The EMBO Journal Vol.16 No.1 pp.173–181, 1997

Staf, a promiscuous activator for enhanced transcription by RNA polymerases II and III

Myriam Schaub, Evelyne Myslinski, Catherine Schuster, Alain Krol and Philippe Carbon[1]

UPR 9002 du CNRS 'Structure des Macromolécules Biologiques et Mécanismes de Reconnaissance', IBMC, 15 rue René Descartes, 67084 Strasbourg Cedex, France

[1]Corresponding author

Staf is a zinc finger protein that we recently identified as the transcriptional activator of the RNA polymerase III-transcribed selenocysteine tRNA gene. In this work we demonstrate that enhanced transcription of the majority of vertebrate snRNA and snRNA-type genes, transcribed by RNA polymerases II and III, also requires Staf. DNA binding assays and microinjection of mutant genes into *Xenopus* oocytes showed the presence of Staf-responsive elements in the genes for human U4C, U6, Y4 and 7SK, *Xenopus* U1b1, U2, U5 and MRP and mouse U6 RNAs. Using recombinant Staf, we established that it mediates the activating properties of Staf-responsive elements on RNA polymerase II and III snRNA promoters *in vivo*. Lastly a 19 bp consensus sequence for the Staf binding site, YY(A/T)CCC(A/G)N(A/C)AT(G/C)C(A/C)YY-RCR, was derived by binding site selection. It enabled us to identify 23 other snRNA and snRNA-type genes carrying potential Staf binding sites. Altogether, our results emphasize the prime importance of Staf as a novel activator for enhanced transcription of snRNA and snRNA-type genes.

Keywords: RNA polymerase II/RNA polymerase III/ snRNA genes/transcriptional activator/zinc finger

Introduction

Genes for vertebrate small nuclear RNAs (snRNAs) are transcribed by either RNA polymerase II (Pol II) or RNA polymerase III (Pol III), depending on the type of promoters they harbor. The basal promoters of both types include an essential proximal sequence element (PSE) located at approximately –59 upstream of the transcription start site. The Pol III-dependent genes also possess a TATA box at –30 which acts as a major determinant of RNA Pol III specificity (Lobo and Hernandez, 1989; Mattaj *et al.*, 1988; see Hernandez, 1992 for a review). A number of other short transcription units, such as the 7SK RNA, Y RNA, MRP RNA and H1 RNA genes have similar basal promoter elements and can be classified as snRNA-type genes. snRNA and snRNA-type genes contain, in addition to the *cis* elements described above, a distal sequence element (DSE). The DSE plays a major role in transcription efficiency, accounting for a 5- to 100-fold level of activation of Pol II or Pol III basal transcription in transfected cells or injected *Xenopus* oocytes.

Numerous Pol II and Pol III DSEs have been dissected and found to be composed of an octamer motif and another, usually close, element (for a review see Hernandez, 1992). Among the latter Sp1 binding sites in the human U2, *Xenopus* U2 and U6 genes (Ares *et al.*, 1987; Janson *et al.*, 1987; Tebb and Mattaj, 1989; Lescure *et al.*, 1992), an AP-2 binding site and a CRE motif in the human U4C gene (Weller *et al.*, 1988), SPH motifs in the chicken U1 and U4B genes (Roebuck *et al.*, 1990; Zamrod and Stumph, 1990; Cheung *et al.*, 1993), CAAT motifs in human and *Xenopus* U3 genes (Ach and Weiner, 1991; Savino *et al.*, 1992), a NONOCT motif in the human U6 gene (Danzeiser *et al.*, 1993), a D2 motif in the *Xenopus* U2 gene (Tebb and Mattaj, 1989) and a CACCC box and octamer-like motifs in the human 7SK gene (Murphy *et al.*, 1989, 1992; Kleinert *et al.*, 1990; Boyd *et al.*, 1995) have been identified.

The Sp1 and octamer motifs contain the recognition sites on the DNA for the well-characterized transcriptional activators Sp1 and Oct-1 respectively (Courey and Tjian, 1988; Sturm *et al.*, 1988; for reviews see Herr, 1992; Hernandez, 1992). However, transcription factors interacting with the other elements described above have not been purified to homogeneity or cloned. Furthermore, owing to the occurrence of octamer or octamer-like sequences in a number of DSEs, it has been tacitly admitted that the activation function of the DSE is mediated essentially by Oct-1 binding at the octamer motif.

The basal promoter of the atypical selenocysteine tRNA gene is principally external to the coding region and comprises a PSE and a TATA motif functionally equivalent to those of vertebrate U6 snRNA genes (Carbon and Krol, 1991; Myslinski *et al.*, 1993a). Additionally, its basal promoter is activated by the activator element (AE), an element which functions without assistance of the octamer (Myslinski *et al.*, 1992, 1993b). Instead, the activation properties of the AE are mediated by Staf, a sequence-specific zinc finger protein that we recently characterized (Schuster *et al.*, 1995). Experimental evidence provided in this work shows that Staf is also involved in transcriptional activation of a large variety of snRNA and snRNA-type genes transcribed by RNA Pol II and Pol III. Our results indicate that AP-2, D2, NONOCT, octamer-like and SPH motifs previously described as being involved in transcriptional activation of a number of these genes are in fact Staf-responsive elements. Staf is thus a key factor for transcriptional activation of snRNA and snRNA-type genes by RNA Pol II and Pol III.

Results

Staf binds specifically to the majority of snRNA and snRNA-type genes

To determine whether Staf is involved in transcriptional activation of snRNA and snRNA-type genes, gel retard-

Table I. snRNA and snRNA-type genes in this study

Gene	Abbreviation	Transcribed by	Reference
Human U1	hU1	Pol II	Lund and Dalhberg (1984)
X.laevis U1b1	xU1b1	Pol II	Krol *et al*. (1985)
X.laevis U1b2	xU1b2	Pol II	Krol *et al*. (1985)
X.laevis U2	xU2	Pol II	Mattaj and Zeller (1983)
X.laevis U3A	xU3A	Pol II	Savino *et al*. (1992)
Human U4C	hU4C	Pol II	Bark *et al*. (1986)
X.laevis U5	xU5	Pol II	Kazmaier *et al*. (1987)
Human U6	hU6	Pol III	Kunkel *et al*. (1986)
Mouse U6	mU6	Pol III	Oshima *et al*. (1981)
X.tropicalis U6	xU6	Pol III	Krol *et al*. (1987)
Human H1 RNA	hH1 RNA	Pol III	Baer *et al*. (1990)
Human Y4	hY4	Pol III	Maraia *et al*. (1994)
Human 7SK	h7SK	Pol III	Murphy *et al*. (1986)
X.laevis MRP RNA	xMRP RNA	Pol III	Bennett *et al*. (1992)

ation assays were used in the first place to examine the ability of Staf to bind the DSEs arising from 14 genes transcribed by RNA Pol II and Pol III (see Table I). Labeled DNA fragments encompassing the various DSEs (see Materials and methods) were incubated with the purified Staf DNA binding domain and then analyzed on non-denaturing polyacrylamide gels. Figure 1 shows that Staf bound to the majority of the 14 DSEs tested. A high yield of binding was detected with the DSEs of xU1b1 (Figure 1, lane 6), xU2 (lane 12), hU4C (lane 20), xU5 (lane 24), hU6 (lane 28), mU6 (lane 32), hY4 (lane 42), h7SK (lane 46) and xMRP RNA (lane 50). The intensities of the retarded complexes observed in these lanes are comparable with that obtained with the $tRNA^{Sec}$ gene (Figure 1, lane 54). In contrast, a very low binding was observed with hU1 (lane 2) and hH1 RNA (lane 38). Lastly, no gel shift at all could be obtained with xU1b2 (lane 10), xU3A (lane 16) and xU6 (lane 36) DSEs. To demonstrate that these retarded complexes were caused by the specific binding of Staf, gel retardation assays were performed in the presence of an excess of two different double-stranded oligodeoxynucleotides acting as competitors. The first contains the AE of the *Xenopus laevis* $tRNA^{Sec}$ gene, which is specifically recognized by Staf, the other carries a mutant AE unable to bind Staf (Schuster *et al.*, 1995). Band shifts were abolished in the presence of the specific competitor (Figure 1, lanes 3, 7, 13, 21, 25, 29, 33, 39, 43, 47, 51 and 55) but unaltered when the mutant AE was used instead (lanes 4, 8, 14, 22, 26, 30, 34, 40, 44, 48, 52 and 56). These results are consistent with a specific binding of Staf to xU1b1, xU2, hU4C, xU5, hU6, mU6, hY4, h7SK and xMRP RNA DSEs.

To localize the Staf binding sites, DNase I footprint analysis was carried out with labeled DNA probes harboring the various DSEs. Those DSEs binding Staf with high yield produced a clear footprint over at least 21 bp (Figure 2A). The protected regions are shown schematically in Figure 2B, together with that obtained on the AE of the $tRNA^{Sec}$ gene (Schuster *et al.*, 1995). Sequence comparisons between the various binding sites revealed homologous sequences, on one strand or the other, allowing derivation of a 20 bp consensus sequence for the Staf binding site, YYTCCCANNRTNCNNYGCRR (Figure 2B).

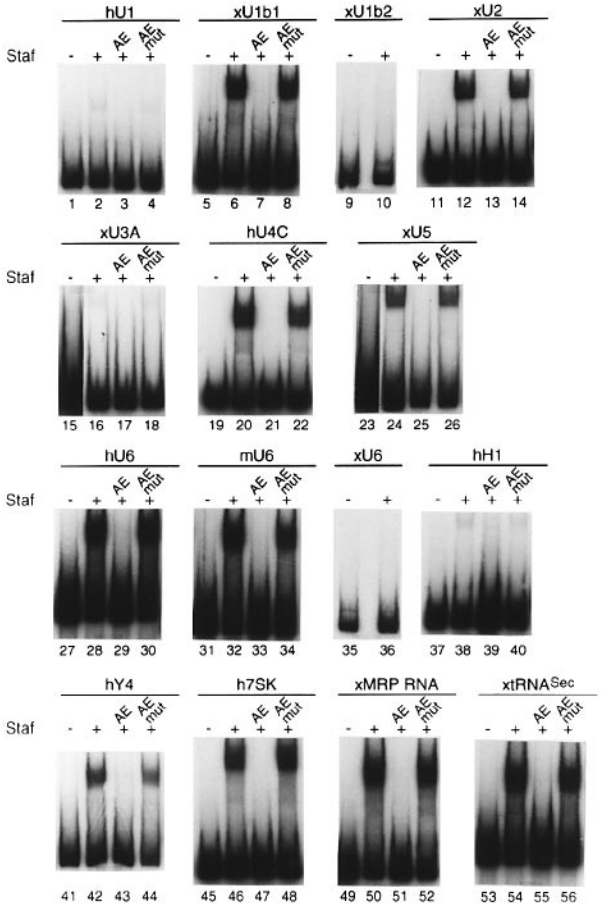

Fig. 1. Staf bound specifically to a majority of DSEs from RNA Pol II and RNA Pol III snRNA and snRNA-type genes. Gel retardation assays with the DSE of 14 snRNA genes or related genes. ^{32}P-End-labeled DNA probes were incubated in the absence (lanes 1, 5, 9, 11, 15, 19, 23, 27, 31, 35, 37, 41, 45, 49 and 53) or presence (other lanes) of the Staf DNA binding domain. Reactions in lanes 3, 7, 13, 17, 21, 25, 29, 33, 39, 43, 47, 51 and 55 contained a 1000-fold molar excess of unlabeled AE competitor DNA. A 1000-fold molar excess of non-specific DNA (AE mut) was added in lanes 4, 8, 14, 18, 22, 26, 30, 34, 40, 44, 48, 52 and 56.

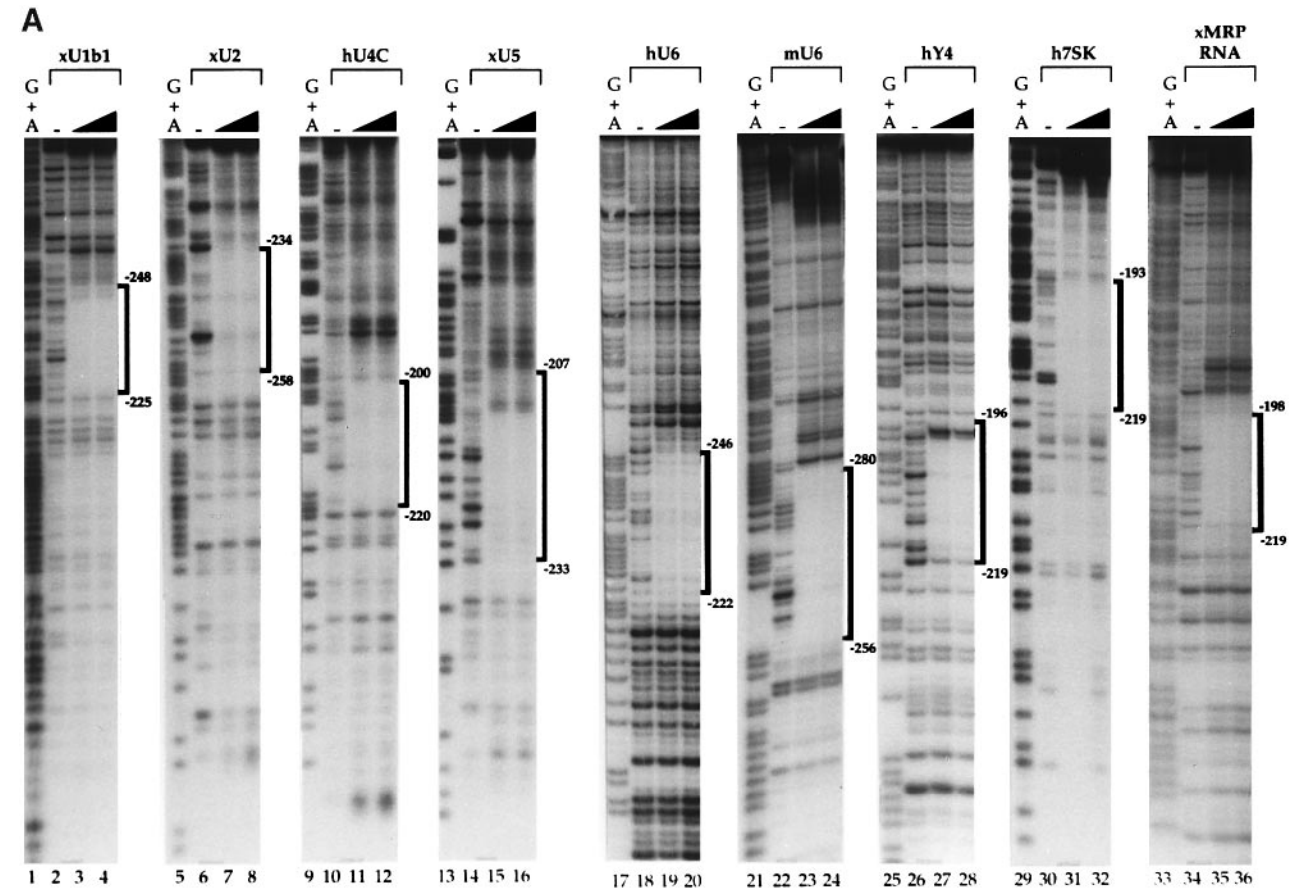

Fig. 2. Identification of Staf binding sites in the DSEs of xU1b1, xU2, hU4C, xU5, hU6, mU6, hY4, h7SK and xMRP RNA genes. (**A**) Footprint analysis of Staf–DNA complexes. DNase I digestion of the xU1b1, xU2, hU4C, xU5, hU6, mU6, hY4, h7SK and xMRP RNA probes in the absence (lanes 2, 6, 10, 14, 18, 22, 26, 30 and 34) or presence of 3 (lanes 3, 7, 11, 15, 19, 23, 27, 31 and 35) or 4 µl (lanes 4, 8, 12, 16, 20, 24, 28, 32 and 36) of Staf DNA binding domain. G+A chemical cleavage is indicated above the lanes. (**B**) Sequence comparisons of the different footprinted Staf DNA binding sites. DNase I-protected regions are diagramed above each sequence by a solid bar. For xU2, hU6 and h7SK, the reported protected regions correspond to the footprint obtained on the opposite strand. The sequence protected by Staf in the *X.laevis* $tRNA^{Sec}$ gene (Schuster *et al*., 1995) is added at the bottom of the figure and the activator element of this gene (Myslinski *et al*., 1992) is indicated in white on a dark background. The numbers on the right and left sides indicate the distance from the transcription initiation site. Residues conserved in at least seven of the 10 mapped Staf binding sites are indicated on a shaded background. The 20 bp consensus sequence is derived at the bottom of the figure. N, R and Y stand for any nucleotide, purine and pyrimidine respectively. Pol II and Pol III, genes transcribed by RNA polymerases II and III.

Functional relevance of the mapped Staf binding sites

We next analyzed the functional relevance of the mapped binding sites by: (i) creating substitution mutants either unable or showing severely reduced abilities to bind Staf (Figure 3A); (ii) assaying their transcription abilities by injection into *Xenopus* oocyte nuclei (Figure 3B). The substitution mutants changed positions 4–7 of the consensus sequence. The conserved CCCA (positions 4–7) in the xU1b1, hU4C, xU5, hU6, mU6, hY4 and xMRP RNA genes was substituted by AAAC. In xU2 and h7SK, CCCG and TCCA (at the same positions) were substituted

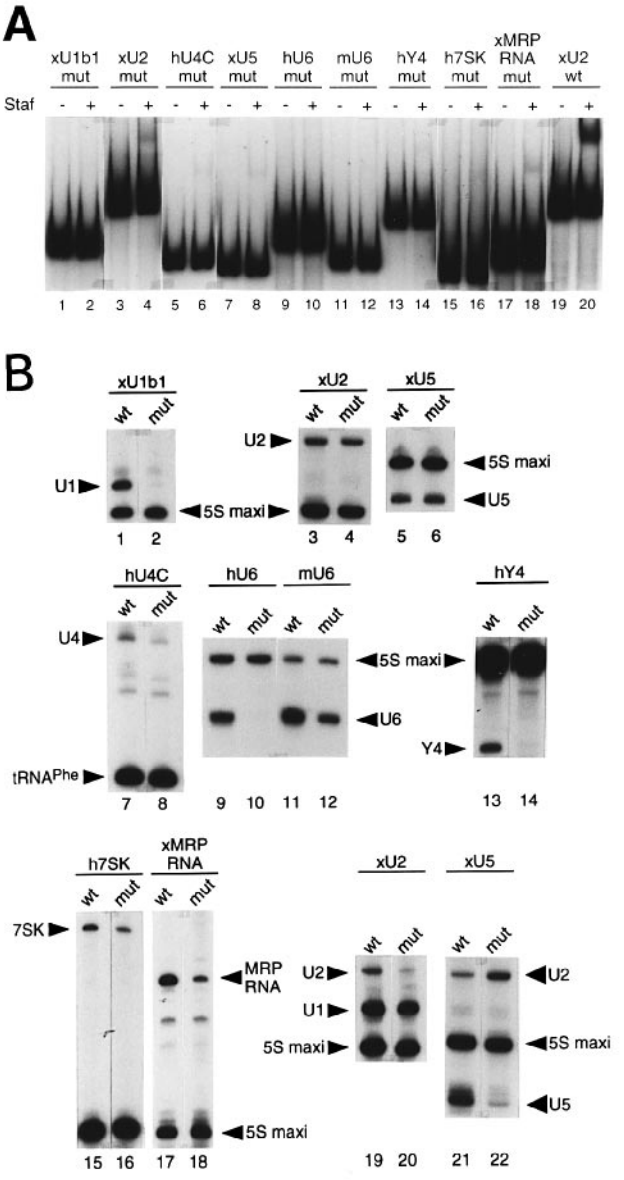

Fig. 3. Binding and template activities of mutant Staf binding sites. (**A**) Gel retardation assays with the mutant Staf binding sites. ^{32}P-End-labeled DNA probes were incubated in the absence (lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17) or presence (other lanes) of the Staf DNA binding domain. Wild-type xU2 labeled DNA probe (lanes 19, 20) was included as a positive control. (**B**) *Xenopus laevis* oocyte nuclei injected with the wild-type (lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17) or mutant (lanes 2, 4, 6, 8, 10, 12, 14, 16 and 18) xU1b1, xU2, xU5, hU4C, hU6, mU6, hY4, h7SK and xMRP RNA templates. Wild-type (lane 19) or mutant (lane 20) xU2 were co-injected with an equal amount of wild-type xU1b1. Lanes 21 and 22, wild-type or mutant xU5 respectively co-injected with an equal amount of wild-type xU2. Positions of the 5S maxi, tRNA^{Phe} and transcription products are indicated.

by AAAT and GAAC respectively. In this injection assay, the transcription activities of seven of the nine mutants dropped considerably (Figure 3B). Normalized residual values, expressed relative to the corresponding wild-type levels, ranged between 2% (hU6 and hY4, lanes 9 and 10, and 13 and 14 respectively), 5% (xU1b1, lanes 1 and 2), 15% (xMRP RNA, lanes 17 and 18), 30% (h7SK, lanes 15 and 16) and 40% (hU4C and mU6, lanes 7 and 8, and 11 and 12 respectively). Since the xU2 and xU5 mutants retained wild-type activity (lanes 3 and 4, and 5 and 6 respectively), they were then tested in a more stringent assay in which the mutant template was co-injected with a competitor gene whose transcription is driven by its wild-type promoter. Here the competitors employed were the wild-type xU1b1 and xU2 genes. The transcription activity of the mutant xU2 was then assessed by competition with wild-type xU1b1, that of the mutant xU5 gene by wild-type xU2. Competitive conditions exacerbated the effects of the mutations, which provoked a marked drop in transcription efficiency to 20% of the wild-type level for xU2 and xU5 (Figure 3B, lanes 19 and 20 and 21 and 22 respectively). These results show that the nine Staf binding sites characterized are functionally important to enhanced Pol II and Pol III transcription of these snRNA or related genes. They will be further referred to as Staf-responsive elements.

Transactivating properties of Staf on Pol II and Pol III snRNA promoters

In order to show that Staf is actually responsible for this activation function, we used the *X.laevis* oocyte expression assay previously developed to establish that Staf mediated transcriptional activation of the tRNA^{Sec} gene (Schuster *et al.*, 1995). In this assay, the endogenous Staf background of the oocyte, which would interfere in the experiment, was eliminated by replacing the Staf DNA binding domain with that of Krox-20 (Krox-20 DBD; Figure 4A). The transcription ability of this chimeric protein, termed Staf–Krox-20, was assayed with wild-type *Xenopus* Pol II U1b2 (Krol *et al.*, 1985) and *Xenopus* Pol III U6 (Krol *et al.*, 1987) reporters (Figure 4B) and mutant versions thereof lacking the DSE (U1·ΔDSE and U6·ΔDSE) or containing instead the Krox-20 binding site E element (U1·3E and U6·3E). The mRNAs of the effectors Staf–Krox-20 and Krox-20 DBD were transcribed *in vitro*, capped and injected separately into oocyte cytoplasm (Schuster *et al.*, 1995). After 20 h incubation, the various U1 and U6 reporters were injected into oocyte nuclei, along with [α-^{32}P]GTP. After a second incubation, labeled RNAs were extracted, the levels of which measure the transactivation properties of the protein tested. In the presence of Staf–Krox-20, the transcription levels of U1wt, U1·ΔDSE, U6wt and U6·ΔDSE (Figure 4C, lanes 7, 8, 16 and 17 respectively) were identical to those observed in the absence of effector (lanes 1, 2, 10 and 11 respectively) or in the presence of Krox-20 DBD only (lanes 4, 5, 13 and 14 respectively). Remarkably, however, comparison of lanes 3 with 9 and 12 with 18 revealed that Staf–Krox-20 could significantly stimulate transcription of U1·3E and U6·3E. Of note, transcription levels varied from 10 (U1·ΔDSE) to 50% (U1·3E) and 0 (U6·ΔDSE) to 10% (U6·3E) of the corresponding wild-type promoter level. Transactivation was not mediated by Krox-20 DBD, since transcription of U1·3E and U6·3E was unaffected by its presence (lanes 6 and 15). These results demonstrate unambiguously the transactivating properties of Staf on Pol II and Pol III snRNA promoters.

Selection of DNA binding sites for Staf

To extend our knowledge of the Staf DNA binding sites, we employed the technique of PCR-mediated amplification of protein-selected random oligonucleotides (Blackwell

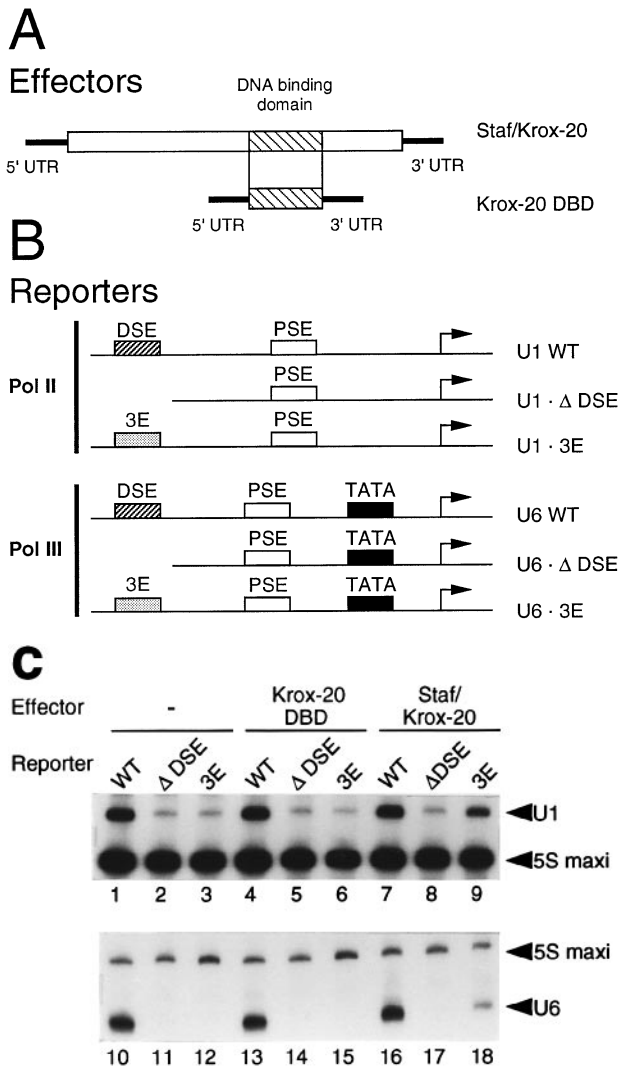

Fig. 4. Staf stimulates transcription from RNA Pol II and III snRNA promoters in *X.laevis* oocytes. (**A**) Schematic organization of the effector mRNAs synthesized *in vitro*. 5′ UTR and 3′ UTR are the corresponding untranslated regions of the *X.laevis* β-globin mRNA. (**B**) Schematic organization of the *Xenopus* U1b2 and U6 snRNA reporter genes. PSE and TATA represent the basal promoter elements of the U1 and U6 genes, DSE their distal sequence elements. 3E indicates three E binding sites of the Krox-20 protein substituting for the wild-type DSE. The bent arrow indicates the start of transcription. (**C**) Enhanced transcription of U1 and U6 snRNAs. Positions of the 5S maxi, U1 and U6 are indicated.

et al., 1990; Chittenden *et al.*, 1991; Delwel *et al.*, 1993). To this end, a chimeric protein was used which consisted of glutathione S-transferase fused to residues 257–475 of the Staf DNA binding domain. The fusion protein was purified by affinity binding to glutathione–Sepharose and the Sepharose-bound protein was used for binding and amplification reactions with a 57 bp oligonucleotide duplex that contained a core of 17 random nucleotides. Seventy three clones chosen from the final pool of selected DNAs were sequenced. Of the 22 positions tabulated, 18 positions (1–7 and 9–19) displayed a significantly higher degree of constraint with respect to base preference (Figure 5). Eleven out of 17 display strong secondary preferences (positions 1–3, 7, 9, 12, 14–17 and 19) when the base of first preference is lacking. The 19 bp consensus sequence thus derived is YY(A/T)CCC(A/G)N(A/C)AT(G/C)C(A/C)YYRCR (Figure 5). Within the consensus, position 8 is degenerate and positions 4–7, 10, 11 and 13 are more highly constrained than bases at positions 1–3, 8, 9, 12 and 14–18. Positions 9, 12, 14 and 15, considered as fully degenerate in the first consensus derived from sequence comparisons of the different footprints (Figure 2B), in fact match the consensus derived from binding site selection. From the selection data it is obvious that position 20 is fully degenerate and not occupied by R, as deduced from Figure 2B.

Twenty three genes with potential Staf binding sites

Lastly, in addition to the 14 genes tested above, we have used the consensus binding site of Figure 5 to search for the presence of potential Staf binding sites in the other 34 vertebrate snRNA and snRNA-type genes found in the database (Gu and Reddy, 1996). Sequences with a high match (at least 14 out of 19) to the Staf consensus sequence occur in 23 Pol II or Pol III genes (Figure 6), residing between –245 and –185, similarly to the positions for the sites characterized experimentally (Figure 2B). In the light of these findings, we consider that the additional 23 sequences also constitute Staf binding sites. Together with the 10 genes for which we provided experimental evidence, our data strongly suggest that Staf is involved in transcriptional activation of at least 70% of the Pol II and Pol III snRNA and snRNA-type genes available up to now in the databases.

Discussion

Staf is a zinc finger protein that was recently identified as the transcriptional activator of the Pol III selenocysteine tRNA gene (Schuster *et al.*, 1995). In the present work, we have demonstrated that enhanced transcription activity provided by Staf is not devoted to the selenocysteine tRNA promoter alone. We have presented several lines of evidence strongly suggesting that Staf is also involved in transcriptional activation of at least 70% of vertebrate snRNA and snRNA-type genes transcribed by RNA Pol II and Pol III. These include the chicken U1 52A and U4B, human U4C, U6, Y4 and 7SK and *X.laevis* U2 and MRP RNA genes, for which various motifs have been attributed a function by others based on sequences which we have here demonstrated to represent in fact Staf binding sites (Figure 7). In chicken U1 52A and U4B, Staf binding sites match perfectly the SPH motifs previously demonstrated to be important for maximal expression of these genes (Roebuck *et al.*, 1990; Zamrod and Stumph, 1990). Thus, it is highly likely that Staf is the *Xenopus* equivalent of the partially purified chicken SBP protein. In the cases of the human U4C, *Xenopus* U2, human U6, 7SK and Y4 and *Xenopus* MRP RNA genes, AP-2, D2, NONOCT and octamer-like motifs have been attributed a function by others (Weller *et al.*, 1988; Tebb and Mattaj, 1989; Murphy *et al.*, 1989, 1992; Bennett *et al.*, 1992; Danzeiser *et al.*, 1993; Maraia *et al.*, 1994; Boyd *et al.*, 1995). In contrast, our data clearly demonstrate that a Staf-responsive element overlaps these motifs (Figure 7).

The high affinity Staf binding site generated by *in vitro* selection is a 19 bp consensus sequence which tolerates a high degree of degeneracy in 12 out of 19 positions

Position	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	+1
A%	27.3	11.9	6.8	55.0	1.6	0.0	0.0	83.9	3.2	54.8	91.8	1.4	0.0	1.4	74.6	23.4	1.9	29.3	3.2	23.1	21.4	7.7
C%	15.2	23.8	20.5	3.3	95.2	100.0	100.0	0.0	38.7	35.6	1.4	0.0	19.2	97.2	15.5	34.4	46.3	2.4	77.4	7.7	35.7	38.5
G%	24.2	11.9	6.8	10.0	0.0	0.0	0.0	14.5	30.7	9.6	6.8	1.4	75.3	1.4	2.9	0.0	3.7	58.5	9.7	61.5	14.3	15.4
T%	33.3	52.4	65.9	31.7	3.2	0.0	0.0	1.6	27.4	0.0	0.0	97.2	5.5	0.0	7.0	42.2	48.1	9.8	9.7	7.7	28.6	38.4
Number of sequences	33	42	44	60	62	62	62	62	62	73	73	73	73	73	71	64	54	41	31	26	14	13
Consensus	N	Y	Y	A/t	C	C	C	A/g	N	A/C	A	T	G/c	C	A/c	Y	Y	R	C	R	N	N

Fig. 5. Derivation of the Staf consensus binding site. Analysis of the Staf binding sequences occurring within affinity-selected oligonucleotides. The frequencies with which the four bases A, C, G and T were selected at positions 1–20 (numbering as in Figure 2B) are listed. Positions flanking the 20 bp oligonucleotide were also subjected to selection, labeled –1 at the 5′-end and +1 at the 3′-end. The consensus nucleotide(s) for each position appears below Number of sequences, with lower case letters indicating bases selected less frequently.

(Figure 5). Such a particularly extended binding site may explain the ability of Staf-responsive elements to accept the substantial number of base changes that occur in the different genes tested, without altering the binding of Staf. This is well illustrated by the example of the Staf-responsive elements in the human U6 and Y4 genes, which lack the 3′-part of the consensus Staf binding site (positions RCR in Figure 7) and yet are recognized efficiently by Staf.

In previous reports, we have shown that Staf possesses the capacity to stimulate CAT expression from a Pol II promoter (Myslinski *et al.*, 1992; Schuster *et al.*, 1995). Therefore, our data collectively demonstrate the particular ability of Staf to activate both snRNA-type and mRNA promoters and thus the whole diversity of Pol II and Pol III promoters. Comparison between Staf and its human homolog ZNF 76 revealed the presence, in addition to the central zinc finger domain, of six conserved motifs (Schuster *et al.*, 1995). We hypothesize that some of these conserved motifs represent promoter-selective activation domains directing the differential activation of snRNA and mRNA promoters. This is currently under investigation.

Although the octamer sequence has been recognized for quite some time as a universal motif in the DSEs of vertebrate snRNA and snRNA-type genes, one major finding of our work is the high prevalence of Staf-responsive elements in the DSEs of these genes. About 70% of the DSEs contain both an octamer motif and a Staf binding site associated or not with a third element. The other DSEs contain either octamer or Staf motifs with or without a second element, depending on the DSE. For example, optimal transcription of the *Xenopus* and human U2 genes is dependent on the three octamer, Staf and Sp1 motifs (Ares *et al.*, 1987; Tebb and Mattaj, 1989; this work). On the other hand, transactivation of the *X.laevis* selenocysteine tRNA promoter, and probably that of the human Y4 and *X.laevis* MRP RNA genes, is dependent on a Staf motif only (Myslinski *et al.*, 1992, 1993b; this work). What might be the reason for the variability in the identity and number of motifs constituting the DSE? The answer(s) may reside in the arrangement and strength of the basal promoter elements, which are known to exert a marked effect on motif composition of the DSE and transcriptional activator function (Myslinski *et al.*, 1993b; Das *et al.*, 1995).

The combined presence of the octamer and Staf motifs in a number of genes indicates that enhanced transcription necessitates the simultaneous presence of Oct-1 and Staf transcription factors. These two motifs are always found

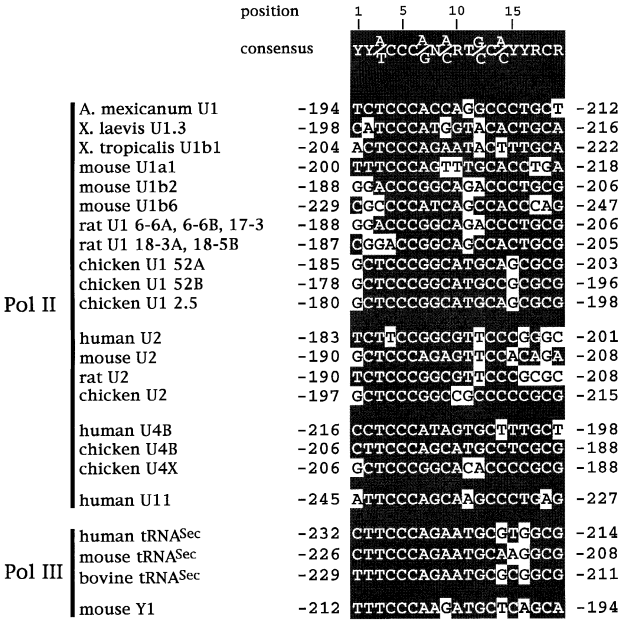

Fig. 6. Potential Staf binding sites in the DSEs of 23 vertebrate snRNA and snRNA-type genes transcribed by Pol II or Pol III. The consensus sequence for the Staf binding site derived from the data on binding site selection (Figure 5) is reported above the Figure. White letters on a dark background show nucleotide identities between the potential Staf binding sites and the experimentally derived consensus sequence. Symbols as in Figure 2B. Sequence references: mouse U1a and U1b (Howard *et al.*, 1986), *X.laevis* U1.3 (Mattaj *et al.*, 1985), *X.tropicalis* U1b1 (P.Carbon and A.Krol, unpublished data), human U11 (Suter-Crazzolara and Keller, 1991), human tRNA^{Sec} (Pavesi *et al.*, 1994), mouse tRNA^{Sec} (Bösl *et al.*, 1995), bovine tRNA^{Sec} (Diamond *et al.*, 1990). For the other genes see Hernandez (1992) and Gu and Reddy (1996).

in close proximity, separated by a maximum of 28 bp. In this regard, we have previously shown that addition of an octamer element in the vicinity of a Staf binding site in the *Xenopus* Pol II U1b2 and Pol III U6 genes produced a synergistic effect on transcriptional activation, with a marked dependence on the spacing between the two motifs (Myslinski *et al.*, 1993b; our unpublished results). Similar results were obtained with chicken U1 52A and U4B (Roebuck *et al.*, 1990; Zamrod and Stumph, 1990). This suggests a functional cooperativity between the two DNA-bound factors, the basis of which is unknown at the present time. Several possibilities can be invoked: (i) Staf and Oct-1 bind cooperatively to the DNA to activate transcription; (ii) the simultaneous presence of Oct-1 and Staf creates a unique surface for interaction with a co-

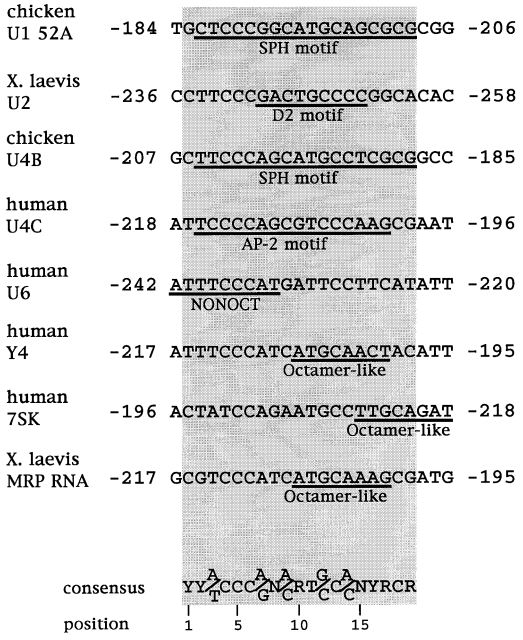

Fig. 7. Staf-responsive elements overlap functional or putative motifs in the activator elements of several snRNA and snRNA-type genes. Staf-responsive elements are indicated on a shaded background. The numbers at the right and left indicate the distance from the transcription initiation site. The sequences of the indicated functional or putative motifs are underlined. References: chicken U1 52A (Roebuck *et al*., 1990), *X.laevis* U2 (Tebb and Mattaj, 1989), chicken U4B (Zamrod and Stumph, 1990), human U4C (Weller *et al*., 1988), human U6 (Danzeiser *et al*., 1993), human Y4 (Maraia *et al*., 1994), human 7SK (Murphy *et al*., 1989, 1992; Boyd *et al*., 1995), *X.laevis* MRP RNA (Bennett *et al*., 1992).

activator or factor(s) of the basal transcription complex; (iii) Staf and Oct-1 each interacts with a distinct co-activator or protein surface of the basal transcription complex. However, the few cases where the DSE function is mediated only by Staf (*X.laevis* selenocysteine tRNA and MRP RNA and human Y4) suggest the interesting possibility that Staf possesses *per se* the capacity to contact alone, or via a co-activator, the basal transcription complex. To the best of our knowledge, it is as yet unknown whether Oct-1 is able to do so in the context of naturally occurring promoters. Further work is required to elucidate this mechanism.

Materials and methods

Preparation of the Staf DNA binding domain

The Staf DNA binding domain was produced using the glutathione S-transferase (GST) gene fusion system. Briefly, the cDNA containing the zinc finger region was inserted into the *Bam*HI and *Eco*RI sites of pGEX-3X (Smith and Johnson, 1988). The resulting plasmid, pGST-Znf 1-7, produces a fusion protein including GST and the zinc finger domain coding sequence between amino acids 256 and 476 (Schuster *et al.*, 1995). The bacterial culture and IPTG induction of GST–Znf 1-7 expression were performed at 25°C. The fusion protein was purified, using glutathione–Sepharose beads, essentially as described in Smith and Johnson (1988).

DNA binding assays

Gel retardation and DNase I footprinting assays were performed essentially as described by Myslinski *et al.* (1992) and Schuster *et al.* (1995). The coding strand of *X.laevis* U1b1 (positions –357/–173) (Krol *et al.*, 1985), *X.laevis* U1b2 (–358/–129) (Krol *et al.*, 1985), *X.laevis* U3A (–310/–160) (Savino *et al.*, 1992), human U6 (–357/–171) (Kunkel and Pederson, 1988), mouse U6 (–315/–220) (Oshima *et al.*, 1981) and *X.tropicalis* U6 (–335/–178) (Krol *et al.*, 1987) were 5′-end-labeled by PCR amplification of the corresponding genes using the proximal ^{32}P-labeled primer. Human U1 (positions –300/–134) (Lund and Dalhberg, 1984), *X.laevis* U2 (–310/–160) (Mattaj and Zeller, 1983), human U4C (–257/–96) (Bark *et al.*, 1986), *X.laevis* U5 (–260/–111) (Kazmaier *et al.*, 1987), human H1 RNA (–279/–130) (Baer *et al.*, 1990), human Y4 (–264/–101) (Maraia *et al.*, 1994), human 7SK (–243/–143) (Murphy *et al.*, 1986), *X.laevis* MRP RNA (–261/–100) (Bennett *et al.*, 1992) and *X.laevis* tRNA^{Sec} (–280/–102) (Lee *et al.*, 1990) were 5′-end-labeled on the non-coding strand by PCR amplification of the corresponding genes using the distal ^{32}P-labeled primer.

Plasmid constructions

Reporter constructs. U1wt, U1·ΔDSE, U6wt and U6·ΔDSE correspond to *X.laevis* U1b2 (Krol *et al.*, 1985), *X.laevis* U1b2·ΔDSE (Murgo *et al.*, 1991), *X.tropicalis* U6 (Krol *et al.*, 1987) and C115 gene constructs (Myslinski *et al.*, 1992) respectively. The U1·3E and U6·3E reporters were obtained by ligating in the inverted orientation the *Bgl*II fragment of PV2-3E (Chavrier *et al.*, 1990) to the *Bam*HI/*Bgl*II-cut *X.laevis* U1b2·ΔDSE and C115 constructs respectively. The E sites map at positions –205/–196, –235/–226, –265/–256 in U1·3E and –219/–210, –249/–240, –279/–270 in U6·3E.

Effector constructs. Construction of pBRN3-Staf/Krox-20 and pBRN3/Krox-20 DBD was as described in Schuster *et al.* (1995).

Oocyte microinjections

In the experiments shown in Figure 3B, *X.laevis* oocytes were co-injected with 4 ng wild-type or mutant templates, 0.2 μCi [α-^{32}P]GTP (800 Ci/mmol) and the 5S RNA maxigene (25 pg for Pol II genes and 100 pg for Pol III genes) as an internal control for oocyte injection and RNA recovery, except for hU4C, where the tRNA^{Phe} gene (100 pg) was used instead. For competition experiments, oocyte nuclei were co-injected with 8 ng each template and 25 pg 5S RNA maxigene. Oocytes were incubated at 19°C for 5 (Pol III transcription) or 16 h (Pol II transcription). RNAs were extracted from batches of 10 oocytes and analyzed as described in Schuster *et al.* (1995). Transcription efficiencies were quantitated with a Fuji Bioimage Analyzer Bas 2000 and normalized relative to 5S RNA maxi or tRNA^{Phe} transcription levels.

In the experiments shown in Figure 4, capped mRNAs (20 nl, 10 ng) were injected into the cytoplasm 20 h before nuclear injection of 20 nl containing the reporter DNA (50 μg/ml), the 5S maxigene (5 μg/ml) as an internal control and [α-^{32}P]GTP (800 Ci/mmol, 0.2 μCi/oocyte). Incubation was for 16 (U1 reporters) or 5 h (U6 reporters). Transcription of the reporter genes was analyzed as described in Schuster *et al.* (1995).

Binding site selection

The 57 bp oligonucleotide used in the binding selection, 5′-CTGGATCCTAAGATTCCCTG(N)$_{17}$AGGCTCAAAGCTGAATTCCT-3′, contained an internal region of 17 degenerate bp flanked on each side by a 20 bp sequence containing *Bam*HI (5′) and *Eco*RI (3′) restriction sites. For PCR amplification, the oligonucleotides 5′-CTGGATCCTAAGATTCCCTG-3′ and 5′-AGGAATTCAGCTTTGAGCCT-3′ served as forward and reverse primers respectively. Selection was performed essentially as described in Delwel *et al.* (1993). After six rounds of binding and amplification by PCR, an additional step was performed to ensure that the majority of the amplified 57 bp oligonucleotides represented perfect duplexes lacking mismatches (Chittenden *et al.*, 1991). To do this, 200 pmol of each primer were added to the reaction and the mixture subjected to an additional PCR cycle. The final oligonucleotide amplification product was purified, *Bam*HI/*Eco*RI digested and ligated to pBS (+). Isolated clones were sequenced by standard methods (Sambrook *et al.*, 1989). In binding site comparisons, to avoid biasing the data, nucleotides recognized by the PCR primers within the defined sequence were excluded.

Acknowledgements

We are grateful to I.W.Mattaj for the gift of the xU2, xU5 and hU6 plasmids, C.Bark and U.Pettersson for hU4C, Y.Oshima and A.Lescure for mU6, R.Maraia for hY4, S.Murphy for h7SK, D.A.Clayton for xMRP RNA, E.Lund for hU1, S.Altman for hH1 RNA and S.Gerbi for xU3. R.Bordonné and F.X.Wilhelm are thanked for critical reading of the manuscript, P.Remy for microinjection facilities and C.Loegler for excellent technical assistance. This work was supported by grants from the ULP Strasbourg, the Association pour la Recherche sur le Cancer

(ARC) and the European Union (EEC Biotech Program BIO2-CT92-0090).

References

Ach,R.A. and Weiner,A.M. (1991) Cooperation between CCAAT and octamer motifs in the distal sequence element of the rat U3 small nucleolar RNA promoter. *Nucleic Acids Res.*, **19**, 4209–4218.

Ares,M.,Jr, Chung,J.-S., Giglio,L. and Weiner,A.M. (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.

Baer,M., Nilsen,T.W., Costigan,C. and Altman,S. (1990) Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. *Nucleic Acids Res.*, **18**, 97–103.

Bark,C., Weller,P., Zabielski,J., and Pettersson,U. (1986) Genes for human U4 small nuclear RNA. *Gene*, **50**, 333–344.

Bennett,J.L., Jeong-Yu,S. and Clayton,D.A. (1992) Characterization of a *Xenopus laevis* ribonucleoprotein endoribonuclease. *J. Biol. Chem.*, **267**, 21765–21772.

Blackwell,T.K., Kretzner,L., Blackwood,E.M., Eisenman,R.N. and Weintraub,H. (1990) Sequence-specific DNA binding by the c-myc protein. *Science*, **250**, 1149–1151.

Bösl,M.R., Seldin,M.F., Nishimura,S. and Taketo,M. (1995) Cloning, structural analysis and mapping of the mouse selenocysteine $tRNA^{[Ser]Sec}$ gene (*Trsp*). *Mol. Gen. Genet.*, **248**, 247–252.

Boyd,D.C., Turner,P.C., Watkins,N.J., Gerster,T. and Murphy,S. (1995) Functional redundancy of promoter elements ensures efficient transcription of the human 7SK gene *in vivo*. *J. Mol. Biol.*, **253**, 677–690.

Carbon,P. and Krol,A. (1991) Transcription of the *Xenopus laevis* selenocysteine $tRNA^{(Ser)Sec}$ gene : a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *EMBO J.*, **10**, 599–606.

Chavrier,P., Vesque,C., Galliot,B., Vigneron,M., Dollé,P., Duboule,D. and Charnay,P. (1990) The segment-specific gene Krox-20 encodes a transcription factor with binding sites in the promoter region of the Hox-1.4 gene. *EMBO J.*, **9**, 1209–1218.

Cheung,C.H., Fan,Q.N. and Stumph,W. (1993) Structural requirements for the functional activity of a U1 snRNA gene enhancer. *Nucleic Acids Res.*, **21**, 281–287.

Chittenden,T., Livingston,D.M. and Kaelin,W.G.,Jr (1991) The T/E1A-binding domain of the retinoblastoma product can interact selectively with a sequence-specific DNA-binding protein. *Cell*, **65**, 1073–1082.

Courey,A.J. and Tjian,R. (1988) Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell*, **55**, 887–898.

Danzeiser,D.A., Urso,O. and Kunkel,G.R. (1993) Functional characterization of elements in a human U6 small nuclear RNA gene distal control region. *Mol. Cell. Biol.*, **13**, 4670–4678.

Das,G., Hinkley,C.S. and Herr,W. (1995) Basal promoter elements as a selective determinant of transcriptional activator function. *Nature*, **374**, 657–659.

Delwel,R., Funabiki,T., Kreider,B.L., Morishita,K. and Ihle,J.N. (1993) Four of the seven zinc fingers of the Evi-1 myeloid transforming gene are required for sequence specific binding to GA(C/T)AAGA(T/C)AAGATAA. *Mol. Cell. Biol.*, **13**, 4291–4300.

Diamond,A.M., Montero-Puerner,Y., Lee,B.J. and Hatfield,D. (1990) Selenocysteine inserting tRNAs are likely generated by tRNA editing. *Nucleic Acids Res.*, **18**, 6727.

Gu,J. and Reddy,R. (1996) Small RNA database. *Nucleic Acids Res.*, **24**, 73–75.

Hernandez,N. (1992) Transcription of vertebrate snRNA genes and related genes. In MacKnight,S.L. and Yamamoto,K.R. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 1, pp. 281–313.

Herr,W. (1992) Oct-1 and Oct-2: differential transcriptional regulation by proteins that bind to the same DNA sequence. In MacKnight,S.L. and Yamamoto,K.R. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 2, pp. 1103–1135.

Howard,E.F., Michael,S.K., Dahlberg,J.E. and Lund,E. (1986) Functional, developmentally expressed genes for mouse U1a and U1b snRNAs contain both conserved and non-conserved transcription signals. *Nucleic Acids Res.*, **14**, 9811–9825.

Janson,L., Bark,C. and Pettersson,U. (1987) Identification of proteins interacting with the human U2 small nuclear RNA genes. *Nucleic Acids Res.*, **15**, 4997–5016.

Kazmaier,M., Tebb,G. and Mattaj,I.W. (1987) Functional characterization of *X.laevis* U5 snRNA genes. *EMBO J.*, **6**, 3071–3078.

Kleinert,H., Bredow,S. and Benecke,B.-J. (1990) Expression of a human 7SK RNA gene *in vivo* requires a novel pol III upstream element. *EMBO J.*, **9**, 711–718.

Krol,A., Lund,E. and Dahlberg,J.E. (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., Carbon,P., Ebel,J.-P. and Appel,B. (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 Pederson,T. (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.

Lee,B.J., Rajagopalan,M., Kim,Y.S., You,K.H., Jacobson,K.B. and Hatfield,D. (1990) Selenocysteine $tRNA^{[Ser]Sec}$ gene is ubiquitous within the animal kingdom. *Mol. Cell. Biol.*, **10**, 1940–1949.

Lescure,A., Tebb,G., Mattaj,I.W., Krol,A. and Carbon,P. (1992) A factor with Sp1 DNA-binding specificity stimulates *Xenopus* U6 snRNA *in vivo* transcription by RNA polymerase III. *J. Mol. Biol.*, **228**, 387–394.

Lobo,S.M. and Hernandez,N. (1989) A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. *Cell*, **58**, 55–67.

Lund,E. and Dahlberg,J.E. (1984) True genes for human U1 small nuclear RNA. *J. Biol. Chem.*, **259**, 2013–2021.

Maraia,R.J., Sasaki-Tozawa,N., Driscoll,C.T., Green,E.D. and Darlington,G.J. (1994) The human Y4 small cytoplasmic RNA gene is controlled by upstream elements and resides on chromosome 7 with all other hY scRNA genes. *Nucleic Acids Res.*, **22**, 3045–3052.

Mattaj,I.W. and Zeller,R. (1983) *Xenopus laevis* U2 snRNA genes: tandemly repeated transcription units sharing 5′ and 3′ flanking homology with other RNA polymerase II transcribed genes. *EMBO J.*, **2**, 1883–1891.

Mattaj,I.W., Zeller,R., Carrasco,A.E., Jamrich,M., Lienhard,S. and De Robertis,E.M. (1985) U snRNA gene families in *Xenopus laevis*. *Oxford Surv. Eukaryotic Genes*, **2**, 121–140.

Mattaj,I.W., Dathan,N.A., Parry,H.D., Carbon,P. and Krol,A. (1988) Changing the RNA polymerase specificity of U snRNA gene promoters. *Cell*, **55**, 435–442.

Murgo,S., Krol,A. and Carbon,P. (1991) Sequence, organization and transcriptional analysis of a gene encoding a U1 snRNA from the axolotl, *Ambystoma mexicanum*. *Gene*, **99**, 163–170.

Murphy,S., Tripodi,M. and Melli,M. (1986) A sequence upstream from the coding region is required for the transcription of the 7SK RNA genes. *Nucleic Acids Res.*, **14**, 9243–9260.

Murphy,S., Pierani,A., Scheidereit,C., Melli,M. and Roeder,R.G. (1989) Purified octamer binding transcription factors stimulate RNA polymerase III-mediated transcription of the 7SK RNA gene. *Cell*, **59**, 1071–1080.

Murphy,S., Yoon,J.-B., Gerster,T. and Roeder,R.G. (1992) Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. *Mol. Cell. Biol.*, **12**, 3247–3261.

Myslinski,E., Krol,A. and Carbon,P. (1992) Optimal $tRNA^{Ser/Sec}$ gene activity requires an upstream SPH motif. *Nucleic Acids Res.*, **20**, 203–209.

Myslinski,E., Schuster,C., Huet,J., Sentenac,A., Krol,A. and Carbon,P. (1993a) Point mutations 5′ to the tRNA selenocysteine TATA box alter RNA polymerase III transcription by affecting the binding of TBP. *Nucleic Acids Res.*, **21**, 5852–5858.

Myslinski,E., Schuster,C., Krol,A. and Carbon,P. (1993b) Promoter strength and structure dictate module composition in RNA polymerase III transcriptional activator elements. *J. Mol. Biol.*, **234**, 311–318.

Ohshima,Y., Okada,N., Tani,T., Itoh,Y. and Itoh,M. (1981) Nucleotide sequences of mouse genomic loci including a gene or pseudogene for U6 (4.8S) nuclear RNA. *Nucleic Acids Res.*, **9**, 5145–5158.

Pavesi,A., Conterio,F., Bolchi,A., Dieci,G. and Ottonello,S. (1994) Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions. *Nucleic Acids Res.*, **22**, 1247–1256.

Roebuck,K.A., Szeto,D.P., Green,K.P., Fan,Q.N. and Stumph,W.E. (1990) Octamer and SPH motifs in the U1 enhancer cooperate to activate U1 RNA gene expression. *Mol. Cell. Biol.*, **10**, 341–352.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Savino,R., Hitti,Y. and Gerbi,S.A. (1992) Genes for *Xenopus laevis* U3 small nuclear RNA. *Nucleic Acids Res.*, **20**, 5435–5442.

Schuster,C., Myslinski,E., Krol,A. and Carbon,P. (1995) Staf, a novel zinc finger protein that activates the RNA polymerase III promoter of the selenocysteine tRNA gene. *EMBO J.*, **14**, 3777–3787.

Smith,D.B. and Johnson,K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*, **67**, 31–40.

Sturm,R.A., Das,G. and Herr,W. (1988) The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeobox subdomain. *Genes Dev.*, **2**, 1582–1599.

Suter-Crazzolara,C. and Keller,W. (1991) Organization and transient expression of the gene for human U11 snRNA. *Gene Expression*, **1**, 91–102.

Tebb,G. and Mattaj,I.W. (1989) The *Xenopus laevis* U2 gene distal sequence element (enhancer) is composed of four subdomains that can act independently and are partly functionally redundant. *Mol. Cell. Biol.*, **9**, 1682–1690.

Weller,P., Bark,C., Janson,L. and Pettersson,U. (1988) Transcription analysis of a human U4C gene: involvement of transcription factors novel to snRNA gene expression. *Genes Dev.*, **2**, 1389–1399.

Zamrod,Z. and Stumph,W.E. (1990) U4B snRNA gene enhancer activity requires functional octamer and SPH motifs. *Nucleic Acids Res.*, **18**, 7323–7330.

Received on July 19, 1996; revised on September 23, 1996