JOURNAL OF VIROLOGY, Nov. 1992, p. 6819–6823
0022-538X/92/116819-05$02.00/0

Vol. 66, No. 11

Recognition Efficiency of the Hepatitis B Virus Polyadenylation Signals Is Tissue Specific in Transgenic Mice

SABINA PERFUMO,[1] LAURA AMICONE,[1] STEFANO COLLOCA,[2] MARCO GIORGIO,[1] LAURA POZZI,[1] AND MARCO TRIPODI[1]*

Dipartimento Biopatologia Umana, Sezione di Biologia Cellulare, Università La Sapienza, Rome,[1] and IRBM P. Angeletti, Pomezia,[2] Italy

Received 18 June 1992/Accepted 10 August 1992

The hepatitis B virus genome contains a unique polyadenylation (TATAAA) signal which is differentially utilized in the formation of the various hepatitis B virus transcripts. A head-to-tail multiple-copy insertion of a viral fragment comprising the viral enhancer, the X promoter, the X open reading frame, and the viral poly(A) signal in transgenic mice allowed us to monitor tissue-specific differences in the expression of transcripts initiating from the X promoter. These transcripts are efficiently processed at the first polyadenylation site in the liver, while in the kidney, the brain, and the testis, a portion of the transcripts covers two copies of the transgene, since only the second polyadenylation site is properly recognized. As discussed in this article, this observation suggests a tissue-specific distribution of cellular factors involved in polyadenylation.

For most eukaryotic genes, two *cis* signals, the poly(A) hexamer (AATAAA) and a GT-rich region located downstream, are sufficient to confer correct 3′ end processing of transcripts (12). It has been previously demonstrated that cellular factors bind to these sequences and form a spatially constrained cooperative interaction (2–9). In the expression of the circular genome of hepatitis B virus (HBV), the regulation of the 3′ end processing of transcripts, an essential step in the viral life cycle, is not understood. These *cis* signals appear insufficient, since the unique viral polyadenylation signal (TATAAA), a variant form of the eukaryotic hexamer, is differentially utilized by subgenomic and genomic viral transcripts (3, 16, 19).

Hepatocytes productively infected by HBV bear transcripts of 3.4, 2.4, 2.0, and 0.7 kb (8–21). The 0.7-kb HBV transcripts code for a protein named X (21), which has been shown to be a transcriptional transactivator of several viral and cellular genes (13, 17, 20, 22). The 2.4- and 2.0-kb transcripts coding for the surface antigen proteins originate from two different promoters and stop about 10 bases downstream of the TATAAA signal (18).

The 3.4-kb class of transcripts comprises all terminally redundant RNAs larger than genomic size. The 5′ ends of genomic HBV RNAs are heterogeneous, and different transcripts encode for the precore and core proteins, with the transcripts that encode for the core proteins also being the templates for reverse transcription of the genomic DNA (3–16). These transcripts initiate about 100 bases upstream of the unique HBV poly(A) signal hexamer, continue throughout the entire genome, and recognize the hexamer only during the second run. Therefore, the polyadenylation sequences are present twice in the resulting RNAs, a feature of transcripts from many terminally redundant viruses (24).

In the present article, we describe the transcription pattern observed with transgenic mice carrying a portion of HBV comprising the viral enhancer, the X promoter, the X open reading frame, and the unique viral poly(A) signal. These animals exhibit promoter-independent, tissue-specific efficiencies of poly(A) site recognition, thereby suggesting a cell-specific distribution of endogenous proteins which might play a role in the polyadenylation process.

Generation and screening of transgenic mice. We isolated an *Acc*I-*Bgl*II (bases 824 to 1984) restriction fragment from the HBV (ayw strain) genome (6) which comprises the viral enhancer, the promoter, and the whole coding region of X, followed by the unique viral termination site. The purified fragment was microinjected into (C56BL/6J × DBA/2)F_1 fertilized mouse eggs, according to the method of Hogan et al. (11). Founders were identified by Southern blotting with the *Acc*I-*Bgl*II viral fragment as a probe. Densitometer scanning of Southern blot analysis of four founders (X1, X4, X6, and X17) demonstrated that the 1.1-kb fragment expected from the digestion with one-cut restriction enzymes (*Taq*I and *Bam*HI) is 3 to 10 times more represented than the other two hybridizing fragments (insert-flanking sequences), indicating also that these families carry a multicopy insertion of the transgene with a head-to-tail structure (see Fig. 1).

X-specific transcripts in transgenic animals. RNA was isolated from various tissues of the four transgenic mouse lines and control mice by the guanidine thiocyanate procedure (4). An RNase protection assay and Northern (RNA) blot analysis revealed that the viral enhancer and the X promoter can direct expression in vivo with broad tissue specificities. The transcriptional activity of the transgene was determined with several members of the families, from the founder to the sixth generation and from 15-day-post-coitum (d.p.c.) embryos to 23-month-old transgenic mice.

Figure 2A shows an RNase mapping analysis of total RNA prepared from various tissues from a 2-month-old transgenic mouse and from control (wild-type) mice. X mRNA was also found in 15-d.p.c. tissues (Fig. 2B).

The RNase protection assay was performed according to the method of Bartenschlager et al. (1) with a P^{32}-labelled antisense RNA probe on 10 μg of total RNA from transgenic and wild-type mouse tissues. The plasmid used as the template for in vitro transcription was a Bluescript vector carrying an HBV *Hin*cII restriction fragment (bases 1682 to 2165) under the control of a T7 promoter. The protected fragments were analyzed by 6% acrylamide–8 M urea denaturing gel.

The arrows in Fig. 2 indicate protected fragments comple-

* Corresponding author.

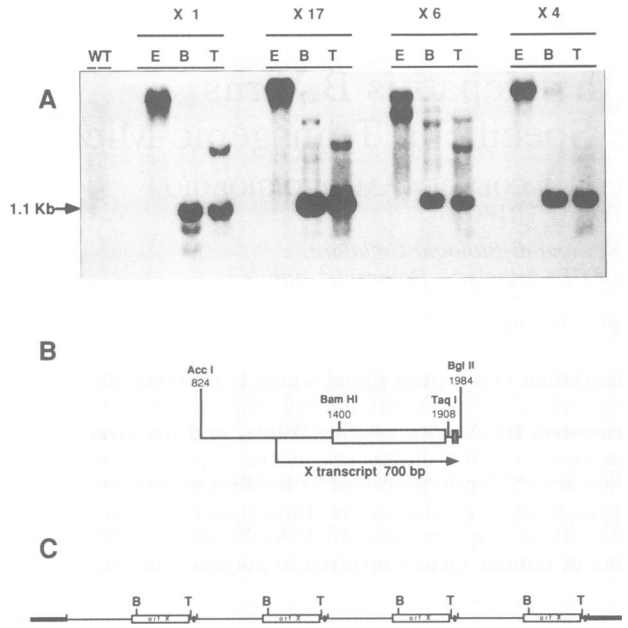

FIG. 1. (A) Southern blot analysis of the transgenic families X1, X17, X6, and X4. Shown is an autoradiograph of a Southern blot analysis of tail DNA from transgenic animals (X1, X17, X6, and X4) and a control wild-type animal (WT). Purified DNA was incubated with the one-cut restriction endonucleases *Taq*I (T) and *Bam*HI (B) and with *Eco*RI (E), which does not cut within the injected fragment. DNA fragments were separated on a 1% agarose gel. The arrow indicates the 1.1-kb fragment (unit-size fragment) resulting from the *Taq*I and *Bam*HI digestions. The copy numbers are 4, 11, 5, and 7, respectively. These numbers were obtained by analyzing Southern autoradiographs at the appropriate exposure with an LKB UltroScan XL Densitometer. (B) Schematic illustration of the microinjected viral fragment comprising regulatory elements (thin line), the promoter (arrow), open reading frame X (open box), the poly(A) site (stippled box), and the restriction enzyme cut sites. (C) Schematic illustration of the tandem head-to-tail integration of the microinjected viral fragment within the mouse genome (thick line).

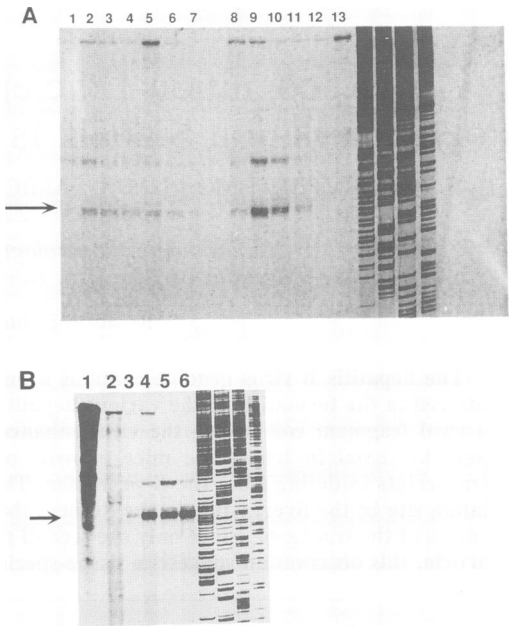

FIG. 2. RNase mapping analysis of the open reading frame X transgene transcripts. The arrows indicate protected fragments detecting X-specific mRNA. The size of the protected fragment predicts that the 3′ ends of transcripts map at base 1932 of HBV (10 bases downstream of the TATAAA hexamer). A second protected fragment is also detected (see the text). (A) Expression pattern of X-specific mRNA in adult tissues. RNA was extracted from the following transgenic tissues. Lanes: 1, brain; 2, cerebellum; 3, salivary glands; 4, muscle; 5, lung; 6, heart; 7, ovary; 8, gut; 9, kidney; 10, liver; 11, spleen. In lanes 12 and 13, RNase mapping analysis was done with kidney and liver RNA from wild-type mice. (B) Expression pattern of X-specific mRNA in embryos. RNA was extracted from a transgenic 15-d.p.c. total embryo (lane 3) and the livers of a 15-d.p.c. embryo (lane 4), a newborn (lane 5), and a 2-month-old animal (lane 6). Also shown are the input probe (lane 1) and wild-type liver RNA (lane 2). The same tissue distribution was found with all the transgenic lines.

mentary to X-specific mRNA in liver as expected but also in a variety of other tissues. The size of the protected fragment predicts that the 3′ ends of transcripts map at HBV base 1932 (10 bases downstream of the TATAAA hexamer). The same tissue distribution was found in all of the transgenic lines, showing that the combined action of HBV enhancer I and the X promoter does not restrict transcription to hepatocytes. This finding is in agreement with previous results obtained by cell transfection experiments in which the enhancer was active in a cell-type-independent manner (5–23). We also detected a second, longer protected fragment, whose relative intensity with respect to the signal described varies from animal to animal and from tissue to tissue (Fig. 2). The size of the longer protected fragment identifies the end of the homology between the RNase mapping probe and the transgene (base 1984), therefore indicating the presence of longer-than-expected transcripts.

Northern blots were performed with Hybond N membranes (Amersham) with 20 μg of total RNA for each analysis. Hybridization, washing, and rehybridization were carried out according to the manufacturer's instructions.

HBV X gene transcripts of the expected size, 0.7 kb, were found in the same tissues described for Fig. 2A above (data not shown). In addition, longer-than-expected transcripts of about 1.9 kb were found in kidney, spleen, brain, testis, and, interestingly, to a much lesser extent, in liver tissue, suggesting tissue-specific distribution of the longer transcript (Fig. 3A). The consistency of this observation has been verified with transgenic lines in animals of different ages and sexes. Figure 3B shows a Northern blot analysis of total RNA from the livers and the kidneys of 12 (B11)-, 17 (B1)-, and 23 (B23)-month-old males and a 23 (B26)-month-old female. Autoradiographs at the appropriate exposures were analyzed with an LKB UltroScan XL densitometer.

In all of these animals, the relative ratio between different transcripts (0.7 and 1.9 kb) is different for the two types of tissues. The longer transcript is almost absent in liver tissue and is much more abundant in kidney tissue.

A quantitative analysis of these transcripts with livers and kidneys of 42 animals of the four different lines revealed that, in liver tissue, transcripts appear to be efficiently processed at the first polyadenylation site (about 95% of the total transcripts were 0.7-kb long), while in kidney tissue, a much larger portion of the transcripts (from 20 to 50%) appears to be processed at the second polyadenylation site. The percentage of the 1.9-kb signal with respect to the total signal is

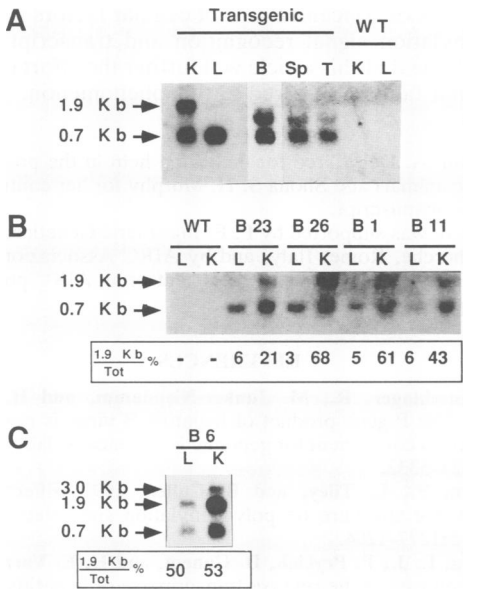

FIG. 3. Tissue-specific expression patterns. Autoradiographs of Northern blot analysis are shown; filters have been probed with HBV X-specific sequences. Arrows indicate the sizes of detected transcripts in kidneys and in livers. (A) RNA from 2-month-old transgenic mice and from a control mouse (WT). Autoradiographs of different Northern blots are shown. K, kidney; L, liver; B, brain; Sp, spleen; T, testis. (B) RNA from a control mouse (WT) and from transgenic mice (B23, B26, B1, and B11). (C) RNA from a transgenic mouse (B6) showing a higher percentage of longer transcripts (see the text); a 3.0-kb transcript probably resulting from double-termination site leakage (see the text) is also detectable in kidney RNA. The percentage of the 1.9-kb signal with respect to the total signal is indicated below each lane (B and C). Note that the 53% value (C, lane K) does not include the 3.0-kb signal, which amounts to 13% of the total signal.

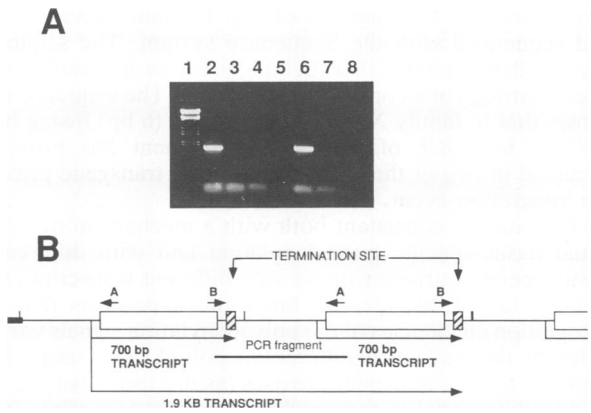

FIG. 4. PCR analysis of X transcripts. (A) In order to distinguish between a single-copy transgene and a direct or inverted tandem of the transgene, PCR was performed using either oligonucleotides A and B (lanes 2, 5, 6, 7, and 8) or oligonucleotide A (lane 3) and oligonucleotide B separately (lane 4). Templates for the reaction are genomic DNA from a transgenic mouse (lanes 2 to 4), the single-copy-injected fragment *Acc*I-*Bgl*II (lane 5), transgenic mouse kidney RNA (lane 6), and mouse wild-type DNA (lane 7) and kidney RNA (lane 8). The amplified fragment is detectable only when the reaction is performed with both oligonucleotides A and B on genomic DNA and on kidney RNA (lanes 2 and 6, respectively) and not when the template is the single-copy fragment (lane 5). (B) Schematic illustration of the structures of the 0.7- and 1.9-kb transcripts and that of the PCR fragment. The arrows indicate the locations of primers A and B. Primer A has an antisense polarity.

shown in Fig. 3. Note that this is not an absolute value, since we cannot establish the number of transcriptionally active transgene units.

Interestingly, RNA from 2 of the 42 animals analyzed gave a notably different pattern; there appears to be relatively less recognition of the first and second polyadenylation signals, both in liver tissue (only 50% of the transcripts were 0.7 kb long) and in kidney tissue, in which a third transcript of about 3 kb, most likely covering three copies of the transgene, is also found (Fig. 3C). Figure 3C shows the transcription pattern of RNA from one of these two animals (B6).

We have so far found no correlation between polyadenylation site recognition and factors such as age, sex, and genetic background. The transcriptional analysis of the transgene which we performed should have allowed us to detect the C promoter activity in our system. Surprisingly, transcripts directed by the HBV enhancer and the C promoter were detectable in none of our transgenic lines.

The longer transcripts are a read-through of one copy of the transgene into the second. Longer transcripts could arise either from the activity of a flanking endogenous promoter driving tissue-specific transcription or from inefficient recognition of the polyadenylation site of the multiple-copy transgene, resulting in transcripts ending at a successive polyadenylation site. In order to confirm one of these hypotheses, we performed reverse transcription and polymerase chain reaction (PCR) with RNA derived from the kidneys of transgenic animals.

PCR was performed with genomic DNA and with cDNA derived from reverse transcription of DNase-treated kidney poly(A)$^+$ RNA. DNase I was removed from the RNA sample by phenol extraction. The reverse transcription with oligo(dT) primer was performed according to the RiboClone cDNA Synthesis System (Promega).

PCR was performed with 50 pmol of oligonucleotides A (a 26-mer reverse primer complementary to HBV sequences 1446 to 1424) and B (a 20-mer direct primer complementary to viral bases 1861 to 1880) and the Perkin-Elmer kit (see Fig. 4B). The reaction mixture was subjected to 30 cycles of PCR (94°C for 30 s, 55°C for 30 s, and 72°C for 30 s).

These oligonucleotides are able to prime PCR only in the presence of a tandem organization of the transgene and its transcripts. Moreover, the size of the amplified fragment distinguishes a tandem in the head-to-tail configuration from an inverted tandem. Single-copy transcripts or transcripts resulting from an endogenous flanking promoter will not function as templates for PCR with this pair of oligonucleotides. Figure 4A shows a 755-bp fragment resulting from a reaction utilizing a mixture of oligonucleotides A and B on genomic DNA (Fig. 4A, lane 2) and on kidney RNA (Fig. 4A, lane 6) from a transgenic animal; neither wild-type DNA (Fig. 4A, lane 7) nor RNA (Fig. 4A, lane 8) nor the single-copy fragment (Fig. 4A, lane 5) is amplified. The lack of an amplified fragment in the reaction performed with oligonucleotides A and B separately (Fig. 4A, lanes 3 and 4) confirms the Southern analysis data that revealed that all copies of the transgene are in a head-to-tail array. As a

further control, the fragment of Fig. 4A, lane 6, was cloned and sequenced with the Sequenase system. The sequence derived demonstrates that the longer transcript results from a read-through of a copy of the transgene. The sequence also shows that in family X1, a small deletion (6 bp) fusing base 1982 to base 828 of the second fragment has probably occurred in two of the four copies of the transgene prior to the integration event.

Our data are consistent both with a mechanism of differential tissue-specific polyadenylation and with differential tissue-specific turnover of the two different transcripts (0.7 and 1.9 kb). However, we favor the hypothesis that the recognition efficiencies of the polyadenylation signals vary in different tissues, since differential poly(A) site usage is a common feature in hepadnaviruses (as discussed below) and is indeed essential in terminally redundant transcripts from retroid elements which include two poly(A) hexamer sequences (15).

We therefore consider it most likely that transcripts in liver cells are polyadenylated mainly at the first polyadenylation site. In other tissues, such as kidney, brain and testis tissues, the first polyadenylation site is largely ignored, and polyadenylation occurs at the second polyadenylation site. Apart from the two cases in which larger transcripts were found (Fig. 3C), there was no detectable leakage in any tissues past the second poly(A) site. The expression of the transgene discussed here would thus demonstrate the occurrence of differential polyadenylation in an in vivo system. It would also demonstrate that polyadenylation is regulated tissue specifically. Our results suggest that polyadenylation in HBV is influenced by the interaction of multiple *cis*-acting elements and cellular factors.

In ground squirrel HBV transcripts, cooperation between the poly(A) hexamer and multiple upstream *cis*-acting elements named PS1 and PS2 is necessary and sufficient in itself for efficient polyadenylation in hepatocytes (15). Accordingly, HBV X transcripts, which cover HBV sequences that include putative signals homologous to the ground squirrel HBV PS1 and PS2, should also be efficiently polyadenylated in hepatocytes. This is in contrast to the finding that in HBV X mRNA, polyadenylation signal recognition is leaky in transfected hepatocytes (10).

The multiple *cis*-acting signals which contribute to the regulation of polyadenylation in ground squirrel HBV (15), still to be demonstrated with HBV, could mediate the efficient polyadenylation signal recognition observed for the livers of transgenic animals. In the other tissues, a possible additional signal located in the region from bases 800 to 1200 could play a relevant cooperative role. This would explain both the observation that leakiness of the second polyadenylation site in our transgenic animals is generally undetectable and the fact that transcripts of the S gene promoter stop efficiently at the HBV polyadenylation site (19).

In some cases, polyadenylation site recognition appears to be influenced by the promoter itself (14). In others, the distance of the promoter from the polyadenylation hexamer seems to play a relevant role (7). The RNA secondary structure is also thought to be a critical factor (2).

Taking advantage of the tandem integration of the transgene as well as its broadly tissue-specific transcriptional activity, we have described a case of putative alternative polyadenylation site selection which is influenced by neither the promoter nor the distance of the promoter from the polyadenylation hexamer.

So far, we can only speculate that the differences in processing observed for different tissues reflect different distributions or concentrations of cellular factors involved in polyadenylation signal recognition and transcript processing. We hope that this article will further the effort to identify the cellular factors involved in this phenomenon.

We thank A. Del Marco for dedicated help in the production of transgenic animals and Shona S. H. Murphy for her contribution to editing the manuscript.

This work was supported by P. F. Ingegneria Genetica, Consiglio delle Richerche, Rome, Italy, and by AIRC Associazione Italiana Ricerca Sul Cancro. S.P. is the recipient of an AIRC postdoctoral fellowship.

REFERENCES

1. **Bartenschlager, R., M. Junker-Niepnamm, and H. Schaller.** 1990. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. J. Virol. **64**:5324–5332.
2. **Brown, P., L. Tiley, and B. Cullen.** 1991. Effect of RNA secondary structure on polyadenylation site selection. Genes Dev. **5**:1277–1284.
3. **Chang, L. J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1989. Biosynthesis of the reverse transcriptase of hepatitis B viruses involves de novo translational initiation, not ribosomal frameshifting. Nature (London) **337**:364–368.
4. **Chomczynsky, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium-thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
5. **Elfassi, E.** 1987. Broad specificity of the hepatitis B enhancer function. Virology **160**:259–262.
6. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. Nature (London) **281**:646–650.
7. **Ganem, D.** Personal communication.
8. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56**:651–693.
9. **Gilmartin, G. M., and J. R. Nevins.** 1989. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev. **3**:2180–2189.
10. **Guo, W., J. Wang, G. Tam, B. Yen, and J. Ou.** 1991. Leaky transcription termination produces larger and smaller than genome size hepatitis B virus X gene transcripts. Virology **81**:630–636.
11. **Hogan, B. L. M., F. Costantini, and E. Lacy.** 1986. Manipulation of the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Levitt, N., D. Briggs, A. Gil, and N. J. Proudfoot.** 1989. Definition of an efficient poly(a) site. Genes Dev. **3**:1019–1025.
13. **Mahè, Y., N. Mukaida, K. Kuno, M. Akiyama, N. Ikeda, K. Matsushima, and S. Marukami.** 1991. Hepatitis B virus X protein transactivates human interleukin-8 gene through acting on nuclear factor kB and CCAAT/enhancer-binding protein-like *cis* elements. J. Biol. Chem. **266**:13759–13763.
14. **Neuman de Vegvar, H., E. Lund, and J. Dahlberg.** 1986. Formation of 3′ end of U1 snRNA is coupled to transcription from snRNA promoter. Cell **47**:259–266.
15. **Russnak, R., and D. Ganem.** 1990. Sequences 5′ to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. Genes Dev. **4**:764–776.
16. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science **232**:477–484.
17. **Seto, E., P. J. Mitchell, and T. S. B. Yen.** 1990. Transactivation of the hepatitis B virus X protein depends on AP-2 and other transcription factors. Nature (London) **344**:72–74.
18. **Siddiqui, A., S. Jameel, and J. Mapoles.** 1986. Transcriptional control elements of hepatitis B surface antigen gene. Proc. Natl. Acad. Sci. USA **83**:566–570.
19. **Simonsen, C., and A. Levinson.** 1983. Analysis of processing and

polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. Mol. Cell. Biol. **3**:2250–2258.

20. **Spandau, D. F., and C. H. Lee.** 1988. *trans*-activation of viral enhancers by the hepatitis B virus X protein. J. Virol. **62**:427–434.
21. **Treinin, M., and O. Laub.** 1987. Identification of a promoter element located upstream from the hepatitis B virus X gene. Mol. Cell. Biol. **7**:545–548.
22. **Twu, J.-S., and R. H. Schloemer.** 1989. Transcription of the human beta interferon gene is inhibited by hepatitis B virus. J. Virol. **63**:3065–3071.
23. **Vannice, J., and A. Levinson.** 1988. Properties of the human hepatitis B virus enhancer: position effects and cell-type nonspecificity. J. Virol. **62**:1305–1313.
24. **Varmus, H.** 1988. Retroviruses. Science **240**:1427–1435.