Volume 15 Number 6 1987 Nucleic Acids Research

Xenopus tropicalis U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes

Alain Krol*, Philippe Carbon, Jean-Pierre Ebel and Bernd Appel[1]

Laboratoire de Biochimie, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue Descartes, 67084 Strasbourg Cedex, France and [1]Max Planck Institut für Molekulare Genetik, 63-73 Ihnestrasse, 1000 Berlin 33, FRG

Received January 22, 1987; Revised and Accepted February 23, 1987

ABSTRACT

We have cloned and sequenced a 977bp DNA fragment, pXTU6-2, that represents the transcription unit for a **Xenopus tropicalis** U6 RNA gene. This basic repeating unit is reiterated ca.500-fold per haploid genome. Oocyte injections of pXTU6-2 led to the transcription of a mature-sized U6 RNA that, however, lacked internal 2'-O-methylations. These posttranscriptional modifications of U6 RNA might be cytoplasmic and could require its association with U4 RNA to be accomplished.The low α- amanitin sensitivity of U6 RNA synthesis in oocytes suggested that U6 RNA is transcribed by RNA polymerase III, consistent with features of the U6 RNA molecule which also contains a Box A- like intragenic control region. Inspection of **X.tropicalis**, mouse and human U6 DNA upstream sequences revealed the presence of a TATA box as well as of the proximal and enhancer (octamer motif) elements contained in snRNA genes transcribed by RNA polymerase II. We propose that U6 RNAs are synthesized by a specialized transcription complex consisting of RNA polymerase III and transcription factors, some of which are very likely shared with RNA polymerase II promoters.

INTRODUCTION

Small nuclear RNAs U1,U2,U4,U5 and U6 (the U series) are metabolically stable molecules ubiquitous in the eukaryotic kingdom (1), although a compact U2-U6 version is found in a single 1kb RNA in S.cerevisiae (2). Whereas U1-U5 RNAs bear a 5' trimethylguanosine cap (1), contain the domain A (3), also called Sm binding site (4), U6 RNA possesses none of these features. It is found associated with U4 RNA in a single ribonucleoprotein particle (5,6) that functions in premRNA splicing (7,8).

U1-U5 RNAs are transcribed by RNA polymerase II (9,10). Although previous experiments performed in cultured cells pointed to a possible role of RNA polymerase III in U6 RNA transcription (11), in vitro studies were hampered by a lack of cloned copies of bona fide U6 genes (12,13). That U4 and U6 RNAs are associated in a single snRNP and may be transcribed by two distinct RNA polymerases raises the intriguing question of how their relative expression is balanced to allow the proper functioning of the U4/U6 snRNP.

This prompted us to isolate true genes for U6 RNA in order to elucidate the nature of the RNA polymerase responsible for U6 RNA transcription.
In this paper, we report on the cloning and sequencing of a 1kb transcription unit coding for a U6 RNA gene in **Xenopus tropicalis** that is highly reiterated in the genome. By α-amanitin inhibition experiments, we showed that U6 RNA is transcribed by RNA polymerase III, a similar conclusion that was reached independently by others for mouse, human and **Drosophila** U6 RNAs (14-16). More interestingly,we identified in the 5'-flanking sequences of the vertebrate U6 RNA genes the presence of signals used by RNA polymerase II : the distal and proximal elements of U RNA genes and a TATA box. These findings suggest that transcription of U6 RNA genes is controlled by an RNA polymerase III and a complex set of factors, some shared with RNA polymerase II transcription units and others which may be unique to snRNA genes.

MATERIALS AND METHODS

ANALYSIS OF DNA

X.tropicalis genomic DNA was prepared by standard procedures from erythrocytes or muscles of single individuals. DNA fragments were fractionated on 0.7 to 1% agarose gels or 6% (60:1) polyacrylamide gels and transferred to nitrocellulose or to Zeta-Probe Bio-Rad nylon membranes (using the manufacturer's recommendations). DNA blots were probed with human U6 RNA that was 3'-end labeled with [^{32}P]pCp. Membranes were hybridized overnight at 50°C in 5 xSSC (1 x SSC is 0.15 M NaCl ; 0.015 M sodium citrate) -0.1% SDS - 1mM EDTA- 5xDenhardt solution - 20mM Tris-HCl pH 7.0 - 50% formamide and washed at 55°C in 5 x SSC and 2 x SSC containing 0.1% SDS.

CLONING OF X.TROPICALIS U6 RNA GENES

X.tropicalis genomic DNA enriched in U6 RNA genes was obtained as previously described (17). Briefly, the total **X.tropicalis** DNA was digested to completion with a mixture of BglII, EcoRI and PstI and size-fractionated on a 5-20% NaCl gradient. Fractions containing large ($>$ 25kb) U6 DNA fragments were pooled and these DNA fragments were redigested with BamHI to generate unit-length fragments of the U6 repeat DNA. These BamHI fragments were purified by preparative agarose gel electrophoresis and cloned into pBR322 DNA by transformation of E.coli HB101. Isolates of the U6 gene (pXTU6-2) were identified by screening ampicillin-resistant colonies with the human U6 RNA probe (18).

DNA SEQUENCING

BamHI DNA fragment from pXTU6-2 was ligated into M13mp9 vector DNA

(19). BamHI-AluI fragments were filled-in with Klenow fragment of DNA polymerase I and blunt-end ligated into SmaI-cut M13mp9 DNA. Sequencing was done using the dideoxy chain termination method (20) with [α-^{35}S] dATP as the label.

XTU6 RNA GENE COPY NUMBER DETERMINATION

X.tropicalis DNA was digested with BamHI and fractionated in a 1% agarose gel. Copy number standards for the 1kb U6 DNA repeat were provided by adjacent lanes containing different amounts of BamHI-digested pXTU6-2 DNA that corresponded to 10, 25, 50, 100, 500 and 1000 copies of XTU6 RNA coding sequences per haploid genome. The various digests were transferred to a nitrocellulose filter and probed with 3'-end labeled human U6 RNA.

TRANSCRIPTION ACTIVITY ANALYSIS

pXTU6-2 DNA was injected into **X.laevis** oocyte nuclei along with 1 to $2x10^6$ cpm of (α-^{32}P)GTP (Amersham, 400 Ci/mmol) in a 20 nl volume. Incubation was at 20°C for 20 hours. When α-amanitin was added to the injection mixture, the final concentration used was in the range of 0.1 to 50 µg/ml (final oocyte concentration). Transcribed U6 RNAs were fractionated in 10% polyacrylamide gels containing 7M urea. RNase T1 fingerprint of eluted U6 RNA was performed as described by Barrell (21) with thin layer homochromatography on PEI plates in the second dimension.

RESULTS

TWO DISTINCT LARGE TANDEMLY REPEATED GENE FAMILIES HYBRIDIZE WITH U6 RNAs IN X.TROPICALIS

Genomic DNA of **X.tropicalis** was assayed by Southern blot analysis using human U6 RNA as a probe. Depending on the enzymes that were used to produce complete digestion of the DNA, U6 DNA migrated either as undigested fragments, or as 1kb fragments, or as 1kb and 1.6kb fragments (Fig.1A). To rule out the possibility that 1.6kb fragments arose from incomplete digestion with BamHI and StuI (the 1kb fragment and a smaller one that either did not hybridize or had run off the gel might add up to 1.6 kb in size), genomic DNA was partially digested with either DraI, SacI or StuI (Fig.1B). In all cases,a ladder of bands of U6 DNA was obtained. The mobilities of the partial DraI and SacI digestion products indicated that they were multimers of the corresponding 1kb fragment. Partial digestion with StuI generated, in addition to the 1kb multimers, fragments whose mobilities corresponded to the 1.6kb multimers. These results, as well as the observation that only those enzymes which released the 1kb fragment alone also left undigested DNA

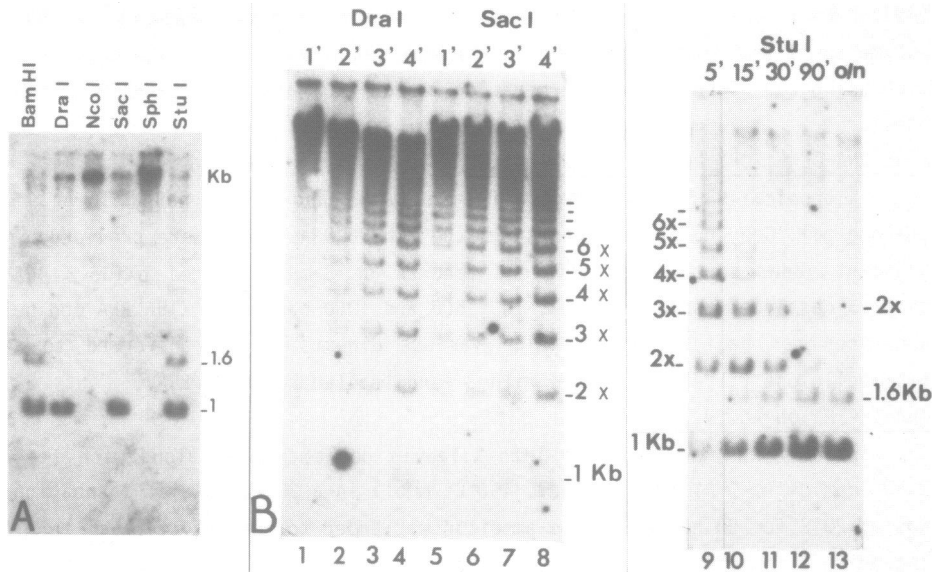

FIGURE 1A : Characterization of X.tropicalis genomic U6 DNA sequences. X.tropicalis DNA (4µg) was digested to completion with different restriction enzymes. The digestion products were fractionated in 1% agarose gels, blotted and hybridized to ^{32}P-labeled human U6 RNA.

FIGURE 1B : Tandem organization of the major repeated U6 RNA genes of X.tropicalis. X.tropicalis DNA was subjected to partial digestion with restriction enzymes DraI, SacI and StuI. The digests were fractionated in 0.7% agarose gels, blotted and hybridized as in Fig.1A. 16µg of DNA were digested for the times indicated at the top of the autoradiogram with 1 unit of DraI (lanes 1 to 4), SacI (lanes 5 to 8) or StuI (lanes 9 to 12) per µg of DNA ; 4µg of DNA were used for overnight (O/N) digestion with 3 units of SacI (see Fig.1A) or StuI (lane 13) per µg of DNA. The ladder of bands observed in lanes 10 to 14 corresponds to integral multimers of the basic 1kb and 1.6kb repeat DNAs.

hybridizing with the probe, demonstrated that the **X.tropicalis** genome contained at least two distinct families of U6 DNA, organized in multiple tandem repeats. One family is constituted by tandem copies of the 1kb repeat unit, the unit length for the other being 1.6 kb. Whether the 1 kb and 1.6 kb U6 DNA units represent functional genes will be discussed.

The 1kb and 1.6kb families will be further referred to as the two major U6 RNA gene families in **X.tropicalis**. A few minor U6 DNA bands showed up in the autoradiogram (Fig.1A), so that the existence of low copy number U6 RNA genes in **X.tropicalis** may not be excluded.

U4 AND U6 RNAs ARE NOT ENCODED BY THE SAME DNA REPEATS IN X.TROPICALIS

Bringmann et al. (5) and Hashimoto and Steitz (6) showed that U4 and U6

RNAs are associated in a single snRNP. We therefore asked whether U4 and U6 RNA genes might also be linked. To this end, a variety of restriction digests of **X.tropicalis** genomic DNA were transferred to nitrocellulose filters by the bidirectional method of Smith and Summers (22). One of the filters was probed with human U6 RNA, the other with 3'-end labeled human U4 RNA (data not shown). No U4 signal showed up at the level of the 1kb and 1.6 kb U6 repeat units. Instead, U4 RNAs were found to be encoded by two distinct tandemly repeated HinfI families, constituted by 0.85 kb and 1.5 kb repeat units (unpublished results). U4 and U6 RNAs are therefore encoded by different tandem arrays in the **X.tropicalis** genome.

CLONING OF A REPEAT-UNIT OF THE 1KB U6 RNA GENE FAMILY

X.tropicalis was enriched in U6 DNA sequences by size fractionation of total genomic DNA which had been digested with enzymes that do not cleave the tandem arrays, according to the technique previously described (17). The 1kb repeat unit was then released from the large DNA fragments by cleavage with BamHI and was cloned into pBR322 DNA. Eight tetracyclin-sensitive colonies contained a 1kb insert that hybridized to the human U6 RNA probe. Restriction-hybridization analyses performed on the recombinant plasmids showed that these sequences appeared homogenous in size and revealed minor or no sequence heterogeneity. Clone pXTU6-2 will be further used for sequencing and transcription studies.

STRUCTURE AND SEQUENCE OF THE DNA

The 1 kb BamHI fragment was subcloned into M13mp9 DNA (19). Coding and non coding single-strands were selected by hybridizing recombinant phages immobilized on nitrocellulose with the human U6 RNA probe. Sequencing both ends of the insert confirmed the position and orientation of the U6 RNA coding region (Fig.2). The sequence of the BamHI fragment was completed as described in Materials and Methods using the sequencing strategy shown in Fig.2. Figure 3 presents the DNA sequence of the non-coding strand (RNA sense) of the 977 bp fragment containing a **X.tropicalis** U6 RNA gene. Although the precise length of this DNA fragment is slightly smaller than 1kb, for reason of convenience it will still be referred to as the 1kb fragment. The repeat harbors one U6 RNA coding region (XTU6 RNA) whose 5' and 3'-ends were tentatively identified by comparison with the rodent U6 RNAS (23,24).The sequence of DNA corresponding to **X.tropicalis** U6 RNA is shown as nucleotides 1 to 107 in Fig.3.The **X.tropicalis** U6 RNA has a single C⟶U change at position 6 from rodent U6 RNAs. This nucleotide is bulged out in the secondary structure model proposed by Rinke et al. (25) and consequently does

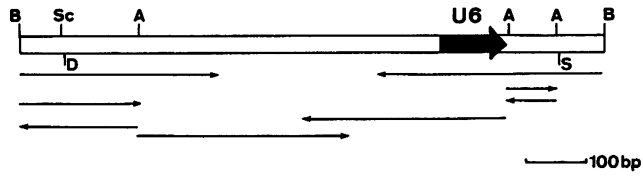

FIGURE 2 : Physical map of the U6 DNA repeat of X.tropicalis. The restriction enzyme cleavage map was obtained by analyses of pXTU6-2 DNA by standard procedures. Position and orientation of the U6 RNA coding region were obtained as described in Results. Arrows show the extent of sequences determined on the non-template and template strands.
A : AluI ; B : BamHI ; D : DraI ; S : SacI ; St : StuI

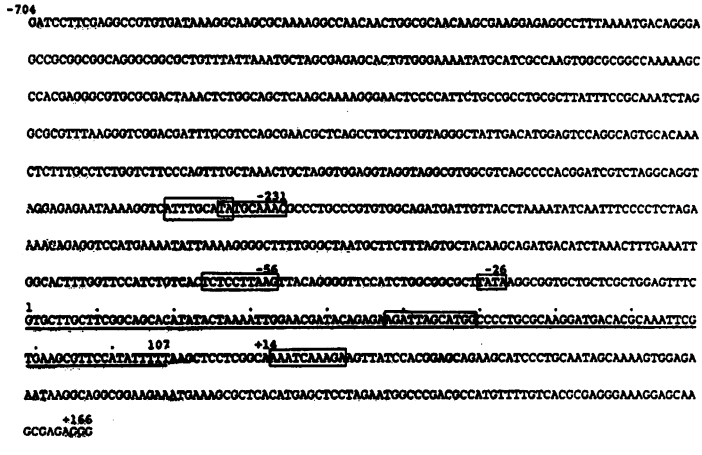

FIGURE 3 : DNA sequence of the non-template strand of the 977bp repeat encoding the X.tropicalis U6 RNA. The U6 RNA coding sequence is underlined. The presumptive 5' and 3'-ends of XTU6 RNA are numbered 1 to 107, deduced from rodent U6 RNA sequences (23,24).Boxed regions are discussed in the text.

not alter the base-pairing scheme. From this, pXTU6-2 appears to be a candidate gene for **X.tropicalis** U6 RNA.

Inspection of the flanking sequences was informative. There is a TATA-box at about -30 from the mature 5' end of the RNA. Centered around -60 is found the decamer TCTCCTTAAG. This sequence matches the consensus TCTCCNTATG (26,27) (with the exception of the point mutation A⟶T) that has been shown to be essential for transcription of **X.laevis** U1 and U2 RNA genes (27-29). An element with a similar sequence and function has been characterized in mammalian U1 and U2 RNA genes (30,31).

Of particular interest is the finding at about -238 of the sequence

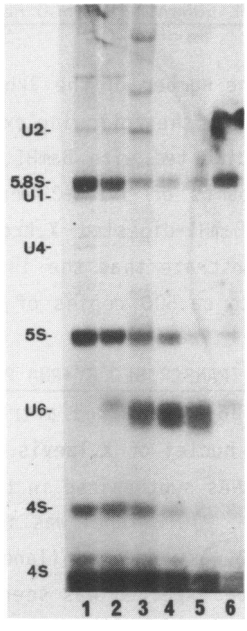

FIGURE 4 : Template activity of X.tropicalis U6 DNA injected into X.laevis oocytes with (α-^{32}P)GTP as the label. The RNAs were separated in 10% polyacrylamide gels, 7M urea. **Lane 1** : no DNA injected. **Lane 2** : Injection of pXTU6-2. **Lanes 3 to 6** : injection of pXTU6-2 along with α-amanitin concentrations of 0.1,1,10,50µg/ml (final oocyte concentration) respectively.

TATGCAAAC. This element has been found at a homologous (distal) position in almost all true snRNA genes sequenced so far where it functions as an enhancer (26,30-33). More appealing is the observation that in **X.tropicalis**, the motif TATGCAAAC is repeated in an inverted orientation (except for the 3' terminal C wich is converted into an A instead of a G), both motifs overlapping each other over two base-pairs. The 1kb family of **X.tropicalis** U6 RNA genes contains the sequence ATTTGCATATGCAAAC that may therefore be considered as a distal-like element.

About 14 nucleotides downstream of the position corresponding to the 3' end of mature U6 RNAs is found a sequence AAATCAAAGA bearing a close resemblance to the 3' box of U snRNA genes that is part of the elements required for correct 3' end formation (34-36).

The possible functional significance of all these elements will be discussed below.

The X.TROPICALIS HAPLOID GENOME CONTAINS Ca.500 REPEAT UNITS ENCODING THE 1KB FAMILY OF U6 RNA GENES

In order to determine the number of the 1kb U6 RNA gene copies in the **X.tropicalis** genome, we performed the following experiment (see Materials and Methods). Genomic DNA was digested with BamHI, yielding the two 1kb and 1.6kb fragments. Various amounts of pXTU6-2 were digested with BamHI and electrophoresed adjacent to BamHI-digested **X.tropicalis** genomic DNA. The results (data not shown) demonstrate that the intensity of the 1kb band in the genomic DNA corresponds to ca.500 copies of the repeat DNA per haploid genome.

X.TROPICALIS U6 RNA GENES ARE TRANSCRIBED BY RNA POLYMERASE III

In order to test if pXTU6-2 contained a true and active U6 RNA gene, the DNA was injected into the nuclei of **X.laevis** stage 6 oocytes. Figure 4, lane 1 shows the spectrum of RNAs synthesized in ten **X.laevis** oocytes when no exogenous DNA but only the (α-^{32}P)GTP label was injected. When a mixture of (α-^{32}P)GTP and 1ng of pXTU6-2 was injected (lane 2), one new band at the position where U6 RNA would be expected was seen. This result showed that pXTU6-2 contains all the information for faithful in vivo synthesis of a true U6 RNA gene. To determine which class of RNA polymerase is responsible for U6 RNA synthesis, **X.laevis** oocytes were injected with a mixture of pXTU6-2 DNA and increasing amounts of α-amanitin(0.1,1,10,50µg/ml, final oocyte concentration). Analysis of the transcripts (Fig.4, lanes 3 to 6) showed that U6 RNA transcription is only inhibited at α-amanitin concentrations higher than 50 µg/ml, paralleling the transcription levels of endogenous 5S RNAs and tRNAs. This low sensitivity to α-amanitin is a typical feature of genes transcribed under the control of RNA polymerase III. We checked the **X.tropicalis** U6 RNA gene for the presence of internal promoter elements, a prerequisite for RNA polymerase III-dependent gene transcription (37). Indeed, XTU6 RNA gene contains the sequence AGATTAGCATGG (positions 48 to 59 of the DNA sequence) which displays a high sequence homology with the Box A consensus sequence of eukaryotic tRNA genes, PuGPyNNAPuPy-GG (37). Sequences equivalent to Box B of tRNA genes or Box C of 5S RNA genes (37) could not be found in the XTU6 RNA gene.

SEQUENCE ANALYSIS OF THE XTU6 RNA TRANSCRIBED UPON INJECTION INTO OOCYTES

To further prove the authenticity of the U6 RNA transcribed in injected oocytes, the RNA that migrated at the level of U6 was eluted and subjected to an RNase T1 fingerprint analysis. Figure 5 shows T1 fingerprints of **X.tropicalis** (panel A) and human U6 RNAs (panel B kindly provided by

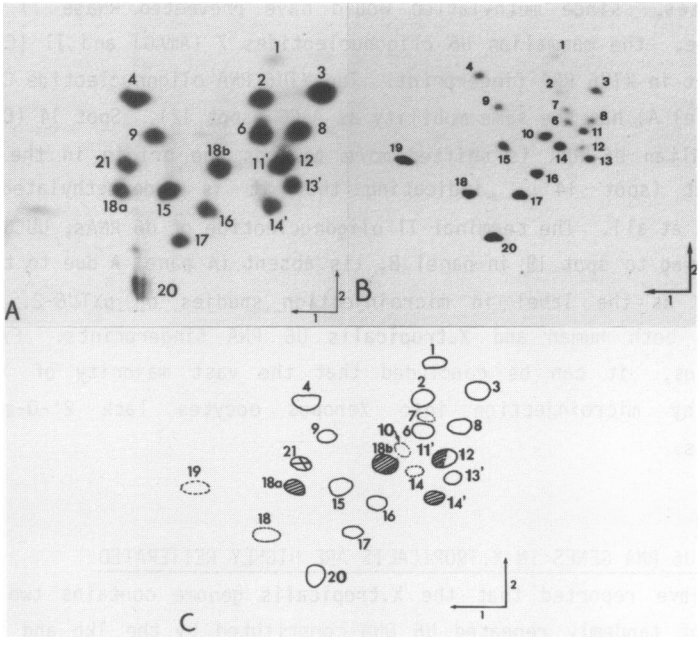

FIGURE 5 : RNase T1 fingerprints of X.tropicalis and human U6 RNAs.
A : Fingerprint of X.tropicalis U6 RNA. RNase T1 digests were separated in the first dimension on cellulose acetate strips, by homochromatography on PEI TLC plates in the second dimension. **B** : Fingerprint of HeLa U6 RNA, as published in (38). **C** : Numbering of XTU6 RNA (as shown in **A**) and human U6 RNA (as shown in **B**) according to (23). 1:G ; 2:AG ; 3:CG ; 4:UG ; 6:AAG ; 8:CAG ; 9:AUG ; 11':CAAG ; 12:AACG ; 13':ACACG ; 14':CCCCUG ; 15:CUUCG ; 16:AUACAG ; 17:CAAAUUCG ; 18a:AUUAG ; 18b:CAUG ; 21:CUUG. ⊘ : New spot in XTU6 RNA fingerprint due to C⟶U transition. ⊘ : New spots in XTU6 RNA fingerprint due to the absence of 2'-O-methylation. ⊘ : Spots absent in XTU6 RNA fingerprint, as discussed in the text.

J. Rinke). The numbering of the XTU6 RNA T1 oligonucleotides is in accordance with fingerprints published for human (38) and rat (23) U6 RNAs.

Due to a single C⟶U nucleotide substitution between the sequences of rat/human and **X.tropicalis** U6 RNAs at position 6 of the DNA sequence, a new T1 oligonucleotide CUUG (spot 21 in panel A) was generated which is not present in the human U6 fingerprint where spot 10 contains the sequence CUCG. Interestingly, two further new spots show up on the XTU6 RNA fingerprint : spot 18a (AUUAG) and spot 18b (CAUG). These oligonucleotides can only be generated by an RNase T1 cleavage of the U6 RNA sequence AUUAGmCAUG (spot 18 in mammalian U6 RNAs). This indicates that the 2'-O-methylated guanosine in spot 18 of mammalian U6 RNAs is absent in the XTU6 RNA produced by injection

into oocytes, since methylation would have prevented RNase T1 cleavage. Furthermore, the mammalian U6 oligonucleotides 7 (AmAG) and 11 (CAAmG) are also absent in XTU6 RNA fingerprint. The XTU6 RNA oligonucleotide CAAG (spot 11' in panel A) has the same mobility as AACG (spot 12). Spot 14 (CmCCmCmUG) from mammalian U6 RNA is shifted more towards the origin in the XTU6 RNA fingerprint (spot 14'), indicating that it is undermethylated or not methylated at all. The terminal T1 oligonucleotide of U6 RNAs, UUCCAUAUUUUU, corresponding to spot 19 in panel B, is absent in panel A due to the use of (α-^{32}P)GTP as the label in microinjection studies of pXTU6-2. Panel C summarizes both human and **X.tropicalis** U6 RNA fingerprints. From these observations, it can be concluded that the vast majority of XTU6 RNAs produced by microinjection into Xenopus oocytes lack 2'-O-methylated nucleotides.

DISCUSSION

THE MAJOR U6 RNA GENES IN X.TROPICALIS ARE HIGHLY REITERATED

We have reported that the **X.tropicalis** genome contains two distinct families of tandemly repeated U6 DNA constituted by the 1kb and the 1.6kb basic units. Restriction-hybridization experiments performed on genomic DNA prepared either from erythrocytes or muscles of two individuals revealed the same organization as well as sequence homogeneity of the repeats. Determination of gene copy number for the 1kb family gave ca.500 repeat units. In the 1.6kb band, however, the intensity of the hybridization signal is reduced about 2- fold (Fig.1A). This decrease can be attributed either to a lower gene copy number for the 1.6kb family or to U6 RNAs exhibiting a lower degree of sequence homology with the human U6 RNA probe that was used. When digesting genomic DNA with BamHI or StuI, only the 1kb and 1.6kb families appear as the major products. Indeed, a few minor bands show up that may correspond to partial digests or to dispersed U6 genes. In **X.laevis**, where a similar situation was observed, the major U1 genes have been shown to encode U1 RNAs that are transcribed at the onset of transcription in the Mid-Blastula Transition of embryonic development (17). By analogy, we propose that both families in **X.tropicalis**, or at least the 1kb transcription unit, may encode embryonic-type U6 RNAs, although no experimental data are available yet.

We have shown that U4 genes are also encoded by two long tandemly repeated gene families that are not harbored by U6 DNA fragments or interspersed with them. Therefore, data from our work and from others point

to the fact that in frogs, the major U1 (17), U2 (29), U4, U5 (I.Mattaj,personal communication) and U6 snRNA genes are organized in long tandemly repeated families. It is tempting to speculate that these highly reiterated genes are responsible for the synthesis of U snRNAs during early embryogenesis.

MODIFIED NUCLEOTIDES AND HIGH EVOLUTIONARY CONSERVATION IN THE U6 RNA MOLECULE

The cloned **X.tropicalis** U6 DNA was used as a template for microinjection studies in **X.laevis** oocytes. T1 RNAse fingerprint analyses showed it was faithfully transcribed, although the transcript made in these oocytes lacked the internal 2'-O-methylnucleotides found in cellular U6 RNAs (23,38).Similar findings were reported by Kunkel et al.(15) and Reddy et al. (14) who observed the absence of internal 2'-O-methylnucleotides in human and mouse U6 RNAs produced in whole cell extracts or in frog oocytes. Two possibilities can be envisaged for explaining this lack of methylation :

1) Some of the modification enzymes cannot cope with the vast amount of plasmid copies that are injected into oocytes. These enzymes may be absent or denatured in whole cell extracts.

2) More interestingly, some of the U6 RNA base modifications are nuclear whereas others are cytoplasmic. This would be reminiscent of the U snRNA series where cap trimethylation is cytoplasmic (4) and of eukaryotic tRNA molecules where only anticodon-loop modifications are cytoplasmic (H. Grosjean, personal communication). If this happens to be the case, it is worth recalling the absence in U6 RNAs of the Sm binding site whose presence is required in U2 RNAs (and most likely in U1, U4 and U5 RNAs as well) for nuclear segregation (39). Therefore, it may well be that U6 RNAs require the preliminary association with U4 RNA in order to be properly matured. The high content of 2'-O-methylated nucleotides in the central part of U6 RNA molecules makes likely the assumption that this domain is of prime importance for the functioning of U6 RNAs.

X.tropicalis U6 RNA has a single C⟶U transition (nucleotide 6 in the DNA sequence, Fig.3) from mammalian U6 RNAs (23,38).Das et al. (16) found 95% sequence homology between **Drosophila** and rat U6 RNAs, whereas U1 and U4 RNAs from these two species exhibit only 75 % sequence homology. Due to the high extent of sequence conservation in U6 RNAs and to the subsequent absence of compensating base changes in the folding proposed by Rinke et al.(25), we could not use phylogeny as a tool for a secondary structure modelling of the U6 RNA molecule. This extreme sequence conservation through evolution is

puzzling and rather uncommon when looking at other RNA molecules of similar sizes. One reason may reside in the putative function of U6 RNAs in premRNA splicing (7) that possibly necessitates not only structural constraints but also a high stringency of nucleotide sequence conservation. With respect to this, it is remarkable that the large RNA in **S.cerevisiae** which shares homology with metazoan U2 RNAs, contains a sequence showing good homology with nucleotides 13-58 of U6 RNAs (2).

PRESENCE OF THREE UPSTREAM REGULATORY POL II SIGNALS IN THE POL III - DEPENDENT U6 RNA GENES

We have assayed template activity as well as polymerase dependence of pXTU6-2 by injection of the DNA into **X.laevis** oocyte nuclei along with α-amanitin. According to the α-aminitin concentration that shut off transcription, we concluded that **X.tropicalis** U6 RNA transcription is RNA polymerase III - dependent. Independent work performed on **Drosophila**, mouse and human U6 RNA genes came to the same conclusion (14-16). However, as described in the Results section, inspection of **X.tropicalis** U6 RNA gene upstream sequences revealed the presence of those three elements that are generally found in RNA polymerase II promoters (cf. Table I) :

- a TATA box at -26 to -29, typically found in mRNA gene promoters (40).
- a TCTC box at -56 to -65, the proximal element typical for Xenopus snRNA gene promoters.
- an inversely repeated enhancer element at -231 to -246, the distal element whose octamer motif ATGCAAAY is found both in mRNA (41,42) and snRNA gene (26,30-33) promoters.

Comparison of upstream sequences in mouse (43) and human (Kunkel, unpublished results) U6 RNA genes showed the presence at homologous positions of the same three Pol II promoter elements, clearly establishing that these sequences do not occur by chance in **X.tropicalis** U6 RNA genes (cf. Table I). In mouse and man, the enhancer-type element is inverted relative to its usual orientation in U1 and U2 genes ; the consensus for the -55 box, TYACCNTAAC is very similar to its human U2 counterpart, TCACCGCGAC (31). **Drosophila** U6 RNA genes (16) contain a TATA box but no sequence homology with distal and proximal elements of vertebrate U1-U6 snRNA genes. In the light of these data, two possibilities emerge as to the nature of the RNA polymerase that transcribes U6 RNA genes :

i) U6 RNA genes are transcribed by RNA polymerase II. U6 RNA coding regions all contain the sequence GAUUAGCA (cf. Table I) that has been

TABLE I

	DISTAL ELEMENT	PROXIMAL ELEMENT	TATA BOX	INTERNAL CONTROL REGION
X. TROPICALIS U6	ATTTGCATATGCAAAC -231	TCTCCTTAAG -56	TATA -26	49 AGATTAGCATGG
X. LAEVIS U1-U2 CONSENSUS	YATGYAAAY[a]	TCTCCNTATG[b]	--	--
MOUSE U6	ATTTGCATA[d] -220	TCACCCTAAC -57	TATA -28	49 AGATTAGCATGG
HUMAN U6			TATA -26	49 AGATTAGCATGG
MAMMALIAN U6 CONSENSUS	ATTTGCATA[d]	TYACCNTAAC		49 AGATTAGCATGG
HUMAN U2	CATGCAAAT[a] -214	TCACCGCGAC -50		--
DROSOPHILA U6	--	--	TATA -26	49 AGATTAGCATGG
BOX A CONSENSUS OF EUKARYOTIC tRNA GENES				RGYNNARY--GG

Sequences in U6 RNA genes showing homology with RNA polymerase II and III promoter elements. The sequences shown were obtained from the following references, as discussed in the text : X.tropicalis U6, this work ; X.laevis U1-U2 consensus (26) ; mouse U6 (43) ; human U6,(15) ; Drosophila U6 (16) ; Box A consensus (37). a,b,c : these sequences have been shown to be part of functional promoter elements, as discussed in the text. d : distal element found in the opposite orientation relative to that found in a. Y, R and N indicate pyrimidines, purines and any nucleotide, respectively.

shown to interact with U4 RNAs (25) in the intact U4/U6 snRNP particle. Therefore, its presence in U6 RNAs is an evolutionary constraint and its resemblance with the eukaryotic tRNA gene Box A consensus sequence (37) is only fortuitous (cf. Table I). Consequently, RNA polymerase III would take over U6 RNA gene transcription only in those conditions where RNA polymerase II is inhibited at 1µg/ml of α-amanitin.

ii) Alternatively, U6 RNA genes are transcribed by an RNA polymerase with the same sensitivity to α-aminitin as RNA polymerase III and the transcription complex requires the presence of the upstream elements used by RNA polymerase II-dependent genes in order to be formed. Consistent with this possibility is the observation that when RNA polymerase II transcription is inhibited by α-amanitin, U6 RNA transcription is stimulated (Fig. 4, lanes 3-5). Requirements of 5'-flanking sequences for RNA polymerase III in vitro synthesis has already been reported for 7SK (44) and 7SL (45) RNA genes, tRNA genes (46) and 5S RNA genes (47). Indeed, in addition to drug sensitivity,

several lines of evidence speak to U6 RNAs being RNA polymerase III-type transcripts :

- They are accurately transcribed in whole cell extracts (14).
- They all contain the sequence AGAUUAGCAUGG (nucleotides 48-59 of the DNA sequence) which is a close match to the Box A of the internal control region of eukaryotic tRNA genes : RGYNNARY-GG (37) (cf. Table I).
- They start with a purine and end with a run of thymines.
- They are transiently associated with a protein that is La-antibody precipitable (38).

Interestingly, while a sequence sharing homology with Box A of tRNA genes is found in U6 RNA genes, no sequence showing homology with Box B of tRNA genes (37) could be detected. Upstream sequences, such as the TATA box and the proximal and distal elements may compensate for a reduction of control functions farther downstream, acting as signals for vertebrate or invertebrate-specific transcription factors, as exemplified by the observation of Das et al. (16) : **Drosophila** U6 RNA genes, which possess a TATA box, but not the distal and proximal elements of vertebrate U snRNA genes, could be transcribed in **Drosophila** but not in mammalian extracts ; a mouse U6 RNA gene could be transcribed in mammalian but not in **Drosophila** extracts ; in contrast, a **Drosophila** $tRNA^{Asp}$ gene could be transcribed and processed in both type of extracts. Therefore, U6 RNA gene transcription may require the Box A-like element together with upstream signals. The proximal and distal elements found upstream of mouse, and **X.tropicalis** U6 RNA genes are identical to those found upstream of vertebrate U1(26-28,30,48), U2(31-33) and U5 (Mattaj, personal communication) RNA genes that are transcribed by RNA polymerase II. Their presence in U6 RNA genes could serve, in addition, the synchronized triggering of U1-U6 expression and the coordinate synthesis of U4-U6 RNAs, the latter two being transcribed by RNA polymerase II (10) and III respectively, by binding the same factors.

We postulate that U6 RNA genes are transcribed by a specialized transcription complex including an RNA polymerase with the same sensitivity to α-amanitin as RNA polymerase III and transcription factors also used by RNA polymerase II. Work is in progress in our laboratory to precisely define the promoter elements that control U6 RNA transcription.

ACKNOWLEDGEMENTS

We are grateful to Drs. Michaïl Fischberg, Georges Spohr and Walter Knöchel for their generous gifts of erythrocyte X.tropicalis genomic DNA and

to Dr. Wolfgang Meyerhof for his generous gift of X.tropicalis frogs. We also thank Ram Reddy, Gary Kunkel and Thoru Pederson for communicating their results prior to publication. This work was supported by grants from the CNRS.

*To whom correspondence should be addressed.

REFERENCES

1. Brunel, C., Sri Widada, J., and Jeanteur, P. (1985) Prog. Mol. Subcell. Biol. 9, 1-52.
2. Ares, M. (1986) Cell 47, 49-59.
3. Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Haendler, B., and Jacob, M. (1982) EMBO J. 1, 1259-1265.
4. Mattaj, I.W. (1986) Cell 46, 905-911.
5. Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H., and Lührmann, R. (1984) EMBO J. 3, 1357-1363.
6. Hashimoto, C., and Steitz, J.A. (1984) Nucl. Acids Res. 12, 3283-3293.
7. Black, D.L., and Steitz, J.A. (1986) Cell 46, 697-704.
8. Berget, S.M., and Robberson, B.L. (1986) Cell 46, 691-696.
9. Chandrasekharappa, S.C., Smith, J.H., and Eliceiri, G.L. (1983) J. Cell. Physiol. 117, 169-174.
10. Hoffman, M.L., Korf, G.M., Mc Namara, K.J., and Stumph, W.E. (1986) Mol. Cell. Biol. 6, 3910-3919.
11. Savouret, J.F., Cathala, G., Eberhardt, N.L., Miller, W.L., and Baxter, J.D. (1984) DNA 3, 365-376.
12. Hayashi, K. (1981) Nucl. Acids Res. 9, 3379-3388.
13. Theissen, H., Rinke, J., Traver, C.N., Lührmann, R., and Appel, B. (1985) Gene 36, 195-199.
14. Reddy, R., Henning, D., Das, G., and Busch, H. (1987) J. Biol. Chem., 262, 75-81.
15. Kunkel, G., Maser, R.L., Calvet, J.P., and Pederson, T. (1986) Proc. Natl. Acad. Sci. USA 83, 8575-8579.
16. Das, G., Henning, D., Busch, H., and Reddy, R. (1987) J. Biol. Chem., in press.
17. Lund, E., Dahlberg, J.E., and Forbes, D.J. (1984) Mol. Cell. Biol. 4, 2580-2586.
18. Grunstein, M., and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
19. Messing, J., and Vieira, J. (1982) Gene 19, 269-276.
20. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
21. Barrell, B.G. (1971) in Procedures in Nucleic Acids Research, Cantoni, G.L., and Davies, D.R., (Ed.) 2, 751-779.
22. Smith, G.E., and Summers, M.D. (1980) Anal. Biochem. 109, 123-129.
23. Epstein, P., Reddy, R., Henning, D., and Busch, H. (1980) J. Biol. Chem. 255, 8901-8906.
24. Harada, F., Kato, N., and Nishimura, S. (1980) Biochem. Biophys. Res. Commun. 95, 1332-1340.
25. Rinke, J., Appel, B., Digweed, M., and Lührmann, R. (1985) J. Mol. Biol. 185, 721-731.
26. Krol, A., Lund, E., and Dahlberg, J.E. (1985) EMBO J. 4, 1529-1535.
27. Zeller, R., Carri, M.T., Mattaj, I.W., and De Robertis, E.M. (1984) EMBO J. 3, 1075-1081.
28. Ciliberto, G., Buckland, R., Cortese, R., and Philipson, L. (1985) EMBO J. 4, 1537-1543.

29. Mattaj, I.W., and Zeller, R. (1983) EMBO J. 2, 1883-1891.
30. 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.
31. Ares, M., Mangin, M., and Weiner, A.M. (1985) Mol. Cell. Biol. 5, 1560-1570.
32. Mattaj, I.W., Lienhard, S., Jiricny, J., and De Robertis, E.M. (1985) Nature 316, 163-167.
33. Mangin, M., Ares, M., and Weiner, A.M. (1986) EMBO J. 5, 987-995.
34. Hernandez, N. (1985) EMBO J. 4, 1827-1837.
35. Neuman de Veguar, H.E., Lund, E., and Dahlberg, J.E. (1986) Cell 47, 259-266.
36. Hernandez, N., and Weiner, A.M. (1986) Cell 47, 249-258.
37. Ciliberto, G., Raugei, G., Costanzo, F., Dente, L., and Cortese, R. (1983) Cell 32, 725-733.
38. Rinke, J., and Steitz, J.A. (1985) Nucl. Acids Res. 13, 2617-2629.
39. Mattaj, I.W., and De Robertis, E.M. (1985) Cell 40, 111-118.
40. Breathnach, R., and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
41. Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A., and Chambon, P. (1986) EMBO J. 5, 387-397.
42. Singh, H., Sen, R., Baltimore, D., and Sharp, P.A. (1986) Nature 319, 154-158.
43. Ohshima, Y., Okada, N., Tani, T., Itoh, Y., and Itoh, M. (1981) Nucl. Acids Res. 9, 5145-5158.
44. Murphy, S., Tripodi, M., and Melli, M.L. (1986) Nucl. Acids Res. 14, 9243-9260.
45. Ullu, E., and Weiner, A.M. (1985) Nature 318, 371-374.
46. Larson, D., Bradford-Wilcox, J., Young, L.S., and Sprague, K.U. (1983) Proc. Natl. Acad. Sci. USA 80, 3416-3420.
47. Selker, E.U., Morzycka-Wroblewska, E., Stevens, J.N., and Metzenberg, R.L. (1986) Mol. Gen. Genet. 205, 189-192.
48. Korf, G., and Stumph, W.E. (1986) Biochemistry 25, 2041-2047.