Transcription analysis of a human U4C gene: involvement of transcription factors novel to snRNA gene expression

Polly Weller, Christina Bark, Lars Janson, and Ulf Pettersson¹

Department of Medical Genetics, Biomedical Center, S-751 23 Uppsala, Sweden

We have investigated the promoter requirements for *in vivo* transcription of a human U4C snRNA gene following transfection into HeLa cells. Two elements required for maximal U4C transcription were identified. The first, located upstream of -50 , provides a basal level of transcription 2–3% of the full activity, and probably corresponds to the previously identified snRNA gene proximal element. The distal element, centered around -220 , acts as a transcriptional enhancer and contains motifs for three previously recognized transcription factors: the octamer-binding protein, NF-A, which binds to motifs in the distal elements of other snRNA genes, and two factors not previously shown to be involved in snRNA gene transcription, cAMP response element binding protein (CREB) and AP-2. The octamer and putative AP-2 motifs are required for maximal transcription of the U4C gene. Specific binding of NF-A and CREB to the motifs in the distal element has been shown *in vitro* by DNase I and DMS methylation protection footprint competition analyses using HeLa nuclear extracts. The presence of a binding motif for the inducible factor CREB, together with the transcriptional requirement for the putative AP-2 motif, suggests a means by which expression of snRNA genes might be regulated.

[**Key Words:** U4 snRNA; snRNA transcription; enhancer-binding proteins; transcription factors]

Received June 6, 1988; revised version accepted September 1, 1988.

Small RNAs (snRNAs) occur at 10^3 – 10^6 copies per cell nucleus as components of ribonucleoprotein particles (snRNPs). It has been established recently that snRNPs U1, U2, U5, and U4/U6 function in the processing of mRNA precursors (for review, see Maniatis and Reed 1987; Sharp 1987). RNA polymerase II is thought to participate in the transcription of snRNA genes U1–U5. This is based on the observations that snRNA synthesis is sensitive to low levels of α -amanitin (Frederiksen et al. 1978) and that the transcripts are capped initially by an m⁷G-cap characteristic of polymerase II transcription, which is only later methylated further to the m₃^{2,2,7}-cap of mature snRNAs (Eliceiri 1980; Skuzeski et al. 1984; Mattaj 1986). In contrast, U6 has been shown to be transcribed by RNA polymerase III (Kunkel et al. 1986; Reddy et al. 1987; Kunkel and Pederson 1988).

A number of different observations, however, combine to suggest that snRNA gene transcription by polymerase II is distinct from the transcription of genes encoding proteins. Transcription is extremely efficient; the transcription rate per gene for U1 and U2 genes has been estimated to be as high as that for rRNA genes (Mangin et al. 1986). Furthermore, the structure of promoter and enhancer elements and also the 3'-processing signals differ between snRNA and protein-coding genes. With

the exception of U6, snRNA gene promoters do not contain TATA boxes, although transcription has been shown to depend on two 5'-flanking sequence elements that contain conserved sequence motifs (Skuzeski et al. 1984; Westin et al. 1984; Ares et al. 1985; Krol et al. 1985; Mattaj et al. 1985; Mangin et al. 1986; Ciliberto et al. 1987; McNamara et al. 1987). The proximal element, located around positions -50 to -60 , is absolutely required for transcription and has a function analogous to that of the TATA box in fixing the site of transcription initiation (Skuzeski et al. 1984; Ares et al. 1985; Ciliberto et al. 1985; Mattaj et al. 1985; Murphy et al. 1987). The second element, the distal element, occurring between positions -200 and -260 , is required for maximal levels of transcription and also has enhancer-like properties (Skuzeski et al. 1984; Ares et al. 1985; Mattaj et al. 1985; Mangin et al. 1986; Murphy et al. 1987).

Given these unusual features of snRNA gene transcription, it might be expected that snRNA gene-specific factors are involved in transcription. However, only two previously identified factors have been implicated to date. The distal element of vertebrate snRNA genes contains (in either orientation) one copy of the octameric motif, or 'octamer,' also found in the 5'-flanking sequences of several protein-coding genes (Harvey et al. 1982; Falkner and Zachau 1984; Parslow et al. 1984; Singh et al. 1986). This motif in the human and *Xenopus*

¹Corresponding author.

U2 distal elements, the SV40 enhancer, human histone H2B promoter, and the murine immunoglobulin heavy-chain enhancer have been shown to bind a common factor in HeLa cell nuclear extracts (Sive and Roeder 1986; Bohmann et al. 1987). A number of snRNA gene distal elements including U2 (Mangin et al. 1986; Ares et al. 1987; Janson et al. 1987), *Xenopus* U2 (Tebb et al. 1987), rat U2 (Tani et al. 1983), and chicken U4B (McNamara et al. 1987) contain consensus binding sites for the transcription factor Sp1 (Briggs et al. 1986). Specific binding of this factor to motifs in the human U2 distal element has been demonstrated (Ares et al. 1987; Janson et al. 1987).

We have described previously the characterization of a human gene for U4C RNA (Bark et al. 1986). We report here an analysis of the 5'-flanking sequence requirements for transcription of this U4C gene following transfection into HeLa cells. We demonstrate by DNase I and dimethylsulfate (DMS) methylation protection footprinting that the distal element of this gene binds

the protein factors CREB and the octamer-binding factor, NF-A in vitro and that binding motifs for NF-A and AP-2 (or a protein recognizing a very similar motif) are required for maximal U4C gene transcription in HeLa cells.

Results

Flanking sequence requirements for U4C transcription

We have described previously the isolation of a human U4C gene cloned with ~340 bp of 5'-flanking sequence and 170 bp of 3'-flanking sequence, which is transcriptionally active when microinjected into *Xenopus laevis* oocytes (Bark et al. 1986). To study transcription following transfection into HeLa cells, a U4C maxigen, U4Cmax, was first constructed by inserting a 17-bp segment of the pGEM2 (Promega-Biotech) polylinker into the *Fsp*I site at +8 in the U4C gene (Fig. 1A). This construct is transcribed equally efficiently as the wild-type U4C gene in *Xenopus* oocytes (not shown).

Figure 1. Structure of U4Cmax plasmids. (A) U4C maxigene. A 17-bp fragment from the pGEM2 polylinker, containing restriction sites for *Bam*HI and *Xba*I, was inserted into the *Fsp*I site at position +8 in the U4C gene. (Open arrow) U4C structural gene, maxi-insert shown filled in; (filled boxes) vector sequences. (B) Structure of the U4C 5'-flanking sequence and 5' deletion constructions. The positions of the putative distal and proximate elements (DE and PE) are indicated by filled boxes. The sequence of these regions is given (DE: -196 to -247; PE: -46 to -63), and putative binding site motifs are indicated. Pertinent restriction sites used in plasmid construction are shown. The transcriptional activities are indicated relative to that of the full-length pU4Cmax. The numbers given represent mean values from at least three and an average of five separate estimations. (C) Comparison of U4C distal

The relative transcription efficiencies of the 5'-deletion constructions are shown in Figures 1B and 2A. Deletion to position -239 in the 5'-flanking sequence has

Figure 2. Transcription of U4C constructions in HeLa cells. Cells were transfected with the appropriate U4C plasmid, together with a reference plasmid pSX β ⁺. Variation in transfection efficiency was normalized by S1 mapping of β -globin transcripts (not shown). U4Cmax transcripts were quantified by S1 mapping of the 5' terminus using a probe 5'-end-labeled at position 125 in the maxigene as described (see Materials and methods). The shortest of the protected fragments corresponds to initiation at +1 in the U4C gene. (A) Deletions of the 5'-flanking sequence; (B) mutation and/or deletion of individual motifs in the distal element; (C) moving the distal element from its natural position.

no effect on the level of transcription, even though the putative CREB-binding site is interrupted; this suggests that the motif is not important for U4C transcription in this system. Plasmid dl-239 was transcribed at the same level as U4Cmax even when much lower amounts of DNA (0.5 μ g, rather than 8 μ g) were used for transfection (not shown). Deletion of the octamer (dl-218) resulted in an approximately eightfold reduction in transcription, indicating that the octamer motif is important for maximal transcription, as in other snRNA genes. Further deletion of 5'-flanking sequence to position -195, including the putative AP-2 motif, reduces transcription again approximately fivefold to a basal level of 2–3%. Loss of sequences between positions -195 and -121 does not result in a further reduction in transcriptional efficiency, but deletion to position -50 abolishes transcription. This suggests that sequences between -50 and -121 are required for any detectable level of U4C expression. It is likely that this element is the snRNA gene proximal element, which has been partially deleted in dl-50.

The transcriptional efficiencies of the constructions shown in Figure 1B indicate that the regions containing the AP-2 and octamer motifs, but not those containing the CRE motif, are important for U4C transcription. To test this, we made an additional four constructions, all of which extend to -251 but are altered specifically in the distal element (Figs. 1D and 2B). The mutated motifs were shown to be unable to bind transcription factors by

their inability to compete for binding in footprint competition experiments (not shown). As anticipated from the deletion constructions, mutation of the CRE motif in DE:octa.AP-2 had no effect on transcriptional activity. Transcription of DE:CRE.AP-2, in which the octamer motif is mutated, however, was reduced to ~14% of the wild-type level, similar to that from plasmid dl-218, in which the octamer and all upstream sequences are deleted. This confirms that the octamer is required for maximal transcription of U4C. The sequence between positions -195 and -216, which consists almost entirely of the AP-2 motif spanning between -216 and -201, has also been deleted in constructs DE:octa and DE:CRE.

With a distal element containing only the octamer, DE:octa, the transcription level is reduced to ~35% of wild type, whereas with a distal element containing only the CRE, transcription is reduced to the basal level. The CRE motif is therefore unable to enhance transcription in the absence of either the octamer or both octamer and AP-2 motifs. In conclusion, mutation or deletion of the octamer or putative AP-2 motifs separately results in an approximately eightfold and three- to fivefold reduction in transcriptional activity, irrespective of the presence of the other motif; this indicates that these two motifs stimulate transcription independently (Fig. 1B,D and Fig. 2).

The distal element of snRNA genes has been shown to have enhancer-like properties (Skuzeski et al. 1984; Ares

et al. 1985; Mattaj et al. 1985; Mangin et al. 1986; Murphy et al. 1987). To test whether this is also true for the human U4C gene, we made another series of constructions in which the distal element is moved relative to the U4C gene (Fig. 1E). Moving the distal element farther upstream from the U4C gene in DE:5'ins (by insertion of a 738-bp *PvuII* fragment of λ DNA at position -195) and also inversion of the distal element in DE:5'inv (by inverting the ~150 bp *EcoRI*-*SspI* fragment between positions -345 and -195) reduced U4C transcription to a low level, although significantly above the basal level of 2–3% (Figs. 1E and 2C). Similarly, moving the distal element, in either orientation, to the 3' side of the gene in DE:3' and DE:3'inv resulted in low but significant stimulation of transcription (Figs. 1E and 2C). Insertion of a *BglII* linker at position -195 in DE:ins.*BglII* had little effect on transcription, although moving the distal element closer to the gene by deleting the sequence between -195 and -121 reduced transcription to ~40%. These results suggest that the U4C distal element does have enhancer-like properties. Small alterations in the position of the element that do not change its orientation have little or no effect on its enhancing activity. However, moving the distal element a long way upstream, inverting it, or placing it on the 3' side of the gene results in only weak stimulation of U4C transcription.

The U4C distal element binds sequence-specific transcription factors

To determine whether the sequence motifs in the U4C distal element were recognized by DNA-binding proteins, we carried out DNase I protection footprint analysis (Galas and Schmitz 1978) using HeLa nuclear extract (Ohlsson and Edlund 1986). A single region of ~30 bp was seen to be protected from DNase I digestion on both DNA strands (Fig. 3A, lane 5; Fig. 3B, lane 5). To establish that this protection was due to sequence-specific protein binding, we carried out footprint competition experiments using unlabeled competitor DNA fragments. Competing with a U4C *HinfI* fragment, which spans from positions -219 to -253, resulted in loss of protection over the whole region, whereas nonspecific DNA failed to compete (Fig. 3A, lanes 7, 8).

The footprint seen over the U4C distal element spans two of the three recognized sequence motifs: the CRE and the octamer. To see whether the observed protection was due to the binding of more than a single factor, we competed the footprint binding with oligodeoxyribonucleotides (oligonucleotides) containing either the U4C octamer or CRE motif (Fig. 3B, lanes 7, 8). In each case, only part of the protection is lost—that part corresponding to the motif in the competing oligonucleotide—therefore indicating that two factors do indeed bind to the U4C distal element. Progressive loss of binding of one factor induced by adding increasing amounts of competitor DNA was found to have no effect on the binding of the other factor, and vice versa (Fig. 3C).

Hence, there is no evidence for cooperativity in the binding of these two factors. To identify whether the factor binding to the U4C octamer sequence was the same as that previously found to bind octamer motifs in a number of other RNA polymerase II promoters and enhancers (Singh et al. 1986; Sive and Roeder 1986; Bohmann et al. 1987), we used a fragment spanning the octamer from the mouse immunoglobulin MOPC-41 κ light-chain gene as a competitor. Binding to the U4C octamer was effectively competed showing that these two octamer motifs are bound by the same factor in HeLa nuclear extracts (Fig. 3A, lane 6; Fig. 3C, lanes 6–9). Similarly, by footprint competition, the putative U4C CRE motif was identified as binding a factor similar or identical to the CREB protein (see below and Fig. 4B).

To confirm the results obtained above and to characterize further the interaction of the individual factors with the DNA template, we carried out DMS footprinting using the method described by Janson et al. (1987). The G residues protected from DMS methylation by protein binding lie within the sequence protected from DNase I digestion and correspond to those within, or immediately adjacent to, the CRE and octamer motifs (Fig. 4). There is, however, one G residue between the two motifs that remains unprotected (Fig. 4A, lane 2). In confirmation of the competition results from DNase I footprinting, the protection of each of the two motifs can be independently competed with oligonucleotides, including the appropriate binding site, and is totally lost when the U4C *HinfI* -253/-219 fragment is used as a competitor (Fig. 4A, lanes 3–5). To determine whether the factor binding to the putative CRE motif in the U4C distal element was the cAMP response element binding protein, CREB (Montminy and Bilezikjian 1987; Montminy et al. 1986), we used as a competitor an oligonucleotide of sequence derived from the rat somatostatin gene (positions -29 to -63), which spans the CRE motif. This oligonucleotide eliminated protection of the putative CRE motif in the U4C distal element, showing that it is most likely the same factor that recognizes the motifs in the U4C and somatostatin genes (Fig. 4B, lane 3). Binding to the U4C CRE was also competed with CRE-containing restriction fragments from the rat somatostatin gene (Montminy et al. 1986; not shown).

Protection of the putative AP-2 motif was not seen with either footprinting method, even when greater amounts of extract were used. However, this is not expected as binding of AP-2 itself has only previously been reported using partially purified extracts. A summary of the footprinting results showing the extent of protection from DNase I and DMS by CREB and NF- κ B is shown in Figure 5.

Discussion

We have used S1 mapping to determine the promoter requirements for *in vivo* transcription of a human U4C gene following transfection into HeLa cells. Progressive deletion of the 5'-flanking sequence identified two functional regions. The first, located between positions

Figure 3. DNase I protection footprinting of the U4C distal element. Fragments 5'-end-labeled at the *EcoRI* (in the polylinker) or *SspI* (–195) sites were incubated with 10 μ g of HeLa nuclear extract in the presence of various competitor DNA fragments before DNase I digestion. The positions of the CRE, octamer, and putative AP-2 motifs, compared with the protected region, are indicated. In each case, lanes 1 and 2 show G and C sequence ladders of the probe fragment, and lanes 3 and 4 (control) are reactions carried out in the absence of extract. Competitor DNAs are indicated (see also Materials and methods). (A) Noncoding strand; (B) coding strand; [C] competition with increasing amounts (~25- to 200-fold molar excess from left to right) of the indicated competitor fragments.

Figure 4. DMS methylation protection footprinting of the U4C distal element. Fragments end labeled at the *Eco*RI (in the polylinker) or *Sau*96 (–123) sites were incubated with extract in the presence of the competitor DNA fragments indicated before treatment with DMS. The positions of the CRE, octamer, and putative AP-2 motifs, compared with the protected region, are indicated. In both cases, reactions shown in lane 1 contained no extract and those in lane 2 showed no competitor DNA. (A) Noncoding strand; (B) coding strand.

–195 and –239, is required for maximal transcription and corresponds well both in position and function to the distal element characterized in other vertebrate snRNA genes (Skuzeski et al. 1984; Westin et al. 1984; Ares et al. 1985; Mattaj et al. 1985; Mangin et al. 1986; Murphy et al. 1987). The second element, which is required for detectable levels of transcription, has a 5' border that lies between positions –50 and –121. Although we have not characterized this element further, it is likely that it corresponds to the snRNA proximal element (Skuzeski et al. 1984; Ares et al. 1985; Ciliberto et al. 1985; Mattaj et al. 1985; Murphy et al. 1987), a motif similar to that present in the U4C gene between positions –46 and –63 (Bark et al. 1986; see Fig. 1).

The snRNA distal elements from a number of genes,

including human U2 (Westin et al. 1984; Ares et al. 1985; Mangin et al. 1986), human U1 (Murphy et al. 1987), *Xenopus* U2 (Mattaj et al. 1985), and *Xenopus* U1 (Ciliberto et al. 1985, 1987; Krol et al. 1985), have been identified as having enhancer-like properties in that they are capable of stimulating transcription independently of location or orientation, although the magnitude of transcriptional stimulation is variable. We have shown that the human U4C gene distal element is also capable of acting as an enhancer, although with variable efficiency (Fig. 1E).

The human U4C gene distal element itself has two functional components, the first located between –195 and –218 and the second between –218 and –239, both of which are required for maximal transcription (Figs. 1 and 2). The sequence between –218 and –239 contains the octamer motif, characteristically found in the distal elements of vertebrate snRNA genes, as well as in the control elements of some protein-coding genes. The binding of a protein factor or factors to the distal elements of snRNA genes has been demonstrated previously *Xenopus* U1, U2, and U5 (Ciliberto et al. 1987; Kazmaier et al. 1987; Tebb et al. 1987), chicken U4B (McNamara et al. 1987), and human U2 (Ares et al. 1987; Janson et al. 1987). In the case of the U2 genes, the distal element has been shown to be bound by the octamer protein, NF-A, which can also recognize the octamer motifs in the SV40 enhancer, human histone H2B promoter, and mouse immunoglobulin heavy-chain enhancer (Sive and Roeder 1986; Ares et al. 1987; Bohmann et al. 1987; Janson et al. 1987). An octamer-binding factor binds to this motif in the 5'-flanking sequence of mouse and *Xenopus* U6 genes (Bark et al. 1987; Carbon et al. 1987). We have shown here that the octamer motif in the U4C gene is also bound in vitro by this factor in HeLa cell extracts.

The other U4C distal element component, between –195 and –218, contains the AP-2 motif. We were unable to see any evidence for protein binding to this region in the footprinting experiments; but on the basis of the observed similarity between U4C and SV40 AP-2 motifs and the short length of the region involved, we consider it highly likely that AP-2, or a protein with a very similar binding motif, is involved in U4C transcription.

The 5'-flanking sequences of a number of snRNA genes also contain recognition sites for transcription

Figure 5. Summary of footprinting results over the U4C distal element. The CRE, octamer, and putative AP-2 motifs are indicated in boxes. Sequences clearly protected from DNase I cleavage on each strand are indicated by a solid line. Dotted lines indicate the first unprotected nucleotide on either side of the protected region (and therefore the maximum extent of the footprint). Sites of enhanced DNase I cleavage are indicated by arrows. G residues protected from methylation by DMS are indicated in black.

factor Sp1 (Briggs et al. 1986), as discussed by Mangin et al. (1986), and this factor has been shown to bind to sites in the human U2 distal element (Ares et al. 1987; Janson et al. 1987). However, there are no Sp1 motifs in the flanking sequence of the human U4C gene.

The requirement for both the octamer and putative AP-2 motif for full transcriptional activation by the U4C distal element, together with the potential involvement of the CRE motif, is not consistent with observations from human U2 and *Xenopus* U1 genes that very short deletions, presumably affecting only a single element, completely abolish distal element function (Ares et al. 1985; Mangin et al. 1986; Ciliberto et al. 1987). The chicken U4B distal element has also been shown to comprise multiple functional elements (McNamara et al. 1987), suggesting that snRNA distal elements need not necessarily be very compact (Ciliberto et al. 1987), but like other enhancers, e.g., SV40, can be made up from multiple elements, each having partial enhancing activity (Herr and Clarke 1986). There is an additional difference between the transcription of human U2 and U4C genes, in that the U4C distal element is not absolutely required for expression (a residual transcription level of 2–3% is still observed when it is deleted), whereas in the absence of the distal element, the U2 gene is completely inactive in HeLa cells (Mangin et al. 1986; L. Janson, P. Weller, and U. Pettersson, in prep.). Considerable transcription activity in the absence of their distal elements has also been observed for the chicken U4B and *Xenopus* U5 genes with levels 27% and 5–10% of wild type, respectively (Kazmaier et al. 1987; McNamara et al. 1987).

This is the first time that transcriptional requirement for a binding site resembling that of transcription factor AP-2 has been demonstrated in a snRNA gene. Interestingly, there is a related sequence in the chicken U4B gene distal element, TTCCCAG, which may be important for transcription but comprises only half of the bipartite motif in the SV40 enhancer and U4C distal element, although it is separated from the octamer by 5 bp in both U4 genes (McNamara et al. 1987). There is also a consensus AP-2-binding site upstream of the third Sp1-binding site in the human U2 distal element (L. Janson, P. Weller, and U. Pettersson, in prep.). Significantly, AP-2-dependent transcription can be induced by phorbol esters, including TPA, which is known to stimulate protein kinase C (Nishizuka 1986), and also by forskolin, which increases intracellular cAMP levels and is thought to stimulate transcription via cAMP-dependent protein kinase (Cohen 1985), thus raising the possibility that snRNA gene expression could be regulated in this manner.

The human U4C gene distal element CRE motif, 6 bp upstream of the octamer, was found to bind a protein factor similar or identical to CREB (Montminy and Bilezikjian 1987). This motif has been found in control elements both from genes known to be regulated transcriptionally by cAMP and from genes where such regulation is not known (for discussion, see Montminy et al. 1986). The CREB protein has been found to be phosphorylated

in vitro by the catalytic subunit of cAMP-dependent protein kinase, suggesting that regulation via cAMP may be a consequence of the phosphorylation of CREB in vivo (Montminy and Bilezikjian 1987). Although the human U4C distal element was found to bind CREB, no deleterious effect on transcription was seen when this motif was deleted, nor does the presence of this motif exert any positive effect in the absence of the octamer and/or AP-2 motifs. However, a specific functional role for this element is suggested by the observation that an identical motif appears in a very similar position 7 bp upstream of the octamer in the chicken U4B gene. In addition, it appears from gel mobility shift data of McNamara et al. (1987) that there may be an additional protein factor in HeLa cell nuclear extract, which binds to sequences upstream of the octamer motif in the chicken U4B gene and might represent binding to the CRE motif. Failure to find any effect on U4C transcription by deletion of the CRE motif might be due to our use of the transient expression assay system, although similar results have been obtained with *Xenopus* oocyte injection experiments for both human U4C and chicken U4B genes (data not shown; McNamara et al. 1987). It is notable in this context that no effect on the basal level of rat somatostatin or human glycoprotein hormone α -subunit gene transcription was seen in transient assays following deletion of the CRE (Montminy et al. 1986; Silver et al. 1987).

We have identified binding sites in the human U4C distal element for two transcription factors: the octamer-binding factor, NF-A, and the inducible transcription factor, CREB. In addition, a motif that is very similar to the AP-2 binding sequence is required for maximal U4C transcription in vivo. Neither CREB nor AP-2 is specific to snRNA genes, suggesting that other factors, perhaps those that recognize the proximal element, may be responsible for the unusual features of snRNA gene transcription. To maintain the full complement of snRNAs per cell in all tissues, snRNA synthesis must, in some way, be coupled to cell growth or metabolism. It remains to be seen whether U4C transcription can be induced by treatment with TPA or forskolin. However, the presence of binding sites for inducible transcription factors in the human U4C gene distal element suggests that regulation might be achieved in response to signaling systems involving both the protein kinase C and cAMP-dependent protein kinase pathways via interaction with such transcription factors as CREB and AP-2.

Materials and methods

Construction of the U4C maxigene

The cloning of the human U4C gene as an ~650-bp fragment between the *Sma*I and *Hinc*II sites of pGEM2 (Promega-Biotech), including ~340-bp 5'-flanking sequence and 170-bp 3'-flanking sequence, has been described (Bark et al. 1986). The maxigene pU4Cmax was constructed by cloning an *Eco*RI-*Fsp*I fragment containing the 5'-flanking sequence and up to +8 in the U4C gene into the *Sma*I site of pGEM2 and then an *Fsp*I-*Hind*III fragment extending 3' from position +8 to the vector

polylinker into the *HincII* site of this plasmid. This results in the insertion of a 17-bp section of the pGEM2 polylinker containing restriction sites for *Bam*HI and *Xba*I into the *Fsp*I site at +8 in the U4C gene.

Construction of pU4Cmax:S1 for S1 mapping of U4Cmax transcripts

The *Ssp*I–*Nla*III fragment from U4Cmax containing the sequence from –195 to +107 in the maxigene was ligated, together with a double-stranded oligonucleotide extending from the *Nla*III site to +125 on both strands, into the Bluescribe vector pBSM13+ (Vector Cloning Systems) cleaved with *Hinc*II. This resulted in plasmid pU4Cmax:S1, extending to a *Hinc*II site at position +125 in the gene, which can be used to generate a probe for S1 mapping the 5' ends of U4Cmax transcripts.

Construction of mutant U4Cmax templates

The 5'-flanking sequence-deleted clones were constructed as follows: dl-269: *Mn*II–*Aha*II from –269 to –239 and *Aha*II–*Hind*III fragments from –239 to the *Hind*III site of pU4Cmax were ligated between *Sma*I–*Hind*III sites of pGEM2. dl-218: pU4Cmax cleaved with *Hin*FI, end repaired (all end-repairing was by fill in with Klenow polymerase unless stated otherwise), and cleaved with *Bam*HI. The fragment (–218 to +11) was ligated between *Eco*RI (end repaired) and *Bam*HI sites of pU4Cmax. dl-239, dl-195, dl-121, and dl-50: pU4Cmax was cleaved with the appropriate enzyme (see Fig. 1), end repaired (dl-239 and dl-50 only), cleaved with *Hind*III, and the resulting U4Cmax-containing fragments were ligated between *Sma*I–*Hind*III of pGEM2.

Templates where separate elements of the U4C distal element are mutated or deleted were constructed using two double-stranded oligonucleotides, each extending from –251 to the *Hin*FI site at –218, with an *Eco*RI cohesive end at the upstream terminus. In one, the octamer motif was mutated to ATGTTTCAG, and in the other, the CRE motif was mutated to TTAAGGCC. Templates DE:octa.AP-2 and DE:CRE.AP-2 were constructed by ligating the appropriate oligonucleotide and the partial *Hin*FI (–218)–*Hind*III pU4Cmax fragment between the *Eco*RI/*Hind*III sites of pGEM2. DE:octa and DE:CRE were constructed by end repairing the oligonucleotides and ligating them, together with the *Ssp*I–*Hind*III pU4Cmax fragment, into pGEM2 cleaved with *Sma*I and *Hind*III. The orientation of the oligonucleotide in each of these constructions was determined by sequencing.

Templates in which the position of the distal element is altered were constructed as follows: DE:5'ins: A 738 bp *Eco*RV fragment of phage λ (position 8086–8834) was inserted into the *Ssp*I site of pU4Cmax. DE:5'inv: The *Sac*I–*Ssp*I fragment containing the U4C distal element was inserted, together with the *Sac*I–*Hind*III U4Cmax-containing fragment from dl-195, between the *Sma*I–*Hind*III sites of pGEM2. DE:3': The *Eco*RI–*Ssp*I fragment from U4Cmax containing the distal element was end repaired (using T4 polymerase) and inserted into the end-repaired *Pst*I site of a U4Cmax construct (pU4Cmax:dl+38), which is also cloned in the *Sma*I–*Hind*II sites of pGEM2 but contains only 38 bp of 3'-flanking sequence. The *Bam*HI–*Hind*III fragment from the resulting plasmid, was then used to replace the *Bam*HI–*Hind*III fragment of dl-195. DE:3'inv: This template was cloned in the same way as DE:3' but with the end-repaired *Eco*RI–*Ssp*I fragment in the reverse orientation. DE:ins.*Bgl*II: An 8-bp *Bgl*II linker was inserted at the *Ssp*I site at position –195. DE:dl(–195/–121): The *Eco*RI–*Ssp*I distal-

element-containing fragment and the *Hae*III–*Hind*III U4C-gene-containing fragment from pU4Cmax were inserted into pGEM2 cleaved with *Eco*RI and *Hind*III.

Transfections

Transfection of HeLa cells was carried out by the calcium phosphate coprecipitation technique (Wigler et al. 1978). A total of 10 μ g of DNA was used per 6-cm plate. To measure transfection efficiency, all U4C constructions were cotransfected with 2 μ g of plasmid pSX β + carrying a rabbit β -globin reference gene driven by the SV40 enhancer (Banerji et al. 1981). To minimize possible competition effects, the ratio between the U4C construction and pSX β + was 4 : 1. Cells were shocked with 15% glycerol 6–15 hr after addition of DNA and harvested 48 hr post-transfection, and RNA was extracted from the cytoplasmic fraction following lysis by IsoB Nonidet P-40 (Akusjärvi and Pettersson 1978).

S1 nuclease mapping

U4C maxigene transcript levels were measured using a probe 5'-end labeled with [γ -³²P]ATP at the *Hinc*II site of pU4Cmax:S1. Isolation of a uniquely end-labeled fragment was not usually carried out. About 10⁵ cpm of probe was annealed with 1–10 μ g of cytoplasmic RNA in 10 μ l of hybridization buffer [80% formamide, 40 mM PIPES (pH 6.4), 0.4 M NaCl, 1 mM EDTA] and incubated at 50°C overnight. Samples were diluted with 400 μ g of S1 buffer [30 mM sodium acetate (pH 4.4), 0.28 M NaCl, 4 mM ZnSO₄, 5% glycerol, 12.5 μ g/ml denatured salmon sperm DNA] and digested with ~110 units S1 nuclease for 30 min at 20°C. Digestion was stopped by phenol extraction, and samples were precipitated with ethanol before resolving by electrophoresis in 6% sequencing gels. Transcript levels were quantified by scanning densitometry. Rabbit β -globin transcripts were measured in a similar way using a probe generated by cleavage of pSX β + with *Bam*HI and 5'-end-labeling with [γ -³²P]ATP. No further cleavage of the probe was carried out. Hybridization (at 56°C) and S1 digestion were carried as above. Relative transfection efficiency was determined by scanning densitometry of the protected 209-nucleotide fragment.

Footprinting analysis

HeLa cells were grown in spinner cultures to a density of 0.5 $\times$ 10⁶ cells per milliliter. Nuclear extract was prepared using the method of Ohlsson and Edlund (1986), as described by Janson et al. (1987). DNA fragments uniquely end labeled on each strand using polynucleotide kinase and [γ -³²P]ATP were generated following cleavage of U4Cmax with *Eco*RI or *Ssp*I and second cleavage with *Bam*HI and *Eco*RI, respectively. DNase I protection footprint reactions were carried out as described by Bark et al. (1987), using 10 μ g of protein extract and 1 $\times$ 10⁴ to 2 $\times$ 10⁴ cpm end labeled fragment (~3 ng). Fragments were resolved in 6% or 10% sequencing gels. Competitor DNA fragments included in some reactions comprise the following: U4C-derived octamer oligonucleotide of sequence CGTTGATTTGCATAAGAT; an octamer-containing 75-bp *Eco*RI–*Pvu*II fragment of the mouse immunoglobulin κ light-chain MPOC-41 gene (gift from H. Singh); a U4C-derived oligonucleotide containing the CRE motif of sequence ACCGAGTCATCTCTGTGACGTCACG; a U4C *Hin*FI fragment spanning from –218 to –254; an oligonucleotide containing the rat somatostatin CRE of sequence GGGCGCCTCCTTGGCTGACGTCAGAGAGAGTTT; and linearized pBR325. DMS methylation protection footprinting was carried out as de-

scribed by Janson et al. (1987), using 10–30 µg of protein and DNA probes kinase labeled at the *EcoRI* (vector polylinker) or *Sau96* (–123) sites of U4Cmax. Competitor DNAs were as described above.

Acknowledgments

We thank H. Hultberg (Kabigen AB) for synthetic oligonucleotides and Britt-Marie Carlberg for excellent technical assistance. We are grateful to H. Singh for a plasmid containing the mouse κ light-chain promoter. This work was supported by the Swedish Medical Research Council, the Swedish National Board for Technical Development, and the Marcus Borgström Foundation at Uppsala University. P.W. was supported by a fellowship in the Royal Society European Science Exchange Programme. L.J. is a fellow of the Swedish Cancer Society.

References

- Akusjärvi, G. and U. Pettersson. 1978. Nucleotide sequence at the junction between the coding region of the adenovirus 2 hexon messenger RNA and its leader sequence. *Proc. Natl. Acad. Sci.* **75**: 5822–5826.
- Ares, M., Jr., M. Mangin, and A.M. Weiner. 1985. Orientation-dependent transcriptional activator upstream of a human snRNA gene. *Mol. Cell. Biol.* **5**: 1560–1570.
- Ares, M., Jr., J.-S. Chung, L. Giglio, and A.M. Weiner. 1987. Distinct factors with Sp1 and NF-A specificities bind to adjacent functional elements of the human U2 snRNA gene enhancer. *Genes Dev.* **1**: 808–817.
- Banerji, J., S. Rusconi, and W. Schaffner. 1981. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**: 299–308.
- Bark, C., P. Weller, J. Zabielski, and U. Pettersson. 1986. Genes for human U4 small nuclear RNA. *Gene* **50**: 333–344.
- Bark, C., P. Weller, J. Zabielski, L. Janson, and U. Pettersson. 1987. A distant enhancer element is required for polymerase III transcription of a U6 RNA gene. *Nature* **328**: 356–359.
- Bohmann, D., W. Keller, T. Dale, H. Schöler, G. Tebb, and I.W. Mattaj. 1987. A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. *Nature* **325**: 268–272.
- Briggs, M.R., J.T. Kadonaga, S.P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* **234**: 47–52.
- Carbon, P., S. Murgo, J.-P. Ebel, A. Krol, G. Tebb, and I.W. Mattaj. 1987. A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II. *Cell* **51**: 71–79.
- Ciliberto, G., R. Buckland, R. Cortese, and L. Philipson. 1985. Transcription signals in embryonic *Xenopus laevis* U1 RNA genes. *EMBO J.* **4**: 1537–1543.
- Ciliberto, G., F. Palla, G. Tebb, I.W. Mattaj, and L. Philipson. 1987. Properties of a U1 RNA enhancer-like sequence. *Nucleic Acids Res.* **15**: 2403–2516.
- Cohen, P. 1985. The coordinated control of metabolic pathways by broad-specificity protein kinases and phosphatases. *Curr. Top. Cell. Regul.* **17**: 23–37.
- Delegeane, A.M., L.H. Ferland, and P.L. Mellon. 1987. Tissue-specific enhancer of the human glycoprotein α -subunit gene: Dependence on cAMP-inducible elements. *Mol. Cell. Biol.* **7**: 3994–4002.
- Eliceiri, G.L. 1980. Formation of low molecular weight RNA species in HeLa cells. *J. Cell. Physiol.* **102**: 199–207.
- Falkner, F.G. and H.G. Zachau. 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. *Nature* **310**: 71–74.
- Fredriksen, S., P. Hellung-Larsen, and E. Gram-Jensen. 1978. The differential inhibitory effect of α -amanitin on the synthesis of low molecular weight RNA components in BHK cells. *FEBS Lett.* **87**: 227–231.
- Galas, D.J. and A. Schmitz. 1978. DNase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5**: 3157–3170.
- Harvey, R.P., A.J. Robins, and J.R.E. Wells. 1982. Independently evolving chicken histone H2B genes: Identification of a ubiquitous H2B-specific 5' element. *Nucleic Acids Res.* **10**: 7851–7863.
- Herr, W. and J. Clarke. 1986. The SV40 enhancer is composed of multiple functional elements that can compensate for one another. *Cell* **45**: 461–470.
- Imagawa, M., R. Chui, and M. Karin. 1987. Transcription factor AP-2 mediates induction by two different signal-transduction pathways: Protein kinase C and cAMP. *Cell* **51**: 251–260.
- Janson, L., C. Bark, and U. Pettersson. 1987. Identification of proteins interacting with the enhancer of human U2 small nuclear RNA genes. *Nucleic Acids Res.* **15**: 4997–5016.
- Kazmaier, M., G. Tebb, and I.W. Mattaj. 1987. Functional characterization of *X. laevis* U5 snRNA genes. *EMBO J.* **6**: 3071–3078.
- Krol, A., E. Lund, and J.E. Dahlberg. 1985. The two embryonic U1 RNA genes of *Xenopus laevis* have both common and gene-specific transcription signals. *EMBO J.* **4**: 1529–1535.
- Krol, A., P. Carbon, J.-P. Ebel, and B. Appel. 1987. *Xenopus tropicalis* U6 snRNA genes transcribed by pol III contain the upstream promoter elements used by pol II-dependent U snRNA genes. *Nucleic Acids Res.* **15**: 2463–2478.
- Kunkel, G.R. and T. Pederson. 1988. Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used. *Genes Dev.* **2**: 196–204.
- Kunkel, G.R., R.L. Maser, J.P. Calvet, and T. Pederson. 1986. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc. Natl. Acad. Sci.* **83**: 8575–8579.
- Mangin, M., M. Ares, Jr., and A.M. Weiner. 1986. Human U2 small nuclear RNA genes contain an upstream enhancer. *EMBO J.* **5**: 987–995.
- Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325**: 673–678.
- Mattaj, I.W. 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell* **46**: 905–911.
- Mattaj, I.W., S. Lienhard, J. Jiricny, and E.M. De Robertis. 1985. An enhancer-like sequence within the *Xenopus* U2 gene promoter facilitates the formation of stable transcription complexes. *Nature* **316**: 163–167.
- McNamara, K.J., R.J. Walker, K.A. Roebuck, and W.E. Stumph. 1987. Transcription signals of a U4 small nuclear RNA gene. *Nucleic Acids Res.* **15**: 9239–9254.
- Mitchell, P.J., C. Wang, and R. Tjian. 1987. Positive and negative regulation of transcription in vitro: Enhancer-binding protein AP-2 is inhibited by SV40 T antigen. *Cell* **50**: 847–861.
- Montminy, M.R. and L.M. Bilezikjian. 1987. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. *Nature* **328**: 175–178.
- Montminy, M.R., K.A. Sevarino, J.A. Wagner, G. Mandel, and R.H. Goodman. 1986. Identification of a cyclic-AMP-re-

- sponse element within the rat somatostatin gene. *Proc. Natl. Acad. Sci.* **83**: 6682–6686.
- Murphy, J.T., J.T. Skuzeski, E. Lund, T.H. Steinberg, R.R. Burgess, and J.E. Dahlberg. 1987. Functional elements of the human U1 promoter. *J. Biol. Chem.* **262**: 1795–1803.
- Nishizuka, Y. 1986. Studies and perspectives of protein kinase C. *Science* **233**: 305–312.
- Ohlsson, H. and T. Edlund. 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell* **45**: 35–44.
- Parslow, T.G., D.L. Blair, W.J. Murphy, and D.K. Granner. 1984. Structure of the 5' ends of immunoglobulin genes: A novel conserved sequence. *Proc. Natl. Acad. Sci.* **81**: 2650–2654.
- Reddy, R., D. Henning, G. Das, M. Harless, and D. Wright. 1987. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J. Biol. Chem.* **262**: 75–81.
- Sharp, P.A. 1987. Splicing of messenger RNA precursors. *Science* **235**: 766–771.
- Silver, B.J., J.A. Bokar, J.B. Virgin, E.A. Vallen, A. Milsted, and J.H. Nilson. 1987. Cyclic AMP regulation of the human glycoprotein hormone α -subunit is mediated by an 18 base-pair element. *Proc. Natl. Acad. Sci.* **84**: 2198–2202.
- Singh, H., R. Sen, D. Baltimore, and P.A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* **319**: 154–158.
- Sive, H.L. and R.G. Roeder. 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin, and U2 small nuclear RNA (snRNA) genes. *Proc. Natl. Acad. Sci.* **83**: 6382–6386.
- Skuzeski, J.M., E. Lund, J.T. Murphy, T.H. Steinberg, R.R. Burgess, and J.E. Dahlberg. 1984. Synthesis of human U1 RNA. II. Identification of two regions of the promoter essential for transcription initiation at position +1. *J. Biol. Chem.* **259**: 8345–8352.
- Tani, T., N. Watanabe-Nagasu, N. Okada, and Y. Ohshima. 1983. Molecular cloning and characterization of a gene for rat U2 small nuclear RNA. *J. Mol. Biol.* **168**: 579–594.
- Tebb, G., D. Bohmann, and I.W. Mattaj. 1987. Only two of the four sites of interaction with nuclear factors within the *Xenopus* U2 gene promoter are necessary for efficient transcription. *Nucleic Acids Res.* **15**: 6437–6453.
- Westin, G., E. Lund, J.T. Murphy, U. Pettersson, and J.E. Dahlberg. 1984. Human U2 and U1 RNA genes use similar transcription signals. *EMBO J.* **3**: 3295–3301.
- Wigler, M., A. Pellicer, S. Silverstein, and R. Axel. 1978. Biochemical transfer of single-copy eukaryotic genes using total cellular DNA as donor. *Cell* **14**: 725–731.

Transcription analysis of a human U4C gene: involvement of transcription factors novel to snRNA gene expression.

P Weller, C Bark, L Janson, et al.

Genes Dev. 1988, **2**:

Access the most recent version at doi:[10.1101/gad.2.11.1389](https://doi.org/10.1101/gad.2.11.1389)

References

This article cites 48 articles, 16 of which can be accessed free at:
<http://genesdev.cshlp.org/content/2/11/1389.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.

USC
SCIENTIFIC
SIDE BY SCIENCE