Identification of proteins interacting with the enhancer of human U2 small nuclear RNA genes

Lars Janson, Christina Bark and Ulf Pettersson*

Department of Medical Genetics, Biomedical Center, Box 589, S-751 23 Uppsala, Sweden

Received April 24, 1987; Revised and Accepted June 8, 1987

ABSTRACT

Protein/DNA interactions in the human U2 RNA gene enhancer have been characterized by DNase I footprint and DMS methylation protection analyses. Nuclear factors present in both HeLa and B cell extracts have been shown to protect an approximately 70 bp region from DNase I digestion. DMS and DNase I footprint competition studies demonstrated that the entire footprint can be accounted for by interactions with two previously identified transcription factors. One of these recognizes the so called octanucleotide-motif ATGCAAT (transcription factor IgNF-A) which has been shown to be essential for transcription. The other is the transcription factor Sp1 which binds to three target sequences located adjacent to the octameric motif. The Sp1 interactions appear to be required for full transcriptional activity. No differences in the DNase I footprint patterns or in the DMS methylation protections were observed when nuclear extracts from HeLa cells, two different B cell lines, or from the adenovirus-transformed 293 cell line were compared.

INTRODUCTION

The genes coding for the snRNAs U1, U2 and U4 are believed to be transcribed by RNA polymerase II (pol II) due to their sensitivity to low concentrations of α -amanitin (1,2,3). The snRNA genes differ from typical class II genes in lacking TATA-boxes, although transcription has been shown to be enhancer dependent (2,4,5,6, Bark et al., manuscript in preparation). Transcription of U2 genes is dependent of two control elements: a basal promoter, which appears to be a functional equivalent of the TATA-box and which is located between the cap site and position -62. A second element, a transcriptional activator, is present within a 72 bp SmaI fragment, located 198-270 bp upstream of the transcriptional start site (see 2,4,6 and Fig. 1a). U RNA gene transcription differs from the transcription of

most other genes in several respects; it occurs with a high efficiency in all cell types and a U gene specific promoter/enhancer seems to be required for proper transcriptional termination (7,8). Thus one might expect that enhancers of U RNA genes interact with a set of unique transcription factors, although so far our knowledge about transcription factors required for U RNA transcription is very limited. Recently Sive et al. (9) reported that a previously identified transcription factor, the so-called octa-factor, interacts with the enhancer of human U2 RNA genes and Bohmann et al. (10) have shown that the same factor interacts with the U2 enhancer of *X.laevis*. This factor, which has been given many different names (IgNF-A, 11; NF III, 12; octamer-binding factor, 10 and H2B consensus sequence-binding factor, 9), interacts with an octameric sequence motif (13,14) present in several different enhancers, promoters, and in the adenovirus replication origin. It seems to stimulate both transcription and replication in in vitro systems.

In the present study we have characterized DNA/protein interactions which occur in vitro between the human U2 enhancer and factors present in HeLa and B cell-derived extracts using DNase I and DMS footprinting.

MATERIALS AND METHODS

Cell cultures and preparation of nuclear extracts

Cells were grown in suspension to a density of 0.5×10^6 cells / ml. HeLa and 293 cells were grown in spinner medium supplemented with 7 % newborn calf serum and the Namalva and BJA-B cells in RPMI 1640, supplemented with 7 % foetal calf serum. Nuclear extracts were prepared as described by Ohlsson and Edlund (15), with the exception that only one nuclear fraction was collected by precipitation with 0.25 g/ml of ammonium sulphate.

Labeling of DNA templates

A 72 bp SmaI fragment located between positions -198 and -270 in the human U2 enhancer was subcloned in both orientations utilizing the HincII site of the vector pGEM2 (Promega Bio-tech). The resulting subclones were designated pEN- and pEN+.

3' end-labeled templates were prepared as follows: pEN+ and pEN- were labeled in the EcoRI site of the plasmid polylinker using (α -³²P)dTTP (Amersham) and Klenow polymerase (New England BioLabs), followed by redigestion with Pvu II to yield a 160 bp fragment. Fragments were purified in 5% polyacrylamide gels. The specific activity of the labeled templates was approximately 10^7 cpm/ μ g. The reactions were performed as described by Maniatis et al. (16).

Competitor DNAs

pBR322 was linearized by HindIII digestion. A 366 bp HindIII-KpnI fragment containing the SV40 promoter/enhancer was used in some experiments. This fragment was also digested further with BglI and FokI to yield an 88 bp fragment containing four Spl binding sites. Double stranded synthetic oligonucleotides were also used in some experiments.

DNase I footprint reactions

Reactions were performed in a final volume of 50 μ l containing: 25 mM Hepes (pH 7.8), 50 mM KCl, 0.05 mM EDTA, 0.5 mM PMSF and 5% glycerol. All reactions contained 2 μ g of poly(dI-dC) (Pharmacia), as nonspecific competitor, and 3 ng of labeled fragment. The amounts of protein extract and competitor DNA used are specified in the text. Reaction mixtures containing competitor DNAs were preincubated at 25°C for 10 min. The labeled fragments were then added, followed by an additional incubation for 15 min at 25°C. DNase I (Pharmacia) digestion was initiated by the addition of $MgCl_2$ to a final concentration of 5 mM. After DNase I digestion for 1 min, the reactions were terminated by the addition of 100 μ l of stop solution (100 mM Tris-HCl pH 8.0, 100 mM NaCl, 1% Sarkosyl, 10 mM EDTA, 100 μ g/ml proteinase K and 25 μ g/ml of DNA). After incubation at 65°C for 20 min, the samples were phenol extracted and analyzed in 10 % polyacrylamide-urea gels.

DMS footprint reactions

Preincubations and incubations were performed as described for the DNaseI footprint reactions, with the exception that no $MgCl_2$ was added and the amount of protein was kept constant at 120 μ g in all the reactions by the addition of bovine serum albumin (BSA). The samples were cooled on ice for approximately

Figure 1.

Organization of the 5' flanking region of a human U2 RNA gene. a) The U2 promoter region contains an enhancer which is present within a 72 bp *Sma*I fragment. The sequence motifs which were found to interact with transcription factors are indicated. The boundaries of the DNase I footprint are marked and enhanced cleavage-sites are indicated with arrows. The Gs protected from DMS methylation are indicated with triangles. Filled triangles indicate strong protection open triangles weak protection. Open circles weak and variable protection. The sequences of corresponding regions from rat (upstream of position -208), chicken (upstream of -207) and *X.laavis* (upstream of -247) U2 RNA genes are shown.

b) Sequences of the three Sp1 binding sites of the U2 enhancer. Binding site Sp1:2 is reversed relative to its position in the enhancer. Methylation protection is indicated as in panel a. The consensus sequence for Sp1 binding sites (Briggs et al., 1986) is also shown.

3 min to stabilize the DNA-protein complexes before the addition of 0.2 μ l of DMS. DMS reactions were incubated for 4 min at room temperature and then stopped by the addition of 300 μ l of 0.5 x G-stop solution (17) followed by phenol extraction. Piperidine treatment was carried out as described by Maxam and Gilbert (17) and the samples were lyophilized three times before electrophoresis.

RESULTS

DNase I footprint analysis of the human U2 enhancer

DNase I footprint experiments (18) were performed to identify factors which interact with the enhancer region of the human U2 RNA genes. In the first set of experiments an AvaII/Sau3A fragment, covering 413 bp of 5'-flanking sequences (Fig. 1a), was incubated with a crude nuclear extract from HeLa cells (15). C and CT sequence ladders (17) were used to determine the boundaries of the resulting footprint. The results revealed an easily detectable, 70 bp footprint extending between position -203 and -272 on the coding strand and between positions -207 and -275 on the noncoding strand (data not shown). Enhanced DNase I cleavage was observed at positions -251, -252, -275 and -276 on the noncoding strand. This footprint covers the region of the human U2 RNA locus to which the U2 enhancer has been mapped (2,4,6). A 72 bp SmaI fragment between positions -198 and -270 (Fig.1a) was subsequently subcloned in both orientations to be used for further characterization of the enhancer. Fig 2a shows DNase I footprints on both strands using 50 μ g of a HeLa-cell derived nuclear extract. The footprint extends from -203 to a residue in the plasmid polylinker, equivalent to position -274 on the coding strand, with no enhanced cleavage sites (Fig. 2a; lanes 1 and 2). This footprint is thus two nt longer than the one obtained with the AvaII/Sau3a fragment. This apparent discrepancy is due to the absence of DNase I cleavage sites in the polylinker sequence equivalent to position -272. The noncoding strand was protected from -207 to -275 with enhanced cleavage sites at the same positions as indicated above (Fig. 2a, lanes 4 and 5).

Figure 2.

DNase I footprint analysis of the coding and noncoding strands, using nuclear extracts from HeLa and Namalva cells.

a) Three control lanes are shown for each strand (C) where the labeled DNA fragments were DNase I digested in the absence of nuclear extract. In lanes 1 to 6 the labeled fragments have been incubated with 50 μ g of HeLa-cell nuclear extract prior to DNase I digestion. Linearized pBR322 DNA (200 ng) was added as nonspecific competitor in lanes 1,2,4 and 5. In lanes 3 and 6 200 ng of a 360 bp SV40 early promoter fragment was used as competitor. The boundaries of the footprints are indicated and enhanced cleavage sites are shown by arrows. The positions of the octa- and Sp1- related sequence-motifs are also shown. b) An identical analysis was performed using 50 μ g of a Namalva cell extract. The lanes are labeled as in panel a.

Footprint analysis of the enhancer region using B cell-derived nuclear extracts were also performed. However, no differences were observed in the footprint patterns obtained with HeLa cell and Namalva (19) cell extracts (Figs. 2a and b).

The only noticeable difference was a slightly stronger protection of the cleavage sites at -219 and -221 by the Namalva extract (Fig. 2b; lanes 1 and 2). Identical footprints (data not shown) were also obtained with extracts from another B cell line, BJA-B (20), the same cell line was used by Davidson et al.(21) to demonstrate cell-specific interactions with the SV40 enhancer. Finally, an identical footprint was obtained with an extract from the human adenovirus-transformed cell line, 293 (22, data not shown).

The U2 RNA gene enhancer and the SV40 early promoter/enhancer interact with common transcription factors

To characterize further the U2 RNA gene enhancer, footprint experiments were performed in the presence of an excess of unlabeled competitor DNA. Thus either 200 ng of linearized pBR322 DNA or the same amount of a 360 bp KpnI/HindIII fragment, containing the SV40 early promoter and its enhancer, were included in the footprint reactions. Figs 2a and 2b (lanes 3 and 6) show that the entire footprint disappeared on competition with the SV40 DNA fragment, whereas no significant changes were noted after competition with pBR322 DNA (Fig. 2; lanes 1,2,4 and 5; Fig. 6; lanes 1 and 4). As the SV40 promoter/enhancer region has been characterized in detail with regard to protein interactions, it was of interest to search for previously characterized transcription factor binding sequence motifs in the nucleotide sequence of the U2 gene enhancer. The analysis revealed four such putative sequence motifs within the region covered by the footprint (Fig.1a). The sequence 5'-ATGCAAAT-3' is located between -221 and -214. This sequence motif is known to interact with the so-called octa-factor (10,11,12). In addition, three sequence motifs, 5'-AGGGCGGGGC-3', 5'-AGGGCGTGGC-3' and 5'-GGGGCGGAGT-3' were identified. These resemble the consensus sequence established for the transcription factor Sp1 (23). Two of the latter motifs are on the coding strand whereas the third is on the noncoding strand (Fig.1a).

As a working hypothesis we therefore assumed that the U2 RNA gene enhancer might contain one octa-binding site and three Sp1 binding sites, designated Sp1:1, Sp1:2 and Sp1:3 (Fig. 1a).

Figure 3.
DMS methylation protection of the coding and noncoding strands. The labeled DNA fragments were incubated with either 120 µg of BSA (0), or with 100 µg of a Namalva nuclear extract (100). The amount of protein in all DMS footprinting reactions was kept constant at 120 µg by the addition of BSA. The separate binding sites are indicated.

DMS footprint analysis of the U2 RNA gene enhancer

DMS footprint analysis was performed to identify the bases within the DNase I-protected region of the U2 enhancer that interact with transcription factors. DMS methylates the N7 of the guanine residues in the major groove of the DNA helix, but the reaction is inhibited if a protein is in close proximity to a G-residue. Fig. 3 shows that several G-residues within the footprint are completely or partially protected from methylation when treated with DMS in the presence of nuclear extract (summarized in Table 1 and Fig. 1a). All the protected Gs were either located within the octa- or Spl- motifs or immediately adjacent to them (Fig. 3; Table 1), thus supporting our hypothesis that these sequence motifs do indeed interact with transcription factors. None of the cleavage sites or methylations

Table 1: Protection of G residues inside the U2 enhancer from DMS methylation

Position of G residues		Competitor DNA			
	Coding strand	Noncoding strand	0	Sp1 ^{a)}	Octa ^{b)}
Octa	-202	-199			
	-206	-200			
	-211	-201			
		-212			
Sp1 ₁	-219	-218	++ ^{c)}	++	
		-222	++	+	++
	-223		+		+
	-224		+		+
Sp1 ₂	-225				
	-226		+		+
		-227	(+)		(+)
	-228		(+)		(+)
Sp1 ₃	-229		++		++
	-230		++		++
	-234				
	-236				
Sp1 ₂		-238	+		+
		-240	++		++
		-241	++		++
		-242			
Sp1 ₂	-243		+		+
		-244	++		++
		-246	++		++
		-247	+		+
Sp1 ₃	-248		+		+
	-249				
	-251				
	-252				
Sp1 ₂		-253			
	-254				
	-255				
		-256			
Sp1 ₃	-261	-259	+		+
	-263		+		+
	-264		++		++
		-265	+		+
Sp1 ₃	-266		++		++
	-267		++		++
	-268		++		++
	-269		++		++
Sp1 ₃	-270		++		++

a) An 88 bp long fragment from the SV40 21 bp repeat.

b) A double stranded synthetic oligonucleotide containing an octa-binding site.

c) ++ = complete protection; + = weak protection; (+) = very weak and variable protection.

Figure 4.

DMS methylation protection of the coding strand in binding sites Sp1:2 and Sp1:3. In lanes C 120 µg of BSA was added whereas 100 µg of a HeLa-extract was added in the other two lanes. The protected Gs in the sequence are boxed.

were caused by nuclease or methylating activities in the extracts (data not shown).

Of particular interest was the observed methylation protection of G-residues within the three Sp1 motifs (Figs. 3; Table 1). However, one G-residue, the one located in the 7th position of the consensus sequence GGGGCGGGC remained unprotected in the Sp1:1 and Sp1:3 motifs (Fig.1; Table 1). This has also been shown to be the case in the Sp1 binding sites of the SV40 21 bp repeats (24). In the Sp1:2 motif this G-residue is replaced by a T. Another difference between the Sp1:2 binding site in the U2 enhancer and the sites in the SV40 21 bp repeats is that the

Figure 5.

DMS footprint competition reactions using octa- or Sp1-specific competitors.

a) Labeled DNA fragments were incubated with 120 µg of BSA (lane C) or 100 µg of Namalva extract (lanes 1-4). No competitor DNA was added in lane 1. In lanes 2 and 4, 50 ng of an octa-specific competitor was added and in lanes 1 and 3 200 ng of the nonspecific oligonucleotide. A G, located in the octa-motif, is pointed out on the coding strand (-219), and the Gs that are located in the octa- and Sp1:1- binding sites are indicated on the noncoding strand (-218, -222 and -227).

b) Labeled DNA fragments were incubated with 120 µg of BSA (lane C) or with 100 µg of Namalva extract (lanes 1 to 4). In lanes 2 and 4, 100 ng of the Sp1-specific competitor was added to the reactions.

4th base in the motif remains unprotected from methylation in the former. In general, the relative degree of protection against methylation varied somewhat in equivalent positions in the different Sp1 binding sites as is outlined in Table 1 and Fig 1b. The DMS protection observed following incubation with nuclear extracts from HeLa cells, the B cell lines Namalva and BJA-B, and 293 cells was found to be identical (Figs. 3, and 5; data not shown).

The consensus sequence for Sp1 binding sites comprises 10 nt, centered around a C residue (Fig 1b; 25). From Figs. 3, 4 and 5 it is apparent that an additional G preceding the Sp1 motif is protected from methylation in the Sp1:2 and Sp1:3 binding sites of the human U2 enhancer, indicating an interaction with these bases. Moreover, a G that follows the 3' end of the consensus sequence shows a decreased methylation in binding site Sp1:3, extending the length of interaction to 12 nt at this binding site (Fig. 1b).

DMS footprinting of the U2 enhancer in the presence of competitor DNAs containing binding sites for Sp1 and the octa-factor

Three different unlabeled DNAs were used as competitors: (a) a double-stranded synthetic oligonucleotide, 5'-ATCTTATGCAAATCAACG-3', containing the octa-motif of the human U4C gene (3) (b) an 88 bp fragment containing four Sp1 binding sites from the SV40 21 bp repeats, (c) a double-stranded synthetic oligonucleotide with the sequence 5'-ACCGAGTCATCTCTGTGACGTCACG-3' this was used as a non-specific competitor.

The results of the competition experiments are shown in Fig. 5 and summarized in Table 1. The unspecific competitor (c) caused no significant alteration of the methylation patterns (Fig. 5a, lanes 1 and 3). Competition with DNAs containing the octa-motif (competitor a) resulted in a decreased protection of the G residues at positions -218 and -219, suggesting that these Gs are protected from methylation by the octa-factor (Figs. 5a, lanes 2 and 4; Table 1). Similarly competition with the DNA fragment containing Sp1 binding sites from SV40 (competitor b) resulted in loss of protection in the regions of the three Sp1 motifs, whereas the G residues within the octa-motif remained protected (Fig 5b, lanes 2 and 4). We therefore conclude that the three putative Sp1 binding sites indeed interact with transcription factor Sp1. The G at position -222 is protected from methylation in the presence of competitors with octa-binding sites, demonstrating that this G interacts with Sp1. No signs of mutual exclusion or interactions between Sp1 and the octa-factor were seen in these experiments.

DNase I footprinting of the U2 enhancer in the presence of competitor DNAs

DNase I footprinting experiments were carried out in the presence of the same unlabeled DNAs as were used in the DMS footprint competition experiments above. As expected, the results showed that in the presence of competitor DNAs containing either the octa-motif or the Spl binding-sites, DNase I cleavage occurs mainly within the corresponding motif in the U2 enhancer. However, an additional interaction between the factors which bind to the octa- and the Spl:1- motifs was found to occur. DNase I cleavage sites protected by the octa-factor on the coding strand became accessible to DNase I on competition with fragments containing Spl-binding sites, while cleavage sites protected by Spl binding to Spl:1 on the noncoding strand became accessible to DNase I on competition with octa-motif containing fragments (Fig. 6, lanes 2 and 3 coding strand; lane 5 noncoding strand). These results may indicate a weak interaction between the factors which bind to the octa- and Spl:1- motifs. This interaction was not observed in the DMS competition experiments probably because of different conditions used for the two assays. An important technical difference between DMS and DNase I footprinting is that the former procedure requires the cooling of the samples on ice before DMS-treatment, which might stabilize weak interactions.

DISCUSSION

Genes for U1 and U2 RNAs have been shown to contain two upstream control elements (2,4,5,7,26,27,28,29) and several of these studies have demonstrated that the more distant of the two has properties typical of enhancers. U RNA gene transcription displays several unusual features: the U RNA genes must be active in all tissues as their products are required for mRNA splicing, they must be transcribed very efficiently as snRNAs are highly abundant in the cell nucleus, transcription is independent of a TATA-box, and U1 and U2 genes are unique in that a U gene-specific enhancer/promoter appears to be required for proper transcriptional termination (7,8). One might therefore expect a complex set of proteins to interact with the U2

Figure 6.

Sp1- and octa-factor DNase I footprint competitions of the coding and noncoding strands. DNase I control digestions are shown in lanes C. In lanes 1 to 6, 50 μ g of a *Namalva* extract was added. In lanes 2 and 5, 50 ng of octa-specific competitor was added. In lanes 3 and 6, 100 ng of Sp1-specific competitor was added. The different binding sites are also indicated.

enhancer. Our results surprisingly revealed that only two types of transcription factors interact with the U2 enhancer, both of which are known to be involved in the expression of several other genes. However, it should be emphasized that our detection methods only will identify relatively strong interactions and, moreover, it seems possible that additional factors could interact with the U2 enhancer via protein-protein interactions. One of the identified factors, the octa-factor, has previously

been shown to interact with enhancers from human (9) and X.laevis (10) U2 enhancers. The present results show, in addition, that three Spl protomers interact with the human U2 enhancer region. This conclusion is based on footprint competition studies using competitor DNA fragments which contained either octa- or Spl- binding sites. Moreover, results from DMS footprinting clearly demonstrate that most of the G residues within the Spl motifs interact with this factor (Fig. 1b, Table 1). Further support comes from the observation that the upstream regions of U2 RNA genes from other species exhibit a similar arrangement of potential Spl binding sites (Fig. 1A). The rat U2 enhancer (30) appears to be organized very similarly, with an octa-motif preceded by two Spl-binding sites. The distances between the different binding sites is exactly the same as in the human U2 enhancer and the equivalent of binding site Spl:2 is also placed in an inverted orientation. In the U2 enhancers of chicken (31) and X.laevis (32) genes there is one base pair inserted between the octa-motif and the adjacent Spl-binding motif (Fig. 1a). It is noteworthy that the DNase I footprint of the X.laevis U2 enhancer (10) is considerably shorter than the footprint of the human U2 enhancer. Based on our model this is an expected finding since the X.laevis only seems to contain one Spl binding site apart from the octa-motif.

A sequence (TGTGAAAG) which closely resembles the so-called enhancer core sequence of SV40 (TGTGGAAAG) is present between -237 and -230 in the human U2 enhancer (Fig. 1a). A factor which interacts with this motif in SV40 has been identified (33). However, no binding has been demonstrated to the corresponding motif in the U2 enhancer, as the Gs at positions -234 and -236 were unprotected from methylation even when Spl-binding was prevented by including competitor DNA.

Several recent reports have demonstrated tissue-specific differences between HeLa and B cells in their repertoire of transcription factors. Differences in the DNase I footprint and DMS methylation patterns have been observed over the SV40 enhancer using extracts from different cell lines (21). In addition, DNA/protein complexes specific to lymphoid cell lines

have been described by Staudt et al. (34). Thus it was of interest to investigate the possibility of tissue-specific interactions in the human U2 enhancer, since U2 genes must be transcriptionally active in many different tissues. However, no significant differences in the DNase I or DMS footprint patterns were observed when the different extracts were tested. The B cell-derived extracts gave a slightly better protection of the octa-motif compared to 293 or HeLa cell extracts. This is likely to be due to concentration variations between the different cell lines as a similar phenomenon has been noticed by other investigators (21).

The sequences of two of the three Sp1-binding sites in the U2 enhancer deviate from the established consensus sequence (Fig. 1b) although the methylation patterns look largely typical for Sp1 interactions. The binding sites Sp1:1 and Sp1:3 show the same lack of protection of the second G following the central C in the Sp1 consensus as has been seen in the Sp1-binding sites of the SV40 21 bp repeats (24). This lack of protection suggests that the protein is not in direct contact with this residue, inferring that this base might be interchangeable. Indeed there is a T at this position in the Sp1:2 binding sites of both the human and rat U2 enhancers (Figs. 1a and b). There is a significant difference in the degree of methylation of the G that precedes the middle C in binding sites Sp1:2 and Sp1:3. This G is protected in binding site Sp1:3, as it is in the binding sites of the 21 bp repeats of SV 40, whereas it is unprotected in binding site Sp1:2. Binding site Sp1:1 shows an intermediate behavior. The results suggest that the protein/DNA interactions vary, probably as a result of differences between the recognition sequences.

Our methylation protection experiments suggest that Sp1 is in contact with several bases inside the Sp1 motifs (Fig 1b). It is noteworthy that some of the protected Gs lie outside the borders of the decanucleotide Sp1 consensus sequence (Fig. 1b). This finding was surprising as it implies that one base is in contact with two Sp1 protomers binding to sites I and II and to sites II and III of the SV40 21 bp repeat. Moreover, it infers that binding site Sp1:1 in the U2 enhancer shares one base pair

with the octa-motif (Fig. 1a). Gidoni et al. (24) have demonstrated that mutual exclusion occurs only between binding sites IV and V in the SV40 21 bp repeats where the decanucleotide sequences happen to overlap. No such relationship appears to exist between the other Sp1-binding sites of SV40 or between the octa-factor and the Sp1:1 binding site of the U2 enhancer. One possible explanation for the extended protection of binding sites Sp1:2 and Sp1:3 is that the sequences of the recognition sites influence the binding properties of the factor, thereby facilitating an interaction with flanking residues. Another possibility is that highly charged amino-acids may occur in close enough proximity to result in a decreased methylation by DMS although the protein is not in direct contact with the G.

The methylation protection studies show that Sp1 interacts with the major groove of the DNA template, although interactions with the minor groove cannot be excluded. The Sp1 protein binds to the same sides of the DNA helix, albeit in opposite orientation, at binding sites Sp1:2 and Sp1:3, separated by two helical turns. Binding sites Sp1:1 and Sp1:2, on the other hand, are separated by 16 bp. Thus the binding to the Sp1:1 site occurs on the other side of the helix relative to sites Sp1:2 and Sp1:3.

Our results also show that the octa-factor protects the Gs at -218 and -219 in the major groove. The octa-factor thus interacts with the template on the same side as Sp1 in two adjacent major grooves, with the recognition domains brought into very close proximity. However, no cooperativity in the binding to the octa- and Sp1:1- sites was observed in the DMS footprint experiments using competitor DNAs which prevented either of the two factors from binding to their targets. Using the DNase I footprint method a slightly different result was obtained; after competition with a DNA fragment containing the octa-motif not only was octa-binding prevented, but partial deprotection was also seen in the region of the Sp1:1 binding site (Fig. 6). This observation was, however, only seen on the noncoding strand. On the coding strand a deprotection of the octa-motif was seen after incubation with a Sp1-specific competitor. Although reproducible the results are difficult to interpret as

the deprotection was only detected on one strand in each case.

Several independent groups have shown that the human U2 5'flanking region located upstream of position -198 has enhancer activity both in HeLa cells (6) and in X.laevis oocytes (2,4,29). A mutant U2 template in which the octa-motif was destroyed has been shown to be transcriptionally inactive in X.laevis oocytes as well as in HeLa cells, indicating that the Sp1 motifs by themselves cannot confer activity, as is also the case for the SV40 promoter (35). The results reported by Westin et al. (2,29) show that all three Sp1 binding sites are not required for transcriptional activity as templates which are terminated at position -258 (thus including only two Sp1-binding sites) are fully active. This truncation results in a correspondingly shorter DNase I footprint (data not shown). However, there is evidence that Sp1 interactions are required for full transcriptional activity. For example, the results of Mattaj et al. (5) revealed a decreased level of transcription when the enhancer of a X.laevis U2 RNA gene was replaced by a synthetic oligonucleotide containing only an octa-binding site. Furthermore, Westin et al. (29) have observed that transcription of a human U2 RNA gene is impaired when the template lacks sequences upstream of position -226. Thus there is strong evidence that the Sp1 interactions are required for a fully active enhancer, as is also the case with the SV40 enhancer/promoter (24,35). From the present results and the sequence organization of the enhancer regions of U2 RNA genes in other species (Fig. 1a), it appears that a combination of sites to which an octa-factor and one or more Sp1 protomers bind is a common theme.

The existence of multiple binding sites for Sp1 in housekeeping genes expressed in all tissues has previously been discussed (36). A gene for a noncoding RNA can now be added to that list. The biological significance of the Sp1 interactions and the functions of Sp1 and the octa-factor are not known. Due to their widespread occurrence, Sp1 and the octa-factor are likely to be general transcription factors present in all tissues. Both factors have been shown to activate transcription in vitro when added to suboptimal transcription reactions

(10,23). One possible interpretation of these results is that Spl plays a structural role in stabilizing the transcription complex while the octa-factor might be the actual transcriptional activator. When odd and even numbers of helical turns were inserted between the 21 bp and 72 bp repeats of SV40, a dramatic effect on transcription was recorded (37), suggesting a steric relationship between the enhancer and the Spl-binding sites.

The lack of unique factors binding to the U2 enhancer may suggest that the unique features of U RNA transcription is the result of interactions with the proximal transcriptional control element rather than with the enhancer.

ACKNOWLEDGEMENTS

We are grateful to Elisabeth Sandberg and Katarina Eriksson for typing the manuscript, and to Göran Akusjärvi, Polly Weller and Jan Zabielski for helpful advice and criticism. This study was made possible by grants from the Swedish Board for Technical Development, the Swedish Medical Research Council and the Marcus Borgström Foundation. L.J. is a fellow of the Swedish Cancer Society.

*To whom correspondence should be addressed

REFERENCES

1. Murphy, J.T., Burgess, R.R., Dahlberg, J.E. and Lund, E. (1982) *Cell* 29, 265-274
2. Westin, G., Lund, E., Murphy, J.T., Pettersson, U. and Dahlberg, J.E. (1984) *EMBO J.* 3, 3295-3301
3. Bark, C., Weller, P., Zabielski, J. and Pettersson, U. (1986) *Gene* 50, 330-344
4. Ares, M.Jr., Mangin, M. and Weiner, A.M. (1985) *Mol.Cell.Biol.* 5, 1560-1570
5. Mattaj, I.W., Lienhard, S., Jiricny, J. and DeRobertis, E.M. (1985) *Nature* 316, 163-167
6. Mangin, M., Ares, M.Jr. and Weiner, A.M. (1986) *EMBO J.* 5, 987-995
7. Hernandez, N. and Weiner, A.L. (1986) *Cell* 47, 249-258
8. Neuman de Vegvar, H.E., Lund, E. and Dahlberg, J.E. (1986) *Cell* 47, 259-266
9. Sive, H.L. and Roeder, R.G. (1986) *Proc.Natl.Acad.Sci. USA* 83, 6382-6386
10. Bohmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G. and Mattaj, I.W. (1987) *Nature* 325, 268-272

11. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) *Nature* 319, 154-158
12. Pruijin, G.J.M., van Driel, W. and van der Vliet, P.C. (1986) *Nature* 322, 656-659
13. Falkner, F.G. and Zachau, H.G. (1984) *Nature* 310, 71-74
14. Parslow, T.G., Blaire, D.L., Murphy, W.J. and Granner, D.K. (1984) *Proc.Natl.Acad.Sci. USA* 81, 2650-2654
15. Ohlsson, H. and Edlund, T. (1986) *Cell* 45, 35-44
16. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory)
17. Maxam, A.M. and Gilbert, W. (1980) *Meth.Enzymol.* 65, 499-560
18. Galas, D. and Schmitz, A. (1978) *Nucl.Acids Res.* 5, 3157-3170
19. Klein, G., Dombos, L. and Gothoskar, B. (1972) *Inst. J. Cancer* 10, 44-57
20. Klein, G., Giovanella, B., Westman, A., Stehlin, J.S. and Mimford, D. (1975) *Intervirology* 5, 319-334
21. Davidson, I., Fromental, C., Augereau, P., Wildeman, A., Zenke, M. and Chambon, P. (1986) *Nature* 323, 544-548
22. Graham, F.L., Abrahams, P.S., Mulder, C., Heijnker, H.L., Warnaar, S.O., deVries F.A.J., Fiers, W. and van der Eb, A.J. (1975) *Cold Spring Harbor Symp. Quant. Biol.* 39: 637
23. Dynan, W.S. and Tjian, R. (1983) *Cell* 32, 669-680
24. Gidoni, D., Kadonaga, J.T., Barrere-Saldaña, H., Tajahashi, K., Chambon, P. and Tjian, R. (1985) *Science* 230, 511-517
25. Brigges, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) *Science* 234, 47-52
26. 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
27. Ciliberto, G., Buckland, R., Cortese, R. and Philipson, L. (1985) *EMBO J.* 4, 1537-1543
28. Krol, A., Lund, E. and Dahlberg, J.E. (1985) *EMBO J.* 4, 1529-1535
29. Westin, G., Zabielski, J., Janson, L. and Pettersson, U. (1987) submitted
30. Tani, T., Watanabe-Nagasu, N., Okada, N. and Ohshima, Y. (1983) *Mol.Biol.* 168, 579-594
31. Korf, G.M. and Stumph, W.E. (1986) *J.Biol.Chem.* 25, 2041-2047
32. Mattaj, I.W. and Zeller, R. (1983) *EMBO J.* 2, 1883-1891
33. Wildeman, A.G., Zenke, M., Schatz, C., Wintzerith, M., Grundström, T., Matthes, H., Takahashi, K. and Chambon, P. (1986) *Mol.Cell.Biol.* 6, 2098-2105
34. Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A. and Baltimore, D. (1986) *Nature* 323, 640-643
35. Benoist, C. and Chambon, P. (1981) *Nature* 290, 304-310
36. Dynan, W.S. (1986b) *Trends Genet.* 8, 196-197
37. Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. and Chambon, P. (1986) *Nature* 319, 121-126