Cis-acting elements required for RNA polymerase II and III transcription in the human U2 and U6 snRNA promoters

Susan M.Lobo, Samantha Ifill and Nouria Hernandez
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Received February 14, 1990; Revised and Accepted April 9, 1990

ABSTRACT

Although the human U2 and U6 snRNA genes are transcribed by RNA polymerases II and III respectively, their promoters are remarkably similar in structure. Both promoters contain a proximal element and an enhancer region with an octamer motif. The U6 promoter contains in addition an A/T rich region that defines it as an RNA polymerase III promoter. We have examined in further detail the contributions of sequences in the human U2 and U6 promoter regions to transcription by RNA polymerase II and III. We find that although the sequences surrounding the U2 cap site favor RNA polymerase II transcription, their presence cannot suppress a shift to RNA polymerase III specificity upon insertion of the U6 A/T box. In the U6 promoter, the 3′ part of the proximal element homology is essential for efficient transcription and is also involved in localizing the start site of transcription. A region downstream of the proximal element homology is required for RNA polymerase II (but not for RNA polymerase III) transcription, both in the U2 promoter and in the U6 promoter. This element may be recognized by an RNA polymerase II transcription factor or by RNA polymerase II itself. The presence of this element in the U6 promoter raises the possibility that the human U6 gene is, under certain circumstances, transcribed by RNA polymerase II.

INTRODUCTION

Most of the RNA polymerase III transcription units examined to date are characterized by a promoter that is at least in part located within the coding sequence of the gene (1). The U6 small nuclear RNA (snRNA) gene, the 7SK gene, and probably the MRP RNA gene are members of a new class of RNA polymerase III transcription units in which all the elements essential for initiation of transcription are located within the sequences 5′ of the genes (2–4). Strikingly, the U6 promoter resembles the RNA polymerase II promoter of the U2 and other snRNA genes both in the sequence and spatial organization of two of its elements (2, 5–7). It is composed of an enhancer region located more than 200 base pairs upstream of the transcriptional start site and characterized, like the U2 snRNA gene enhancer, by the presence of a perfect octamer motif (2, 6–12). It also contains a 17 base pair snRNA proximal element which in the human gene matches the U2 proximal element in 13 out of 17 positions and which is absolutely required for efficient transcription by RNA polymerase III (2, 6, 13, 14). Interestingly, an 11 nucleotide sequence overlapping 5 nucleotides of the downstream portion of the human U6 proximal element is also found in roughly the same position relative to the start site of transcription in the Saccharomyces cerevisiae U6 gene (15). The significance, if any, of this similarity is unknown. The third element of the U6 promoter is not present in the U2 promoter and consists of an A/T rich region (A/T box) located 18 to 26 nucleotides upstream of the transcription start site (13, 16). The A/T box is reminiscent of the TATA box of mRNA promoters, although its function is quite different.

In the human U6 promoter, the A/T box located in the 5′ flanking region is a dominant element that defines the promoter as an RNA polymerase III promoter (13). When the A/T box is mutated, the U6 promoter is converted into an RNA polymerase II promoter. This RNA polymerase II promoter is of the snRNA type, in that it directs the formation of a transcription complex capable of recognizing the 3′ box (13), the termination signal located downstream of the RNA polymerase II snRNA genes (7, 14, 17–19). This is in contrast to mRNA promoters, which direct the formation of RNA polymerase II transcription complexes that ignore the 3′ box (19, 20). More significantly, introduction of the U6 A/T box into the human U2 promoter extending from −247 to −7, i.e. lacking the U2 cap site, converts it into an RNA polymerase III promoter (13). This observation suggests that the U2 and U6 promoters may bind a common set of transcription factors, and that the RNA polymerase specificity may be determined by the absence or presence of the factor(s) binding to the A/T box. However, in the Xenopus U6 gene, transcription by RNA polymerase III requires, in addition to the A/T box, the sequences surrounding the start site of transcription (16). This observation raises the possibility that in the complete human U2 promoter with an A/T box inserted, the sequences surrounding the U2 cap site just upstream of the potential RNA polymerase III start site could supress RNA polymerase III transcription.

Here, we have pursued three aspects of our study of the U2 and U6 promoters. First, we have examined the role of the U2 cap site region in determining RNA polymerase specificity of a U2 promoter containing an A/T box. We find that although it influences to a certain extent the relative levels of RNA polymerase II and polymerase III transcription, this effect is minimal. Second, we have further characterized the properties of a minor start site located around position −16 in the U6 promoter, which is activated by mutations in the 3′ portion of

the U6 proximal element. Third, we have examined in further detail the promoter region immediately downstream of the proximal element homology in a mutated U2 promoter containing an A/T box and therefore recognized by both RNA polymerases II and III. We find that a mutation in this region is very deleterious for RNA polymerase II transcription but has only a minor effect on RNA polymerase III transcription. Interestingly, mutation of the region downstream of the proximal element in a U6 promoter lacking an A/T box also inhibits RNA polymerase II transcription, although this region is not conserved in sequence between the U2 and U6 promoters. Thus, this region may bind an RNA polymerase II specific transcription factor with a wide DNA sequence recognition flexibility, or perhaps RNA polymerase II itself.

MATERIALS AND METHODS

Constructs

The constructs pU2/−247/RA.2, pU2/TA, U6/Hae/RA.2, LS3/Hae, LS4/Hae, LS5/Hae, have been described in (13), and the constructs pU2/−247, LS−34/−43, and LS−33/−42 have been described in (14).

All the new constructs were generated by oligonucleotide-directed mutagenesis. pU2/U2Cap and pU2/LS4/RA.2 were derived from pU2/−247/RA.2; pU2/TA/U2Cap, pU2/TA/(LS−34/−43) and pU2/TA/LS4 from pU2/TA; LS3b/Hae and LS24/Hae from U6/Hae/RA.2. The oligonucleotides used were:

pU2/U2Cap	GGCGCGAAGGCGAGCGCATCG
pU2/TA/U2Cap	CTTAGAGTCGAGATCCACG
pU2/TA/(LS−34/−43)	GCGACTTGAATGCCCTCGAGG
	GTGGG
pU2/TA/LS4	CTCACCGCGACTTCTCGAGGGATG
pU2/LS4/RA.2	
LS3b/Hae	CTTACCGTAACTTGTCGAGTCTCGAATTTC
LS24/Hae	TACCGCTCGAGCTCGAGATTTCG

The transfections into HeLa cells, the *in vitro* transcriptions and the RNA analysis were performed as described previously (13).

RESULTS

Figure 1A shows the constructs and probes used in this study. pU2/−247/RA.2 contains the U2 promoter from position −247 to position −7 followed by a 137 bp fragment of spacer DNA derived from the rabbit β-globin gene. The spacer sequence is followed by three signals for RNA 3′ end formation which are, in order: (i) the last six nucleotides of the human U1 coding sequence and the U1 3′ box, which together direct the formation of discrete 3′ ends on transcripts derived from RNA polymerase II snRNA promoters, (ii) a run of six T residues which constitutes a termination signal for RNA polymerase III, and (iii) the L3 polyadenylation site from Adenovirus 2. pU6/Hae/RA.2 is similar to pU2/−247/RA.2, except that the U2 promoter fragment is replaced by a U6 promoter fragment extending from position −241 to position +1.

Transcription from the different constructs was analyzed *in vitro* by incubation in a HeLa cell nuclear extract or *in vivo* by transient expression in HeLa cells. As depicted in Figure 1B, the hybrid genes were inserted into a derivative of pUC119 in which they are preceded by two tandem polyadenylation sites derived from the early region of SV40 (labeled SV40pA) and followed by a fragment containing the SV40 origin of replication (labeled SV40 ORI). The presence of the tandem SV40 polyadenylation sites reduces considerably the amount of stable RNA extending through the U2 or U6 promoters in transient transfections, presumably because RNA derived from cryptic promoters within the vector becomes polyadenylated at these sites. The SV40 origin of replication allows the vector to replicate when T antigen is provided in *trans*, and this results in higher levels of expression from both the U2 and U6 constructs.

A U2 promoter including the U2 cap site acquires RNA polymerase III specificity after insertion of a A/T box

We have previously shown that insertion of the U6 A/T box into a U2 promoter fragment extending from positions −247 to −7 upstream of the U2 cap site converts the U2 promoter into a predominantly RNA polymerase III promoter (13). This construct (pU2/TA) did not contain the natural U2 cap site. Thus, the possibility remained that the U2 cap site region was inhibitory to RNA polymerase III transcription or contained a determinant for RNA polymerase II transcription which, if present in pU2/TA, would have counteracted the effect of the A/T box and maintained RNA polymerase II specificity. To exclude these possibilities, we modified the parent construct, pU2/−247/RA.2, as well as the pU2/TA construct by site-directed mutagenesis to create pU2/U2cap and pU2/TA/U2cap (Figure 2A). In these modified constructs the natural U2 sequence surrounding the U2 cap site was restored, so that the U2 promoter sequences now extend to position +7 downstream of the U2 cap site. The four constructs were transfected into HeLa cells, and initiation of transcription was analyzed by primer extension. As shown in Figure 2B (lane 5), transcription derived from pU2/−247/RA.2 initiates mainly at a C residue, two nucleotides further upstream than in the natural U2 promoter (compare the upper bands of lanes 1 and 5). A minor start site is also detected at an A residue at position +4 relative to the wild type cap site (lane 5). When the wild type U2 sequences from −6 to +7 are restored, initiation shifts to the normal U2 cap site (lane 1). In both constructs, insertion of the A/T box results in the appearance of a new start site, located at an A residue at position +8 relative to the natural U2 cap site (+10 relative to RNA polymerase II initiation in pU2/−247/RA.2). In pU2/TA, this new start site represents 80% or more of the total signal (lane 6), and in pU2/TA/U2Cap it represents about 50% of the signal (lane 2). The downstream start site is the only one used *in vitro* (lanes 3, 4, and 7), suggesting that it is directed by RNA polymerase III since RNA polymerase II snRNA promoters are not active in our extracts (13). Indeed, the downstream start site in the pU2/TA/U2cap construct is resistant to low, but not high concentrations of α-amanitin *in vitro*, whereas transcription from an α-globin gene control is as expected inhibited by low concentrations of α-amanitin (Figure 2C lanes 2−4). Thus, although the U2 cap site sequences influence the relative levels of RNA polymerase II and III transcription, the U6 A/T box is nevertheless the dominant element.

RNA polymerase III transcription complexes derived from a U6 promoter or from pU2/TA read through the 3′ box and terminate transcription in a run of T residues located downstream (13). To determine whether the same is true for the new construct pU2/TA/U2cap, RNA was analyzed by RNase T_1 protection of

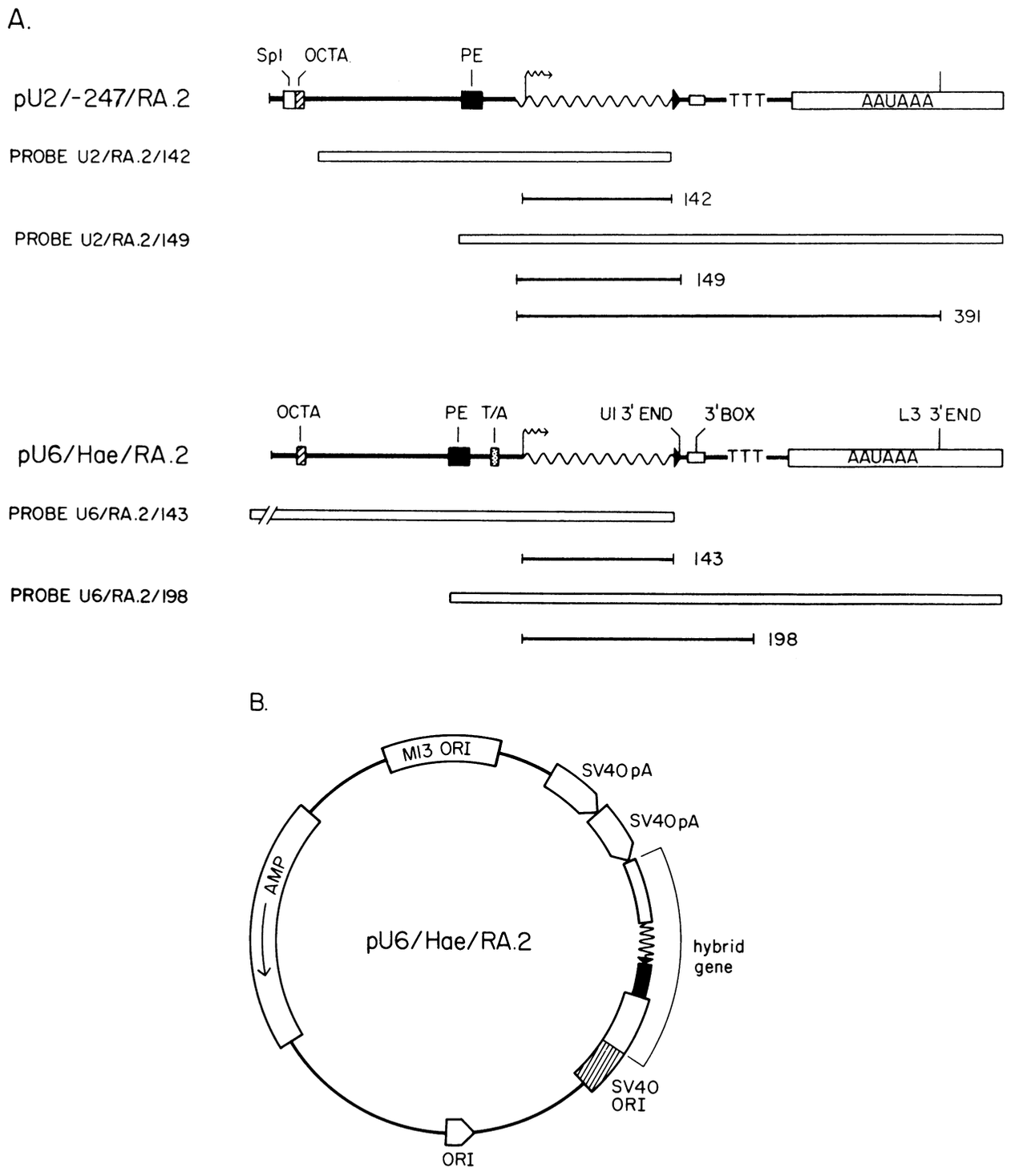

Figure 1. A: Structures of the U2 and U6 hybrid genes, the RNA antisense probes, and the expected protected fragments. The pU2/−247/RA.2 insert consists of U2 5′ flanking sequences from −247 to −7 (thick line), a fragment derived from the rabbit β-globin gene and cloned in the reverse orientation (wavy line), the last six nucleotides (nt) of the U1 coding region (arrowhead), 90 nt of U1 3′ flanking sequences containing the U1 3′ box (small box) and a run of 6 Ts (labeled 'TTT'), and a fragment containing the adenovirus L3 polyadenylation site (open box). The U6 hybrid gene is similar except that the U2 5′ flanking region is replaced by U6 5′ flanking sequences from −241 to +1. The thin boxes represent antisense RNA probes, and the lines represent expected protected fragments. B: Structure of the vector carrying the U2 and U6 hybrid genes. The region corresponding to the hybrid genes as shown in A is indicated by a bracket.

an antisense RNA probe extending from upstream of the initiation site of transcription to downstream of the last 3′ end formation signal, the Ad2 L3 polyadenylation site (probe U2/RA.2/149, see Figure 1A). The results are shown in Figure 3. The U2 parent constructs give rise predominantly to a doublet diagnostic of termination of transcription at the 3′ box (lanes 1 and 3). However, the constructs containing the A/T box give rise to a new band, which corresponds to RNA initiated at position +7 downstream of the normal U2 cap site and terminated at the run of 6 Ts located downstream of the U1 3′ box (lanes 2 and 4).

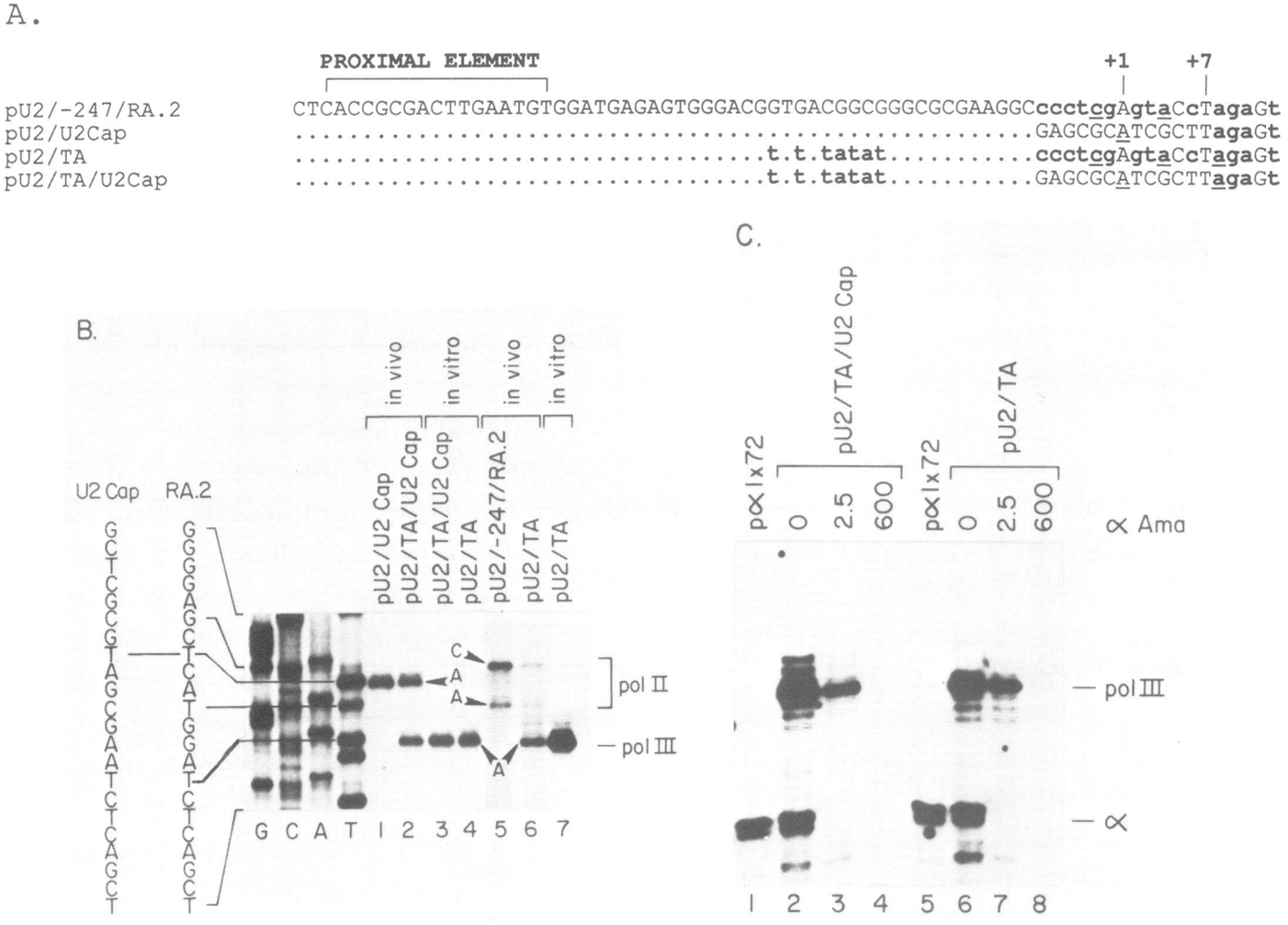

Figure 2. A: 5′flanking sequence in different U2 constructs. The nt that differ from the wild type U2 sequence are represented by bold lower case characters. The different start sites of transcription are indicated by underlined characters. B: The indicated constructs were incubated in an HeLa cell extract under transcription conditions (lanes labeled 'in vitro') or transfected into HeLa cells together with pSVEori-, a SV40 large T Antigen expressing plasmid (lanes labeled 'in vivo'). RNA was analyzed by primer extension with a primer hybridizing to the β-globin portion of the hybrid genes. The sequencing ladders were obtained by extension of the same primer on the pU2/−247/RA.2 template in the presence of dideoxynucleotides, and the sequence of the non-RNA strand is indicated on the left of the panel (labeled RA.2), together with the sequence of the pU2/U2Cap construct (labeled U2Cap). The primer extension bands are labeled with the corresponding residue on the RNA strand. C: The indicated constructs together with a plasmid carrying the α-globin gene (pα1×72) to serve as a control for transcription by RNA polymerase II were incubated in a HeLa nuclear extract in the presence of 0, 2.5, and 600 μg/ml of α-amanitin, as indicated. RNA was analyzed by primer extension with a mixture of two primers, one hybridizing to the β-globin sequences of the hybrid RNA and one complementary to the α-globin RNA. In lanes 1 and 5, the α-globin gene was the only template in the *in vitro* reaction.

In pU2/TA/U2cap (lane 4), some RNA still ends at the 3′ box, confirming that this mutant directs some RNA polymerase II transcription.

The 3′ part of the U6 proximal element and the region immediately downstream are involved in localizing the start site of transcription

Previous studies have shown that in the human U1 and Xenopus U6 snRNA genes, the proximal element is involved in localizing the start site of transcription (16, 21, 22). In our previous linker scanning analysis of the U6 promoter we noticed that some mutants (LS4 and LS3) in which the downstream portion of the proximal element was mutated gave rise to a weak aberrant trancriptional start site around position −16 (13). This suggested that the downstream portion of the proximal element might harbor the localization function. However, it was unclear whether the −16 start site was dependent on snRNA promoter elements and whether it was directed by RNA polymerase II or III. To characterize this start site further, we analyzed expression from the clustered point mutations shown in Figure 4A. LS4 mutates 6 nucleotides in the 3′ part of the proximal element whereas LS3 mutates 5 nucleotides in the 3′ part of the proximal element and the 6 adjacent nucleotides immediately downstream, thus modifying a region of the U6 5′ flanking sequence which is conserved in Saccharomyces cerevisiae (15). LS3b mutates exactly the same region as LS3, but the substituted sequence is different. Because the 16 start site is barely detectable in transfections performed in the absence of the SV40 large T antigen (see Figure 3B in (13)), the different constructs were transfected into HeLa cells together with an SV40 large T antigen expression vector. Under these conditions, the transfection assay was saturated for U6 expression, and thus the levels of expression of mutant U6 promoters appeared artificially high (13). However, this facilitated the analysis of the −16 start site. The RNA was

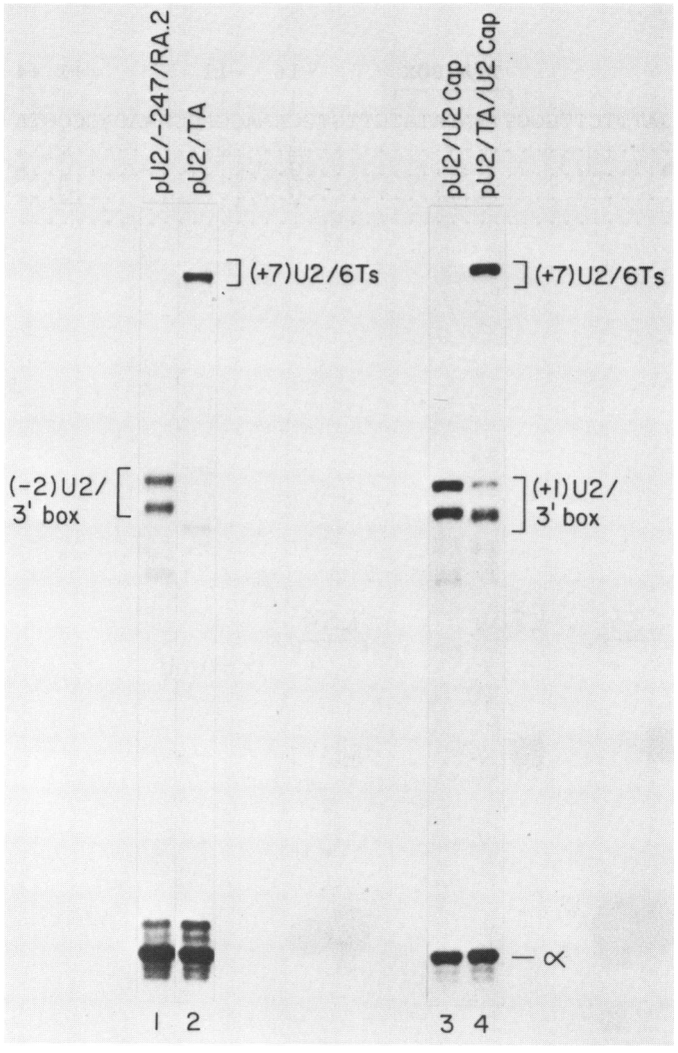

Figure 3. RNA from cells transfected with the indicated constructs and the plasmids pSVEori and pα1×72 as an internal control for transfection efficiency and RNA recovery was analyzed by RNase T_1 protection of a mixture of two probes: probe U2/RA.2/149 (see Figure 1A) (or the corresponding probe derived from the construct pU2/U2Cap in lanes 3 and 4) and probe α98, which protects correctly initiated RNA derived from the α-globin internal control over 98 nt. The different bands are labeled by the initiation and termination sites of the corresponding RNAs. For example, the bands labeled (−2)U2/3'box are initiated at position −2 relative to the wild type U2 cap site and terminated at the 3' box.

analyzed by RNase T_1 protection of the antisense RNA probe U6/RA.2/143 (Figure 1A), which extends from upstream of the start site of transcription to within the hybrid gene. As shown in Figure 4B (lanes 2−4), these three constructs all generate a start site around position −16, although with different efficiencies. LS3b is the mutant that produces the highest relative levels of these 5' elongated molecules, while correct initiation is significantly reduced.

To determine whether initiation around the −16 start site depends on the same promoter elements as initiation at the correct start site, we truncated the LS3 promoter mutant at an NdeI site just upstream of the proximal element, thus removing the whole enhancer region. Both the correct and the −16 start sites were greatly reduced (lane 5), indicating that like the +1 start site, the −16 start site is dependent on upstream elements. Analysis of the mutant LS24 (Figure 4A), in which the upstream region of the proximal element is mutated in the context of the LS4 mutation, indicated that the −16 start site was dependent on the proximal element (data not shown).

To evaluate the contributions of RNA polymerase II and III to the start sites around position −16, HeLa cells transfected with LS3, LS3b and LS4 were incubated in medium containing 50 μg per ml of α-amanitin, a concentration which inhibits RNA polymerase II, but not RNA polymerase III, in whole cells (23). RNA was collected 5h later and analyzed by primer extension. This method allows the precise mapping of the different start sites, in contrast to RNase T_1 mapping whose resolution is limited by the enzyme specificity for cleavage at G residues. As shown in Figure 4C, the 16 start site in all the mutants is an RNA polymerase II start site, because it is suppressed in the presence of α-amanitin (lanes 2, 4, and 6). LS3 directed in addition a minor, α-amanitin resistant, start site at position 11, which thus probably corresponds to an RNA polymerase III start site (compare lanes 1 and 2). In LS3b, initiation at +1 was strongly reduced, and an new RNA polymerase III start site appeared at position +4 (lanes 3 and 4).

In all the mutants analyzed so far (see Figure 3 and (13)), transcription by RNA polymerase III correlated with termination at a run of Ts, whereas transcription by RNA polymerase II correlated with termination at the 3' box. Analysis of RNA derived from the LS3, LS3b and LS4 mutants with the riboprobe U6/RA.2/198 (Figure 1A) confirmed that the same holds true for the minor −16 and −11 start sites. As shown in Figure 5, lane 1, the mutant LS3 gives rise to a doublet and a single band whose sizes suggested that they corresponded to RNA initiated at least ten nt upstream of the correct U6 start site and terminated at the 3' box (labeled (−16)U6/3'box) and the run of Ts (labeled (−11)U6/6Ts), respectively. Indeed, the doublet (−16)U6/3'box disappeared with α-amanitin (lane 2) and was more intense in the mutant LS3b (lane 4), thus correlating completely with the 16 signal observed in the primer extensions (Figure 4C), whereas the (−11)U6/6Ts band was resistant to αamanitin (lane 2) and was absent or very weak in the LS3b and LS4 mutants (lanes 4 and 5), thus correlating with the −11 signal in the primer extension (Figure 4C). In conclusion, mutations encompassing the downstream portion of the proximal element and the sequence immediately downstream generate new aberrant RNA polymerase II and RNA polymerase III start sites. This region may therefore be involved in localizing both the RNA polymerase II and III transcription start sites. As shown below, mutations encompassing only the region immediately downstream of the proximal element also affect the efficiency of RNA polymerase II transcription from the correct start site.

The region immediately downstream of the proximal element homology is required for efficient RNA polymerase II, but not for RNA polymerase III transcription from the normal start site

In the human U2 promoter, a linker scanning mutation that modified mainly the region directly downstream of the proximal element homology and overlapped with the proximal element homology itself only by one bp (Figure 6A, mutant LS−34/−43) was as detrimental to transcription as mutations that modified the central part of the proximal element homology (14). In addition, a mutation offset by one nucleotide towards the 3' side and therefore not affecting the proximal element homology (Figure 6A, mutant LS−33/−42) suppressed RNA polymerase II transcription in a promoter lacking sequences between the enhancer region and the proximal element, although it had no

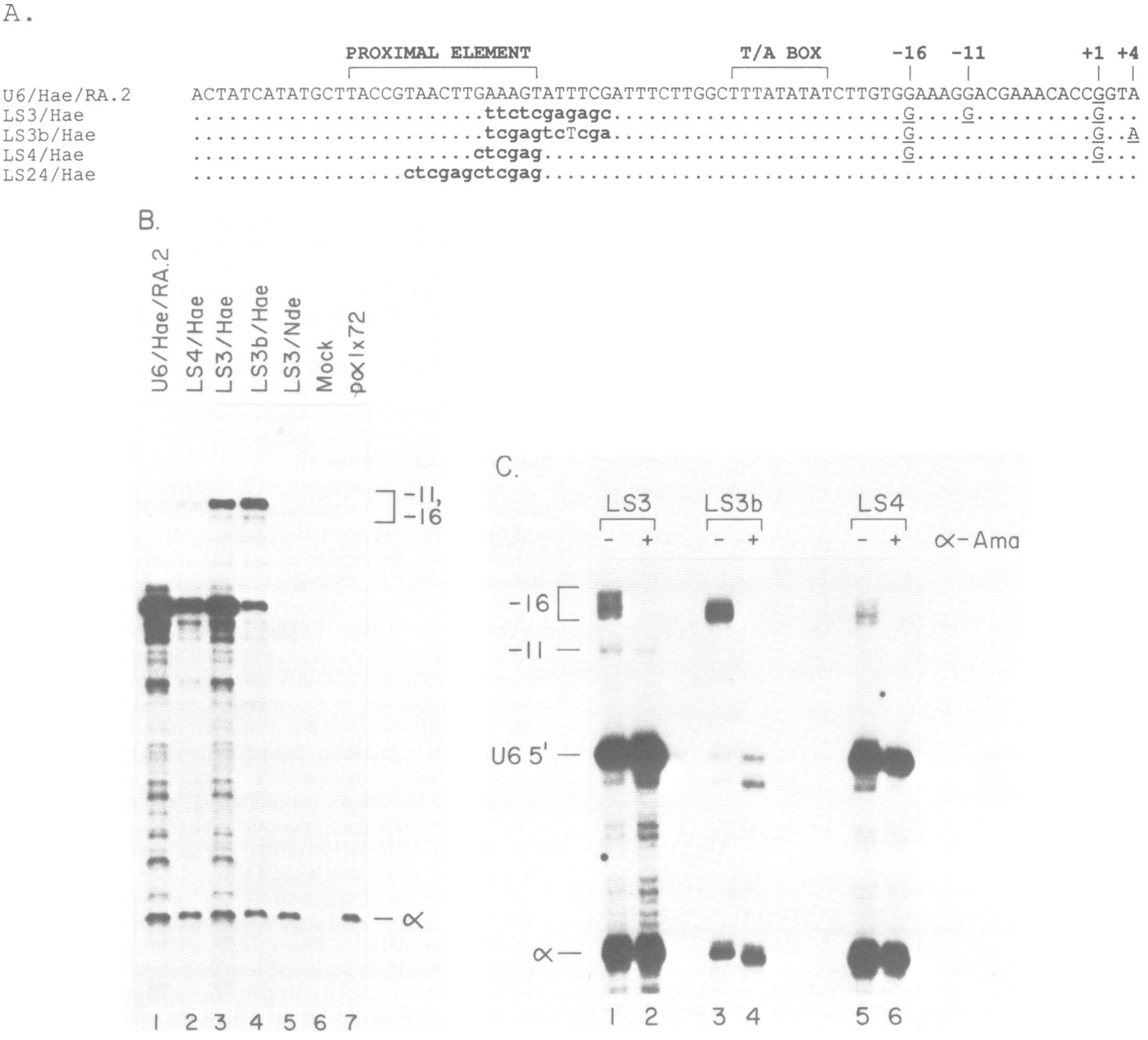

Figure 4. A. 5′ flanking sequences in different U6 constructs. The sequences that differ from the wild type sequence are indicated by bold lower case characters. The proximal element and the A/T box are indicated. The start sites of transcription directed by the different constructs are indicated by underlined characters. B. RNA from HeLa cells transfected with the indicated constructs together with the pSVEori plasmid and the pα1×72 internal control, or with the pSVEori and pα1×72 constructs alone (lane 7) was analyzed by RNase T_1 protection with the probes α98 and U6/RA.2/143. The band corresponding to start sites at positions −11 and −16 are indicated. The major band corresponds to RNA initiated at +1. C: HeLa cells were transfected with the indicated constructs and the constructs pSVEori- and pα1×72. 43h after transfection, half of the cells was collected, and the other half was incubated for another 5 h in the presence of 50 μg of α-amanitin per ml of medium. RNA was analyzed by primer extension with a mixture of two primers as in Figure 2C. Note that the α-globin RNA is stable over 5h, and therefore the α-globin signal is not diminished by the α-amanitin treatment.

effect in the context of the wild type U2 promoter (14). These observations suggested that the region immediately downstream of the proximal element contained some element required for RNA polymerase II transcription, although this region is not conserved in different RNA polymerase II snRNA promoters. In the human U6 promoter, a clustered point mutation that modified essentially the same region (LS5, see Figure 6A) without modifying the proximal element had little effect on RNA polymerase III transcription. However, when combined with the LS7 mutation, which mutates the A/T box and switches the U6 promoter to a predominantly RNA polymerase II promoter, the LS5 mutation had a marked deleterious effect, again suggesting that this region was required for RNA polymerase II transcription (13). However, because the mutations were different and were introduced in the context of different promoters, the above data could not be rigorously interpreted.

To determine whether the region downstream of the proximal element represents an element required mainly for RNA polymerase II transcription, we introduced the LS−34/−43 mutation in the context of the pU2/TA promoter which directs

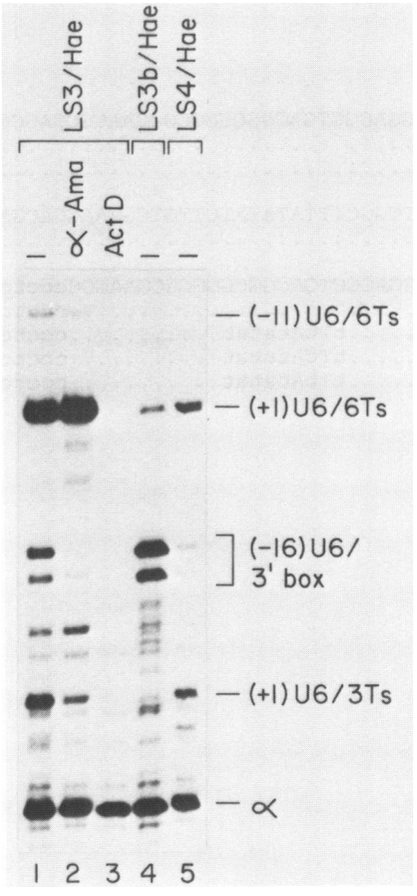

Figure 5. HeLa cells were transfected with the indicated constructs and pα1×72. In the case of the LS3/Hae construct, the cells were then divided into three batches. The first one was collected 43h after transfection, the second one was incubated for another 5h in medium containing 50 μg per ml of α-amanitin, and the third one was incubated in medium containing 2.5 μg per ml of actinomycin D, a concentration that inhibits all transcription. The RNA was analyzed by RNase T_1 protection with the probes α132 and U6/RA.2/198 (see Figure 1A). α132 protects correctly initiated RNA derived from the α-globin gene internal control over 132 nt. The bands are labeled by the sites at which the corresponding RNAs initiate and terminate. For example, the band labeled (−16)U6/3'box corresponds to RNA initiated at position −16 upstream of the normal U6 start site of transcription and terminated at the 3' box.

both RNA polymerase II and RNA polymerase III transcription, thus creating the pU2/TA/(LS−34/−43) construct (Figure 6A). This construct was transfected into HeLa cells and RNA was analyzed by RNase T_1 protection of the antisense RNA probe U2/RA.2/142 (Figure 1A). As shown in Figure 6B, lane 1, RNA derived from the parent construct pU2/−247/RA.2 gives rise to a protected doublet, which corresponds to RNA synthesized by RNA polymerase II (Figures 2B and 3). Introduction of the U6 A/T box resulted in the appearance of a second protected doublet (lane 3), which corresponds to the RNA polymerase III start site at position +8 (Figure 2C). Introduction of the mutation LS−34/−43 reduced RNA polymerase II transcription to undetectable levels, whereas RNA polymerase III transcription is only slightly affected (lane 4). The mutation LS4, which changes the proximal element itself, severely reduces both RNA polymerase II and III transcription, as expected (lanes 2 and 5). Thus, at least one element in the U2 and U6 promoters appears to be required by RNA polymerase II but not by RNA polymerase III.

DISCUSSION

The U2 and U6 promoters represent a unique system in which the determinants of RNA polymerase specificity can be studied. These determinants are located within a compact, 64 nucleotide long region, extending from upstream of the proximal element homology to the start site of transcription. It is therefore of interest to dissect this region in detail. We have previously shown that the U6 A/T box represents a dominant element, which when inserted into a U2 promoter extending from −247 to −7 upstream of the U2 cap site changes the U2 promoter specificity from RNA polymerase II to RNA polymerase III (13). In the Xenopus U6 promoter, transcription by RNA polymerase III is extremely sensitive to changes near the start site of transcription (16). This observation led us to ask whether the presence of the U2 cap site region in a U2 promoter with the U6 A/T box might abolish transcription by RNA polymerase III. We found that restoring the natural U2 sequence from positions −7 to +6 did not abolish activation of RNA polymerase III transcription by insertion of the A/T box. In this construct, RNA polymerase III transcription still represented 50% of the total transcription, demonstrating that the U2 cap site cannot counteract the effect of the U6 A/T box, which therefore is indeed a dominant element. The effects of nucleotides at the RNA polymerase III start site of transcription itself were not examined in detail. However, we note that in all our mutant promoters, as in the Xenopus U6 promoter (16) and other RNA polymerase III promoters (24), RNA polymerase III transcription starts at purines: in the LS3b mutant the minor −11 start site is at a G residue, and in the U2 promoters with a U6 A/T box inserted, RNA polymerase III transcription starts with an A residue. This suggests that replacing the U6 start site of transcription by a pyrimidine might be deleterious to RNA polymerase III transcription.

The mutations LS3, LS3b, and LS4 cause the appearance of new start sites at positions −16 and −11. These start sites are weak, and can be easily examined only when the levels of expression are increased by inclusion of a SV40 large T Antigen expressing vector in the transfections. Under these conditions, we could determine that the −16 start site was an RNA polymerase II start site whereas the −11 start site was an RNA polymerase III start site. Both are of the snRNA type, in that they are dependent on the central part of the proximal element and the RNA polymerase II RNAs initiated at −16 end at the 3' box. The proximal elements of the human U1 (21,22) and Xenopus U6 (16) snRNA genes have been shown to be involved in localizing the start site of transcription. The above results suggest that this localization function may lie in the downstream part of the proximal element. The main function of this region is, however, to increase transcriptional efficiency, since in the absence of SV40 large T Antigen, transcription from positions +1, −11 and −16 is very low in the LS3, LS3b, and LS4 mutants (13, and data not shown). It will be interesting to determine whether the corresponding conserved region in the 5' flank of the Saccharomyces cerevisiae U6 gene performs similar functions. It is remarkable that the RNA polymerase II start sites are always located a few nucleotides upstream from the RNA polymerase III start sites. Thus, in the human U6 promoter, mutation of the A/T box activates RNA polymerase II start sites 3 and 4 nucleotides upstream of the RNA polymerase III start site (13). The RNA polymerase II 16 start site is 5 nucleotides upstream of the RNA polymerase III 11 start site, and in the U2 promoter containing the U6 A/T box, transcription by RNA

A.

```
                                  PROXIMAL ELEMENT
                                  ┌──────────────┐
pU2/-247          GGC TGGGGCTCTCACCGCGACTTGAATGTGGATGAGAGTGGGACGGTGACGGCGGGCGCGAAGGCGAGCGCATCGCT
LS-34/-43         ... ..........................ccctc...g.....................................
LS-33/-42         ... ..........................ccc.cgag.g....................................

U6/Hae/RA.2       ACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG
LS5/Hae           ..............................tctcgagaa.....................................G

pU2/-247/RA.2     GGC TGGGGCTCTCACCGCGACTTGAATGTGGATGAGAGTGGGACGGTGACGGCGGGCGCGAAGGCccctcgAgtaCcTa
pU2/LS4/RA.2      ... ....................ctcgag.....................................ccctcg.gta.c.a
pU2/TA            ... .........................................tTtAtatat...........ccctcg.gta.c.a
pU2/TA/(LS-34/-43) ... ........................ccctc...g........tTtAtatat...........ccctcg.gta.c.a
pU2/TA/LS4        ... ....................ctcgag................tTtAtatat...........ccctcg.gta.c.a
```

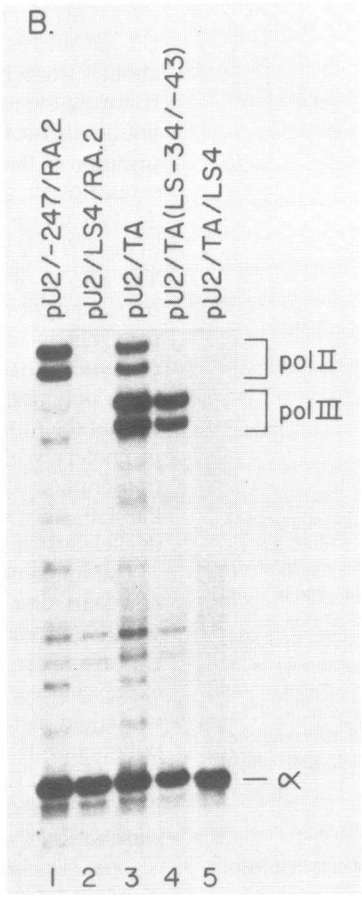

Figure 6. A: 5′ flanking sequences in different U2 and U6 constructs. The nt differing from the wild type sequence are indicated in bold lower case characters. Start sites of transcription are indicated by underlined characters. The proximal element is indicated. B: RNA from cells transfected with the different constructs and pα1×72 was analyzed by RNase T_1 protection mapping with the probes α98 and U2/RA.2/142. Because of slight overdigestion with RNase T1, doublets of bands rather than single bands were obtained in this experiment.

polymerase II again starts 9 nucleotides upstream from transcription by RNA polymerase III. Perhaps this reflects a difference in the size or the spatial arrangement of the RNA polymerase II and III subunits that contact the DNA.

Our analysis of the pU2/TA(LS−34/−43) construct demonstrates that mutation of the region downstream of the U2 proximal element has little effect on RNA polymerase III transcription but suppresses RNA polymerase II transcription. The corresponding region in the U6 promoter is functionally equivalent, since a mutation in this region (mutation LS5) has little effect on RNA polymerase III transcription but severely inhibits RNA polymerase II transcription from a U6 promoter in which the A/T box is mutated (13). Surprisingly, this promoter region is conserved neither among mammalian RNA polymerase II snRNA promoters nor between the U2 and U6 promoters. This region may thus bind different RNA polymerase II transcription factors of similar function, or a single factor with a wide DNA sequence recognition flexibility. Examples of transcription factors with such DNA binding flexibility include the Oct-1 and HAP1 proteins (25, 26). The absence of RNA polymerase II specific factors in nuclear HeLa cell extracts might explain the apparent paradox of active RNA polymerase III but inactive RNA polymerase II transcription from snRNA promoters *in vitro*. A third possibility, however, is that this region binds RNA

polymerase II itself. Binding of RNA polymerase II to the DNA template may be compatible with a large number of, but not with all sequences. Perhaps the substitutions in LS−34/−43 and in LS5, which are quite similar to each other, are detrimental for binding of RNA polymerase II, but not RNA polymerase III.

Although mutations downstream of the U6 proximal element mainly affect RNA polymerase II transcription (in a construct in which the A/T box is debilitated), the mutation LS3b suggests that there is a certain sequence constraint in this region for transcription by RNA polymerase III. The LS3b mutation extends the LS4 mutation by 6 nucleotides downstream of the proximal element, and as revealed by the analysis in the presence of SV40 large T Antigen (Figure 4), is more deleterious to RNA polymerase III transcription from the +1 start site than LS4. It remains, however, that transcription by RNA polymerase II and III is affected differently by mutations in this region. Why, then, does the U6 gene contain a sequence compatible with RNA polymerase II transcription in this region? This element may represent a remnant from when the snRNA promoters may have been transcribed by the same RNA polymerase. Alternatively, there may be slightly divergent U6 gene copies in the human genome that are transcribed by RNA polymerase II. Little is known about the genomic organization and the copy number of the true U6 genes in the human genome, although the total number of U6 loci including pseudogenes has been estimated to be 200 (27). Finally, the U6 gene we have analyzed may itself be transcribed by RNA polymerase II under circumstances not reproduced in our assay.

ACKNOWLEDGEMENTS

We thank W. Herr for comments on the manuscript and L. Johal for technical assistance in the later stages of this work. We also thank M. Ockler, J. Duffy, and P. Renna for preparing the figures. This work was supported by grant RO1 GM38801 from the National Institutes of Health. N.H. is a Rita Allen Foundation Scholar.

REFERENCES

1. Geiduschek, E.P., and Tocchini-Valentini G.,P. (1988). Ann. Rev. Biochem. 57, 873−914.
2. Das, G., Henning, D., Wright, D., and Reddy, R. (1988). EMBO J. 7, 503−512
3. Murphy, S., Di Liegro, C., Melli, M. (1987). Cell 51, 81−87.
4. Chang, D. D., and Clayton, D. A. (1989). Cell 56, 131−139.
5. Krol, A., Carbon, P., Ebel, J.-P., Appel B. (1987). Nucl. Acids Res. 15, 2463−2477
6. Kunkel, G. R., and Pederson, T. (1988). Genes and Development 2, 196−204.
7. Dahlberg, J. E., and Lund, E. (1988). In: Structure and function of major and minor small nuclear ribonucleoprotein particles, M. Birnstiel, ed. (Heidelberg: Springer-Verlag), pp 38−70.
8. Carbon, P., Murgo, SW., Ebel, J.-P., Krol, A., Tebb, G., and Mattaj, I. W. (1987). Cell 51, 7179.
9. Mattaj, I. W., Lienhard, S., Jiricny, J., and De Robertis, E. M. (1985). Nature 316, 163−167.
10. Ares, M. Jr., Mangin, M., and Weiner, A.M. (1985). Mol. and Cell. Biology, 5, 1560−1570.
11. Mangin, M., Ares, M. Jr., and Weiner, A.M. (1986). EMBO J. 5, 987−99.
12. Bark, C., Weller, P., Zabielski, J., Janson, L., and Pettersson, U. (1987). Nature 328, 356−359.
13. Lobo, S.M., and Hernandez, N. (1989). Cell 58, 55−67.
14. Hernandez, N. and Lucito, R. (1988). EMBO J. 7, 3125−3134.
15. Brow, D.A., and Guthrie, C. (1988). Nature 334, 213−218.
16. Mattaj, I.W., Dathan, N.A., Parry, H.D., Carbon, P. and Krol, A. (1988). Cell 55, 435−442.
17. Hernandez, N. (1985). EMBO J. 4, 1827−1837.
18. Yuo, C., Ares, M., Jr., and Weiner, A.M. (1985). Cell 42, 193−202.
19. Neuman de Vegvar, H. E., Lund, E., and Dahlberg, J. E. (1986). Cell 47, 259−266.
20. Hernandez, N. and Weiner, A. M. (1986). Cell 47, 249−258.
21. Skuzeski, J. M., Lund, E., Murphy, J. T., Steinberg, T. H., Burgess, R.R., and Dahlberg, J.E. (1984). J. Biol. Chem. 259, 8345−8352.
22. Murphy, J.T., Skuzeski, J.T., Lund, E., Steinberg, T.H., Burgess, R.R., and Dahlberg, J.E. (1987). J. Biol. Chem. 262, 1795−1803.
23. Zieve, G., Benecke, B.-J., and Penman, S. (1977). Biochemistry 16, 4520−4525.
24. Sakonju, S., Bogenhagen, D.F., and Brown, D.D. (1980). Cell 19, 13−25.
25. Baumruker, T., Sturm, R., and Herr, W. (1988). Genes & Development 2, 1400−1413.
26. Pfeifer, K., Prezant, T., and Guarente, L. (1987). Cell 49, 19−27.
27. Hayashi, K. (1981). Nucl. Acids Res. 14, 3379−3388.