MOLECULAR AND CELLULAR BIOLOGY, Jan. 1994, p. 437–445
0270-7306/94/$04.00+0

Vol. 14, No. 1

Human Adenovirus Encodes Two Proteins Which Have Opposite Effects on Accumulation of Alternatively Spliced mRNAs

KATARINA NORDQVIST,† KARIN ÖHMAN, AND GÖRAN AKUSJÄRVI*

Department of Cell and Molecular Biology, The Medical Nobel Institute, Karolinska Institutet, S-171 77 Stockholm, Sweden

Received 23 June 1993/Returned for modification 17 August 1993/Accepted 30 September 1993

All mRNAs expressed from the adenovirus major late transcription unit have a common, 201-nucleotide-long 5′ leader sequence, which consists of three short exons (the tripartite leader). This leader has two variants, either with or without the i-leader exon, which, when present, is spliced between the second and the third exons of the tripartite leader. Previous studies have shown that adenovirus early region 4 (E4) encodes two proteins, E4 open reading frame 3 (E4-ORF3) and E4-ORF6, which are required for efficient expression of mRNAs from the major late transcription unit. These two E4 proteins appear to have redundant activities, and expression of one has been shown to be sufficient for efficient major late mRNA accumulation during a lytic virus infection. In this report, we provide evidence that E4-ORF3 and E4-ORF6 both regulate major late mRNA accumulation by stimulating constitutive splicing. Moreover, we show that the two proteins have different effects on accumulation of alternatively spliced tripartite leader exons. In a DNA transfection assay, E4-ORF3 was shown to facilitate i-leader exon inclusion, while E4-ORF6 preferentially favored i-leader exon skipping. In addition, E4-ORF3 and E4-ORF6 had the same effects on accumulation of alternatively spliced chimeric β-globin transcripts. This finding suggests that the activities of the two proteins may be of more general relevance and not restricted to splicing of major late tripartite leader-containing pre-mRNAs. Interestingly, E4-ORF6 expression was also shown to stimulate i-leader exon skipping during a lytic virus infection.

Alternative RNA splicing plays a central role in the posttranscriptional control of gene expression in higher eukaryotes. By combining different 5′ and 3′ splice sites in a pre-mRNA, multiple structurally related, but functionally distinct, proteins can be generated from a single gene (reviewed in reference 30). This results in a more economic use of the genetic information, a consideration of particular importance for viruses which have small genomes and therefore need to compress the coding information to more efficiently support replication. Thus, viruses have proven to be valuable model systems with which to study mechanisms involved in the regulation of alternative splicing.

The adenovirus system provides an excellent opportunity to study regulatory events which are manifested at the level of RNA splicing. Structural studies of mRNAs expressed at different stages of the infectious cycle have shown that most adenovirus transcription units encode two or more alternatively spliced mRNAs, many of which give rise to proteins with unique biological activities (see reference 3 for a review). Furthermore, the relative concentrations of alternatively spliced viral mRNAs are subjected to a temporal control and vary during the infectious cycle. This temporal shift in mRNA expression is of vital importance for virus growth since it allows for selective expression of specific subclasses of viral proteins at different time points during infection, for example production of mRNAs encoding the structural proteins of the viral capsid late after infection.

Most capsid proteins are translated from mRNAs expressed from the adenovirus major late transcription unit (Fig. 1). This unit generates a primary transcript of approximately 28,000 nucleotides that is processed into a minimum of 20 cytoplasmic mRNAs which are grouped into five families (late region 1 [L1] to L5), each of which consists of multiple alternatively spliced mRNAs with coterminal 3′ ends (reviewed in reference 3). An important consequence of the processing pathway is that all of the major late mRNAs receive a common tripartite leader sequence of 201 nucleotides at the 5′ end (Fig. 1). The accumulation of the major late mRNAs has been shown to be regulated at multiple levels during the infectious cycle (18), including a temporal control of alternative RNA splice site choice. For example, the i-leader exon (Fig. 1) is efficiently retained between the second and third tripartite leader exons in major late mRNAs expressed early after infection, while it is efficiently skipped late after infection (2, 23).

By using a short-term DNA transfection assay, we have previously shown that adenovirus early region 4 (E4) stimulates both nuclear and cytoplasmic tripartite leader RNA accumulation posttranscriptionally (24). This enhancement was intron dependent and required either of two E4 proteins, E4 open reading frame 3 (E4-ORF3) or E4-ORF6 (25). We have further demonstrated that E4 gene expression is necessary for normal accumulation of alternatively spliced late mRNAs during the lytic infection (25). The E4-ORF3 and E4-ORF6 proteins appear to have redundant activities also in virus growth, and expression of one seems to be sufficient to substitute for the whole E4 region for establishment of an essentially wild-type productive infection (4, 11, 12, 14). Adenovirus mutants that carry large deletions in E4 are severely defective in accumulation of nuclear and cytoplasmic RNA derived from the major late transcription unit (9, 29, 38). This deficiency is not due to a reduced transcriptional activity of the major late promoter in E4 mutant-infected cells (29). More likely, E4 gene expression is

* Corresponding author. Mailing address: Dept. of Cell and Molecular Biology, Karolinska Institutet, S-171 77 Stockholm, Sweden. Phone: 46-8-728 7330. Fax: 46-8-34 88 19. Electronic mail address: goran.akusjarvi@postmac.cmb.ki.se.

† Present address: Laboratory of Developmental Genetics, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom.

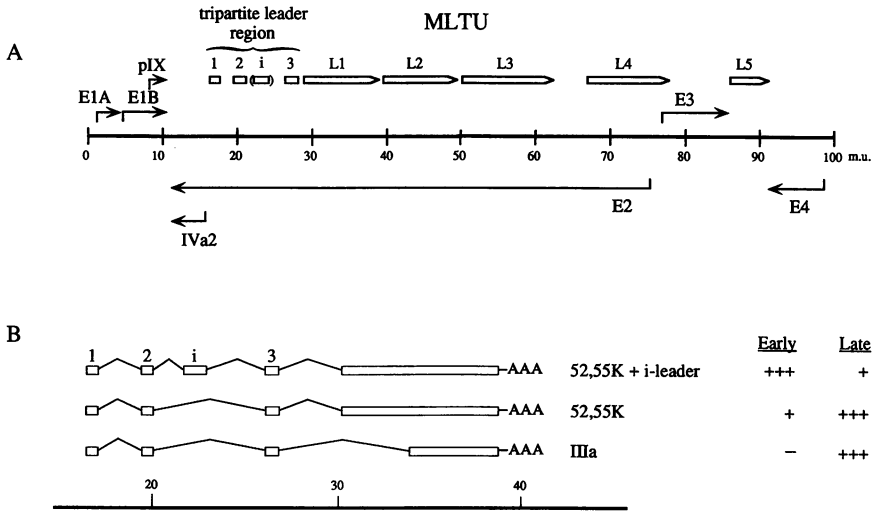


FIG. 1. Organization of the adenovirus genome. (A) Schematic representation of the transcription units encoded by the adenovirus type 2 genome. Thin lines illustrate transcription units expressed at early and intermediate times of infection, whereas thick arrows designate regions expressed late after infection. The major late transcription unit (MLTU) encodes five families of mRNAs (L1 to L5). All mRNAs from this unit receive a common set of 5′ leaders through splicing, the tripartite leader region. m.u., map units. (B) Spliced structures of the major mRNAs expressed from the L1 region and their relative expression during infection.

required for stable nuclear accumulation of the major late pre-mRNA (29). Collectively, the available data are compatible with the hypothesis that the E4-ORF3 and E4-ORF6 products regulate major late gene expression at the level of tripartite leader assembly.

In this report, we show that the E4-ORF3 and E4-ORF6 products have different effects on accumulation of alternatively spliced mRNAs from the major late transcription unit. In a transient DNA transfection assay, E4-ORF3 was found to facilitate i-leader exon inclusion, whereas E4-ORF6 predominantly facilitated i-leader exon skipping. In agreement with the hypothesis that the E4 products are important for alternative splicing of viral transcripts, we also show that E4-ORF6 expression modulates i-leader exon skipping during lytic virus growth. Interestingly, the opposite effects of E4-ORF3 and E4-ORF6 on accumulation of alternatively spliced mRNA were also reproduced on a chimeric β-globin transcript containing multiple exons, suggesting that the biological activities of the two proteins may be of more general relevance in the processing of viral and cellular RNAs.

MATERIALS AND METHODS

Plasmid constructions. pWtCAT was constructed by introducing the *Bam*HI-*Eco*RI fragment (converted to a *Hin*dIII fragment), containing a truncated version of the rabbit β-globin large intron with flanking exon sequences, from pBSAL4 (17) into the *Hin*dIII site upstream of the chloramphenicol acetyltransferase (CAT) coding sequence in pL1CAT (24). pDon*CAT, pAcc*CAT, and pDmCAT were similarly constructed, using the *Bam*HI-*Eco*RI fragments from pBSAL5, pBSAL8, and pBSAL11, respectively (17). The new plasmids were verified by DNA sequencing. pDDCAT, pDDDCAT, and pD*DCAT were constructed by introducing multiples of a *Hin*dIII-*Apa*I fragment from pWtCAT or pDon*CAT into the *Hin*dIII site upstream of the β-globin sequence in pWtCAT. The *Apa*I site, located 10 bp upstream of the natural β-globin branch site, was converted to *Hin*dIII by linker addition. Plasmids preTripCAT, pCMVE4, pCMVE4-ORF3, and pCMVE4-ORF6 have previously been described (24, 25).

Virus infection and RNA blot analysis. Subconfluent monolayers of HeLa cells were infected at a multiplicity of 20 fluorescence-forming units per cell (26) with wild-type adenovirus type 2, H2*dl*808 (38), H5ilE4I (11), or H5*dl*366-ORF3 (12). Total cytoplasmic RNA was prepared 16, 20, and 24 h postinfection (32). Five micrograms of cytoplasmic RNA from H2*dl*808-infected cells, 0.75 μg of cytoplasmic RNA from H5*dl*366-ORF3-infected cells, and 0.25 μg of cytoplasmic RNA from adenovirus type 2- and H5ilE4I-infected cells were electrophoresed in a 1% agarose gel containing 1.8% formaldehyde, transferred to a nitrocellulose filter, and hybridized to a ^{32}P-labelled L1-specific probe as previously described (18). The results were quantitated by PhosphorImager scanning.

DNA transfection and RNA preparation. Conditions for transfection of 293 cells by the calcium phosphate coprecipitation technique and isolation of cytoplasmic RNA were as previously described (25, 33).

S1 endonuclease analysis. Total cytoplasmic RNA (20 μg) was hybridized to the 5′-end-labelled DNA probes specified in the figure legends. S1 nuclease cleavage and gel electrophoresis were performed as previously described (24, 33). Quantitative results were obtained by PhosphorImager scanning.

PE reaction. Before the primer extension (PE) reaction, RNA was subjected to DNase treatment to remove contaminants of transfected plasmid DNA. Ten micrograms of cytoplasmic RNA was incubated, at 37°C for 20 min, with 20 U of RNasin (Promega) and 1 U of RQ1 RNase-free DNase (Promega) in a buffer consisting of 40 mM Tris (pH 8.0), 10 mM NaCl, 6 mM $MgCl_2$, and 5 mM dithiothreitol. The reaction mixture was phenol-chloroform extracted twice and ethanol precipitated, using KCl as the salt. DNase-treated cytoplasmic RNA (1.5 μg) was incubated in 70°C for 10 min with a 20-nucleotide-long CAT primer (5′-CAACGGTGG TATATCCAGTG-3′) which hybridizes to nucleotides +15

to +34 relative to the AUG in the CAT coding sequence. The PE reaction was performed at 44°C for 1 h in a buffer containing 40 mM Tris (pH 7.5), 30 mM KCl, 3 mM $MgCl_2$, 10 mM dithiothreitol, 0.5 mM deoxynucleoside triphosphates (dNTPs), and 200 U of Superscript reverse transcriptase (GIBCO BRL). The reaction was terminated by phenol-chloroform extraction, and the cDNA was concentrated by ethanol precipitation.

DNA amplification. PE reactions and PCRs were done under conditions in which both primers and enzymes were in excess. The PCRs were also titrated to minimize the risk of reaching saturated conditions. To visualize PCR products, a fraction of the CAT primer was 5′ end labelled before being added to the PCR reaction. Half of the PE reaction was used for PCR. The cDNA was amplified in a final volume of 50 μl containing the following reagents: 15 pmol of each primer, 0.8 mM dNTPs, 2.5 mM $MgCl_2$, PCR buffer supplied with *Taq* DNA polymerase, and 5 U of *Taq* DNA polymerase (Promega). Samples were overlaid with mineral oil and amplified by 23 cycles of 94°C (1 min), 50°C (2.5 min), and 72°C (2 min). The primers used were the CAT primer described above and the L1 primer (5′-TCTCTTCCGC ATCGCTGT-3′), which hybridizes to nucleotides +3 to +20 relative to the major late cap site. The reaction products were purified by phenol-chloroform extraction, ethanol precipitated, and resolved on a 2.5% agarose gel or a 5% native polyacrylamide gel. The most abundant PCR-amplified fragments were purified from a low-melting-point agarose gel, using the Magic PCR Preps DNA purification system (Promega), and sequenced by using the Taq Dye Deoxy Terminator Cycle sequencing kit and a model 373A automated DNA sequencer (Applied Biosystems).

RESULTS

E4 enhancement of mRNA accumulation requires wild-type splicing signals. We have previously shown that E4 enhancement of tripartite leader mRNA accumulation requires an intron. This enhancement was not specific to viral introns since the wild-type rabbit β-globin intron also was responsive, although with a slightly reduced efficiency compared with the tripartite leader region (24). To determine whether the requirement for an intron correlated with RNA splicing or was dependent on other intron signals, we constructed a set of CAT reporter plasmids based on the rabbit β-globin large intron (17). This experimental strategy was chosen because point mutants that destroy the activity of the 5′ and 3′ splice sites in this intron have been extensively characterized both in vitro and in vivo (1, 17). We constructed four CAT reporter plasmids with a truncated version of the wild-type β-globin large intron (Fig. 2A): pWtCAT, containing wild-type 5′ and 3′ splice sites; pDon*CAT, containing a G-to-A mutation at the β-globin 5′ splice site; pAcc*CAT, containing an A-to-U mutation at the 3′ splice site; and pDmCAT, containing both point mutations. Except for the specific point mutations, the structures of the CAT reporter constructs were identical.

As expected from previous results (24), cotransfection of pKGO-895, encoding the entire E4 transcription unit, stimulated mRNA accumulation in pWtCAT-transfected cells about fivefold (Fig. 2B). Interestingly, mutations inactivating both the wild-type 5′ and 3′ splice sites completely abolished the E4 effect (pDmCAT; Fig. 2B), suggesting that E4 enhancement requires bona fide RNA splice sites. The 5′ splice site mutant (pDon*CAT; Fig. 2B) demonstrated the same phenotype as the double mutant, suggesting that the β-globin

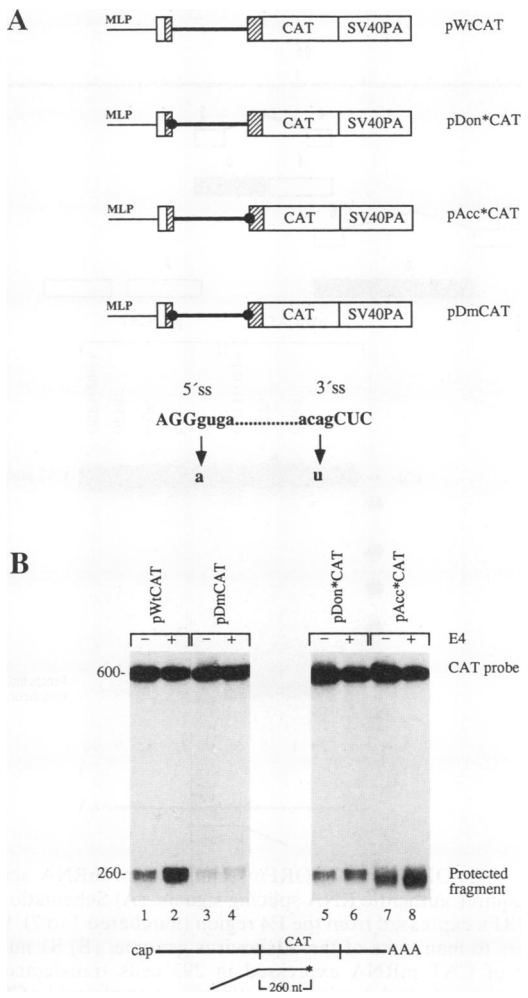


FIG. 2. E4 enhancement requires wild-type splicing signals. (A) Schematic drawing of plasmids used in the experiment. MLP, major late promoter; CAT, CAT coding sequence; SV40PA, simian virus 40 early poly(A) addition sequence. Hatched boxes represent rabbit β-globin sequences; dots represent locations of point mutations. The lower part shows the nucleotide sequence at the 5′ and 3′ splice sites (ss) and the nucleotide change present in the respective point mutant. (B) S1 nuclease analysis of CAT mRNA levels in 293 cells transfected with reporter plasmids and carrier DNA (lanes −) or pKGO-895 encoding adenovirus type 2 E4 (lanes +). Numbers on the left represent sizes of protected fragments in nucleotides (nt). The lower part shows a schematic picture of the S1 probe used in the experiment.

5′ splice site was critical for E4 responsiveness. In contrast, the mutant that inactivated the wild-type β-globin 3′ splice site (pAcc*CAT; Fig. 2B) was still E4 responsive, although efficiency was reduced 2.5-fold compared with pWtCAT. Sequencing of PCR products was used to verify RNA structures. As expected, unspliced RNA predominantly accumulated in pDon*CAT-transfected cells, whereas the 3′ splice site mutant (pAcc*CAT) RNA was processed to a cryptic 3′ splice site with a suboptimal sequence composition located +49 relative to the wild-type β-globin 3′ splice site (data not shown). From these results, we conclude that the E4-mediated enhancement of mRNA accumulation requires bona fide RNA splice sites.

E4-ORF3 and E4-ORF6 stimulation of mRNA accumulation

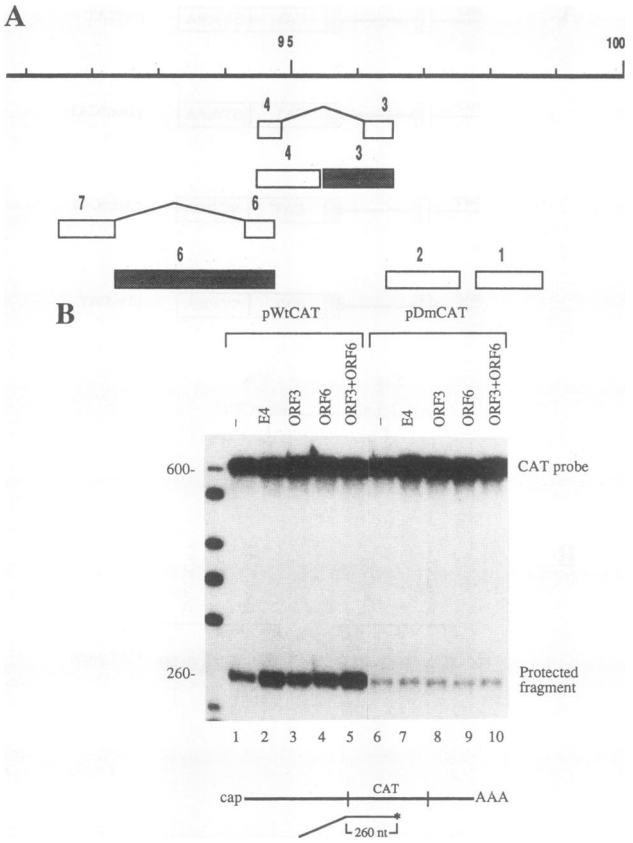


FIG. 3. E4-ORF3 and E4-ORF6 stimulation of mRNA accumulation requires authentic RNA splicing signals. (A) Schematic drawing of ORFs expressed from the E4 region (numbered 1 to 7). 95 and 100 refers to map units of the adenovirus genome. (B) S1 nuclease analysis of CAT mRNA expressed in 293 cells transfected with reporter plasmids and carrier DNA (lanes −) or plasmid pCMVE4 (lanes E4), pCMVE4-ORF3 (lanes ORF3), pCMVE4-ORF6 (lanes ORF6), or pCMVE4-ORF3 plus pCMVE4-ORF6 (lanes ORF3+ORF6). Numbers to the left represent sizes of protected fragments in nucleotides (nt).

requires the presence of authentic RNA splice sites. The E4 transcription unit is very complex. As many as 24 alternatively spliced mRNAs are produced. These encode a minimum of seven E4 proteins (Fig. 3A) (7, 35, 37). Of these, E4-ORF3 and E4-ORF6 appear to be the key proteins since expression of one is sufficient to establish an essentially wild-type virus infection (4, 11, 12, 14). We have recently used cytomegalovirus expression vectors encoding single E4-ORFs to show that the same two E4 proteins, E4-ORF3 and E4-ORF6, are the E4 products that facilitate accumulation of tripartite leader containing mRNAs during transient transfections (25).

To determine whether E4-ORF3 and E4-ORF6 stimulation of mRNA accumulation required authentic splice sites, we transfected pCMVE4-ORF3 and/or pCMVE4-ORF6 together with reporter plasmid pWtCAT or pDmCAT. As shown in Fig. 3B, both ORFs were individually able to stimulate cytoplasmic wild-type mRNA production approximately two- to threefold (lanes 3 and 4). In contrast, they had no stimulatory effect on cytoplasmic mRNA accumulation in pDmCAT-transfected cells (lanes 8 and 9). The same results were obtained with nuclear RNA (data not shown). The effects of E4-ORF3 and E4-ORF6 appear to be additive since coexpression of both gave the fivefold stimulatory activity observed with a plasmid encoding all E4 products (compare lanes 2 and 5). Collectively, these results suggest that E4-ORF3 and E4-ORF6, directly or indirectly, may be involved in RNA splice site recognition.

E4-ORF6 promotes exon skipping during a virus infection. L1 encodes two predominant cytoplasmic mRNAs, the 52,55K and IIIa mRNAs, which differ in the use of alternative 3′ splice sites (Fig. 1 and 4). The 52,55K mRNA exists in two major forms that vary in the composition of 5′ leaders: 52,55K + i (species a), containing the i-leader exon, and 52,55K (species b), lacking the i-leader exon. The abundance of the two mRNAs varies during the infectious cycle. Thus, species a, containing the i-leader exon, is the major L1 mRNA expressed early after infection. In contrast, at late time points, the relative concentration of this mRNA diminishes in favor of the tripartite leader containing 52,55K mRNA (species b) and the late-specific IIIa mRNA (species c).

Since E4 products appear to regulate tripartite leader assembly during a lytic virus infection (25), we determined whether E4-ORF3 and E4-ORF6 were individually able to promote i-leader exon skipping. In this experiment, HeLa cells were infected with wild-type adenovirus type 2 or E4 deletion mutant H2*dl*808 (E4$^-$ [38]), H5ilE4I (encoding only E4-ORF6 and E4-ORF6/7 [11]), or H5*dl*366-ORF3 (encoding only E4-ORF3 [12]). Cytoplasmic RNAs were prepared 16, 20, or 24 h postinfection, and the structures of the L1 mRNAs were analyzed by Northern (RNA) blot hybridization (Fig. 4). Note that 20-fold more RNA was loaded from H2*dl*808-infected cells compared with the other infections. This was necessary because the absence of E4 expression results in a drastic reduction in late RNA accumulation (9, 38). As shown in Fig. 4A and quantitated in Fig. 4B, the absence of E4 expression (H2*dl*808) resulted, at all time points tested, in an incomplete splicing of the tripartite leader, with a large fraction of the 52,55K mRNA population retaining the i-leader exon (Fig. 4A, species a). In contrast, the late-specific 52,55K mRNA (Fig. 4A, species b) was predominantly expressed in H5ilE4I-infected cells, which encode only E4-ORF6 and E4-ORF6/7. In fact, i-leader exon skipping was slightly more efficient than during wild-type virus infection (Fig. 4B), possibly because the mRNA encoding the E4-ORF6 protein is a minor E4 species expressed during a wild-type infection (37). In H5ilE4I-infected cells, it is a major E4 RNA species (11). Since we have previously shown that the E4-ORF6/7 product has no effect on accumulation of tripartite leader mRNA (25), we attribute the enhanced efficiency of i-leader exon skipping in H5ilE4I-infected cells to the expression of the E4-ORF6 product. Collectively, these results suggest that E4 expression is necessary for efficient i-leader exon skipping during lytic virus growth and, furthermore, that E4-ORF6 expression is sufficient to establish the wild-type efficiency of i-leader exon skipping.

The phenotype of the E4-ORF3-expressing virus (H5*dl*366-ORF3) was somewhat more complex. Expression of the E4-ORF3 protein resulted in an overall increase in L1 mRNA expression. However, the fraction of 52,55K mRNAs retaining the i-leader exon was still elevated compared with a wild-type infection at the earliest time points tested. At 16 and 20 h postinfection, the proportion of i-leader exon containing 52,55K mRNA was significantly higher compared with wild-type adenovirus type 2 (Fig. 4B), suggesting that the E4-ORF3 protein does not induce as

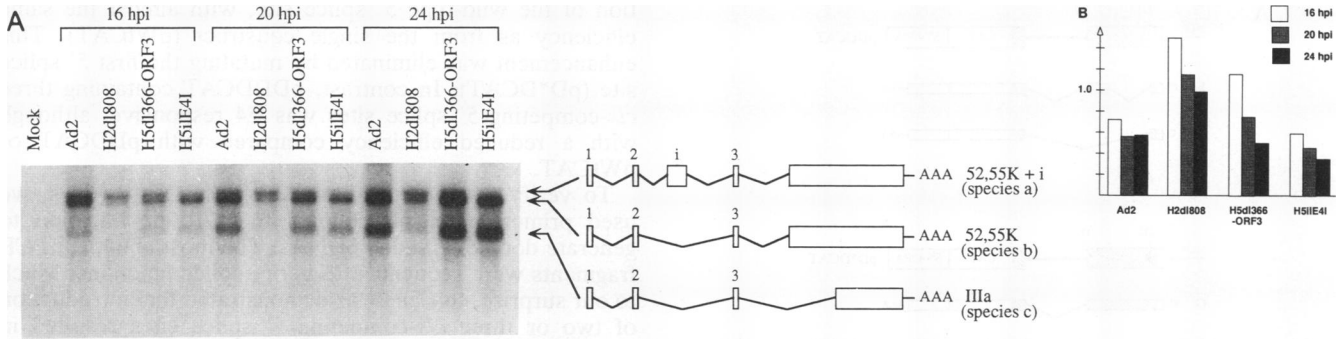


FIG. 4. E4-ORF6 promotes exon skipping during a virus infection. (A) RNA blot analysis of L1 mRNAs expressed in wild-type adenovirus type 2 (Ad2)-, H2*dl*808-, H5*dl*366-ORF3-, or H5ilE4I-infected cells. Note that 20-fold more RNA from H2*dl*808-infected cells and 3-fold more RNA from H5*dl*366-ORF3-infected cells was loaded compared with the other virus infections. The spliced structures of the major L1 mRNAs are shown schematically to the right. (B) Quantitation of 52,55K + i (species a) and 52,55K (species b) mRNA expression in virus infected cells (panel A). The relative efficiency of i-leader exon skipping is presented as the ratio between the level of 52,55K + i and 52,55K mRNA expression. Data represent the mean values from two independent infections. hpi, hours postinfection.

efficient i-leader exon skipping as does E4-ORF6. However, at 24 h postinfection, the i-leader exon was skipped with the same efficiency as seen during a wild-type infection (Fig. 4B). Thus, a prolonged late phase appears to overcome the requirement for E4-ORF6 for efficient i-leader skipping. Obviously, the virus has a backup system to produce alternatively spliced L1 mRNAs. This finding was not unexpected since the H5*dl*366-ORF3 virus grows with almost wild-type virus efficiency in HeLa cells (12) and thus must be able to induce alternative splicing of viral transcripts.

E4-ORF3 and E4-ORF6 have different effects on i-leader exon skipping in a transient transfection assay. The interpretation of results obtained by using virus infection is complex since multiple regulatory proteins are expressed. Thus, it is not obvious whether the E4-ORF6 protein works directly to facilitate i-leader exon skipping or whether it provides a function at an earlier step in a cascade of events. Therefore, we decided to use a transient transfection system to assay more directly for activities mediated by E4-ORF3 and E4-ORF6 on i-leader exon skipping.

For this experiment, 293 cells were cotransfected with cytomegalovirus E4-ORF-expressing plasmids (25) and the reporter plasmid preTripCAT, containing the wild-type tripartite leader region (Fig. 5). To determine the structure of 5′ leaders, we used primer extension coupled to PCR amplification to generate double-stranded cDNAs. The identity of the major PCR products was verified by restriction endonuclease cleavage (data not shown). As shown in Fig. 5, pCMVE4 (encoding the entire E4 region) cotransfection resulted in a stimulation of accumulation of both the early (1,2,i,3) and late (1,2,3) types of tripartite leader mRNA. In contrast, the E4-ORF3 and E4-ORF6 products were shown to have different effects on accumulation of alternatively spliced tripartite leaders. Whereas pCMVE4-ORF3 cotransfection resulted in a specific i-leader exon inclusion (1,2,i,3), pCMVE4-ORF6 cotransfection predominantly stimulated accumulation of the late specific (1,2,3) tripartite leader combination. Collectively these results suggest that E4-ORF3 has an activity as an exon inclusion factor whereas E4-ORF6 predominantly stimulates exon skipping. This result agrees with the observation that E4-ORF6 expression stimulates i-leader skipping during a virus infection (Fig. 4).

E4 enhances exon skipping in a nonviral pre-mRNA. To test whether the effects of E4-ORF3 and E4-ORF6 on accumulation of alternatively spliced mRNAs were general properties of the proteins or restricted to the major late tripartite region, we constructed three reporter plasmids containing multiples of a 300-bp fragment encompassing the rabbit β-globin 5′ splice site: pDDCAT, which contains duplicated wild-type 5′ splice sites; pD*DCAT, which contains a mutated 5′ splice site followed by the wild-type 5′ splice site; and pDDDCAT, which contains three wild-type 5′ splice sites connected together (Fig. 6A). This experimental strategy was chosen to determine the effect of E4 on *cis*-competing 5′ splice sites which all have the same strength. Previous results have shown that in a context in which 5′ splice sites of comparable strength compete, the proximal site is preferred (27). From our findings that an intron must be located close to the 5′ end of the pre-mRNA to be E4 responsive (24) and that the wild-type β-globin 5′ splice site was of critical importance for E4 responsiveness (Fig. 2), one might predict that E4 cotransfection would result in alternative 5′ splice site usage.

As shown in Fig. 6B, E4 cotransfection stimulated mRNA accumulation from pDDCAT, containing a tandem duplica-

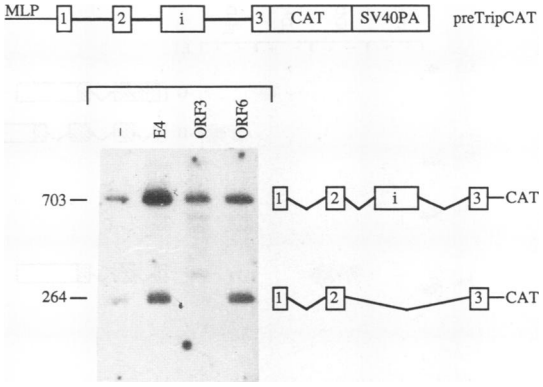
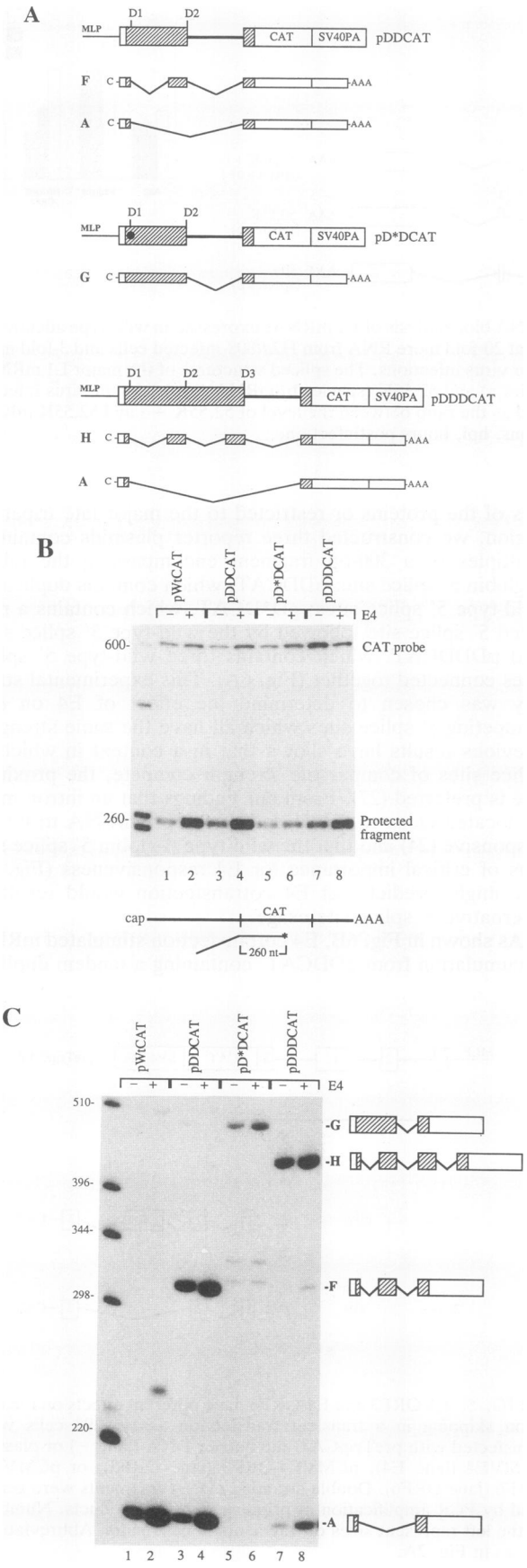


FIG. 5. E4-ORF3 and E4-ORF6 have different effects on i-leader exon skipping in a transient transfection assay. 293 cells were transfected with preTripCAT and carrier DNA (lane −) or plasmid pCMVE4 (lane E4), pCMVE4-ORF3 (lane ORF3), or pCMVE4-ORF6 (lane ORF6). Double-stranded cDNA fragments were generated by PCR amplification of primer extension products. Numbers to the left represent sizes of fragments in base pairs. Abbreviations are as in Fig. 2A.



tion of the wild-type 5′ splice site, with almost the same efficiency as from the single construct (pWtCAT). This enhancement was eliminated by mutating the first 5′ splice site (pD*DCAT). In contrast, pDDDCAT containing three *cis*-competing 5′ splice sites was E4 responsive, although with a reduced efficiency compared with pDDCAT or pWtCAT.

To verify the structures of various mRNA species, we used primer extension coupled to PCR amplification to generate double-stranded cDNAs. The most abundant PCR fragments were sequenced to verify splice junctions. Much to our surprise, this analysis demonstrated that introduction of two or three *cis*-competing 5′ splice sites resulted in splicing of transcripts encoding novel exons (species F and H; Fig. 6A and C). This occurred because a weak 3′ splice site, located 120 nucleotides upstream of the second and third wild-type 5′ splice sites, became activated. Other studies have also shown that weak 3′ splice sites sometimes become activated by introduction of functional 5′ splice sites a short distance downstream (see Discussion for further details). Interestingly, this multiexonic RNA was the most predominant RNA species accumulating in both pDDCAT- and pDDDCAT-transfected cells (Fig. 6C). However, additional RNAs in which the first 5′ splice site was joined to the wild-type 3′ splice site were also detectable (species A; Fig. 6A and C). Most importantly, E4 cotransfection stimulated accumulation of species A sixfold and accumulation of species F and H only twofold. This result is therefore consistent with the hypothesis that the E4 region, besides stimulating accumulation of constitutively spliced RNAs, also enhances exon skipping in a nonviral pre-mRNA.

In agreement with our previous finding that an intron must be located close to the pre-mRNA 5′ end to be E4 responsive (24), mRNA levels from plasmid pD*DCAT were not enhanced by E4 cotransfection (Fig. 6B). This plasmid produced a mRNA in which the functional 5′ splice site was located 350 nucleotides downstream of the cap nucleotide.

E4-ORF3 and E4-ORF6 have different effects on accumulation of alternatively spliced chimeric β-globin RNAs. To determine whether the E4-ORF3 and E4-ORF6 proteins also could promote exon inclusion and exon skipping on a nonviral pre-mRNA, we cotransfected pCMVE4-ORF3 and pCMVE4-ORF6 together with the reporter plasmid pDDCAT. The structures of the resulting mRNA species were determined by primer extension coupled to PCR amplification. Interestingly, the two E4 proteins were shown to have different effects on accumulation of alternatively spliced mRNAs from this transcript as well. As shown in

FIG. 6. The E4 region enhances exon skipping in nonviral pre-mRNAs. (A) Schematic drawing of plasmids used in the experiment. Below each plasmid, the structure of individual mRNA species is depicted. Boldface letters to the left refer to cDNA fragments shown in panel C. The dot shows the location of the 5′ splice site mutation. Hatched boxes represent rabbit β-globin sequences. D, position of 5′ splice site. Abbreviations are as in Fig. 2A. (B) S1 nuclease analysis of CAT mRNA levels in 293 cells transfected with reporter plasmids and carrier DNA (lanes −) or pCMVE4 (lanes +). Numbers to the left represent sizes of fragments in nucleotides (nt). (C) Analysis of the structures and abundances of CAT mRNA species expressed from reporter plasmids. Double-stranded cDNA fragments were generated by primer extension followed by PCR amplification. Fractions of the same RNA samples used in panel B were used. Positions and structures of the different mRNA species are shown to the right. Sizes of pBR marker fragments are shown at the left in nucleotides.

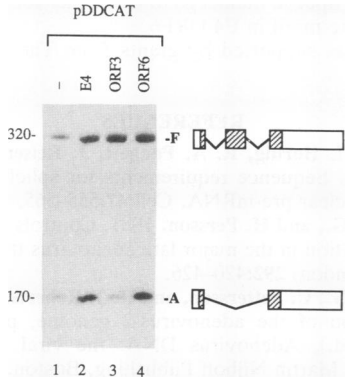


FIG. 7. E4-ORF3 and E4-ORF6 have different effects on accumulation of alternatively spliced chimeric β-globin RNAs. 293 cells were transfected with pDDCAT and carrier plasmid (lane −) or the plasmid pCMVE4 (lane E4), pCMVE4-ORF3 (lane ORF3), or pCMVE4-ORF6 (lane ORF6). Double-stranded cDNA fragments were generated by PCR amplification of primer extension products. Boldface letters to the right refer to mRNA species depicted in Fig. 6A. Numbers to the left represent size of fragments in base pairs.

Fig. 7, E4-ORF3 enhanced exon inclusion (species F, two- to threefold), whereas E4-ORF6 predominantly enhanced exon skipping (species A, sixfold, compared with species F, two- to threefold). Thus, we conclude that the opposite effects of E4-ORF3 and E4-ORF6 on accumulation of alternatively spliced RNA are not restricted to the major late tripartite leader region.

DISCUSSION

In this report, we have examined the effects of the adenovirus E4-ORF3 and E4-ORF6 proteins on the accumulation of alternatively spliced mRNAs from the viral major late transcription unit. We show that E4 expression is necessary for efficient tripartite leader assembly during virus growth (25) (Fig. 4). This activity of E4 was mediated by two proteins, E4-ORF3 and E4-ORF6, which had different effects on accumulation of alternatively spliced mRNA. In a transient transfection assay, the E4-ORF3 protein was shown to facilitate constitutive splicing, resulting in inclusion of the i-leader exon, whereas E4-ORF6 protein expression was shown to preferentially favor skipping of the i-leader exon. Interestingly, E4-ORF6 also stimulates i-leader exon skipping during lytic virus growth. Collectively, our data are compatible with the hypothesis that the E4 proteins are virus-encoded factors that regulate viral gene expression at the level of accumulation of alternatively spliced viral mRNAs.

The i-leader exon encodes an abundant 16-kDa protein that is expressed predominantly at intermediate and late times of infection (34). Since this protein is dispensable for lytic virus growth in tissue culture cells (31), it seems excessive that the virus would devote two proteins to only control its expression. Indeed, two of our experiments suggest that the E4 products may also be required for accumulation of alternatively spliced mRNAs from other viral transcription units. First, the activities of E4-ORF3 and E4-ORF6 are not limited to mRNAs expressed from the major late transcription unit. As shown here, E4-ORF3 and E4-ORF6 also stimulate accumulation of alternatively spliced mRNAs from a chimeric β-globin construct (Fig. 7). This finding suggests that the effects observed may be general properties of the two proteins. Second, E4 expression appears to be necessary for the switch from the early- to the late-specific profile of E1B mRNA expression during lytic virus growth (data not shown). The E1B mRNAs are alternatively spliced by using a common 3′ splice site and alternative 5′ splice sites (3).

Several lines of evidence support the hypothesis that the E4-ORF3 and E4-ORF6 proteins regulate major late mRNA accumulation at the level of RNA splicing, although alternative interpretations are still possible. First, E4 stimulates RNA accumulation posttranscriptionally (24, 29). Second, the E4 enhancement is detectable in the nuclear RNA fraction, suggesting that the E4 products do not promote mRNA accumulation by facilitating nuclear-to-cytoplasmic RNA transport (24; unpublished data). Third, E4 enhancement requires the presence of an intron with a wild-type 5′ splice site (Fig. 2 and 3) close to the 5′ end of the pre-mRNA (24) (Fig. 6). Fourth, E4-ORF3 and E4-ORF6 have different effects on accumulation of alternatively spliced tripartite leader containing mRNAs: E4-ORF3 facilitates i-leader exon inclusion, whereas E4-ORF6 predominantly favors i-leader exon skipping (Fig. 5). Collectively, these observations are consistent with the hypothesis that the E4-ORF3 and E4-ORF6 proteins are viral factors which, directly or indirectly, regulate major late pre-mRNA splicing.

We had not expected that creating transcription units with multiples of the wild-type rabbit β-globin 5′ splice site would result in the activation of a silent 3′ splice site present in the truncated β-globin intron (Fig. 6A and C). However, in retrospect, it is clear that similar results have previously been observed. For example, natural point mutations in the human β-globin second intron that creates novel 5′ splice sites result in the activation of otherwise silent upstream 3′ splice sites and thus production of aberrant four exonic hemoglobin mRNAs (5, 36). Furthermore, several studies have shown that increasing the length of the polypyrimidine tract at the 3′ splice site and/or enhancing the ability of U1 small nuclear RNA base pairing to the downstream 5′ splice site will increase the efficiency of exon inclusion (6, 16). Results of this kind have been taken as support for the so-called exon definition model (28). This model postulates that splicing factors binding to a 3′ splice site communicate across an exon with factors binding to a downstream 5′ splice site during spliceosome formation, thus resulting in a concerted recognition of both the 3′ and 5′ splice sites bordering an exon. Our results are consistent with this model; i.e., a strong 5′ splice site activates a weak upstream 3′ splice site and thereby results in inclusion of a novel exon.

The exon skipping activity of E4-ORF6 appears not to be universal. For example, inappropriate skipping of the second major late tripartite leader exon (leader exon 2; Fig. 1) does not occur under conditions in which the i-leader exon is efficiently skipped (Fig. 5). Most likely, natural constitutively spliced exons have evolved a combination of appropriate strength of splice signals and structural context to avoid improper exon skipping (21). With this hypothesis in mind, we can provide a reasonable explanation for the observed cases of E4-ORF6 exon skipping. Internal exons in eukaryotic genes are usually less than 350 nucleotides in length. This has been taken as supporting evidence that factors binding to 5′ and 3′ splice sites bordering an exon communicate across the exon (28). The i-leader exon is 440 nucleotides long and thus would be expected to exceed the maximal length for efficient exon definition. In contrast, the leader exon 2 is only 79 nucleotides long and, furthermore, is bordered by efficient 3′ and 5′ splice site sequences. Thus,

the difference in length could provide the explanation of why the i-leader exon is skipped under conditions in which we do not observe skipping of leader exon 2 (Fig. 5). Similarly, the 3′ splice site of the internal exons in the chimeric β-globin constructs (Fig. 6 and 7) is very weak and therefore probably makes this exon a target for the exon skipping activity of the E4-ORF6 protein. Thus, these experiments suggest a model whereby E4-ORF3 facilitates splice site communication across an exon whereas E4-ORF6 causes exon skipping by disrupting weak 5′ and 3′ splice site interactions.

An interesting analogy may exist between E4-ORF6 and E4-ORF3 and the cellular splicing activities heterogeneous nuclear ribonucleoprotein (hnRNP) A1 (22)/DSF (10) and ASF/SF2 (8, 15). Recently, ASF/SF2 has been shown to modulate alternative splicing by preventing exon skipping on complex pre-mRNAs (21). This activity of ASF/SF2 resembles the exon inclusion activity of E4-ORF3 that we describe here (Fig. 5 and 7). In vitro, the effect of ASF/SF2 is counteracted by hnRNP A1, which causes exon skipping in transcripts with improper exon length or in transcripts with a weak 3′ splice site preceding an internal exon (21). This is reminiscent of the activity of E4-ORF6, which causes skipping of unusually long internal exons and internal exons with weak 3′ splice sites (Fig. 5 and 7). Similar to the E4-ORF6 protein, hnRNP A1 does not appear to cause inappropriate exon skipping on natural constitutively spliced pre-mRNAs containing multiple exons (21). The balanced expression of ASF/SF2 and hnRNP A1 has been suggested to play a role in regulation of alternative splicing in eukaryotic cells. Similarly, the level of E4-ORF3 and E4-ORF6 expression during virus infection may have important consequences on alternative splicing in an adenovirus-infected cell. A direct sequence comparison of E4-ORF3 against ASF/SF2 and E4-ORF6 against hnRNP A1 shows only a limited path homology between the two protein pairs. Without genetic experiments pointing out the important features of the E4 proteins, it is not meaningful to draw conclusions concerning the possible structural relationship between these proteins.

A search through the EMBL data base has not identified any extensive homologies between E4-ORF3 or E4-ORF6 and other cloned genes except for the analogous proteins in other adenovirus serotypes. Several metazoan RNA splicing factors have been shown to have a set of common sequence motifs important for biological activity. These involve one, or several, RNA binding domains (including RRM motifs and Zn fingers [reviewed in references 13 and 20]) and an arginine/serine-rich sequence, which appears to be relevant for biological function (39). The E4-ORF3 protein appears to be devoid of both types of signals. The E4-ORF6 protein does not encode an extensive arginine/serine domain. However, we note that the C terminus of E4-ORF6, which is conserved between sequenced adenovirus E4-ORF6 proteins, is high in basic amino acids. Furthermore, the E4-ORF6 sequence shows a weak homology to the RRM motif (13) and encodes multiple cysteine and histidine residues that are highly conserved between E4-ORF6 proteins. These cysteines and histidines show homology to a zinc finger-like motif that is found in several other proteins involved in RNA processing (PRP6, PRP9, and PRP11 [19]). Clearly, the activity of the E4-ORF3 and E4-ORF6 proteins in RNA splicing should be subjected to a direct genetic dissection to pinpoint features important for biological activity.

ACKNOWLEDGMENTS

We are grateful to Angus Lamond for the generous gift of β-globin plasmids and to P. Hearing and E. Bridge for kindly providing E4 mutant viruses. Special thanks go to G. Ketner for first pointing out the Zn finger-like motif in E4 ORF6.

This work was supported by grants from The Swedish Cancer Society.

REFERENCES

1. **Aebi, M., H. Hornig, R. A. Padgett, J. Reiser, and C. Weissmann.** 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. Cell **47**:555–565.
2. **Akusjärvi, G., and H. Persson.** 1981. Controls of RNA splicing and termination in the major late adenovirus transcription unit. Nature (London) **292**:420–426.
3. **Akusjärvi, G., U. Pettersson, and R. J. Roberts.** 1986. Structure and function of the adenovirus-2 genome, p. 53–95. *In* W. Doerfler (ed.), Adenovirus DNA: the viral genome and its expression. Martin Nijhoff Publishing, Boston.
4. **Bridge, E., and G. Ketner.** 1989. Redundant control of adenovirus late gene expression by early region 4. J. Virol. **63**:631–638.
5. **Dobkin, C., and A. Bank.** 1985. Reversibility of IVS 2 missplicing in a mutant human β-globin gene. J. Biol. Chem. **260**:16332–16337.
6. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. **12**:2108–2114.
7. **Freyer, G. A., Y. Katoh, and R. J. Roberts.** 1984. Characterization of the major mRNAs from adenovirus 2 early region 4 by cDNA cloning and sequencing. Nucleic Acids Res. **12**:3503–3519.
8. **Ge, H., P. Zuo, and J. L. Manley.** 1991. Primary structure of human splicing factor ASF reveals similarities with Drosophila regulators. Cell **66**:373–382.
9. **Halbert, D. N., J. R. Cutt, and T. Shenk.** 1985. Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. J. Virol. **56**:250–257.
10. **Harper, J. E., and J. L. Manley.** 1991. A novel protein factor is required for use of distal alternative 5′ splice sites in vitro. Mol. Cell. Biol. **11**:5945–5953.
11. **Hemström, C., K. Nordqvist, U. Pettersson, and A. Virtanen.** 1988. Gene product of region E4 of adenovirus type 5 modulates accumulation of certain viral polypeptides. J. Virol. **62**:3258–3264.
12. **Huang, M.-M., and P. Hearing.** 1989. Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. J. Virol. **63**:2605–2615.
13. **Kenan, D. J., C. C. Query, and J. D. Keene.** 1991. RNA recognition: towards identifying determinants of specificity. Trends Biochem. Sci. **16**:214–220.
14. **Ketner, G., E. Bridge, A. Virtanen, C. Hemström, and U. Pettersson.** 1989. Complementation of adenovirus E4 mutants by transient expression of E4 cDNA and deletion plasmids. Nucleic Acids Res. **17**:3037–3048.
15. **Krainer, A. R., A. Mayeda, D. Kozak, and G. Binns.** 1991. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. Cell **66**:383–394.
16. **Kuo, H.-C., F.-U. H. Nasim, and P. J. Grabowski.** 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science **251**:1045–1050.
17. **Lamond, A., M. M. Konarska, and P. A. Sharp.** 1987. A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. Genes Dev. **1**:532–543.
18. **Larsson, S., C. Svensson, and G. Akusjärvi.** 1992. Control of adenovirus major late gene expression at multiple levels. J. Mol. Biol. **225**:287–298.
19. **Legrain, P., and A. Choulika.** 1990. The molecular characterization of PRP6 and PRP9 yeast genes reveals a new cysteine/histidine motif common to several splicing factors. EMBO J. **9**:2775–2781.
20. **Mattaj, I. W.** 1993. RNA recognition: a family matter? Cell **73**:837–840.

21. **Mayeda, A., D. M. Helfman, and A. R. Krainer.** 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol. Cell. Biol. **13**:2993–3001.
22. **Mayeda, A., and A. R. Krainer.** 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68**:365–375.
23. **Nevins, J. R., and M. C. Wilson.** 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. Nature (London) **290**:113–118.
24. **Nordqvist, K., and G. Akusjärvi.** 1990. Adenovirus early region 4 stimulates mRNA accumulation via 5′ introns. Proc. Natl. Acad. Sci. USA **87**:9543–9547.
25. **Öhman, K., K. Nordqvist, and G. Akusjärvi.** 1993. Two adenovirus proteins with redundant activities in virus growth facilitates tripartite leader mRNA accumulation. Virology **194**:50–58.
26. **Philipson, L.** 1961. Adenovirus assay by the fluorescent cell counting procedure. Virology **15**:263–268.
27. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46**:681–690.
28. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10**:84–94.
29. **Sandler, A. B., and G. Ketner.** 1989. Adenovirus early region 4 is essential for normal stability of late nuclear RNAs. J. Virol. **63**:624–630.
30. **Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23**:527–577.
31. **Soloway, P. D., and T. Shenk.** 1990. The adenovirus type 5 i-leader open reading frame functions in *cis* to reduce the half-life of L1 mRNAs. J. Virol. **64**:551–558.
32. **Svensson, C., and G. Akusjärvi.** 1984. Adenovirus VA RNAI: a positive regulator of mRNA translation. Mol. Cell. Biol. **4**:736–742.
33. **Svensson, C., U. Pettersson, and G. Akusjärvi.** 1983. Splicing of adenovirus 2 early region 1A mRNAs is non-sequential. J. Mol. Biol. **165**:475–495.
34. **Symington, J. S., L. A. Lucher, K. H. Brackmann, A. Virtanen, U. Pettersson, and M. Green.** 1986. Biosynthesis of adenovirus type 2 i-leader protein. J. Virol. **57**:848–856.
35. **Tigges, M. A., and H. J. Raskas.** 1984. Splice junctions in adenovirus 2 early region 4 mRNAs: multiple splice sites produce 18 to 24 RNAs. J. Virol. **50**:106–117.
36. **Treisman, R., S. H. Orkin, and T. Maniatis.** 1983. Specific transcription and RNA splicing defects in five cloned β-thalassemia genes. Nature (London) **302**:591–596.
37. **Virtanen, A., P. Gilardi, A. Näslund, J. M. LeMoullec, U. Pettersson, and M. Perricaudet.** 1984. mRNAs from human adenovirus 2 early region 4. J. Virol. **51**:822–831.
38. **Weinberg, D. H., and G. Ketner.** 1986. Adenoviral early region 4 is required for efficient viral DNA replication and for late gene expression. J. Virol. **57**:833–838.
39. **Zamore, P. D., J. G. Patton, and M. R. Green.** 1992. Cloning and domain structure of the mammalian splicing factor U2AF. Nature (London) **355**:609–614.