The different positioning of the proximal sequence element in the Xenopus RNA polymerase II and III snRNA promoters is a key determinant which confers RNA polymerase III specificity

Alain Lescure, Philippe Carbon and Alain Krol*
Laboratoire de Biochimie, Institut de Biologie Moleculaire et Cellulaire du CNRS, 15 rue Rene Descartes, 67084 Strasbourg Cedex, France

Received November 27, 1990; Revised and Accepted January 3, 1991

ABSTRACT

We and others have previously described the TATA motif as a major determinant for Pol III specificity of the U6 promoter. Surprisingly, however, the data documented here show that the sole introduction of a TATA sequence into a U1 Pol II snRNA gene is not sufficient to confer Pol III transcription. Rather, this promoter element can mediate optimal Pol III transcription only if the PSE, the second promoter element, is shifted 4 bp upstream of the position it occupies in Pol II snRNA genes. As a result, the PSE-TATA-start site spacing introduced into the U1 Pol II gene is identical to that of the U6 gene and is strictly required to produce properly initiated Pol III transcripts. Thus, Pol II and Pol III PSEs, although similar in sequence, are not positionally equivalent. Competitive experiments raise the possibility that vertebrate U6 genes contain other, as yet unidentified, promoter elements.

INTRODUCTION

Most of the spliceosomal RNAs, those small nuclear RNAs (snRNAs) that participate in pre-mRNA splicing, are transcribed by an RNA polymerase that displays the same α-amanitin sensitivity as RNA polymerase II (Pol II), which transcribes protein-coding genes. They constitute, however, a specialized class of RNA polymerase II transcription units. In vertebrates, two upstream motifs, the Distal Sequence Element (DSE) and the Proximal Sequence Element (PSE), contribute to the expression of spliceosomal U1-U5 snRNA genes. The DSE is an enhancer-like element for U snRNA transcription, which normally is located 220−250 bp upstream of the initiation site and contains, usually in combination with other motifs, one copy of the octamer sequence ATGCAAAT. The PSE, centered on −55, is required for accurate start-site selection. Starting between 14 and16 bp downstream of the coding region exists a consensus sequence, called the 3'box, that is used as a termination signal for Pol II transcription (see 1,2 for a review on U snRNA transcription). In addition to these cis-element dissimilarities, the transcription complex made on U snRNA genes differs from that made on mRNA genes at least in some of its constituent transcription factors (3).

U6 snRNA, another spliceosomal RNA, and 7SK RNA, a functionally unrelated species, are members of a novel class of RNA polymerase III (Pol III) genes which differs from that constituted by tRNA and 5S RNA genes in that the promoter elements of U6 and 7SK genes are external to the coding region (4−8; reviewed in 9,10). Extensive work has been carried out by several laboratories on U6 genes from various vertebrates. U6 has a functional DSE and an essential PSE-like sequence (5,11−15) similarly to Pol II snRNA genes. In contrast to these latter, it possesses a TATA-like motif around −30, as do mRNA genes in which this sequence provides a binding site for the transcription factor TFIID (reviewed in 16). Surprisingly, however, the TATA-like sequence occurring in U6 genes was shown to be a major determinant conferring RNA polymerase III specificity to the U6 promoter (14,15,17). In a recent contribution, Simmen and Mattaj (18) showed that the sequence requirements of the TATA motif are different whether it functions in the context of Pol II mRNA or Pol III snRNA promoters, strengthening the possibility that U6 transcription is mediated through a Pol III-specific TATA-binding factor.

By compiling the extent of sequence similarities between PSEs contained in Pol II and Pol III PSE-containing genes, we made the observation that this element is not positionally conserved between the genes in question. This led us to reinvestigate the separate contributions of the PSE and TATA elements to Pol III specificity of a U6 snRNA promoter. This was approached by analyzing in vivo the effects of sequentially modifying or adding cis-elements into a Xenopus U1 Pol II snRNA gene in order to induce its promoter to be efficiently recognized by RNA polymerase III. Here, we show that the sole introduction of a TATA sequence in a position which keeps the distance from the start site equal to that of U6 genes is not sufficient to confer Pol III specificity to a U1 snRNA gene. Rather, the combined effects of an upstream shift of the PSE element (relative to its position

* To whom correspondence should be addressed

in a Pol II snRNA gene) with the presence of a TATA element lead to optimal Pol III transcription of U snRNA promoters conserving the natural transcriptional start point.

MATERIALS AND METHODS

Constructs

L1 construct containing the X. laevis U1 gene was derived by subcloning the 696 bp *Pst* I fragment containing the wt U1B2 (19) into pBluescribe (−) (Stratagene). Site-directed mutagenesis was performed using the in vitro kit from Amersham. The presence of the desired mutation was checked by sequencing the single-stranded DNAs. L18 construct was obtained as follows. The 977 bp *Bam* HI fragment containing the X. tropicalis U6 gene (4) was digested with *Fsp* I and the resulting 770 bp *Bam* HI/*Fsp* I fragment was ligated to *Bam* HI/*Sma* I cut pBS (+) vector. A 205 bp *Sac* I fragment originating from the X. laevis U1 minigene and containing the U1 termination signal, was introduced into the *Sac* I site of the preceding construct, downstream of the U6 coding region.

Oocyte microinjection

X. laevis oocyte nuclei were microinjected with 20 nl of DNA at 600 μg/ml. In coinjection experiments, concentrations of the L11 or L28 mutants and C1 competitor were 300 μg/ml each (1:1 ratio). The 5S maxigene (20) was added at a concentration of 5 μg/ml as an internal standard for microinjection and RNA recovery. In the Pol II inhibition experiments, α-amanitin was added to the DNA at a concentration of 2 μg/ml. Transcription efficiencies of the various templates were normalized relative to 5S maxi : scintillation counting and scanning with a Shimadzu CS-9000 densitometer were performed with exposures different from those displayed in the Figures (non-saturating transcripts).

RESULTS

The sole introduction of a TATA sequence in an otherwise Pol II snRNA gene does not confer Pol III specificity to the promoter

The rationale of the experiments was to analyze the effects of introducing into a Pol II U snRNA gene the TATA motif, an element that has been shown to be determinant for conferring Pol III specificity to the Xenopus U6 promoter (14). We chose the X. laevis U1B2 snRNA gene (19,21), now called L1 construct (Figure 1A, lane 1), which is transcribed by RNA polymerase II as the starting template to be converted into a Pol III transcript. In a first step, the U6 TTATAA sequence (positions −30/−25) was introduced to substitute these same positions of the U1 wt sequence (Figure 1A, construct L3). If introduction of a TATA sequence results in the Pol III transcription of U1, then we can expect the transcript to end at a natural run of 6 Ts occurring at position 74 downstream of the mature U1 3′end (5 or more Ts in a row constitute an efficient Pol III terminator, ref. 22). Injection of L3 into Xenopus oocyte nuclei (Figure 1B, lane 2) shows that no such product appears. The transcript observed in lane 2 has the same mobility as U1 wt (lane 1) which uses the 3′box signal for termination of Pol II transcription (1). In order to eliminate the possibility that the external run of Ts is ignored because of the presence of the 3′box, a run of 5 Ts was created by introducing a point mutation (C130 to T130) into the Sm binding site of the U1 gene (Figure 1A, construct L6). Figure 1B lane 3 shows that L6 is not able to drive Pol III transcription either, since the product ending at the engineered Pol III terminator should be shorter than U1 wt. Knowing that efficient Pol III transcription of the Xenopus U6 gene demands a G as the start point (14), the normal U1 start site which is an A was replaced by a G, giving rise to construct L7 (Figure 1A, lane 4). This template now possesses all the elements that were shown to be important for Pol III transcription of the U6 gene. Again, Figure 1B lane 4 only shows the persistence of a transcript migrating at the level of U1 wt, but not of a shorter Pol III transcript. The Pol II nature of the transcript seen in lane 4 is confirmed by an injection experiment with an α-amanitin concentration inhibiting RNA polymerase II transcription (lane 5). The identity of the bases upstream of the U1 start site is not likely to be responsible for the lack of Pol III trancription since they are constituted by pyrimidines in both U1 and U6 genes. In fact, over the 23 bp extending from positions −13 to +9 of both genes, 48% of strict sequence similarity is found (4, 19). This figure is as high as 86% if one also includes transitions.

The simple introduction of the Xenopus U6 TATA sequence into a U1 gene, in a position identical to that of U6, does not lead to Pol III specificity of the U1 promoter. Accurate densitometric analysis of the transcripts shown in Figure 1B indicates that a 50% inhibition of Pol II transcription occurs when constructs carry a TATA sequence. This is not due to instability of the transcripts containing an altered Sm binding site

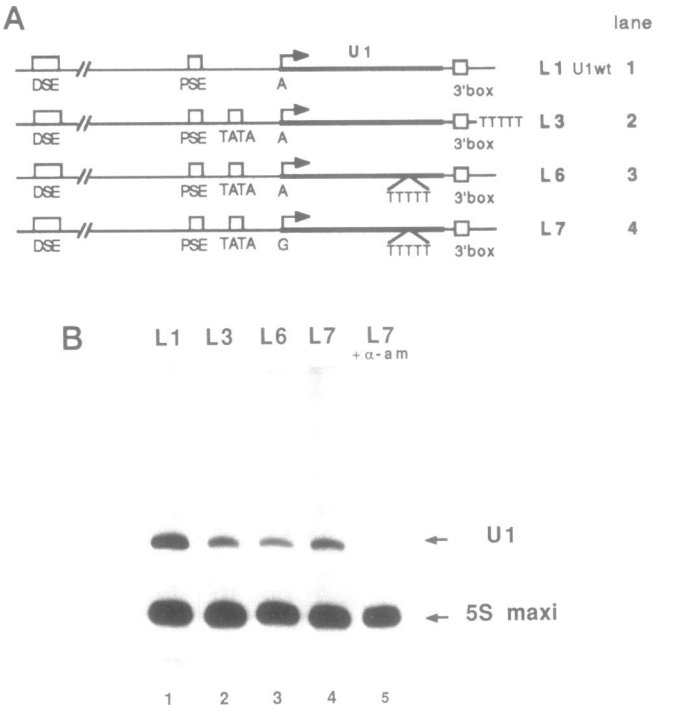

Figure 1. Template activity of the TATA sequence-containing U1 snRNA gene promoter. **(A)** Structure of the templates used in this analysis. DSE, PSE, TATA and 3′box elements are represented by open boxes. In construct L3, the string of Ts represents the 6 Ts naturally occurring in the DNA. In constructs L6 and L7, the 5 Ts were engineered. In all three cases, they represent the location of the RNA polymerase III terminator. A heavy line represents the U1 coding region, an arrow the start of transcription which is an A or a G. Drawings are not to scale. **(B)** Injection into Xenopus oocyte nuclei (see Materials and Methods) of the templates shown in (A) with (α-^{32}P) GTP (Amersham, 400 Ci/mmol) and 5S maxigene (20) as an internal control. Lanes 1 to 4 correspond to constructs shown in (A). In lane 7, construct L7 was injected in the presence of α-amanitin (+ α-am). Transcripts (RNA from 2 oocyte equivalents) were analyzed on 10% sequencing-type gels. Positions of the U1 and 5S maxi transcripts are indicated.

(where the C to T transition creates the run of 5 Ts), since the drop in transcription (normalized relative to 5S maxi transcription) is observed wherever the location of the Pol III terminator. Although we cannot interpret these results unequivocally, a steric hindrance between a PSE factor and a TATA factor resulting from too close a distance between the PSE and TATA cis-elements could provide a straightforward explanation.

A Pol II snRNA promoter is efficiently recognized by RNA polymerase III if introduction of a TATA sequence is concomitant with a 4 bp shift of the proximal element

The Pol II and Pol III PSE alignments displayed in Table I stress the strict conservation of the two C residues. Substitution of C-62/C-61 by two G abolishes transcription of the Xenopus U6 gene (Figure 2B, lane 2), as does substitution of the entire U6 PSE element (Figure 2B, lane 3; see also ref. 5,14). The result of this experiment indicates that C-62/C-61 play a crucial role in the U6 PSE, in much the same way as what was observed for C-58/C-57 in Xenopus U1 and U2 Pol II snRNA genes (our unpublished results and ref. 23, respectively). Sequence comparisons of the proximal elements revealed also another interesting feature. While maximal sequence similarity is found for Pol II PSEs between −62 and −52 (with a positional flexibility of 1 or 2 base-pairs), those which are elements of Pol III gene promoters lie between −68 and −54 (with a 1 base-pair positional flexibility). They, however, still keep the sequence similarity to Pol II PSEs (compare Xenopus U1 and U6 PSEs in Table I, for instance). In order to know whether this differential positioning might be another determinant for RNA polymerase specificity, the wt U6 PSE was moved 4 bp downstream of its natural position so that C-62/C-61 were shifted towards the C-58/C-57 Pol II snRNA configuration (construct C194 in Figure 2A). Injection of construct C194 leads to a complete loss of U6 transcription (Figure 2B, lane 4). From these results it seems clear that, eventhough the PSEs of Pol II and Pol III U snRNAs show a high degree of sequence similarity and may even be interchangeable between U2 and U6 genes (15, 23), their positions in the promoters are not functionally equivalent.

It remained to determine whether upstream shifts of the PSE could trigger a change in polymerase specificity of the U1.TATA promoter carrying a G at the start site, as in U6 genes. To test this hypothesis, the distance between the PSE and the TATA motifs was varied by increments of 2 to 14 bp introduced by site-directed mutagenesis between positions −44 and −43, as indicated in Figure 3A. Template activities of the various constructs are shown in Figure 3B. They revealed what we anticipated but also unexpected results. A product migrating below 5S maxi appears when insertions of 4 and 6 bp are introduced and the 4 bp-arising product has the maximal intensity (compare lanes 4 and 5). Both are transcribed by RNA polymerase III since they resist to α-amanitin concentrations inhibiting RNA polymerase II (this is shown in lane 9 for the product directed by construct L11), start at +1 (as deduced from a primer extension experiment) and most likely end at the engineered run of Ts. This would generate a ~130 nucleotide long molecule which effectively corresponds to the observed electrophoretic mobility of the transcripts in lanes 4 and 5. That these transcripts are properly terminated is corroborated by the following (not shown) experiments. When the U1.TATA construct carrying the 4 bp shift, but no artificial run of Ts, is injected into oocytes, two types of Pol III transcripts appear: one ending at the external natural run of 6 Ts and the other at the U1 Sm binding site TTTCT (which then acts as a potent Pol III terminator). However, substituting C by T in the Sm binding site of construct L11 leads to complete conversion of the

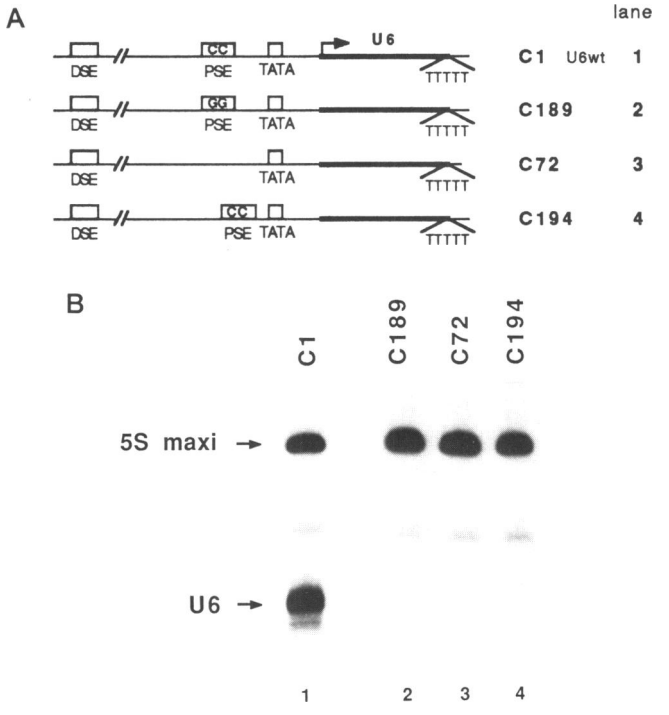

Figure 2. Structure and template activities of various U6 mutant constructs. **(A)** Structure of the mutants used in this analysis. Symbols are as in Figure 1A. Nucleotides inside the PSE open boxes indicate those mutations that were introduced. The arrow indicates the start of transcription. Drawings are not to scale. **(B)** Transcription and gel conditions are as in Figure 1B. Positions of the 5S maxi and U6 transcripts are indicated. Injection and electrophoresis arising from C1 (lane 1) were done in a separate experiment.

Table I. Sequence comparisons of the Proximal Sequence Elements of various vertebrate Pol II (PSE Pol II) or Pol III (PSE Pol III) genes.

PSE Pol II				CC positions	
U1	X.laevis	CTCTCCTTATG	−52	−58/−57	(19)
	human	GTGACCGTGTG	−52	−58/−57	(31)
	chicken	GTCGCCGTGCG	−51	−57/−56	(32)
U2	X.laevis	CTCTCCCCATG	−52	−58/−57	(23)
	human	CTCACCGCGAC	−52	−58/−57	(33)
U4B	chicken	CTCGCCGTGAG	−50	−56/−55	(34)

PSE Pol III				CC positions	
U6	X.tropicalis	CTCTCCTTAAGTT	−54	−62/−61	(4)
	human	CTTACCGTAACTT	−54	−62/−61	(15)
	mouse	CTCACCCTAACCT	−55	−63/−62	(12)
7SK	human	TTGACC−TAAGTG	−55	−63/−62	(7)

Numbers refer to the position of the last base of each sequence with respect to the transcription start site. Bold letters refer to the two conserved C residues whose positions in the PSE are indicated on the right. Sequences were taken from the literature cited in parenthesis. Only the X. laevis and human U2, human U6 and 7SK PSEs have been mapped with accuracy. All PSE sequences were aligned to maximize sequence similarities.

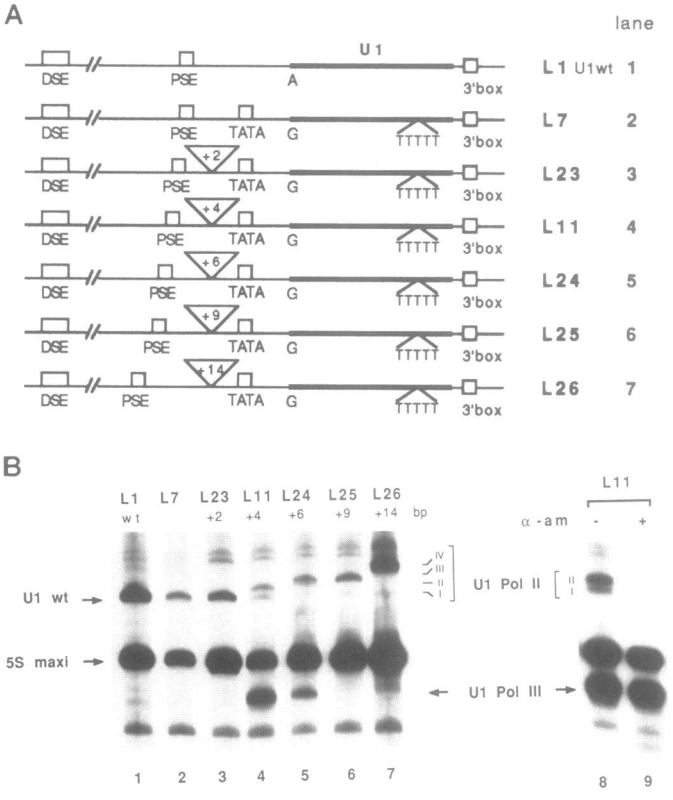

Figure 3. Effects on U1 Pol III transcription of incrementing the PSE-TATA distance. **(A)** The length of DNA sequence that was added between the PSE and TATA elements is indicated in the insert. This was done by site-directed insertion of the sequences AG (+2), AGAT (+4), AGTGAT (+6), AGCTGAGAT (+9), AGCTGTCTAGAGAT (+14) between positions −44 and −43 in the respective constructs. Other symbols are as in Figure 1A. Drawings are not to scale. **(B)** Effects of the insertions on U1 transcription as assayed by oocyte injection. The lane numbers 1 to 7 correspond to lane numbers given in (A). Lanes 8 and 9 correspond to injection of construct L11 from lane 4 in the absence (−) or presence (+) of α-amanitin (α-am) at 2 μg/ml, respectively. +2 to +14 bp correspond to the insertion length displayed in (A). Transcripts were analyzed as in Figure 1B. Positions of U1 wt, 5S maxi, U1 Pol III and U1 Pol II bands I to IV described in the text are indicated. The lowest band that shows up in all lanes corresponds to endogenous 5S RNA and the doublet in the upper part of the gel to endogenous 5.8S RNA.

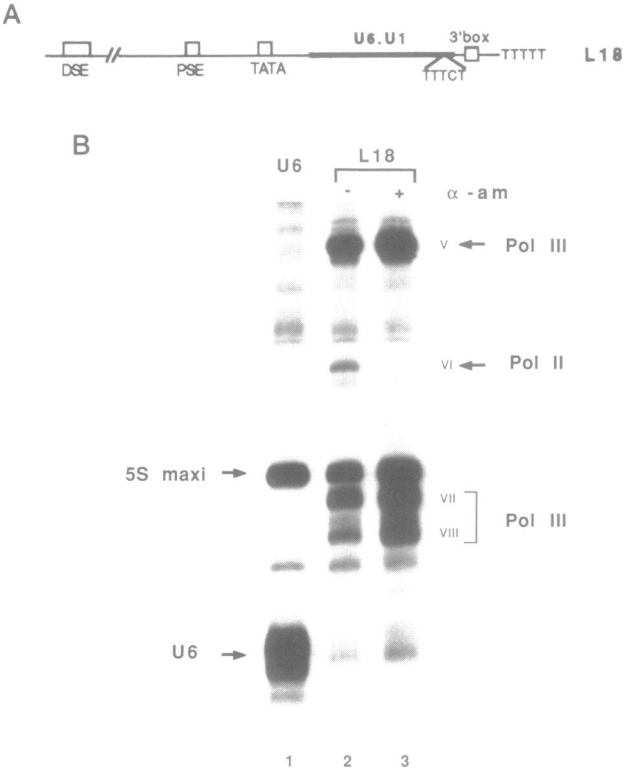

Figure 4. Transcriptional activity of the U6.U1 fusion construct. **(A)** The heavy line represents the U6.U1 fusion construct (see Materials and Methods). TTTCT is the sequence of the X. laevis U1 snRNA Sm binding site, acting as an efficient Pol III terminator. Other symbols are as in Figure 1A. Drawings are not to scale. **(B)** Template activity of construct L18 depicted in (A). Lane 1 shows the transcription activity of the X. laevis wt U6 (same as C1 in Figure 2A and B). Lanes 2 and 3 show the result of the injection into Xenopus oocyte nuclei of L18 in the absence (−) or presence (+) of α-amanitin (α-am), respectively. Gels are as in Figure 1B. The positions of 5S maxi, U6, U6.U1 Pol II (band VI) and U6.U1 Pol III (band V,VII,VIII) are indicated. The band between band VIII and U6 corresponds to endogenous 5S RNA.

transcripts terminated at the external run of 6 Ts to those ending at the engineered run of Ts. The intensity of the Pol III product decreases according to the insertion length since in lane 5, where 6 bp were introduced, transcription is reduced 4-fold. In lane 7, a lower intensity product appears which is slightly different in size from that observed in lanes 4 and 5. As its electrophoretic mobility showed small variation during the various injection experiments, we are unable to state whether it is also a specific Pol III transcript or a degradation product arising from the high intensity band IV. This series of spacing mutants showed us that templates in which either only 2 bp or more than 6 bp were inserted between the PSE and TATA elements are inactive in Pol III transcription (lane 3, 6 and 7). In contrast, the 4 bp and to a lesser extent the 6 bp insertion allow the U1.TATA promoter to be efficiently recognized by RNA polymerase III. The 4 bp insertion is in keeping with the natural positioning of the U6 PSE. The surprising result is that the Pol II transcription still persists upon injection of construct L11 (lane 4), generating the two bands I and II which are both sensitive to low α-amanitin concentrations (compare lanes 8 and 9). Band I corresponds to a U1 wt start-site and the more intense band II contains a 5′ extended transcript starting at −4 (primer extension not shown), presumably resulting from the upstream shift of the PSE which serves to position the normal start-site of transcription in Pol II U snRNA genes (1,2). The other larger products (bands III and IV) observed in lanes 5, 6 and 7 also contain Pol II transcripts (α-amanitin experiments not shown) and it appears that the farther the PSE from its natural position, the more remote the initiation site from its wt position, although there is no direct correlation since a 9 bp insertion (lane 6) apparently results in the same upstream start-site selection as a 6 bp insertion (lane 5). The prominent feature, however, is that template L26 (lane 7) gives a Pol II transcript which is 3 times as much intense as transcripts contained in bands III or in the sum of bands I and II. In addition, it only shows a 40% drop in transcription relative to that exhibited by U1 wt (lane 1) despite the fact that the two important C residues in the wt U1 and L26 PSEs (the two Cs are mentioned here as positional markers) are located 14 bp apart.

Pol II transcription of the snRNA type can always arise from a chimeric Pol III snRNA template in injected X. laevis oocytes

The persistence of U1 Pol II transcripts obtained upon injection of construction L11 (Figure 3B, lane 4) led us to ask what the converse experiment would yield. Would a Xenopus U6 gene be also efficiently expressed by RNA polymerase II, in addition

to its being normally transcribed by RNA polymerase III ? It has been shown that Pol II transcripts also arise from the injected Xenopus U6 gene. These transcripts occur at a low level, however, they can only be detected by primer extension and are presumably terminated like mRNAs since there is no signal for Pol II snRNA termination in the U6 gene (14). The construct L18, designed for the purpose indicated above, contained a fused U6.U1 gene harboring the U1 3'box downstream of the fused U6.U1 coding region (Figure 4A). L18 therefore contains not only the 3'box but also the 3'terminal stem-loop of U1 snRNA which is another important structural feature for termination of transcription of Pol II snRNA genes (24). Its template activity is shown in Figure 4B, lane 2 (lane 1 shows the wt U6 for comparison). L18 gives rise to four bands denoted V to VIII (Figure 4B). Electrophoretic mobilities of bands V, VI and VII are consistent with their being transcripts originating from templates using the expected termination signals displayed in Figure 4A : the external run of Ts for V, the U1 3'box for VI and the U1 Sm binding site (acting as an efficient Pol III terminator) for VII. However, the important point to make is the disappearance of band VI in the presence of α-amanitin (compare lanes 2 and 3), inferring that the fusion U6.U1 transcript it contains is clearly Pol II-dependent. We have no explanation for the origin of band VIII whose size does not correspond to termination at any expected termination signal. It is definitely a Pol III transcript since it persists in the presence of α-amanitin conditions leading to the disappearance of band VI (Figure 4B, lane 3).

The above experiments showed that the Pol II transcripts of the snRNA type directed by the Xenopus U6 promoter can be directly detected in transcription experiments upon injection into Xenopus oocytes, provided that a Pol II snRNA termination signal is added. Hence, in the light of the data showing that the Xenopus U6 Pol III promoter always directs Pol II transcription that can or cannot be directly detected on gels, it is not surprising that Pol II snRNA transcription arising from the U1 Pol III promoter cannot be suppressed. The question now arises as to whether it is feasable to force the U1 Pol III promoter to drive solely Pol III transcription. This is examined in the following experiments.

A tRNA B box or a U2 snRNA PSE are elements that influence the levels of Pol II and Pol III snRNA type transcriptions from the U1 Pol III promoter

The constructs that were used are displayed in Figure 5A. Introduction of a B box, one of the two internal elements promoting Pol III transcription of tRNA genes (25), into the Xenopus U6 gene was shown to increase the level of U6 Pol III transcription of this chimeric gene (26). A similar result was obtained with the $tRNA^{(Ser)Sec}$ gene (P. Carbon and A. Krol, in press), a naturally occurring version of the box B-containing U6 gene. Injection of L28, a L11 construct containing a B box substituting positions 60 to 70 of the U1 coding region, leads to a 2-fold increased Pol III transcription, compared to the activity of L11 (Figure 5B, lane 3 and 4). The neutral effect on U1 RNA stability of introducing a B box into the U1 coding region was verified prior to mutagenize the L11 construct (Figure 5B, lane 1 and 2). Introduction of this element thus only augments the level of the U1 Pol III transcript but does not lower the Pol II transcription level. Substitution of the 13 bp of the U1 wt PSE by that of the Xenopus U2 gene (23) was also tested. This is because it was previously reported that a fused construct carrying the DSE and PSE elements of the Xenopus U2 gene but retaining

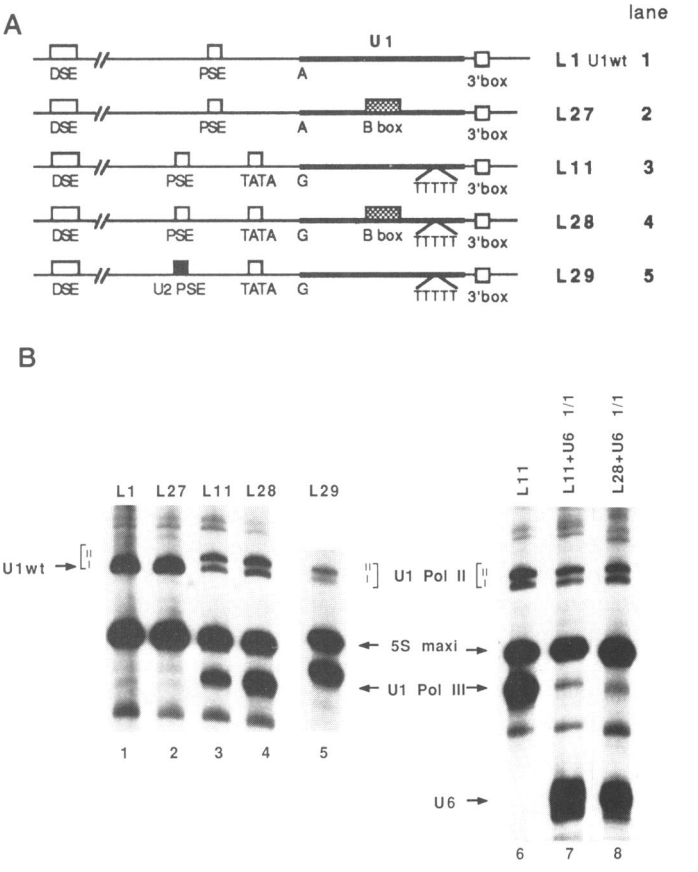

Figure 5. Analysis of box B and Xenopus U2 PSE containing U1 Pol III templates. **(A)** Structure of the constructs used in this analysis. GGTTCAATTCC (30) substituted the U1 coding sequence from 60 to 70. The solid square represents 11 bp and 2 bp of 3' flanking region of the X. laevis U2 PSE sequence (23) which substitutes −62 CTCTCCTTATGTT −52 of the U1 PSE to −62 CTCTCCCCATGGA −52. Both substitutions were performed by site-directed mutagenesis. Other symbols are as in Figure 1A. Drawings are not to scale. **(B)** Lanes 1 to 5 correspond to oocyte injections of the constructs shown in (A). Injection and electrophoresis of the transcripts shown in lane 5 were done in a separate experiment. Lane 6 corresponds to injection of L11 as in lane 3. Lanes 7 and 8 show coinjection of U6 with either L11 or L28 (1/1 ratio), respectively. Gels are as in Figure 1B. Positions of the U1 wt, U1 Pol II (bands I and II), U1 Pol III, U6 and 5S maxi transcripts are indicated. The lower U6 transcription activity in lane 8 compared to that of lane 7 results from variability of the U6 transcription pattern in oocytes of different frogs as injections shown in lanes 6, 7 and 8 were performed with different batches of oocytes. The lowest band observed in lanes 1−4 corresponds to endogenous 5S RNA which migrates between U1 Pol III and U6 in lanes 6−8. The upper doublet in lanes 1−4 and 6−8 is endogenous 5.8S RNA.

the U6 coding and 5' flanking regions up to the TATA element surprisingly produced only Pol III transcripts (14). Injection of the construct of Figure 5A lane 5 showed that a Xenopus U2 PSE, in the context of the U1 promoter, is not sufficient either to fully suppress Pol II snRNA transcription initiated from this promoter (Figure 5B, lane 5). The ratio of Pol II versus Pol III transcription changes, however, as observed by comparing lanes 3 and 5 in Figure 5B: transcription of the U1 Pol III product is increased to a level similar to that obtained with the introduction of a B box into L11 (lane 4) and there is a simultaneous marked decrease in the level of the Pol II products I and II.

In summary, while it is possible to enhance the level of Pol III transcription directed by the U1 Pol III promoter by either introducing an internal B box or by substituting another PSE, none of these two elements taken separately enable the complete Pol II to Pol III conversion of the U1 promoter.

Introduction of a TATA sequence and a shifted PSE is not sufficient to confer competitive ability to the U1 Pol III promoter

We have introduced into the U1 wt promoter all the cis-elements that have been known so far to be essential and sufficient for the Xenopus U6 transcription by RNA polymerase III. Therefore, the U1 Pol III template should stand competition over the Xenopus U6 wt gene in coinjection experiments (1/1 ratio) with either L11 or L28. Instead, lanes 7 and 8 in Figure 5B show that this is not the case since even the L11 construct carrying a B box shows a severely reduced template activity for Pol III transcription, the intensity of bands I and II (Pol II transcripts) remaining constant. The same holds true for a competition experiment performed between L29 (L11 carrying the U2 PSE) and U6 (data not shown).

Thus, the introduction of a TATA box into a Pol II snRNA promoter and the subsequent 4 bp shift of the PSE are indispensable elements that allow recognition of this promoter by RNA polymerase III. However, they do not seem to be sufficient to confer competitive ability in coinjection experiments with a natural Pol III template. Even introduction of a B box into the U1 coding region, while resulting in an increased Pol III activity, does not change the competitive strength of the U1 Pol III promoter. These unexpected results will be discussed below.

DISCUSSION

The reason which has motivated the beginning of this study was driven by the need of analyzing the effects of introducing a TATA motif into a natural Pol II snRNA gene. Previous studies pointed out the dominant role of the TATA sequence for conferring Pol III specificity to U6 promoters. However, this was deduced either by deleting this sequence in a U6 gene or by fusing a U2 promoter (containing the DSE and the PSE) to a U6 coding fragment carrying the TATA sequence (14) or by converting to Pol III a U2 promoter driving a β-globin coding region (15, 17). In addition, this latter construct did not produce correctly initiated Pol III transcripts. Instead, our strategy was different. It consisted in adding or modifying sequentially in an X. laevis U1 snRNA gene, a regular Pol II-transcribed gene, all the elements that have been known so far to be responsible for efficient Pol III transcription of the Xenopus U6 gene (5, 14). We showed in this report that the sole introduction of the U6 TTATAA sequence (positions $-30/-25$) in place of the corresponding U1 wt sequence is without effect since no U1 Pol III product appears while the U1 template is still transcribed by RNA polymerase II (albeit at a lower level). It must be emphasized that in this construct the distance existing between the TATA element and the start of transcription was kept constant. We could be successful in obtaining Pol III transcription initiated by the U1 Pol III promoter and ending at the engineered run of Ts by simply moving the PSE 4 or 6 bp upstream of its Pol II position. In fact, the optimal spacing results from the 4 bp shift. The shifts were measured relative to positions C-58/C-57, two nucleotides contained in the PSE and that were shown to be crucial for its functioning in Xenopus U1 and U2 genes (our unpublished results and ref. 23, respectively). Interestingly, sequence alignments (Table I) and our results clearly show that, although PSEs are interchangeable at the sequence level between Pol II and Pol III snRNA genes (15, 23), they function in a position-dependent manner in Pol II or Pol III snRNA type promoters. Therefore, in order to render the promoter of a typical Pol II snRNA gene efficiently recognized by RNA polymerase III, it is not only necessary to introduce a TATA sequence (at a distance from the start of transcription identical to that of U6 genes) but also to displace the PSE to the C-62/C-61 configuration, corresponding to the 4 bp upstream shift relative to its Pol II (C-58/C-57) configuration. Only if these conditions are respected can a Pol III transcript appear that is correctly initiated at $+1$. This is in marked contrast to what was reported for the Pol II-Pol III conversion of the human U2 gene (15) in which the TATA sequence was added at a distance from the U2 Pol II PSE equal to that observed in the human U6 gene. This most likely explains the start site reported by these authors to occur inside the β-globin coding region and not at the expected $+1$ position. The rather strict dependence on a small shift of the PSE (4 or 6 bp) to obtain Pol III transcription strongly argues in favor of a need for the PSE-binding and TATA-binding factors to contact each other.

In the Xenopus U6 gene, the region of initiation of transcription was also found to be important for the Pol III specificity of this gene since changing G, the normal initiation nucleotide, to an A results in a 3-fold drop in the transcription level (14). Surprisingly, this same transition, when carried out in the U1 Pol III promoter, did not alter the Pol III transcription level at all (unpublished results). This observation is even more striking if one considers that this G to A transition was performed in the context of the U1 initiation region which is extremely similar in sequence (over the 23 bp spanning positions -13 to $+9$) to that of the Xenopus U6 gene. The upstream three base-pairs are pyrimidines in both genes and it was suggested that YYYG is the preferred initiation sequence for the U6 gene (14). We are currently unable to interpret the different behavior of the U6 and U1 Pol III genes.

While addition of a TATA sequence and the concomitant shift of the PSE served to convert the U1 Pol II promoter into a promoter efficiently recognized by RNA polymerase III, these processes have not revealed sufficient to fully suppress the U1 Pol II-snRNA type transcription which uses the 3'box termination signal. The U1 Pol II transcription is nevertheless only residual, since in construct L11 it represents one tenth of the U1 Pol III transcription level. A dual transcription seems even to represent an intrinsic property of U snRNA Pol III promoters : we could also produce with a high rate from the Xenopus U6 wt Pol III promoter a discrete band corresponding to Pol II transcription of the snRNA type by the simple addition of a 3'box termination signal. As U6 naturally lacks a 3'box, the Pol II transcripts described in a previous report (14) were of the mRNA type with regard to termination, therefore containing long heterogenous 3'ends only detectable by primer extension. Two discrete U1 Pol II transcripts were actually generated by construct L11, carrying the C-62/C-61 PSE. Band I corresponds to the wt U1 start site, band II contains a transcript initiating at -4 upstream of the wt start site, in a position also occupied by an A. As the PSE is moved 4 bp upstream of its wt position normally leading to a $+1$ start, the start at -4 correlates well with the PSE being involved in start site selection in U snRNA genes (1, 2). A surprising result in these spacing experiments is that the highest intensity (almost that of U1 wt) among all of the observed U1 Pol II products occurred when the PSE was moved the farthest from its Pol II configuration (Figure 3B, lane 7). This suggests that, whereas the PSE imposes a constraint on the distance downstream of which initiation should start, it is also likely to work in conjunction with another element(s) that may reside around the transcription start site and that fortuitously exists in this construct.

We tried to artificially force the U1 Pol III promoter to be uniquely Pol III dependent and this was done by either substituting a B box (one of the two internal promoter elements of tRNA genes) inside the U1 coding region or the U1 PSE by the U2 PSE. Both constructs resulted in a higher U1 Pol III expression than that obtained with the parental construct. This was also observed by others in a U6 construct carrying a B box (26). As to the effects on the U1 Pol II expression level, they were different according to whether we are looking at B box or U2 PSE-containing constructs. Only substitution by the U2 PSE led to an inversion in the balance of the Pol II and Pol III transcripts since the former dropped simultaneously to the increase of the latter. This is in keeping with a previous observation which showed that a construct containing a U6 coding region and its own TATA element fused to U2 upstream flanking regions carrying both the DSE and the PSE completely suppressed the Pol II-dependent U6 transcription (14). Here, the L29 construct (carrying the U2 PSE) retained almost all of the U1 PSE flanking regions and this might provide one reason to explain the residual level of U1 Pol II transcription, but the mechanism by which the U2 PSE functions in augmenting the U1 Pol III transcription level remains unknown. Addition of a B box does not seem to result in a larger number of assembled transcription complexes since the increased Pol III transcription level is not accompanied by a concomitant drop in the Pol II transcription level. Rather, the likely binding of TFIII C to the created B box (27) and the subsequent binding of the general Pol III transcription factor TFIII B (28) could stabilize an already assembled transcription complex, thereby enabling more rounds of Pol III transcription to occur.

The B box and the U2 PSE serve to augment Pol III transcription on the U1 promoter without providing a competitive advantage over a U6 gene. A possible influence of the two different U1 and U6 DSEs may be eliminated since the U1 Pol III transcription is almost abolished while the U1 Pol II transcription arising from the same template remains constant. This abolition in the presence of a competitor could be due to a U6 element(s) that has not been mapped so far and therefore not introduced into the U1 Pol III promoter: either sequences flanking the PSE (since U1 and U6 PSEs differ by only 1 bp) or sequences residing between the TATA element and the start site or a motif in the U6 coding region itself. In all likelihood, the Pol II and Pol III transcripts that arise from the U1 Pol III template are caused by the assembly of independent Pol II and Pol III transcription complexes.

In plant U snRNA genes, a spacing between promoter elements have also been shown to be crucial for determining Pol II or Pol III specificity (29). The plant system is distinguishable, however, from the vertebrate system we have analyzed in that in plants both Pol II and Pol III genes contain the same two promoter elements, a TATA sequence and a USE. The USE is structurally and maybe functionally different from the PSE. This data and ours nevertheless reinforce the idea that these systems derive from a common ancestor and that the border between the current Pol II and Pol III U snRNA gene transcription systems is less tight than expected.

ACKNOWLEDGMENTS

We are grateful to J. P. Ebel for support in this work, to P. Remy for microinjection facilities and to C. Sturchler for critical reading of the manuscript. A. Hoeft is thanked for oligonucleotide synthesis. This work was supported by an INSERM grant (Contrat de Recherche Externe 891009).

REFERENCES

1. Dahlberg, J. and Lund, E. (1988) In *'Structure and function of small nuclear ribonucleoprotein particles'*, Birnstiel, M.L. (Ed.) Springer Verlag Berlin Heidelberg, pp. 38–70.
2. Parry, H.D., Scherly, D. and Mattaj, I.W. (1989) *Trends Biochem.* **14**, 15–19.
3. Tanaka, M., Grossniklaus, U., Herr, W. and Hernandez, N. (1988) *Genes Dev.* **2**, 1764–1778.
4. Krol, A., Carbon, P., Ebel, J.P., Appel, B. (1987) *Nucl. Acids Res.* **15**, 2463–2478.
5. Carbon, P., Murgo, S., Ebel, J.P., Krol, A., Tebb, G. and Mattaj, I.W. (1987) *Cell* **51**, 71–79.
6. Murphy, S., Di Liegro, C. and Melli, M. (1987) *Cell* **51**, 81–87.
7. Murphy, S., Pierani, A., Scheidereit, C., Melli, M. and Roeder, R.G. (1989) *Cell* **59**, 1071–1080.
8. Brow, D. and Guthrie, C. (1990) *Genes Dev.* **4**, 1345–1356.
9. Sollner-Webb, B. (1988) *Cell* **52**, 153–154.
10. Murphy, S., Moorefield, B. and Pieler, T. (1989) *Trends Genet.* **5**, 122–126.
11. Bark, C., Weller, P., Zabielski, J., Janson, L. and Pettersson, U. (1987) *Nature* **328**, 356–359.
12. Das, G., Henning, D., Wright, D. and Reddy, R. (1988) *EMBO J.* **7**, 503–512.
13. Kunkel, G.R. and Pederson, T. (1988) *Genes Dev.* **2**, 196–204.
14. Mattaj, I.W., Dathan, N.A., Parry, H.D., Carbon, P. and Krol, A. (1988) *Cell* **55**, 435–442.
15. Lobo, S.M. and Hernandez, N. (1989) *Cell* **58**, 55–67.
16. Lewin, B. (1990) *Cell* **61**, 1161–1164.
17. Lobo, S.M., Ifill, S. and Hernandez, N. (1990) *Nucl. Acids Res.* **18**, 2891–2899.
18. Simmen, K.A. and Mattaj, I.W. (1990) *Nucl. Acids Res.* **18**, 5649–5657.
19. Krol, A., Lund, E. and Dahlberg, J.E. (1985) *EMBO J.* **4**, 1529–1535.
20. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980) *Cell* **19**, 27–35.
21. Lund, E., Bostock, C.J. and Dahlberg, J.E. (1987) *Genes Dev.* **1**, 47–56.
22. Bogenhagen, D.F. and Brown, D.D. (1981) *Cell* **24**, 261–270.
23. Parry, H.D., Tebb, G. and Mattaj, I.W. (1989) *Nucl. Acids Res.* **17**, 3633–3644.
24. Ciliberto, G., Dathan, N., Frank, R., Philipson, L. and Mattaj, I.W. (1986) *EMBO J.* **5**, 2931–2937.
25. Galli, G., Hofstetter, H. and Birnstiel, M.L. (1981) *Nature* **294**, 626–631.
26. Parry, H.D. and Mattaj, I.W. (1990) *EMBO J.* **9**, 1097–1104.
27. Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) *Ann. Rev. Biochem.* **57**, 873–914.
28. Kassavetis, G.A., Braun, B.R., Nguyen, L.H. and Geiduschek, E.P. (1990) *Cell* **60**, 235–245.
29. Waibel, F. and Filipowicz, W. (1990) *Nature* **346**, 199–202.
30. Sharp, S.F., Schaack, J., Cooley, L., Burke, D.J. and Söll, D. (1985) *CRC Crit. Rev. Biochem.* **19**, 107–145.
31. 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.
32. Earley, J.M., Roebuck, K.A. and Stumph, W.E. (1984) *Nucl. Acids Res.* **12**, 7411–7421.
33. Hernandez, N. and Lucito, R. (1988) *EMBO J.* **7**, 3125–3134.
34. Mc Namara, K.J. and Stumph, W.E. (1990) *J. Biol. Chem.* **265**, 9728–9731.