Small Nuclear RNA Genes: a Model System to Study Fundamental Mechanisms of Transcription*

Published, JBC Papers in Press, June 4, 2001, DOI 10.1074/jbc.R100032200

Nouria Hernandez‡
From the Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

The human small nuclear RNA (snRNA)1 genes, which encode snRNAs that are involved in RNA processing reactions such as mRNA splicing, serve as prototypes for a family of genes whose promoters are characterized by the presence of a proximal sequence element (PSE) and a distal sequence element (DSE). From a transcription point of view, this family of genes is highly interesting because all of its members have very similar promoters, even though some of them are transcribed by RNA polymerase (pol) II and others by pol III. As a result, the snRNA genes have served as a model system to explore how RNA polymerase specificity is determined and, in general, to compare the pol II and III transcription machineries. This has led to the concept that the pol II and III transcription machineries use common factors, the best known of which is the TATA box binding protein (TBP). In addition, the relative simplicity of these promoters has also made them an attractive system to study how transcriptional activators perform their function.

The Structure of snRNA Promoters

Fig. 1 shows the structures of snRNA promoters from Homo sapiens (Hs), Arabidopsis thaliana (At), and Drosophila melanogaster (Dm) and serves to illustrate the remarkable fact that although snRNA promoters have diverged during evolution, the close similarity between those recognized by pol II and those recognized by pol III has been conserved. In fact, in each of the examples in Fig. 1, RNA polymerase specificity can be changed by altering a single parameter, indicated in red on the figure.

In the human genes, the U1 and U2 snRNA promoters serve as the prototypic pol II snRNA promoters, and the U6 snRNA promoter serves as the prototypic pol III snRNA promoter (see Ref. 1 for a review). The human pol II snRNA core promoters contain only one essential element, the PSE, whereas the pol III snRNA core promoters consist of two elements, the PSE and a TATA box located at a fixed distance downstream. The DSE serves to enhance transcription from the core promoter. Both the DSE and the PSE can be interchanged between pol II and III snRNA promoters with no effect on RNA polymerase specificity, which is determined by the presence or absence of the TATA box. The A. thaliana pol II and III snRNA promoters contain an upstream sequence element (USE) and a TATA box, which are both interchangeable between the pol II and III snRNA promoters. RNA polymerase specificity is determined in this case by the exact spacing between the USE and the TATA box, which is 33–34 base pairs (bp) and 23–24 bp in the pol II and III snRNA promoters, respectively (2).

The D. melanogaster pol II snRNA promoters contain two elements referred to as the PSEA and the PSEB spaced by 8 bp, and the pol III snRNA promoters contain a PSEA and a TATA box spaced by 12 bp. The PSEA is quite conserved in various pol II and III snRNA promoters, but positions 19 and 20 of the 21-bp elements are always g/a in the pol II and T in the pol III snRNA promoters. RNA polymerase specificity is determined by the precise sequence of the PSEA element, with the base pairs at positions 19 and 20 playing a major role (Ref. 3, and references therein).

The number of snRNA promoters have been characterized in various sea urchins. As in other species, the pol II and III snRNA promoters are closely related in structure. They all have a PSE and some have a TATA box, but the presence of the TATA box does not correlate with RNA polymerase specificity. The PSEs in different snRNA promoters show little sequence identity and yet can be exchanged with no effect on polymerase specificity. The determinants of RNA polymerase specificity are not known (Ref. 4, and references therein). In Saccharomyces cerevisiae, only the pol III U6 snRNA promoter has been studied. It consists of a TATA box located upstream of the transcription start site and A and B boxes typical of gene-internal tRNA promoters. The A box is located, as in tRNA genes, within the RNA coding region, but the B box is located at an anomalous position 3’ of the gene (5–7).

The PSE Binding Factors

The snRNA promoters in vertebrates, A. thaliana, D. melanogaster, and sea urchins all contain an element, variously called the PSE, PSEA, or USE, centered 50–70 bp upstream of the transcription start site. The factor binding to this element has been best characterized in the human system and is variously known as PBP, PTF, or SNAP, SNAPc. It is a complex containing five types of subunits, SNAP190, SNAP50, SNAP45, PTF (PTF1, PTF2, PTF3, PTF4, and PTF5), and SNAP19 (see Ref. 8, and references therein). SNAP190 forms the backbone of the complex, with SNAP19 and SNAP45 associating toward the N and C terminus, respectively, of the molecule. SNAP43 can associate with the same region of SNAP190 as SNAP19, and SNAP50 joins the complex by associating with SNAP43 (see Fig. 3 for an illustration of SNAP190).

The ability to assemble recombinant SNAPc, and thus mutant forms of SNAPc, has allowed a study of the role of SNAPc subunits for binding to DNA and for basal and activated transcription by RNA polymerases II and III. The smallest subassembly of SNAPc subunits tested that binds to the PSE with the same specificity as the complete complex consists of SNAP190 aa 84–505, SNAP43 aa 1–268, and SNAP50 (9). This observation is consistent with UV cross-linking experiments that suggest that within SNAPc, both SNAP190 and SNAP50 are in close contact with the DNA (see Ref. 8, and references therein). The specific binding of SNAPc to the PSE is mediated in part by an unusual Myb domain extending from aa 263 to 503 within SNAP190 and containing a half-repeat followed by four repeats (8, 10).

In the human system, the very same SNAPc is involved in transcription by pol II and pol III (11). The polypeptide composition of PSE binding factors from other species is unknown, but there are indications that at least in some cases, the same PSE binding factor is also recruited to both pol II and III snRNA promoters. Thus, both in sea urchins and D. melanogaster, similar complexes bind to the PSEs of pol II and III snRNA promoters as judged from electrophoretic mobility shift assays (3, 4), and in the latter case, site-specific protein-DNA photo-cross-linking experiments reveal the same set of polypeptides in close proximity to the DNA in both cases. Interestingly, however, the precise cross-linking patterns of these polypeptides to the U1 and U6 PSEAs are significantly different. Thus, in D. melanogaster, RNA polymerase specificity may

* This minireview will be reprinted in the 2001 Minireview Compendium, which will be available in December, 2001. The work on snRNA gene transcription is funded in part by National Institutes of Health Grant GM38810.
‡ Supported by the Howard Hughes Medical Institute.
1 The abbreviations used are: snRNA, small nuclear RNA; PSE, proximal sequence element; DSE, distal sequence element; pol, polymerase; TBP, TATA box binding protein; USE, upstream sequence element; bp, base pairs; aa, amino acid(s); hP, human P; hBRF, human BRF; hBRFU, human BRFU.
ultimately be determined by different conformations of the same factor, which are dictated by the exact PSEA sequence (12).

**Factors Besides SNAP, Required for pol II Transcription of snRNA Genes**

Transcription from TATA box-containing mRNA promoters can be reconstituted with a combination of recombinant and well-defined factors, as shown in Fig. 2A. *In vitro*, these factors can be added sequentially to the promoter to form a functional transcription initiation complex, and each step can be monitored by electrophoretic mobility shift assay. TBP or the TBP-containing complex TFIIID binds first to the TATA box, followed by TFIIIB, a TFIIIP-pol II complex, TFIIIE, and TFIIH. TFIIIA can join the initiation complex at any stage of assembly, and its main role for mRNA core promoter function appears to be counteracting repressors that associate with TBP and prevent its binding to DNA (13).

In the case of the pol II snRNA promoters, the transcription initiation complex has not yet been assembled in a stepwise fashion *in vitro*, but many of its components have been identified functionally by depletion of transcription extracts with specific antibodies and reconstitution of transcription with recombinant factors. Thus, many of the players are known, but their mode of assembly on snRNA promoters is not, and thus their location in Fig. 2A is arbitrary. Depletion and reconstitution experiments indicate that recombinant TBP (but not the TBP-containing complexes TFIIID or TFIIIB), TFIIIB, TFIIIA, TFIIIE, and TFIIIE are required (Ref. 14, and references therein). TFIIIA appears to perform a more direct function in snRNA transcription complex assembly than just counteracting TBP-associated repressors. The role, if any, of TFIIH in pol II transcription of snRNA genes is not clear. Depletion and reconstitution experiments suggest that U1 transcription either does not require TFIIH or requires much lower levels than transcription from a mRNA promoter. If TFIIH is indeed not required, this raises the interesting question of how open complex formation is achieved at snRNA promoters. A combination of all the general transcription factors and SNAP, does not initiate transcription from the U1 promoter, suggesting that additional as yet unidentified factors are required (14).

**Factors Besides SNAP, Required for pol III Transcription of snRNA Genes**

The key player for recruitment of pol II to a promoter is the factor TFIIIB because it contacts the RNA polymerase directly. In the case of pol III, this role is played mainly by the multisubunit factor TFIIIB. TFIIIB was completely defined first in *S. cerevisiae* and consists of three subunits, TBP, a tightly associated subunit referred to as the TFIIIB-related factor BRF1 (PCF4/TD84) (see Ref. 15, and references therein), and a more loosely associated polypeptide called B° (TFIIIB90/TFC5/TFC7) (16, 17). This TFIIIB complex is involved in transcription from all types of yeast pol III promoters tested including the gene-internal tRNA-type promoters and the U6 promoter, which, as described above, contains a TATA box and A and B boxes (18).

TBP was shown to be required for pol III transcription of vertebrate snRNA genes before TBP was known to be a subunit of TFIIIB (see Ref. 15 for a review). Ironically, however, the composition of mammalian TFIIIB and the role of TFIIIB polypeptides other than TBP in snRNA gene transcription have been determined only recently. A human homologue of yeast B° was recently cloned (19). The protein shows strong similarity to the yeast protein within and around a 59-aa domain called the SANT domain, which is essential for transcription in yeast. Depletion of human B° (hB°) from transcription extracts debilitates transcription from the U6 promoter and a tRNA-type promoter, and transcription can be restored in both cases by addition of recombinant hB° (19). hB° is, therefore, shown as part of the initiation complex assembled on both U6 and tRNA-type promoters in Fig. 2B.

The first human homologue of yeast BRF cloned was called TFIIIB90 (20) or human BRF (hBRF) (21). Like its yeast counterpart and like its cousin TFIIIB, the protein has a zinc binding domain at its N terminus followed by a core domain consisting of two degenerate repeats. The C-terminal half of the protein is poorly conserved with the yeast protein and has no counterpart in TFIIIB. Depletion and reconstitution experiments have shown that human BRF is required for transcription from tRNA-type promoters but not from transcription from the U6 snRNA promoter (21). Remarkably, as shown in Fig. 2B, the U6 snRNA promoter uses another homologue of yeast BRF, human BRFU (19), also called TFIIIB50 (22). hBRFU represents another member of the TFIIIB-related family of proteins and has conserved zinc and core domains and a divergent C-terminal domain.

Human BRFU was cloned both through database searching of proteins similar to hBRF (19) and through purification of a complex consisting of BRFU and four tightly associated polypeptides (22). The role of the BRFU-associated polypeptides is presently not clear. In one series of experiments, U6 transcription could be restored in a BRFU-depleted extract by addition of recombinant BRFU expressed in *Escherichia coli* (19), and a combination of partially purified pol III and recombinant SNAP, TBP, hB°, and hBRFU could direct efficient U6 transcription (23). In another
series of experiments, depletion of BRFU (TFIIB50) debilitated transcription from a snRNA-type promoter, but transcription could not be reconstituted by addition of recombinant BRFU. Instead, transcription could only be reconstituted by addition of a BRFU-containing complex immunopurified from cells expressing tagged BRFU (22). Thus, it is not clear whether the BRFU-associated polypeptides in the BRFU-containing complex are essential for U6 transcription.

BRF2, another factor encoded by an alternatively spliced BRF pre-mRNA, may also be required for U6 transcription (24). BRF2 lacks the zinc finger domain and the first repeat that are present in BRF and conserved in the other proteins of the TFIIB family, as well as the C-terminal domain present in BRF. Depletion of extracts with antibodies recognizing all BRF variants debilitated U6 transcription, and transcription could be specifically reconstituted by addition of material immunopurified from cells expressing tagged BRF2 (24). These results raise the possibility that the U6 transcription-complex contains two proteins related to BRF, BRFU, and BRF2.

The discovery of BRFU is an important step toward a complete understanding of how RNA polymerase specificity is determined at the human snRNA promoters. Indeed, in mRNA promoters, TFIIB associates with TBP bound to the TATA box and, in a manner absolutely dependent on the zinc ribbon domain, with the pol II-TFIIF complex (13). Thus, at least for mRNA promoters, TFIIB can be viewed as the key factor that bridges DNA-associated TBP or TFIIID with the polymerase, and it seems likely that TFIIB performs the same role in the pol II snRNA promoters. Like TFIIB, BRF also contacts directly the RNA polymerase, which is pol III in this case. The precise BRF domain required for this function is not known, but it is not the zinc ribbon domain because deletion or mutation of the zinc ribbon does not affect RNA polymerase recruitment but instead affects open complex formation (see Ref. 19, and references therein). It seems likely that in the human U6 promoter, the recruitment of pol III is accomplished by BRFU, either through the zinc domain as for TFIIB and pol II or through another part of the protein as for BRF and pol III at trnA-type promoters. Thus, the determination of RNA polymerase specificity may ultimately depend on whether TFIIB or BRFU is recruited to the promoter.

Activation of snRNA Gene Transcription

The human snRNA promoters are activated by a DSE. The DSE is composed of various protein binding sites, but one of them is almost certainly the octamer sequence ATGCAAAAT. In addition, it has become clear that the DSEs of many snRNA genes contain an element referred to as the SPH element. The SPH (for “Sp1 postoctamer homology”) element was first identified in the chicken snRNA gene enhancer (Refs. 25 and 26, and references therein) and in the enhancer of the selenocysteine trnA gene, whose promoter contains a PSE and TATA box (27, 28). It then became clear that a functionally important element located immediately upstream of the octamer sequence in the human U6 snRNA promoter called the NONOCT element (29) corresponds, in fact, to an SPH element (30) and that SPH elements are present in the enhancers of many snRNA promoters (31). In the human U6 snRNA promoter, both the octamer and SPH elements stimulate the formation of preini-

![Fig. 3](http://www.jbc.org/)

**Fig. 3.** Assembly of a stable snRNA transcription initiation complex. **A.** TBP and SNAP$_3$ contain built-in dampers of DNA binding that down-regulate their binding to DNA. **B.** Assembly of the U6 initiation complex. For a description see “Assembly of a Stable snRNA Transcription Initiation Complex.”

as suggested by (i) the broad expression of Oct-1, which parallels the broad expression of snRNA genes; (ii) the presence of snRNA-specific transcription activation domains within Oct-1; and (iii) the localization of Oct-1 to snRNA promoter sequences in vivo by chromatin immunoprecipitation experiments (Ref. 39, and references therein) (40). Oct-1 activates snRNA gene transcription not only through its activation domains but also through its POU domain, a bipartite DNA binding domain consisting of two helix-turn-helix-containing DNA binding structures: an N-terminal POU-specific (POU) domain and a C-terminal POU-homeo (POU$_h$) domain, joined by a flexible linker (41). As described further below, this results from the ability of the Oct-1 POU domain to bind cooperatively with SNAP$_3$, and thus recruit SNAP$_3$ to the PSE.

Assembly of a Stable snRNA Transcription Initiation Complex

The characterization of many of the factors that bind to snRNA promoters has allowed a study of how these factors interact with each other to form a stable transcription initiation complex. Our current understanding of this process is summarized in Fig. 3. Both TBP and SNAP$_3$, have built-in mechanisms that prevent their efficient binding to DNA on their own. In the case of TBP, this “damp-

and in this case it is clear that cooperative binding is dependent on a direct protein-protein contact between the two proteins, which involves a glutamic acid at position 7 within the POU$_S$ domain (blue triangle in Fig. 3B) and a lysine at position 900 within SNAP$_{190}$ (Oct-1 interacting region (OIR) in Fig. 3B) (43). Thus, cooperative binding of these factors probably involves conformational changes that convert the dampers of DNA binding into handles that contact and stabilize the factor binding to a neighboring site. This intricate mode of DNA binding ensures, in effect, that factors bind to sites located in promoter sequences rather than to inappropriate isolated sites.

The cooperative binding of the Oct-1 POU domain and SNAP, was originally characterized on probes containing closely spaced octamer and PSE. In the natural snRNA promoters, however, the octamer sequence and the PSE are separated by about 150 bp, and this distance prevents cooperative binding of Oct-1 and SNAP, on naked DNA probes. However, mapping of DNase I and micrococcal nuclease cleavage sites in chromatin suggests the presence of a positioned nucleosome between the octamer sequence and the PSE such that SNAP, and the Oct-1 POU domain can contact and recruit each other to the DNA, as illustrated in Fig. 3B. Thus, this is a case where a nucleosome does not repress transcription but instead is a functional component of the transcription activation process.

Acknowledgment—I thank Winship Herr for comments on the manuscript.

REFERENCES
1. Lobo, S. M., and Hernandez, N. (1994) in Transcription, Mechanisms and Regulation (Conaway, R. C., and Conaway, J. W., eds) pp. 127–159, Raven Press, Ltd., New York.
2. Wabele, F., and Filipowicz, W. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 9960–9964.
3. Jensen, R. C., Wang, Y., Hardin, S. B., and Stumph, W. E. (1998) J. Biol. Chem. 276, 3015–3026.
Small Nuclear RNA Genes: a Model System to Study Fundamental Mechanisms of Transcription
Nouria Hernandez

J. Biol. Chem. 2001, 276:26733-26736.
doi: 10.1074/jbc.R100032200 originally published online June 4, 2001

Access the most updated version of this article at doi: 10.1074/jbc.R100032200

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 44 references, 30 of which can be accessed free at http://www.jbc.org/content/276/29/26733.full.html#ref-list-1