MOLECULAR AND CELLULAR BIOLOGY, Mar. 1990, p. 939–946
0270-7306/90/030939-08$02.00/0

Vol. 10, No. 3

Capping of Mammalian U6 Small Nuclear RNA In Vitro Is Directed by a Conserved Stem-Loop and AUAUAC Sequence: Conversion of a Noncapped RNA into a Capped RNA

RAVINDER SINGH, SHASHI GUPTA, AND RAM REDDY*

Department of Pharmacology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

Received 5 October 1989/Accepted 4 December 1989

The cap structure of U6 small nuclear RNA (snRNA) is γ-monomethyl phosphate and is distinct from other known RNA cap structures (R. Singh and R. Reddy, Proc. Natl. Acad. Sci. USA 86:8280–8283, 1989). Here we show that the information for capping the U6 snRNA in vitro is within the initial 25 nucleotides of the U6 RNA. The capping determinant in mammalian U6 snRNA is a bipartite element—a phylogenetically conserved stem-loop structure and an AUAUAC sequence, or a part thereof, following this stem-loop. Wild-type capping efficiency was obtained when the AUAUAC motif immediately followed the stem-loop and when the γ-phosphate of the initiation nucleotide was in close proximity to the capping determinant. Incorporation of a synthetic stem-loop followed by an AUAUAC sequence is sufficient to covert a noncapped heterologous transcript into a capped transcript. Transcripts with the initial 32 nucleotides of *Saccharomyces cerevisiae* U6 snRNA are accurately capped in HeLa cell extract, indicating that capping machinery from HeLa cells can cap U6 snRNA from an evolutionarily distant eucaryote. The U6-snRNA-specific capping is unusual in that it is RNA sequence dependent, while the capping of mRNAs and other U snRNAs is tightly coupled to transcription and is independent of the RNA sequence.

U6 small nuclear RNA (snRNA), a member of the U snRNA family, is a component of eucaryotic spliceosomes (13, 26, 32) and is required for splicing nuclear pre-mRNAs (2, 3). In the budding yeast *Saccharomyces cerevisiae*, U6 is a single-copy gene essential for cell viability (4). Although the exact role of U6 snRNA in pre-mRNA splicing is not known, it has been suggested that U6 RNA may be directly involved in catalysis during nuclear pre-mRNA splicing (5).

U6 RNA, besides being the most conserved of all U snRNAs, is unique in many respects. First, while trimethylguanosine-cap-containing U snRNAs are transcribed by RNA polymerase II (for reviews, see references 6 and 7), U6 RNA is the only known capped RNA transcribed by pol III (15, 24). Second, U6 RNA is the only U snRNA that does not contain either the trimethylguanosine cap present on other U snRNAs (10) or the Sm antigen binding site found in other nucleoplasmic U snRNAs (17); instead, the 5′ end of human U6 snRNA is unique and is blocked by γ-monomethyl phosphate (30).

In eucaryotes, RNAs transcribed by RNA polymerase II, such as the mRNAs, snRNAs (U1 to U5), and most viral RNAs, are blocked on their 5′ termini by a guanosine cap: m^7GpppN in the case of mRNAs (for reviews, see references 1 and 27) and $m_3^{2,2,7}$GpppN in the case of snRNAs U1 to U5 (for reviews, see references 6 and 23). Trimethylation of the monomethylguanosine cap in U2 snRNA occurs posttranscriptionally in cytoplasm and requires the presence of an Sm binding site (18). While most mRNAs have an m^7G cap, certain viral RNAs have protein covalently attached to their 5′ termini (for a review, see reference 1). The cap structure in mRNAs is known to be important in mRNA biogenesis, including transcription initiation, generation of primers for mRNA synthesis, mRNA splicing, 3′-end formation, and mRNA stability, as well as in the translation of mRNAs (for reviews, see references 1 and 28). However, the function(s) of the $m_3^{2,2,7}$G cap structure in snRNAs U1 to U5 and of the methyl phosphate cap structure in U6 snRNA is not known.

Since U6 RNA has an unusual cap structure (γ-monomethyl phosphate), it offers a model system for elucidating the fundamental differences in the mechanism(s) of cap formation on U6 RNA versus mRNAs and other U snRNAs. The results from this study show that the information for the capping of U6 snRNA resides within the U6 RNA and is not inherent in the U6 transcription complex. The initial 25 nucleotides of U6 snRNA were found to be both necessary and sufficient for efficient capping. We further characterized the identity of capping determinants within this region by linker scan, substitution and deletion mutagenesis. The information obtained from these experiments was used to convert an otherwise noncapped heterologous transcript into a capped transcript by incorporating an AUAUAC sequence following a sequence capable of forming a synthetic stem-loop. The HeLa cell extracts were capable of capping U6 snRNAs from evolutionarily distant species.

MATERIALS AND METHODS

Construction of U6 mutants. The construction of altered U6 genes was done by using standard recombinant DNA techniques. The details of the construction of each plasmid are given in Table 1. Linker scan mutants were constructed by oligonucleotide-directed mutagenesis (16, 36).

Transcription and analysis for cap structure. HeLa cell extract was prepared as described previously (35). In a typical reaction, a 50-μl reaction mixture containing 2 μg of supercoiled plasmid DNA was incubated at 30°C for 3 h. The reaction mixture contained 50% (vol/vol) cell extracts, 0.6 mM unlabeled CTP, 0.6 mM unlabeled UTP, 0.6 mM unlabeled ATP, 0.025 mM GTP, 5 μCi of [α-^{32}P]GTP, 6 mM creatine phosphate, 20 mM KCl, 1 mM dithiothreitol, and 10 mM Tris (pH 8.1). The RNAs were phenol extracted, precipitated with ethanol, and loaded on 5 or 10% polyacrylamide gels. U6 RNA or other RNA transcripts were ex-

* Corresponding author.

TABLE 1. Characteristics of mutants

Mutant	Characteristics[a]
U6 (wild-type)	
1-106	Mouse U6 gene containing 315 bp upstream, 106 bp corresponding to U6 RNA, and 286 bp downstream in Bluescript vector (8)
1-100	Mouse U6 gene in Bluescript vector containing 150 bp upstream and 100 bp corresponding to U6 RNA; RNA analyzed was 146 nt long (100 nt original length + 46 nt from vector)
1-65	Mouse U6 gene in Bluescript vector containing 286 bp upstream and 65 bp corresponding to U6 RNA; RNA analyzed was 116 nt long (65 nt original length + 51 nt from vector)
1-25	Mouse U6 gene in Bluescript vector containing 315 bp upstream and 25 bp corresponding to U6 RNA; RNA analyzed was 175 nt long (25 nt original length + 150 nt from vector)
1-19	Mouse U6 gene in SP65 vector containing 315 bp upstream and 19 bp corresponding to U6 RNA; RNA analyzed was 445 nt long (19 nt original length + 426 nt from vector)
1-13	Mouse U6 gene in Bluescript containing 315 bp upstream and 13 bp corresponding to U6 RNA; RNA analyzed was 50 nt long (13 nt original length + 37 nt from vector)
1 or −315/1	Mouse U6 gene in Bluescript containing 315 bp upstream and 1 bp corresponding to U6 RNA; RNA analyzed was 71 nt long (1 nt original length + 70 nt from vector)
LS1/8	6 bp between nt 1 and 8 in U6 gene substituted using CTGCCGAAGggtaccCAAACAAGG
LS13/20	6 bp between nt 13 and 20 in U6 gene substituted using TTAGTATATggtaccCCGAAGCGA
LS19/26	6 bp between nt 19 and 26 in U6 gene substituted using CCAATTTTAggtaccGTGCTGCCG
LS19/26A	6 bp between nt 19 and 26 in U6 gene substituted using CCAATTTTAggtatcGTGCTGCCG
LS25/32	6 bp between nt 25 and 32 in U6 gene substituted using ATCGTTCCAggtaccGTATATGTG
LS13/20kl	LS13/20 construct digested with *Asp* 718, blunt ended with Klenow fragment of the DNA polymerase, and religated
dl(k)20	Constructed by inserting *Bam*HI-*Asp* 718 fragment of LS1/8 in place of *Bam*HI-*Asp* 718 region of LS13/20
dl(k)26	Constructed by inserting *Bam*HI-*Asp* 718 fragment of LS1/8 in place of *Bam*HI-*Asp* 718 region of LS19/26
dl3(k)26	Constructed by inserting *Bam*HI-*Asp* 718 fragment of LS13/20 in place of *Bam*HI-*Asp* 718 region of LS19/26
−315/1 E/S	Obtained by replacing *Eco*RI-*Sal*I polylinker sequence of −315/1 construct with synthetic oligonucleotides AATTCTGCAGCCCATATACGGG and TCGACCCGTATATGGGCTGCAG
−315/1 E/S/P	−315/1 E/S DNA was digested with *Pst*I and religated
hU6 1(+17)6	Same as pMarkedU6$^-$, a gift from N. Hernandez and S. Lobo at Cold Spring Harbor, N.Y.; contains a 17-bp polylinker inserted in place of nt 2 to 5 of human U6 gene
hU6 1(+8)6	Derived by digesting hU6 1(+17)6 with *Xba*I, blunt ending with SI nuclease, and religating; contains an 8-bp polylinker instead of nt 2 to 5 of human U6 gene
hU6 1(31)1-106	Constructed by cloning the −23/+84 *Bam*HI-*Hin*dIII fragment of mouse U6 clone (8) under human U6 promoter of hU6 1(+17)6 clone
LS1(Sc 2-32)26	*Bam*HI-*Hgi*AI fragment corresponding to −315/1 bp of mouse U6 gene was blunt ended at *Hgi*AI site and cloned along with synthetic oligonucleotides in dl(k)26 construct digested with *Bam*HI-*Asp* 718; oligonucleotides corresponding to *S. cerevisiae* nt 2 to 32 were TTCGCGAAGTAACCCTTCGTGGAC ATTTG and GTACCAAATGTCCACGAAGGGTTACTTCGCGAA.

[a] bp, Base pairs; nt, nucleotides.

tracted from the gels and digested with nuclease P1 (Sigma Chemical Co.), unless otherwise indicated (10). The nuclease P1 digest was used for electrophoresis on DEAE-cellulose paper at pH 3.5 (10). In some instances, the nuclease P1 digests were also treated with bacterial alkaline phosphatase at pH 8 before electrophoresis on DEAE-cellulose paper. The autoradiography was done for 48 to 72 h at −70°C with XAR-5 film and Cronex Lightning-Plus screens.

Capping efficiency. The autoradiograms were scanned with a laser beam densitometer (LKB Instruments, Inc.). The capping efficiencies of various mutants were expressed in terms of percentage: i.e., capping efficiency of mutant RNAs relative to capping efficiency of wild-type U6 RNA (taken as 100%) (Table 2).

RESULTS

U6 snRNA cap structures synthesized in vivo and in vitro are identical. To establish that an in vitro system faithfully caps U6 RNA, plasmid DNAs containing the 5S or U6 gene were transcribed in vitro with the whole HeLa cell extracts prepared by the method of Weil et al. (35). The resulting RNA transcripts were purified on polyacrylamide gels, digested with nuclease P1, and analyzed for their 5′ termini. The 5S RNA contained pppG (Fig. 1A, lane 3), and U6 RNA contained γ-monomethylguanosine triphosphate (CH_3 OpppG) in addition to pppG (Fig. 1A, lanes 4, 10, and 14). When these nuclease P1 digests were also treated with alkaline phosphatase, the phosphatase-resistant cap structure was found in U6 RNA (Fig. 1A, lane 2) but not in 5S RNA (Fig. 1A, lane 1). The cap structures synthesized in vivo and in vitro and synthetic CH_3OpppG were analyzed in a two-dimensional chromatographic system (29), and the mobilities of all were identical (results not shown). On the basis of these analyses, we conclude that the cap structure formed on U6 RNA in the in vitro system is the same as that found in vivo. Therefore, we used the in vitro system to delineate the substrate requirements for U6 snRNA capping. The conversion of the 5′ pppG to CH_3OpppG on U6 snRNA synthesized in vitro was in the range of 20 to 40%; the efficiency of capping differed with different batches of HeLa cell extracts. The nuclear extract prepared from HeLa cells by the method of Dignam et al. (9) was also tested; the efficiency of capping with nuclear extracts was significantly less (<10%) than that with whole-cell extracts (results not shown).

Initial 25 nucleotides of U6 RNA are necessary and sufficient for U6 RNA capping. Our earlier studies showed that the promoter for the mouse U6 gene is exclusively in the 5′-flanking region (8). To test whether the capping informa-

TABLE 2. Capping efficiencies of various U6 gene mutants[a]

Mutant	% Capping efficiency[b]
Controls	
5S	ND
U6 wild-type	100
3′ Deletion	
1-100	100
1-65	120
1-25	125
1-19	12
1-13	ND
1	ND
Internal deletion-insertion	
d1(k)26	ND
d13(k)26	ND
d1(k)20	8
hU6 1(+17)6	<10
hU6 1(+8)6	<10
Substitution	
LS1/8	50
LS13/20	30
LS13/20kl	<5
LS19/26A	<10
LS19/26	<10
LS25/32	100
Other	
−315/1 E/S	42
−315/1 E/S/P	ND
LS1(Sc2-32)26	15

[a] All the mutants are described in Table 1, and the results are shown in Fig. 1 to 5. The intensities of radioactive spots were quantitated by scanning the autoradiograms on an LKB laser beam densitometer.

[b] Capping efficiency of mutant RNA = (fraction of termini capped for mutant U6 RNA/fraction of termini capped for wild-type U6 RNA) × 100. The fraction of capped termini was calculated as (peak area under CH_3OpppG curve/total peak area [CH_3OpppG+pppG]) × 100. The capping of wild-type U6 varied between 20 and 40%. The results given are averages of three or more independent experiments. ND, Not detectable.

tion is U6 promoter specific, heterologous transcript (vector polylinker sequence) synthesized under the U6 promoter was analyzed for capping. This heterologous transcript was not capped (Fig. 1A, lane 9), while U6 RNA was capped (Fig. 1A, lanes 4, 10, and 14). These data suggest that the information necessary for capping is within the U6 RNA rather than in the U6 gene promoter. In order to delineate the capping determinant, we sequentially deleted the U6 gene from the 3′ end. Transcripts containing 100 (Fig. 1A, lane 5), 65 (lane 6), or 25 (lanes 7 and 13) nucleotides corresponding to the 5′ end of U6 RNA were all capped. The efficiency of capping with these constructs was equal to or sometimes more than that observed for wild-type U6 RNA (Table 2). Transcripts containing nucleotides 1 to 13 remained uncapped (Fig. 1A, lanes 8 and 12), while transcripts with nucleotides 1 to 19 of U6 RNA were capped 12% as efficiently as those of wild-type U6 RNA (lane 11). These data show that 1 to 25 nucleotides of U6 snRNA are necessary and sufficient for optimal capping efficiency (Fig. 1A). The data obtained from the 3′ deletion mutants also show that the minimal sequence required for capping extends to either nucleotide 19 or nucleotide 25 of U6 snRNA. Further studies are necessary to see whether the entire sequence between nucleotides 19 and 26 is necessary for efficient capping of U6 snRNA.

To see whether the U6 RNA sequence downstream of nucleotide 25 has any capping determinant, we deleted nucleotides 2 to 24 and 14 to 24 of U6 snRNA such that a *Kpn*I site (ggtacc) was generated in place of the deleted sequences; both the deletions abolished the capping (Fig. 1B, lanes 2 and 4). Transcripts containing 20 to 106 nucleotides of U6 RNA from the construct d1(K)20 were capped with low (8%) efficiency (Fig. 1B, lane 3) (Table 2). These data show that nucleotides 26 to 106 of U6 RNA cannot direct capping.

Capping determinant in U6 snRNA is bipartite. In all the secondary-structure models proposed for U6 snRNA, the 5′ region of U6 RNA can potentially form a stem-loop structure (Fig. 2A) (for a review, see reference 23). The chemical modification data for U6 snRNA support this secondary

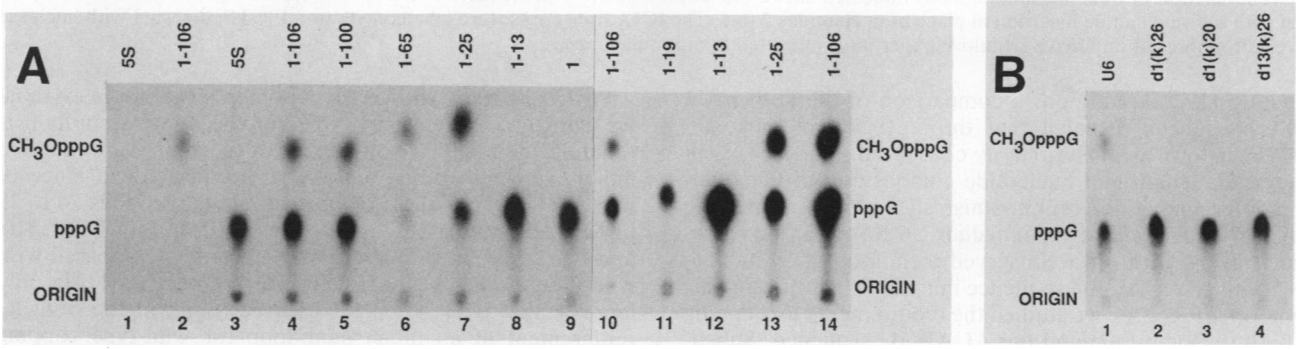

FIG. 1. Analysis of 5′ termini of transcripts from 3′-deletion and internal-deletion mutants of U6 RNA. The 3′-deletion mutants of the U6 gene were transcribed in vitro, and the transcripts were analyzed for 5′ termini by electrophoresis on DEAE-cellulose paper. In addition to a portion of the U6 gene, these transcripts also contained part of the vector DNA transcribed (Table 1). (A) Lanes 1 and 3, 5S DNA; lanes 2, 4, 10 and 14, mouse U6 gene; lanes 5 to 9 and 11 to 13, 3′-deletion mutants of U6 gene containing 100, 65, 25, 19, 13, or 1 base pair corresponding to the 5′ end of U6 RNA as indicated above each lane. The transcripts analyzed in lanes 3 to 14 were digested with nuclease P1, and samples in lanes 1 and 2 were digested with alkaline phosphatase following nuclease P1 digestion. The electrophoretic migration of CH_3OpppG in lanes 6 and 7 is different from that in other lanes, probably because of different sample volumes loaded and consequent lane-to-lane pH differences. Radioactive spots corresponding to pG and Pi, which migrate faster (24) in this electrophoretic system, are not shown in this and subsequent figures. (B) Lane 1, Mouse U6 gene; lanes 2 to 4, U6 DNA in which sequences between 1 and 26, 1 and 20, and 13 and 26 nucleotides, respectively, were replaced with linker ggtacc. The RNA transcripts were labeled with [α-^{32}P]GTP, digested with nuclease P1, electrophoresed on DEAE-cellulose paper, and subjected to autoradiography.

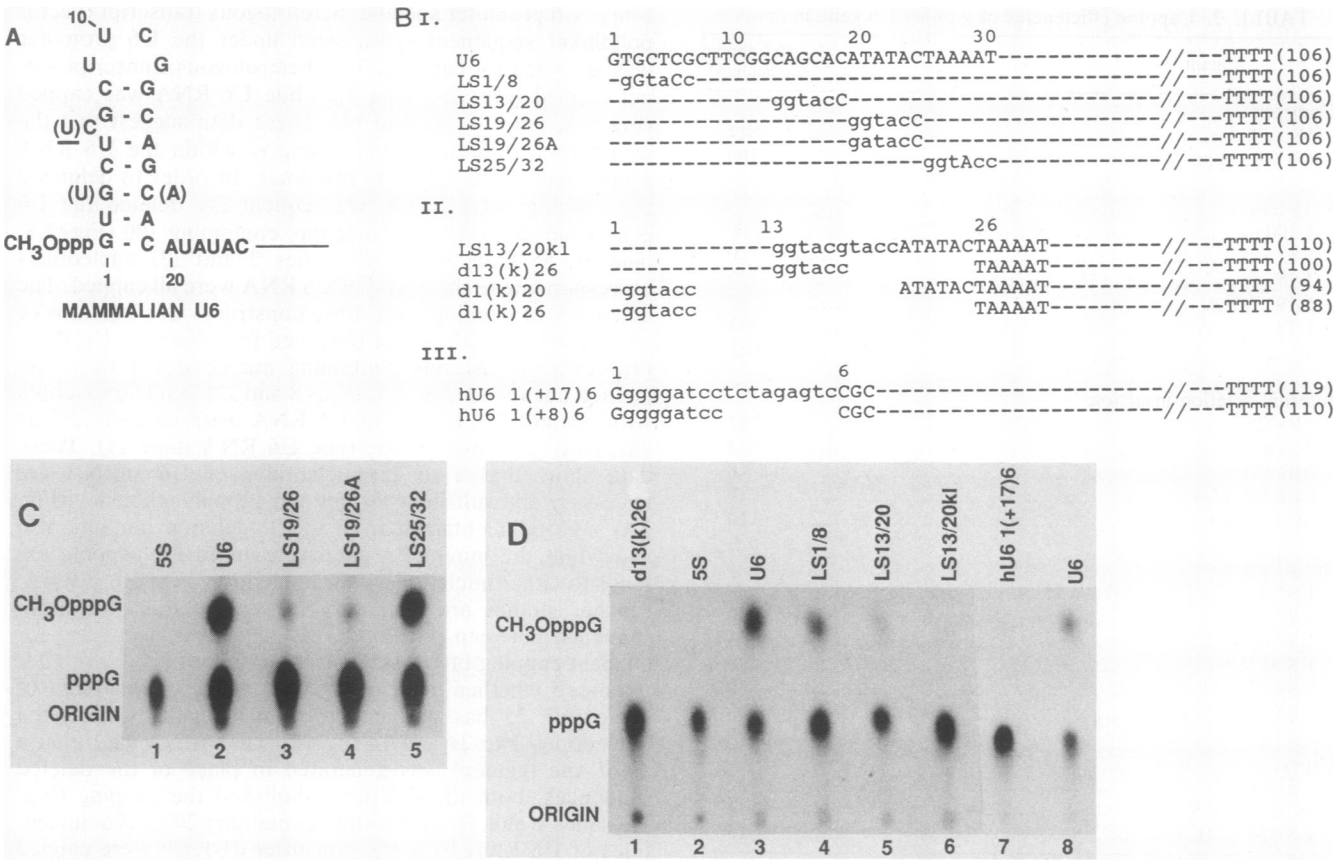

FIG. 2. (A) Conserved secondary structure of 5′ end of mammalian U6 RNA (10). The nucleotide substitutions in *Drosophila* U6 RNA are shown in parentheses, and the substitution in the *Xenopus* U6 snRNA is underlined. Some U6 snRNA sequences, like bean U6 RNA (14), do not fit into the secondary structure. (B) Schematic representations of linker scan, deletion, and spacer mutations. The wild-type U6 DNA sequence is shown on the top line, and the linker-spacer sequences are shown in small letters. Dashes indicate nucleotide identity. Gaps indicate the deletions. Slash marks indicate breaks in the U6 gene not shown in the figure. The expected and observed transcript lengths for each clone are in parentheses. The numbers above the sequences correspond to the transcribed portion of mammalian U6 snRNA (10). (C) Analysis of 5′ terminus of transcripts from linker scan mutants of U6 RNA. The linker scan mutants were transcribed in vitro, and the 5′ termini were analyzed as described in the legend to Fig. 1A. Lane 1, 5S gene; lane 2, wild-type U6 gene; lanes 3 to 5, different linker scan mutants as indicated above the lanes. The RNA transcripts were labeled with [α-^{32}P]GTP, digested with nuclease P1, electrophoresed on DEAE-cellulose paper, and subjected to autoradiography. (D) Analysis of 5′ termini of transcripts from linker scan mutants of U6 RNA. The linker scan mutants were transcribed in vitro, and the 5′ termini were analyzed as described in the legend to Fig. 1A, lanes 1 and 4 to 6, Different deletion or linkerscan mutants as indicated above the lanes; lane 2, 5S gene; lanes 3 and 8, wild-type U6 gene; lane 7: human U6 gene with a 17-nucleotide insertion in place of nucleotides 2 to 5. The RNA transcripts were labeled with [α-^{32}P]GTP, digested with nuclease P1, electrophoresed on DEAE-cellulose paper, and subjected to autoradiography.

structure (19). In addition, a comparison of the known U6 RNA sequences revealed that the 5′ region of most U6 RNAs can form an evolutionarily conserved stem-loop, even though the length and nucleotide composition of the stem-loop differ among different organisms (21, 25, 33). Since the initial 25 nucleotides of mammalian U6 RNA can be represented in the form of a conserved stem-loop of 19 nucleotides with an AUAUAC sequence immediately following this stem-loop (Fig. 2A), we studied the requirements for capping of both the stem-loop and the AUAUAC sequence. Substitution mutants corresponding to both the stem-loop and the single-stranded regions were made as shown in Fig. 2B. The substitution of AUAUAC (nucleotides 20 to 25) with either GGUACC (Fig. 2C, lane 3) or GAUACC (Fig. 2C, lane 4) decreased the capping efficiency by over 90%; the substitution of nucleotides 26 to 31 did not have any significant effect on the capping efficiency (Fig. 2C, lane 5) (Table 2). Therefore, all or a part of the AUAUAC sequence in the single-stranded region following the stem-loop is an important component of the capping determinant.

Phylogenetically conserved stem-loop structure is essential for capping. To determine whether the intact stem-loop is required for capping, both halves of the stem were mutated individually by placing a *Kpn*I restriction site in place of nucleotides 2 to 7 in LS1/8 mutant and nucleotides 14 to 19 in LS13/20 mutant (Fig. 2B). These substitutions in LS1/8 and LS13/20 reduced the capping efficiency (compared with wild-type U6 RNA) by 50 and 70%, respectively (Fig. 2D, lanes 4 and 5) (Table 2). To further confirm the obligate requirement of an intact stem-loop for wild-type capping efficiency, the stem-loop was disrupted by the insertion of a 17-mer in place of nucleotides 2 to 5 (Fig. 2B). This resulted in a reduction in the capping efficiency of approximately 90% (Fig. 2D, lane 6). A similar effect on capping efficiency was observed when only 8 nucleotides, instead of 17, were inserted (data not shown) (Table 2). On the basis of these insertion, deletion, and substitution mutagenesis studies, an intact stem-loop structure appears to be an integral part of the capping determinant necessary for wild-type capping efficiency. This and the observation that nucleotides 20 to 25

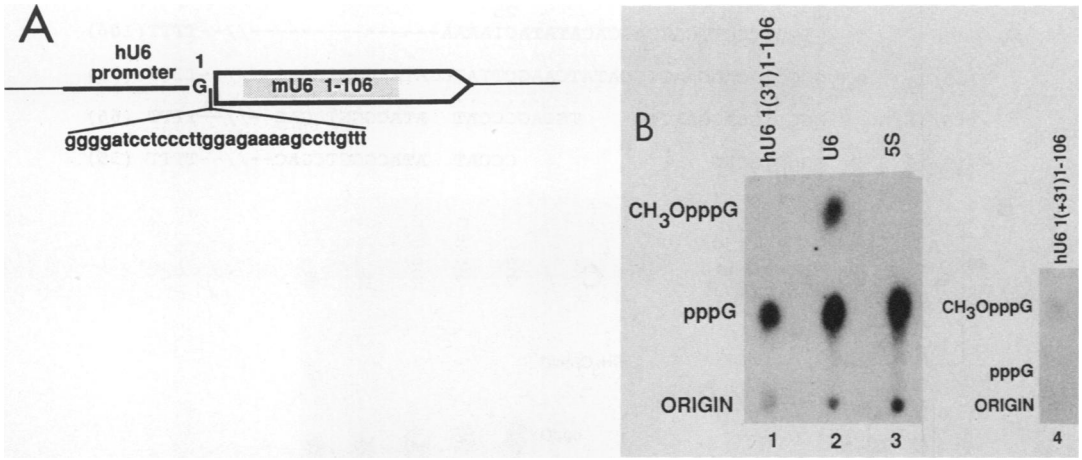

FIG. 3. (A) Schematic representation of a DNA construct in which the U6 capping signal is internal. A 31-base-pair linker sequence is inserted between the complete U6 sequence and the U6 promoter with the initiation nucleotide G. The U6 sequence is in a horizontal box. hU6, Human U6; mU6, mouse U6. (B) Analysis of cap structure from the 5′ spacer mutant. The U6 mutant was transcribed and analyzed for cap structure (lane 1). Lane 2, Wild-type U6 gene; lane 3, 5S gene used as control; lane 4, same as lane 1, except that more sample was used, and the sample was also digested with alkaline phosphatase.

of U6 snRNA are important show that the capping determinant is a bipartite element comprising an intact stem-loop structure and an AUAUAC sequence following this stem-loop.

If there is a stringent requirement that the AUAUAC sequence be next to the stem-loop, the insertion of a spacer sequence between this stem-loop and the AUAUAC motif should result in a reduction in capping efficiency. The insertion of GUACC between the stem-loop of LS13/20 and the AUAUAC sequence in LS13/20kl reduced the capping efficiency from 30% (Fig. 2D, lane 5) to <5% (Fig. 2D, lane 6). These data indicate that the AUAUAC sequence, as a part of the capping determinant, functions efficiently only when it immediately follows the stem-loop.

Close proximity of the initiation nucleotide to the stem-loop favors capping. To seek further insight into the structure of substrate necessary for U6 RNA-specific capping, we made a construct—hU6 1(31) 1-106—such that 32 nucleotides were 5′ to the whole U6 snRNA (Fig. 3A). There was no readily detectable cap structure in these transcripts (Fig. 3B, lane 1). However, when more radioactive transcripts were used and subjected to autoradiography for a longer time, there was some detectable cap structure (Fig. 3B, lane 4) (Table 2). These data show that the γ-phosphate of the initiation nucleotide needs to be in close proximity to the capping signal for wild-type capping efficiency.

Noncapped vector polylinker RNA converted into a capped RNA. We attempted to convert a noncapped heterologous transcript (vector polylinker sequence) into a capped RNA. To achieve this, we used the construct −315/1, in which a 71-nucleotide-long vector polylinker sequence is transcribed under the U6 promoter (8); this transcript was not capped (Fig. 4C, lane 1). We replaced a portion of this polylinker at positions 16 to 36 with another sequence (Fig. 4A) such that the polylinker could form a heterologous stem-loop structure (a 9-base-pair stem and a 5-nucleotide loop, compared with the 7-base-pair stem and a 4-nucleotide loop of mammalian U6 RNA [Fig. 2A]) and an AUAUAC sequence would appear immediately following this stem-loop (Fig. 4B). This transcript was capped (Fig. 4C, lane 2), and the efficiency of capping was 42% compared with that of wild-type U6 RNA (Fig. 4C, lane 5) (Table 2); 5S RNA, used as a control, was not capped (Fig. 4C, lane 4). Although the primary sequences of U6 RNA and the transcripts from the −315/1 E/S construct are altogether different, both are capped presumably because they share a common structural feature(s). To know whether there are any requirements for the loop and minimum stem length, we deleted nucleotides 7 to 17 of −315/1 E/S (Fig. 4B). The deletion of the whole loop and a part of the stem abolished capping (Fig. 4C, lane 3), suggesting that there may be a structural requirement for the loop, or a minimum stem length, or both, for capping.

HeLa cell extracts can cap U6 snRNAs from evolutionarily distant species. The secondary structures corresponding to the 5′ ends of U6 snRNAs from *S. cerevisiae* (Fig. 5A) and from mammals (Fig. 2A) are similar. To see whether U6 snRNAs from other species are capped in HeLa cell extracts, we transcribed DNA containing the 5′ 32 nucleotides of the *S. cerevisiae* U6 gene in HeLa cell extract. The RNA synthesized from this DNA construct was 116 nucleotides long and contained 1 to 32 nucleotides corresponding to *S. cerevisiae* U6 RNA followed by an acc sequence and nucleotides 26 to 106 from mouse U6 snRNA. These transcripts were capped, though at a reduced efficiency compared with that of mammalian U6 RNA (Fig. 5B, lanes 2, 3, 4, 6, and 7) (Table 2). As expected, 5S RNA (Fig. 5B, lane 1) and U6 RNA in which nucleotides 2 to 25 were deleted (Fig. 5B, lane 5) were not capped. In addition, U6 snRNAs synthesized from *Xenopus* or *Drosophila* U6 genes in HeLa cell extracts were efficiently and accurately capped (data not shown). These data show that HeLa cell extracts can cap U6 snRNAs from both lower and higher eucaryotes.

DISCUSSION

We report here that capping of mammalian U6 snRNA in vitro is dependent on the U6 snRNA sequence. Wild-type capping efficiency requires the following three elements in a particular context: an intact stem-loop, an AUAUAC sequence immediately following this stem-loop, and pppG in close proximity to the stem-loop. This AUAUAC sequence, when present following a synthetic stem-loop generated from a vector polylinker sequence, was sufficient to convert a noncapped heterologous transcript into a capped tran-

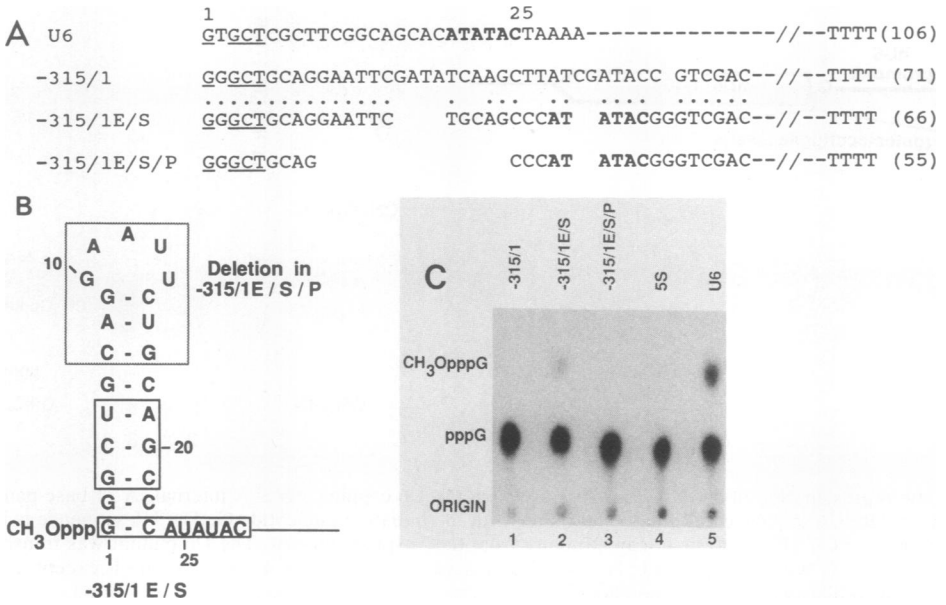

FIG. 4. Conversion of a noncapped vector polylinker transcript into a capped transcript. (A) Comparison of the 5′ portions of wild-type U6 RNA (U6), the polylinker sequence in Bluescript vector (−315/1), and the modified Bluescript vectors (−315/1E/S and −315/1E/S/P). The underlined nucleotides are similar in mammalian U6 and transcripts from the vector; the dots represent nucleotide identity between −315/1 and −315/1E/S; the numbers in parentheses represent transcript lengths. Slash marks indicate breaks in the sequence not shown. (B) Predicted secondary structure of the 5′ portion of RNA transcribed from the −315/1E/S DNA construct; the conserved nucleotides between RNA from the −315/1E/S DNA construct and the mammalian U6 snRNAs are shown in solid boxes. The region shown in the large wavy box (nucleotides 7 to 17) is absent in the −315/1E/S/P construct. (C) Analysis of cap structure from the altered vector transcripts. The DNA constructs shown in panel A were transcribed and analyzed for the cap structure. Lanes 1 to 5, Transcripts from DNA constructs indicated above the lanes.

script. U6 RNAs from *Xenopus tropicalis*, *Drosophila melanogaster*, and a lower eucaryote, *S. cerevisiae*, were capped in a mammalian extract.

The results presented here indicate that the mechanism of U6 snRNA capping is distinct from that of mRNAs and snRNAs U1 to U5. There are several lines of evidence to suggest that transcription and capping of mRNAs and snRNAs U1 to U5 are concurrent. Detailed studies with viral and cellular mRNAs (12; for a review, see reference 1) as well as snRNAs U1 to U5 (18, 31) showed that capping and transcription by pol II are coupled. Furthermore, over 95% of the transcripts synthesized by heterologous T7 RNA polymerase in mammalian cells remain uncapped, presumably because of the inability of the capping machinery to interact with heterologous T7 polymerase (11). These data further confirm that capping is tightly linked to transcription

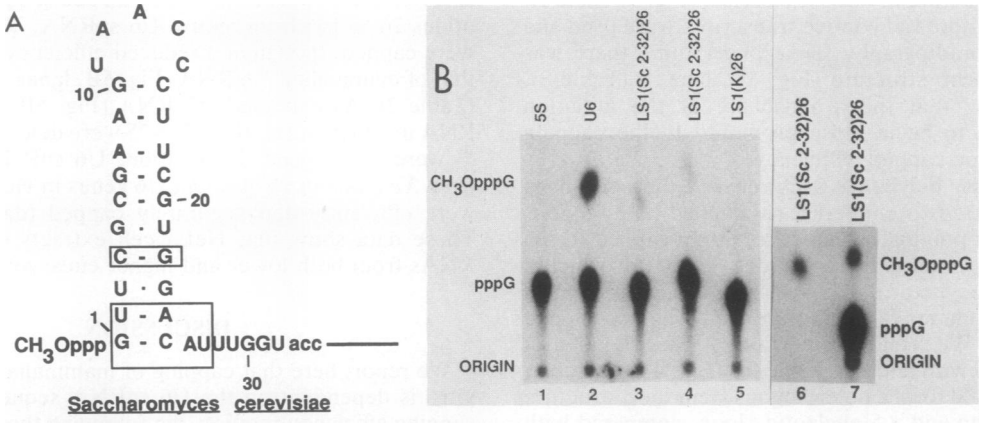

FIG. 5. (A) Predicted secondary structure for 5′ end of yeast (*S. cerevisiae*) U6 snRNA. The nucleotides conserved between mammalian and *S. cerevisiae* U6 RNAs (4, 25) are shown in boxes. The acc sequence is from the linker sequence, and the solid line corresponds to the mammalian U6 RNA nucleotide sequence from nucleotide 26 to 106. (B) Analysis of cap structure. The transcripts from the DNA construct LS1(Sc2-32)26 from three different experiments were digested with nuclease P1 and analyzed for cap structure (lanes 3, 4, and 7). Lane 1, 5S gene; lane 2, wild-type U6 gene; lane 5, RNA from LS1(K)26 DNA as a noncapped control in which the capping determinant, nucleotides 2 to 25, is deleted from the mouse U6 gene; lane 6, aliquot of sample analyzed in lane 7 digested with alkaline phosphatase and subjected to electrophoresis.

by pol II. Since the transcripts lacking the initial 25 nucleotides of U6 RNA synthesized under U6 promoter remained uncapped, it would be reasonable to argue that capping of U6 RNA is not inherent in the U6 transcription complex.

The pol II-associated capping enzyme has not been shown to exhibit sequence specificity (at least not qualitatively), because RNAs from different sources and even homopolyribonucleotides have been capped (20, 34). However, the information is still insufficient for us to assume that all RNAs in a mixture will be capped with equal efficiencies. In contrast, our data on the characterization of U6 capping substrate revealed the importance of two elements—a stem-loop structure at the 5′ end of U6 RNA and an AUAUAC motif immediately following this stem-loop. RNA-sequence-dependent accurate methylation of internal adenosine residues in mRNA was reported recently (22); however, this is the first instance of RNA-sequence-dependent capping.

The 5′ region of most U6 snRNAs can be represented in the form of an evolutionarily conserved stem-loop structure, although nucleotide composition and the length of this stem-loop are rather variable (25). The integrity of this stem-loop is important for U6 RNA function, since deletion of the stem-loop is lethal in *S. cerevisiae*. The stem-loop of *Schizosaccharomyces pombe*, which is half the length of that of *S. cerevisiae*, can rescue these mutants (C. Guthrie, personal communication). It appears that elements for capping and the essential functional domain(s) of U6 RNA in the 5′ region are the same or at least overlapping.

The capping determinant in U6 snRNA has rather stringent requirements with respect to its distance from the initiation nucleotide. This is in contrast to Sm-binding-site-dependent trimethylation of the m^7G cap of U2 snRNA, in which the distance between the Sm antigen binding site and the cap site is variable (18). The Sm-binding site is present at different places with respect to the initiation nucleotide in different nucleoplasmic U snRNAs (17). A comparison of the available U6 snRNA sequences shows that a stem-loop is present near the 5′ end; this possibly reflects that any insertions before the stem-loop were selected against during evolution to position the initiation nucleotide very close to the capping determinant.

Incorporation of an AUAUAC sequence following a synthetic stem-loop derived from a vector polylinker sequence was sufficient to convert a noncapped vector polylinker RNA transcribed under the U6 promoter into a capped transcript (Fig. 4). Although the nucleotide sequences of the mammalian U6 RNA, the heterologous capped transcript, and the *S. cerevisiae* U6 RNA are markedly unrelated, all these transcripts are capped in HeLa cell extracts, presumably because they have a common structural element(s) which acts as the capping determinant. Further studies are necessary to determine whether conserved base pairs in the stem-loop have any effect on capping efficiency or U6 RNA function or both. It is likely that another cellular RNA(s) with this kind of cap structure, if such an RNA exists, would share this structural motif characterized for U6 RNA.

ACKNOWLEDGMENTS

This investigation was supported by Public Health Service grant GM-38320 from the National Institute of Health.

We thank Harris Busch for encouragement; Michael Barattain for the use of the LKB densitometer; Y. Ohshima, S. Lobo, and W. Folk for mouse U6, human U6, and Chinese hamster 5S genes, respectively; L. Rothblum for synthesizing deoxyoligonucleotides; Minyone Finley for HeLa cells; and S. Chirala, S. Natarajan, Dale Henning, and Gokul Das for valuable discussions.

LITERATURE CITED

1. **Banerjee, A. K.** 1980. 5′-Terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol. Rev. **44**:175–205.
2. **Berget, S. M., and B. L. Robberson.** 1986. U1, U2, and U4/U6 small nuclear ribonucleoproteins are required for in vitro splicing but not for polyadenylation. Cell **46**:691–696.
3. **Black, D. L., and J. A. Steitz.** 1986. Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein. Cell **46**:697–704.
4. **Brow, D. A., and C. Guthrie.** 1988. Spliceosomal RNA U6 is remarkably conserved form yeast to mammals. Nature (London) **334**:213–218.
5. **Brow, D. A., and C. Guthrie.** 1989. Splicing a spliceosomal RNA. Nature (London) **337**:14–15.
6. **Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi.** 1982. snRNAs, snRNPs and RNA processing. Annu. Rev. Biochem. **51**:617–654.
7. **Dahlberg, J. E., and E. Lund.** 1988. The genes and transcription of the major small nuclear RNAs, p. 38–70. *In* M. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag KG, Berlin.
8. **Das, G., D. Henning, D. Wright, and R. Reddy.** 1988. Upstream regulatory elements are necessary and sufficient for transcription of U6 RNA gene by RNA polymerase III. EMBO J. **7**:503–512.
9. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
10. **Epstein, P., R. Reddy, D. Henning, and H. Busch.** 1980. The nucleotide sequence of nuclear U6 (4.7S) RNA. J. Biol. Chem. **255**:8901–8906.
11. **Fuerst, T. R., and B. Moss.** 1989. Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells. Importance of 5′ untranslated region. J. Mol. Biol. **206**:333–348.
12. **Furuichi, Y.** 1978. "Pretranscriptional capping" in the biosynthesis of cytoplasmic polyhedrosis virus mRNA. Proc. Natl. Acad. Sci. USA **75**:1086–1090.
13. **Grabowski, P., and P. A. Sharp.** 1986. Affinity chromatography of splicing complexes: U2, U5, and U4+U6 small nuclear ribonucleoprotein particles in the spliceosome. Science **233**: 1294–1299.
14. **Kiss, T., M. Antal, and F. Solymosy.** 1987. Plant small nuclear RNAs. II. U6 RNA and 4.5SI-like RNA are present in plant nuclei. Nucleic Acids Res. **15**:543–560.
15. **Kunkel, G. R., R. L. Maser, J. P. Calvet, and T. Pederson.** 1986. U6 small nuclear RNA gene is transcribed by RNA polymerase III. Proc. Natl. Acad. Sci. USA **83**:8575–8579.
16. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
17. **Liautard, J. P., J. Sri-Widada, C. Brunel, and P. Jeanteur.** 1982. Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. J. Mol. Biol. **162**:623–642.
18. **Mattaj, I. W.** 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell **46**: 905–911.
19. **Miura, K., S. Tsuda, F. Harada, and T. Ueda.** 1983. Chemical modification of cytosine residues of U6 snRNA with hydrogen sulfide. Nucleic Acids Res. **11**:5893–5901.
20. **Mizumoto, K., and F. Lipmann.** 1979. Transmethylation and transguanylation in 5′-RNA capping system isolated from rat liver nuclei. Proc. Natl. Acad. Sci. USA **76**:4961–4965.
21. **Mottram, J., K. L. Perry, P. M. Lizardi, R. Lührmann, N. Agabian, and R. G. Nelson.** 1989. Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4, and U6 RNA analogs. Mol. Cell. Biol. **9**:1212–1223.
22. **Narayan, P., and F. Rottman.** 1988. An in vitro system for accurate methylation of internal adenosine residues in messenger RNA. Science **242**:1159–1162.
23. **Reddy, R., and H. Busch.** 1988. Small nuclear RNAs: RNA

sequences, structure and modifications, p. 1–37. *In* M. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag KG, Berlin.

24. **Reddy, R., D. Henning, G. Das, M. Harless, and D. Wright.** 1987. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. J. Biol. Chem. **262**:75–81.
25. **Roiha, H., E. O. Shuster, D. A. Brow, and C. Guthrie.** 1989. Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive size variation. Gene **82**:137–144.
26. **Ruby, S. W., and J. Abelson.** 1988. An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. Science **242**:1028–1035.
27. **Shatkin, A. J.** 1976. Capping of eukaryotic mRNAs. Cell **9**: 645–653.
28. **Shatkin, A. J.** 1985. mRNA cap binding proteins: essential factors for initiating translation. Cell **40**:223–224.
29. **Silberklang, M., A. M. Gillum, and U. L. RajBhandary.** 1979. Use of in vitro ^{32}P labelling in the sequence analysis of nonradioactive tRNAs. Methods Enzymol. **59**:58–109.
30. **Singh, R., and R. Reddy.** 1989. γ-Monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. Proc. Natl. Acad. Sci. USA **86**:8280–8283.
31. **Skuzeski, J. M., E. Lund, J. T. Murphy, T. H. Steinberg, R. R. Burgess, and J. E. Dahlberg.** 1984. Multiple elements upstream of the coding region are required for accumulation of human U1 RNA in vivo. J. Biol. Chem. **259**:8345–8352.
32. **Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, A. Kramer, D. Frendeway, and W. Keller.** 1988. Functions of major small nuclear RNAs, p. 115–154. *In* M. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag KG, Berlin.
33. **Tani, T., and Y. Ohshima.** 1989. The gene for the U6 small nuclear RNA in fission yeast has an intron. Nature (London) **337**:87–90.
34. **Venkatesan, S., and B. Moss.** 1980. Donor and acceptor specificities of HeLa cell mRNA guanylyltransferase. J. Biol. Chem. **255**:2835–2842.
35. **Weil, P. A., J. Segall, B. Harris, S.-Y. Ng, and R. G. Roeder.** 1979. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. J. Biol. Chem. **254**:6163–6173.
36. **Zoller, M., and M. Smith.** 1982. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucleic Acids Res. **10**:6487–6500.