The EMBO Journal vol.3 no.13 pp.3295-3301, 1984

Human U2 and U1 RNA genes use similar transcription signals

Gunnar Westin, Elsebet Lund[1], James T.Murphy[2], Ulf Pettersson and James E.Dahlberg[1]

Medical Genetics and Microbiology, Box 589, Biomedical Center, S-751 23 Uppsala, Sweden, [1]Department of Physiological Chemistry, and [2]McArdle Laboratory, University of Wisconsin-Madison, Madison, WI 53706, USA

Communicated by U.Pettersson

We have analyzed the requirements for human U2 RNA transcription by injection of cloned U2/6 RNA genes into nuclei of *Xenopus laevis* oocytes. Two forms of human U2 RNAs accumulate, a major species corresponding to mature-sized U2 RNA and a minor species corresponding to a 3'-extended precursor. This RNA polymerase II transcription requires only 258 and 94 bp of 5'- and 3'-flanking region sequences, respectively. Efficient U2 RNA synthesis depends on a promoter element located between positions −258 and −198. This region contains a 12-bp direct repeat which strongly resembles a comparable upstream promoter element of the human U1 RNA genes. Sequences between −258 and −198 also confer on the U2 RNA template the ability to compete with co-injected U1 RNA templates for a snRNA gene-specific transcription factor(s). Transcription of U2 RNA is reduced off templates containing an active RNA polymerase III transcription unit, presumably because of relaxation or sequestration of the DNA. *In vitro* transcription of the U2 RNA gene, like that of the U1 RNA gene, is initiated upstream of the point corresponding to the 5' end of *in vivo* synthesized RNA.

Key words: human U2 RNA synthesis/oocyte injection/RNA polymerase II transcription signals/small nuclear RNA genes

Introduction

The small nuclear RNAs (U1−U6) are abundant RNA species that are present, in the form of ribonucleoprotein particles, in the nuclei of all higher eucaryotes (for reviews, see Reddy and Busch, 1981; Busch *et al.*, 1982). U2 RNA, which is the subject of this study, is 188 nucleotides long and, like most of the U RNAs, contains a trimethylated cap structure and several base and ribose modifications (Reddy *et al.*, 1981; Nohga *et al.*, 1981; Van Arsdell and Weiner, 1984; Westin *et al.*, 1984). A precursor of U2 RNA, 10−12 nucleotides longer at the 3' end than the mature form, has been observed as a transient cytoplasmic intermediate in human and rat cells (Eliceiri and Sayavedra, 1976; Tani *et al.*, 1983).

The genes for human U2 RNA have been cloned by others and ourselves (Van Arsdell and Weiner, 1984; Westin *et al.*, 1984) who found that they are clustered in a large tandem array consisting of ~20 repeating units, each of which is 6.2 kb long. Sequence analyses of cloned U2 RNA genes from a variety of species (Mattaj and Zeller, 1983; Nojima and Kornberg, 1983; Tani *et al.*, 1983; Beck *et al.*, 1984; Van Arsdell and Weiner, 1984; Westin *et al.*, 1984) have shown that they have several features in common with other vertebrate snRNA genes (Roop *et al.*, 1981; Manser and Gesteland, 1982; Marzluff *et al.*, 1983; Watanabe-Nagasu *et al.*, 1983; Zeller *et al.*, 1984; Lund and Dahlberg, 1984; Skuzeski *et al.*, 1984). For example, they lack a TATA-box that is normally present ~30 nucleotides upstream of the position corresponding to the 5' end of RNA polymerase II transcripts; instead they share a consensus sequence located around positions −50 to −60, which is functionally equivalent to a TATA-box (Skuzeski *et al.*, 1984). Thus it is likely that all small nuclear RNAs are transcribed by the same general mechanism, including the use of common signals for transcription initiation and termination.

We and others have previously used *Xenopus laevis* oocytes to study transcription of cloned snRNA genes (Murphy *et al.*, 1982; Nojima and Kornberg, 1983; Mattaj and Zeller, 1983; Zeller *et al.*, 1984; Lund *et al.*, 1984; Skuzeski *et al.*, 1984). The occytes are capable of faithful transcription and of carrying out almost all of the post-transcriptional nucleotide modifications of the snRNAs. In this way we have determined that transcription of a human U1 RNA gene is initiated at the 5' cap site, but that sequences upstream of position −203 are required for template activity (Skuzeski *et al.*, 1984). In contrast, Zeller *et al.* (1984) reported that only 149 bp of 5'-flanking region sequences are sufficient for efficient transcription of a *X. laevis* U1 RNA gene. It is therefore of interest to determine the transcriptional requirements for sequences far upstream in human U2 RNA templates.

We report here that cloned human U2 RNA genes are capable of directing the synthesis of U2 RNA in *X. laevis* oocytes. We observe two forms of U2 RNA, one corresponding to the mature sized U2 RNA and another (usually minor species), corresponding to the precursor of U2 RNA observed in mammalian cells. We find that 258 bp of upstream sequences are sufficient for transcription and that a region located between positions −258 and −198 is required for both efficient U2 RNA synthesis and competition between co-injected U1 and U2 snRNA gene templates but not for transcription in cell-free extracts.

Results

Transcription of a human U2 RNA gene in X. laevis oocytes

We have previously described the isolation and characterization of several genes for human U2 RNA (Westin *et al.*, 1984). To test whether these genes were transcriptionally active, fragments containing the U2 RNA coding region sequences and ~3 kb and 1 kb of 5'- and 3'-flanking region sequences, respectively, (see Figure 1A) were injected into the nuclei of *X. laevis* occytes. Total RNA synthesized during the subsequent 20 h incubation was analyzed by 10% polyacrylamide gel electrophoresis.

As shown in Figure 2 (lanes 1 and 2) gel analysis of the RNA products showed that the synthesis of two U2 RNA-sized molecules (U2 and preU2) and several other large species (A−D) was dependent on injection of the cloned DNA. The U2 and preU2 RNAs were identified as human U2

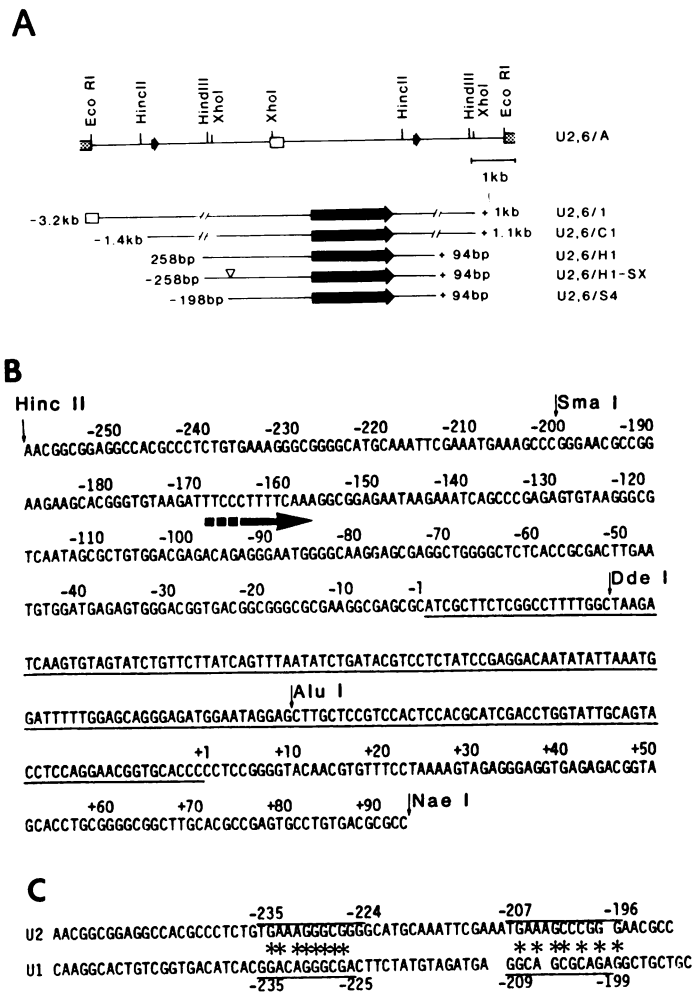

Fig. 1. Templates used for *in vivo* and *in vitro* U2 RNA transcription. **(A)** A restriction enzyme cleavage map of clone U2,6/A is shown at the top of the figure. This clone contains one-and-a-half repeat units of the human U2 RNA gene (Westin *et al.*, 1984). The U2 coding region is shown by the solid horizontal arrow and the 5′- and 3′-flanking sequences are denoted by thin lines. The unfilled box indicates the position at which repetitive sequence elements related to the Alu family are found. The lower five lines illustrate, on an expanded scale, five different subclones containing fragments from the U2/6 locus. Subclone H1-SX contains a *Xho*I linker inserted in the *Sma*I cleavage site at position −198 (▽). **(B)** The nucleotide sequence of a gene for human U2 RNA and its immediate flanking sequences. The nucleotide sequence between positions −213 and +94 has been published before (Westin *et al.*, 1984, Hammarström *et al.*, 1984). The sequence between positions −213 and −258 was determined by sequencing upstream from the *Ava*I cleavage site at position −197 and downstream from an *Eco*RI cleavage site located within the vector. The U2 RNA coding region is underlined and relevant restriction sites are indicated. The site of *in vitro* transcription initiation is shown by the arrow. **(C)** A sequence comparison between 5′-flanking sequences in the U1 and U2 genes. The nucleotide sequence between −258 and −190 is shown for both genes. Repetitive sequence elements which are present in both genes are indicated by the lines. Homologous nucleotides are indicated by asterisks. The U1 gene sequence is from Lund and Dahlberg (1984).

RNA transcripts by RNase T1 fingerprinting (see Figure 3, below). Low concentrations of α-amanitin (1−2 μg/ml) inhibited the synthesis of the U2 transcripts but not of the larger A−D RNAs (lanes 8−10). These results indicated that the transcription of human U2 RNA was catalyzed by RNA polymerase II whereas that of the larger RNAs was probably by RNA polymerase III.

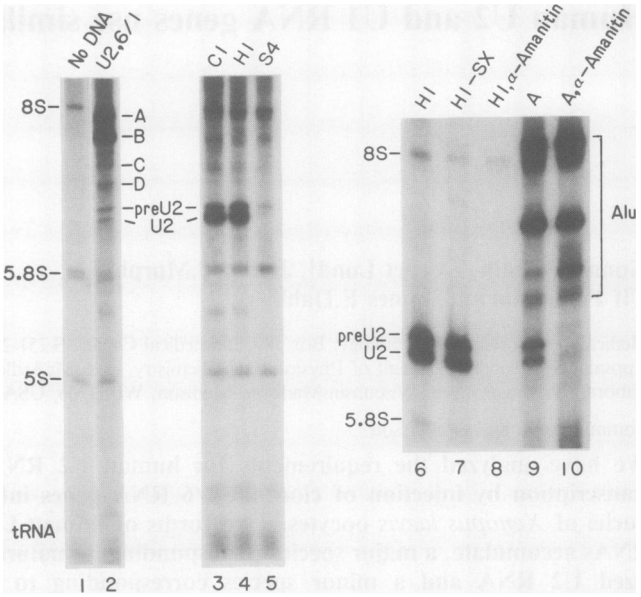

Fig. 2. Human U2 RNA and Alu-type transcripts synthesized in *X. laevis* oocytes. Gel electrophoresis of total RNAs from pooled oocytes in 8% (30:0.8) polyacrylamide containing 7 M urea. Oocytes were injected 20 h prior to RNA extraction with [α-^{32}P]GTP and the following DNA templates: U2,6/1 **(lane 2)**, U2,6/C1 **(lane 3)**, U2,6/H1 **(lanes 4, 6 and 8)**, U2,6/S4 **(lane 5)** or U2,6/A **(lanes 9 and 10)**. In addition, α-amanitin (to give a final concentration of 1−2 μg/ml) was co-injected into oocytes represented in **lanes 8 and 10**. For clarity, only the portion of the gel containing U2 RNA and Alu-type transcripts (A−D) is shown in **lanes 6−10**.

Characterization of the U2 RNA transcripts

The RNase T1 fingerprint of the RNA-labeled U2 in Figure 2 (Figure 3A) is identical to that of human U2 RNA (Nogha *et al.*, 1981) and it matches the fingerprint expected for a transcript synthesized off the human U2/6 gene. Base modifications of the RNA were studied by RNase T2 digestion of gel-purified RNAs. As shown in Figure 3B, the mature form had all of the expected modified nucleotides that could be labeled by [α-^{32}P]GTP, including the 5′ trimethylated cap structure.

The fingerprint of preU2 RNA was very similar to that of U2 RNA except for the presence of one major (arrow) and several minor (a−d) new oligonucleotides. In addition, oligonucleotides 17, 22 and 26 (stars), which contain 2′ O-methyl cytidylic and guanylic acids (Reddy *et al.*, 1981; cf. Figure 3C), were decreased in yield. The major new oligonucleotide (an 11-mer designated e) corresponded to the sequence that overlapped the 3′ end of the mature U2 RNA; RNase T2 redigestion of this oligonucleotide yielded labeled Cp and Gp (data not shown) indicating that preU2 RNA was elongated by a least seven nucleotides at its 3′ end (cf. Figure 3C). The mobilities and redigestion analyses of the other new, minor oligonucleotides (a, b, c and d) showed that they were derived from the unmethylated forms of the sequences of oligonucleotides 17, 22 and 26 (data not shown; cf. legend to Figure 3C).

The reduction in the level of ribose methylation in the preU2 RNA is evident in the RNase T2 digest shown in Figure 3B. We note, however, that the 5′ cap structure, which includes two methylated ribose moieties (Reddy *et al.*, 1981), is fully modified in preU2 RNA. In human cells, preU2 RNA has been shown to be a precursor of U2 RNA (Elicieri and

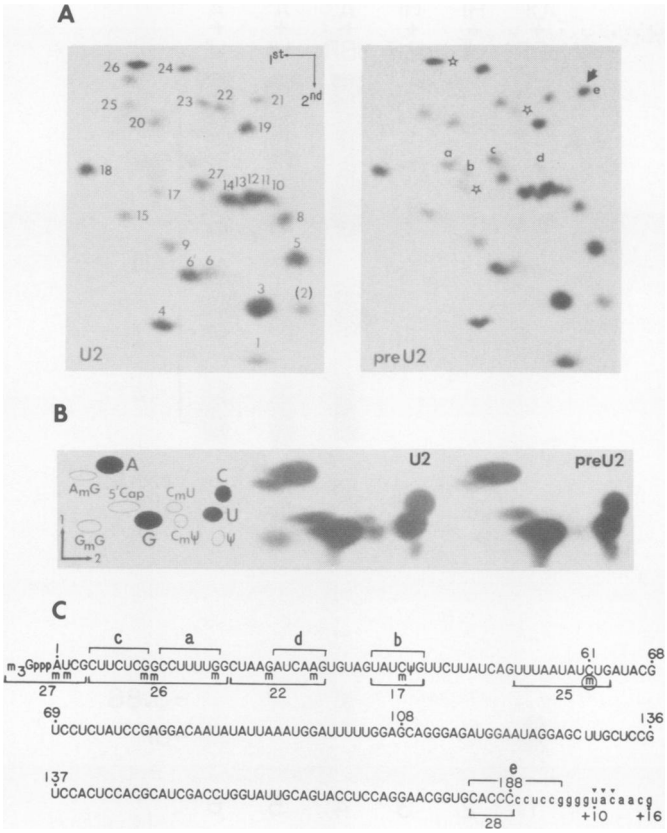

Fig. 3. Fingerprint analysis of human U2 RNA synthesized in oocytes. **(A)** Two-dimensional RNase T1 fingerprint analysis of the RNAs eluted from bands labeled U2 and preU2 in Figure 2. Oligonucleotides are numbered according to Nohga *et al.* (1981). In preU2 RNA a, b, c, d and e indicate new oligonucleotides which are precursor-specific; the heavy arrow indicates the 3′ overlap oligonucleotide and stars indicate oligonucleotides # 17, 22 and 26 which are present in less than molar amounts (see text). The oligonucleotide labeled (2) is C-Gp which is an abundant contaminant of U2 RNA preparations (cf. Reddy *et al.*, 1981). The first dimension was high volt electrophoresis on cellulose acetate at pH 3.5 and the second was homochromatography on PEI thin layer plates. **(B)** Modified nucleotide analysis of U2 and preU2 RNAs. RNAs were digested with RNase T2 and the resulting nucleotides were separated by two-dimensional chromatography. The left hand panel is a schematic representation of the pattern observed with uniformly labeled U2 RNA. The first dimension was solvent a and the second was solvent c of Silberklang *et al.* (1979). **(C)** Nucleotide sequence of human U2 RNA transcripts indicating the positions of the oligonucleotides discussed in the text. Only the nucleotide modifications that could be observed using [α-^{32}P]GTP are noted. Oligonucleotides a – d (shown above the sequence) contained no ribose-methylated nucleotides, but b does contain ψp. The (m) of residue 61 indicated that this modification was not observed in preU2 RNA. Triangles above the sequence indicate the proposed 3′ end of preU2 RNA. The transcripts were identified as human rather than frog U2 RNAs since they did not produce the oligonucleotide A-C-C-A-G-p; that oligonucleotide results from a Cp rather than Gp at position 108 in *X. laevis* U2 RNA (Zeller and Mattaj, 1983; E.Lund, unpublished).

Sayavedra, 1976). By pulse labeling (with [α-^{32}P]GTP followed by a chase in the presence of α-amanitin) we found that preU2 RNA synthesized in oocytes was converted to U2 RNA with a half-life of several hours (data not shown). We conclude that preU2 RNA represents the U2 RNA precursor previously identified in human cells (Eliceiri and Sayavedra, 1976).

Characterization of the α-amanitin-resistant transcripts

The abundant, α-amanitin-resistant transcripts (A – D in lane 2; cf. Figure 2) were also analyzed by RNase T1 fingerprinting. As shown in Figure 4, transcripts A and B have identical fingerprints with the exception of one major oligonucleotide (open arrow in Figure 4A) that remains at the origin of the second dimension. Very similar fingerprints were obtained from C and D RNAs (data not shown) indicating that they probably arose from the same transcription unit as the A and B RNAs. However, these fingerprints bore no relation to that of U2 RNA; instead, they resembled fingerprints of RNAs transcribed by RNA polymerase III, using Alu-type repetitive DNA sequences as template (Elder *et al.*, 1981; Pan *et al.*, 1981). Presumably these transcripts were synthesized from the Alu-like sequences located ~3 kb upstream of the U2/6 RNA gene (Westin *et al.*, 1984). Consistent with this conclusion is the fact that clone C1, which contains only 1.4 kb of 5′-flanking region sequences, does not direct the synthesis of these Alu-type RNAs (Figure 2, lane 3).

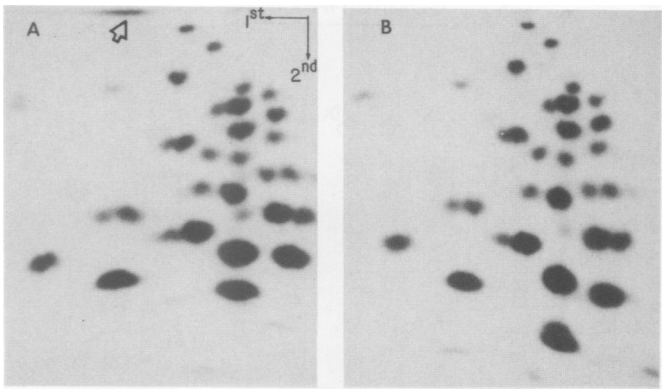

Fig. 4. RNase T1 fingerprints of the Alu-type transcripts A and B indicated in lane 2 of Figure 2. The open arrow indicates the only detectable oligonucleotide difference between these two RNAs.

Requirements for 5′-flanking sequences

To determine which upstream sequences were necessary for template activity a series of subclones of the U2 RNA gene, with progressively smaller 5′-flanking regions (cf. Figure 1A and B), were injected. As shown in Figure 2 (and analyzed in more detail below) DNAs lacking Alu-type sequences were more efficient templates for U2 RNA transcription than were DNAs containing these sequences (compare lanes 2 and 9 with lanes 3, 4, 6 and 7). A clone with 258 bp of upstream flanking region sequences (H1) appeared to be as efficient a template as clone C1 (Figure 2, lanes 3 and 4). In contrast, a clone with only 198 nucleotides of upstream sequence (S4) was a very poor template for human U2 RNA synthesis (Figure 2, lane 5). Oocytes injected with this latter clone did accumulate a small amount of preU2 RNA (estimated to be ~1 – 2% of the normal level of total U2 RNA transcripts). Therefore, we conclude that at least some of the sequences between positions – 258 and – 198 are required for highly efficient transcription, although they are not absolutely essential for template activity. A template with an 8-bp insertion at position – 198 (H1-SX) functioned like wild-type DNA (Figure 2, lanes 6 and 7). Thus, there appears to be no requirement for a fixed distance between sequences upstream of position 198 and the point at which RNA synthesis starts; also, the precise nucleotide sequence of the region around position – 198 does not appear to be essential.

Competition between U1 and U2 RNA genes

We have recently found that sequences upstream of position

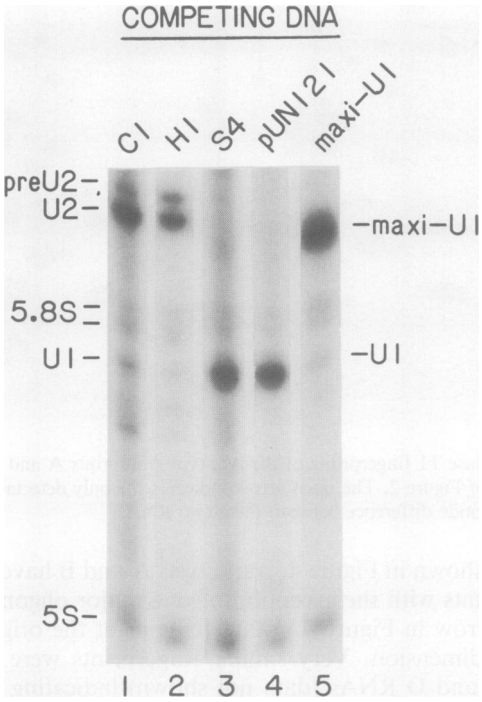

Fig. 5. Competition between co-injected human U2 and U1 RNA templates. 1 ng of the HU1-1D template containing a functional U1 RNA gene (Murphy *et al.*, 1982) and 10 ng of the competing DNA were co-injected with [α-^{32}P]GTP into *X. laevis* oocytes. After 20 h of incubation, the RNAs from single oocytes were isolated and analyzed by electrophoresis in a 10% polyacrylamide gel containing 7 M urea. The competing DNAs were U2,6/C1 **(lane 1)**, U2,6/H1 **(lane 2)**, U2,6/S4 **(lane 3)** the vector of the U2,6-clones, pUN121 **(lane 4)** and a 'maxi-U1 RNA' gene (Skuzeski *et al.*, 1984) **(lane 5)**. An autoradiograph of a portion of the gel is shown.

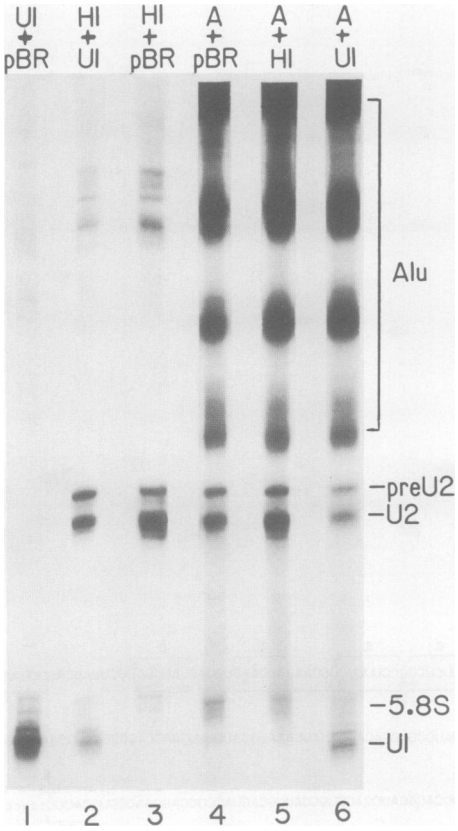

Fig. 6. Effects of Alu-type sequences on the template activity of U2 RNA genes. Equal amounts (5 ng each) of two DNA templates were co-injected into oocytes and the RNAs synthesized were analyzed as in Figure 5. (A total of 5 ng DNA was sufficient to give maximal RNA synthesis; J.T.Murphy, unpublished.) The RNAs were displayed by electrophoresis in an 8% polyacrylamide ultra-thin (sequencing) gel containing 7 M urea. This method separates human U1 and U2 transcripts into multiple bands, resulting from 3′ end-heterogeneity of the RNAs synthesized in oocytes. The DNAs injected were HU1-1D (U1), pBR322 (pBR), U2,6/H1 (H1) and U2,6/A (A). The mobilities of relevant RNA species are indicated on the right.

−203 in the human U1 RNA gene confer on the template the ability to compete with other U1 DNA templates for factor(s) that are limiting in oocytes. This factor is specific for snRNA genes rather than for RNA polymerase II transcription units in general, since a template with the major late promoter of adenovirus 2 does not compete (J.M.E.Lund, J.T.Murphy and J.E.Dahlberg, in preparation). To determine whether the human U2 RNA gene also contained such sequences, we tested the ability of the U2 RNA templates described above to compete with the wild-type human U1 RNA gene. In these experiments, a 10-fold excess of the U2 RNA templates (cf. Figure 1) was mixed with wild-type U1 RNA templates prior to injection into oocytes. As shown in Figure 5, U2 RNA genes containing 258 or more nucleotides of upstream sequence (lanes 1 and 2) competed well with the wild-type U1 RNA gene. However, the U2 RNA template with only 198 nucleotides of 5′-flanking region (lane 3) did not affect transcription of the U1 RNA gene, Thus, a functional element exists between positions −258 and −198 that provides the U2 RNA template with the ability to usurp one or more factors needed for snRNA transcription. We note that this is the same region that is required for efficient transcription of the U2 RNA gene in frog oocytes (cf. Figure 2 above). Lanes 4 and 5 (cf. Figure 5) show that the vector DNA alone does not compete with the U1 RNA gene, but a template with a U1 'maxi-RNA' gene (Skuzeski *et al.*, 1984) does.

Effect of Alu-type sequences on snRNA gene transcription

As noted above in Figure 2, the presence of Alu-type sequences appeared to reduce the overall level of U2 RNA transcripts in injected oocytes (compare lanes 2 and 9 with lanes 3, 4, 6 and 7). To determine whether that reduction resulted from competition for transcription factors and/or nucleotide precursors or from alterations of the template structures, we analyzed the effects of injecting mixtures of U2 RNA templates that did or did not contain Alu-type sequences.

The results represented in Figure 6 (lanes 1−3) indicate that snRNA genes compete with each other for a factor(s) that appears to be limiting when oocytes are injected with 10 ng of template DNA (cf. Figure 5). The levels of U1 or U2 RNAs accumulating in oocytes co-injected with both templates is significantly lower than the levels accumulating in oocytes that received only one or the other DNA. Comparison of the levels of RNAs accumulated using templates with or without Alu-type sequences demonstrates that the presence of these sequences in the U2 template decreases the level of U2 RNAs (lanes 3 and 4). This reduction is seen to work in *cis* but not *trans* since the level of U2 RNA is not significantly reduced when U2,6/H1 DNA (lacking Alu-type sequences) is co-injected with U2,6/A DNA (containing Alu-type sequences (compare lanes 3 and 5). Likewise, we observe that the level of accumulated U1 RNA is the same regardless

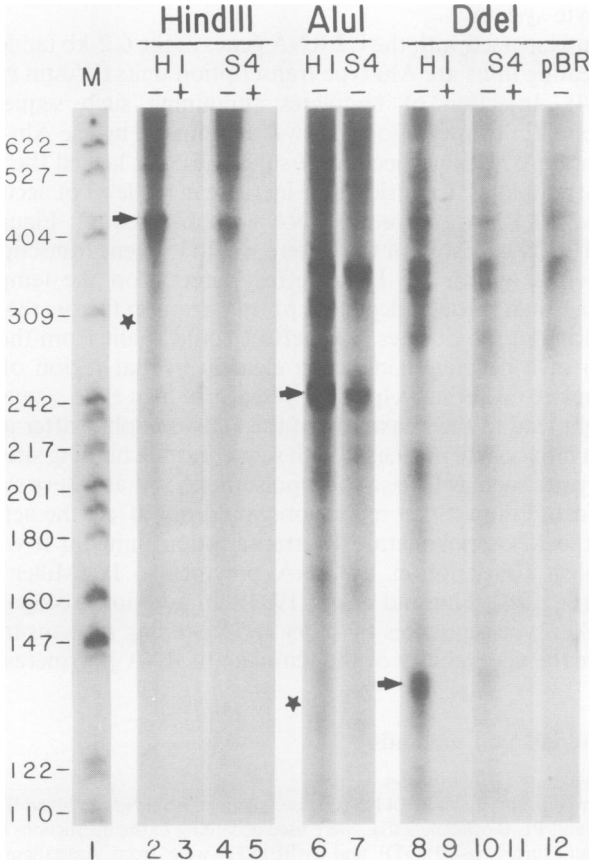

Fig. 7. *In vitro* transcription of U2 RNA templates. RNA 'run off' transcripts synthesized in HeLa cell extracts in the absence (−) or presence (+) of 1 μg/ml α-amanitin were displayed in an 8% (30:0.8) polyacrylamide gel containing 7 M urea. The RNAs analyzed were transcripts made using the following templates: U2,6/H1 DNA cut with *Hind*III **(lanes 2,3)**, *Alu*I **(lane 6)** or *Dde*I **(lanes 8,9)**, U2,6/S4 DNA cut with *Hind*III **(lanes 4,5)**, *Alu*I **(lane 7)**, *Dde*I **(lanes 10,11)** or pBR322 vector DNA cut with *Dde*I **(lane 12)**. Size markers of end-labeled *Msp*I cut pBR322 DNA are shown in **lane 1**. Arrows indicate the major U2 RNA specific 'run off' transcripts, estimated to be 380 (*Hind*III), 220 (*Alu*I) and 115 (*Dde*I) nucleotides long. Asterisks indicate the positions of transcripts if initiation had occurred at position −1, calculated to be 285 (*Hind*III) and 127 (*Alu*I) nucleotides long. The gel mobility of single-stranded RNA is ~14% slower than that of single-stranded DNA of the same length.

of whether or not the co-injected U2 RNA template contained Alu-type sequences (lanes 2 and 6).

Transcription of the U2 RNA gene in vitro

We have shown elsewhere that transcription of human U1 RNA genes is not initiated at the same position of the template *in vivo* as it is *in vitro*. In injected oocytes, initiation occurs at the site corresponding to the 5′ nucleotide of normal U1 RNA, whereas in cell-free extracts transcription is initiated 183 nucleotides upstream of that point (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984).

As shown in Figure 7, transcription of human U2 RNA genes resembled that of U1 RNA genes in that initiation in a cell-free extract occurred only at a position considerably upstream of the point corresponding to the 5′ end of normal U2 RNA. This transcription is sensitive to 1 μg/ml α-amanitin. In the case of U2 RNA 'run off' transcripts, we mapped the start site to around position −95. This was confirmed by nuclease S1 mapping of the 5′ ends of transcripts made from circular U2 RNA templates (data not shown).

It was expected that the *in vitro* transcriptional activity of U2,6/S4 (−198) would be decreased ~100-fold relative to U2,6/H1 (−258) because this is their relative activity in oocytes. Surprisingly, both of the templates, U2,6/H1 and U2,6/S4, were very active in the cell-free transcription extracts although U2,6/S4 is ~3-fold less active *in vitro* than U2,6/H1 (compare lanes 2, 6 and 8 with lanes 4, 7 and 10 of Figure 7). This is in contrast to the case with the U1 gene where there is a close correspondence between *in vivo* and *in vitro* activities. Thus some sequences located between positions −198 and −258 of the U2 gene are required for *in vivo* but not *in vitro* transcription.

Discussion

We have demonstrated here that the cloned human U2 RNA gene, U2/6 can be transcribed in *X. laevis* oocytes. Two forms of the U2 RNA transcript were observed, corresponding to the mature and precursor forms seen in human cells. The RNAs were identified as being derived from the human gene since RNase T1 fingerprints of the U2 RNA and the preU2 RNA (Figure 3A) are entirely consistent with the published fingerprint for human U2 RNA, but not with the fingerprint of U2 RNA synthesized in oocytes from endogenous *X. laevis* U2 RNA genes (E.Lund, unpublished results; see also Mattaj and Zeller, 1983). Also, in the polyacrylamide gels used here, the human U2 RNAs migrated significantly slower than frog U2 RNA (data not shown).

Both of the U2 RNA transcripts contained the 5′ end trimethyl-G-cap as well as the methyl groups on the first and second ribose units in the polynucleotide chain. Since preU2 RNA appeared to be generally deficient in methylation of internal ribose groups but not of the ribose moieties adjacent to the 5′ cap, we conclude that different methylases are probably operative and that 5′ end maturation occurs prior to internal maturation.

The 3′ ends of the U2 transcripts were deduced by analysis of the sizes of the RNAs and their RNase T1 fingerprints; we were unable to determine the locations exactly since the RNAs were labeled only with [α-^{32}P]GTP. The accumulated U2 RNA lacked the oligonucleotide that would overlap the 3′ end of mature U2 RNA and extend to position 194 (cf. Figure 3C); therefore the end point must be downstream of position 194. The fingerprint of preU2 RNA contained all oligonucleotides encoded by sequences through position 195 but not up to position 204; thus the 3′ end of that RNA is located between positions 195 and 204. In agreement with the oligonucleotide analysis, the electrophoretic mobilities of U2 RNA and preU2 RNA indicated that they were ~188 and 200 nucleotides long, respectively, comparable with the lengths of U2 RNA and preU2 RNA synthesized in human cells (Eliceiri and Sayvedra, 1976).

In some experiments (Figures 5 and 6) in which we analyzed the RNAs isolated from individual oocytes, rather than from pools of oocytes, we noted considerable variation in the abilities of individual oocytes to accumulate U2 RNA as opposed to preU2 RNA. Although the level of preU2 RNA was approximately constant, the level of U2 RNA ranged from being undetectable to being very high (data not shown). It is unclear why particular oocytes appear to be unable to produce mature U2 RNA. Nevertheless, that variation could account for some of the differences in the relative levels of preU2 RNA and U2 RNA we have observed in analyses of RNAs pooled from several oocytes (e.g., compare lanes 4 and 6 of Figure 2).

Like the expression of human U1 RNA genes, the efficient transcription of human U2 RNA genes in *X. laevis* oocyte requires a considerable amount of 5'-flanking region sequence. Templates containing 258 nucleotides of this region were very active in transcription whereas templates with only 198 nucleotides were inefficient, although not entirely inactive. Another property of the region between positions −258 and −198 is that it confers on the template the ability to compete effectively with other snRNA templates for one or more transcription factors that appear to be limiting in the frog oocytes (Figure 5). It is unclear how this ability to compete for factors is related to the overall efficiency of transcription.

We have previously shown that U1 DNA templates lacking elements between positions −231 and −203 are inactive in transcription of U1 RNA (Skuzeski *et al.*, 1984). It was therefore of interest to search for homologous sequences in this region of U1 and U2 genes. Skuzeski *et al.* (1984) noted that the region between −235 and −199 in the human U1 gene contained a direct repeat of an 11 nucleotide long sequence that shared properties with enhancer sequences. Examination of the corresponding region in the U2 gene reveals the presence of a 12 nucleotide long direct repeat in the corresponding position (see Figure 1C). The sequence TGAAAG (G,C)(G,C)CGGG occurs between −235 and −224 as well as between −207 and −196. It is noteworthy that this repeated sequence is closely related to the core sequence, found to be of critical importance for viral enhancer sequences (Weiher *et al.*, 1983). It thus seems likely that the far upstream sequence, both in U1 and U2 genes, has an enhancer-like function.

In vitro transcription of both U1 and U2 RNA genes initiates in the 5'-flanking regions rather than at the sites corresponding to the 5' end of *in vivo* synthesized RNA (Figure 7 and Murphy *et al.*, 1982). Thus, correct initiation of transcription of both genes requires a factor(s) that is missing from the cell-free extracts (Skuzeski *et al.*, 1984). Although there are no extensive regions of homologous sequence in comparable locations upstream of the *in vitro* start points, the two genes share essential sequences closer to their coding regions. Between positions −41 and −62 in both genes is a sequence that is necessary for fixing the point of initiation *in vivo* (at least in the case of U1 RNA; Skuzeski *et al.*, 1984). Transcription of the U1 and U2 RNA templates differs in that a deletion mutant of the U2 RNA gene which is almost inactive in injected oocytes still functions efficiently as a template *in vitro*; in contrast, deletions in the 5'-flanking region of a U1 RNA gene affect *in vitro* and *in vivo* transcription to comparable extents. It is still unclear, however, how the *in vitro* transcription might be related to that observed *in vivo*.

Van Arsdell and Weiner (1984) noted two regions of sequence homology between human and mouse U1 and U2 RNA genes which they termed Regions I and II. The two blocks of homology are separated from each other by about the same distance (22−25 nucleotides). Regions I and II are found relatively close to the coding region of U2 RNA genes (~35 bp) but far upstream of position −1 of U1 RNA genes (~210 bp). U1 RNA template with only 231 nucleotides of 5'-flanking sequences lacks the entire Region I but is still transcribed efficiently in oocyte nuclei (Skuzeski *et al.*, 1984); thus Region I is not required for transcription. In contrast, the U2 RNA template S4 contains both Regions I and II but is poorly transcribed in this system. Therefore, these conserved regions are not sufficient for efficient transcription in the oocyte system.

Interspersed with the U2 RNA genes in the 6.2-kb tandemly repeating units are Alu-type transcription units (Westin *et al.*, 1984). Injection of templates containing such sequences resulted in transcription that was dominated by the Alu-type transcripts. Upon injection of subclones that lacked the Alu-type templates, a considerable increase in the level of accumulated U2 RNA and preU2 RNA was observed (cf. Figures 2 and 6). The reason for the differences in U2 gene transcription activities appears to be an effect directly on the template rather than through depletion of transcription factors that are limiting in the oocytes. This effect could result from the action of a negative controlling element in that region of the template. Alternatively, the presence of Alu-type sequences might lead to the relaxation of the DNA template after injection into oocyte nucleus if such sequences are binding sites for enzymes such as DNases or topoisomerases that relieve superhelical tension. Such relaxation would not affect the activity of the RNA polymerase III transcription unit but it would abolish transcription by RNA polymerase II (Miller and Mertz, 1982; Harland *et al.*, 1983). In addition, recognition of Alu-type sequences by other DNA binding proteins might alter the accessibility of the template to RNA polymerase II.

Materials and methods

Enzymes and triphosphates

Restriction enzymes and T4 DNA ligase, obtained from New England Biolabs Inc. and PL-Biochemical Co., were used according to the instructions of the manufacturers. [α-^{32}P]GTP and [γ-^{32}P]ATP were from Amersham Ltd. Nuclease S1 was from Boehringer-Mannheim.

Construction of DNA templates

The templates were derived from a clone of human fetal liver DNA in bacteriophage λ DNA, designated U2/6 (Westin *et al.*, 1981, 1984). This clone contains three U2 gene copies. Smaller fragments from this template were subcloned using either pBR322 (Bolivar *et al.*, 1977) or pUN121 (Nilsson *et al.*, 1983) DNAs as vectors, to make subclones U2,6/1, U2,6/H1-SX and U2,6/H1, U2,6/S4, U2,6/C1 and U2,6/A, respectively. The structures of the U2/6 locus and the cloned fragments used as templates are illustrated in Figure 1. Clone U2,6/1 contains a 4.4-kb long *Pst*I cleavage fragment (Westin *et al.*, 1984). Clones U2,6/A and U2,6/C1 contain a 9.5-kb long *Eco*RI and a 2.7-kb long *Eco*RI/*Hind*III fragment, respectively. In U2,6/H1, a *Hinc*II/*Nae*I fragment (extending from position −258 to +94) was inserted into the pUN121 vector using *Hind*III linkers. In U2,6/S4, a *Sma*I/*Nae*I fragment (between positions −198 and +94) was inserted, again using *Hind*III linkers. In clone U2,6/H1-SX, a *Xho*I linker (5'-C-C-T-C-G-A-G-G-3') was inserted into the *Sma*I cleavage site (at position −198) of clone U2,6/H1 after that fragment had been recloned into pBR322.

The U1 RNA gene was HU1-ID DNA containing 392 and 35 nucleotides of 5'- and 3'-flanking region sequences (Murphy *et al.*, 1982; Lund and Dahlberg, 1984). The 'maxi-U1' RNA gene (called Bi 33+ in Skuzeski *et al.*, 1984) was the same as HU1-ID DNA except that the coding region had a duplication of 33 nucleotides at its 5' end, resulting in the production of an RNA 33 nucleotides longer than normal U1 RNA.

DNA sequence analysis

The protocol of Maxam and Gilbert (1980) was used for sequencing and 5' end-labeling. The sequence strategy is described in the legend to Figure 1.

Transcription activity analysis

X. laevis oocyte injections were performed as previously described (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984) using [α-^{32}P]GTP as label and supercoiled DNA templates containing both human and vector DNA sequences. Total nucleic acids, extracted from pooled or individual oocytes 20 h after injection, were analyzed by polyacrylamide gel electrophoresis.

Gel-purified RNAs were analyzed by RNase T1 fingerprinting (Sanger *et al.*, 1965). Modified nucleotide analysis was done according to the methods of Nishimura (1972) and Silberklang *et al.* (1979).

In vitro transcription of U2 RNA was carried out as previously described (Murphy *et al.*, 1982) using *Alu*I-, *Dde*I- or *Hind*III cleaved U2,6/H1 and U2,6/S4 DNAs (cf. Figure 1) as templates.

Acknowledgements

We are grateful to Ms Jeanette Backman, D.Stalsberg and P.Tregloan for typing the manuscript. Ms Elsy Johnsen and Mr Thomas Steinberg provided skilled technical assistance. This study was made possible by grants from the Swedish Medical Research Council, the Swedish National Board for Technical Development, the Marcus Borgstrom Foundation and NIH grant GM-30220.

References

Beck,E., Jorcano,J.L. and Alonso,A. (1984) *J. Mol. Biol.*, **173**, 539-542.
Bolivar,F., Rodriguez,R.L., Green,P.J., Betlach,M.C., Heynecker,H.L., Boyer,H.W., Crosa,J.H. and Falkow,S. (1977) *Gene*, **2**, 95-113.
Busch,H., Reddy,R., Rothblum,L. and Choi,Y.C. (1982) *Annu. Rev. Biochem.*, **51**, 617-654.
Elder,J.T., Pan,J., Duncan,C.H. and Weissman,S.M. (1981) *Nucleic Acids Res.*, **9**, 1171-1189.
Eliceiri,G.L. and Sayavedra,M.S. (1976) *Biochem. Biophys. Res. Commun.*, **72**, 507-512.
Hammarström,K., Westin,G., Bark,C., Zabielski,J. and Pettersson,U. (1984) *J. Mol. Biol.*, **179**, 157-169.
Harland,R.M., Weintraub,H. and McKnight,S.L. (1983) *Nature*, **302**, 38-43.
Lund,E. and Dahlberg,J.E. (1984) *J. Biol. Chem.*, **259**, 2013-2021.
Lund,E., Dahlberg,J.E. and Forbes,D. (1984) *Mol. Cell. Biol.*, in press.
Manser,T. and Gesteland,R.F. (1982) *Cell*, **29**, 257-264.
Marzluff,W.F., Brown,D.T., Lobo,S. and Wang,S.S. (1983) *Nucleic Acids Res.*, **9**, 6255-6270.
Mattaj,I. and Zeller,R. (1983) *EMBO J.*, **2**, 1883-1891.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Miller,T.J. and Mertz,J.E. (1982) *Mol. Cell. Biol.*, **2**, 1595-1607.
Murphy,J.T., Burgess,R.R., Dahlberg,J.E. and Lund,E. (1982) *Cell*, **29**, 265-274.
Nilsson,B., Uhlen,M., Josephson,S., Gatenbeck,S. and Philipson,L. (1983) *Nucleic Acids Res.*, **11**, 8019-8030.
Nishimura,S. (1972) *Progr. Nucleic Acids Res. Mol. Biol.*, **12**, 50-86.
Nohga,K., Reddy,R. and Busch,H. (1981) *Cancer Res.*, **41**, 2215-2220.
Nojima,H. and Kornberg,R.D. (1983) *J. Biol. Chem.*, **258**, 8151-8155.
Pan,J., Elder,J.T., Duncan,C.H. and Weissman,S.M. (1981) *Nucleic Acids Res.*, **9**, 1151-1170.
Reddy,R. and Busch,H. (1981) *Cell Nucleus*, **8**, 261-306.
Reddy,R., Henning,D., Epstein,P. and Busch,H. (1981) *Nucleic Acids Res.*, **9**, 5645-5658.
Roop,D.R., Kristo,P., Stumph,W.E., Tsai,M.J. and O'Malley,B. (1981) *Cell*, **23**, 671-680.
Sanger,F., Brownlee,G.G. and Barrell,B.G. (1965) *J. Mol. Biol.*, **13**, 373-398.
Silberklang,M., Gillum,A.M. and RajBhandary,U.L. (1979) *Methods Enzymol.*, **59**, 58-109.
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.
Tani,T., Watanabe-Nagasu,N., Okada,N. and Oshima,Y. (1983) *J. Mol. Biol.*, **168**, 579-594.
Van Arsdell,S. and Weiner,A.M. (1984) *Mol. Cell. Biol.*, **4**, 492-499.
Watanabe-Nagasu,N., Itoh,Y., Tani,T., Okano,K., Koga,N., Okada,N. and Oshima,Y. (1983) *Nucleic Acids Res.*, **11**, 1791-1801.
Weiher,H., Konig,M. and Gruss,P. (1983) *Science (Wash.)*, **219**, 626-631.
Westin,G., Monstein,H.-J., Zabielski,J., Philipson,L. and Pettersson,U. (1981) *Nucleic Acids Res.*, **9**, 6323-6338.
Westin,G., Zabielski,J., Hammarström,K., Monstein,H.-J., Bark,C. and Pettersson,U. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3811-3815.
Zeller,R., Carri,M.T., Mattaj,I.W. and DeRobertis,E.M. (1984) *EMBO J.*, **3**, 1075-1081.

Received on 17 October 1984