-ribosomal RNA interaction. Indeed, for the *in vivo* U3 snRNA/pre-rRNA crosslinks obtained for various species, the U3 snRNA crosslinked residues were located in the 5'-domain adjacent to Box A (16, 20). It was thus of high importance, to get a definitive answer as to the possible evolutionary divergence of the U3 snRNA 5'-terminal region.

Another question to be solved concerning the U3 snRNA secondary structure was the possibility of additional base-pairing interactions between the two internal single-stranded regions of the Y-shaped structure. Such interactions were previously mentioned, but they were discarded because their positions vary between species (24). Nevertheless, a strong base-pairing is possible in the case of tomato U3 snRNA (27). An experimental study of these putative additional interactions was also of interest, since they concern two of the U3 snRNA segments, Box B and C, whose primary structure has been highly conserved throughout evolution (22). Furthermore, Box C was recently shown to be essential for fibrillarin binding to U3 snRNA (29). In the nucleolus, U3 snRNA does not act as a naked RNA but as an snRNP particle, which contains at least 6 proteins (24). One of them, the 34 kD fibrillarin (30), is common to other nucleolar snRNPs (U8, U13, U14, X, Y) (31). The three Boxes B, C and D were found to be protected against ribonuclease action in the U3 snRNP (24) and they are probably all interacting with proteins.

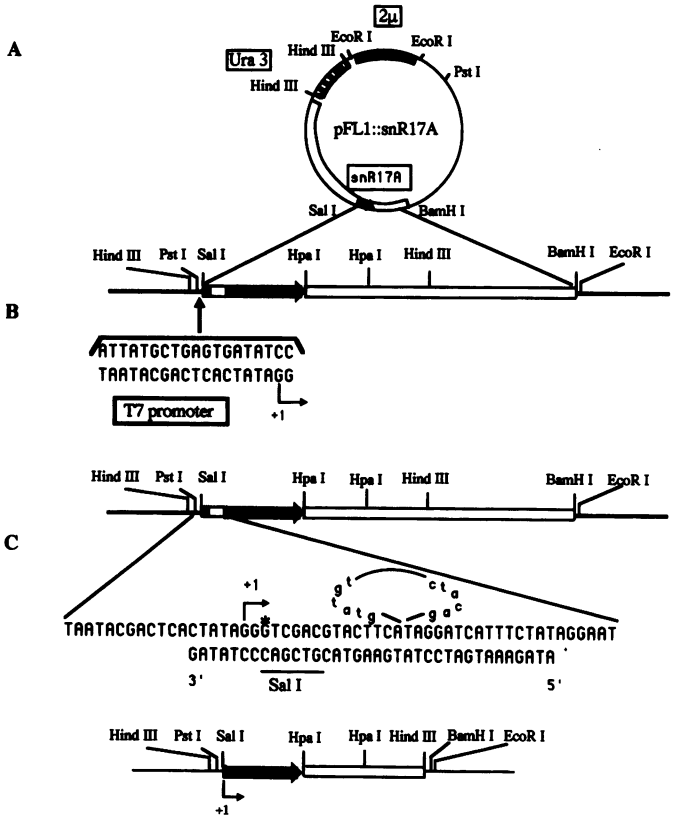


Figure 1. Construction of an snR17A RNA gene without intron under control of a promoter for T7 RNA polymerase. (**A**) The SalI–BamHI fragment from plasmid pFL1::snR17A (28) containing the snR17A RNA gene, was inserted in bacteriophage M13mp9 cleaved by SalI and BamHI nucleases. The dark thick line represents snR17A RNA coding region, the dashed thick line the intron, and the thin line M13mp9 sequences. (**B**) A 19 nucleotide sequence corresponding to a T7 RNA polymerase recognition site (35) was inserted upstream the SalI site by directed mutagenesis. (**C**) The intronic sequence was deleted by site-directed mutagenesis using the oligonucleotide represented at the bottom of the figure. The position of the 5'-nucleotide in the authentic snR17A RNA is indicated by a star.

In this paper, we describe an experimental study of the *S.cerevisiae* U3 snRNA secondary structure. In addition to the above mentioned motivations for getting additional information on the U3 snRNA secondary structure, this study was of interest for two reasons: i) *S.cerevisiae* U3 snRNA contains an additional 71 nucleotide sequence, relative to all other U3 snRNAs sequenced (25), that is inserted between the 3' and the 5' domains of the other U3 snRNAs. Two alternative structures were proposed for this additional segment (25) and an experimental study was necessary to choose among them. ii) It was important to get information on the structure of mature *S.cerevisiae* U3 snRNA in order to compare it with that of the intron-containing U3 pre-snRNA.

In order to achieve this study, the intron was deleted from the U3 gene (snR17A gene) that was placed under the control of a promoter for T7 bacteriophage RNA polymerase. By using this genetic construction, snR17A RNA was produced *in vitro*. Its secondary structure was studied by using enzymatic and chemical probes, and by the reverse transcriptase method for analysis. As a control, the same methods were applied to the authentic snR17A RNA from a cell extract. We concluded that two short helical structures exist at the 5'-end of *S.cerevisiae* snR17A RNA and that, in naked snR17A RNA, a higher order structure exists in Box B and C. The results concerning the 5'-domain are discussed taking into account recent results of Beltrame and Tollervey (20).

MATERIALS AND METHODS

Preparation of *S.cerevisiae* small RNA mixture

S.cerevisiae cells (strain ATCC 28383) were grown in standard YPD medium [1% (w/v) yeast extract, 1% (w/v) bactopeptone and 2% (w/v) D glucose], for 10 hours at 30°C. They were ground with glass beads in the presence of phenol and 10 mM Tris–HCl pH 7.6. Total RNA was fractionated on a 15–30% sucrose gradient made up in 100 mM NaCl, 10 mM Tris–HCl, pH 7.6 buffer. Centrifugation was for 24 hours, in a SW28 rotor, at 27000 rpm and 4°C, RNA with a sedimentation coefficient between 4S and 8S was ethanol precipitated, and redissolved in water.

Construction of an intron-depleted snR17A gene

The SalI site which delimites the coding region from its upstream non-coding sequence (28) was used for the construction (Figure 1A). The SalI–BamHI fragment from plasmid pFL1::snR17A (28), which contained the coding region for RNA snR17A, was inserted in a SalI–BamHI cleaved M13mp9 bacteriophage. Two successive site-directed mutageneses were performed: i) the first one inserted a T7 promoter sequence at the 5'-end of the snR17A coding region (Figure 1B); ii) the second one deleted the intronic

sequence (Figure 1C). The HindIII fragment containing the chimeric snR17A gene obtained was transferred to a pUC18 plasmid denoted pVS1::snR17A. The mutageneses were performed using the commercial kit of Amersham. Deoxyoligonucleotides were synthesized on a Pharmacia automated DNA synthesizer and purified on a polyacrylamide gel.

In vitro transcription of RNA snR17A

We again took advantage of the presence of a HpaI restriction site (located at the 3'-end of the RNA snR17A coding region) in plasmid pVS1::snR17A (28). Transcription reactions were carried out in 250 μl volumes containing 5 μg of plasmid pVS1::snR17A linearized by HpaI nuclease, 75 nmol of each ribonucleoside triphosphate, 125 units RNase Guard™ (Pharmacia), 140 units T7 RNA polymerase (Amersham) in 10 mM NaCl, 10 mM $MgCl_2$, 10 mM DTT, 40 mM Tris−HCl, pH 8 buffer. After 2 hours of incubation at 37°C, nucleic acids were phenol extracted and ethanol precipitated. The template DNA was digested with 5 units of RNase-free DNase I (Boehringer) in 250 μl of 5 mM $MgCl_2$, 100 mM sodium acetate, pH 5 buffer, for 30 min at 37°C. After phenol extraction and ethanol precipitation, the RNA was dissolved in 120 μl of sterile water; 1 μl of this solution was used for each chemical reaction or enzymatic digestion.

Chemical modifications

Two chemical reagents were used: dimethylsulfate (DMS, Aldrich) and 1-cyclohexyl-3-(2-morpholino ethyl)-carbodiimide-metho-p-toluene sulfonate (CMCT, Merck). Modifications were performed under non-denaturing and semi-denaturing conditions,

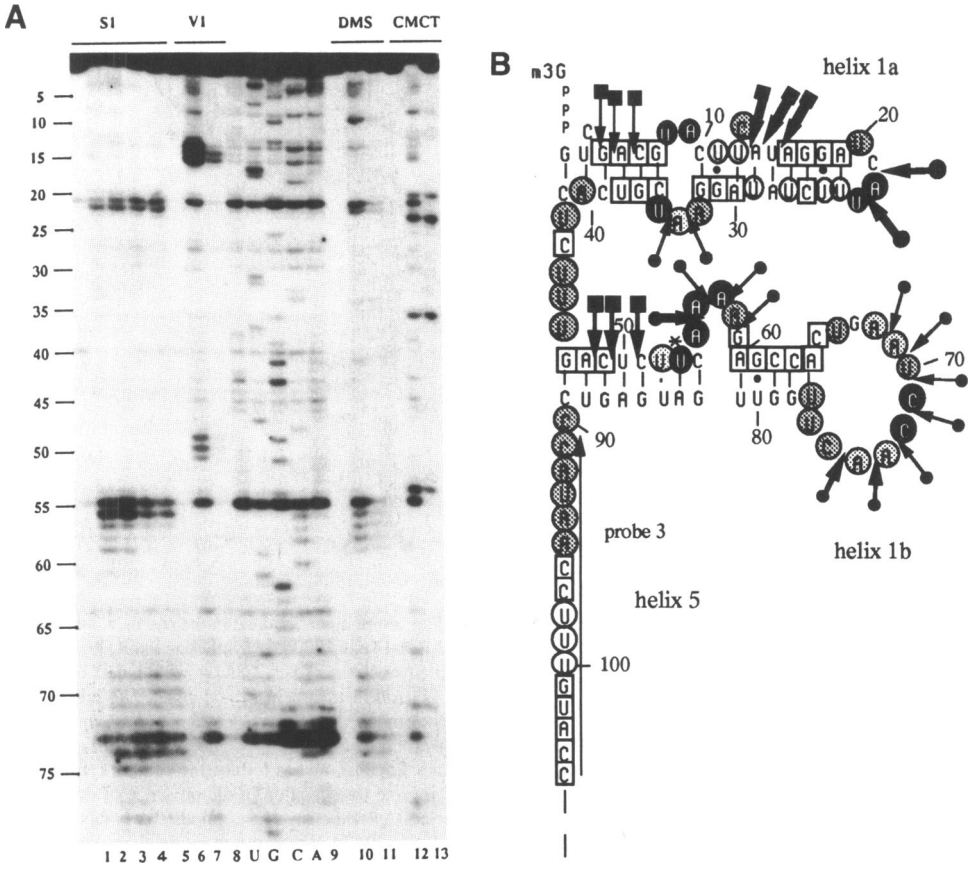


Figure 2. Structural analysis of the snR17A RNA 5'-domain. (**A**) An example of autoradiogram of a gel obtained in the chemical and enzymatic probing experiments in non-denaturing conditions. Both authentic (lanes 1, 4, 5, 6, 8, 11 and 13) and *in vitro* synthesized snR17A RNA (lanes 2, 3, 7, 10 and 12) were treated with chemical and enzymatic probes, as described in Materials and Methods. Lanes 2, 3, 4, 5: treatment with S1 nuclease (lanes 2 and 4 with 1 unit of enzyme per assay, lanes 3 and 5 with 2 units); lanes 7 and 8: treatment with V1 RNase (1.4 unit per assay), lanes 10 and 11: chemical treatment conditions with DMS (2 μl of pure DMS per assay), lanes 12 and 13 with CMCT (4.2 mg per assay). Positions of modifications and cleavages were identified using oligodeoxynucleotide probe 3 and reverse transcriptase. The synthesized cDNAs were fractionated by electrophoresis on a 6% polyacrylamide−urea gel, cDNAs obtained by reverse transcription of RNAs treated as for enzymatic digestion but without enzyme added were fractionated in parallel (lanes 1 and 6) as a control; cDNA obtained by reverse transcription in the presence of deoxy/dideoxynucleotide mixtures (lanes U, G, C, A) were used for sequence analysis and to identify the modified nucleotides and cleaved phosphodiester bonds. The elongation reaction products (lane 9) obtained in the absence of dideoxynucleotide was used as a control in order to identify reverse transcriptase stops on unmodified RNA. (**B**) Derived secondary structure model. The results of several series of chemical modifications in non-denaturing and semi-denaturing conditions and enzymatic digestions in non-denaturing conditions of the snR17A RNA are schematically represented on the proposed secondary structure. ●→: bond cleaved by S1 nuclease, ■→: bond cleaved by V1 RNase. The thickness of the arrows indicates the yield of cleavage. □: nucleotide modified by none of the chemical reagents tested in both semi-denaturing and non-denaturing conditions. ○: modified in semi-denaturing conditions, only. ◍: modified in both semi-denaturing and non-denaturing conditions. ●: strongly modified in non-denaturing and semi-denaturing conditions. The degree of modification of the nucleotide marked with an asterisk varied from one experiment to the other. Nucleotides which are neither circled nor squared were not accessible to analysis due to a parasite band in the gel or to their proximity with the probe. Nucleotides are numbered starting from the first nucleotide after the cap structure in authentic RNA (28). Accessibility of nucleotides 90 to 105 was determined using oligonucleotide 2 as a primer (Figure 3B).

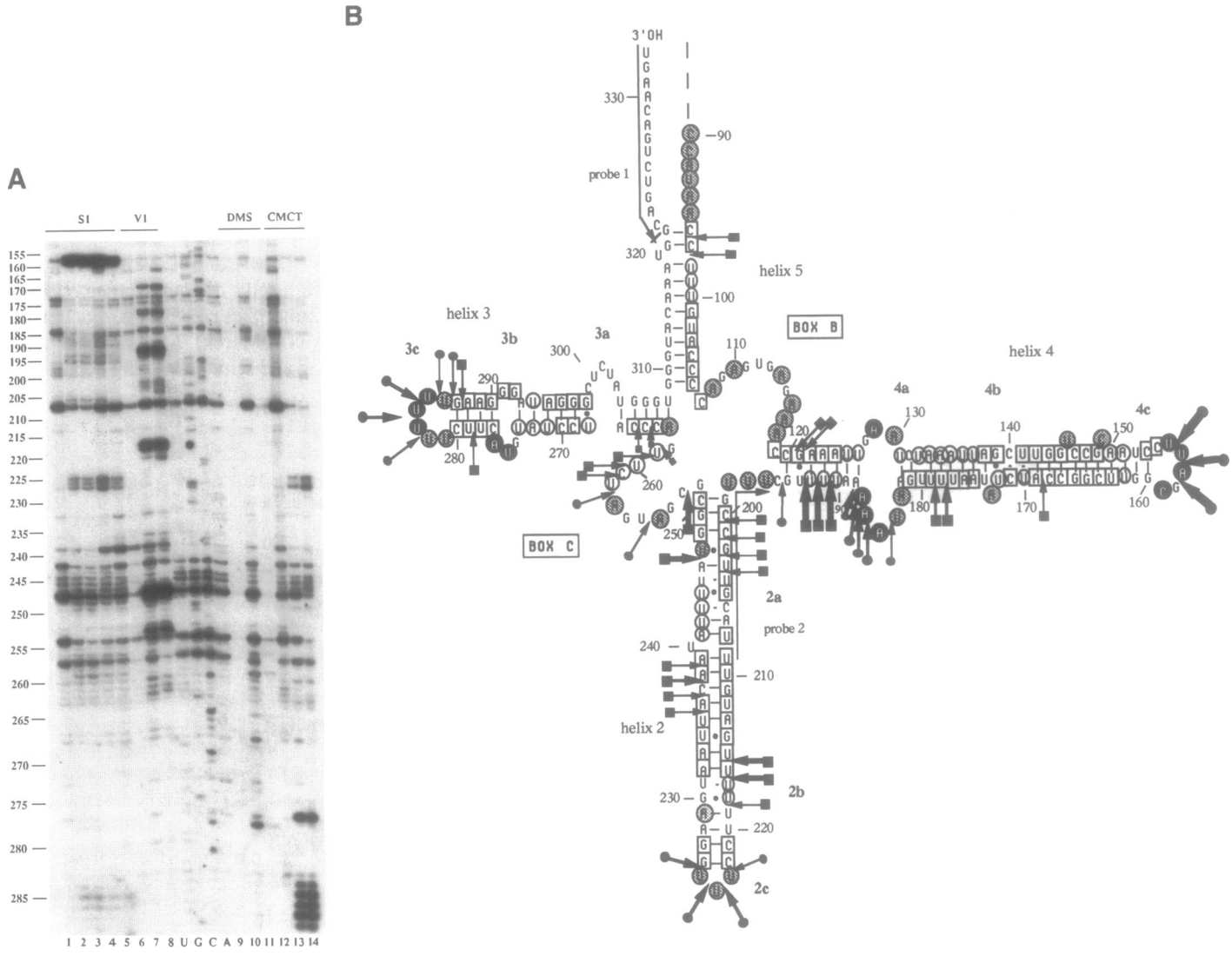


Figure 3. Structural analysis of the snR17A RNA 3′-domain. (**A**) An example of autoradiogram of a gel obtained in the chemical and enzymatic probing experiments in non-denaturing conditions. Here again, both authentic (lanes 1, 4, 5, 6, 8, 9, 11, 12 and 14) and *in vitro* synthesized snR17A RNA (lanes 2, 3, 7, 10 and 13) were treated with chemical and enzymatic probes. Lanes 2, 3, 4 and 5: treatment with S1 nuclease (lanes 2 and 4 with 1 unit of enzyme per assay, lanes 3 and 5 with 2 units per assay); lanes 7 and 8: treatment with V1 RNase. Lanes 10 and 11 : chemical treatment in non-denaturing conditions with DMS (2 μl of pure DMS per assay). Lanes 13 and 14: treatment with CMCT (4.2 mg per assay). Positions of modifications and cleavages were identified using oligodeoxynucleotide probes 1 and 2 and reverse transcriptase. The synthesized cDNAs were fractionated by electrophoresis on a 6% polyacrylamide−urea gel. cDNAs obtained by reverse transcription of RNAs treated as for enzymatic digestion but without enzyme addition (lanes 1 and 6), and as for chemical modifications but without chemical reagent addition (lanes 9 and 12), were fractionated in parallel as controls; cDNA obtained by reverse transcription in the presence of deoxy/dideoxynucleotides mixtures (lanes U, G, C, A) were used for sequence analysis and identification of the modified nucleotides and cleaved phosphodiester bonds. (**B**) Derived secondary structure model. The results of several series of chemical modifications and enzymatic digestions of the snR17A RNA are schematically represented on the proposed secondary structure, using the same rules for representation as in Figure 2B. Helix structures have been numbered as previously proposed (25). 2a, 2b, 3a, 3b, 4a and 4b designate putative internal loops of helical structures 2, 3 and 4, respectively, and 2c, 3c and 4c designate their terminal loops. Box B and Box C are the evolutionarily conserved nucleotide sequences according to (22, 25).

as described by Ehresmann *et al.* (32). For a given condition, experiments were repeated at least twice to obtain consistent data. The modifications were carried out either on 2.5 μg of the 4S−8S *S.cerevisiae* small RNA mixture or on *in vitro* synthesized snR17A RNA, in which 2.5 μg of a commercial yeast tRNA mixture (Boehringer) was added to get the same RNA/chemical reagent ratio as for the 4S−8S RNA mixture. Prior to the chemical reaction, the RNA was dissolved in 100 μl of the buffer used for chemical modifications, heated 5 min at 50°C and then slowly cooled to 20°C, which is the temperature used for all the modifications.

DMS modifications under non-denaturing conditions were performed in 50 mM KCl, 10 mM $MgCl_2$, 100 mM sodium cacodylate, pH 7.5 buffer and under semi-denaturing conditions in 2 mM EDTA, 100 mM sodium cacodylate pH 7.5 buffer. Two DMS concentrations were used, either 2 μl of pure DMS or 2 μl of a DMS solution 1/1 or 1/3 (v/v) (DMS/EtOH) per assay. The 2 μl were added to 100 μl of preheated RNA. Samples were incubated for 15 min, the reaction was immediately stopped by addition of 20 μl of 3 M NaOAc and ethanol precipitation. The ethanol precipitate was washed with 70% ethanol and dissolved in 12 μl of H_2O, 6 μl of this solution were used for a reverse transcriptase elongation assay.

CMCT modifications under non-denaturing conditions were

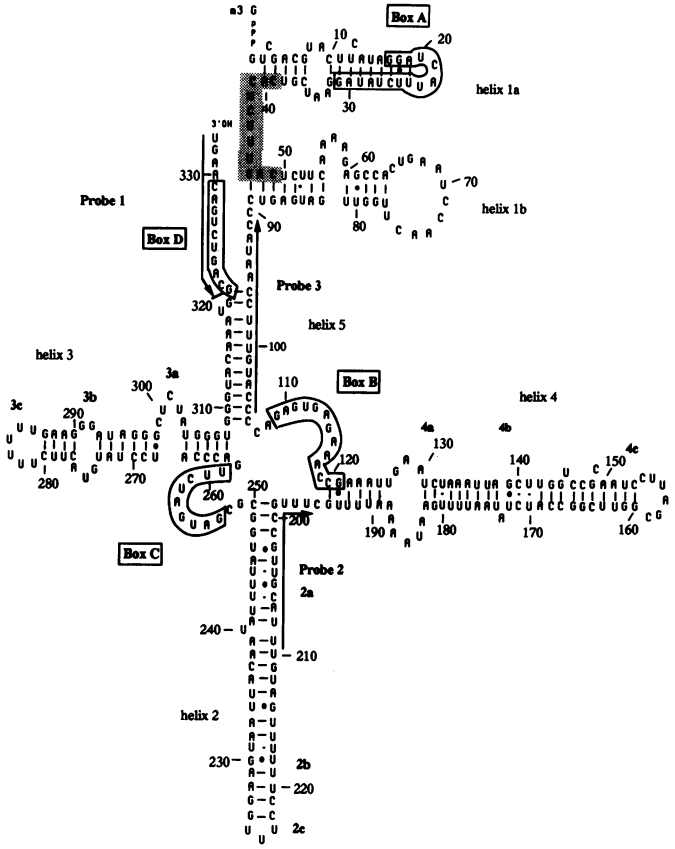


Figure 4. The secondary structure of snR17A RNA in solution. The evolutionarily conserved sequences, denoted Box A, B, C and D, are indicated, as well as the segment proposed to be base-paired with the 5-ETS of pre-rRNA *in vivo* (20) (sequence in the grey box).

performed in 50 mM KCl, 10 mM $MgCl_2$, 100 mM sodium borate, pH 8 buffer, and under semi-denaturing conditions, in the same buffer except that the 10 mM $MgCl_2$, 50 mM KCl were replaced by 2 mM EDTA. Various CMCT concentrations were used: 50 μl of a 21, 42 or 84 mg/ml solution of CMCT were added to 100 μl of preheated RNA. Incubation was for 40 min. At the end of the reaction, the same procedure was followed as for DMS modification.

Enzymatic cleavages

Two nucleases were used: S1 Nuclease (Pharmacia) and V1 RNase prepared from cobra *Naja oxiana* venom according to (33) (the venom was a generous gift of Dr S.Vassilenko). The reaction conditions were adapted from those previously described by Branlant *et al.* (34). The enzymatic digestions were also performed, either on 2.5 μg of *S.cerevisiae* 4S−8S RNA mixture or on *in vitro* synthesized snR17A RNA mixed with 2.5 μg of commercial tRNA.

V1 RNase digestions were carried out in 10 μl of 350 mM KCl, 10 mM $MgCl_2$, 10 mM Tris−HCl, pH 7.5 buffer, at 0°C. Before addition of the enzyme, the RNA was preincubated 10 min at 0°C in the reaction buffer. An enzyme/RNA ratio of 1.4 unit/μg was used. Incubations for 2 min and 7 min, were always performed in parallel. They were stopped by the addition of 5 μl of 100 mM EDTA, immediately followed by phenolic extraction of the RNA. The digested RNAs obtained from the 2 min and 7 min incubations were mixed, washed with 70% ethanol and dissolved in 12 μl of water and 6 μl of this solution was used for a reverse transcription elongation assay.

S1 digestions were performed in 10 μl of 50 mM KCl, 10 mM $MgCl_2$, 1 mM $ZnCl_2$, 25 mM Na acetate, pH 4.5 buffer, at 37°C. Prior to enzyme addition the RNA was preincubated 10 min at 37°C, in the reaction buffer. Two digestions were always performed in parallel: one using 1 unit of enzyme for 2.5 μg of RNA with a 12 min incubation, and the other one using 2 units of enzyme for 2.5 μg of RNA with a 2 min incubation. Both reactions were stopped by adding 2 μl of 100 mM EDTA, followed by phenolic extraction. The RNA issued from each digestion condition was analyzed by reverse transcription.

Reverse transcription

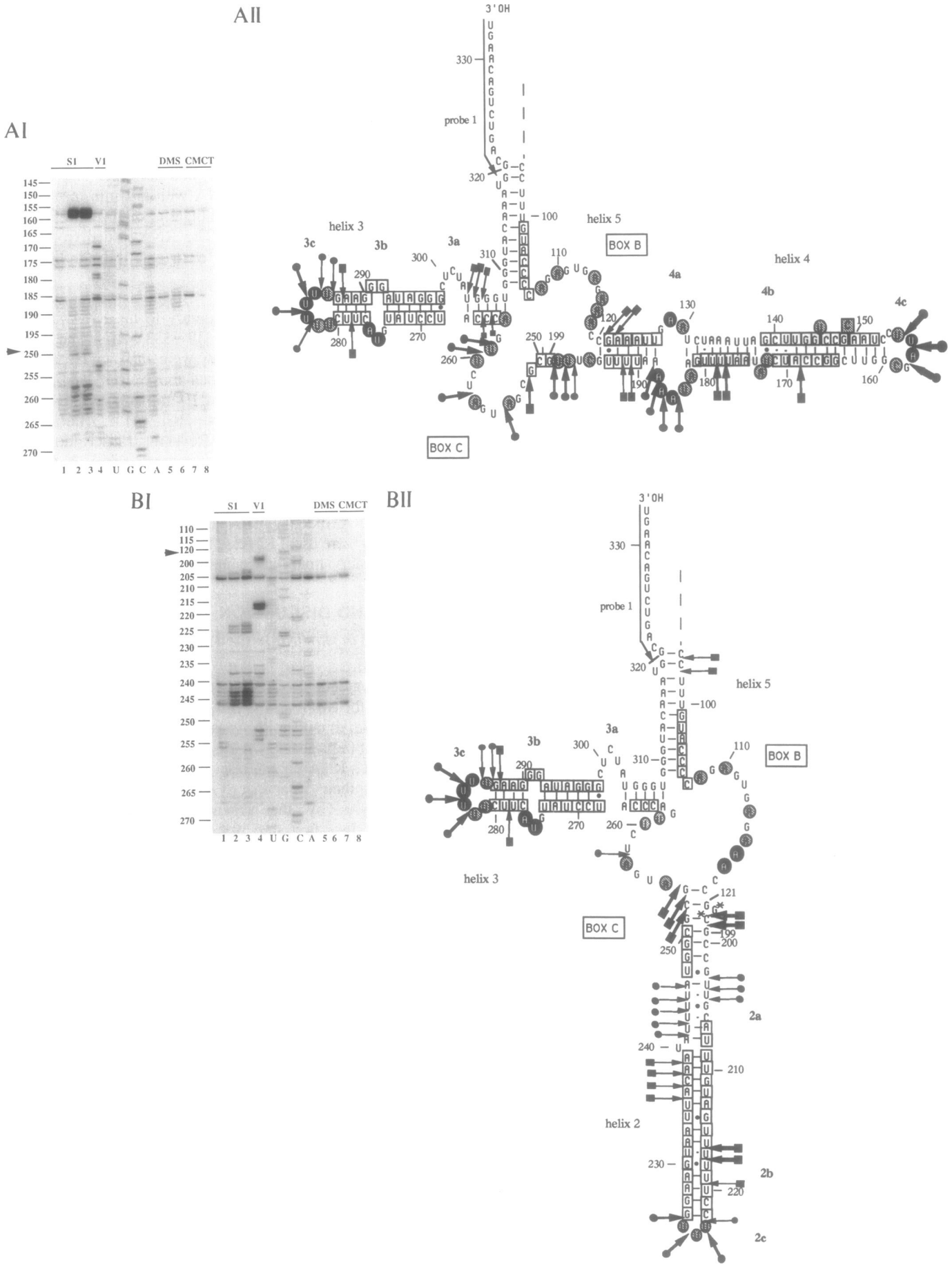
Oligonucleotide primers complementary to the RNA snR17A nucleotides 91−105, 196−210 and 319−333 were 5′-end labeled with [γ-^{32}P] ATP (Amersham). They were annealed (separately) to 6 μl of the RNA samples described above. RNA and labeled oligonucleotides were incubated for 10 min at 65°C in 40 mM KCl, 6 mM $MgCl_2$, 50 mM Tris−HCl, pH 8.3 buffer and returned to room temperature. Reverse transcription was performed with 1 unit of Rous sarcoma virus 2 reverse transcriptase (Amersham), for 30 min at 45°C, in the presence of 250 μM of each dNTP. To prepare sequencing ladders of unmodified RNA, dideoxynucleotide:deoxynucleotide mixtures (in a 1/5 ratio) were used. Reverse transcripts were fractionated on a 6% polyacrylamide gel containing 8 M urea.

RESULTS AND DISCUSSION

Authentic and *in vitro* synthesized snR17A RNAs have similar conformations

The use of synthetic RNA for secondary structure studies is advantageous for two reasons: i) large scale production of pure RNA is possible; ii) mutant RNAs can be obtained by site-directed mutagenesis of their genes. On the other hand, synthetic RNAs produced by T7 RNA polymerase, generally display additional nucleotides at their 5′ and/or 3′-ends and also differ from authentic RNAs by the absence of post-transcriptional modifications. In a first step, the structures of snR17A RNA produced *in vitro* and of authentic snR17A RNA were compared.

For this purpose, the intronic sequence of the snR17A gene was deleted by site-directed mutagenesis (Figure 1C). To minimize the presence of additional nucleotides at the 5′-end of the synthetic RNA, the T7 RNA polymerase promoter was introduced, immediately upstream the snR17A coding region, by site-directed mutagenesis (Figure 1B). The promoter sequence was chosen according to Chapman *et al.* (35). Due to the GGG sequence requirement at the T7 RNA polymerase initiation site (35), the presence of two additional guanosine residues at the 5′-end of the synthetic RNA could not be avoided. Obviously, the synthetic RNA was missing the 5′-cap structure that the authentic RNA contains. Cleavage of the template DNA with HpaI nuclease prior to RNA polymerization limited the additional residues at the RNA 3′-end to a single extra U residue (28). To check for possible differences between synthetic and authentic snR17A RNAs, a mixture of yeast *S.cerevisiae* 4S−8S RNAs containing snRN17A RNA was prepared. Chemical modifications and enzymatic digestions were performed in parallel on both synthetic RNA and 4S−8S RNA mixture. The similar patterns of modifications and cleavages obtained for both RNAs (Figure


AII
AI
S1
V1
DMS
CMCT
BI
BII
3'OH
probe 1
helix 3
helix 4
helix 5
helix 2
BOX B
BOX C
3a
3b
3c
4a
4b
4c
2a
2b
2c

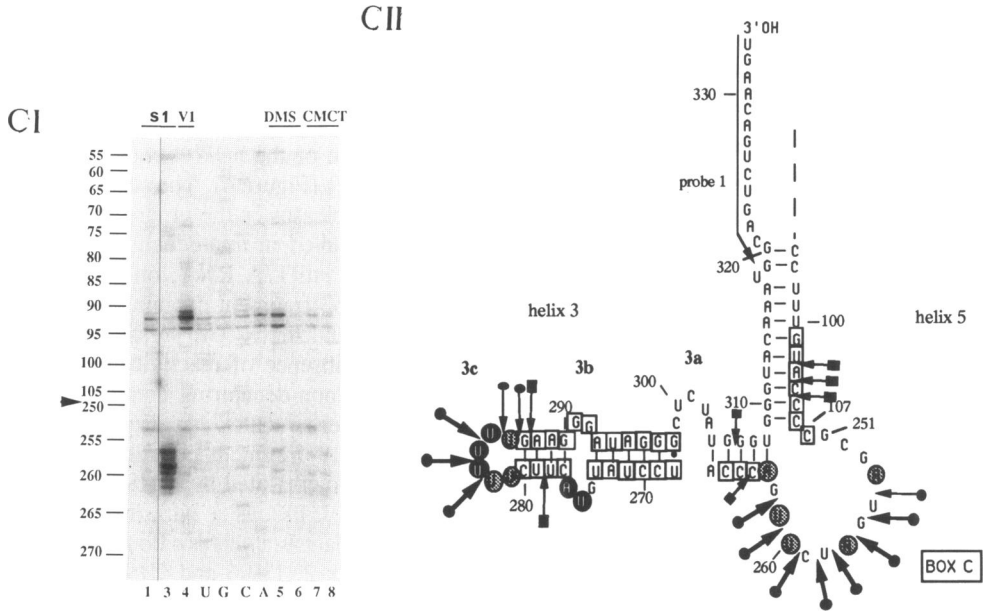


Figure 5. Chemical and enzymatic probing of the 3′-domain of mutant snR17A RNAs in non-denaturing conditions. For each mutant, an example of the electrophoretic fractionation of the reverse transcripts obtained by reverse transcription of chemically modified RNA and enzymatically digested RNA is given (I) as well as a schematic representation of the results obtained (II) using the same rules for representation as in Figure 2B, except that squared nucleotides are nucleotides found unreactive in non-denaturing conditions, as experiments have not been performed in semi-denaturing conditions; cDNAs obtained by reverse transcription of RNAs treated as for enzymatic digestion but without enzyme addition (lane 1), and as for chemical modifications but without chemical reagent addition (lanes 5 and 7), were fractionated in parallel, as controls. (**A**) Mutant Δ1 where the 200−249 fragment has been deleted. (**B**) Mutant Δ2 where the 122−198 fragment has been replaced by a GCp dinucleotide marked with asterisks. (**C**) Mutant Δ3 where the 108−250 fragment has been deleted. Chemical modifications and enzymatic digestions were performed as described in Figure 2A on the autoradiograms, the deletion sites are indicated by arrows (I).

2) proved that they have very similar conformations and this opened the possibility to use *in vitro* synthesized snR17A RNA for further studies.

snR17A RNA secondary structure analysis

We used two chemical probes: i) DMS modifies N7-G, N1-A and more slowly N7-A and N3-C, but only the N1-A and N3-C modifications were detected in our analysis conditions since N7-G and N7-A methylation do not stop reverse transcription (32), ii) CMCT reacts with N3-U and very slowly with N1-G, and both modifications arrest reverse transcription (32). For each reagent, modifications were performed in both non-denaturing and semi-denaturating conditions. Methylation of N1-A or N3-C by DMS, and modification of N3-U and N1-G by CMCT, are indicative of the absence of classical Watson−Crick base-pairing. Absence of reactivity may also be indicative of base stacking or of non-canonical interactions. Chemically modified nucleotides were identified by reverse transcription, using 3 oligonucleotide primers (Figures 2B and 3B). Representative examples of DMS and CMCT modification analyses are shown in Figures 2A and 3A. Since, no N1-G was found to be modified by CMCT in the conditions we used, additional information on single-stranded nucleotides was gained from a study of the S1 nuclease cleavage sites. V1 RNase was used to obtain information on Watson−Crick interactions and stacked nucleotides. Enzymatic cleavage sites were identified by reverse transcription using the same 3 oligonucleotide primers as for the chemical modification analyses. Representative examples of S1 and V1 RNase digestion analyses are shown in Figures 2A and 3A. Results obtained with chemical modifications and enzymatic digestions are schematically represented in Figures 2B and 3B. The complete structure deduced for snR17A RNA is represented in Figure 4.

The snR17A 3′-domain (nucleotides 95−333) has the cruciform structure previously proposed (25) (Figure 3B). Helical structures 3 and 5 are the counterparts of two of the helical structures of the vertebrate Y-shaped domain. It is difficult to know which of the two helices 2 and 4 corresponds to the third one. Our experiments allowed to define which of the two previously proposed structures for helix 4 (25) is present in the RNA. According to our previous (28) and present results, a guanine residue at position 194 was omitted in the previously published snR17A sequence (25). As a consequence, compared to the previous model (25), helix 4 contains two additional base-pairs: a G·C and a G·U pair.

In solution, the 5′-terminal region of naked snR17A RNA is folded into a two helical structure (Figure 2B), instead of a single helical structure as in vertebrate RNA. Helix 1b fits perfectly with that previously proposed (26). For the reasons mentioned above, only part of the previously proposed helix 1a was correct. In the revised helical structure 1a, Box A encompasses the top loop and part of the right strand of the stem, as was proposed for *S.pombe, D.discoïdeum* and tobacco U3 snRNAs (26, 27). Surprisingly Box A has a different geometry in vertebrate RNAs; it encompasses the left strand of the stem and contains an internal loop (24). It is also noticeable that, whereas helix 1a is rather well conserved in size, shape and sequence, among lower eukaryotes and plants, helix 1b and especially its two bordering single-stranded sequences are quite variable. One explanation for the observation of poor conservation of the secondary structure of the 5′-terminal region of U3 snRNA could be that, in its functional state, it is base-paired to another RNA likely the pre-rRNA. A base-pairing with another small RNA species, as is the case for U4 and U6 snRNAs, can also be postulated, but no association between U3 snRNA and another small RNA species

has been detected yet, whereas U3 snRNA/pre-rRNA complexes have already been observed. In light of this hypothesis, the low degree of nucleotide sequence conservation (except Box A) of the 5'-domain of U3 snRNA may be a consequence of the poor sequence conservation of the pre-rRNA 5'-ETS (36). In the case of *S.cerevisiae*, strong evidence for a possible base-pairing between the 5'-terminal domain of snR17A RNA and the 5'-ETS of pre-rRNA was recently obtained from both crosslinking and genetic experiments (20, 21). Interestingly, the snR17A segment involved in the suggested intermolecular base-pairing (39–49) corresponds to the segment linking helix 1a to helix 1b and a few nucleotides within these two helices (Figure 4). Initiation of the intermolecular base-pairing may be favoured by the single-stranded state of segment 3949 in snR17A RNA.

Deduced information on snR17A RNA tertiary structure

In the conditions used for chemical modification of snR17A RNA in its native conformation, most of the postulated single-stranded adenine residues were the target of DMS methylation. During prolonged incubation, incubation with higher amounts of DMS or incubation under semi-denaturing conditions, a few base-paired adenine residues were modified at low yield (Figures 2 and 3). Under non-denaturing conditions, almost all bulged pyrimidines and pyrimidines in terminal loops were reactive to chemical probes, but as a result of base-stacking or non-canonical base-pairing interactions only a limited number of those in internal loops were modified. By using longer times for incubation or larger amounts of CMCT, new sites of modification appeared in the internal single-strands and, as previously reported (32), uridine residues in stems with low stability were modified under semi-denaturing conditions (Figures 2 and 3). Thus, information on RNA overall conformation could be obtained from an inspection of the pyrimidine reactivity and also from a study of the relative sensitivity of single-strands and double-strands to S1 and V1 nuclease, respectively.

In the 5'-terminal domain (Figure 2B), internal and terminal loops of helices 1a and 1b are true loops without base-stacking or non-canonical interactions since almost all nucleotides are the target of both S1 nuclease and chemical modifications. Single-stranded segments 42–46 and 90–95, although modified by chemical reagents, were not cleaved by S1 nuclease. They may be involved in some weak interactions in naked RNA which may be disrupted upon association of snR17A RNA with snRNP proteins or with the pre-rRNA 5'-ETS.

The most S1 nuclease-sensitive single-strand in the 3'-domain (Figure 3B) is the terminal loop of helix 4. It should be very accessible within the tertiary structure of the molecule and should have an open conformation. Terminal loops of helices 2 and 3, whereas highly susceptible to CMCT, are poorly digested by S1 nuclease. All putative internal loops of this domain were not cleaved or cleaved at low yield by S1 nuclease. Base-stacking or non-canonical base-pairings probably occur in these loops (Figure 3). In helix 4, the unreactive C140 and C171 on one hand, and the unreactive U141 and U170 (only modified in semi-denaturing conditions) on the other hand, may form pyrimidine pairs. In helix 2, three pyrimidine pairs are predicted by the absence of chemical modification under non-denaturing conditions and by the presence of V1 RNase cleavages U204–U245, C206–U243 and U217·U231. Interestingly, the two strands of helix 2, which have a discontinuous complementarity, form a rather continuous helical structure that is very sensitive to V1 nuclease.

Based on the results of wild-type snR17A RNA analysis, it was difficult to conclude anything about possible interactions between the two single-stranded segments linking helix 5 to helix 4 and helix 2 to helix 3. These two segments contain Box B and C, respectively. Mutant snR17A genes deleted of either the helix 2 coding region or the helix 4 coding region or both had been constructed (37) (Figure 4). The susceptibility towards chemical probes and nucleases of the *in vitro* transcripts of these mutant genes were studied in non-denaturing conditions (Figure 5).

In wild-type snR17A RNA, only two phosphodiester bonds of the 251–262 fragment containing Box C (Figure 3B) were cleaved by S1 nuclease. The presence of 4 V1 RNase cleavage sites and the absence of reactivity of U 258, 260 and 261 to CMCT under non-denaturing conditions suggest the existence of a higher order structure. No enzymatic cleavage was observed in the 107–119 fragment containing Box B, but all adenine residues were methylated by DMS at low yield. Nothing could be said about U112 as a parasite band obscure the reverse transcriptase analysis at this position. In the short fragment linking helix 4 to helix 2, U196 and U197 were modified by CMCT and S1 nuclease cleaved bond 194–195.

When helix 2 was deleted (transcript Δ1) (Figure 5A), no significant variation of susceptibility to chemical or enzymatic probes were detected in helix 4 and helix 5. As expected, new S1 nuclease sites appeared in the 196–198 uridine stretch. We noticed the following changes in helix 3: a slight increase of V1 RNase cleavages at bonds 264–265 and the appearance of V1 RNase cleavages in the 303–306 segment. In the Box C segment, S1 nuclease cleavages at bonds 254–255, 257–258 were reinforced and, instead of being cleaved by V1 RNase, bonds 260–261 and 261–262 were cleaved by S1 nuclease. The V1 cleavage at bond 251–252 was conserved. The differences observed between wild-type RNA and Δ1 transcript could be explained by the existence in the wild-type RNA of an interaction between helices 2 and 3 also involving the 3'-extremity of their linking fragment. The persistence of V1 RNase cleavage at bond 251–252 in the Δ1 transcript was surprising. Indeed, this cleavage of the phosphodiester bond flanking helix 2 was very similar to cleavages found in tRNA (38) and U1 snRNA (39), which have been interpreted as cleavages of stacked nucleotides bridging coaxial helices. The strong V1 cleavage at bond 251–252 should therefore result either from internal interaction in the 251–262 fragment or from interaction of this fragment with sequences outside helix 2.

In transcript Δ2 (Figure 5B), helix 4 was deleted and helix 2 was linked to the 107–121 segment by an inserted additional GC sequence. This sequence allowed formation of two additional base-pairs in helix 2 and as a consequence strong additional V1 cleavages are observed. Interestingly, the V1 RNase cleavages at bonds 258–259, 259–260, 260–261 disappear, as well as the S1 cleavage at bond 254–255. And the putative internal loop 2a becomes sensitive to S1 nuclease. Therefore, here again deletion of one hairpin structure has an effect on the higher order structure of the segment containing Box C and on the bottom part of helix 2. Nucleotides 251 and 252 being involved in additional base-pairings of helix 2 may be the reason for the disappearance of the higher order structure in the segment containing Box C. This suggests that the V1 cleavages at bonds 251–252, 258–259 and 259–260 in wild-type RNA are related to a higher order structure involving nucleotides 251 to 261.

In transcript Δ3 (Figure 5C), where the sequence 108–250 was deleted, the segment 251–262 that now links helix 5 to helix

3 has a completely different structure, as evidence by the disappearance of V1 cleavages and the appearance of S1 cleavages at all bonds between residues 254 and 262. We also note the appearance of new V1 cleavages in the 102–105 fragment.

Altogether, these deletion experiments demonstrate the existence in the 3′-domain of a complex conformation of the single-stranded segments joining helices and the bottom parts of these helices. As a consequence, the fragment containing Box C has a higher order structure that is altered by deletion of any of the helices. In contrast, the structures at the extremity of each helice are completely independent from each other. A short base-pairing between the UCU sequence (258–260) of Box C and the AGA sequence (108–110) of Box B is compatible with our results. This would explain the V1 cleavages in segment 258–260. However in the present state of the study, we have no absolute proof for it. The possibility to form a three base-pair interaction, including a G·U pair, is conserved throughout evolution (24, 27). It is extended to 5 base-pairs in tomato RNA, which may be to compensate for the very low stability of helix 3 in this RNA. Clearly, this possible base-pairing is not sufficient to account for the overall reactivity of segment 251–262 to chemical modification and nuclease digestion. Other higher order interactions not yet identified should exist and experiments are underway to get additional information.

Concluding remarks

i) snR17A RNA extracted from yeast cells and *in vitro* synthesized snR17A RNA, were found to have similar structures, suggesting that post-transcriptional modifications have a limited influence on snR17A RNA structure.

ii) The 3′-domain of snR17A RNA has a cruciform shaped secondary structure, with a complex network of tertiary interactions in the central region, involving the bottom parts of helices and the single-strands linking them. The central role of nucleotides at the junction of helical domains, in determining the coaxial stacking interactions and tertiary structure of RNA, was already described in both tRNA (38) and 5S rRNA (40). It should be noticed that in U3 snRNA these single-stranded junction segments have a particularity compared to those in tRNA, 5S rRNA, and U1 snRNA: they are quite longer, they correspond to the most highly evolutionarily conserved segments of U3 snRNA, and they are probably interacting with the U3 snRNP proteins, especially the fibrillarin. Clearly among RNAs whose conformations have been studied in detail, the 3′-domain of U3 snRNA displays a different mode of tertiary folding that involves interactions between single-stranded segments linking helices but no long-range interactions between terminal loops.

iii) The 5′-terminal region of naked snR17A RNA has a two-helical structure that is different from that found in human U3 snRNA. The loose tertiary structure of the 5′-terminal region of U3 snRNA and the great variability of its length and secondary structure can be explained by a base-pairing with the pre-ribosomal RNA. This is in accord with recent results obtained from crosslinking and genetic experiments (20, 21).

ACKNOWLEDGEMENTS

K.Tanner is thanked for helpful discussion. M.Chevrier-Miller is thanked for providing the mutant snR17A genes, which have been used in this study. The expert technical assistance of J.Bayeul and E.Habermacher is acknowledged. V.Ségault and A.Grégoire were supported by grants of the french Ministry of Research and Technology.

REFERENCES

1. Zieve,G. and Penman,S. (1976) *Cell* **8**, 19–31.
2. Reddy,R. and Busch,H. (1988) In Birnstiel,M.L. (ed.) *Structure and function of major and minor small nuclear ribonucleoprotein particles*, Springer-Verlag, Berlin, pp. 1–37.
3. Lührmann,R., Kastner,B. and Bach,M. (1990) *Biochim. Biophys. Acta* **1087**, 265–292.
4. Kass,S., Tyc,K., Steitz,J.A. and Sollner-Webb,B. (1990) *Cell* **60**, 897–908.
5. Tiollais,P., Gallibert,F. and Boiron,M. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1117–1120.
6. Hadjiolov,A.A. (1985) *The Nucleolus and Ribosome Biogenesis*, Springer-Verlag, New York.
7. Miller,K.G. and Sollner-Webb,B. (1981) *Cell* **27**, 165–174.
8. Gurney,T.,Jr (1985) *Nucleic Acids Res.* **13**, 4905–4919.
9. Kass,S., Craig,N. and Sollner-Webb,B. (1987) *Mol. Cell. Biol.* **7**, 2891–2898.
10. Prestayko,A.W., Tonato,M. and Busch,H. (1969) *J. Mol. Biol.* **47**, 505–515.
11. Epstein,P., Reddy,R. and Busch,H. (1984) *Biochemistry* **23**, 5421–5425.
12. Bachellerie,J.P., Michot,B. and Raynal,F. (1983) *Mol. Biol. Rep.* **9**, 79–86.
13. Crouch,R.J., Kanaya,S. and Earl,P.L. (1983) *Mol. Biol. Rep.* **9**, 75–78.
14. Tague,B.W. and Gerbi,S.A. (1984) *J. Mol. Evol.* **20**, 362–367.
15. Maser,R.L. and Calvet,J.P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6523–6527.
16. Stroke,I.L. and Weiner,A.M. (1989) *J. Mol. Biol.* **210**, 497–512.
17. Parker,K.A., Bruzik,J.P. and Steitz,J.A. (1988) *Nucleic Acids Res.* **16**, 10493–10509.
18. Tollervey,D. (1987) *EMBO J.* **6**, 4169–4175.
19. Savino,R. and Gerbi,S.A. (1990) *EMBO J.* **9**, 2299–2308.
20. Beltrame,M. and Tollervey,D. (1992) *EMBO J.* **11**, 1531–1542.
21. Hughes,J.M.X. and Ares,M.,Jr (1991) *EMBO J.* **9**, 4231–4239.
22. Wise,J.A. and Weiner,A.M. (1980) *Cell* **22**, 109–118.
23. Bernstein,L.B., Mount,S.M. and Weiner,A.M. (1983) *Cell* **32**, 461–472.
24. Parker,K.A. and Steitz,J.A. (1987) *Mol. Cell. Biol.* **7**, 2899–2913.
25. Hughes,J.M.X., Konings,D.A.M. and Cesareni,G. (1987) *EMBO J.* **6**, 2145–2155.
26. Porter,G.L., Brennwald,P.J., Holm,K.A. and Wise,J.A. (1988) *Nucleic Acids Res.* **16**, 10131–10151.
27. Kiss,T. and Solymosy,F. (1990) *Nucleic Acids Res.* **18**, 1941–1949.
28. Myslinski,E., Ségault,V. and Branlant,C. (1990) *Science* **247**, 1213–1216.
29. Baserga,S.J., Gilmore-Herbert,M., Yang,X.D. and Steitz,J.A. (1991) *RNA Processing*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, NY.
30. Ochs,R.L., Lischwe,M.A., Spohn,W.H. and Busch,H. (1985) *Biol. Cell.* **54**, 123–134.
31. Tyc,K. and Steitz,J.A. (1989) *EMBO J.* **8**, 3113–3119.
32. Ehresmann,C., Baudin,F., Mougel,M., Romby,P., Ebel,J.-P. and Ehresmann,B. (1987) *Nucleic Acids Res.* **15**, 9109–9127.
33. Vassilenko,S.K. and Babkina,G.T. (1965) *Biokhymiya* **30**, 705–712.
34. Branlant,C., Krol,A., Ebel,J.-P, Gallinaro,H., Lazar,E. and Jacob,M. (1981) *Nucleic Acids Res.* **9**, 841–857.
35. Chapman,K.A. and Burgess,R.R. (1987) *Nucleic Acids Res.* **15**, 5413–5431.
36. Michot,B. and Bachellerie,J.-P. (1991) *Eur. J. Biochem.* **195**, 601–609.
37. Banroques,J., Ségault,V., Mougin,A., Grégoire,A., Chevrier-Miller,M. and Branlant,C. Manuscript in preparation.
38. Romby,P., Moras,D., Bergdoll,H., Dumas,P., Vlassov,V.V., Westhof,E., Ebel,J.-P. and Giegé,R. (1985) *J. Mol. Biol.* **184**, 455–471.
39. Krol,A., Westhof,E., Bach,M., Lührmann,R., Ebel,J.P. and Carbon,P. (1990) *Nucleic Acids Res.* **18**, 3803–3811.
40. Baudin,F., Romaniuk,P.J., Romby,P., Brunel,C., Westhof,E., Ehresmann,B. and Ehresmann,C. (1991) *J. Mol. Biol.* **218**, 69–81.