MOLECULAR AND CELLULAR BIOLOGY, Aug. 1987, p. 2899–2913
0270-7306/87/082899-15$02.00/0

Vol. 7, No. 8

Structural Analyses of the Human U3 Ribonucleoprotein Particle Reveal a Conserved Sequence Available for Base Pairing with Pre-rRNA

KATHERINE A. PARKER* AND JOAN A. STEITZ

Howard Hughes Medical Institute, Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Received 26 February 1987/Accepted 20 May 1987

The human U3 ribonucleoprotein (RNP) has been analyzed to determine its protein constituents, sites of protein-RNA interaction, and RNA secondary structure. By using anti-U3 RNP antibodies and extracts prepared from HeLa cells labeled in vivo, the RNP was found to contain four nonphosphorylated proteins of 36, 30, 13, and 12.5 kilodaltons and two phosphorylated proteins of 74 and 59 kilodaltons. U3 nucleotides 72-90, 106-121, 154-166, and 190-217 must contain sites that interact with proteins since these regions are immunoprecipitated after treatment of the RNP with RNase A or T_1. The secondary structure was probed with specific nucleases and by chemical modification with single-strand-specific reagents that block subsequent reverse transcription. Regions that are single stranded (and therefore potentially able to interact with a substrate RNA) include an evolutionarily conserved sequence at nucleotides 104-112 and nonconserved sequences at nucleotides 65-74, 80-84, and 88-93. Nucleotides 159-168 do not appear to be highly accessible, thus making it unlikely that this U3 sequence base pairs with sequences near the 5.8S rRNA-internal transcribed spacer II junction, as previously proposed. Alternative functions of the U3 RNP are discussed, including the possibility that U3 may participate in a processing event near the 3′ end of 28S rRNA.

U3 RNA is small (217 nucleotides in the human cells [59]), abundant (about 2×10^5 copies per mammalian cell [64]), conserved from yeasts to vertebrates (1, 30, 47, 57, 59, 65; C. Jeppesen, M. Ares, and S. A. Gerbi, personal communication; J. Hughes, D. Konings, and G. Cesareni, EMBO J., in press), and the only vertebrate nucleolar RNA known to have an m_3G cap (7). The U3 ribonucleoprotein (RNP) can be selectively immunoprecipitated by certain patient autoantibodies (34) and a monoclonal antibody (G. Reimer, K. M. Pollard, C. A. Penning, R. L. Ochs, M. A. Lischwe, H. Busch, and E. M. Tan, Arthritis Rheumatism, in press) that react with an associated 34-kilodalton (kDa) protein. The nucleolar location of U3 (7, 34, 43, 45, 64; Reimer et al., in press) and the observations that it can be found associated with a >60S RNP particle (16) and with a 28 to 35S RNA (after deproteinization) (8, 45, 67) suggest an involvement in rRNA processing, ribosome assembly, or transport.

Specific theories regarding the function of the U3 RNP (3, 13, 60) have been based primarily on assumptions concerning the identity of the 28 to 35S RNA with which it associates (8, 45, 67) and on complementarity between U3 and pre-rRNA sequences. In mammalian cells, a 47S precursor is processed in several steps to yield 18S, 5.8S, and 28S rRNAs (reviewed by Hadjiolov [24]). One abundant processing intermediate is a 32S RNA, which contains 5.8S and 28S covalently linked by the internal transcribed spacer II (ITS-II). Bachellerie et al. (3) and, later, Tague and Gerbi (60) noted that the sequence at the extreme 5′ end of the ITS-II was conserved among vertebrate species and was complementary to nucleotides 158-167 of rat U3 RNA. They therefore postulated that U3 may base pair to this sequence and mediate the cleavage that separates 5.8S from the ITS-II.

Here we report structural studies and a derived model for the secondary structure of the RNA in the human U3 RNP. The protein constituents of the U3 RNP and protein-RNA interaction sites have also been analyzed. We conclude that U3 nucleotides 159 to 168 (human) are not highly accessible in the RNP and therefore probably do not base pair with ITS-II sequences. In contrast, a highly conserved sequence at nucleotides 104-112 is available for interacting with pre-rRNA. Complementarity between this region of U3 and conserved pre-rRNA sequences suggests that U3 may be involved in processing events occurring at or beyond the 3′ end of 28S rRNA.

MATERIALS AND METHODS

Materials. Monoclonal antibody 72B9 was provided by G. Reimer and E. Tan from the Scripps Institute (Reimer et al., in press), and patient antisera were donated by R. Sontheimer from the University of Texas, Dallas (LS, JH, and RB), by T. Mimori and M. Homma from Keio University, Tokyo (MK), and by J. Hardin and J. Craft from Yale University (SC, AG, and ON).

The chemical reagents used include dimethyl sulfate (DMS; A>C specific), kethoxal (G specific), and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluenesulfonate (CMCT; U>G specific).

Methods. (i) Gel electrophoresis. All RNA samples were analyzed on 8 M urea-polyacrylamide gels (19:1 acrylamide-bisacrylamide) with 1× TBE (90 mM Tris base, 90 mM boric acid, 2.5 mM EDTA).

(ii) Buffers. Modification buffers were altered from those of Moazed et al. (41): CMK (80 mM potassium cacodylate, pH 7.2, 100 mM KCl, 5 mM $MgCl_2$) for the DMS and kethoxal reactions; BMK (80 mM potassium borate, pH 8.1, 100 mM KCl, 5 mM $MgCl_2$) for the CMCT reaction; or TMK

* Corresponding author.

(80 mM Tris hydrochloride [pH 7.5], 100 mM KCl, 5 mM $MgCl_2$) for the cobra venom reaction.

(iii) Preparation of extracts and naked RNA. HeLa cells grown to a density of 4×10^5 to 5×10^5 cells per ml were washed with 1/10 the original volume of TBS (40 mM Tris hydrochloride, pH 7.5, 150 mM NaCl) and suspended in 300 μl (DMS, kethoxal, and cobra venom reactions) or 200 μl (CMCT reaction) of modification buffer per 1.6×10^7 cells. Cells were sonicated on ice six times for 30 s each at setting 1 with a Branson microtip. The chromatin was then removed by two centrifugations for 2 min each at $12,000 \times g$ in a microfuge. Half of the extract was used for the RNP modification reactions. The rest was immediately treated with proteinase K (with 1% sodium dodecyl sulfate [SDS]), phenol extracted twice, and chloroform extracted once. Ethanol-precipitated RNA derived from 1.6×10^7 cells was suspended in the same volume of modification buffer as was used for that particular RNP reaction and was subsequently used for the naked-RNA reaction.

Modification reactions and cobra venom cleavage. (i) DMS. For the RNP reaction, 0 to 2.0 μl of DMS (gold label; Aldrich) was added to 300 μl of freshly prepared extract, and the samples were incubated at 25°C for 15 min. Then, 150 μl of cold DMS stop buffer (1.5 M sodium acetate, 1 M 2-mercaptoethanol, 1 M Tris acetate, pH 7.5, 0.1 mM EDTA), 300 μl of cold CMK buffer, and 80 μl of 10% SDS were added. Samples were immediately extracted twice with 400 μl of cold phenol and once with 400 μl of chloroform, and the RNA was ethanol precipitated. The naked-RNA reaction was identical except that the reaction was stopped with 75 μl of DMS stop solution and the RNA was ethanol precipitated without phenol extraction. Ethanol precipitates from the DMS and cobra venom reactions (see below) were suspended in 40 μl of TE (10 mM Tris hydrochloride, pH 7.5, 0.1 mM EDTA).

(ii) Kethoxal. For the RNP reaction, 300 μl of extract was mixed with 0 to 20 μl of a kethoxal (U.S. Biochemical Corp.) solution (37 mg/ml in 25% ethanol) and incubated for 15 min at 25°C. The reaction was stopped by adding 25 μl of 0.5 M potassium borate (pH 7.0), 400 μl of cold CMK buffer, 80 μl of 10% SDS, and 80 μl of 3 M sodium acetate, and then the RNA was isolated as in the DMS reaction. For the naked RNA, the highest amount of modifying reagent was empirically decreased to 5 μl to equalize the degree of modification of the RNP and RNA samples. A 15-μl volume of 0.5 M potassium borate was added at the end of the reaction period, and the samples were precipitated with ethanol. Ethanol precipitates from the kethoxal and CMCT reactions were suspended in 40 μl of 25 mM potassium borate (pH 7.0) to stabilize the adduct.

(iii) CMCT. For the RNP reaction, 200 μl of extract was mixed with 0 to 200 μl of CMCT (Aldrich) solution (42 mg/ml in BMK buffer) and BMK buffer (to a total added volume of 200 μl) and incubated for 15 min at 25°C. The reaction was stopped by decreasing the pH to 6.5 with 500 μl of cold 0.5 M potassium borate (pH 6.1), and the RNA was isolated as described for the DMS reaction. The naked-RNA reaction was identical except that 25 μl of 0.5 M potassium borate (pH 7.0) was added at the end of the reaction period.

(iv) Cobra venom. For the RNP reaction, 0 to 3.2 μl of cobra venom (Calbiochem; 10 U/mg) at 1 U/ml was added to 300 μl of extract and incubated for 15 min at 25°C. The reaction was stopped by adding 500 μl of cold TMK buffer and 80 μl of 10% SDS, and the RNA was isolated as described for the DMS reaction. For the naked-RNA reaction, samples were digested as described for the RNP reaction and then extracted once with phenol and once with chloroform, and the RNA was ethanol precipitated.

Reverse transcription. Oligonucleotides complementary to nucleotides 63-77, 106-120, 151-167, and 203-217 of U3 were 5′ end labeled. These were annealed (separately) to 10 μl of the RNA samples described above, and reverse transcription (avian myeloblastosis virus reverse transcriptase; Promega) was performed (56) after hybridization at 42°C for 3 h. To prepare sequencing ladders of unmodified RNA, 0.6 μl of dideoxynucleotide (3.3 mM ddATP, 2.0 mM ddCTP, 3 mM ddGTP, or 3 mM ddTTP) was added to four different reverse transcriptase reactions.

Enzymatic digestion of 3′ end-labeled U3. Immunoprecipitated U3 was 3′ end labeled with [^{32}P]pCp and RNA ligase (15) (Pharmacia) and gel purified. U3 (1,000 cpm; specific activity, 4.6×10^7 cpm/μg) was digested with increasing amounts of a fresh dilution of pancreatic (5,561 U/mg; Cooper), T_1 (2,941 U/mg; Calbiochem), or cobra venom nuclease for 15 min at 25°C in 10 μl of TMK buffer containing 2.5 μg of unlabeled total HeLa RNA. A 1-μl (4-μg) sample of proteinase K was added and incubated for 15 min at 25°C. A 15-μl volume of 99% formamide containing 0.05% xylene cyanol and 0.05% bromophenol blue was added, and samples were heated at 65°C for 3 min and loaded directly on a 20% polyacrylamide gel. RNA sequencing ladders were prepared enzymatically (14) (Pharmacia).

Immunoprecipitation of proteins labeled in vivo. Proteins were immunoprecipitated as described by Mimori et al. (40) except that NET-2 buffer (40 mM Tris hydrochloride, pH 7.5, 150 mM NaCl, 0.05% Nonidet P-40) was used throughout and each sample received either extract from 100 ml of HeLa cells labeled for 20 h with 1 mCi of [^{35}S]methionine or extract from 10 ml of cells labeled with 0.1 mCi of $^{32}PO_4$. After washing, all samples labeled with $^{32}PO_4$ were incubated for 15 min at 25°C with RNases A and T_1 at final concentrations of 2 mg/ml and 3,000 U/ml, respectively, and then loaded directly on the gel after boiling unless otherwise stated.

RNase protection. (i) In vitro label. For each sample, 1 μl of LS antiserum was prebound to 1 mg of protein A-Sepharose (Pharmacia) (40). The immunoprecipitate from the extract of 2×10^6 cells was suspended in 500 μl of NET-2. RNase (1.1 μl of A at 4.7 mg/ml or 5 μl of T_1 at 30 U/μl) was added, and the mixture was incubated at 37 or 25°C, respectively, for 15 to 60 min. Samples were then washed 10 times with NET-2 and once with NET (without Nonidet P-40). The RNA fragments were 5′ end labeled with the digested RNP still on the protein A-Sepharose beads, as follows. A 3-μl volume of 10× kinase buffer (700 mM Tris hydrochloride, pH 7.5, 100 mM $MgCl_2$, 50 mM dithiothreitol), 100 μCi of [γ-^{32}P]ATP (7,000 Ci/mmol), 20 μl of H_2O, and 3 μl (24 U) of polynucleotide kinase (Biolab) were added to the protein A-Sepharose pellet and incubated for 15 min at 37°C. The samples were then washed 5 times with 1 ml of NET-2 and 10 times with NET-2 containing 200 μg of yeast RNA per ml. The RNA was extracted with phenol-chloroform-isoamyl alcohol (50:49:1) and analyzed on a polyacrylamide gel.

(ii) In vivo label. Each protein A-Sepharose–antibody sample (prepared as described above) received extract derived from 10^6 HeLa cells labeled with 50 μCi of $^{32}PO_4$ for 14 h. They were treated in parallel with samples labeled in vitro except for the omission of the kinase reaction.

Computer methods. All software was in University of Wisconsin programs. Secondary structures and their free energy were determined by the FOLD program (68). To

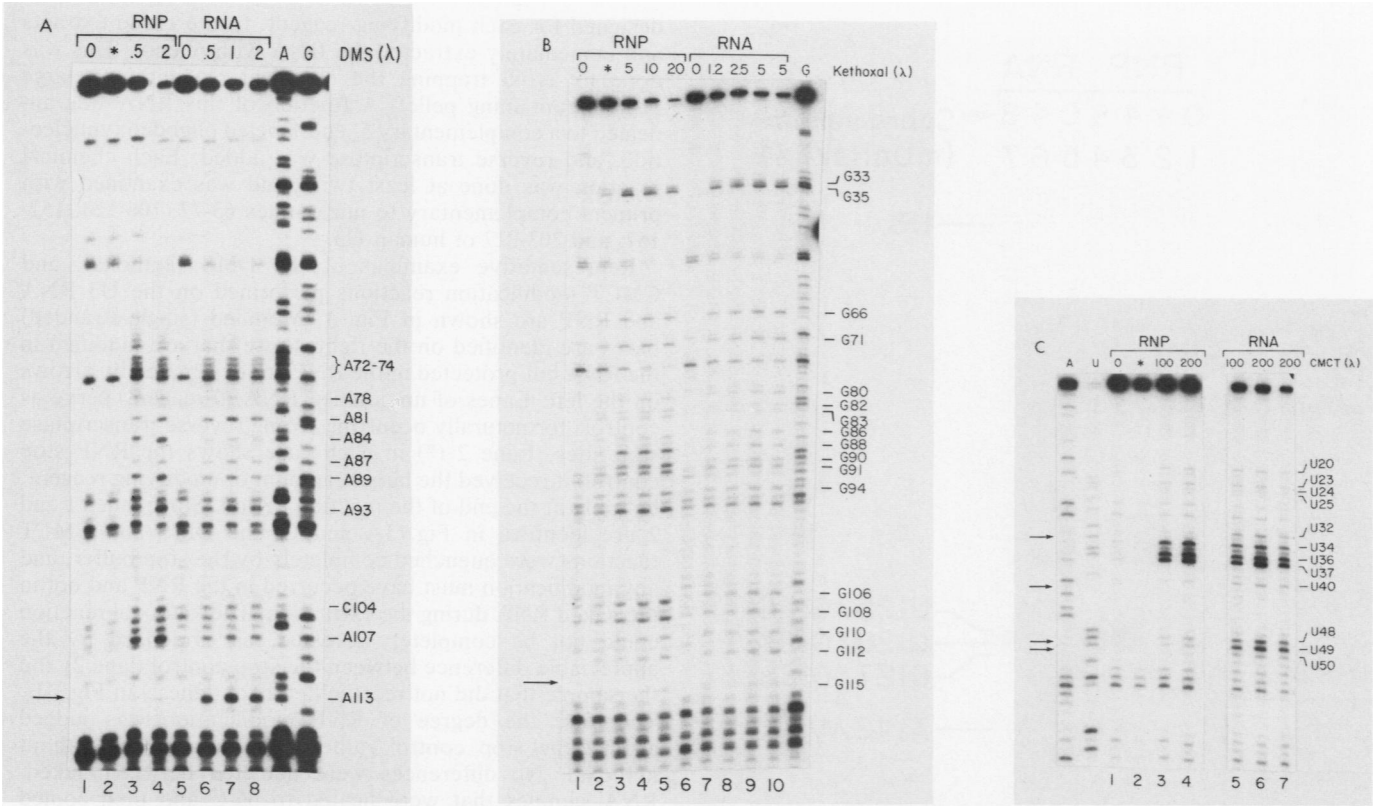

FIG. 1. Chemical modification of the U3 RNA and RNP. (A) DMS reaction. RNA or RNP samples were modified with the designated amount of DMS (in microliters), and then the extracted RNA was reverse transcribed with an oligodeoxynucleotide primer complementary to nucleotides 151-167 of U3 and electrophoresed on an 8% polyacrylamide-urea gel. Dideoxynucleotide sequencing lanes are labeled according to the base of the RNA they detect (e.g., the "A" lane received ddTTP). Modified bases identified at the right are one position below that base in the sequencing lane. Bases modified in the RNA but not the RNP are denoted by an arrow on the left. The * (lane 2) designates the control RNP sample, which received 2 μl of DMS after the addition of stop buffer. (B) Kethoxal reaction. Samples were modified with the designated amount of kethoxal stock solution (in microliters) and then analyzed with a primer complementary to nucleotides 151-167 of U3. The * (lane 2) designates the control RNP sample, which received 20 μl of kethoxal after the addition of stop buffer. The samples in lanes 6 through 9 were modified immediately after suspension in modification buffer, whereas the sample in lane 10 was heated to 65°C and cooled slowly before the addition of kethoxal. (C) CMCT reaction. Samples were modified with the designated amount of CMCT stock solution (in microliters) and then analyzed with a primer complementary to nucleotides 106-120 of U3. The * (lane 2) designates the control RNP sample, which received 200 μl of CMCT after the addition of stop buffer. The naked-RNA sample that received no CMCT was lost. The samples in lanes 5 and 6 were modified immediately after suspension in modification buffer, whereas the sample in lane 7 was heated to 65°C and then slowly cooled before the addition of CMCT.

incorporate experimental data, sites of major chemical modification were prevented from base pairing. For comparative analysis, the GAP program was used to align each sequence with the human sequence. Bases that aligned with known single-stranded regions of human U3 were prevented from base pairing in the FOLD program to obtain secondary structures similar to that of the human U3. To search for complementarity between U3 RNA and pre-rRNA sequences, the programs WORDSEARCH and SEGMENTS were used to find matches of ≥5 nucleotides within 100 nucleotides of the sites shown in Fig. 9A. The U3 sequences that were evaluated are listed in Table 1 and were chosen by using the GAP function to align U3 regions from the nonhuman sequences with the single-stranded regions of human U3. Human (17, 20, 61; J. L. Gorski and R. D. Schmickel, personal communication for sequences downstream of the 3′ end of 28S), rat (9–11, 49, 58, 66), *Xenopus laevis* (25, 32, 35, 51, 52, 55), *Dictyostelium discoideum* (44), and *Saccharomyces cerevisiae* (4, 18, 50, 54, 63) pre-rRNA sequences were compared with the homologous U3 sequence; mouse pre-rRNA sequences (19, 21, 31, 36–38, 46, 62) were compared with the rat U3 sequence.

RESULTS

Analysis of RNA secondary structure. The secondary structure of U3 RNA was determined by chemically modifying single-stranded nucleotides in their base-pairing moieties and detecting the modifications by their ability to block reverse transcription from a downstream primer (27, 41). Because the altered nucleotide cannot base pair normally, a band 1 nucleotide below the position of that base in a dideoxynucleotide sequencing lane of unmodified RNA is produced. This method is particularly powerful because the use of specific primers for reverse transcription allows a single RNA species present in a complex mixture to be selectively analyzed. The chemical reagents used are described in Materials and Methods.

In determining the U3 RNA secondary structure, we analyzed the RNP (which is most likely the biologically active form) as well as the naked RNA. In previously reported studies of the secondary structure of an RNA (16S) in a purified RNP particle (30S ribosomes) (41), the chemical modification reactions were terminated by ethanol precipitation. We instead added a cold stop buffer, specifically

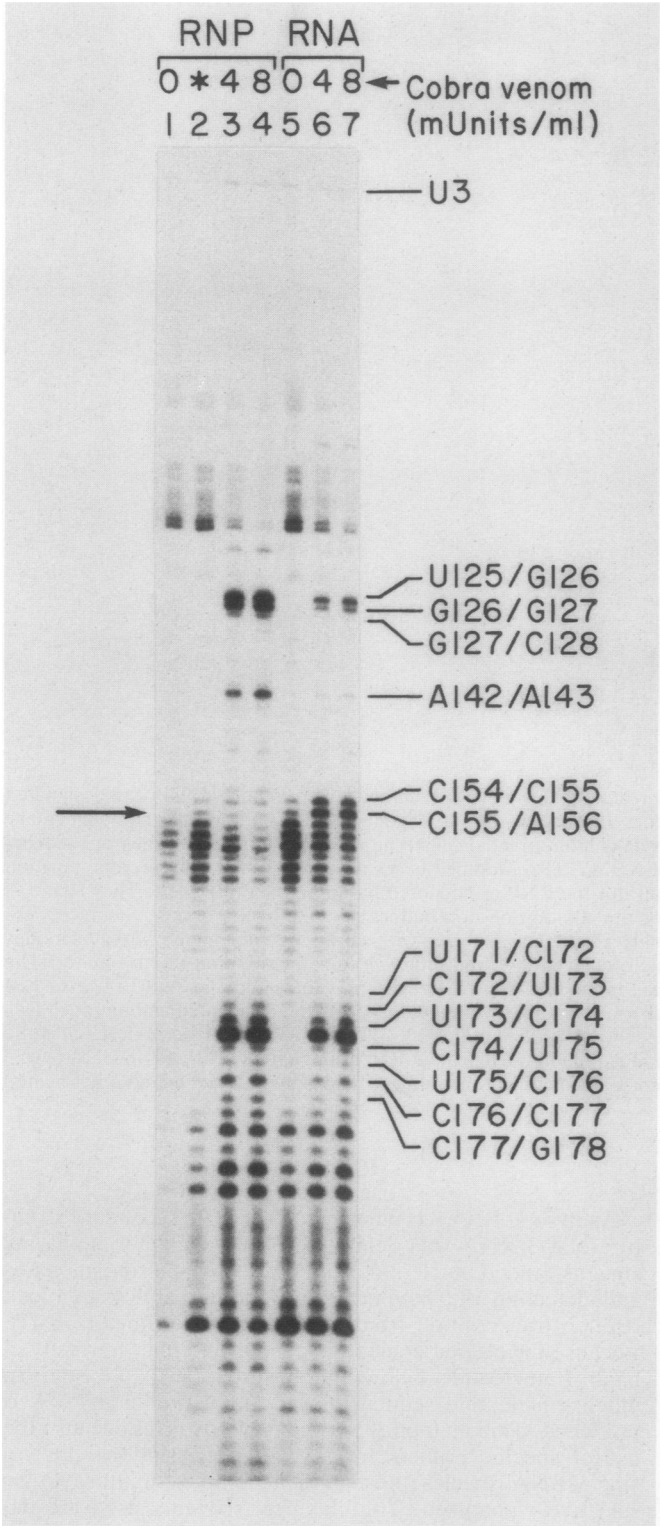

FIG. 2. Cobra venom cleavage of the U3 RNA and RNP. Samples treated with the designated final concentration of cobra venom nuclease were phenol extracted, and the RNA was reverse transcribed with a primer complementary to nucleotides 203-217 of U3. Cleavage sites are identified at the right by the bases between which the cut occurs. The * (lane 2) denotes the control RNP sample, which received 8 mU of cobra venom nuclease per ml (final concentration) only after the addition of stop buffer.

designed for each modifying reagent, to the crude extracts and immediately extracted the RNA with phenol. This was done to avoid trapping the chemical reagent in a large protein-containing pellet. A fraction of this RNA was annealed to a complementary 5′ end-labeled oligodeoxynucleotide, and reverse transcriptase was added. Each chemical reaction was done at least twice and was examined with primers complementary to nucleotides 63-77, 106-120, 151-167, and 203-217 of human U3.

Representative examples of the DMS, kethoxal, and CMCT modification reactions performed on the U3 RNA and RNP are shown in Fig. 1. Modified (single-stranded) bases are identified on the right; those that are modified in the RNA but protected in the RNP are designated by arrows on the left. Lanes of unmodified RNA (0 reagent) serve as controls for naturally occurring strong reverse transcriptase stop sites. Lane 2 (*) in each case shows the RNP stop control: it received the largest amount of modifying reagent, but only at the end of the reaction period. Since lanes 1 and 2 are identical in Fig. 1A and C, the DMS and CMCT reactions were quenched completely by the stop buffer, and any modification must have occurred in the RNP and not in the naked RNA during the extraction. The kethoxal reaction could not be completely arrested, as evidenced by the appreciable difference between the stop control (lane 2) and the sample that did not receive kethoxal (lane 1) in Fig. 1B. Therefore the degree of RNP modification was judged against the stop control rather than the control without kethoxal. No differences were detected between naked-RNA samples that were heated to 65°C and then cooled slowly (Fig. 1B, lane 10; Fig. 1C, lane 7) and RNAs that were simply suspended after ethanol precipitation (Fig. 1B, lane 9; Fig. 1C, lane 6). It is important to note that it is the relative, not the absolute, intensities of bands in the RNP versus the naked-RNA reactions that must be compared (see, for example, Fig. 1A and B, where there is a generalized increase in the intensity of the RNP as compared with the RNA lanes at the sites of modified bases).

To define further the bases involved in secondary structure, RNP and RNA samples were treated with the double-strand-specific RNase cobra venom nuclease (53). Cleavage sites were again located by reverse transcription with four primers. Figure 2 shows a representative example of cobra venom cleavage probed with the 203-217 oligonucleotide. The ability to stop the RNP reaction completely is demonstrated by the absence of new bands in the stop control (lane 2) as compared with the sample that did not receive cobra venom (lane 1). Apparent differences between the extent of cleavage in the RNP versus the naked RNA at position A142/A143 and beyond are due to the difficulty in evaluating reactivities at sites distant from the annealed primer; no differences between the RNP and the RNA were noted for most of this region when a primer complementary to nucleotides 151-167 was used.

The use of reverse transcription to detect chemical modifications or cobra venom cleavages fails to give information regarding the secondary structure of the extreme 3′ end of U3. Therefore, uniquely 3′ end-labeled naked U3 was incubated with pancreatic, T_1, or cobra venom nuclease under nondenaturing conditions. The samples were then separated on a polyacrylamide gel alongside enzymatic sequencing reactions (performed under denaturing conditions) (Fig. 3). More than 80% of the U3 remained full length after digestion with the single-strand-specific enzymes, making it unlikely that any cleavages were secondary events. The fragments created by cobra venom have 5′ phosphates, which makes

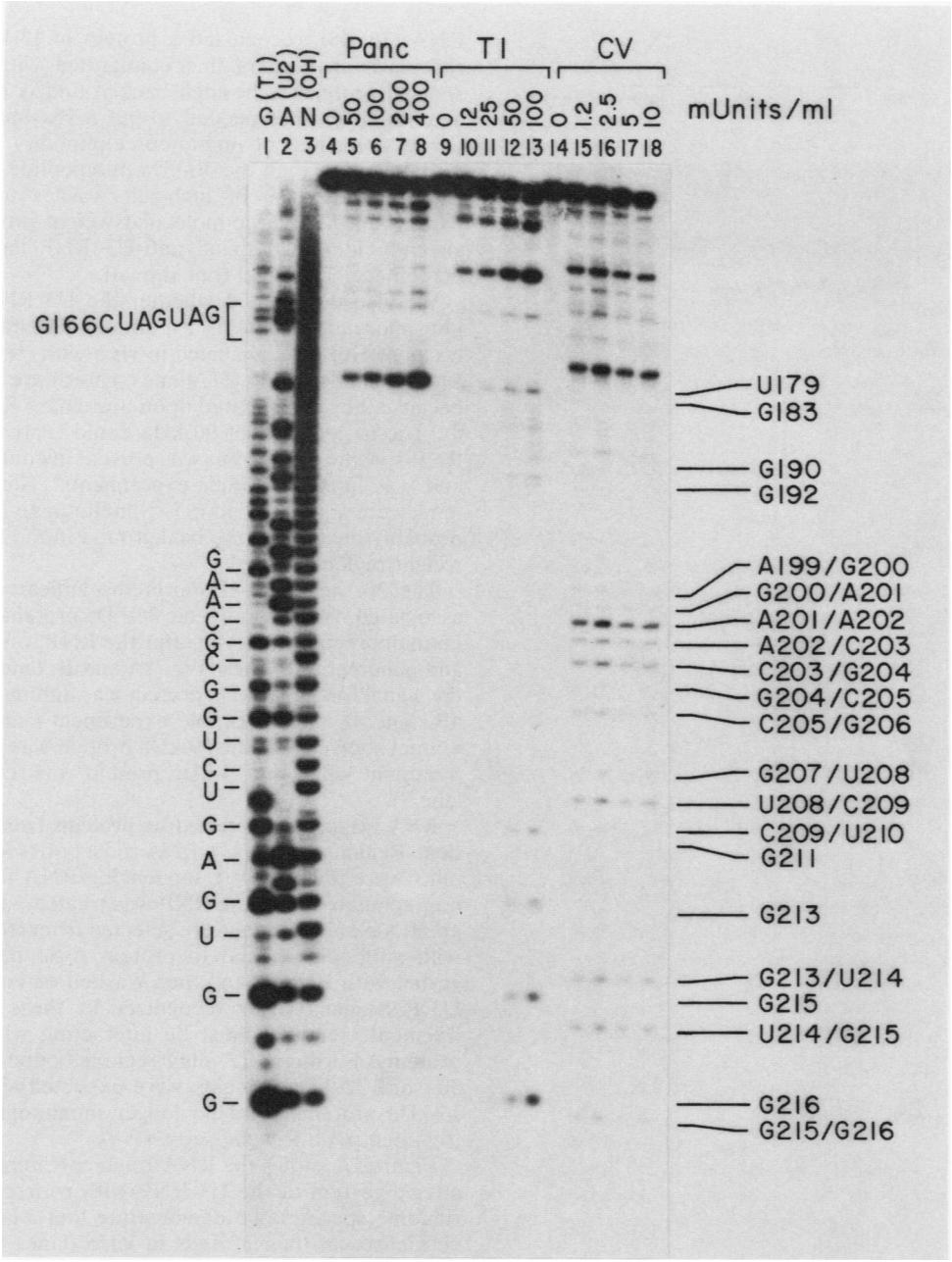

FIG. 3. Enzymatic cleavage of 3′ end-labeled RNA. Gel-purified, 3′ end-labeled U3 was digested for 15 min at 25°C with the designated final concentration of pancreatic, T_1, or cobra venom (CV) nuclease under nondenaturing conditions. Proteinase K was added, and the samples were electrophoresed directly on a 20% polyacrylamide-urea gel alongside enzymatic sequencing reactions incubated under denaturing conditions (14). On the right, positions cut by a single-strand-specific enzyme are denoted by a single base and sites of cleavage by the double-strand-specific enzyme cobra venom nuclease are denoted by two bases separated by /.

them migrate as if they were half a nucleotide shorter than the fragments created by T_1 or pancreatic RNase, which have 5′ hydroxyls. Data on the RNP were not obtained because U3 could not be end labeled while still associated with proteins.

Identification of the protein constituents of the U3 RNP. Four patient antisera that specifically immunoprecipitate the U3 RNA were identified; a U3 RNP-specific mouse monoclonal antibody, 72B9 (Reimer et al., in press), was generously provided by Reimer and Tan. All these antibodies recognized a protein of 36 kDa on immunoblots (not shown) of HeLa cell nuclear proteins (which must be the protein determined to be 34 kDa by others [34; Reimer et al., in press]) and failed to immunoprecipitate deproteinized U3 (not shown).

Figures 4A and B show the proteins immunoprecipitated by different anti-U3 RNP antibodies from an extract of HeLa cells labeled in vivo with [^{35}S]methionine. In addition to the 36-kDa polypeptide, which contained the antigenic determinant(s), proteins of 30, 13, and 12.5 kDa were immunoprecipitated with all anti-U3 RNP antibodies (lanes 5 to 9) but not with a normal human (lane 1) or an anti-U1 RNP (lane 2) serum. An anti–7-2S/8-2S RNP serum (lane 3), which immunoprecipitates a different nucleolar RNP containing 7-2S

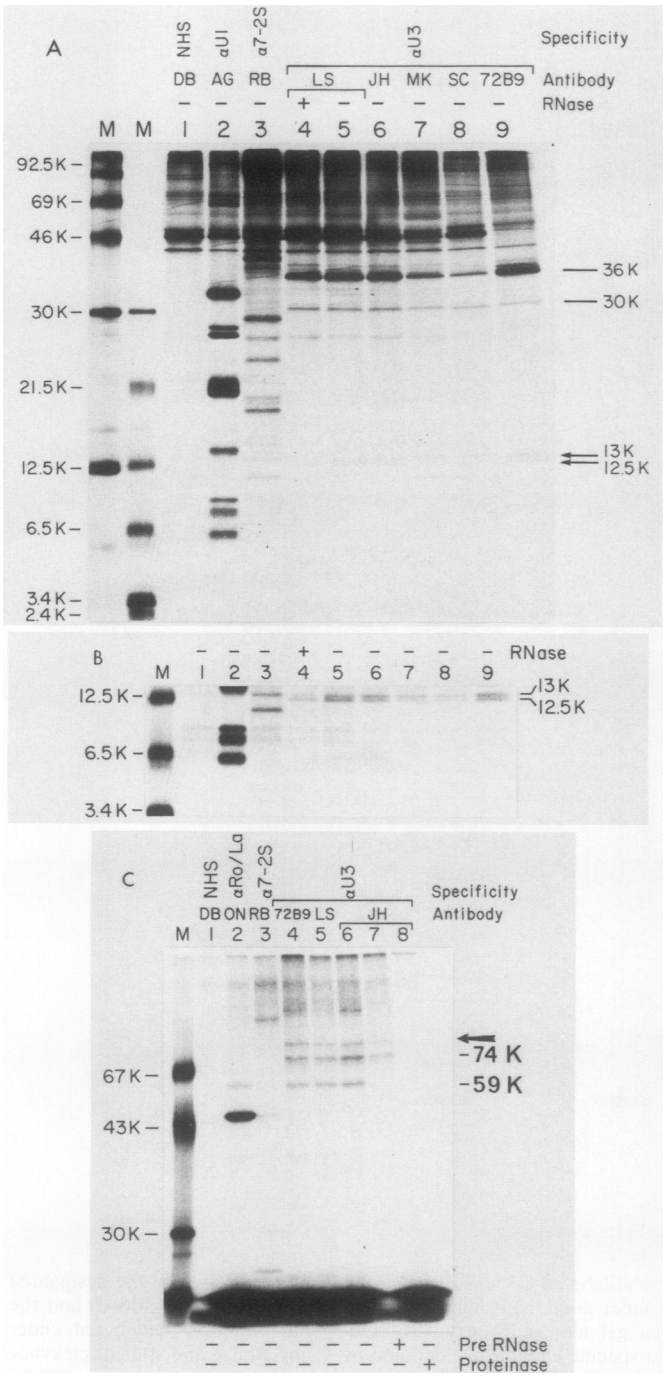

FIG. 4. Protein constituents of the U3 RNP. (A) Proteins labeled in vivo with [^{35}S]methionine were immunoprecipitated from a HeLa cell extract with various anti-U3 RNP and control antibodies and then analyzed on a 15% SDS-polyacrylamide gel (33). NHS refers to a normal human serum. On the right, the U3 RNP-specific proteins are indicated; their molecular weights were calculated from a semi-log plot of the ^{14}C-protein markers (Amersham) shown in lanes M. Lane 4 contains a sample that was treated with T_1 (660 U/ml) and pancreatic (0.5 mg/ml) RNases during the immunoprecipitation (RNase +). (B) Low-molecular-weight proteins labeled with [^{35}S]methionine are shown by a longer exposure of the gel shown in panel A. (C) Proteins labeled in vivo with $^{32}PO_4$ were immunoprecipitated from a HeLa cell extract and analyzed on a 10% SDS-polyacrylamide gel (33). All samples were treated with pancreatic and T_1 RNases after immunoprecipitation; the sample in lane 7 then received three additional washes (+ pre-RNase) while the others did not. The sample in lane 8 was also digested with proteinase K (2 mg/ml) for 10 min before being loaded on the gel. Definite U3 RNP-specific proteins are identified on the right by their size; a possible U3 RNP-specific protein of 90 kDa is designated by an arrow. Lane M had iodinated small-molecular-weight markers (Pharmacia).

RNA (26, 48), recognized a protein of 13 kDa but did not recognize any others that comigrated with the U3 RNP-specific proteins. The high background is due to the large amount of extract needed to get a reasonable signal (see Discussion); loss of immunoprecipitability of the U3 RNP proteins other than the 36-kDa polypeptide in 0.25 M NaCl precluded the use of high-salt washes to decrease this background. No large-molecular-weight proteins were consistently identified in all anti-U3 RNP lanes on a lower percentage (7.5%) gel (not shown).

We also determined whether the U3 RNP contains any phosphorylated proteins by immunoprecipitating from extracts of HeLa cells labeled in vivo with $^{32}PO_4$. The 74- and 59-kDa bands in Fig. 3C (lanes 4 to 6) are clearly proteins because they disappeared upon proteinase K treatment (lane 8). The larger band of 90 kDa could represent another U3 RNP-specific protein but was present in control lanes (1 to 3) and was fainter in other experiments. These phosphoproteins were not identified in [^{35}S]methionine-labeled extracts, probably because of the background in the large-molecular-weight region.

The 30- and 12.5-kDa proteins appeared to be closely associated with the antigenic 36-kDa protein since they were immunoprecipitated even after the RNP was treated with T_1 and pancreatic RNases (Fig. 4A and B, lane 4); conversely, the signal for the 13-kDa protein was diminished >80% (Fig. 4B, lane 4). A comparable experiment using $^{32}PO_4$-labeled extract showed that the 59-kDa protein was lost after RNase treatment while the 74-kDa protein was retained (Fig. 4C, lane 7).

RNA fragments protected by proteins from nuclease digestion. Regions of the U3 RNA closely associated with proteins were identified by sequencing RNA fragments immunoprecipitated after the RNP was treated with T_1 or pancreatic RNase. U3 RNPs were selected from HeLa cell extracts with antibodies bound to protein A-Sepharose beads, digested with RNase, and then washed extensively. Because U3 RNA itself is not recognized by these antibodies, any fragments retained must be interacting with one or more proteins. Furthermore, only regions bound by the 74-, 36-, 30-, or 12.5-kDa proteins were expected since the 59- and 13-kDa proteins were no longer immunoprecipitated after treatment with RNase (see above).

Figure 5A shows the RNA fragments immunoprecipitated after digestion of the U3 RNP with pancreatic RNase (pyrimidine specific). To demonstrate that all fragments could be kinased at their 5′ ends in vitro (lanes 8 through 13), a parallel experiment with RNA labeled in vivo (lanes 1 through 7) was done. Lanes 1 and 8 show controls that did not receive RNase; as expected, intact U3 did not appear in lane 8 because the m_3G cap structure blocked the 5′ kinase reaction. Anti-U3 RNP antibodies did not protect (lane 3) or immunoprecipitate (lanes 2 and 9) large RNA fragments in the absence of U3 proteins, and U3 fragments were not immunoprecipitated by sera with other specificities (not shown). The same protection pattern as seen in lanes 4 through 7 and 10 through 13 was obtained with all five anti-U3 RNP antibodies including the monoclonal (not

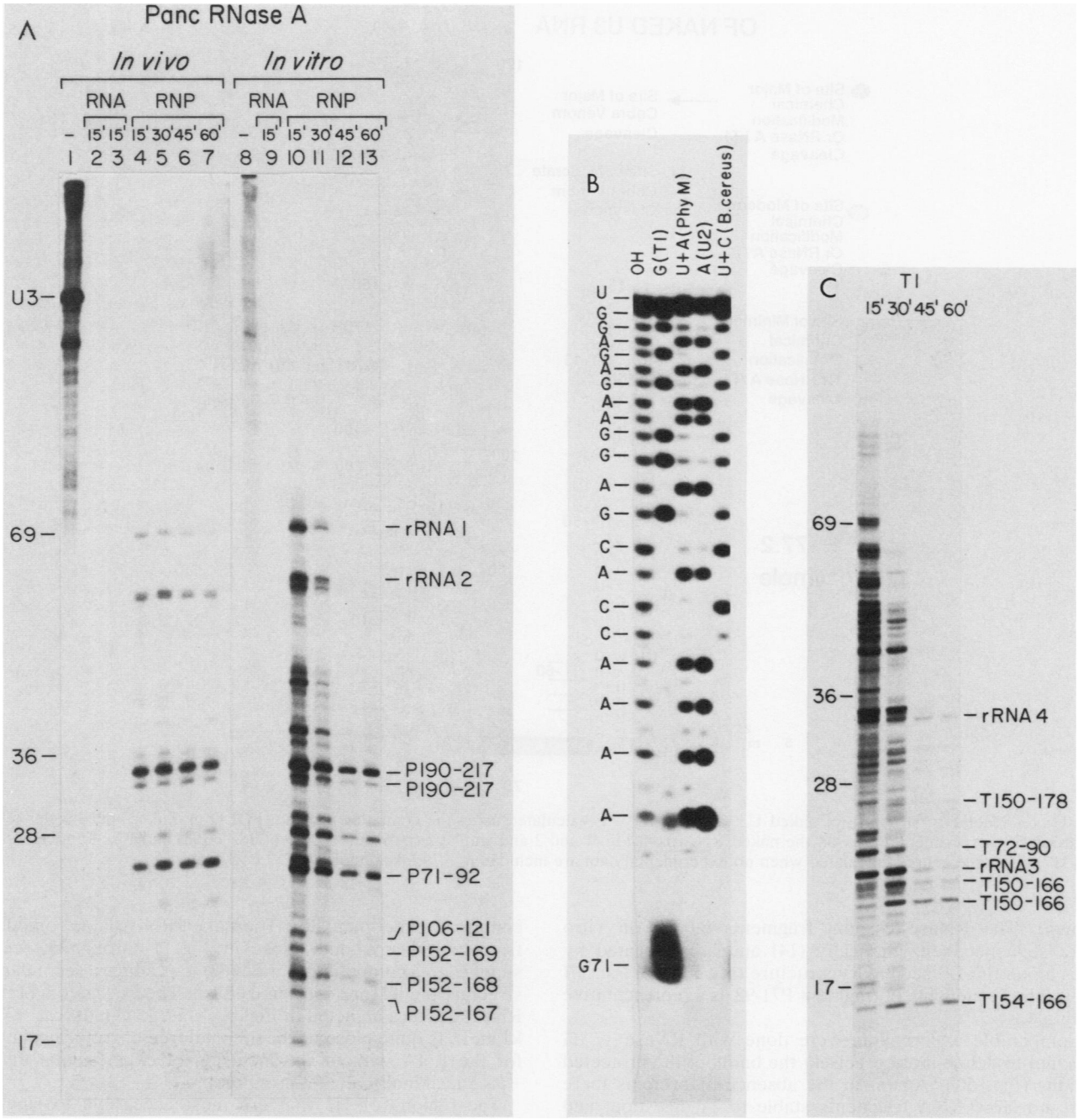

FIG. 5. U3 fragments protected from pancreatic or T_1 RNase by proteins. (A) U3 RNPs were immunoprecipitated, digested with pancreatic RNase at 37°C for the indicated time, and washed as described, and the extracted RNA fragments were analyzed on a 12% polyacrylamide gel. Lanes 1 through 7, RNAs labeled in vivo with $^{32}PO_4$ (48-h autoradiograph); lanes 8 through 13, RNAs 5′ end labeled in vitro with polynucleotide kinase after RNase digestion (30-min autoradiograph). Lanes 1 and 8, Immunoprecipitated U3 which did not receive any enzyme. For lanes 2, 3, and 9, samples were immunoprecipitated, deproteinized with phenol, added back to the same amount of antiserum and protein A-Sepharose as was used for the initial immunoprecipitation, and then digested with pancreatic RNase for 15 min; lanes 2 and 9 were subsequently washed like the other samples, while lane 3 was phenol extracted and ethanol precipitated immediately after digestion. U3 fragments are labeled on the right indicating the included nucleotides; the prefix P indicates that they were produced by pancreatic RNase digestion. The band migrating between fragments P71-92 and P190-217 contained at least two different components, neither of which appeared to be derived from U3; the bands labeled rRNA1 and rRNA2 contain nucleotides 2541-2602 and 2617-2660 of 28S rRNA (20), respectively. The location of DNA markers (filled-in *Hpa*II digest of pBR322) is marked on the left. (B) Fragment P71-92 was analyzed by enzymatic sequencing (14) and electrophoresed on a 20% polyacrylamide gel. (C) U3 RNPs were treated as for panel A except that T_1 RNase at 25°C was used and fragments labeled in vitro were fractionated on a 20% polyacrylamide gel. The identity of the protected fragments is again indicated on the right; the prefix T indicates that they were produced by T_1 digestion. The bands labeled rRNA3 and rRNA4 contain nucleotides 2623-2639 and 255-283 of 28S rRNA (20), respectively. DNA markers were as for panel A.

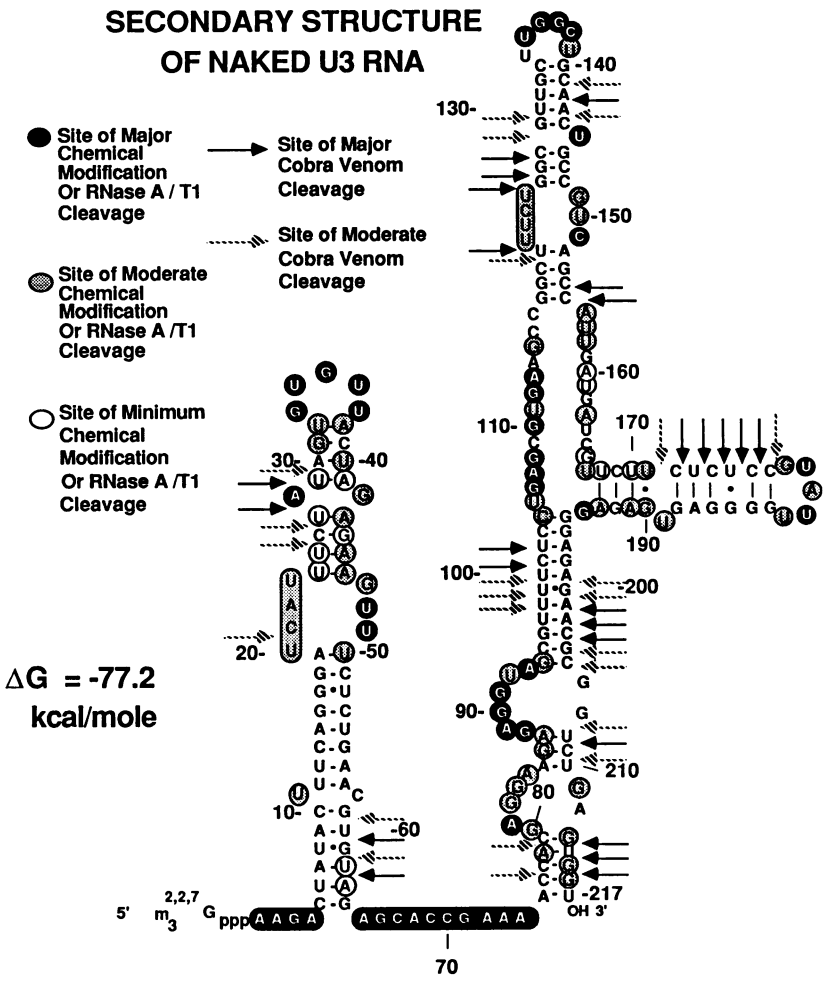

FIG. 6. Secondary structure of naked U3. A structure was calculated using the computer program FOLD or Zuker and Stiegler (68), incorporating experimental data on the naked RNA from Fig. 1 and 2 and similar experiments (nucleotides 1-194) and Fig. 3 (nucleotides 157-217). The free energy calculated when no experimental data are included is −79.5 kcal/mol (ca. 332.6 kJ/mol).

shown). The RNase-resistant fragments labeled in vitro were sequenced enzymatically (14) and are indicated by lines along the U3 secondary structure (see Fig. 7); Fig. 5B shows the sequencing of fragment P71-92 as a representative example.

Comparable experiments were done with RNase T_1 (G specific) to define more precisely the limits of the protected regions (Fig. 5C). Again, in the absence of proteins there were no large RNA fragments stable to T_1 digestion, and samples labeled in vivo gave the same pattern of protected bands as samples labeled in vitro (not shown).

In addition to the fragments derived from the U3 RNA, there were some minor bands whose sequences were identical to various regions of 28S rRNA (20; see legend, Fig. 5). We deduce that these were nonspecific contaminants because they were also obtained when RNase protection experiments were performed with a number of different serum specificities, including antisera to Sm, U1 RNP, U2 RNP, Ro/La, 5S, and 7-2S/8-2S RNP (not shown).

DISCUSSION

Secondary structure of U3. The secondary structures of deproteinized U3 RNA and of U3 in its RNP form have been compared experimentally. The naked-RNA data, compiled from experiments like those in Fig. 1 through 3, were supplied to a computer program (68) to produce a secondary-structure model for deproteinized U3. The free energy of the RNA structure depicted in Fig. 6, −77.2 kcal/mol (ca. −323 kJ/mol), is quite close to the minimal free energy calculated for the RNA without the inclusion of experimental data, −79.5 kcal/mol (ca. −332.6 kJ/mol).

The regions of U3 that are most accessible to single-stranded probes in the RNP include nucleotides A65-A74 (Fig. 1A and 1B) and C104-G112 (the failure to identify C109 as a modified base may be due to the lower efficiency of the DMS reaction with cytidines [41]). A third region, G80-A93, may be single stranded since all positions except A85 are moderately or strongly modified. Yet, G86 and A87 (Fig. 1A and B) are less strongly modified than the surrounding bases in both the RNA and the RNP. Furthermore, the naked RNA is cleaved at the complementary nucleotides, U208-U210, by cobra venom but not by pancreatic RNase (Fig. 3), and nucleotides 190-217 are highly resistant to pancreatic RNase digestion in the RNP (Fig. 5A and 7). Thus, A85-A87 and U208-U210 may form a short, but rather unstable stem.

Short loops at the ends of stems and internal stretches of ≤4 nucleotides are the only other regions in the U3 RNP that

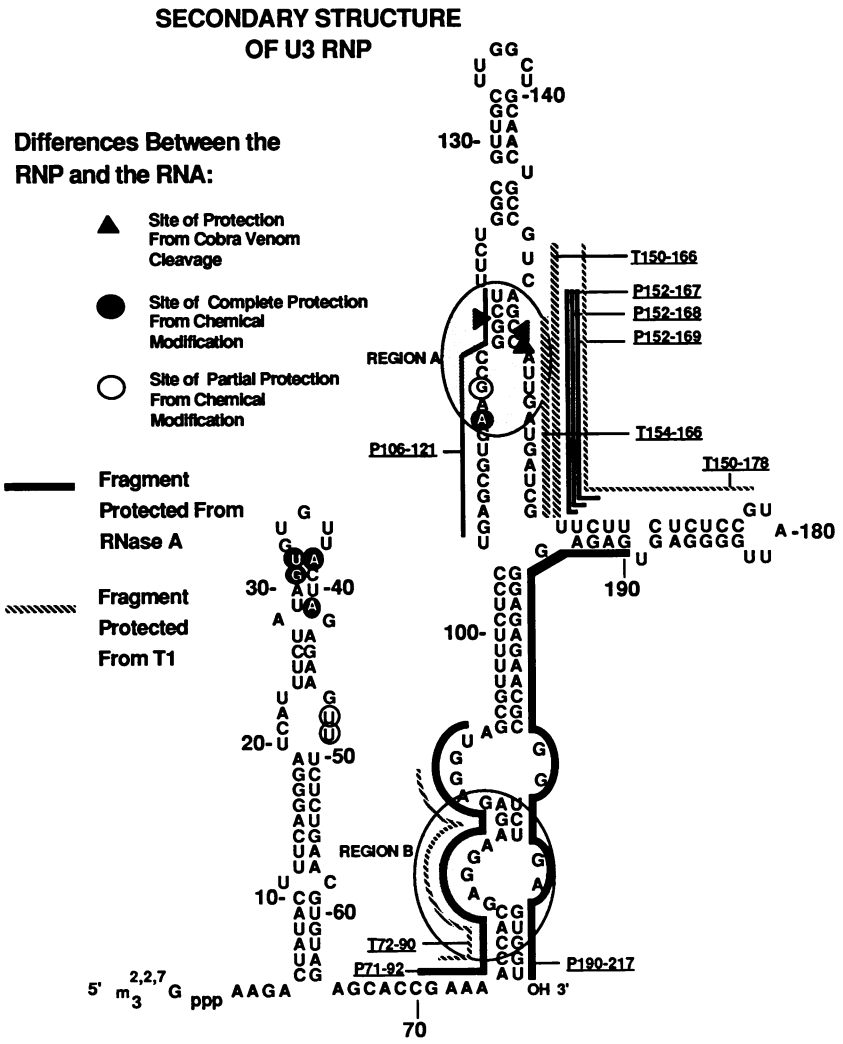

FIG. 7. Secondary structure of U3 in the RNP. Sites of differences in susceptibility to chemical modification or cobra venom cleavage between the RNA and RNP are indicated (from Fig. 1 and 2 and similar experiments). RNA fragments that are immunoprecipitated with U3 RNP-specific antibodies after RNase digestion (from Fig. 5) are depicted as a line parallel to the nucleotides each contains. The prefix P or T denotes fragments created with pancreatic or T_1 RNase, respectively. The thickness of each line is roughly proportional to the relative representation of the fragment.

appear to be single stranded. The minimal to moderate modification (or single-strand-specific enzyme cleavage) of some bases depicted as double stranded in Fig. 6 can most likely be ascribed to "breathing" of slightly unstable stems. This phenomenon, which is most frequent in regions rich in G:U or A:U base pairs, has been noted by others using this technique (27, 41). The absence of cobra venom cleavage at some bases that are clearly double stranded has been reported for other RNAs (27, 42, 53). Additional evidence for the stem at nucleotides A75-C79/G213-U217 was previously provided by Bernstein et al. (5), who showed that naked U3 alone can prime the synthesis of a 74-nucleotide cDNA by reverse transcriptase, implying that the sequence at the 3' end must base pair with the region downstream of and including A75.

Nucleotides A156-U167 (which include the region of U3 postulated to base pair with ITS-II sequences) are depicted in Fig. 6 and 7 to be single stranded simply because they have no nearby complement. The middle nucleotides, however, showed little or no modification above a background of strong reverse transcriptase stops in experiments done on gel-purified U3, naked U3 in a crude RNA fraction, or U3 in its RNP form (see Fig. 2, 6, and 7; data not shown). The background of strong stops does not simply reflect a highly accessible region that is easily nicked, since there are no bands corresponding to G159-G166 in Fig. 3 (see the top set of labeled nucleotides), where gel-purified U3 has been cleaved with T_1 or pancreatic RNase. Instead, the situation may be similar to several loops of 16S rRNA, for which Moazed et al. (41) have postulated that protected bases (which frequently cause naturally occurring reverse transcriptase stops) flanked by modified bases are involved in tertiary structure. It is possible that this region base pairs with distant nucleotides (e.g., A160-A163 is perfectly complementary to U20-U23). An attempt to identify tertiary interactions of the sort determined for 5S rRNA (6), using naked U3 RNA and UV light, was unsuccessful (H. Robertson, personal communication).

We considered the possibility that a region of U3 was protected from modification because it was base paired with another RNA during the chemical reaction, but we felt it to be unlikely for the following reasons. First, although 65% of

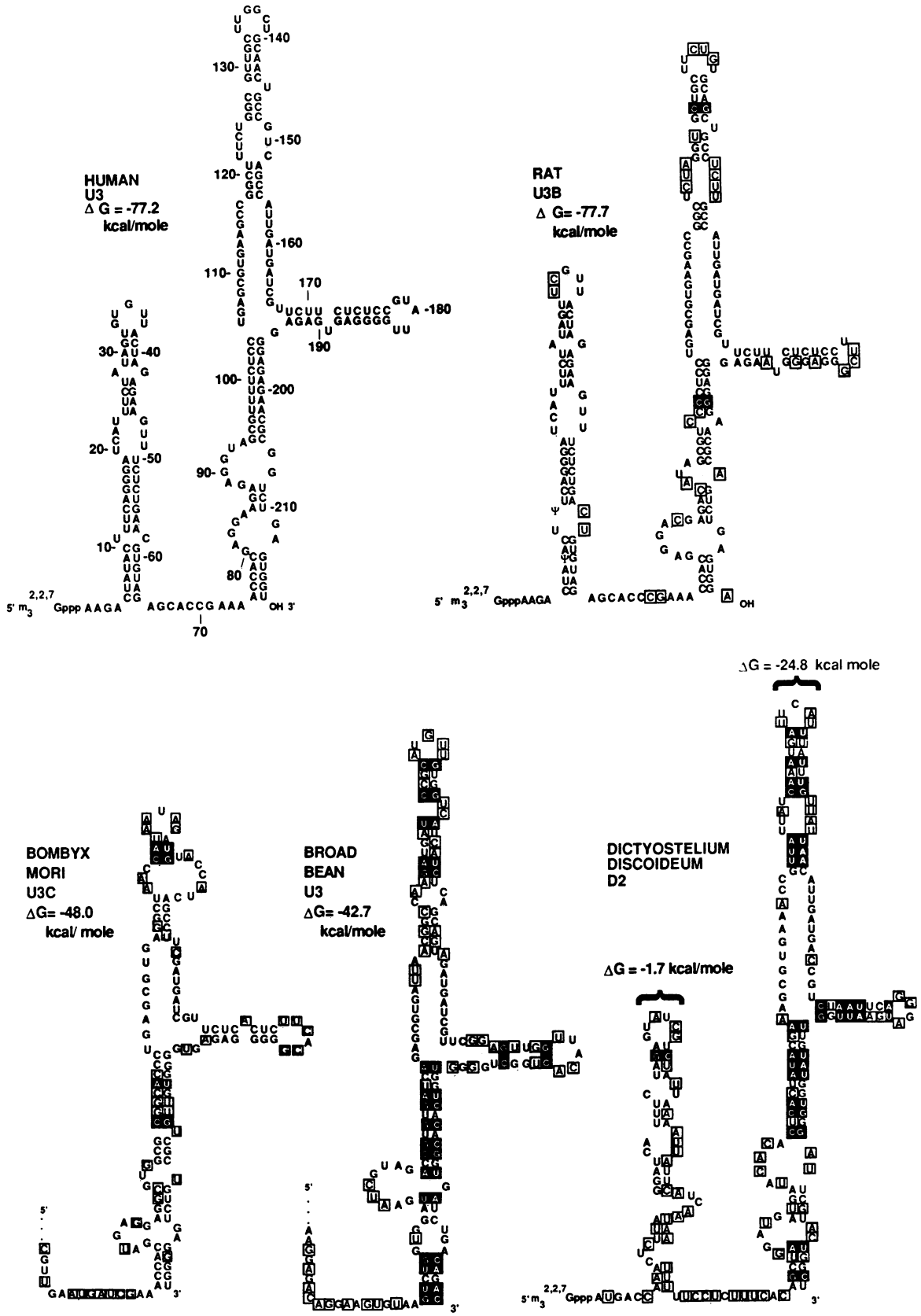

TABLE 1. U3 sequences evaluated for complementarity with pre-rRNA[a]

U3 source (reference)	65–74 region	80–93 region	104–115 region
Human (59)	-AGCACCGAAA-	GAGGAAGA-GAGGUA	CUGAGCGUGAAG
Rat U3B (47, 59)	-AGCACCcgAA-	GAGGAcGA-GAcaUA	CUGAGCGUGAAG
X. laevis[b]	cuGaACucAcAA	GAGGAAGA-GcGucA	CUGAGCGUGAAG
D2 (65)	ccuCuuucAcA-	GAuGAuGA-uAcacA	gaaAGCGUGAAa
S. cerevisiae snR17A[c]	GAGCcaCugAAu	uuGGuuGAuGAGuc	CccAGaGUGAga

[a] Capital letters are nucleotides conserved with the human U3 at that position; lowercase letters are nonconserved.
[b] Jeppesen et al., personal communication.
[c] Hughes et al., in press.

U3 has been reported to be associated with a >60S particle before deproteinization of nucleolar extracts (16), we observed that 90 to 95% of the U3 in our whole-cell extracts was in a 10-11S particle (not shown) and therefore cannot be base paired with pre-rRNA. Second, even though a significant fraction of U3 can be found associated with a 28-35S RNA after deproteinization of nucleolar RNA (45, 67), modification experiments with gel-purified U3 RNA gave virtually the same results as experiments with naked U3 present in a crude RNA fraction (not shown).

Oligodeoxynucleotide-directed cleavage by RNase H (specific for the RNA strand of an RNA-DNA hybrid) supports the secondary structure model in Fig. 6 and 7. Oligonucleotides complementary to U3 nucleotides 63-77, 80-94, or 104-120 led to greater than 50% cleavage of U3 in the RNP, whereas those complementary to nucleotides 1-15, 17-32, 151-167, or 203-217 had no effect (not shown).

Phylogenetic evidence for the proposed secondary structure. Phylogenetic comparison is a powerful method for verifying predicted RNA secondary structures. Figure 8 shows that the U3 RNA from several species can be folded into a secondary structure similar to that determined experimentally for human U3. Two independent sets of compensating base changes are considered to be proof of the existence of an RNA helix (42). Compensatory changes that maintain standard Watson-Crick base pairing are boxed in black in Fig. 8, and single changes (many of which also maintain stems) are indicated by unfilled boxes. The inclusion of two A:G base pairs within these stems is consistent with data on similar noncanonical base pairs found in tRNA and rRNA (reviewed by Noller [42]) and with nuclear magnetic resonance studies of oligodeoxynucleotides (28). Our proposed secondary structure for the 3′ domain (nucleotides 75-217) is very well supported by phylogenetic evidence, while that of the 5′ domain (nucleotides 1-64) must be considered more tentative.

One of the two regions (C104-G112) of human U3 that we determined to be highly accessible to single-strand-specific probes is extraordinarily conserved, while the other (A65-A74) is remarkably divergent. The sequence we suspect is involved in tertiary structure, A160-A163, is absolutely conserved across species, as is its potential complement, U20-U23.

The U3 RNP contains at least six proteins. By immunoprecipitation from HeLa cell extracts we have identified four nonphosphorylated proteins of 36, 30, 13, and 12.5 kDa and two phosphorylated proteins of 74 and 59 kDa which appear to be constituents of the U3 RNP. Although considerably fainter, the lower-molecular-weight bands do not seem to represent degradation products of the 36-kDa protein because they appeared after the cells had been labeled with [^{35}S]methionine for only 6 h, whereas the 36-kDa protein could not be detected until cells had been labeled for 20 h (not shown). We believe instead that the unequal intensity is due to dissociation of the 36-kDa protein (which carries the antigenic determinants) from the rest of the U3 RNP. In our extracts, only 10 to 20% of the U3 RNA was immunoprecipitable after the 3 min of sonication necessary to release the particle from the nucleolus. Furthermore, 10 min of sonication, high-salt washes (400 mM NaCl), heparin washes (0.5 mg/ml), or incubation at 35°C for 60 min further decreased the immunoprecipitability to <1% of the total U3 in the extract.

Protein-RNA interactions. Although some sites of interaction between proteins and U3 RNA could have been disrupted during the preparation of extract, others were clearly well preserved. Two regions of interaction between the proteins closely associated with the antigenic 36-kDa polypeptide (74, 30, and 12.5 kDa; see Fig. 4) and the RNA (see Fig. 7) were clustered in two loosely delimited regions (A and B), both located in the 3′ domain of U3.

In region A, the protein(s) may interact with both single- and double-stranded RNA since A113 (Fig. 1A) and G115 (Fig. 1B) were protected from chemical modification in the RNP but not in the naked RNA, and the stem between G118-U121/A152-C155 contains some bases that were protected from cobra venom cleavage only in the RNP (Fig. 2). The strong protection of nucleotides U157-G166 from RNases may be due simply to higher order structure instead of to a direct protein-RNA interaction, since these bases were protected in the naked RNA as well (see Fig. 3).

Region B contains no nucleotides that were protected from cobra venom cleavage or chemical modification in the RNP as compared with the RNA (with the caveat that we have no RNP data for the nucleotides at the 3′ end of U3). It is possible that the protein(s) binds the RNA backbone and

FIG. 8. Phylogenetic comparison of U3 secondary structures. The free energy values of most of these secondary structures were obtained with the FOLD program of Zuker and Steigler (68) by mandating that certain bases be single stranded (determined by alignment with single-stranded regions in human U3); the 5′ domain of *D. discoideum* was modified slightly from results of Bernstein et al. (5). In the case of *Bombyx mori* (1) and broad bean plant (30), only incomplete sequences are available. We have not included a *Drosophila* sequence (2) which is not yet known to be expressed. The theoretically most stable structure for each sequence was also determined with the FOLD program by allowing all nucleotides the possibility of base pairing. They have the following free energies: rat U3B (47, 57), $\Delta G = -83.4$ kcal/mol (ca. 348.9 kJ/mol); *Bombyx mori* U3C (1) (nucleotides 52-179), $\Delta G = -50.6$ kcal/mol (ca. 211.7 kJ/mol); broad bean plant U3 (30) (nucleotides 155-1), $\Delta G = -45.2$ kcal/mol (ca. 189.1 kJ/mol); and *Dictyostelium* D2 (65) (5′ domain, nucleotides 1-63), $\Delta G = -15.4$ kcal/mol (ca. 64.4 kJ/mol) and (3′ domain, nucleotides 75-210) $\Delta G = -32.8$ kcal/mol (ca. 137.2 kJ/mol). All differences from the human sequence (determined by the GAP program), including insertions, are boxed; double mutations that maintain Watson-Crick base pairing in a stem are boxed in black.

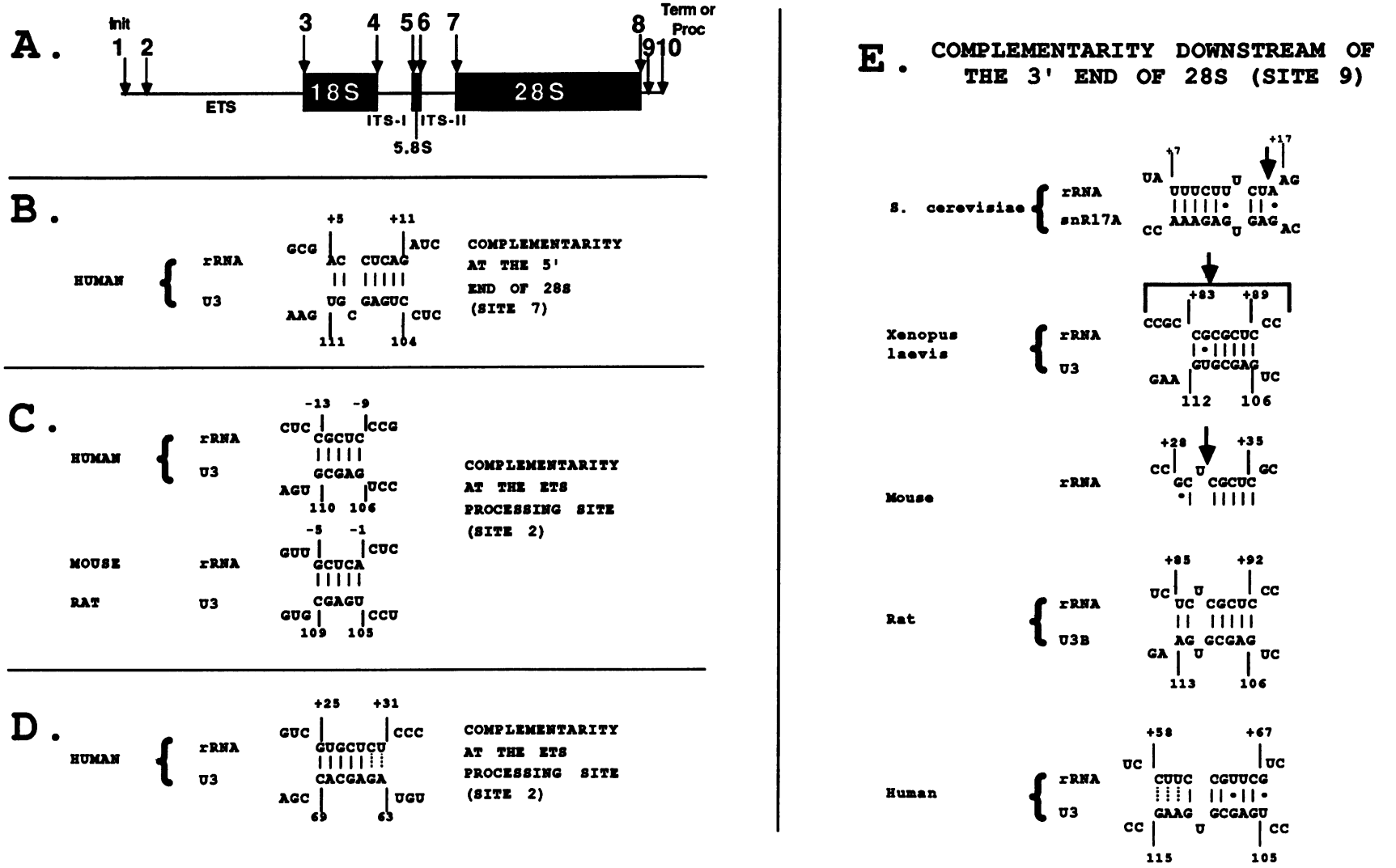

FIG. 9. Complementarity between U3 and pre-rRNA sequences. (A) Sites of initiation, termination, and processing of pre-rRNA are depicted in relation to the final products, 18S, 5.8S, and 28S, on an approximate scale. The sites shown include: (1) initiation of transcription; (2) a processing site several hundred nucleotides downstream of the initiation site; (3 through 8) processing sites at the ends of mature rRNAs; (9) a processing site 15 to 100 nucleotides downstream of the 3′ end of 28S; and (10) termination (or possibly processing [32]) site 210 to 565 nucleotides downstream of the 3′ end of 28S rRNA. (B through E) Complementarity between U3 and sequences near processing sites. Pre-rRNA sequences are numbered according to the nucleotide number upstream (−) or downstream (+) of that particular processing site. U3 sequences are numbered according to the distance from the 5′ end. Base pairing for the mouse pre-rRNA sequences assumes that U3 sequences are conserved between rat and mouse sequences. Solid lines indicate standard Watson-Crick base pairs; dots indicate G:U or A:G pairs; dotted lines indicate base pairs involving U3 nucleotides that are not highly available in the RNP (Fig. 7). (E) Cleavage sites in mouse and *S. cerevisiae* RNA have been precisely mapped and are marked by a large arrow. In *X. laevis*, an approximate cleavage site is identified by a bracket centered at the estimated site.

allows small-molecular-weight probes to reach and modify the base-pairing moieties even in the RNP. Conversely, an interaction between the protein(s) and base-pairing moieties may have been disrupted in most of the RNPs during the preparation of extract and thus not be detectable.

A third region, nucleotides 31-49, contains bases that were protected from chemical modification in the RNP but not the RNA (Fig. 1C and 7). The protection of bases in the putative stem between U29-U32/A38-A41 could reflect stabilization of the stem by a protein(s), as has been hypothesized to occur with 16S rRNA in ribosomes (41). The failure to immunoprecipitate protected RNA fragments from this region may simply be due to separation of the 5′ domain from the 3′ domain (where the antigenic 36-kDa protein appears to reside) during the RNase digestion. The 59- and 13-kDa proteins may interact with the RNA in this region, as they were not immunoprecipitated with the rest of the U3 RNP proteins after treatment with RNases.

Possible functions of the U3 RNP. The idea that G159-U168 of U3 RNA might base pair with the ITS-II at its junction with 5.8S (3, 60) originally seemed reasonable in view of the extensive complementarity between the U3 and ITS-II sequences known at the time and the predicted accessibility of that region in U3 RNA. Our data, however, indicate that these nucleotides are not highly available to chemical modification, to single-strand-specific nucleases, or to RNase H cleavage in the presence of a complementary oligonucleotide. In addition, although this region of U3 is highly conserved across species (including *D. discoideum* [65] and *S. cerevisiae* [J. Hughes et al., in press]), the sequence at the 5′ end of the ITS-II diverges considerably in these lower eucaryotes (44, 63).

Alternative functions for U3 were formulated based on three assumptions: (i) that U3 acts near a known initiation, termination, or processing site of pre-rRNA; (ii) that the interaction involves base pairing between experimentally determined single-stranded regions of U3 and pre-rRNA sequences; and (iii) that significant complementarity is conserved across species. The highly accessible sequences (A65-A74 and C104-G112) and the partially accessible region (G80-A93) of human U3 RNA were searched for ≥5 nucleotides of complementarity to human pre-rRNA sequences (see Materials and Methods). The analogous sequences of rat U3B, *X. laevis* U3, *D. discoideum* D2, and *S. cerevisiae* snR17A are shown in Table 1. Each was compared with pre-rRNA sequences of the same species (see Methods) except that rat U3B was compared with mouse pre-rRNA.

Figure 9 shows the processing sites considered and the most compelling complementarities with U3. A 5- to 7-nucleotide complementarity between U3 nucleotides C104-U111 and a mature rRNA sequence located 5 nucleotides downstream of the 5′ end of 28S rRNA (Fig. 9B) is conserved in all species, but most secondary-structure models for eucaryotic large rRNAs predict that these 28S rRNA nucleotides are stably base paired with the 3′ end of 5.8S rRNA (reviewed by Hadjiolov [24]). Two short complementarities between U3 and pre-rRNA sequences are found near site 2, which is a processing site several hundred nucleotides downstream of the transcription initiation site (12, 39, 62) (Fig. 9C and D), but conservation of this potential interaction cannot be judged, as this processing site has not been identified in any of the other species for which U3 sequences are available.

The best complementarity found is between the highly conserved, single-stranded sequence at nucleotides 104-112 of U3 and sequences near processing site 9 in pre-rRNA (Fig. 9E). Recent experiments with mouse (21, 23), *X. laevis* (32), and *S. cerevisiae* (29) pre-RNAs show that RNA polymerase I transcription continues at least a few hundred nucleotides past the 3′ end of 28S rRNA and that formation of mature 3′ ends requires one or more processing reactions. In the mouse pre-rRNA, processing site 9 is 30 nucleotides downstream of the 3′ end of 28S rRNA (22, 31), within a sequence that is complementary to nucleotides 106-113 of rat U3 (see Fig. 9B). Yeast 26S has a well-documented precursor with 15 additional nucleotides at the 3′ end (4, 29); adjacent sequences are complementary to the homologous region of snR17A, which is the yeast U3 homolog. Although this processing site has not been precisely mapped in any other species, pre-rRNA sequences which could base pair with the homologous U3 sequences are within 100 nucleotides of the 3′ end of 28S rRNA in *X. laevis*, rat, and human genomes (as shown in Fig. 9B). In *X. laevis*, RNase protection experiments suggest a processing event in the region downstream of the 3′ end of 28S rRNA (32), and primer extension analysis shows a band mapping to +80 to +100 (P. Labhart, personal communication).

In summary, complementarity between single-stranded regions of U3 and pre-rRNA sequences suggests that U3 could be involved in a processing event near the 3′ end of mature 28S rRNA. Studies to identify such an interaction are under way.

ACKNOWLEDGMENTS

We thank Manny Ares, Doug Black, Jon Glickman, David Solnick, Ilana Stroke, and other members of the Steitz lab for helpful advice and discussions. We especially thank J. Hughes; D. Suh, H. Busch, and R. Reddy; C. Jeppesen, M. Ares, and S. Gerbi; J. Gorski and R. Schmickel; P. Labhart; and H. Robertson for sharing unpublished results. We are most grateful to Harry Noller for advice on the secondary structure analysis and phylogenetic models. We also thank G. Reimer and E. Tan, J. Craft and J. Hardin, T. Mimori and M. Homma, and R. Sontheimer for generously providing antibodies. Perusal of the manuscript by Ellen Gottlieb, Volker Gerke, and Kim Mowry is appreciated.

K.A.P. is a Physician Scientist Award (National Institutes of Health) recipient. This work was supported in part by Public Health Service grant GM26154 from the National Institutes of Health.

LITERATURE CITED

1. **Adams, D. S., R. J. Herrera, R. Lührmann, and P. M. Lizardi.** 1985. Isolation and partial characterization of U1-U6 small RNAs from *Bombyx mori*. Biochemistry **24**:117–125.
2. **Akao, M., R. Reddy, and H. Busch.** 1986. Multiple sequences in the *Drosophila melanogaster* U3 RNA gene are homologous to vertebrate U3 RNA. Biochem. Biophys. Res. Commun. **138**: 512–518.
3. **Bachellerie, J.-P., B. Michot, and F. Raynal.** 1983. Recognition signals for mouse pre-rRNA processing. Mol. Biol. Rep. **9**:79–86.
4. **Bayev, A. A., O. I. Georgiev, A. A. Hadjiolov, N. Nikolaev, K. G. Skryabin, and V. M. Zakharyev.** 1981. The structure of the yeast ribosomal RNA genes. 3. Precise mapping of the 18S and 25S rRNA genes and structure of the adjacent regions. Nucleic Acids Res. **9**:789–799.
5. **Bernstein, L. B., S. M. Mount, and A. M. Weiner.** 1983. Pseudogenes for human small nuclear RNA U3 appear to arise by integration of self-primed reverse transcripts of the RNA into new chromosomal sites. Cell **32**:461–472.
6. **Branch, A. D., B. J. Benenfeld, and H. D. Robertson.** 1985. Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa 5S RNA. Proc. Natl. Acad. Sci. USA **82**:6590–6594.

7. **Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi.** 1982. SnRNAs, snRNPs, and RNA processing. Annu. Rev. Biochem. **51**:617–654.
8. **Calvet, J. P., and T. Pederson.** 1981. Base-pairing interactions between small nuclear RNAs and nuclear RNA precursors as revealed by psoralen cross-linking *in vivo*. Cell **26**:363–370.
9. **Cassidy, B. G., C. S. Subrahmanyam, and L. I. Rothblum.** 1982. The nucleotide sequence of the 5′ region of rat 18S rDNA and adjoining spacer. Biochem. Biophys. Res. Commun. **107**: 1571–1576.
10. **Chan, Y.-L., R. Gutell, H. F. Noller, and I. G. Wool.** 1984. The nucleotide sequence of a rat 18S ribosomal ribonucleic acid gene and a proposal for the secondary structure of 18S ribosomal ribonucleic acid. J. Biol. Chem. **259**:224–230.
11. **Chan, Y.-L., J. Olvera, and I. G. Wool.** 1983. The structure of rat 28S ribosomal ribonucleic acid inferred from the sequence of nucleotides in a gene. Nucleic Acids Res. **11**:7819–7831.
12. **Craig, N., S. Kass, and B. Sollner-Webb.** 1987. Nucleotide sequence determining the first cleavage site in the processing of mouse precursor rRNA. Proc. Natl. Acad. Sci. USA **84**: 629–633.
13. **Crouch, R. J., S. Kanaya, and P. L. Earl.** 1983. A model for the involvement of the small nucleolar RNA (U3) in processing eukaryotic ribosomal RNA. Mol. Biol. Rep. **9**:75–78.
14. **Donis-Keller, H., A. M. Maxam, and W. Gilbert.** 1977. Mapping adenines, guanines, and pyrimidines in RNA. Nucleic Acids Res. **4**:2527–2538.
15. **England, T. E., A. G. Bruce, and O. C. Uhlenbeck.** 1980. Specific labeling of 3′ termini of RNA with T4 RNA ligase. Methods Enzymol. **65**:65–74.
16. **Epstein, P., R. Reddy, and H. Busch.** 1984. Multiple states of U3 RNA in Novikoff hepatoma nucleoli. Biochemistry **23**: 5421–5425.
17. **Financsek, I., K. Mizumoto, Y. Mishima, and M. Muramatsu.** 1982. Human ribosomal RNA gene: nucleotide sequence of the transcription initiation region and comparison of three mammalian genes. Proc. Natl. Acad. Sci. USA **79**:3092–3096.
18. **Georgiev, O. I., N. Nikolaev, and A. A. Hadjiolov.** 1981. The structure of the yeast ribosomal RNA genes. 4. Complete sequence of the 25S rRNA gene from Saccharomyces cerevisiae. Nucleic Acids Res. **9**:6953–6958.
19. **Goldman, W. E., G. Goldberg, L. H. Bowman, D. Steinmetz, and D. Schlessinger.** 1983. Mouse rDNA: sequences and evolutionary analysis of spacer and mature RNA regions. Mol. Cell. Biol. **3**:1488–1500.
20. **Gonzalez, I. L., J. L. Gorski, T. J. Campen, D. J. Dorney, J. M. Erickson, J. E. Sylvester, and R. D. Schmickel.** 1985. Variation among human 28S ribosomal RNA genes. Proc. Natl. Acad. Sci. USA **82**:7666–7670.
21. **Grummt, I., U. Maier, A. Öhrlein, N. Hassouna, and J.-P. Bachellerie.** 1985. Transcription of mouse rDNA terminates downstream of the 3′ end of 28S RNA and involves interaction of factors with repeated sequences in the 3′ spacer. Cell **43**:801–810.
22. **Grummt, I., H. Sorbaz, A. Hofmann, and E. Roth.** 1985. Spacer sequences downstream of the 28S RNA coding region are part of the mouse rDNA transcription unit. Nucleic Acids Res. **13**:2293–2304.
23. **Gurney, T., Jr.** 1985. Characterization of mouse 45S ribosomal RNA subspecies suggests that the first processing cleavage occurs 600 ± 100 nucleotides from the 5′ end and the second 500 ± 100 nucleotides from the 3′ end of a 13.9 kb precursor. Nucleic Acids Res. **13**:4905–4919.
24. **Hadjiolov, A. A.** 1985. The nucleolus and ribosome biogenesis. Springer-Verlag, New York.
25. **Hall, L. M. C., and B. E. H. Maden.** 1980. Nucleotide sequence through the 18S-28S intergene region of a vertebrate ribosomal transcription unit. Nucleic Acids Res. **8**:5993–6005.
26. **Hashimoto, C., and J. A. Steitz.** 1983. Sequential association of nucleolar 7-2 RNA with two different autoantigens. J. Biol. Chem. **258**:1379–1382.
27. **Inoue, T., and T. R. Cech.** 1985. Secondary structure of the circular form of the Tetrahymena rRNA intervening sequence: a technique for RNA structure analysis using chemical probes and reverse transcriptase. Proc. Natl. Acad. Sci. USA **82**:645–652.
28. **Kan, L.-S., S. Chandrasegaren, S. M. Pulford, and P. S. Miller.** 1983. Detection of a guanine · adenine base pair in a decadeoxyribonucleotide by proton magnetic resonance spectroscopy. Proc. Natl. Acad. Sci. USA **80**:4263–4265.
29. **Kempers-Veenstra, A. E., J. Oliemans, H. Offenberg, A. F. Dekker, P. W. Piper, R. J. Planta, and J. Klootwijk.** 1986. 3′-End formation of transcripts from the yeast rRNA operon. EMBO J. **5**:2703–2710.
30. **Kiss, T., M. Tóth, and F. Solymosy.** 1985. Plant small nuclear RNAs: nucleolar U3 snRNA is present in plants: partial characterization. Eur. J. Biochem. **152**:259–266.
31. **Kominami, R., Y. Mishima, Y. Urano, M. Sakai, and M. Muramatsu.** 1982. Cloning and determination of the transcription termination site of ribosomal RNA gene of the mouse. Nucleic Acids Res. **10**:1963–1979.
32. **Labhart, P., and R. H. Reeder.** 1986. Characterization of three sites of RNA 3′ end formation in the *Xenopus* ribosomal gene spacer. Cell **45**:431–443.
33. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
34. **Lischwe, M. A., R. L. Ochs, R. Reddy, R. G. Cook, L. C. Yeoman, E. M. Tan, M. Reichlin, and H. Busch.** 1985. Purification and partial characterization of a nucleolar scleroderma antigen (M_r = 34,000; pI, 8.5) rich in N^G, N^G-dimethylarginine. J. Biol. Chem. **260**:14304–14310.
35. **Maden, B. E. H., M. Moss, and M. Salim.** 1982. Nucleotide sequence of an external transcribed spacer in Xenopus laevis rDNA: sequences flanking the 5′ and 3′ ends of 18S rRNA are non-complementary. Nucleic Acids Res. **10**:2387–2398.
36. **Michot, B., J.-P. Bachellerie, and F. Raynal.** 1982. Sequence and secondary structure of mouse 28S rRNA 5′ terminal domain. Organization of the 5.8S-28S rRNA complex. Nucleic Acids Res. **10**:5273–5282.
37. **Michot, B., J.-P. Bachellerie, and F. Raynal.** 1983. Structure of mouse rRNA precursors. Complete sequence and potential folding of the spacer regions between 18S and 28S rRNA. Nucleic Acids Res. **11**:3375–3391.
38. **Michot, B., J.-P. Bachellerie, F. Raynal, and M.-H. Renalier.** 1982. Sequence of the 3′ terminal domain of mouse 18S rRNA: conservation of structural features with other pro- and eukaryotic homologs. FEBS Lett. **142**:260–266.
39. **Miller, K. G., and B. Sollner-Webb.** 1981. Transcription of mouse rRNA genes by RNA polymerase 1: *in vitro* and *in vivo* initiation and processing sites. Cell **27**:165–174.
40. **Mimori, T., M. Hinterberger, I. Pettersson, and J. A. Steitz.** 1984. Autoantibodies to the U2 small nuclear ribonucleoprotein in a patient with scleroderma-polymyositis overlap syndrome. J. Biol. Chem. **259**:560–565.
41. **Moazed, D., S. Stern, and H. F. Noller.** 1986. Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. J. Mol. Biol. **187**: 399–416.
42. **Noller, H. F.** 1984. Structure of ribosomal RNA. Annu. Rev. Biochem. **53**:119–162.
43. **Ochs, R. L., M. A. Lischwe, W. H. Spohn, and H. Busch.** 1985. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. Biol. Cell **54**:123–134.
44. **Ozaki, T., Y. Hoshikawa, Y. Iida, and M. Iwabuchi.** 1984. Sequence analysis of the transcribed and 5′ non-transcribed regions of the ribosomal RNA gene in *Dictyostelium discoideum*. Nucleic Acids Res. **12**:4171–4184.
45. **Prestayko, A. W., M. Tonato, and H. Busch.** 1970. Low molecular weight RNA associated with 28S nucleolar RNA. J. Mol. Biol. **47**:505–515.
46. **Raynal, F., B. Michot, and J.-P. Bachellerie.** 1984. Complete nucleotide sequence of mouse 18S rRNA gene: comparison with other available homlogs. FEBS Lett. **167**:263–268.
47. **Reddy, R., D. Henning, and H. Busch.** 1979. Nucleotide sequence of nucleolar U3B RNA. J. Biol. Chem. **254**:11097–11105.

48. **Reddy, R., E. M. Tan, D. Henning, K. Nohga, and H. Busch.** 1983. Detection of a nucleolar 7-2 ribonucleoprotein and a cytoplasmic 8-2 ribonucleoprotein with autoantibodies from patients with scleroderma. J. Biol. Chem. **258**:1383–1386.
49. **Rothblum, L. I., R. Reddy, and B. Cassidy.** 1982. Transcription initiation site of rat ribosomal DNA. Nucleic Acids Res. **10**:7345–7362.
50. **Rubtsov, P. M., M. M. Musakhanov, V. M. Zakharyev, A. S. Krayev, K. G. Skryabin, and A. A. Bayev.** 1980. The structure of the yeast ribosomal RNA genes. I. The complete nucleotide sequence of the 18S ribosomal RNA gene from Saccharomyces cerevisiae. Nucleic Acids Res. **8**:5779–5794.
51. **Salim, M., and B. E. H. Maden.** 1980. Nucleotide sequence encoding the 5′ end of Xenopus laevis 18S rRNA. Nucleic Acids Res. **8**:2871–2884.
52. **Salim, M., and B. E. H. Maden.** 1981. Nucleotide sequence of *Xenopus laevis* 18S ribosomal RNA inferred from gene sequence. Nature (London) **291**:205–208.
53. **Shelness, G. S., and D. L. Williams.** 1985. Secondary structure analysis of apolipoprotein II mRNA using enzymatic probes and reverse transcriptase. J. Biol. Chem. **260**:8637–8646.
54. **Skryabin, K. G., M. A. Eldarov, V. L. Larionov, A. A. Bayev, J. Klootwijk, V. C. H. F. de Regt, G. M. Veldman, R. J. Planta, O. I. Georgiev, and A. A. Hadjiolov.** 1984. Structure and function of the nontranscribed spacer regions of yeast rDNA. Nucleic Acids Res. **12**:2955–2968.
55. **Sollner-Webb, B., and R. H. Reeder.** 1979. The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in X. laevis. Cell **18**:485–499.
56. **Solnick, D.** 1985. Alternative splicing caused by RNA secondary structure. Cell **43**:667–676.
57. **Stroke, I. L., and A. M. Weiner.** 1985. Genes and pseudogenes for rat U3A and U3B small nuclear RNA. J. Mol. Biol. **184**:183–193.
58. **Subrahmanyam, C. S., B. Cassidy, H. Busch, and L. I. Rothblum.** 1982. Nucleotide sequence of the region between the 18S rRNA sequence and the 28S rRNA sequence of rat ribosomal DNA. Nucleic Acids Res. **10**:3667–3680.
59. **Suh, D., H. Busch, and R. Reddy.** 1986. Isolation and characterization of a human U3 small nucleolar RNA gene. Biochem. Biophys. Res. Commun. **137**:1133–1140.
60. **Tague, B. W., and S. A. Gerbi.** 1984. Processing of the large rRNA precursor: two proposed categories of RNA-RNA interactions in eukaryotes. J. Mol. Evol. **20**:362–367.
61. **Torzcynski, R. M., M. Fuke, and A. P. Bollon.** 1985. Cloning and sequencing of a human 18S ribosomal RNA gene. DNA **4**:283–291.
62. **Urano, Y., R. Kominami, Y. Mishima, and M. Muramatsu.** 1980. The nucleotide sequence of the putative transcription initiation site of a cloned ribosomal RNA gene of the mouse. Nucleic Acids Res. **8**:6043–6058.
63. **Veldman, G. M., J. Klootwijk, H. von Heerikhuizen, and R. J. Planta.** 1981. The nucleotide sequence of the intergenic region between the 5.8S and 26S rRNA genes of the yeast ribosomal RNA operon. Possible implications for the interaction between 5.8S and 26S rRNA and the processing of the primary transcript. Nucleic Acids Res. **9**:4847–4862.
64. **Weinberg, R. A., and S. Penman.** 1968. Small molecular weight monodisperse nuclear RNA. J. Mol. Biol. **38**:289–304.
65. **Wise, J. A., and A. M. Weiner.** 1980. Dictyostelium small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. Cell **22**:109–118.
66. **Yavachev, L. P., O. I. Georgiev, E. A. Braga, T. A. Avdonina, A. E. Bogomolova, V. B. Zhurkin, V. V. Nosikov, and A. A. Hadjiolov.** 1986. Nucleotide sequence analysis of the spacer regions flanking the rat rRNA transcription unit and identification of repetitive elements. Nucleic Acids Res. **14**:2799–2810.
67. **Zieve, G., and S. Penman.** 1976. Small RNA species of the HeLa cell: metabolism and subcellular localization. Cell **8**:19–31.
68. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.