MOLECULAR AND CELLULAR BIOLOGY, Aug. 1987, p. 3018–3020
0270-7306/87/083018-03$02.00/0

Vol. 7, No. 8

The Natural 5′ Splice Site of Simian Virus 40 Large T Antigen Can Be Improved by Increasing the Base Complementarity to U1 RNA

YUAN ZHUANG, HELIOS LEUNG, AND ALAN M. WEINER*

Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Received 18 March 1987/Accepted 5 May 1987

The use of alternative 5′ splice sites in the simian virus 40 early-transcription unit controls the ratio of large T to small t antigen during viral infection. To study the regulation of these alternative 5′ splice sites, we made two mutants which improve the match of the large-T-antigen 5′ splice site to the 5′ splice site consensus sequence. Whether these mutants were assayed in vitro or in vivo, we found that the efficiency of large-T splicing is increased by improving the match of the large-T-antigen 5′ splice site to the consensus. We conclude that the match of a 5′ splice site is an important determinant of 5′ splice site utilization and that the simian virus 40 large-T-antigen 5′ splice site is almost certainly recognized by the U1 small nuclear RNA component of the U1 small nuclear ribonucleoprotein particle.

The splicing of mRNA precursors requires base pairing between the 5′ splice site and U1 RNA (7, 11, 14). However, the consensus 5′ splice site extends over 9 nucleotides (nt), but the match of natural 5′ splice sites to consensus varies from 5 to 9 nt with an average of only 7 nt (9). One possible explanation for the existence of suboptimal 5′ splice sites is that the strength of U1 binding is a key factor in determining the utilization of alternative 5′ splice sites. However, characterization of the alternative 5′ splice sites in the simian virus 40 (SV40) early transcript indicates that this explanation cannot be true in its simplest form. The large-T-antigen (large-T) splice site (with only a 5-nt match to the consensus) is utilized much more efficiently than the alternative small-t-antigen (small-t) 5′ splice site (with a 7-nt match to the consensus) both in vivo and in vitro (5, 10, 13). In addition, utilization of the large-T 5′ splice site is relatively insensitive to treatments that are thought to inactivate U1 small nuclear ribonucleoprotein particles, such as injection of RNP antibodies into the oocyte (4) or digestion of the 5′ end of U1 RNA with RNase H and a complementary oligodeoxynucleotide (2). Finally, Fu and Manley (5) have recently reported that the ratio of large-T to small-t splicing depends on cell type, suggesting the involvement of a *trans*-acting factor(s) other than U1 small nuclear ribonucleoprotein particles.

Taken together, the preceding observations suggest the formal possibility that the large-T 5′ splice site might not be recognized by U1 RNA. To test whether U1 recognizes the large-T 5′ splice site, we made two mutations that increase the match of the large-T 5′ splice site to the consensus sequence but leave the small-t 5′ splice site unchanged as a reference site. We then assayed the efficiency of large-T splicing both in vitro and in vivo.

Table 1 shows the 5′ splice site sequences of wild-type small t and large T and two large-T splice site mutants. The mutants were generated by oligonucleotide-directed mutagenesis of the appropriate SV40 *Hin*dIII fragment cloned in mp10 (6). We designated the wild-type large-T 5′ splice site T5(wt) because it matched the consensus at five contiguous positions. Mutant T6 matched at six contiguous positions, and mutant T8 matched at eight. If U1 recognition were important in determining the use of alternative 5′ splice sites, we would have expected to see the efficiency of large-T splicing increase in the order of T5(wt), T6, and T8.

We first tested the mutants in an in vitro splicing reaction with a HeLa cell nuclear extract (10, 13). In vitro splicing of the SV40 T-antigen transcript has been well studied, and it is known that the large-T 5′ splice site is much more efficiently used than the small-t 5′ splice site (10, 13). In our experiments, substrate RNAs were made by SP6 runoff transcription from each mutant SP6 clone, starting at the *Hin*dIII site (265 nt upstream of the large-T 5′ splice site) and extending to the *Sty*I site (162 nt downstream of the 3′ splice site) (Fig. 1B). We determined the efficiency of splicing at the large-T 5′ splice site by measuring the ratio of large-T splice products to remaining precursors. Figure 1A shows the time course of in vitro splicing. Although T6 had little or no effect on the splicing reaction, T8 dramatically increased the efficiency of large-T 5′ splice site utilization compared with T5(wt).

The advantage of the in vitro splicing assay is that one can see the direct effect of mutations on the kinetics of the reaction. However, in vitro splicing is not always a faithful representation of splicing reactions inside a living cell (1). For example, the small-t 5′ splice site is normally used in vivo but not in vitro (2, 10). For this reason, we next tested the mutants in vivo by using a transient expression assay. The plasmids we used for transfection contain the SV40 early region (from *Hpa*II to *Bam*HI) inserted in the linker region of pUC13 (Fig. 2). For technical reasons, these

TABLE 1. Wild-type and mutant SV40 early gene 5′ splice site sequences

Type of mRNA	5′ Splice site sequence	No. of contiguous matches to the consensus
t	AAG/GTAAAT	7
T5(wt)	GAG/GTATTT	5
T6	GAG/GTAAAT	6
T8	GAG/GTAAGT	8
Consensus	$^{C}_{A}$AG/GT$^{A}_{G}$AGT	9

* Corresponding author.

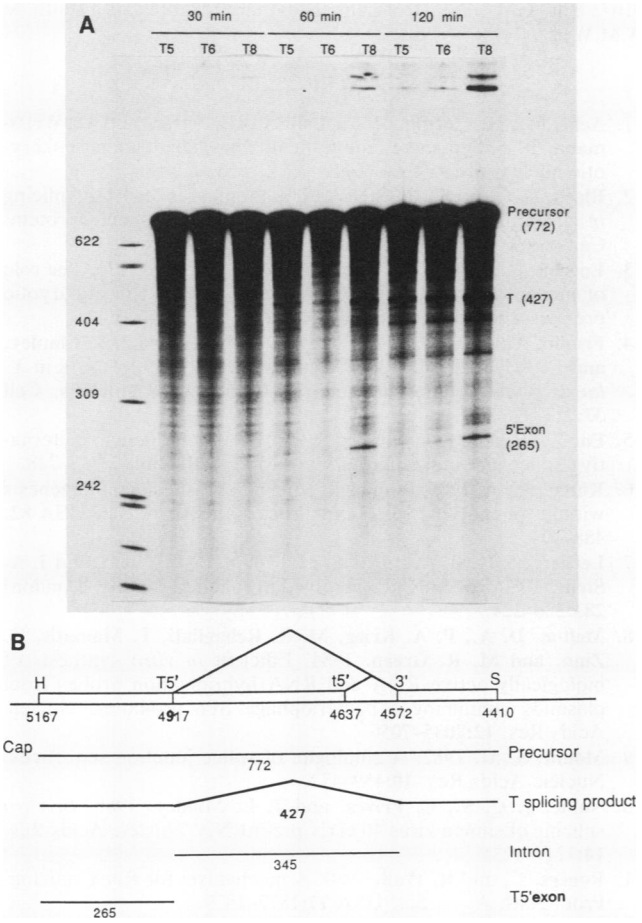

FIG. 1. Splicing of wild-type and mutant SV40 early mRNA precursors in vitro. (A) Capped and uniformly labeled SP6 RNA transcripts were spliced in vitro as described elsewhere (2) for 2 h, the products were resolved on a 10% sequencing gel, and the gel was autoradiographed. The leftmost lane is an *Hpa*II digest of pBR322, with resulting fragments used as size markers. (B) The mutagenized *Hin*dIII fragments were subcloned into the *Hin*dIII site (H) of pSP64. The plasmids were linearized by cutting with *Sty*I (S) and transcribed with SP6 polymerase in the presence of GpppG and [^{32}P]rGTP (8). Sizes are given in nucleotides.

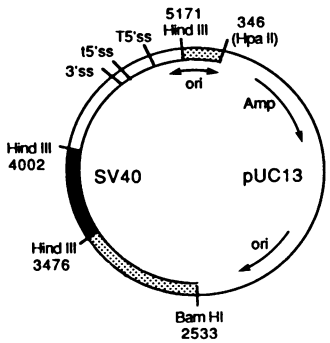

FIG. 2. Plasmid used for in vivo transfection assay. The SV40 fragment starts at the *Hpa*II site, continues through the replication origin and the entire early region, and stops at the *Bam*HI site. All sites within the SV40 fragment are numbered by the method of Tooze (12). ——, pUC13 plasmid opened at the *Bam*HI site; ▬, *Hin*dIII fragment which was deleted in the mutant clones; ▭, *Hin*dIII fragment which was subcloned into mp10 vector for mutagenesis.

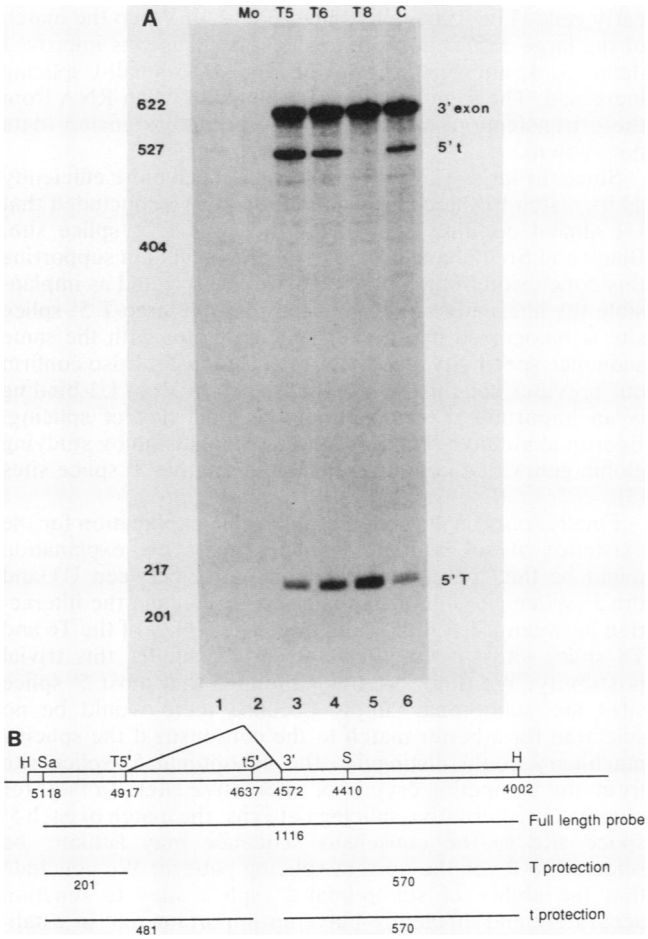

FIG. 3. RNase protection assay for spliced products in transfected cells. (A) HeLa cells were transfected by the calcium phosphate precipitation method. Total cytoplasmic RNA was isolated and used directly in the RNase protection assay (14). RNase mapping was carried out by the standard procedure described by Melton et al. (8) except that T1 ribonuclease alone was used in the digestion. The protected fragments of the RNA probe were analyzed on a 4% sequencing gel. Lane 1, Marker DNA as described in the legend to Fig. 1; lane 2, mock transfection control; lane 6, wild-type control which has an intact SV40 early gene. The ratio of large-T to small-t splicing is slightly lower than that described previously (5), perhaps because of the pUC13 vector used here. Lanes 3, 4, and 5 correspond to mutants T5, T6, and T8, respectively; for convenience, the *Hin*dIII fragment within the 3′ exon was deleted in each of these constructions. (B) The RNA probe was made from the pSP64 construction described in the legend to Fig. 1B except that the *Hin*dIII fragment (H) was inserted in the opposite orientation. The probe extended from the *Hin*dIII site (downstream of the 3′ splice site) to the *Sau*96I site (Sa; upstream of the large-T 5′ splice site). The sizes of the full-length probe and the protected exon regions are indicated on the map. S, *Sty*I.

plasmids also lack a *Hin*dIII fragment that lies entirely within the 3′ exon. As described below, this internal deletion had no effect on splicing. For convenience, we used SP6 probes and RNase protection to assay the splicing products from cytoplasmic RNA (8). Since we could not detect mRNA precursors in this kind of assay, we used the ratio of large-T to small-t splicing as a measure of the efficiency of splicing at the large-T 5′ splice site. As a control for the plasmid constructions, we showed that deletion of the internal *Hin*dIII fragment had no effect on splicing of the SV40

early gene (Fig. 3; compare lanes 3 and 6). When the match of the large-T 5′ splice site to the consensus was improved (lanes 3, 4, and 5), the ratio of large-T to small-t splicing increased. The same results were obtained when RNA from these transfections was assayed by primer extension (data not shown).

Since the large-T 5′ splice site was utilized more efficiently as its match to the consensus improved, we concluded that U1 almost certainly recognizes the large-T 5′ splice site. Black and Steitz have provided biochemical data supporting this conclusion (data not published). We rejected as implausible the alternative interpretation that the large-T 5′ splice site is recognized by another splicing factor with the same sequence specificity as U1 RNA itself. The data also confirm our previous conclusion (14) that the strength of U1 binding is an important determinant of the efficiency of splicing. Eperon et al. have reached the same conclusion by studying globin genes that contain artificial alternative 5′ splice sites (3).

Finally, our results suggest a plausible explanation for the existence of suboptimal 5′ splice sites. One explanation might be that increased complementarity between U1 and the 5′ splice site inhibits splicing by prolonging the interaction between U1 and the substrate. The ability of the T6 and T8 splice sites to function efficiently excludes this trivial possibility. We therefore favor the idea that most 5′ splice sites are suboptimal simply because there would be no selection for a better match to the consensus if the splicing machinery could distinguish the suboptimal 5′ splice site from any competing cryptic or alternative sites. In fact, for complex or alternative splicing patterns, the match of each 5′ splice site to the consensus sequence may actually be selected to favor the correct splicing pattern. We conclude that the ability of suboptimal 5′ splice sites to function accurately and efficiently plays an important role in establishing the subtle balance between the many potential 5′ splice sites within an mRNA precursor.

We thank Doug Black for advice about in vitro splicing and A. Ruusala for synthesizing all the oligonucleotides.

This research was supported by Public Health Service grants GM 31073 and GM 31335 from the National Institutes of Health to A.M.W.

LITERATURE CITED

1. **Aebi, M., H. Hornig, R. A. Padgett, J. Reiser, and C. Weissmann.** 1986. Sequence requirements for splicing higher eukaryotic nuclear pre-mRNA. Cell **47**:555–565.
2. **Black, D. L., and J. A. Steitz.** 1986. Premessenger RNA splicing *in vitro* requires intact U4/U6 small nuclear ribonucleoprotein. Cell **46**:697–704.
3. **Eperon, L. P., J. P. Estibeiro, and I. C. Eperon.** 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. Nature (London) **324**:280–282.
4. **Fradin, A., R. Jove, C. Hemenway, H. D. Keiser, J. L. Manley, and C. Prives.** 1984. Splicing pathways of SV40 mRNAs in *X. laevis* oocytes differ in their requirements for snRNPs. Cell **37**:927–936.
5. **Fu, X.-Y., and J. L. Manley.** 1987. Factors influencing alternative splice site utilization in vivo. Mol. Cell. Biol. **7**:738–748.
6. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**: 488–492.
7. **Lerner, M. R., J. A. Boyle, S. M. Mount, S. L. Wolin, and J. A. Steitz.** 1980. Are snRNPs involved in splicing? Nature (London) **283**:220–224.
8. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
9. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
10. **Noble, J. C. S., C. Prives, and J. L. Manley.** 1986. *In vitro* splicing of simian virus 40 early pre-mRNA. Nucleic Acids Res. **14**:1219–1235.
11. **Rogers, J., and R. Wall.** 1980. A mechanism for RNA splicing. Proc. Natl. Acad. Sci. USA **77**:1877–1879.
12. **Tooze, J.** 1981. DNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **van Santen, V. L., and R. A. Spritz.** 1986. Alternative splicing of SV40 early pre-mRNA *in vitro*. Nucleic Acids Res. **14**:9911–9926.
14. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.