

A U6 snRNA gene with an internal promoter is juxtaposed to an snRNP protein sequence within an intron of a human G protein gene

Eric D. Wieben, Anne M. Vrabel, Eileen L. Holicky, Ivana Klisak¹, Robert S. Sparkes¹ and David R. Stanford

Department of Biochemistry and Molecular Biology, Mayo Clinic/Foundation, Rochester, MN 55905 and ¹Department of Medicine, UCLA Center for the Health Sciences, UCLA, Los Angeles, CA 90024, USA

Received March 19, 1991; Revised and Accepted May 7, 1991

EMBL accession no. X54048

ABSTRACT

A complex locus on human chromosome 1 brings together sequences homologous to a G protein and two components of the RNA processing machinery of eukaryotic cells. Specifically, the seventh intron of the human G_{i3} α gene contains a fusion of a partial snRNP E protein pseudogene to a variant U6 snRNA gene. The novel U6 sequence contains nine point mutations and a one nucleotide deletion relative to the major U6 gene from humans. Unlike all other vertebrate U6 genes characterized to date, the variant U6 gene is efficiently transcribed by RNA polymerase III even in the absence of all natural flanking sequences. The union of elements from the signal transduction pathway and the RNA processing machinery suggests the possibility of functional interplay.

INTRODUCTION

Small nuclear ribonucleoprotein (snRNP) complexes containing snRNAs U1–U13 are integral components of the RNA processing machinery of eukaryotic cells (1–3). The spliceosomal snRNPs (those containing snRNAs U1, U2, U5, and U4/U6) have each been shown to be required for efficient excision of introns from primary gene transcripts (4). The protein components of the spliceosomal snRNPs have been studied extensively in the past few years, and it is now clear that the U1, U2, U5, and U4/U6 snRNPs each have a distinctive complement of snRNP proteins. However, all of these snRNPs do share a subset of common snRNP proteins, including a 'core' of small proteins that have been designated D, E, F, and G (5).

The U4/U6 snRNP is unique among this group of snRNPs that are required for mRNA splicing in that it contains two base-paired snRNAs (6,7). Additionally, the U6 component of the U4/U6 snRNP is atypical in that it has a unique cap structure, lacks the binding site for the common snRNP proteins, and is synthesized by RNA polymerase III rather than RNA polymerase II (8–10). Surprisingly, analysis of the regulation of transcription of the major U6 genes in mammals has demonstrated that sequence elements **upstream** of U6 genes are both necessary and sufficient

for correct initiation of transcription (11,12). Thus, the internal promoter elements common to many other polymerase III transcripts do not appear to play an important role in the expression of the most abundant U6 RNAs found in mammalian cells.

Further interest in U6 has been stimulated by studies of U6 genes in yeast. Sequence analysis reveals that U6 is the most conserved spliceosomal snRNA through evolution, suggesting that it must have a particularly crucial role in eukaryotic cells (4). Particularly interesting is the finding that the U6 gene from the fission yeast, *Schizosaccharomyces pombe*, contains an intron (13). This finding has led to the speculation that U6 might be directly involved in catalysis during splicing. The discovery of a transcriptionally-active variant U6 gene within a G protein intron raises the possibility that U6 snRNAs might have additional roles in gene regulation.

MATERIALS AND METHODS

Cloning, sequencing, and mapping of LH87

Genomic clone LH87 was isolated from a human genomic library as described previously (14). The human genomic library was purchased from Clontech Laboratories, Inc. (Palo Alto, CA) and was generated from a partial MboI digest of human leukocyte DNA. Clone LH87 and several other independent isolates were subcloned into plasmid and M13 vectors for restriction and sequence analysis. Sequencing was as described previously (14,15). The in situ hybridization data were collected and analyzed as described (16).

In Vitro transcription

U6 transcription reactions were carried out in HeLa cell S100 extracts prepared according to Weil *et al.*, (17). Transcription reactions were assembled as described by Kunkel *et al.* (8), with minor modifications. Specific reaction conditions are given in the figure legends. Uncut plasmid DNA (20 μ g/ml unless otherwise noted) was used as template. Reactions were incubated at 30°C for 90 min. Transcription products were purified by phenol extraction and either analyzed directly by electrophoresis

and autoradiography or further purified by hybridization selection as described (18).

In Vitro mutagenesis

A pair of oligonucleotides (5'-cctgcagaattcGTGCTTGCTTTG-GTAG-3' and 5'-ccccggaattcAAAAAATATGGAACGC-TTC-3') were used as primers for amplification of a DNA fragment corresponding precisely to the 87U6 gene. The amplified DNA fragment was cut with EcoRI and cloned into pBluescribe- and pUC18 to generate 87U6 clones lacking human flanking sequences.

RESULTS

Genome structure

While screening a human genomic DNA library for sequences homologous to one of the common snRNP proteins, the 11,000 dalton E protein (15), we isolated a group of clones that hybridized only to probes from the 3' end of the E protein cDNA. Sequencing studies on this group of genomic clones revealed the presence of a putative U6 snRNA pseudogene immediately upstream of DNA sequences coding for the 3' terminal 213 nucleotides of the E protein mRNA. Further sequence analysis indicated that this unusual combination of sequences was located in an intron of the human gene for a signal transduction G protein, G_i3 α (Fig. 1A).

LH87 is one of the group of 23 independent isolates containing this gene arrangement. All have been restriction mapped, and four of the 23 clones have been analyzed at the DNA sequence level. Since the organization of the U6, E protein and G protein sequences is the same in all clones, further discussion will be limited to the prototypic clone of this group, LH87.

The E protein homologous region of LH87 includes the last 64 nucleotides (nts) of protein coding sequence, and contains eleven mismatches relative to the parental sequence (Fig. 1C). The U6 sequences found immediately upstream of the E protein pseudogene sequence (87U6) have 91% homology to an expressed human U6 gene, pGem/U6 (8). Half of the ten differences between the U6 sequences are clustered in the middle of the U6 coding region, which is thought to be important for interaction of U6 with U4 snRNA (19).

There is some limited similarity between the 5' flanking sequence of 87U6 and the human U6 gene isolated by Kunkel *et al.*, (8) including the presence of a 12 nt purine-rich sequence between 150 and 200 nts upstream of the transcription start site in both U6 genes (Fig. 1C, asterisks). However, none of the upstream transcriptional control elements (the octamer motif, proximal sequence element (PSE) and TATA box) found in the major vertebrate U6 genes are present in the immediate 5' flanking region of 87U6. Furthermore, there are no similarities between the 3' flanking regions of 87U6 and those of previously characterized mammalian E protein or U6 genes.

Given the preponderance of pseudogenes in both the E protein gene family and genes for snRNAs, it is noteworthy that pairs of short direct repeats flank both the U6 sequence (8/9 or 10/13), and the entire U6/E protein assemblage (13/14). There is also a perfect 9/9 direct repeat that defines the boundaries of the E protein pseudogene region. Although this complex pattern of repeats obscures the evolutionary history of this unusual locus, the sequence data are consistent with the hypothesis that the 87U6 and E protein segments were inserted into the G_i3 α intron as

a single transposition event. If the absence of flanking sequence homologies and the poly(A) stretch at the end of the E sequence are interpreted to suggest an RNA-mediated transposition, this hypothesis would imply that at some point there was a single transcript containing both U6 and E protein sequences. Since U6 has been proposed to be active at the catalytic site of splicing reactions, this arrangement of sequences might reflect a transient splicing intermediate that involved covalent attachment of U6 to an E protein transcript. In this regard, it is curious that the one base deletion in 87U6 relative to the consensus human U6 sequence occurs precisely at the site of the intron in the U6 gene of *S. pombe*. However, this hypothesis would not explain the structure of the U6-E protein junction, which is near the splice junction of exons 4 and 5 in the E gene (14), but includes the last 11 nts of the 79 nts in exon 4.

The E protein and 87U6 sequences are located within the seventh intron of the G_i3 α gene (Fig. 1A). As depicted in Figure 1A, the G protein gene runs in the opposite orientation from both the 87U6 gene and E protein pseudogene. Exon 8 of the G_i3 α protein begins at nucleotide position -404 and ends at position -616 (Fig. 1B). This exon and the immediate flanking intron regions are identical to that reported by Itoh *et al.* (20), (Fig. 1B). The possibility that the LH87 group of clones represents an unusual cloning artifact peculiar to a single genomic library has been excluded by Southern blot analysis of one of the original G_i3 α clones isolated by Itoh *et al.* (20). These studies confirmed the presence of E protein and U6 sequences in a 622 nt EcoRI fragment of clone β HGi3-81 (data not shown).

A portion of the remaining 87U6 5' upstream sequence that we have analyzed extends beyond the limits of the published genomic sequence for G_i3 α , but does match that of human cDNAs for G_i3 α (21-23). This similarity defines the beginning of exon 9, which is separated from exon 8 by a 95 nucleotide intron (Fig. 1B). Exon 9 contains a single mismatch in the 3' untranslated region compared to the cDNA sequence reported by Codina *et al.* (21), and maintains homology to the longer G_i3 α cDNA sequenced by Didsbury and Snyderman (22) throughout the region we have sequenced.

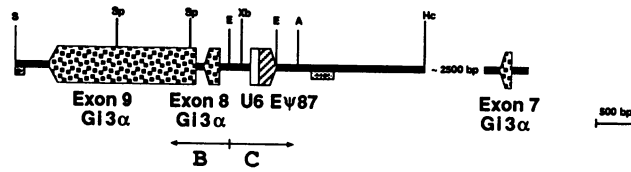
Mapping

The unique functional E protein gene is located on chromosome 1 at 1q25-q43, (16) and previous efforts at mapping G proteins have also localized the human G_i3 α gene to an undetermined locus on chromosome 1 (24). Although it seems likely that an RNA intermediate was involved in the dispersion of the E pseudogene sequence in the human genome, the sequence data do not exclude the possibility that the LH87 locus overlaps that of the functional E protein gene.

To evaluate the hypothesis that the E protein sequence in LH87 represented an alternative exon of the E protein gene and to clarify the relationship of the G_i3 α sequence in LH87 to previously characterized G protein genes, we used *in situ* hybridization to map the location of the LH87 sequence in the human genome.

An 880 nucleotide SphI-SphI fragment from exon 9 of the G_i3 α gene (Fig. 1A) was used to make a nick-translated probe for *in situ* hybridization. The results of the *in situ* hybridization are shown in Figure 2A and 2B. The LH87 probe mapped to a region near the centromere of chromosome one, at 1p13-21 (Fig. 2B). As depicted in Figure 2B, this places LH87 on the opposite arm of chromosome 1 from the expressed E protein gene, and rules out the possibility that the E protein homologous sequences are a functional part of the major E protein

A. LH87



B. G₁₃α

Sequence alignment for G₁₃α protein exon 8 and exon 9. The top section shows the LH87 genomic DNA sequence for intron 7 and exon 8. The middle section compares LH87 exon 8 with human cDNA sequences Humgip3a8, Humgtpbp, and Humgkas. The bottom section compares LH87 exon 9 with Humgtpbp, Humgkas, and Humgiab. Nucleotide positions are indicated on the left and right of each sequence block.

C. 87U6-Eψ87

Sequence alignment for the 87U6-Eψ87 region. The top section shows the OCTAMER region with Human U6, Mouse U6, and 87 U6 sequences. The middle section shows the U6 Codine Region with Human U6, Mouse U6, and 87 U6 sequences, including a diagram of the A box. The bottom section shows the E Protein Coding Region with Human U6, Mouse U6, and E cDNA sequences, including a diagram of the B box? and intron 4. Nucleotide positions are indicated throughout the alignment.

Figure 1.A. Sequence organization of genomic clone LH87. The solid line represents genomic DNA. Stippled boxes represent exons for the G₁₃α protein, the cross-hatched box represents the E protein pseudogene, and the open box represents the 87U6 gene. The portion of genomic DNA representing the G₁₃α protein exon 7 is not present in clone LH87. The shaded boxes containing arrows represent Alu repeats. Restriction enzymes are: A = *Acc* I; E = *Eco* RI; Hc = *Hinc* II; S = *Sma* I; Sp = *Sph* I; Xb = *Xba* I. **B.** Sequence comparison of clone LH87 with G₁₃α protein nucleotide sequences. Humgip3a8 is the sequence presented by Itoh *et al.* (20), for the human G₁₃α protein exon 8 and immediate flanking sequences. Humgtpbp is sequence of a human cDNA for a G₁₃α protein (23). Humgkas is cDNA sequence for a human G_k protein (21). Humgiab is a cDNA sequence for a G₁ protein from human lymphocytes (37). Nucleotide number 1 is located at the start of the U6 gene while nucleotide -1 is the first nucleotide upstream of the U6 gene. A (-) indicates identity with LH87 while a (.) indicates a deletion. Lower case letters indicate intron sequences. The amino acid sequence of the carboxyl-terminus of the G₁₃α protein is shown above the nucleotide sequence. **C.** Sequence comparison of clone LH87 with U6 and E protein sequences. Human (8) and mouse (12) U6 genes have been aligned above the U6 gene from clone LH87, while the E protein cDNA sequence is aligned below the sequence of LH87. Direct repeats flanking the U6/E protein pseudogene region and other sequence features mentioned in the text are indicated.

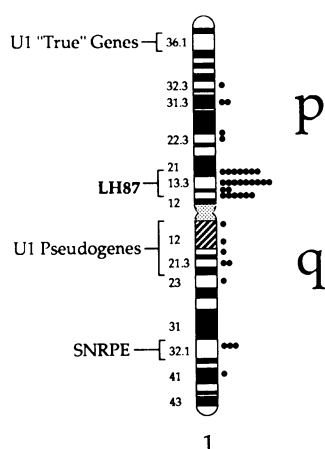


Figure 2. Regional location of LH87 on chromosome 1. The location of other snRNP components are also indicated. SNRPE is the functional gene for the snRNP E protein. The closed circles depict a summary of the localization of silver grains on chromosome 1. No more than 3 grains were found at any other locus in the genome.

transcription unit. These data also provide the first regional localization of the $G_3\alpha$ gene on the human genome.

Transcription of the 87U6 gene *in vitro*

To determine if the variant U6 gene found in LH87 could be expressed despite the lack of homology to known U6 regulatory sequences, a construct containing 87U6 plus 1.4 kilobases (kb) of 5' and 1 kb of 3' flanking sequences (p87 Hc) was incubated in a HeLa cell S100 extract capable of supporting transcription by RNA polymerase III (17). As shown in lane 3 of Figure 3, the LH87 construct directed the synthesis of a labeled product that co-migrated with the U6 RNA produced from another human U6 gene, pGem/U6 (lane 1; ref. 8). The RNA synthesized from both the consensus and variant U6 templates specifically hybridized to single stranded DNA containing sequence complementary to human U6 RNA, but failed to hybridize to single stranded DNA complementary to U2 snRNA. Vector sequence alone did not yield any products that are recovered by hybrid selection (lanes 5 and 6). These results suggest that the variant U6 gene from LH87 can be efficiently and accurately transcribed *in vitro*.

To further investigate the regulation of transcription of the variant 87U6 gene, increasing amounts of α -amanitin were added to the *in vitro* transcription reactions. As shown in Figure 3B (lane 3), the transcription of 87U6 was not affected by α -amanitin at a concentration of 1 μ g/ml, but transcription was abolished when the concentration of α -amanitin was increased to 200 μ g/ml (lane 4). These results are consistent with those obtained for the major U6 genes from human and mouse, and suggest that the variant 87U6 gene was being transcribed by RNA polymerase III in the HeLa extract.

87U6 Transcription in the absence of 5' flanking sequence

The results shown in Figure 3 suggest that the efficiency of transcription of 87U6 by RNA polymerase III *in vitro* was similar to that of pGem/U6, even though 87U6 lacks all known U6 transcriptional regulatory sequences. These findings establish that the transcription of these two U6 genes relies upon different sequence elements, even though they are transcribed by the same

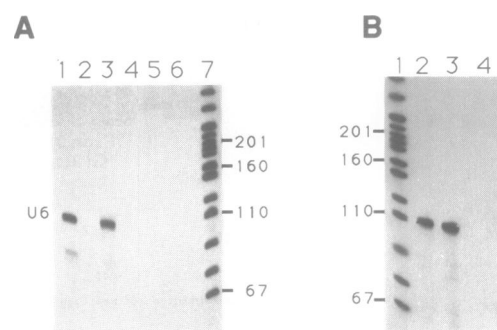


Figure 3. *In vitro* transcription of a variant U6 gene. Following phenol extraction and ethanol precipitation, transcription products were hybrid selected on filters containing single stranded M13 clones with inserts complementary to either U6 (lanes 1, 3, and 5) or U2 (lanes 2, 4, and 6) snRNAs (18). Final reaction volumes were 50 μ l, final concentrations of ATP, UTP, and CTP were 0.4 mM, and the final concentration of GTP was 20 μ M. A 30 μ l aliquot of HeLa cell S100 extract and 10 μ Ci of 32 P-GTP were used for each reaction. Template DNAs for the transcriptions were pGem/U6 (lanes 1 and 2), p87Hc, (lanes 3 and 4), or vector only (Bluescribe-; lanes 5 and 6). Lane seven contains 32 P-labeled DNA markers. **B.** *In vitro* transcription of p87-112Hc in the presence of α -amanitin. *In vitro* transcription reactions were supplemented with 0 (lane 2), 1 (lane 3), or 200 μ g/ml (lane 4) α -amanitin. Transcription products were further purified by hybridization selection on M13-U6 filters as described above. Lane 1 contains 32 P-labeled DNA markers.

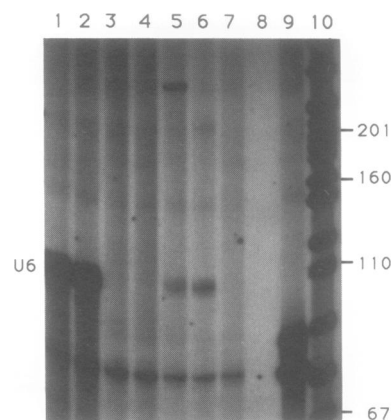


Figure 4. Transcription efficiency of 87U6 in the absence of human flanking DNA. Lanes 1-4: Mouse U6 constructs mU6 -89/+84, mU6 -79/+84, mU6 -23/+84, and mU6 6/84 (12). Lane 5: p87U6 1/107, Lane 6: Human U6 gene pGem/U6 (including 470 nts of 5' flanking sequence; ref. 11). Lane 7: vector only (Bluescribe -) Lane 9: *Xenopus laevis* tRNA^{Met} gene (pXlmet1; ref. 38). Lane 10: α - 32 P-labeled DNA markers. Lane 8 is blank. Final transcription reactions for this experiment contained 0.6 mM ATP, UTP and CTP, 0.025 mM cold GTP, 1 mM dithiothreitol, 6 mM creatine phosphate, 20 mM KCl, 10 mM Tris-HCl, pH 8.0, 10 μ Ci α - 32 P-GTP, 2 μ g of plasmid DNA, and 10 μ l HeLa S-100 in a final volume of 20 μ l. Products of *in vitro* transcriptions were purified by phenol extraction and ethanol precipitation and then directly analyzed by electrophoresis and autoradiography.

polymerase. Since many polymerase III genes utilize internal promoters rather than elements in flanking sequences, we used the polymerase chain reaction (PCR) to produce plasmids that contained the entire coding region of 87U6 but lacked any human 5' or 3' flanking sequences. As shown in lane 5 of Figure 4, the minimal 87U6 constructs continued to be suitable templates for transcription *in vitro*, yielding approximately as much U6-sized product as the complete pGem/U6 with all known 5'

regulatory elements (lane 6). Similar results were obtained with minimal 87U6 constructs cloned in both directions in two different vectors (pUC 18 and pBS-; data not shown). Some of the 87U6 minimal PCR constructs also yielded a higher molecular weight product as shown in lane 5. This larger band was not selected on filters containing U6 DNA, and has not been characterized further.

Transcription of both human U6 genes was considerably less than that obtained with two constructs containing a murine U6 gene, mU6 -89/+84 and mU6 -79/+84 (ref. 12; lanes 1 and 2, Figure 4). Both of the functional mouse U6 constructs contained sufficient 5' flanking sequence to support transcription by polymerase III in our HeLa cell extracts. However, in agreement with the results of Das *et al.* (12), no U6 RNA was synthesized from two mouse U6 constructs that contain lesser amounts of 5' flanking sequences, mU6 -23/+84 and mU6 6/+84 (lanes 3 and 4). **Thus, the transcription of the minimal 87U6 construct cannot be attributed simply to the inability of our extracts (or incubation conditions) to discriminate between active and inactive U6 templates.**

The low molecular weight band that is detectable in lanes 3 and 4 was also present in every other lane, including the vector only control (lane 7). Since this labeled species was also found when template was omitted from the incubation (not shown), it probably represents end-labeling of an RNA present in the extract by the α -³²P-GTP.

DISCUSSION

External vs. internal U6 promoters

The variant U6 gene in LH87 is unique among all previously characterized vertebrate U6 genes in that its transcription by RNA polymerase III *in vitro* is not dependent upon flanking DNA sequences. In fact, transcription of previously characterized mammalian U6 genes requires only the initiating G nucleotide of the U6 coding sequence and the PSE and TATA elements found in the immediate 5' flanking sequence (11,12,25,26). Elimination of the remainder of the U6 coding sequence does not influence the efficiency of transcription by RNA polymerase III *in vitro*. The 87U6 gene represents the opposite extreme. Following the paradigm of the 5S RNA and tRNA-type RNA polymerase III transcription units, all required promoter elements reside **within** the 87U6 coding sequence. Thus, one or more of the ten mutations that distinguish 87U6 from the major human U6 gene is responsible for converting an entirely external RNA polymerase III promoter into an entirely internal promoter.

Although we have not yet mapped the internal sequences responsible for the function of 87U6, it seems unlikely that this feature of 87U6 arose by chance during divergence of this gene copy from a parental U6 gene that required flanking sequences for expression. It is more likely that some (or all) of the sequence differences between 87U6 and the wild type U6 were necessary for the function of an ancestral, internally controlled U6 gene. In this regard, it is interesting that 3 of the 10 sequence differences between 87U6 and the consensus human U6 reside within the previously noted homology to the A box of tRNA genes. While two of these changes are neutral, one of these sequence changes (G to A at position 57) does improve the homology to the A box consensus (Fig. 5, top).

In contrast to the typical tRNA-like polymerase III transcription unit, the coding region of the 87U6 gene does not contain a recognizable homology to the tRNA B box. Studies of tRNA gene

function suggest that this element is not absolutely required for RNA polymerase III function (27), and the 87U6 coding region may be an example of a polymerase III transcription unit that lacks a functional B box.

However, it may be significant that there is a good match to the B box consensus within the adjacent E protein sequence (Fig 5, bottom and Fig. 1C). The B box-like sequence embedded in the E protein sequence begins 70 nts downstream of the U6-E protein junction. We have not directly tested the influence of the E protein segment on transcription of 87U6. It is clear that the E sequences are **not** absolutely required for *in vitro* transcription of the 87U6 gene, since constructs lacking these sequences yield as much U6 sized product as the wild type human U6 gene. However, as evidenced by the influence of a downstream B box on the transcription of a U6 gene in *S.cerevisiae* (28), the external location of the B box sequences does not exclude the possibility of interaction with other RNA polymerase III promoter elements.

Thus, while only the internal A box might be required to obtain basal expression of 87U6, the B box homology in the 3' flanking region might influence the rate of expression of this gene *in vivo*. Selective pressure based upon an influence of the E sequences on 87U6 function would also provide a plausible explanation for why the 87U6 and E protein sequences at this locus have diverged only minimally from the parent sequences.

Speculations on the origin and function of the unusual sequence arrangement in LH87

The additional factor that must be considered in discussions of the evolution of this locus is the location of the 87U6 and E protein sequence in the G₃ α intron. Although the linkage of snRNP protein, U6 and G protein sequences has not been reported before, an association between the U4/U6 snRNP and a G-protein related polypeptide (PRP4) has been proposed in yeast (29-31).

Possible interactions between the 87U6 and the E pseudogene sequences at the transcriptional level have been discussed above. A more important biological question is whether the concurrent expression of the G₃ α gene and the 87U6 gene might be involved in the regulation of expression and/or function of one or both products.

There is little doubt that the G₃ α gene at this locus is functional. Transcription of this gene will of necessity produce a primary gene transcript that contains sequences capable of base pairing to both cellular U6 RNA and E protein mRNA. Thus, one possibility is that there might be antisense interactions between the RNA processing machinery and the G₃ α transcripts. This hypothesis would predict an inverse relationship between the levels of G₃ α transcription and the levels of U6

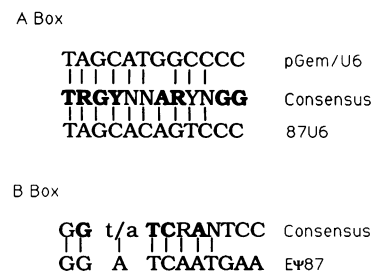


Figure 5. Comparison of 87U6 and E pseudogene sequences to consensus A box and B box sequences from tRNA genes. Invariant residues in the consensus sequences for RNA polymerase III A and B boxes (26,39) are shown in bold print.

RNA and E protein in human cells. With this possibility in mind, it is interesting that most cells seem to contain a molar excess of U6 over U4, and that metabolically distinguishable subpopulations of U6 RNA have been described in human cells (32,33). The possibility that U6 RNAs not found in complete U4/U6 snRNPs might have additional roles in cellular regulation has not been addressed experimentally.

It is also possible that the variant 87U6 gene product could be involved in the splicing of the $G_3\alpha$ transcript. Two of the ten differences that distinguish between 87U6 and the consensus human U6 sequences occur at positions that have been absolutely conserved during evolution (4,34). The consensus C at position 11 is a U in 87U6, and the conserved U at position 64 is a C in 87U6 (thereby conserving the run of four C's at this location). However, neither of these changes occurs at sites that have been demonstrated experimentally to be important for viability, assembly with U4, or spliceosome formation. In fact, mutational analysis in yeast has established that substitution of the conserved C at the 5' end of U6 does not compromise viability (34,35). The phenotype produced by mutation of the conserved U in the U4 interaction domain has not been reported, but none of the changes in 87U6 in this region would be predicted to abolish the proposed base pairing interactions between U4 and U6 (4,19). Thus, there is no reason to believe that 87U6 RNA could not be assembled into a functional U4/U6 snRNP.

Discussion of possible roles for the 87U6 product in cellular regulation makes the assumption that the 87U6 gene is expressed *in vivo*. The evidence presented here demonstrates that the 87U6 gene contains a functional promoter. Since the $G_3\alpha$ gene is actively expressed in a variety of tissues, the 87U6 gene should be accessible to the transcriptional machinery and is likely to be expressed in human cells *in vivo*. While fingerprinting of the major U6 from HeLa cells suggests that the 87U6 product is not a predominant product of these cells (8), there is no evidence that excludes the possibility that other forms of U6 are expressed at low levels in HeLa cells or at high levels in other human cells.

Assessment of the relative contribution of each of the 200 or more U6-related sequences in the human genome (36) to the total pool of human U6 RNA will require further studies of U6 gene structure and function.

ACKNOWLEDGMENTS

We would like to thank Drs. G.Kunkel and R.Reddy for advice and for the gifts of m13 and plasmid clones containing consensus human and mouse U6 genes and the control template tRNA gene. We would also like to thank Dr. Yoshito Kaziro for the gift of the $G_3\alpha$ clone HGi3-81 and Dr. Birgit Knerer for producing some of the PCR constructs of 87U6. Finally, we thank Drs. T.Pederson, J.Calvet, and P.Camacho for stimulating discussions, and Dr. K.Neiswanger for coordinating the mapping studies. This work was supported by grant AR38554 from the National Institutes of Health and in part by NSF grant DMB 8505716 and the Mayo Foundation.

REFERENCES

1. Tyc, K., and Steitz, J.A. (1989) *EMBO J.* **8**, 3113–3119
2. Montzka, K.A., and Steitz, J.A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8885–8889
3. Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Kramer, A., Frendewey, D., and Keller, W. (1988) in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, ed. Birnstiel, M.L., (Springer-Verlag, New York) pp. 115–154
4. Guthrie, C., and Patterson, B. (1988) *Ann. Rev. Genet.* **22**, 387–419
5. Luhrmann, R. (1988) in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, ed. Birnstiel, M.L., (Springer-Verlag, New York) pp. 71–99
6. Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H., and Luhrmann, R. (1984) *EMBO J.* **3**, 1357–1363
7. Hashimoto, C., and Steitz, J.A. (1984) *Nucleic Acids Res.* **12**, 3283–3293
8. Kunkel, G.R., Maser, R.L., Calvet, J.P., and Pederson, T. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8575–8579
9. Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D. (1987) *J. Biol. Chem.* **262**, 75–8
10. Singh, R., and Reddy, R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8280–8283
11. Kunkel, G.R., and Pederson, T. (1988) *Genes Dev.* **2**, 196–204
12. Das, G., Henning, D., Wright, D., and Reddy, R. (1988) *EMBO J.* **7**, 503–512
13. Tani, T., and Ohshima, Y. (1989) *Nature* **337**, 87–90
14. Stanford, D.R., Perry, C.A., Holicky, E.L., Rohleder, A.M., and Wieben, E.D. (1988) *J. Biol. Chem.* **263**, 17772–17779
15. Stanford, D.R., Kehl, M., Perry, C.A., Holicky, E.L., Harvey, S.E., Rohleder, A.M., Rehder, K., Luhrmann, R., and Wieben, E.D. (1988) *Nucleic Acids Res.* **16**, 10593–10605
16. Neiswanger, K., Stanford, D.R., Sparkes, R.S., Nishimura, D., Mohandas, T., Klisak, I., Heinzmann, C., and Wieben, E.D. (1990) *Genomics* **7**, 503–508
17. Weil, P.A., Segall, J., Harris, B., Ng, S.Y., and Roeder, R.G. (1979) *J. Biol. Chem.* **254**, 6163–6173
18. Wieben, E.D., Rohleder, A.M., Nenninger, J.M., and Pederson, T. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7914–7918
19. Bindereif, A., Wolff, T., and Green, M.R. (1990) *EMBO J.* **9**, 251–255
20. Itoh, H., Toyama, R., Kozasa, T., Tsukamoto, T., Matsuoka, M., and Kaziro, Y. (1988) *J. Biol. Chem.* **263**, 6656–6664
21. Codina, J., Olate, J., Abramowitz, J., Mattera, R., Cook, R.G., and Birnbaumer, L. (1988) *J. Biol. Chem.* **263**, 6746–6750
22. Didsbury, J.R., and Snyderman, R. (1987) *FEBS Lett.* **219**, 259–363
23. Kim, S., Ang, S-L., Bloch, D.B., Bloch, K.D., Kawahara, Y., Tolman, C., Lee, R., Seidman, J.G., and Neer, E.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4153–4157
24. Blatt, C., Eversole-Cire, P., Cohn, V.H., Zollman, S., Fournier, R.E.K., Mohandas, L.T., Nesbitt, M., Lugo, T., Jones, D.T., Reed, R.R., Weiner, L.P., Sparkes, R.S., and Simon, M.I. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7642–7646
25. Kunkel, G.R., and Pederson, T. (1989) *Nucleic Acids Res.* **17**, 7371–7379
26. Lescure, A., Carbon, P., and Krol, A. (1991) *Nucleic Acids Res.* **19**, 435–441
27. Geiduschek, E.P., and Tocchini-Valentini, G.P. (1988) *Ann. Rev. Biochem.* **57**, 873–914
28. Brow, D., and Guthrie, C. (1990) *Genes Dev.* **4**, 1345–1356
29. Banroques, J., and Abelson, J.N. (1989) *Molec. Cell. Biol.* **9**, 3710–3719
30. Bjorn, S.P., Solyk, A., Beggs, J.D., and Friesen, J.D. (1989) *Molec. Cell. Biol.* **9**, 3698–3709
31. Dalrymple, M.A., Petersen-Bjorn, S., Friesen, J.D., and Beggs, J.D. (1989) *Cell* **58**, 811–812
32. Rinke, J., and Steitz, J.A. (1985) *Nucleic Acids Res.* **13**, 2617–2629
33. Choudhury, K., Choudhury, I., Jones, R.W., Thirunavukkarasu, C., and Eliceiri, G.L. (1988) *J. Cell. Physiol.* **137**, 529–536
34. Madhani, H.D., Bordonne, R., and Guthrie, C. (1990) *Genes Dev.* **4**, 2264–2277
35. Fabrizio, P., and Abelson, J. (1990) *Science* **250**, 404–409
36. Hayashi, K. (1981) *Nucleic Acids Res.* **9**, 3379–3388
37. Beals, C.R., Wilson, C.B., and Perlmutter, R.M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7886–7890
38. Clarkson, S.G., Kurer, V., and Smith, H.O. (1978) *Cell* **14**, 713–724
39. Galli, G., Hofstetter, H., and Birnstiel, M.L. (1981) *Nature* **294**, 626–631