Site-directed Mutational Analysis of a U4 Small Nuclear RNA Gene Proximal Sequence Element

LOCALIZATION AND IDENTIFICATION OF FUNCTIONAL NUCLEOTIDES*

(Received for publication, February 8, 1990)

Kathleen J. McNamara and William E. Stumph†

From the Department of Chemistry and the Molecular Biology Institute, San Diego State University, San Diego, California 92182-0328

The genes that encode the small nuclear RNAs (snRNAs) are unusual RNA polymerase II transcription units in that 5'-flanking DNA sequences more than 50 base pairs upstream of snRNA genes are essential for specifying the transcription initiation site. The relevant cis-acting DNA sequence, termed the proximal sequence element (PSE), is required for both transcription initiation and 3'-end formation of snRNAs. We have used site-directed mutagenesis and expression in *Xenopus* oocytes to map nucleotides important for the function of the chicken U4B snRNA gene PSE. The results indicate that nucleotide sequences upstream of position -65 are not required for U4B PSE activity. However, nucleotides lying within a region 53-65 base pairs upstream of the U4B gene are essential for obtaining a detectable level of U4B gene expression. Six nucleotides between positions -53 and -59 were identified at which base substitutions reduced the transcriptional activity of the U4B gene.

The small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6 function in the splicing of mRNA precursors (Steitz et al., 1988). These small RNAs are capped but not polyadenylated and (with the exception of U6) are synthesized by an RNA polymerase II-type activity (Dahlberg and Lund, 1988; Parry et al., 1989a). However, the genes that encode these snRNAs are unusual RNA polymerase II transcription units. Even though they lack TATA boxes, transcription in vivo is always initiated at a well-defined nucleotide position. In addition, the formation of the 3'-ends of snRNAs is dependent upon transcription being initiated from an snRNA gene promoter (Neuman de Vegvar et al., 1986; Hernandez and Weiner, 1986).

Comparative sequence analysis and functional expression of a number of snRNA genes have revealed two distinct regions in the 5'-flanking DNA important for snRNA gene expression. The more distal region, generally positioned between nucleotides -250 and -180 relative to the snRNA cap site, functions as a transcriptional enhancer (Dahlberg and Lund, 1988; Parry et al., 1989a). A more compact proximal sequence element (PSE) specifies the transcription initiation site (Dahlberg and Lund, 1988; Parry et al., 1989a) and is required for the formation of a stable transcription complex (Tebb and Mattaj, 1988). The PSE is therefore considered to be functionally analogous to a TATA box. In addition, however, the PSE also plays an important role in 3'-end formation of the RNA polymerase II-transcribed snRNAs (Neuman de Vegvar et al., 1986; Hernandez and Lucito, 1988; Parry et al., 1989b; Neuman de Vegvar and Dahlberg, 1989). Only transcription complexes formed using an snRNA gene PSE are able to recognize a signal in the 3'-flanking DNA (the 3'-box) which is essential for snRNA 3'-end formation. No mutations in the 5'-flanking DNA have been found that decrease the efficiency of 3'-end formation while leaving the efficiency of transcription initiation unaffected (Hernandez and Lucito, 1988; Parry et al., 1989b; Neuman de Vegvar and Dahlberg, 1989). Thus, both activities (transcription initiation and ability to recognize the 3'-box) seem to require the same DNA sequence in the PSE, meaning that correct initiation and termination of snRNA synthesis cannot be uncoupled. Transient expression studies have localized functional DNA sequences of the human and *Xenopus* U1 and U2 RNA gene PSEs to nucleotides located -50-60 base pairs (bp) upstream of the snRNA cap site (Ciliberto et al., 1985; Murphy et al., 1987; Hernandez and Lucito, 1988; Parry et al., 1989b).

The gene that codes for chicken U4B RNA (Hoffman et al., 1986) is a typical vertebrate snRNA gene in that it contains both distal and proximal regulatory elements. We previously showed that deletion of the enhancer region, which resides upstream of position -180, resulted in a 3-5-fold decrease in U4B transcriptional activity (McNamara et al., 1987). A U4B template with 117 base pairs of 5'-flanking DNA still retained a considerable level of basal activity, whereas a template with a 5'-truncation to position -38 had no detectable activity (McNamara et al., 1987). This indicated that sequences located between positions -117 and -38 are required for U4B gene expression. Similar experiments performed with a human U4 gene indicated that sequences between positions -121 and -50 are required for activity (Weller et al., 1988). However, neither detailed mapping nor point mutational analysis of a U4 gene PSE has yet been reported.

Therefore, we generated a number of constructs containing truncations and base substitutions in the chicken U4B gene 5'-flanking DNA, and we assayed these constructs for transcriptional activity by injection into *Xenopus laevis* oocytes. Our results indicate that DNA sequences located between positions -65 and -53 are required for the transcriptional activity of the U4B gene template. Moreover, we have identified six nucleotide positions in this region at which base substitutions have a detrimental effect on the activity of the U4B gene PSE.
Materials and methods

Template constructions—All U4 constructions contain the chicken U4B gene (Hoffman et al., 1986) cloned into the polylinker region of the pUC19 vector in the same transcriptional orientation as the lacZ gene. In the constructions pU4BI-227, pU4BΔ-117, pU4BΔ-65, and pU4BΔ-52 (truncated at AarII, KpnI, Sau961, and Bsp1286 sites, respectively) each contain the designated amount of wild-type 5' flanking DNA sequence, the U4B gene coding region, and 114 bp of 3' flanking DNA. The templates designated in this report as pU4BΔ-227 and pU4BΔ-117 differ from earlier constructions of the same gene in that the previous constructions (McNamara et al., 1987) contained 283 bp of 3' flanking DNA.

Template pU4B/S is identical to pU4BΔ-227 except for an A to T base change at position -48 in the 5' flanking DNA. This change converts a natural Bsp1286 recognition sequence (GAGCAC) to a T base change at position -48. This template pU4B/S is similar to the mutant pU4B/S except that it contains an SstI site due to an A to T change at position -48 (which cleaves in the polylinker 3' of the gene), isolating the U4B gene-containing fragment, and then cloning it between the SstI and EcoRI sites of pUC19. (Although the Bsp1286 and SstI sticky ends are not identical in this instance, they ligate together with a fairly high efficiency.) Clones were then screened, and one was selected that had an SstI site due to an A to T change at position -48. After digesting this construct with HindIII and SstI, it was cloned into the oocyte nuclei of X. laevis together with [α-32P]GTP. RNA containing 227, 117, 65, or 52 bp of 5' flanking DNA were injected for U4B gene PSE activity. Chicken U4B gene plasmid templates were used as a positive internal control to normalize for efficiency of injection into the oocyte nuclei and for the recovery of RNA. Each oocyte was injected with ~30 nl of solution containing 88 mM NaCl, 10 mM Tris Cl (pH 7.5), 400 ng of U4B gene plasmid/μl, and 7 μCi of [α-32P]GTP (410 Ci/mmol)/μl. For the transcription competition experiments, the injection mixture contained the above components except that the U4B plasmid and the U1 competitor template (pU1A-246; Roeback et al., 1987) were also included at a concentration of 200 ng/μl.

Results

Functional localization of U4B PSE—To localize the functional DNA sequences of the proximal regulatory region of the U4B RNA gene, an initial series of 5' deletion constructions was prepared and then analyzed for template activity by injection into X. laevis oocytes. The relevant portions of the constructions are diagramed at the bottom of Fig. 1. Previous work had indicated that 227 bp of 5' flanking DNA is sufficient for a high level of expression of the U4B gene in Xenopus oocytes (McNamara et al., 1987). Transcription results using this "wild-type" template (pU4BΔ-227) are shown in Fig. 1 (lanes 1 and 5). A deletion to position -117 (lane 2), which removed the U4B enhancer, resulted in a ~3-4-fold reduction in transcriptional activity. A very slight further reduction in activity was observed with the template pU4BΔ-65 (lane 3); this suggests that sequences between positions -117 and -65 provide a small degree of positive modulation of U4B gene expression. Several potential GC boxes (Sp1 factor-binding sites) are located within this region and may account for this activity (Hoffman et al., 1986). Because the effect of the deletion at positions -117 to -65 was minor in the expression assay, the relevant sequences have not been mapped or studied further. More importantly, there was no detectable expression from the template truncated at position -52 (lane 4). This indicates that sequences essential for a basal level of U4B gene activity lie within a region 53-65 bp upstream of the gene.

To further map the sequences required for the function of the proximal regulatory region and to facilitate site-directed...
Site-directed Mutational Analysis of a U4 Gene PSE

Mutagenesis, two single base substitutions were introduced into the 5′-flanking DNA of the wild-type construction to generate unique restriction sites flanking the PSE. An A to T change at position −48 created a unique SstI site, and a G to C change at position −64 created a unique ApaI site. Fig. 1 (right) shows that both the single mutant (pU4B/S; lane 7) and the double mutant (pU4B/AS; lane 8) were expressed at the same level as the wild-type template (lanes 6 and 9). Therefore, the template pU4B/AS was utilized as a parental construction to generate additional point mutations within the proximal regulatory region.

Effects of Point Mutations within U4B PSE—The pentanucleotide sequence CCGTG is perfectly conserved in the proximal region of the chicken U4B and U2 RNA genes and the four chicken U1 RNA genes that have been cloned (Hoffman et al., 1986). The U4X gene contains the similar sequence CTGTG (Hoffman et al., 1986). Therefore, for the first series of experiments, single point mutations were generated in this conserved pentanucleotide sequence. These constructions (pU4B/AS1 through pU4B/AS6) are shown at the bottom of Fig. 2. Each of these mutant templates included the wild-type U4B enhancer since each extended to position −227 in the 5′-flanking DNA. When these constructions were injected individually into Xenopus oocytes, only one of the mutations (a C to A transversion at position −56) significantly lowered the level of expression in the oocyte expression assay (data not shown).

We and others have found that coinjection of a wild-type competing snRNA gene template into oocytes increases the sensitivity of the assay (Murphy et al., 1987; Tebb and Mattaj, 1988; Parry et al., 1989b; Roebuck et al., 1990). The competition for transcription factors makes it possible to observe changes in the transcriptional efficiency of mutant templates that are not measurable when the mutant templates are injected by themselves. Therefore, in the remaining experiments, the U4B mutant templates were coinjected together with a plasmid containing a wild-type chicken U1 RNA gene as a competitor (Roebuck et al., 1987). Under these conditions, a much wider variation in template efficiency was observed among the various U4B templates (Fig. 2).

Two single point mutations considerably reduced U4B gene expression: a C to A change at position −56 and a T to G change at position −54 (Fig. 2, lanes 1 and 5, respectively). A G to T substitution at position −52 (lane 6) had a nearly negligible effect. Interestingly, a C to T transition at position −55 decreased expression (lane 2), whereas a C to A transversion at the same position (−55) appeared to cause a slight increase in template activity (lane 3). Similarly, the construct with a G to T change at position −54 (lane 4) was also expressed somewhat better than the parental pU4B/AS construction.

Because the single transversion mutations at positions −55 and −54 resulted in templates that appeared more active than the parental construct, we also studied the effect of the double transversion mutation at positions −55 and −54 (Fig. 3). This double point mutant was expressed as well as the pU4B/AS parental construct (Fig. 3, compare lanes 2 and 8). A triple point mutant that incorporated an additional T to G change at position −53 (lane 1) exhibited a reduced level of expression. This is consistent with the down-effect observed in the single point mutant pU4B/AS5 (Fig. 2, lane 5). However, the magnitude of the effect in the context of the triple mutant was not as great as in the single substitution mutant, consistent with the apparent stimulatory effect of the single point transversions at positions −55 and −54.

We also studied the effects of mutating the three nucleotide...
positions immediately upstream of the CCGTG sequence. The results shown in Fig. 3 indicate that these three positions (−59 to −57) are also important for the function of the proximal region. A single change (G to T) at position −57 caused a reduction of U4B expression (Fig. 3, lane 6). The two double mutants had template activities that were further reduced relative to the single point mutant (lanes 3 and 5). Consistent with the above results, transcription of the triple point mutant (with changes at positions −59 to −57) was barely detectable (lane 4). Finally, a template with a deletion of 19 bp encompassing the entire proximal region was totally inactive (lane 7), indicating that a gross deletion of 19 bp encompassing the proximal region completely destroyed template activity.

**DISCUSSION**

Our results can be summarized as follows. 5′-Truncation experiments indicate that sequences upstream of position −52 are essential for U4B PSE function, whereas sequences upstream of position −65 can be deleted with little or no loss of PSE activity. Moreover, a 19-bp internal deletion of nucleotides between positions −66 and −48 completely eliminated detectable U4B gene expression. However, point mutations at positions −64 and −48 had no effect on expression.

We also examined the effects of a number of specific point mutations at positions −52 to −59. This region is well-conserved in sequence among chicken U1, U2, and U4 RNA genes, whereas nucleotides flanking this region exhibit considerably more variability (Hoffman et al., 1986). The effects of these mutations at positions −52 to −59 are summarized in Fig. 4. In initial experiments (Fig. 2), we made single base changes in the pentanucleotide sequence CCGTG, which is 100% conserved among the chicken U4B gene and the chicken U1 and U2 genes that have been cloned (Hoffman et al., 1986). Somewhat surprisingly, only the mutations at positions −53 and −56 in this sequence had major effects on U4B gene activity. Our data indicate that the three not-as-highly-conserved nucleotide positions (−50 to −57) immediately upstream of the CCGTG pentanucleotide also make a substantial contribution to PSE function. Indeed, the triple point mutation at positions −59 to −57 resulted in a PSE that was nearly nonfunctional.

However, none of the single, double, or triple point mutations in the region at positions −52 to −59 completely eliminated PSE activity. Thus, there is certainly sequence flexibility permitted in the U4B PSE. In fact, two of the PSE mutations resulted in a higher level of activity in the oocyte expression system than was observed with the wild-type chicken U4B PSE sequence. One possibility is that the U4B PSE is designed to function at less than optimal efficiency to allow for additional levels of regulation. However, another explanation is that there may be some degree of species specificity involved in the recognition of the PSE by the transcriptional machinery of the frog oocyte versus the homologous chicken machinery. Indeed, sequence comparisons have resulted in the derivation of somewhat different PSE consensus sequences for amphibian, avian, and mammalian snRNA genes (Hoffman et al., 1986; Bark et al., 1986; Dahlberg and Lund, 1988; Parry et al., 1989b), suggesting that snRNA gene PSEs and the factors that recognize them may have co-evolved during vertebrate evolution. In this case, the exact effect of any specific point mutation may depend upon the expression system chosen for the assay.

To our knowledge, Parry et al. (1989b) have published the only other study that has examined the effect of single, double, and triple point mutations in an snRNA gene PSE. Their results on the Xenopus U2 gene PSE are difficult to compare directly with our results even though both studies were carried out in an oocyte expression system. Although the chicken U4B gene is efficiently expressed in frog oocytes, the PSEs of the Xenopus U2 and chicken U4B genes differ substantially in sequence; this makes it impossible to unambiguously align the functional nucleotides in the two PSEs. However, in comparing the data, there seems to be a central point of consensus. In both cases, substitutions at positions −59 to −56 were almost always detrimental, whereas changes at positions −55 to −52 seemed to have variable effects. Thus, the strict conservation of nucleotides at these latter positions among chicken snRNA genes may represent an additional requirement specific to the chicken system.

**REFERENCES**

K J McNamara and W E Stumph


Access the most updated version of this article at http://www.jbc.org/content/265/17/9728

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/17/9728.full.html#ref-list-1