Human U2 small nuclear RNA genes contain an upstream enhancer

Marguerite Mangin, Manuel Ares, Jr and Alan M. Weiner

Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, 333 Cedar Street, PO Box 3333, New Haven, CT 06510, USA

Communicated by E. De Robertis

The human U1 and U2 snRNA genes lack an obvious TATA box, but are extremely powerful RNA polymerase II transcription units capable of accurately initiating at least one transcript per gene every 2–4 s. We have investigated the location of *cis*-acting regulatory elements within the flanking sequences of human U2 and U1 genes. By introducing marked human U2 genes into HeLa cells on SV40- and pUC13-based vectors, we found that transient expression of the marked U2 gene did not require the SV40 enhancer. The U2 promoter element responsible for SV40 enhancer-independent U2 expression was localized within the 5'-flanking sequence of the gene, and shown to stimulate transcription from the U2 basal promoter in an orientation- and position-independent fashion. In addition, the U2 element could be functionally replaced by either the SV40 enhancer or by distal sequences from the human U1 promoter. We conclude that the human U2 and U1 genes contain functionally equivalent enhancer elements. Moreover, since the human U2 enhancer sequences resemble the *Xenopus* U2 enhancer-like element, enhancers appear to be a general feature of vertebrate snRNA promoter structure. Key words: human U2 genes/small nuclear RNA/enhancer sequences

Introduction

The vertebrate small nuclear RNAs (snRNAs) U1 and U2 are thought to be transcribed by RNA polymerase II because the snRNAs themselves are capped, and their synthesis is sensitive to the drugs α -amanitin (Jensen *et al.*, 1979; Murphy *et al.*, 1982) and DRB (Tamm *et al.*, 1980). The U1 and U2 promoters lack a TATA box, as has been found previously for certain other viral (Brady *et al.*, 1982; Fire *et al.*, 1981; Cattaneo *et al.*, 1983) and cellular (Reynolds *et al.*, 1984; Lee *et al.*, 1983; Melton *et al.*, 1984; Crouse *et al.*, 1982; Giguere *et al.*, 1985) promoters for RNA polymerase II. However, two considerations suggest that the promoters for abundant structural RNAs such as U1 and U2 may differ from those for most species of mRNA.

First, the U1 and U2 promoters must be sufficiently powerful to initiate transcription as often as every 2–4 s. Human U1 RNA is encoded by only 30 genes (Lund *et al.*, 1983; Bernstein *et al.*, 1985) and U2 by only 10–20 genes (Van Arsdell and Weiner, 1984; Westin *et al.*, 1984a), but U1 and U2 snRNA are nearly 10% as abundant as rRNA on a molar basis (Weinberg and Penman, 1968). Human rRNA is encoded by 200–400 genes and is transcribed by a specialized RNA polymerase (for a review,

see Wilson, 1982). Thus, on a per gene basis, the snRNA promoters are as powerful as the rRNA promoters.

Second, the promoters for these stable structural snRNAs must be extremely accurate. The known viral and cellular mRNA promoters that lack a TATA box direct initiation at heterogeneous sites spanning as much as 200 bp; however, multiple transcription start sites need not interfere with the choice of an initiator AUG by ribosomes. In contrast, the U1 and U2 promoters direct initiation at unique sites. At least for U1, multiple start sites would presumably interfere with the ability of the 5' end of the RNA in the U1 small nuclear ribonucleoprotein particle (U1-snRNP) to recognize the 5' splice junction in mRNA precursors (Mount *et al.*, 1983; Kramer *et al.*, 1984).

The promoters for human U1 (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984) and U2 snRNA (Westin *et al.*, 1984b; Ares *et al.*, 1985), as well as *Xenopus* U1 (Ciliberto *et al.*, 1985; Krol *et al.*, 1985) and U2 (Mattaj *et al.*, 1985) have been characterized in some detail. We found by analysis of the human U2 promoter in *Xenopus* oocytes that an upstream 'activator' element was important for efficient transcription, although the 'basal promoter' within the first 62 bp of 5'-flanking sequence was able to direct accurate initiation of U2 snRNA at a low level (Ares *et al.*, 1985). The U2 activator stimulated correct initiation of U2 RNA in the natural orientation from several positions upstream of the basal promoter, but in the opposite orientation the activator stimulated only aberrant upstream starts. Mattaj *et al.* (1985) have recently shown that the *Xenopus* U2 gene also contains two promoter elements when assayed in the homologous oocyte system: a basal promoter, and an enhancer-like upstream element which activates transcription through formation of stable transcription complexes. The upstream element identified by Mattaj *et al.* (1985) is part of a significant sequence homology between the human and *Xenopus* U2 genes which we previously designated region III (Ares *et al.*, 1985).

The *Xenopus* oocyte transcription system has given a faithful picture of the structure of heterologous promoters such as the herpes virus thymidine kinase promoter (Eisenberg *et al.*, 1985), and oocytes can synthesize both human U1 (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984) and human U2 snRNA (Westin *et al.*, 1984b; Ares *et al.*, 1985) after microinjection with the heterologous genes. However, a significant fraction of the human U2 transcripts synthesized in oocytes initiate aberrantly ~180 nucleotides upstream from the correct start site (Ares *et al.*, 1985), suggesting that the human U2 promoter might not function altogether normally in the heterologous environment. We therefore investigated the structure of the human U2 promoter in human cells, using the transient expression system of Hernandez (1985). Here we show that the upstream promoter element of the human U2 gene previously defined as an 'activator' in the oocyte system (Ares *et al.*, 1985) can function as a true enhancer in HeLa cells. We also show that both the SV40 enhancer and an upstream element from the human U1 promoter can functionally replace the U2 enhancer.

Results

Transient expression of the marked U2 gene

We first constructed a marked U2 gene by inserting a 12-bp *Xba*I linker at position 28 within the U2 coding region in the U2 template dl-556 (Ares *et al.*, 1985). The marked dl-556 template contains 556 bp of 5'-flanking sequence, 200 bp of U2 coding region (rather than 188 bp as in wild-type), and 152 bp of 3'-flanking sequence. This template produces stable marked U2 snRNA after injection into *Xenopus* oocytes (data not shown) and after transfection into HeLa cells (see below). We measured synthesis of marked human U2 snRNA in HeLa cells by primer extension using a synthetic 15-mer complementary to an internal U2 sequence just downstream from the *Xba*I linker. Thus the reverse transcripts derived from the marked U2 snRNA are 12 nucleotides longer than those derived from endogenous U2 snRNA. This marked U2 gene was used for all subsequent constructions (see Figure 1, and Materials and methods).

We initially cloned the marked U2 gene into the pSVE expression vector (Hernandez, 1985) containing the SV40 early region and the complete transcriptional enhancer (positions 346–2533 on the SV40 genome). The wild-type U2 gene was also cloned into the same vector as a control. The constructs were introduced into HeLa cells by calcium phosphate precipitation, along with the rabbit β -globin gene on a separate SV40 plasmid as an internal reference gene. RNA was extracted 48 h after transfection, and S1 mapping was performed to measure globin mRNA in each preparation (see Materials and methods; data not

shown). U2 expression was then assayed by primer extension using RNA samples which had been normalized according to the amount of globin mRNA, in order to correct for variations in the efficiency of transfection (Figure 2). Endogenous U2 RNA produced the expected 42-nucleotide cDNA when the synthetic 15-mer was used to prime reverse transcription of total RNA from untransfected HeLa cells (lane 2), or from HeLa cells transfected with pSVE alone (lane 3) or pSVE containing the wild-type U2 gene (lane 4). An additional 54-nucleotide cDNA, derived from marked U2 RNA, was seen when the marked U2 gene was introduced into HeLa cells with (lane 5) or without (lane 6) the rabbit β -globin gene. Since the level of marked U2 RNA was unchanged when the rabbit β -globin reference gene was omitted, the U2 and β -globin templates do not compete under our transfection conditions.

Figure 2 also demonstrates that no major upstream initiation sites are seen when human U2 genes are transcribed in HeLa cells (Figure 2, lanes 2–8). This contrasts with the many aberrant upstream starts seen in the vicinity of position -180 when human U2 genes are assayed by injection into heterologous frog oocytes (Ares *et al.*, 1985).

In other experiments, we found that the orientation of the U2 dl-556 template relative to the pSVE vector had little effect on U2 transcription (data not shown). For consistency, however, we placed the U2 gene in the opposite orientation from the SV40 early transcription unit in all subsequent constructions using the pSVE vector.

Fig. 1. Structure of the U2 templates. Although many of the U2 templates were cloned in both the pUC13 and pSVE vectors, for consistency all the templates are shown here as pSVE clones. Restriction sites which are destroyed by the cloning protocol are enclosed in parentheses. Gaps indicate a deletion, dotted lines the site of an insertion, and arrowheads the orientation of inserted fragments. Regions of homology between the 5'-flanking sequences of various vertebrate snRNA genes are denoted by horizontal black boxes. A vertical black box indicates the position of the *Xba*I linker in the marked U2 gene.

Fig. 2. Expression of the marked dl-556 gene and the marked 6-kb U2 repeat unit in the vector pSVE. HeLa cells were transfected with 8 μ g of the U2 test plasmid ('- globin'), or 6 μ g of test plasmid and 2 μ g of the SV40 construct (pSK⁺) carrying the rabbit β -globin gene ('+ globin'). RNA was extracted 48 h later and, after normalization to the amount of globin mRNA, was annealed with the U2 primer oligonucleotide. For the mock reaction the primer was annealed with 1 μ g of yeast tRNA. Extension products were separated on an 8% sequencing gel, with kinases *Hpa*II fragments of pBR322 as markers (lane M). Arrows indicate the extension products derived from normal U2 RNA (42 nucleotides) and marked U2 RNA (54 nucleotides).

The intact 6-kb U2 repeat unit does not contain additional promoter elements

The 10–20 human U2 genes are organized in a tandem array with a 6-kb repeat unit (Van Arsdel and Weiner, 1984; Westin *et al.*, 1984a) on chromosome 17 (Lindgren *et al.*, 1985). To test whether the marked U2 dl-556 gene contained all the information necessary for maximal U2 expression in human cells, we replaced the wild-type U2 gene with the marked U2 gene in a 6-kb *Hind*III fragment corresponding to one complete repeat unit. (The U2 coding region is flanked by ~5 kb of upstream and 1 kb of downstream DNA in this fragment.) We then recloned the marked 6-kb repeat unit into the pSVE vector, and transfected the construction into HeLa cells. As shown in Figure 2, the presence of 6 kb of natural flanking sequence did not affect the synthesis of marked U2 RNA (compare lanes 5 and 7).

The U2 activator stimulates the U2 basal promoter in the presence of a distant SV40 enhancer

Using the frog oocyte assay system, we previously demonstrated that the *Sma*I fragment between positions -198 and -270 contains a transcriptional activator (Ares *et al.*, 1985). The U2 basal promoter, located downstream of position -62, was able to initiate U2 transcription accurately, but was stimulated by the upstream activator in an orientation-dependent and partially position-independent fashion. To test whether the activator element would behave similarly in a homologous transcription system, we introduced mutations into the promoter of the marked human U2 gene in the pSVE vector and assayed for U2 transcription after transfection into HeLa cells. As shown in Figure 3A, the synthesis of U2 RNA was reduced by a factor of 8–10 by deletion of all but 62 bp of 5'-flanking sequence (compare lane 4 with lane 8), or by deletion of only 4 bp between -223 and -218 within the *Sma*I fragment (lane 9). We tested a large number of U2 promoter constructions (see Figure 1) carried on the pSVE vector before finding that the SV40 enhancer and early region were not required for transient expression of U2 genes (see below). In all cases pSVE-based templates lacking the 72-bp *Sma*I fragment between -198 and -270 were expressed 10% as efficiently as wild-type. Introduction of the fragment downstream from the gene in either orientation or upstream in the opposite orientation stimulated transcription from the basal promoter to 20–30% of wild-type levels. Wild-type levels of transcription were achieved only when the *Sma*I fragment was present upstream of the basal promoter in the natural orientation (data not shown). Deletion of a sequence called region I, which is conserved between the human U1 and U2 promoters (Van Arsdel and Weiner, 1984), had no detectable effect on U2 transcription (Figure 1 and data not shown).

The U2 activator is required for U2 expression in the absence of SV40 sequences

To test whether the human U2 gene could function independently of SV40 sequences, we introduced the marked U2 gene and several promoter mutations into the plasmid vector pUC13. After transfection into HeLa cells, the complete U2 gene (dl-556) was expressed as well on the pUC13 vector (Figure 3A, lane 4) as on the pSVE vector containing the entire SV40 early region and enhancer (lane 5). Moreover, although the U2 basal promoter (dl-62) exhibited a low level of activity on the pSVE vector (lane 8), the basal promoter was completely inactive on the pUC13 vector (Figure 3A, lane 6 and Figure 3B, lane 1). Similarly, the 4-bp deletion (dl-223/-218) which abolishes activator function was transcribed at a low level on the pSVE vector (Figure 3A, lane 9) but was inactive on pUC13 (lane 7). These

Fig. 3. Comparative expression of mutant U2 genes in the SV40 vector pSVE and plasmid vector pUC13. (A) Promoter mutants of the marked dl-556 U2 gene in either pUC13 or pSVE were co-transfected into HeLa cells together with the pSK⁺ reference plasmid, and normalized aliquots of RNA were analyzed by primer extension. As a control (lane 10), primer extension was performed on a mixed RNA sample containing 10% of the amount of pSV/dl-62 RNA analyzed in lane 8 mixed with 90% of the HeLa RNA analyzed in lane 2. (B) Promoter mutants in the marked U2 gene were cloned into pUC13 and transfected into HeLa cells with the globin reference plasmid. Expression of marked U2 RNA was assayed by primer extension as before. Extension products >54 bases are not reproducibly obtained and can be observed using RNA from untransfected HeLa cells (data not shown).

results suggest that the distant SV40 enhancer accounts for the residual transcription from the U2 basal promoter observed in pSVE constructions.

When one copy of the *Sma*I fragment was introduced in the natural orientation just upstream of the U2 basal promoter in mutant dl-62, wild-type levels of U2 synthesis were restored (Figure 3B, lane 2). Introduction of the *Sma*I fragment at the same site in the opposite orientation also increased U2 transcription from the basal promoter, but 2- to 3-fold less than the natural orientation (Figure 3B, lane 3). Introduction of the *Sma*I fragment at position +314, 502 bp downstream from the U2 cap site, also stimulated transcription from the U2 basal promoter in both the natural orientation (lane 5) and the opposite orientation (lane 6). Thus, in the absence of SV40 sequences, the U2 activator contained on the *Sma*I fragment is required for U2 expression from the basal promoter.

Human U1 genes also require upstream sequences for SV40-independent expression

An important element of the human U1 promoter has been identified between positions -230 to -203 upstream from the U1 990

coding region (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984). To determine whether this element of the U1 promoter might be functionally similar to the U2 activator, we used a marked U1 gene (Eco⁺) cloned into a pBR322 derivative called pBRN/S (Haltiner *et al.*, 1985) and containing an *Eco*RI linker in the *Bcl*I site at position 31 of the U1 coding region (Hernandez, 1985). The marked U1 gene was transfected into HeLa cells, and synthesis of marked U1 RNA was assayed by primer extension using a synthetic 20-mer complementary to an internal U1 sequence (See Materials and methods). The marked U1 gene was as efficiently expressed on the pBRN/S vector (Figure 4, lane 1) as on pSVE (lane 5), indicating that the complete U1 gene, like the complete U2 gene, does not require SV40 sequences for efficient transcription. Deletion of sequences upstream from -160 abolished U1 expression (lane 3). Sequences between -106 and -80 contributed to promoter efficiency, but were not absolutely required for transient expression in HeLa cells (lane 2). However, neither this mutant nor a similar mutant with a deletion between positions -92 and -70 produced correctly initiated U1 RNA after injection into frog oocytes (M.Mangin, unpublished results). Sequences between -160 and -6 were also essen-

Fig. 4. Expression of the marked U1 gene in the SV40 vector pSVE or the plasmid pBRN/S. The pSK⁺ reference gene was included in all transfections except that shown in **lane 5**, and RNA from the transfected cells was analyzed by primer extension using the U1 primer oligonucleotide. The RNA analyzed in **lane 5** was kindly donated by N.Hernandez as a control for expression of the marked Eco⁺ U1 gene. Arrows indicate the extension products derived from marked U1 RNA (61 nucleotides) and normal U1 RNA (47 nucleotides).

tial for U1 transcription (lane 4), but we did not detect the aberrant upstream starts observed when similar mutants were injected into frog oocytes or transcribed *in vitro* (Murphy *et al.*, 1982; Skuzeski *et al.*, 1984).

The U2 basal promoter can be activated by the SV40 enhancer or the U1 upstream promoter element

The U2 basal promoter (dl-62) was completely inactive in

Fig. 5. Expression of the U2 basal promoter linked to heterologous enhancers in the vector pUC13.

pUC13 (Figure 3) but exhibited a lower level of activity in pSVE constructions having a distant viral enhancer sequence (Figure 3A). These results suggested that the SV40 enhancer might be able to substitute functionally for the U2 activator. To test this possibility, we introduced the SV40 *NcoI* fragment (containing the 21- and 72-bp repeats) in each orientation at positions -62 upstream from the basal U2 promoter in pUC13. Although the U2 basal promoter by itself was inactive in pUC13 (Figure 5, lane 9), the SV40 enhancer region was able to restore high levels of U2 transcription in either orientation upstream from the gene (lanes 3 and 4). Restoration of U2 synthesis by the SV40 *NcoI* fragment was completely independent of orientation (compare lanes 3 and 4) whereas the U2 *SmaI* fragment worked better in the natural orientation (compare lanes 6 and 7). U2 transcription from the basal promoter could also be restored by fusing U1 promoter sequences upstream from -160 to position -198 in the U2 gene (lane 5).

We also found that the low level of U2 synthesis observed from the basal promoter in pSVE constructions having a distant SV40

Table I. Conserved sequence elements upstream from vertebrate snRNA genes

[illegible]

snRNA enhancer consensus   G G G C G ^G_A G N N₀₋₂ ^T_C A T G C A A A T

Conserved sequence elements upstream from vertebrate snRNA genes. The sequences were taken from the following sources: human U1 (Murphy *et al.*, 1982); human U2 (Ares *et al.*, 1985); chicken U1 (Early *et al.*, 1984); chicken U2 (Korf and Stumph, 1986); rat U1 (Watanabe-Nagasu *et al.*, 1983); rat U2 (Tani *et al.*, 1983); rat U3 (Stroke and Weiner, 1985); *Xenopus* U1 (Krol *et al.*, 1985; Ciliberto *et al.*, 1985); *Xenopus* U2 (Mattaj and Zeller, 1983). Region III is overlined in the human U2 sequence, and the synthetic 14-mer of Mattaj *et al.* (1985) is underlined in the *Xenopus* U2 sequence. Bold letters indicate the 4-bp deletion at the *Sph*I site in the human U2 mutant dl-223/-218.

enhancer could be restored to normal levels by inserting either the SV40 *NcoI* fragment or a U1 *AvaII* fragment (positions -389 to -133, spanning the U1 upstream element) in place of the U2 *SmaI* fragment deleted in pSV/dl-270/-198 (data not shown). This suggests that the SV40 enhancer works better immediately upstream from the U2 basal promoter at position -62 than when located 3 kb away, as expected if the SV40 enhancer is not completely position-independent (Moreau *et al.*, 1981; Mosthaf *et al.*, 1985). Alternatively, the SV40 early promoter may preempt the SV40 enhancer in certain pSVE constructions (Wasylyk *et al.*, 1983).

Discussion

We previously found that sequences located upstream from the 5' end of the human U2 gene were both necessary and sufficient for U2 synthesis in the heterologous frog oocyte system (Ares *et al.*, 1985). The U2 basal promoter, located downstream of position -62, was able to direct accurate initiation of U2 transcription, albeit at only 2-5% of the wild-type level. An upstream activator located between -270 and -198 on a 72-bp *Sma*I fragment was required for maximal transcription, and behaved in some respects like an enhancer. In the natural orientation the activator stimulated the U2 basal promoter from different positions upstream, but in the opposite orientation the activator stimulated only aberrant upstream starts.

To determine whether the human U2 promoter elements defined in the heterologous frog oocyte behaved in a similar fashion in a homologous assay system, we introduced human U2 templates into HeLa cells using SV40 and pUC13 vectors. All experiments were performed with U2 genes marked by a linker

insertion within the coding region to distinguish the products of the exogenous gene from endogenous U2 snRNA. The U2 promoter mutants were constructed using the U2 template dl-556 (Ares *et al.*, 1985), because transfection of a complete 6-kb repeat unit from the U2 tandem array (Van Arsdel and Weiner, 1984; Westin *et al.*, 1984a; Lindgren *et al.*, 1985) did not reveal any additional promoter elements upstream of position -556 or downstream of +152 (Figure 2).

The results presented in this paper demonstrate that the same regulatory elements characterized in the *Xenopus* system are also required for U2 transcription in human cells. However, there are two major differences between the behavior of the promoter elements in the heterologous oocyte and the homologous HeLa system. First, although the basal U2 promoter alone directed a low level of U2 transcription in oocytes, it was completely inactive in HeLa cells unless linked to the U2 activator, the SV40 enhancer or distal promoter element from a human U1 gene. Second, although the human U2 activator was strictly orientation-dependent in oocytes, it behaved like a true transcriptional enhancer in human cells.

Transient expression of U2 RNA was equally efficient when the complete human U2 gene (dl-556) was introduced into HeLa cells on either a pUC13 plasmid vector or a pSVE vector containing the SV40 viral enhancer and an intact early region (Figure 3). Thus SV40 sequences are not necessary for transient expression of U2 genes. The U2 basal promoter (dl-62) was completely inactive in pUC13 constructions lacking the U2 activator but exhibited a low level of activity in pSVE constructions having a distant SV40 enhancer (Figure 3). This suggested that the SV40 enhancer might be able to replace functionally the U2 activator. Both the *Sma*I fragment containing the U2 activator, and

the *NcoI* fragment containing the SV40 enhancer, restored high levels of U2 synthesis when inserted in either orientation at position -62 upstream from the U2 basal promoter (Figure 5). Although the SV40 enhancer was completely orientation-independent at this upstream position, the U2 activator worked 2- to 3-fold better in the natural orientation.

The U2 activator stimulated the basal promoter in a orientation-independent fashion at four different positions: 552 bp upstream from the gene (mutants Δ ins-552/Sma $\rightarrow$ and $\rightarrow$; data not shown), at its natural position 198 bp upstream from the gene (dl-556 and inv-198/-270; Figure 3A), 63 bp upstream from the gene (ins+62/Sma $\rightarrow$ and $\rightarrow$; Figures 3B and 5), and 314 bp downstream from the U2 gene (ins+314/Sma $\rightarrow$ and $\rightarrow$; Figure 3B; Δ ins+314/Sma $\rightarrow$ and $\rightarrow$, data not shown). Although the U2 enhancer works best when upstream of the basal promoter in the natural orientation, it does work both upstream and downstream from the basal U2 promoter in either orientation. Such position effects (Moreau *et al.*, 1981; Mosthaf *et al.*, 1985; Serfling *et al.*, 1985) and orientation effects (Hen *et al.*, 1983; Rosenthal *et al.*, 1983; Kenney *et al.*, 1984; Laimins *et al.*, 1984) have been reported previously for other enhancers. We conclude that the U2 activator is a true enhancer.

A number of early studies indicated that enhancer-dependent transcription during a transient expression assay does not depend on plasmid replication (reviewed in Serfling *et al.*, 1985). In a systematic re-investigation of this phenomenon, Treisman *et al.* (1983) confirmed the earlier results and also made the surprising observation that template replication does not augment enhancer-dependent transcription during a transient expression assay. By comparing U2 expression on replicating (pSVE) and presumably non-replicating vectors (pUC13, Figure 3; and origin-defective and enhancerless pSVE vector mutants, data not shown), we found that U2 expression is independent of template replication and is not affected by T-antigen supplied in *trans* by the SV40 construct containing the rabbit β -globin reference gene (data not shown). Thus, in this respect as well, the U2 enhancer behaves like other enhancers.

Skuzeski *et al.* (1984) demonstrated that transcription of the human U1 gene in *Xenopus* oocytes required sequences located between positions -231 and -203, and a TATA-like element between positions -105 and -6 which specifies the initiation site but cannot direct U1 transcription by itself. We found that transient expression of the human U1 gene, like that of the human U2 gene, does not require the SV40 enhancer (Figure 4). Moreover, sequences between -420 and -160 (containing an essential U1 promoter element as defined in the oocyte system) can functionally replace the U2 enhancer (Figure 5). These observations suggest that both the human U1 and U2 promoters may be functionally divisible into distal enhancer and proximal promoter elements. Others have reached similar conclusions for the *Xenopus* U1 (Ciliberto *et al.*, 1985; Krol *et al.*, 1985) and U2 (Mattaj *et al.*, 1985) genes, and for the human U2 gene as assayed in oocytes (Ares *et al.*, 1985). However, it remains to be seen why the human U1 proximal promoter element by itself fails to direct transcription in oocytes, and whether the human U1 activator can in fact function as a true enhancer.

The U2 enhancer appears to be entirely contained on a 72-bp *SmaI* fragment spanning positions -270 to -198. This *SmaI* fragment contains a block of homology called region III (Ares *et al.*, 1985; overlined in Table I) that is highly conserved between the frog, rat and human U2 promoters. Region III is not restricted to U2 genes, and is found 230-280 bp upstream from the RNA coding region in other sequenced vertebrate U-snRNA

genes (Table I). The region III consensus can be divided into two modules which have been noted separately or in part by others in previous sequence comparisons (Early *et al.*, 1984; Ciliberto *et al.*, 1985; Krol *et al.*, 1985; Mattaj *et al.*, 1985; Korf and Stumph, 1986). The upstream module is fairly well conserved and strongly resembles the GGGCGG binding site for the Sp1 transcription factor (Dyner and Tijian, 1985). The downstream module is highly conserved and is identical to the octamer 5'-ATGCAAAT-3' found in all human and mouse immunoglobulin heavy and light chain gene promoters and in the heavy chain enhancer (Parslow *et al.*, 1984; Falkner and Zachau, 1984). This octamer has been shown to be important for binding a factor present in crude extracts of HeLa cell nuclei (Singh *et al.*, 1985). We have shown that deletion of just 4 bp from the *SphI* site in the downstream module of the human U2 gene (dl-223/-218) abolishes U2 enhancer function both in oocytes (Ares *et al.*, 1985) and in human cells (Figure 3). Moreover, Mattaj *et al.* (1985) have recently shown that a synthetic 14-bp oligonucleotide spanning only the downstream module of the *Xenopus* U2 region III (underlined in Table I) can partially restore enhancer function in a competition assay using homologous frog oocytes. These experiments suggest that the U2 enhancer may overlap or coincide with region III.

We do not yet know what role the U2 enhancer might play in regulating U2 expression. One possibility is that enhancers might regulate tissue-specific expression of different snRNAs or snRNA variants. In fact, an embryo-specific U1 variant has been observed in the mouse (Lund *et al.*, 1985). However, the homogeneity of the human U1 (Manser and Gesteland, 1982; Bernstein *et al.*, 1985) and U2 genes (Van Arsdel and Weiner, 1984; Westin *et al.*, 1984a) suggests that mammalian snRNAs are not generally expressed in a tissue-specific manner. Another possibility is that snRNA enhancers might play a role in coordinating the biosynthesis of the snRNA and protein components of the U-snRNPs. The various species of U-snRNPs appear to differ in abundance by as much as two orders of magnitude, and some proteins are shared by all U-snRNPs while others are unique to a particular snRNP species. We have shown previously that U1 gene expression is subject to dosage-compensation in mouse cells (Mangin *et al.*, 1985), but we do not know whether regulation occurs at the level of snRNA transcription, snRNP assembly or snRNA degradation. U-snRNA enhancers represent a potential control point in regulating the relative abundance of both the snRNA and protein components of the different U-snRNPs.

Materials and methods

Construction of the marked U2 gene

The dl-556 plasmid contains 556 bp of 5'-flanking sequence, the 188-bp U2 coding region and 152 bp of 3'-flanking sequence in plasmid pML2d (Ares *et al.*, 1985). Using the following procedure, we inserted a 12-bp *XbaI* linker in the unique *Sau3A* site at position 28 in the U2 coding region of dl-556. The 580-bp *EcoRI*-*Sau3A* fragment of dl-556 containing the 5'-flanking sequence and the first 28 bp of the gene was cloned between the *EcoRI* and *BamHI* sites of the plasmid vector pUC13, bringing the *Sau3A* site next to the *XbaI* site in the polylinker (5' half gene). The 494-bp *Sau3A* fragment of dl-556 containing the 3' end of the gene, the 3'-flanking sequence and some pML2d sequences was cloned into the *BamHI* site of pUC13 in the opposite orientation (3' half gene). The 3' half gene was then excised using the *XbaI* and *Sall* sites within the polylinker, and recloned downstream from the 5' half gene between the *XbaI* and *Sall* sites in the polylinker. The presence of the *XbaI* linker was confirmed by DNA sequence analysis (Sanger *et al.*, 1977).

Construction of mutant U2 templates

The marked U2 gene was introduced into the complete 6-kb U2 repeat unit in four steps. First, a 4.6-kb *HindIII*-*SphI* fragment containing the U2 5'-flanking

sequences (positions -4600 to -220) was cloned between the *HindIII* and *SphI* sites of the vector pBRN/S (Haltiner *et al.*, 1985). Second, the marked U2 gene was introduced between the *SphI* and *SalI* site in the vector as an *SphI*-*SalI* fragment (positions -219 to +152). The 3'-flanking sequences downstream of +152 were introduced into the *SalI* site of this construction as two fragments (Van Arsdel and Weiner, 1984): a *BamHI*-*PstI* fragment excised from a clone of the 750-bp *AccI* fragment in the *SmaI* site of the vector M13mp8, and a *PstI*-*SalI* fragment excised from a clone of the 250-bp *AccI*-*SalI* fragment in the *SalI* site of the vector pUC13. The resulting marked 6-kb repeat contained small insertions of polylinker at both *AccI* sites of the wild-type repeat.

The wild-type and marked dl-556 U2 templates, as well as the marked 6-kb repeat and derivatives of the marked dl-556 template, were cloned into the *EcoRI* site, the *SalI* site or between the *EcoRI* and *Clal* sites of the SV40 vector pSVE (Hernandez, 1985). For all constructions shown in Figure 1 except dl-223/-218, we first joined complementary ends between the insert and vector by ligation for 2 h, then filled in non-complementary ends by incubation with the Klenow fragment of DNA polymerase I (Joyce and Grindley, 1983) in the presence of 50 mM each of the deoxyribonucleotide triphosphates, and finally joined the resulting blunt ends by an overnight ligation. Mutant dl-223/-218 was generated by digestion with *SphI* followed by treatment with S1 nuclease and re-ligation, and the extent of deletion confirmed by DNA sequence analysis (Sanger *et al.*, 1977).

The U1 templates were derived from the 681-bp *HaeIII* fragment of the HSD4 gene (Manser and Gesteland, 1979). This fragment was fitted with *BamHI* linkers, cloned into pBRN/S (Haltiner *et al.*, 1985), and marked by insertion of an *EcoRI* linker in the original *BclI* site at position 27 within the U1 coding region (Hernandez, 1985). The template encodes a marked U1 RNA 14 nucleotides longer than wild-type U1. Sequences between -160 and -6 were deleted by complete digestion with *BglII* and partial digestion with *XmnI*. Sequences between -106 and -80 were deleted by digestion with *PvuII* and *AvaI*. Sequences upstream of -160 were deleted by subcloning the *XmnI*-*BamHI* fragment spanning the U1 gene. Finally, mutant ins-198/*XmnI* was constructed as follows. An *EcoRI*-*XmnI* fragment spanning U1 promoter sequences between -420 and -160 was cloned between the *EcoRI* site in the vector and the *SmaI* site at position -198 upstream from the U2 gene.

Transfections

Transfections were performed as described by Hernandez (1985), using a total of 8 µg DNA per 9 cm plate. Between 20 and 40% of the cells were routinely transfected, as judged by indirect immunofluorescence with a monoclonal antibody directed against the SV40 T-antigen (Harlow *et al.*, 1981). To measure the efficiency of transfection accurately, we co-transfected the U2 or U1 constructs with a separate SV40 construction (pSK⁺) containing the rabbit β-globin gene (Banerji *et al.*, 1981). To minimize the possible effect of competition between RNA polymerase II transcription units as described by Schoeler and Gruss (1984), the ratio of the test plasmid to reference plasmid was always 3:1. Control experiments indicated that transcription of the U2 dl-556 template was unaffected by even a 50-fold molar excess of the β-globin reference template in the transfecting DNA (data not shown). RNA was extracted from the post-nuclear supernatant of cells lysed with NP40 (Hernandez, 1985). Globin mRNA was measured by S1 mapping as described below.

S1 nuclease mapping of globin mRNA

A 550-bp *EcoRI*-*PvuII* fragment extending from position 418 in the third exon of the rabbit β-globin gene to position +380 in the 3'-flanking sequence (Banerji *et al.*, 1981) was labeled at the 3' end with the Klenow fragment of DNA polymerase I (Joyce and Grindley, 1983) and [α -³²P]dATP (sp. act. 3000 Ci/mmol). About 10⁵ c.p.m. of double-stranded probe or 3 × 10⁴ c.p.m. of single-stranded probe were annealed with 20 ng of RNA from transfected cells (~2-3% of the RNA from a 9 cm dish) in 15 µl of hybridization buffer (80% formamide, 400 mM NaCl, 40 mM Pipes pH 6.4, 1 mM EDTA) and incubated overnight at 50°C. Samples were then diluted with 200 µl of S1 buffer (30 mM Na acetate pH 4.5, 3 mM ZnSO₄, 300 mM NaCl, 10 µg/ml denatured salmon sperm DNA) containing 150 units of S1 nuclease and digested for 90 min at 25°C. S1 digests were precipitated with ethanol, redissolved, and resolved by electrophoresis on a 6% sequencing gel. The 171 nucleotide protected fragment corresponding to the expected polyadenylation site was excised from the gel and quantified by Cerenkov radiation.

Primer extension analysis of U2 and U1 snRNA

The primers were synthetic oligonucleotides complementary to the first loop of the proposed U1 and U2 RNA secondary structures. The U2 primer was the L15 15-mer described by Black *et al.* (1985) and used previously by Ares *et al.* (1985). The U1 primer was a 20-mer with sequence 5'-GGAAAACACCTTCGTG-ATC-3'. For both the marked U1 and U2 genes, the linker insertion lies upstream from the primer site. Thus the reverse transcript derived from the marked U2 snRNA is 12 nucleotides longer than the normal 42-nucleotide extension product, and the reverse transcript derived from the marked *Eco*⁺ U1 gene is 14

nucleotides longer than the normal 47-nucleotide extension product.

Primers were labeled at the 5' end by kinasin. A normalized amount of RNA from transfected cells (usually corresponding to 1-5% of a 9 cm dish) was mixed with an excess of labeled primer (usually 10⁶ c.p.m.) in 10 µl of 0.1 mM EDTA, heated to 65°C for 5 min and quenched on ice. After addition of 2.5 µl of 5 × reverse transcription buffer (1 × buffer: 100 mM Tris-HCl pH 8.3, 10 mM MgCl₂, 14 mM KCl), the primer was annealed to the template snRNA for 15 min at 65°C and then cooled to room temperature. Samples were made 2 mM in each deoxyribonucleotide triphosphate, 1 mM in dithiothreitol (DTT), and 5 units of avian myeloblastosis virus reverse transcriptase (Life Sciences or Promega Biotec) were added.

Reactions were incubated sequentially at 30°C for 10 min, 37°C for 10 min and finally 42°C for 1 h. Reactions were stopped with 10 mM EDTA, brought to 0.3 M sodium acetate and precipitated with ethanol. Samples were re-dissolved in 10 mM Tris-HCl, 1 mM EDTA, pH 7.5 containing 1 ng/ml of DNase-free pancreatic RNase. The labeled extension products were resolved on an 8% sequencing gel and relative amounts of radioactivity in the different bands were measured by densitometric scanning of the autoradiograph.

Acknowledgements

We thank Doug Black for the U2 oligonucleotide, Nouria Hernandez for the marked U1 gene and Vernita Hoffarth Ares for sequencing the U2 clones. We are grateful to Nouria Hernandez and Elisabetta Ullu for helpful advice and criticism and to Bill Stumph for communicating unpublished results. This work was supported by NIH grants GM 31335 and GM 31073 to A.M.Weiner and an NIH postdoctoral fellowship (GM 01948) to M.Ares, Jr.

References

- Ares, M., Mangin, M. and Weiner, A.M. (1985) *Mol. Cell. Biol.*, **5**, 1560-1570.
- Banerji, J., Rusconi, S. and Schaffner, W. (1981) *Cell*, **27**, 299-308.
- Bernstein, L.B., Manser, T. and Weiner, A.M. (1985) *Mol. Cell. Biol.*, **5**, 2159-2171.
- Black, D.L., Chabot, B. and Steitz, J.A. (1985) *Cell*, **42**, 737-750.
- Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J. and Salzman, N.P. (1982) *Cell*, **31**, 625-633.
- Cattaneo, R., Will, H., Hernandez, N. and Schaller, H. (1983) *Nature*, **305**, 336-338.
- Ciliberto, G., Buckland, R., Cortese, R. and Philipson, L. (1985) *EMBO J.*, **4**, 1537-1543.
- Crouse, G.F., Simonsen, C.S., McEwan, R.N. and Schimke, R.T. (1982) *J. Biol. Chem.*, **257**, 7887-7897.
- Dynan, W.S. and Tjian, R. (1985) *Nature*, **316**, 774-778.
- Early, J.M., III, Roebuck, K.A. and Stumph, W.E. (1984) *Nucleic Acids Res.*, **12**, 7411-7421.
- Eisenberg, S.P., Coen, D.M. and McKnight, S.L. (1985) *Mol. Cell. Biol.*, **5**, 1940-1947.
- Falkner, F.G. and Zachau, H.G. (1984) *Nature*, **310**, 71-74.
- Fire, A., Baker, C.C., Manley, J.L., Ziff, E.B. and Sharp, P.A. (1981) *J. Virol.*, **40**, 703-719.
- Giguere, V., Isobe, K. and Grosfeld, F. (1985) *EMBO J.*, **4**, 2017-2024.
- Haltiner, M., Kempe, T. and Tjian, R. (1985) *Nucleic Acids Res.*, **13**, 1015-1025.
- Harlow, E., Crawford, L.V., Pim, D.C. and Williamson, N.M. (1981) *J. Virol.*, **39**, 861-869.
- Hen, R., Borrelli, E., Sassone-Corsi, P. and Chambon, P. (1983) *Nucleic Acids Res.*, **11**, 8747-8760.
- Hernandez, H. (1985) *EMBO J.*, **4**, 1827-1837.
- Jensen, E.G., Hellung-Larsen, P. and Frederiksen, S. (1979) *Nucleic Acids Res.*, **6**, 321-330.
- Joyce, C. and Grindley, N. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1830-1834.
- Kenney, S., Natarajan, V., Strike, D., Khoury, G. and Salzman, N.P. (1984) *Science*, **226**, 1337-1339.
- Kramer, A., Keller, W., Appel, B. and Luhrmann, R. (1984) *Cell*, **38**, 299-307.
- Krol, A., Lund, E. and Dahlberg, J.E. (1985) *EMBO J.*, **4**, 1529-1535.
- Laimins, L.A., Gruss, P., Pozzatti, R. and Khoury, G. (1984) *J. Virol.*, **49**, 183-189.
- Lee, M.G., Lewis, S.A., Wilde, C.D. and Cowan, N.J. (1983) *Cell*, **33**, 477-487.
- Lindgren, V., Ares, M., Weiner, A.M. and Francke, U. (1985) *Nature*, **314**, 115-116.
- Lund, E., Bostock, C., Robertson, M., Christie, S., Mitcheu, J. and Dahlberg, J. (1983) *Mol. Cell. Biol.*, **3**, 2211-2220.
- Lund, E., Kahan, B. and Dahlberg, J.E. (1985) *Science*, **229**, 1271-1274.
- Mangin, M., Ares, M. and Weiner, A.M. (1985) *Science*, **229**, 272-275.
- Manser, T. and Gesteland, R. (1982) *Cell*, **29**, 257-264.
- Mattaj, I.W. and Zeller, R. (1983) *EMBO J.*, **2**, 1883-1891.
- Mattaj, I.W., Lienhard, S., Jiricny, J. and De Robertis, E.M. (1985) *Nature*, **316**, 163-167.

- Melton, D.W., Konecki, D.S., Brennand, J. and Caskey, C.T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2147–2151.
- Moreau, P., Hen, R., Wasylyk, B., Gaub, M.P. and Chambon, P. (1981) *Nucleic Acids Res.*, **9**, 6047–6068.
- Mosthaf, L., Pawlita, M. and Gruss, P. (1985) *Nature*, **315**, 597–600.
- Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983) *Cell*, **33**, 509–518.
- Murphy, J.T., Burgess, R.R., Dahlberg, J.E. and Lund, E. (1982) *Cell*, **29**, 265–274.
- Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2650–2654.
- Reynolds, G.A., Basu, S.K., Osborne, T.F., Chin, D.J., Gil, G., Brown, M.S., Goldstein, J.L. and Luskey, K.L. (1984) *Cell*, **38**, 275–285.
- Rosenthal, N., Kress, M., Gruss, P. and Khoury, G. (1983) *Science*, **222**, 749–755.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Schoeler, H.R. and Gruss, P. (1984) *Cell*, **36**, 403–411.
- Serfling, E., Jasin, M. and Schaffner, W. (1985) *Trends Genet.*, **1**, 224–230.
- Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1985) *Nature*, **319**, 154–158.
- 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.
- Stroke, I.L. and Weiner, A.M. (1985) *J. Mol. Biol.*, **184**, 183–193.
- Tamm, I., Kikuchi, T., Darnell, J.E., Salditt-Georgiev, M. (1980) *Biochemistry*, **19**, 2743–2748.
- Tani, T., Watanabe-Nagasu, N., Okada, N. and Oshima, Y. (1983) *J. Mol. Biol.*, **168**, 579–594.
- Treisman, R., Green, M.R. and Maniatis, T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7428–7432.
- Van Arsdell, S.W. and Weiner, A.M. (1984) *Mol. Cell. Biol.*, **4**, 492–499.
- Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) *Cell*, **32**, 503–514.
- Watanabe-Nagasu, N., Itoh, Y., Tani, T., Okano, K., Koga, N., Okada, N. and Oshima, Y. (1983) *Nucleic Acids Res.*, **11**, 1791–1801.
- Weinberg, R. and Penman, S. (1968) *J. Mol. Biol.*, **38**, 289–304.
- Westin, G., Zabielski, J., Hammarstrom, K., Monstein, H.-J., Bark, C. and Pettersson, U. (1984a) *Proc. Natl. Acad. Sci. USA*, **81**, 3811–3815.
- Westin, G., Lund, E., Murphy, J.T., Pettersson, U. and Dahlberg, J.E. (1984b) *EMBO J.*, **3**, 3295–3301.
- Wilson, G.N. (1982) In Busch, H. and Rothblum, L. (eds), *The Cell Nucleus XA*. Academic Press, NY, pp. 287–318.

Received on 19 December 1985; revised on 17 February 1986