Identification of Proximal Sequence Element Nucleotides Contributing to the Differential Expression of Variant U4 Small Nuclear RNA Genes*

Hirdeypal S. Bhathal†, Zulkeflie Zamrodi, Takaaki Tobaru, and William E. Stumph‡

From the Department of Chemistry and Molecular Biology Institute, San Diego State University, San Diego, California 92182-1030

The two U4 genes in the chicken genome code for distinct sequence variants of U4 small nuclear RNA that are differentially expressed during development. Whereas U4B RNA is constitutively expressed, U4X RNA is specifically down-regulated relative to U4B in a tissue-specific manner during development. To investigate mechanisms controlling the differential expression of the U4B and U4X genes, chimeric U4 genes were constructed and their transcriptional activities assayed by injection into Xenopus oocytes or by transfection of CV-1 cells. The proximal regulatory region of the U4B gene and the 5′-flanking regions of both the U4B and U4X genes functioned efficiently in each expression system. However, the proximal region of the U4X gene was inactive. To localize and identify the responsible nucleotides, reciprocal point mutations were introduced into the U4X and U4B proximal regulatory regions. The results indicate that the U4X gene contains a suboptimal proximal sequence element, and that this results primarily from the identities of the nucleotides at positions −61 and −57 relative to the transcription start site.

The small nuclear RNAs (snRNAs) of the U family (U1, U2, U3, etc.) are evolutionarily conserved and metabolically stable RNAs that function in pre-mRNA and pre-rRNA processing in the nuclei of eukaryotic cells. The genes for vertebrate snRNAs possess similar promoter structures and, with the exception of U6, are transcribed by RNA polymerase II. Two distinct evolutionarily conserved regions in the 5′-flanking DNA contribute to the expression of vertebrate snRNA genes (1–4). The more distal region, generally positioned between nucleotides −250 and −180 relative to the transcription start site, functions as a transcriptional enhancer. It normally contains at least one copy of the octamer motif (recognized by the Oct-1 transcription factor), and it usually contains functional binding sites for one or more additional proteins, such as a GC box, SRF motif, CCAAT box, cyclic AMP response element, etc. (5–12).

The proximal sequence element (PSE) is an essential promoter element centered near position −55 and is required for accurate initiation of snRNA transcription. The PSE is recognized by a transcription factor variously named PSE binding protein (13, 14), PSE transcription factor (15), or snRNA activating protein complex (16). The PSE binding factor is believed to play a role in recruiting the TATA box-binding protein to snRNA promoters (16–19).

In the chicken, there are two and only two genes that code for U4 snRNA (20, 21). These genes, designated U4B and U4X, are closely linked within 500 base pairs (bp) of each other in the chicken genome. The two genes encode distinct sequence variants of U4 RNA that differ at seven nucleotide positions. Moreover, the relative amount of U4X and U4B RNA is differentially regulated during chicken development. The U4X and U4B genes are expressed at nearly equal levels during early embryogenesis, but U4X gene expression is specifically down-regulated in a tissue-specific manner relative to U4B as development proceeds (22).

Although there is no evidence linking the differential expression of the chicken U4X and U4B RNA genes to transcriptional control, their promoter structures are consistent with such a mechanism. First there is no obvious similarity between the U4B and U4X gene 5′-flanking DNA other than the PSE and the enhancer region. Second, both the enhancer and PSE contain a number of base differences in crucial regulatory regions that could affect the relative expression of the two U4 genes (20, 23). Third, just 5′ of, and partially overlapping, the U4X PSE there exists a unique dyad symmetry element that is not found in the U4B proximal region. This palindrome sequence is recognized by a sequence-specific DNA binding factor, proximal palindrome binding factor, whose DNA binding activity is inversely related to the accumulation of U4X RNA (23).

To gain further insight into the molecular basis for the differential expression of the U4X and U4B genes, we have generated constructs that possess chimeric U4B/U4X promoters, including families of constructs with reciprocal point mutations in the PSE and dyad symmetry elements. We find that the proximal regulatory region, not the enhancer, is responsible for the differential activity of the U4B and U4X genes in transient expression assays. Moreover, we identify two specific nucleotide positions that differ between the U4X and U4B PSEs that are crucial for U4 snRNA gene expression in these assays.

MATERIALS AND METHODS

Construction of Plasmids—All constructs were in the pUC19 vector. Plasmid BBB contained the chicken U4B gene, 280 bp of 5′-flanking DNA, and 283 bp of 3′-flanking DNA. Plasmid XXX contained the U4X gene, 323 bp of 5′-flanking DNA, and 116 bp of 3′-flanking DNA. These plasmids were previously designated pUCB(Sst) and pUCX(Sst) since the cloned wild type fragments were flanked by natural SstI sites (20).
Template BBX contained the U4B gene 5'-flanking DNA but the coding region and 3'-flanking DNA of the U4X gene; it was constructed by utilizing an AluI site that occurs at the beginning of the U4 coding region. The reciprocal exchange of SstI-AluI fragments produced the template XBB.

To construct template BXX, U4B sequences from -227 to -142 were joined to position -132 of the U4X gene. Template BXX was constructed by deleting U4B sequences from positions -280 to -118 in the 5'-flanking region of the BBX construct. Template XBB contained, in order, U4X sequences from -323 to -97, two nucleotides (AT) of linker sequence, and U4B 5'-flanking DNA from position -117 to the beginning of the gene, followed by the U4X RNA coding region and 3'-flanking DNA.

To construct the U4B maxigene, 17 nucleotides from the pUC19 polylinker were inserted between positions 8 and 9 of the U4B gene as follows. The construct U4B 3' to -227 (9) was cut with EcoRI and HindIII, and the fragment containing the U4B gene sequence was isolated. This fragment was then cut with restriction endonuclease FspI, giving blunt ends between nucleotides 8 and 9 of the U4B DNA coding region. These fragments were then cloned between the EcoRI and SmaI sites and between the HindIII and SmaI sites of the pUC19 vector. Thus, the plasmid polylinker sequence between the HindIII and SmaI sites became a part of the U4B maxigene transcript. The U4mB/AS template was constructed by exchanging the 5'-flanking DNA of U4mB with that of U4B/AS (24). The remaining U4X and U4B constructs that contained point mutations were generated by site-directed mutagenesis using polymerase chain reaction or by cloning synthetic oligonucleotides between pre-existing restriction sites.

Expression Assays in Xenopus Oocytes—Microinjection of oocytes and analysis of labeled RNA were performed as described previously (20). Oocytes were each injected with ~30 nl of a solution containing 400 μg/ml U4 template DNA, 1 μg/ml Xenopus 5 S gene plasmid, and 8 μCi/ml [α-32P]GTP. After an overnight incubation, RNA was isolated from oocytes, analyzed by electrophoresis in denaturing polyacrylamide gels, and autoradiographed. Normalization for efficiency of injection and recovery of RNA was carried out by adjusting the amounts of sample loaded in subsequent gels to give a uniform 5 S RNA signal. Relative transcription efficiencies were determined by cutting out the bands from the gels and quantitating in a liquid scintillation counter.

Transfection of CV-1 Cells—Monkey kidney CV-1 cells were grown in Dulbecco's modified Eagle's medium from Life Technologies, Inc. supplemented with 5% bovine serum for transfection experiments. Plates containing confluent cells were split 1:6 on 100-mm plates 24 h prior to transfection. Transfection was carried out using the Profection Mammalian Transfection System and accompanying protocols from Promega. Each plate was transfected with 15 μg of the U4 plasmid DNA together with 5 μg of pCDNA3-CT Plasmid (Invitrogen) as control for transfection efficiency. The cells were exposed to the precipitate for 3 h after which the media was replaced, and the cells were further incubated for 24 h before harvesting. RNA from the transfected CV-1 cells was isolated using the RNAgent Total RNA Isolation System employing guanidinium isothiocyanate (Promega). U4 transcripts were measured by primer extension as described below, and chloramphenicol acetyltransferase transcripts were quantified by an RNAase protection assay using a kit purchased from Ambion (Austin, TX). Transfections were repeated a minimum of three times and typical results are presented in the figures.

Primer Extension Assays—All primers were labeled by phosphorylation with γ-[32P]ATP and T4 polynucleotide kinase. Primer extension of U4X RNA employed a labeled U4X primer (5'-CCTCCGGGGAGCCTCTCAA-3') that anneals specifically to U4X RNA from position 19 to +35. The primer extension reactions were performed as follows: 5 μg of the specific primer (~500,000 cpm) was mixed with 10 μl of total RNA in 10 μl of 1× annealing buffer solution (250 mM KCl, 10 mM Tris-Cl, pH 8.3). The mixture was heated at 80°C for 3 min and transferred to a 45°C water bath for 45 min. After 45 min of annealing, 10 μl of ice-cold nuclease-free water was added to each sample. To 10 μl of the annealed primer/RNA mixture, 12 μl of reverse transcription mixture (24 mM Tris-Cl, pH 8.3, 16 mM MgCl2, 0.8 mM each dNTP, 8 mM dithiothreitol, 125 ng actinomycin D/μl, and 1.5 units of avian myeloblastosis virus reverse transcriptase/μl) were added. The mixture was incubated at 45°C for 45 min to allow extension of the primer. The reaction was stopped by adding 8 μl of stop solution (100% formamide with 0.3% bromphenol blue and 0.3% xylene cyanol FF). The samples were boiled for 3 min and then loaded on either a 20 or 10% denaturing polyacryl-
wild type U4B gene was efficiently expressed in the oocyte assay, whereas the wild type U4X gene was not detectably expressed (Fig. 1A, lanes 2, 4, and 6). However, when the U4B gene 5'-flanking DNA was placed upstream of the U4X coding region, U4X RNA was efficiently synthesized (BBX, lanes 1 and 9). In contrast, the U4X gene 5'-flanking DNA was unable to drive efficient expression from the U4B coding region (XXB, lane 5). (An extremely low level of expression detectable from this construct suggests that the U4B gene coding region or 3'-flanking DNA makes a very minor contribution to the greater efficiency of U4B gene expression. However, since expression of the construct was very low, this observation has not been further investigated.)

More definitive conclusions were obtained from experiments employing the chimeric constructs that combined the U4B enhancer with the U4X proximal region, and vice versa. The U4B enhancer region was unable to activate U4X gene expression from a template containing the U4X proximal region (BBX, Fig. 1A, lane 3). In contrast, the U4X enhancer fully supported a high level of transcriptional activity from a template containing the U4X proximal sequence (XBX, cf. lanes 8 and 9 versus lane 7). Taken together, these results indicate that the U4X proximal region, not the enhancer, is the regulatory region primarily responsible for the inactivity of the U4X gene in frog oocytes.

We next examined the relative activities of the enhancer and proximal regions of the U4X and U4B genes in a different expression system. Four of the chimeric constructs that contain the U4X gene coding region were transfected into monkey CV-1 cells, total RNA was prepared, and the accumulation of U4X RNA was assayed by primer extension using a 17-mer primer complementary to the 3' end of U4X RNA. This primer was chosen because it does not cross-hybridize with any form of U4 RNA found in CV-1 cells (21, 22). As shown in Fig. 1B, the two constructs that contained the U4X proximal region, XXX and BXX, were not detectably expressed (lanes 1 and 3). In contrast, the two constructs that contained the U4B proximal region, BBX and XBX, were efficiently expressed (lanes 2 and 4).

Overall, the results in Fig. 1 indicate that the U4X proximal region is unable to support transcription in either frog oocytes or monkey CV-1 cells, whereas the U4B gene proximal region is an efficient promoter in both expression systems.

Base Alterations That Activate the U4X PSE—To further localize and identify the nucleotides responsible for the differential activities of the U4X and U4B promoters, we introduced a series of point mutations into the region near and within the U4X PSE that, in a stepwise fashion, made the U4X PSE more similar to the U4B PSE. The ability of these specific base alterations to rescue U4X expression was measured by transient transfection of CV-1 cells. As previously observed, a chimeric construct (XBX) that carried the wild type U4B proximal region between the U4X enhancer and the U4X coding region was efficiently expressed and served as the positive control in this experiment (Fig. 2, lane 1). A 3 bp change in the U4X PSE (at positions −49, −51, and −52) was not sufficient to activate U4X expression (Fig. 2, lane 2). The conversion of 8 additional bp between position −37 and −48, bringing them into agreement with the U4B sequence, resulted in a low level of U4X expression (lane 3). Interestingly, nearly full activity was obtained from a construct in which the two bases at positions −57 and −61 were changed to match the U4B sequence (lane 4). These two changes not only increased the similarity of the U4X PSE to the U4B PSE, but also destroyed the perfect dyad symmetry of the palindromic sequence that overlaps the U4X PSE. A marginal further increase in activity was observed from a construct that combined the entire set of mutations extending from −37 to −61 (lane 5). These results indicate that the nucleotides at positions −57 and −61 play an especially important role in determining the activities of the U4X and U4B promoters in the transient expression assay.
Fig. 3. Transcriptional activity of U4maxiB gene constructs having mutations within or near the PSE. Total RNA was prepared from transfected CV-1 cells and primer extension analysis was carried out using a U4maxiB specific primer. The sequence alignment at the bottom shows the alterations from the U4mB/AS sequence shown in the top line. The underlined regions in the U4mB/pal and U4mB/pal59 sequences indicate a 16-bp palindromic sequence identical to that found upstream of the U4X PSE. Expression levels relative to U4mB/AS (taken as 100%) are shown in the column to the right.

DNA sequences in the 3' portion of the PSE (−37 to −52) appear to make only a minor contribution toward the relative strengths of the U4B and U4X promoters.

Reciprocal Base Alterations That Inactivate the U4B PSE—To further investigate and confirm the importance of the bases at positions −61 and −57, we made the reciprocal changes in the U4B PSE. However, it is important to note that, when the U4X and U4B PSEs are aligned for maximum similarity, the U4B PSE is located two nucleotides nearer to the transcription start site than the U4X PSE. Thus positions −59 and −55 in the U4B PSE are the counterparts of positions −61 and −57 in the U4X PSE (note nucleotide numbering in Fig. 2). For the following studies, the parent construct was U4mB/AS which contains 17 nucleotides from the pUC19 polylinker inserted between positions 8 and 9 in the U4B coding region. This allowed the resultant U4maxiB product to be distinguished from the endogenous U4 RNA in primer extension assays. The U4mB/AS construct also has two point mutations at positions −64 and −48 that introduce Apal and SstI sites, respectively. These two point mutations have previously been shown to have no detectable effect on the activity of the U4B gene in an oocyte expression assay (24).

The U4mB/AS construct was efficiently expressed in CV-1 cells (Fig. 3, lanes 1 and 7). However, a single C to T point mutation at position −55 resulted in a 10-fold reduction in promoter activity (lane 2) and a single T to G change at position −59 resulted in a 6-fold reduction (lane 3). When the two point mutations were combined (lane 4), the residual level of expression was very low and was detectable only on long exposures or by PhosphorImager analysis.

Role of the Palindromic Sequence—We next investigated, in the context of the U4B promoter, the role of the gene-specific dyad symmetry element that normally overlaps the U4X PSE. Additional point mutations were introduced into the U4B proximal region at positions −62, −66, and −70. This construct, U4mB/pal, contains the entire 16-bp U4X palindromic sequence at a position in the U4B promoter homologous to its location in the U4X promoter. It overlaps the U4B PSE and includes the two strong down-mutations at positions −55 and −59. When this construct was transfected into CV-1 cells, there was no detectable expression (Fig. 3, lanes 5 and 8). Restoring the nucleotides at positions −55 and −59 back to the wild type (which consequently also destroys the dyad symmetry element), restored transcription to high levels (lanes 6 and 9).

In an attempt to divorce the effect of the nucleotide substitutions at positions −55 and −59 from potential regulatory effects of the palindromic sequence, a construct was made in which the palindrome was displaced 4 bp in the 5’ direction. In this construct (U4mB/pal59), the wild type U4B PSE sequence was maintained to position −63, and the complete palindromic sequence extended 16 bp upstream beginning at position −59. Fig. 3, lane 10, shows that this construct was expressed at a wild type level in CV-1 cells. Moreover, a construct that was mutated at six positions within the palindromic sequence was expressed at a similar level (lane 11). Thus, expression of the U4B gene in CV-1 cells was not significantly affected by either the presence or the absence of the palindromic sequence near, but not extensively overlapping, the U4B PSE.

DISCUSSION

We have investigated the cis-acting DNA sequences important for the differential expression of the U4X and U4B genes in frog oocytes and monkey CV-1 cells. The chicken U4B gene was efficiently expressed in both oocytes and CV-1 cells, whereas the U4X gene was not detectably expressed in either system. To localize the DNA sequences responsible for this differential activity, enhancer and promoter swap experiments were carried out using chimeric constructs. The results indicated that the proximal region of the U4X gene was inactive in both transient expression systems.

To define the specific nucleotides responsible for the difference in activity, point mutations were introduced into the U4X proximal region to determine whether specific changes (designed to match the sequence of the U4B PSE) could rescue U4X expression. Indeed, the alteration of two nucleotides at positions −57 and −61 was sufficient to raise the activity of the U4X PSE to a level comparable to that of the U4B PSE (Fig. 2). Moreover, some increased level of activity was alternatively obtained by replacing the U4X sequences between −52 and −37 with the corresponding nucleotides of the U4B PSE. These findings strongly suggest that the U4X gene contains a suboptimal PSE sequence, and that the nucleotides responsible for this effect are to some extent delocalized throughout the PSE. However, it is clear that the greatest down-effects on expres-
PSE-mediated Differential Expression of U4 snRNA Genes

A key finding from the presence of a G at position -61 and a T at position -57.

The importance of the bases of these two positions was confirmed by making the reciprocal changes at the homologous positions in the U4B PSE (positions -59 and -55). These two nucleotide changes resulted in 10- and 6-fold reductions in the U4B PSE activity, respectively, when the mutations were assayed individually, and when present together in the same construct, resulted in essentially complete loss of activity (Fig. 3). It is perhaps significant that none of the other cloned chicken U1, U2, or U4B genes contain either a G or a T at these respective positions (20). Our data localizing the critical nucleotides to the proximal region are reminiscent of studies of developmentally-regulated sea urchin U2 RNA genes performed by Stefanovic and Marzluff (25). In those studies, the sequences responsible for the differential expression of the early and late U2 genes was localized to a 14-bp region within the U2 PSE.

The two critical nucleotide positions identified above are not only located within the U4X PSE, but they also are members of a 16-bp dyad symmetry element that overlaps the U4X PSE. Thus, in addition to disrupting DNA recognition by the PSE-binding factor, mutations at these positions could also affect DNA recognition by another factor, proximal palindromic binding factor, that binds specifically to the palindromic element (23). In an attempt to separate and distinguish between these two potential effects, a construct was prepared that contained an intact palindromic sequence upstream of the U4B PSE (U4mB/pal59, Fig. 3). In this construct, the wild type U4B PSE was maintained, but the palindromic sequence was displaced 4 bp upstream relative to its homologous position in the wild type U4X proximal region. As shown in Fig. 3, the palindromic sequence at this position had no significant effect on expression of the U4mB transcript. Thus, in this transient expression system, U4 gene transcription was more closely correlated with the identities of the nucleotides at U4B positions -59 and -55 (U4X positions -61 and -57) rather than with the presence or the absence of the intact palindromic sequence. Nonetheless, it remains a possibility that the palindromic sequence in its natural chromosomal context in chicken cells may play a role in regulating the developmental expression of the U4X gene in vivo. Such effects may not be observable in the heterologous transient expression systems investigated here. Another possibility, not addressed by our experiments, is that U4X gene activity conceivably could be regulated in response to the methylation state of the multiple Cpg dinucleotides in the palindromic sequence (CGCCGGGGCGG) (26).

Acknowledgments—We thank Kathleen McNama Schroeder for excellent technical assistance, and J. Miyake for many helpful discussions throughout the course of this work.

REFERENCES

1. Dahlberg, J. E., and Lund, E. (1988) in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles (Birnstiel, M. L., ed) pp. 38–70. Springer Verlag KG, Heidelberg, Federal Republic of Germany
2. Parry, H. D., Scherly, D., and Mattaj, I. W. (1989) Trends Biochem. Sci. 14, 15–19
3. Reddy, R. and Singh, R. (1991) in Progress in Molecular and Subcellular Biology (Venter, P., Kuchino, Y., Muller, W. E. G., and Paine, P. L., eds) pp. 1–36, Springer-Verlag, Berlin
4. Hernandez, N. (1992) in Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R., eds) pp. 281-313, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
5. Ares, M., Jr., Chung, J.-S., Giglio, L., and Weiner, A. M. (1987) Genes Dev. 1, 808–817
6. Janson, L., and Pettersson, U. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4732–4736
7. Roebuck, K. A., Walker, R. J., and Stumph, W. E. (1987) Mol. Cell. Biol. 7, 4185–4193
8. Roebuck, K. A., Szeto, D. P., Green, K. P., Fan, Q. N., and Stumph, W. E. (1990) Mol. Cell. Biol. 10, 341–352
9. Zamrod, Z., and Stumph, W. E. (1990) Nucleic Acids Res. 18, 7323–7330
10. Cheung, C. H., Fan, Q. N., and Stumph, W. E. (1993) Nucleic Acids Res. 21, 281–287
11. Ach, R. A., and Weiner, A. M. (1991) Nucleic Acids Res. 19, 4209–4218
12. Weller, P., Bark, C., Janson, L., and Pettersson, U. (1988) Genes Dev. 2, 1389–1399
13. Waldschmidt, R., Wanandi, I., and Seifart, K. H. (1991) EMBO J. 10, 2595–2603
14. Wanandi, I., Waldschmidt, R., and Seifart, K. H. (1993) J. Biol. Chem. 268, 6629–6640
15. Murphy, S., Yoon, J.-B., Gerster, T., and Roeder, R. G. (1992) Mol. Cell. Biol. 12, 3247–3261
16. Sadowski, C. L., Henry, R. W., Lobo, S. M., and Hernandez, N. (1993) Genes Dev. 7, 1535–1548
17. Zamrod, Z., Tyree, C. M., Song, Y., and Stumph, W. E. (1993) Mol. Cell. Biol. 13, 5918–5927
18. Yoon, J.-B., Murphy, S., Bai, L., Wang, Z., and Roeder, R. G. (1995) Mol. Cell. Biol. 15, 2019–2027
19. Henry, R. W., Sadowski, C. L., Kobayashi, R., and Hernandez, N. (1995) Genes Dev. 9, 1389–1399
20. Hoffmann, M. L., Korf, G. M., McNamara, K. J., and Stumph, W. E. (1986) Mol. Cell. Biol. 6, 3910–3919
21. McNamara, K. J., and Stumph, W. E. (1989) Nucleic Acids Res. 17, 6748
22. Korf, G. M., Botros, I. W., and Stumph, W. E. (1988) Mol. Cell. Biol. 8, 5566–5569
23. Miyake, J. H., Botros, I. W., and Stumph, W. E. (1992) Gene Exp. 2, 161–173
24. McNamara, K. J., and Stumph, W. E. (1990) J. Biol. Chem. 265, 9728–9731
25. Stefanovic, B., and Marzluff, W. F. (1994) Gene Exp. 4, 1–18
26. Caoeres, J. F., McKenzie, D., Thimmapaya, R., Lund, E., and Dahlberg, J. E. (1992) Nucleic Acids Res. 20, 4247–4254

457

150

348

57.
Identification of Proximal Sequence Element Nucleotides Contributing to the Differential Expression of Variant U4 Small Nuclear RNA Genes
Hirdeypal S. Bhathal, Zulkeflie Zamrod, Takaaki Tobaru and William E. Stumph

J. Biol. Chem. 1995, 270:27629-27633.
doi: 10.1074/jbc.270.46.27629

Access the most updated version of this article at http://www.jbc.org/content/270/46/27629

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 23 references, 10 of which can be accessed free at http://www.jbc.org/content/270/46/27629.full.html#ref-list-1