Volume 13 Number 16 1985 Nucleic Acids Research

Novel deletion mutants that enhance a distant upstream 5' splice in the E3 transcription unit of adenovirus 2

Susan L.Deutscher[1], Bheem M.Bhat, Michael H.Pursley, Christos Cladaras and William S.M.Wold*

Institute for Molecular Virology and [1]The Edward A. Doisy Department of Biochemistry, Saint Louis University School of Medicine, St. Louis, MO 63110, USA

Received 29 May 1985; Accepted 24 July 1985

ABSTRACT

Region E3 of adenovirus is a "complex" transcription unit: i.e. different mRNAs and proteins arise by differential RNA 3' end selection and differential splicing of the primary transcript. We are using viable virus mutants to understand the controls that dictate the specificity and efficiency of the RNA processing signals. We describe a novel class of deletion mutations that enhance a natural 5' splice site located ~740 nucleotides (nt) upstream. In particular, deletions within nt 1691-2044 in the E3 transcription unit result in a 5-fold enhancement of the 5' splice site at nt 951 (as reflected in steady-state mRNA). The effect is specific, because the deletions do not affect the 5' splice site at nt 372, and because deletions within nt 2044-2214 do not affect either the 951 or the 372 5' splice sites. As a consequence of the enhanced splicing at the 951 5' site, synthesis of the major E3 mRNA and the major E3 protein (gp19K) are dramatically reduced. At least one of the natural 3' splice sites, located at nt 2157, is the recipient of the enhanced splicing at the 951 5' splice site. We conclude that sequences located within nt 1691-2044 affect (probably in cis) splicing at the 951 5' splice site. We speculate that nt 1691-2044 includes a splicing control region which functions to suppress splicing at the 951 5' splice site.

INTRODUCTION

In "complex" transcription units, a nuclear RNA precursor can yield alternate mRNAs and proteins depending on the selection of the RNA 3' end site and/or the splice site (1). Most transcription units in DNA and RNA tumor viruses are complex, and more than a dozen cellular complex transcription units have now been identified. Complex transcription units are important because they compress genetic information and because they provide for gene regulation at the level of RNA processing (reviewed in 1, 2). In some cases, the choice of the RNA 3' end or splice site is a function of the differentiated state of the cell. In other cases, the 3' end and splice site selection remains constant but nevertheless this selection controls the quantity of each mRNA and protein product produced.

We are studying early region E3 of adenovirus as a model complex transcription unit (3, 4). A schematic representation of region E3 of adenovirus 2

(Ad2) is shown in Fig. 1; the schematic is explained and references are given in the legend. All E3 mRNAs (a-j in Fig. 1) initiate from a single promoter, but their 3' ends are formed at one of two sites, E3A or E3B. The mRNAs in the E3A and E3B families each have a different spliced structure. The selection of the 3' end and splice site determines which mRNA and protein will be made. For example, mRNAs a (the major E3 mRNA), b, and c are made if splicing does not occur at the 5' splice site at nucleotide (nt) 951 (Fig. 1). These mRNAs encode gp19K, an abundant well-characterized glycoprotein (5-7). If the 951 5' splice is used, mRNAs d-h are formed and gp19K cannot be synthesized.

We wish to understand the principles that underlie the organization of the E3 genes and that apply to the selection of the RNA 3' end and splice sites. Controls of some sort must exist to dictate both the specificity and relative efficiency of these RNA processing signals. Furthermore, since most E3 RNA 3' end and splice sites are not used during "late" stages of infection (8), despite the fact that E3 is embedded in the major late transcription unit, the E3 RNA processing signals must be responsive to the differentiated state of the cell. We are using a genetic approach whereby mutations are introduced into E3 and their consequences analyzed in vivo in mutant virus-infected cells. In this report we describe the isolation of a novel class of deletion mutants that specifically enhance the 951 5' splice site. We speculate that these deletions define a cis-acting regulatory region which normally functions in wild type virus to specifically suppress the 951 5' splice site. This supressor function has been destroyed by the deletions. Suppression of the 951 5' splice site in wild type virus permits synthesis of mRNAs a, b, and c at the expense of the other mRNAs. This regulatory region lies somewhere within nt 1691-2044 in the 3' untranslated portion of mRNAs a, b, and c.

MATERIALS AND METHODS

Viruses, Cells, Protein Labeling, and Protein Analysis

The procedures have been described previously, including maintenance of KB cells, adenovirus preparation and plaque assays (9), labeling of virus-infected cell proteins with [^{35}S]Met, extraction of proteins, immunoprecipitation, and SDS-PAGE (7). The cycloheximide enhancement procedure (10) was used to label gp19K: 25 μg/ml of cycloheximide was present from 3-6 h post-infection, and labeling was with 5 μCi/ml of [^{35}S]Met (New England Nuclear, ~1000 Ci/mmol) from 6-9 h postinfection in the presence of 20 μg/ml of

1-β-D-arabinofuranosylcytosine. The antisera against gp19K and the E1B-176R protein are described in refs. 7 and 11, respectively.

Cell free translation was done using a rabbit reticulocyte extract (Bethesda Research Laboratories) with [^{35}S]Met as label and 6 μg of polyadenylated cytoplasmic RNA prepared by the cycloheximide enhancement procedure (4). Immunoprecipitation of gp19K from the cell free translation products was done as described (12).

RNA Extraction and Nuclease-Gel Analysis

Suspension cultures of KB cells were infected with mutants and maintained in cycloheximide (25 μg/ml) from 3-9 h postinfection. Cytoplasmic RNA was extracted and poly(A)$^+$ RNA was prepared by two passages through oligo(dT)-cellulose (4).

RNA samples were analyzed using the nuclease-gel procedure (13), with a ^{32}P-labeled RNA probe prepared by in vitro transcription using the SP6 RNA polymerase (14). One plasmid substrate (Figs. 4 and 5) was prepared from an Ad2 EcoRI-D/pBR322 clone containing a deletion at the BclI (1825) site (nt 1695-2076 are deleted) and with an XbaI linker inserted into the deletion. The SmaI/XbaI fragment from this mutant, representing nt -36 to 1694 in the Ad2 E3 transcription unit, was cloned between the SmaI and XbaI sites of pSP64 (14). The other plasmid substrate (Fig. 6) was a clone of the Ad2 SmaI/EcoRI fragment (representing nt -36 to 2443 in the E3 transcription unit) inserted between the SmaI and EcoRI sites of pSP65 (14). Synthesis of the ^{32}P-labeled RNA probe was essentially as described (14) except that the reaction contained 80 μM [α-^{32}P]UTP (New England Nuclear) as label. RNAsin, SP6 polymerase, pSP64, and pSP65 were obtained from Promega Biotech, and RNAse-free DNAse from Worthington.

For S1 nuclease analysis (Figs. 4 and 5), 30 μg or 15 μg of cytoplasmic RNA, or 1 μg of poly(A)$^+$ RNA, was mixed with 5 x 10^6 cpm of ^{32}P-labeled RNA probe and annealed for 18 h at 40° in 80% formamide, 0.4 M NaCl, 1 mM EDTA, and 40 mM PIPES (pH 6.7). Samples were digested with 500 units of S1 nuclease (Bethesda Research Laboratories) and electrophoresed on 5% sequencing gels as described (4). For RNAse analysis (Fig. 6), 30 μg of cytoplasmic RNA was mixed with 10^6 cpm of probe, hybridized, then digested with 40 μg/ml of RNAse A plus 10 units/ml of RNAse T1 (both from Sigma) at 30° for 1 h in a buffer containing 0.3 M NaCl (14). Samples were electrophoresed on a 6% sequencing gel.

Isolation of Plasmid and Virus Mutants

Mutations were made in the Ad2 EcoRI-D fragment cloned at the EcoRI site

in pBR322. We used a mutant of pBR322, isolated in our laboratory, which has the sequences between the ClaI and PvuII sites deleted. Deletions were made as described (15), by cleaving at the BclI (1825) or StuI (2076) sites, deleting with Bal 31 exonuclease, and recircularizing. With dl708, dl710, and dl712, a BamHI linker, d(CGGGATCCCG) (New England Biolabs) was ligated to the plasmid prior to recircularizing. The deletions were sequenced by the procedure of Maxam and Gilbert (16).

To transfer the plasmid mutations to the viral genome (see Fig. 2), plasmids were cleaved with EcoRI, digested with bacterial alkaline phosphatase, and ligated to EcoRI-cleaved Ad5 terminal protein-DNA complex prepared as described (17). The ligation mixture was transfected into KB cells (18,

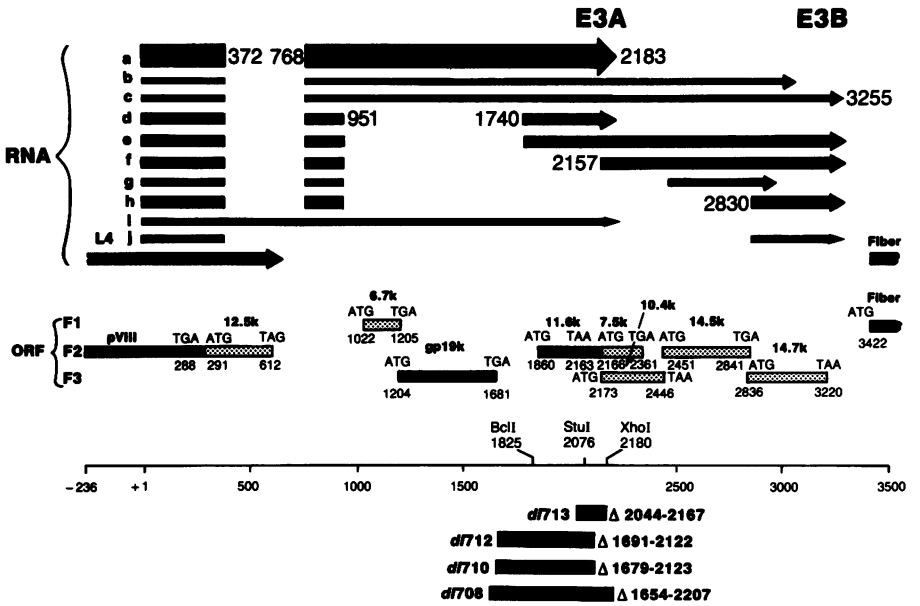


Figure 1. Schematic representation of the E3 transcription unit of Ad2, and of deletion mutants dl708, dl710, dl712, and dl713. The numbers represent nucleotide numbers in the transcription unit (3, 20). Nucleotide +1 is the first major transcription initiation site. Open reading frames (ORFs) 1 (F1), 2 (F2), and 3 (F3) are shown. Black bars indicate proven proteins, and stippled bars indicate proteins that we propose to exist (3). Proteins pVIII and fiber are "late" structural proteins coded by the L4 and fiber mRNAs, respectively. The arrows a-j represent the spliced structures of the E3A and E3B 3' coterminal families of mRNAs as determined by electron microsocopy (41, 42). The RNA 3' end and splice sites have been determined by sequence analysis of cDNA clones of E3 mRNAs (21, 45, 46) or by nuclease-gel analysis (4, 13, 47). Evidence in support of this schematic is discussed in refs. 3 and 4.

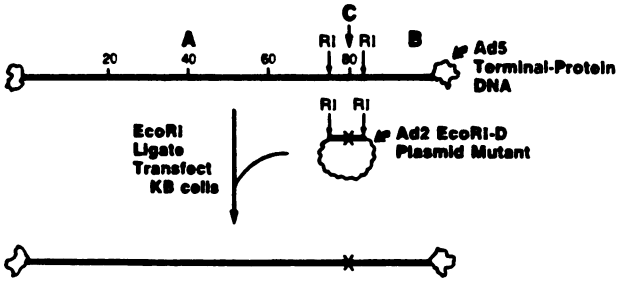


Figure 2. Method used to construct virus mutants. Mutations were made in a plasmid consisting of the Ad2 EcoRI-D fragment (map position 76-83) cloned in pBR322. After DNA sequencing the mutation, the mutated EcoRI-D fragment was ligated between the Ad5 EcoRI-A and EcoRI-B fragments, then transfected into KB cells. Plaques were picked, expanded into seed stocks, and screened by HindIII digestion (Fig. 3) for the expected structure.

19) and plaques were allowed to form. Well-separated plaques were picked and expanded into virus stocks. Virus stocks were screened for the expected deletion by isolating virus ^{32}P-labeled DNA from the Hirt supernatant (17), cleaving with HindIII, electrophoresing on 0.8% agarose gels, and autoradiographing the dried gel.

The deletion mutants, as well as their "wild type" control (rec700) are Ad5-Ad2-Ad5 recombinants. Accordingly, the complete name for the mutants has the prefix rec dl, e.g. rec dl713. For convenience, we will omit the rec prefix.

RESULTS

Construction of Virus Mutants

Deletion mutants were generated in the Ad2 EcoRI-D fragment cloned in pBR322 (see Experimental Procedures). EcoRI-D is located at position 76-83 in the Ad2 genomic map, and it represents nt -236 to 2437 in the Ad2 E3 transcription unit (3, 20). Deletions were made at the BclI, StuI, and XhoI sites at nt 1825, 2076, and 2180, respectively. Four mutants were selected for characterization; the sequences deleted in these mutants are indicated in Fig. 1.

The plasmid deletion mutations were introduced into the virus genome as shown in Fig. 2. That is, the Ad2 EcoRI-D fragment containing the mutation was ligated between the Ad5 EcoRI-A and -B fragments containing the adenovirus terminal protein. Plaques were developed on KB cells, expanded into stocks, and analyzed for the presence of the deletion. In this analysis, viral DNA

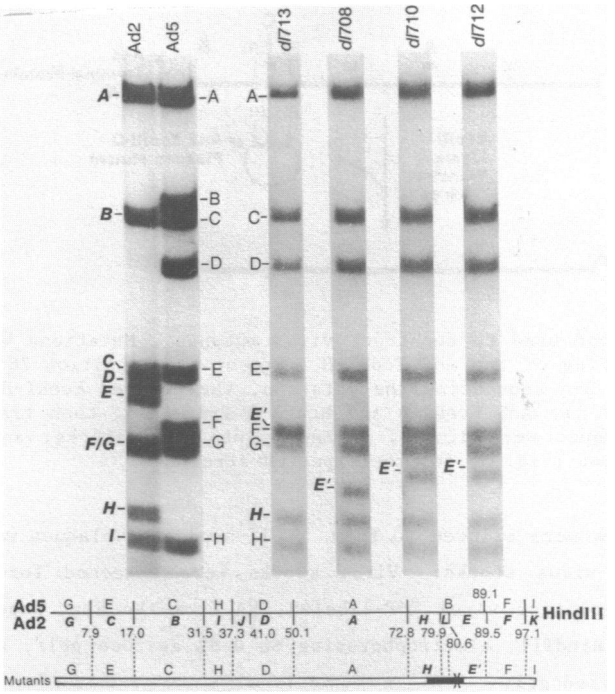


Figure 3. All mutants have the expected genomic structure. Virus DNA was labeled in vitro with ^{32}P, extracted from a Hirt supernatant, digested with HindIII, electrophoresed on a 0.8% agarose gel, and autoradiographed. Ad5 fragments are in Roman type and Ad2 fragments are in bold italics. Restriction maps and the expected mutant structures are indicated at the bottom. With mutants, Ad2 sequences are shown by black, and the deletion is indicated by the "X". The results are interpreted in the text.

was labeled in vivo with ^{32}P, and DNA from the Hirt supernatant was digested with HindIII and electrophoresed on agarose gels. The results are shown in Fig. 3. As expected, all mutants contained Ad5 fragments except HindIII-B. The mutants had a fragment that comigrates with Ad2 HindIII-H; this fragment has Ad5 sequences from map position 72.8 to 76 and Ad2 sequences from map position 76 to 79.9. The mutants had another fragment, HindIII-E', which varies in size in each mutant because it contains the deletion. This fragment has Ad2 sequences from map position 80.6 to 83 and Ad5 sequences from map position 83 to 89.1. Thus, all mutants had the expected HindIII fragments. Another virus, rec700, was isolated in a similar manner; this virus has wild type Ad2 sequences at map position 76-83 (the EcoRI-D fragment) and wild type Ad5 sequences in the remainder of the genome. We will refer to rec700 as "wild type" because it was indistinguishable from Ad2 in the studies reported

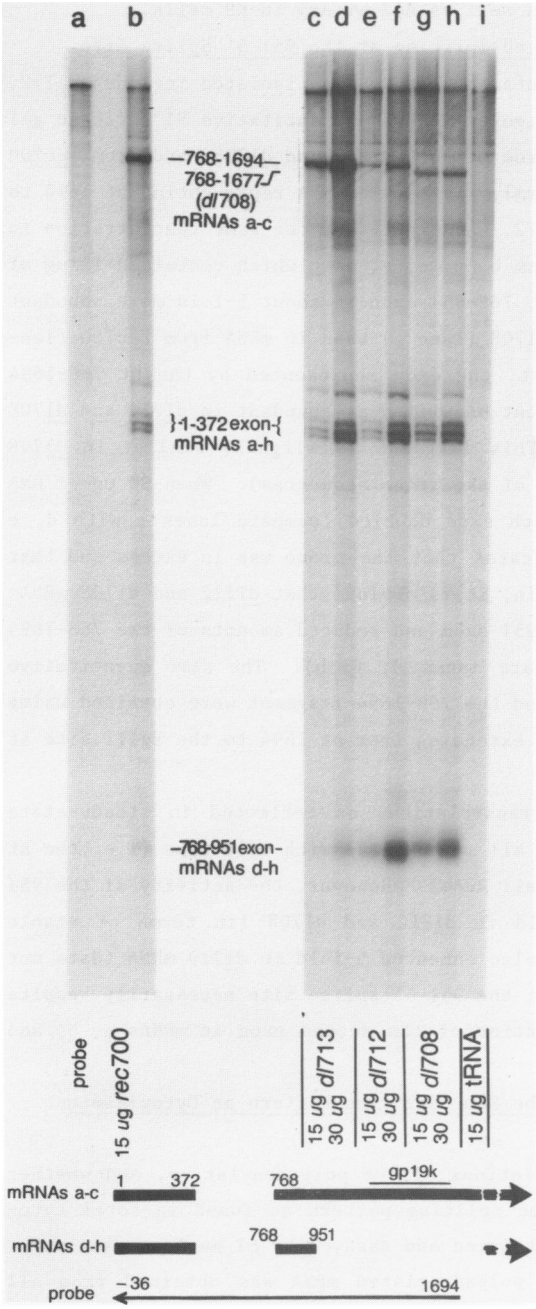


Figure 4. S1 nuclease-gel analysis of cytoplasmic mRNA from mutant-infected cells. The probe is uniformly labeled (using [α-^{32}P]UTP) RNA of 1730 nt, representing nt 1694 to -36 in the antisense strand of the E3 transcription unit. (a) probe. (b) 15 µg of rec700 RNA. (c and d) 15 µg and 30 µg of dl713 RNA. (e and f) 15 µg and 30 µg of dl712 RNA. (g and h) 15 µg and 30 µg of dl708 RNA. (i) 15 µg of tRNA. See Fig.1 for complete mRNA structures and for mutant deletions.

in this paper. All mutants grew as well as Ad2 or Ad5 in KB cells.

Mutants dl712 and dl708 Have Enhanced Splicing at the 951 5' Splice Site

The novel class of splicing mutants that we have isolated includes dl712, dl710, and dl708 (see Fig. 1). Figure 4 shows a quantitative S1 nuclease gel analysis of cytoplasmic E3 mRNAs from mutants dl712 and dl708, and from rec700 (wild type). The probe is a uniformly ^{32}P-labeled RNA representing nt 1694 to -36. The 5'-proximal exon, nt 1-372, was present at the same concentration in all RNA preparations (compare lanes b, c, e, g, each which contained 15 μg of cytoplasmic RNA). However, the nt 768-951 exon is about 5-fold more abundant in mRNA from dl712 (lane e) and dl708 (lane g) than in mRNA from rec700 (lane b) or dl713 (lane c). In contrast, the exon represented by the nt 768-1694 fragment (768-1677 in dl708) is about one-fifth as abundant in dl712 and dl708 compared to rec700 and dl713. (This fragment is slightly smaller in dl708 because the deletion includes some of the probe sequences). When 30 μg of RNA were assayed, the intensity of each exon doubled (compare lanes c with d, e with f, and g with h), which indicates that the probe was in excess and that the results are quantitative. Again, it is obvious that dl712 and dl708 mRNAs have elevated amounts of the 768-951 exon and reduced amounts of the 768-1694 fragment (768-1677 in dl708)(compare lanes d, f, h). The same quantitative differences in the 768-951 exon and the 768-1694 fragment were obtained using a different ^{32}P-labeled RNA probe, extending from nt 1694 to the BglII site at nt 566 (data not shown).

We conclude that total E3 transcription, as reflected in steady-state cytoplasmic mRNA, is the same with all mutants and with wild type (i.e. the nt 1-372 exon is equally abundant in all RNAs). However, the activity of the 951 5' splice site is enhanced 5-fold in dl712 and dl708 (in terms of stable mRNA). The 951 5' splice site is also enhanced 5-fold in dl710 mRNA (data not shown). This increased splicing at the 951 5' splice site necessarily results in a 5-fold reduction in concentration of the second exon in mRNAs a, b, and c.

Polyadenylated Mutant mRNAs Have the Same Splicing Pattern as Cytoplasmic mRNAs

We determined whether the deletions affect polyadenylation, and whether polyadenylated mRNAs have the same splicing pattern as found in total cytoplasmic mRNA. Poly(A)$^+$ mRNA was isolated and assayed by S1 nuclease analysis. As shown in Fig. 5, E3-specific polyadenylated mRNA was obtained from all mutants. Furthermore, all poly(A)$^+$ mRNA preparations had the same splicing pattern as did cytoplasmic mRNA. That is, compared to rec700 (lane a) and

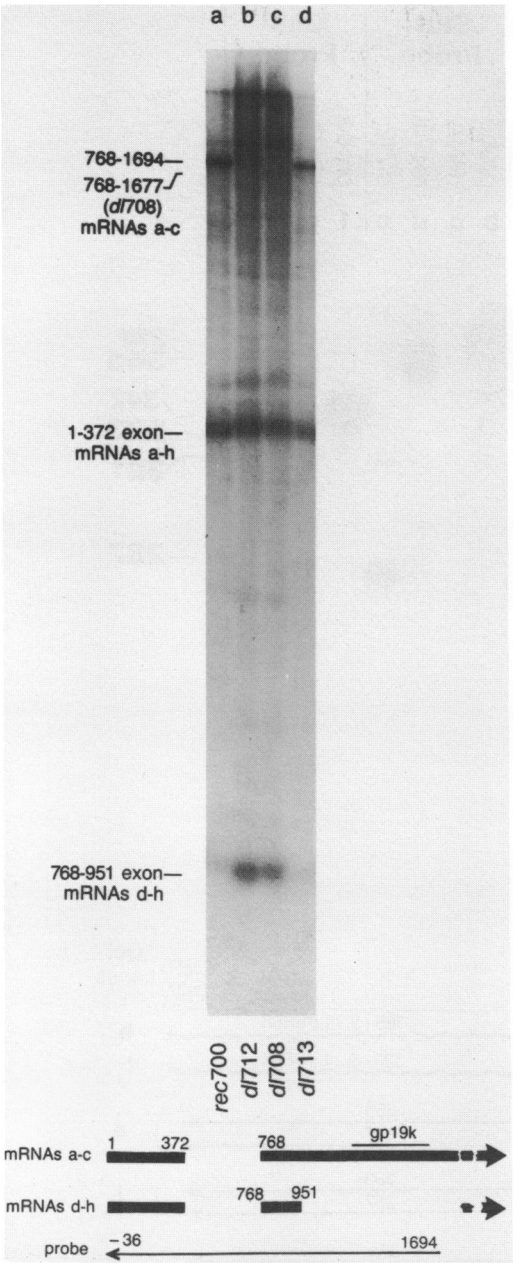
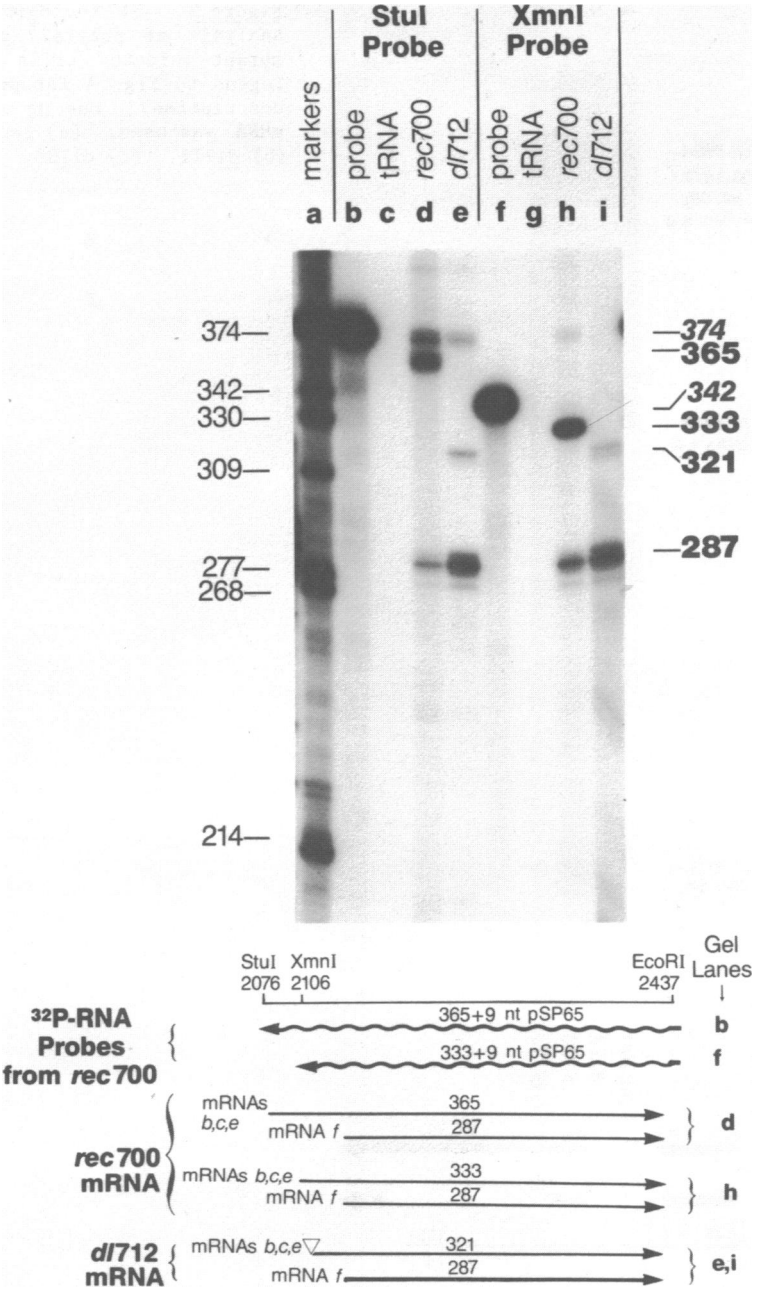


Figure 5. S1 nuclease-gel analysis of poly(A)$^+$ mRNA from mutant infected cells (see legend to Fig. 4 for probe description). One μg of poly(A)$^+$ mRNA was used. (a) rec700. (b) dl712. (c) dl708. (d) dl713.

dl713 (lane d), dl712 (lane b) and dl708 (lane c) had equivalent levels of the 1-372 exon, had 5-fold higher levels of the 768-951 exon, and had 5-fold lower levels of the 768-1694 fragment. Poly(A)$^+$ mRNA from dl710 had the same


StuI Probe
XmnI Probe
markers
probe
tRNA
rec700
dl712
probe
tRNA
rec700
dl712
a b c d e f g h i
374
342
330
309
277
268
214
374
365
342
333
321
287
StuI 2076
XmnI 2106
EcoRI 2437
Gel Lanes
32P-RNA Probes from rec700
365+9 nt pSP65
333+9 nt pSP65
b
f
rec700 mRNA
mRNAs b,c,e
mRNA f
365
287
d
333
287
h
dl712 mRNA
mRNAs b,c,e
321
287
e,i

splicing pattern as dl712 and dl708 (data not shown). Since dl708 and dl713 have the E3A 3' end sites deleted, we presume that the 3' ends of the mRNAs from these mutants were formed at the E3B site. We conclude that mRNAs from dl712, dl708, and dl710 are polyadenylated, including those mRNAs that result from the enhanced splicing at the 951 site. Furthermore, the absence of the E3A RNA 3' end signal in dl713 (the ATTAAA motif is deleted) did not affect the activity of the 372 and 951 5' splice sites or the 768 3' splice site.

Mutant dl712 Has Enhanced Splicing at the 2157 3' Splice Site

Since poly(A)$^+$ mRNAs from dl712, dl708, and dl710 have enhanced splicing at the 951 5' splice site, it follows that they must also have enhanced splicing at one or more 3' splice sites. We have used the nuclease-gel procedure to examine the abundance of the 2157 3' splice site in cytoplasmic mRNA from rec700 and dl712. The 2157 3' splice site is used by mRNA f (Fig. 1). Two ^{3}P-labeled RNA probes were used (see Fig. 6, bottom). The "StuI" probe contains 365 nt of viral sequences extending from the EcoRI (2437) site to the StuI (2076) site. The "XmnI" probe contains 333 nt of viral sequences, from the EcoRI (2437) site to the XmnI (2106) site. Both probes have 9 nt of pSP65 sequences at the 5' end.

When the StuI probe was hybridized to rec700 mRNA, two RNAse-resistant bands were observed, of 365 and 287 nt (Fig. 6, lane d). The XmnI probe gave bands of 333 and 287 nt (lane h). The 365 and 333 nt bands represent mRNAs that are co-linear with the virus sequences in the probe (mRNAs b, c, and e in Fig. 1). These bands are 9 nt smaller than the probes. The 287 nt band represents RNA that extends from the 2157 3' splice site to the EcoRI (2437) site. This is mRNA f in Fig. 1. With dl712 mRNA, the 287 nt band as well as a band of 321 nt was observed (Fig. 6, lanes e and i). The 287 nt band is several-fold more abundant in dl712 mRNA than in rec700 mRNA. The 321 nt band represents co-linear RNA extending from the deletion in dl712 (nt 2122) to the EcoRI (2437) site. We conclude that the 2157 3' splice site is enhanced in

Figure 6. RNAse-gel analysis of the 2157 3' splice site in mRNA from rec700 and dl712. The schematic at the bottom depicts the ^{32}P-labeled RNA probes used and the RNAse-resistant fragments observed. Restriction sites are numbered at the 5' base in the recognition sequence, not at the actual site of cleavage. The StuI and XmnI probes were made by cleaving at those sites, respectively, prior to transcription with SP6 polymerase. The triangle at the end of the 321 fragment arrow indicates the deletion in dl712. For the gel, heavy Roman type indicates RNAse-resistant fragments, italics indicate probes, and light Roman type indicates RNA size markers. 30 μg of cytoplasmic RNA were used. (a) markers. (b) StuI probe. (c) StuI probe versus tRNA. (d) StuI probe versus rec700 RNA. (e) StuI probe versus dl712 RNA. (f) XmnI probe. (g) XmnI probe versus tRNA. (h) XmnI probe versus rec700 RNA. (i) XmnI probe versus dl712 RNA.

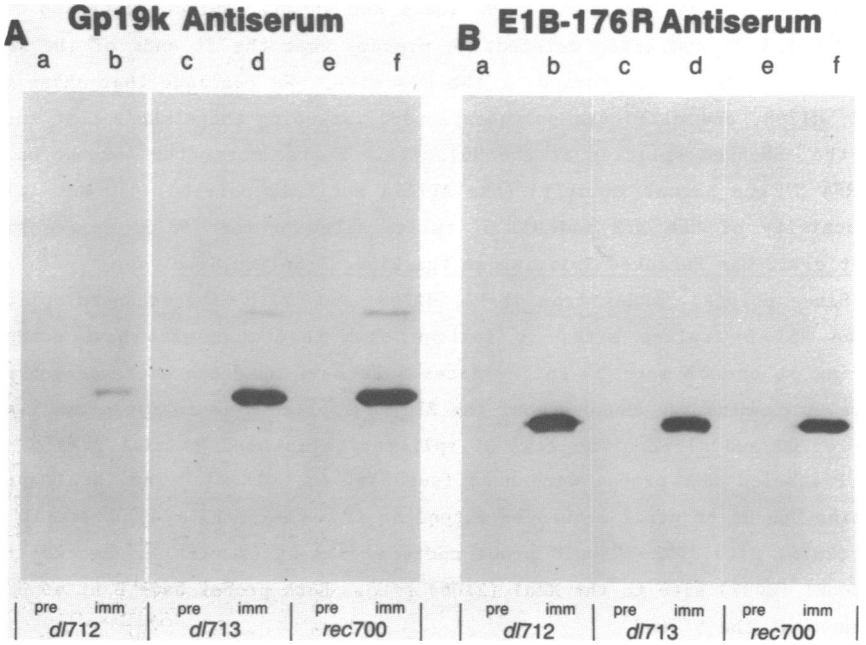


Figure 7. dl712 synthesizes reduced levels of gp19K but dl713 synthesizes wild type levels of gp19K. Gp19K and E1B-176R proteins were immunoprecipitated from the same [^{35}S]Met-labeled extract for each mutant. (A) (a and b) preimmune and anti-gp19K sera vs dl713. (c and d) preimmune and antigp19K sera vs dl713. (e and f) preimmune and anti-gp19K sera vs rec700. (B) Same as A but with anti-176R in place of anti-gp19K.

dl712 mRNA as compared to rec700 mRNA. The enhancement of the 2157 3' splice site must be an effect, not a cause, of the enhanced 951 5' splice site. This is because the 2157 3' splice site is deleted in dl708, yet the 951 5' splice site is enhanced.

Mutants dl712 and dl710 are Defective in the Synthesis of the gp19K Protein

As noted above, the mutants with enhanced splicing activity at the 951 5' splice site have reduced levels of the second exon in mRNAs a, b, and c (see Fig. 1). Since the second exon in these mRNAs encodes the major E3 protein, gp19K (6, 21), we might expect that the splicing mutants would synthesize reduced levels of gp19K. This was the case. Two representative experiments are shown in Fig. 7A and Fig. 8. In Fig. 7A, compare lanes b, d, and f. In Fig. 8, compare lane b (rec700, wild type) with lane e (dl710, splice enhancing mutant). As judged by the cpm's of [^{35}S]Met-labeled immunoprecipitable protein in several independent experiments, dl712 and dl710 synthe-

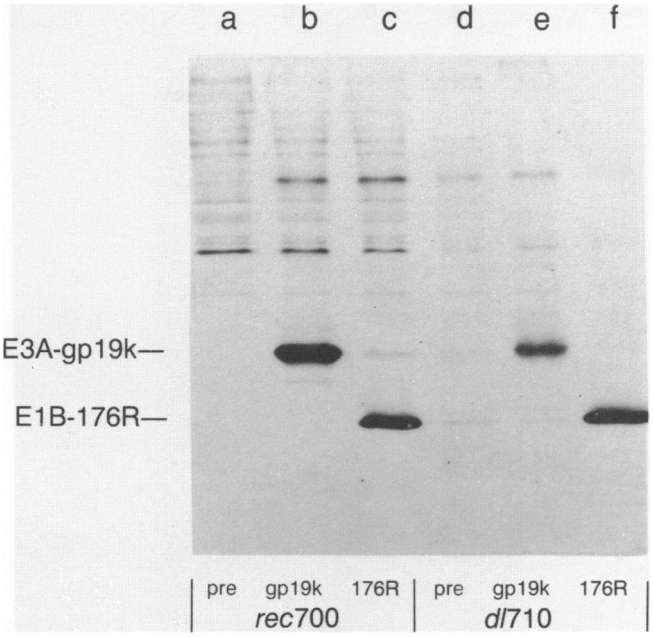


Figure 8. dl710 synthesizes reduced levels of gp19K compared to rec700 (wild type) virus. [^{35}S]Met-labeled gp19K and E1B-176 proteins were immunoprecipitated from the same extract of mutant- infected KB cells and were resolved by SDS-PAGE. (a, b, c) preimmune, anti-gp19K, anti-E1B-176R sera vs rec700. (d, e, f) preimmune, anti-gp19K, anti-E1B-176R sera vs dl710.

sized only 1/7th to 1/25th as much gp19K as did rec700 or dl713. However, all mutants synthesized about the same quantity of the 176R protein encoded by the E1B transcription unit (compare lanes b, d, and f in Fig. 7B, and compare lanes c and f in Fig. 8). The E1B-176R protein was immunoprecipitated from the same extract as was gp19K in Fig. 7 and 8. The E1B-176R data indicate that an active infection had occurred, and therefore that dl712 and dl710 in fact synthesized less gp19K. Cells infected by dl708 had the same quantitative patterns of gp19K and E1B-176R synthesis as did those infected by dl712 or dl710 (data not shown). Although the deletions in dl710 and dl708 slightly impinge on the COOH- terminus of gp19K, the deletion in dl712 is 8 bp downstream from TGA_{1681} for gp19K. Therefore, the paucity of gp19K in dl712 and presumably the other mutant extracts is not the result of unstable gp19K. Rather it indicates a true reduction in the synthesis of the protein.

mRNA from rec700, dl708, and dl712 Can be Translated In Vitro into gp19K

The splicing results discussed above argue that dl712, dl710, and dl708

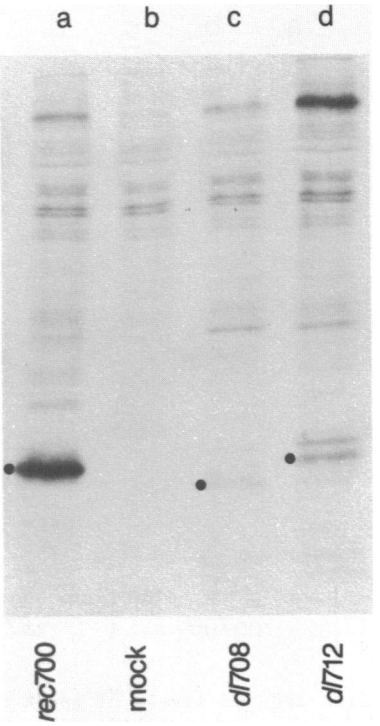


Figure 9. Cell free translation of gp19K from poly(A)$^+$ mRNA. After translation, gp19K was immunoprecipitated and analyzed by SDS-PAGE. (a) rec700 mRNA. (b) mock infected cell mRNA. (c) dl708 mRNA. (d) dl712 mRNA.

synthesized reduced levels of gp19K because the concentration of the gp19K-coding sequences in mRNA is reduced by the enhanced splicing at the 951 5' splice site. However, an additional possibility is that the mutant mRNAs are defective in translation of gp19K. As shown in Fig. 9, gp19K could be translated in a cell free system programmed with mRNA from rec700 (lane a), dl708 (lane c), and dl712 (lane d). The gp19K from dl708 was slightly smaller than gp19K from rec700 and dl712. This is expected because gp19K from dl708 is lacking eight amino acids at its COOH-terminus (wild type gp19K lacking the NH_2-terminal signal sequence has 142 amino acids). Gp19K from dl708-infected cells was also slightly smaller than wild type (data not shown). The quantity of gp19K translated from rec700 mRNA was much higher than that from dl708 or dl712 mRNA. Most likely this is a reflection of the relative abundance of gp19K-specific mRNA in the poly(A)$^+$ RNA preparations. Although these data do not exclude a gp19K translation defect in vivo for dl708 and dl712, they do exclude a gross translation defect in vitro.

DISCUSSION

Our results establish that the deletions in mutants dl712, dl708, and dl710 enhance the activity of the 951 5' splice site. This increased splicing activity reduces the concentration of mRNAs a, b, and c which encode gp19K. Accordingly, less gp19K is synthesized. The acceptors for this increased splicing are probably the natural 3' splice sites; consistent with this, we have shown that splicing at the 2157 3' splice site is enhanced in dl712 mRNA. Mutant dl713 is indistinguishable from wild type (rec700) in its splicing activity at the 951 5' splice site and in its synthesis of gp19K. Thus, the critical sequences that affect the 951 5' splice site, as defined by our deletion mutants, includes the region from nt 1691 to 2044. That is, within the 5' boundaries of the deletions in dl712 (splice enhancing mutant) and dl713 (non-enhancing mutant).

The reason that the 951 5' splice site is enhanced by deletions in the nt 1691-2044 region is unknown. However, it is remarkable that a deletion 740 nt downstream in dl712 can affect the activity of a 5' splice. Furthermore, the effect is specific for the 951 site, because the 372 5' splice site is unaffected. Indeed, since the deletions in dl712, dl710, and dl708 all remove the 3' splice site at nt 1740 (Fig. 1), we might have expected that the 951 splice would be reduced rather than elevated. It seems that our deletion mutants have defined a cis-acting region that affects the efficiency of the 951 5' splice site. A trans-acting effect, e.g. by the E3A-11.6K protein (Fig. 1), is unlikely because the deletion in dl713 (non-enhancing mutant) removes 40 amino acids from the COOH-terminus of the 101-amino acid 11.6K protein.

Our results are relevant to the mechanism of splicing. Splicing of pre-mRNA in metazoa (22-24) and yeast (25, 26) occurs by a lariat mechanism. The intron is excised in one piece, with the 5' end of the intron joined to a sequence located ~20-40 nt upstream from the 3' splice site. In yeast, this sequence, known as a TACTAAC box, is absolutely required for splicing (27-29). In metazoa, although a branchpoint site is required for splicing, a specific natural branchpoint site apparently is not essential. For example, synthetic 5' and 3' splice sites are capable of splicing together (30). Other cis-acting sequences that are required for splicing are located at and near the 5' and 3' splice sites, as indicated by the conservation of these sequences (31) and by genetic studies (e.g. see refs. 32 and 33). No other sequences appear to be obligatory for splicing, because deletions throughout introns do not prevent splicing (e.g. 33, 34), and because chimeric (35, 36)

and synthetic (30) 5' and 3' splice sites are accurately spliced. Nevertheless, our data establish that distant downstream sequences can affect 5' splicing efficiency. With E3 mRNAs, therefore, splicing does not occur by simple juxtapositioning of 5' and 3' splice sites, with other portions of the RNA having no effect. Rather, sequences within the nt 1691-2044 region apparently interact with the 951 5' splice site. Of course, we do not know if the interaction is direct through RNA-RNA and/or RNA-protein-RNA contacts, or indirect through an influence on overall RNA tertiary structure. There is considerable evidence that specific intron secondary and tertiary structure is important for splicing of rRNA and mitochondrial pre-mRNA of lower eukaryotes (32). However, if RNA secondary and tertiary structure does affect splicing efficiency and specificity in metazoa (34, 37) a specific structure is clearly not essential.

There are other reports that sequences distant from splice sites can influence splicing. Deletions near the 3' end of the first exon for SV40 small t antigen mRNA (34) or in the first exon for SV40 late mRNAs (38) reduced the splicing of these mRNAs. Deletions in the intron for MLV env RNA inhibited splicing (39), and a base substitution in the large intron of β-globin RNA activated an upstream cryptic 3' splice site (40). However, our mutations are unique because they elevate rather than decrease splicing.

Our results may be relevant to the control of differential splicing in complex transcription units. One model depicts differential splicing as a stochastic process, with different splices occurring at a certain frequency as a consequence of sequence context at the splice sites and of RNA tertiary structure in the splicing ribonucleoprotein complex. Thus, a mutation that perturbs one of the splicing pathways would be expected to perturb other splicing pathways. Our results are not consistent with this model in its simplest form because the effects that we observed are specific to the 951 5' splice site and to the nt 1691-2044 region. The 1691-2044 deletions did not affect the 372 5' splice. Furthermore, the deletion in dl713 had no effect on either 5' splice site, even though dl713 has the 2157 3' splice site deleted, and even though dl713 fails to synthesize RNA 3' ends at the E3A site (B. Bhat, H. Brady, and W. Wold, unpublished results). That is, most of the RNA precursors in dl713 are processed into mRNAs c and h (see Fig. 1). Although the pattern of E3 mRNAs is grossly altered in dl713, the 951 5' splice site is not enhanced.

Another model of differential splicing which is particularly intriguing holds that complex transcription units contain cis-acting regions that regu-

late the efficiency of splice site selection. Perhaps we have discovered one of the elements in the E3 splicing control scheme. Specifically, the function of the nt 1691-2044 region in wild type virus would be to suppress the 951 5' splice site. The deletions in dl712, dl710, and dl708 destroy this suppressor function. In wild type, suppression of the 951 5' splice site results in the synthesis of mRNAs a, b, and c at the expense of the other E3 mRNAs. This may explain why mRNA a is the most abundant E3 mRNA (41, 42), and why gp19K which is encoded by mRNA a (5, 21) is the most abundant E3 protein (5, 7, 43, 44). We note that these 951 suppressor sequences are located within the 3' untranslated region of mRNAs a, b, and c; this is a logical location because their presence there would ensure that the 951 5' splice site remains suppressed in the final mRNA product. For example, without suppression, mRNA a might always be spliced into mRNA d, or mRNA c into mRNAs e-h. mRNAs a and c both contain the 5' and 3' splice sites required to carry out these splices.

ACKNOWLEDGEMENTS

S. Magie isolated some of the plasmid mutants. The virus mutant plaques were developed by N. Takemori. We thank D. Rosser and A. Berk for the anti-E1B-176R sera. This work was supported by Public Health Service grants CA24710 and GM31276, and by Research Career Development Award CA00650 (to W.S.M.W.) S.L.D. was supported by Public Health Service grant GM30365.

*To whom correspondence should be addressed

REFERENCES

1. Darnell, J.E., Jr. (1982) Nature (London) 297, 365-371.
2. Nevins, J.R. (1983) Annu. Rev. Biochem. 52, 441-466.
3. Cladaras, C., and Wold, W.S.M. (1985) Virology 140, 28-43.
4. Cladaras, C., Bhat, B., and Wold, W.S.M. (1985) Virology 140, 44-54.
5. Persson, H., Jansson, M., and Philipson, L. (1980) J. Mol. Biol. 136, 375-394.
6. Persson, H., Jörnvall, H., and Zabielski, J. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 6349-6353.
7. Wold, W.S.M., Cladaras, C., Deutscher, S.L., and Kapoor, Q.S. (1985) J. Biol. Chem. 260, 2424-2431.
8. Chow, L.T., Broker, T.R. (1978) Cell 15, 497-510.
9. Green, M., and Wold, W.S.M. (1979) Methods Enzymol. 58, 425-435.
10. Harter, M.L., Shanmugam, G., Wold, W.S.M., and Green, M. (1976) J. Virol. 19, 232-242.
11. Spindler, K.R., Rosser, D.S.E., and Berk, A.J. (1984) J. Virol. 49, 132-141.
12. Wold, W.S.M., Cladaras, C., Magie, S.C., and Yacoub, N. (1984) J. Virol. 52, 307-313.
13. Berk, A.J., and Sharp, P.A. (1977) Cell 12, 721-732.

14. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. (1984) Nucl. Acids Res. 12, 7035-7056.
15. Lee, D.C., Roeder, R.G., and Wold, W.S.M. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 41-45.
16. Maxam, A.M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
17. Chinnadurai, G., Chinnadurai, S., and Brusca, J. (1979) J. Virol. 32, 623-628.
18. Graham, F.L., and van der Eb, A.J. (1973) Virology 52, 456-467.
19. Frost, E., and Williams, J. (1978) Virology 91, 39-50.
20. Hérissé, J., Courtois, G., and Galibert, F. (1980) Nucl. Acids Res. 8, 2173-2192.
21. Ahmed, C.M.I., Chanda, R.S., Stow, N.D., and Zain, B.S. (1982) Gene 20, 337-344.
22. Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984) Science 225, 898-903.
23. Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984) Cell 38, 317-331.
24. Zeitlin, S., and Efstratiadis, A. (1984) Cell 39, 589-602.
25. Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E., and Abelson, J. (1984) Cell 39, 611-621.
26. Rodriguez, J.R., Pikielny, C.W., and Rosbash, M. (1984) Cell 39, 603-610.
27. Langford, C.J., and Gallwitz, D. (1983) Cell 33, 519-527.
28. Langford, C.J., Klinz, F.-J., Donath, C., and Gallwitz, D. (1984) Cell 36, 645-653.
29. Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983) Cell 34, 395-403.
30. Rautmann, G., Matthes, H.W.D., Gait, M.J., and Breathnach, R. (1984) EMBO J. 3, 2021-2028.
31. Mount, S.M. (1982) Nucl. Acids Res. 10, 459-472.
32. Cech, T.R. (1983) Cell 34, 713-716.
33. Wieringa, B., Hofer, E., and Weissmann, C. (1984) Cell 37, 915-925.
34. Khoury, G., Gruss, P., Dhar, R., and Lai, C.-J. (1979) Cell 18, 85-92.
35. Khoury, G., Alwine, J., Goldman, N., Gruss, P., and Jay, G. (1980) J. Virol. 36, 143-151.
36. Chu, G., and Sharp, P.A. (1981) Nature (London) 289, 378-382.
37. Kühne, T., Wieringa, B., Reiser, J., and Weissmann, C. (1983) EMBO J 2, 727-733.
38. Ghosh, P.K., Piatak, M., Mertz, J.E., Weissman, S.M., and Lebowitz, P. (1982) J. Virol. 44, 610-624.
39. Hwang, S.L.-H., Park, J., and Gilboa, G. (1984) Molec. Cellul. Biol. 11, 2289-2297.
40. Dobkin, C., Pergolizzi, R.G., Bahre, P., and Bank, A. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 1184-1188.
41. Chow, L.T., Broker, T.R., and Lewis, J.B. (1979) J. Mol. Biol. 134, 265-303.
42. Kitchingman, G.R., and Westphal, H. (1980) J. Mol. Biol. 137, 23-48.
43. Wold, W.S.M., and Green, M. (1979) J. Virol. 30, 297-310.
44. Ross, S.R., Flint, S.J., and Levine, A.J. (1980) Virology 100, 419-432.
45. Ahmed, C.M.I., Chanda, R., Stow, N., and Zain, B.S. (1982) Gene 19, 297-301.
46. Stalhandske, P., Persson, H., Perricaudet, M., Philipson, L., and Pettersson, U. (1983) Gene 22, 157-165.
47. Berk, A.J., and Sharp, P.A. (1978) Cell 14, 695-711.