JOURNAL OF VIROLOGY, June 1979, p. 899-912
0022-538X/79/06-0899/14$02.00/0

Vol. 30, No. 3

Comparison of Late mRNA Splicing Among Class B and Class C Adenoviruses

BILL A. KILPATRICK,* RICHARD E. GELINAS, THOMAS R. BROKER, AND LOUISE T. CHOW
Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Received for publication 16 January 1979

Adenovirus class B (Ad3 and Ad7) and class C (Ad1, Ad5, and Ad6) late *r*-strand mRNA's were found to have segmented 5′ leaders. These leaders were very similar among serotypes within a class but differed in sequence from the leaders on late mRNA's of a different class. However, the leader components of class B viruses mapped at essentially the same map coordinates as those of class C viruses. The 5′ coordinates of the main bodies of class B messages to which the tripartite leaders are attached as well as the map positions of several of their early mRNA's were very similar to those of Ad2 transcripts. Infrequent examples of late *r*-strand polysomal RNAs of Ad3 and Ad7 had, in addition to the three common leader segments, a fourth leader segment derived from RNA encoded at various sites between the second and third leaders. The extra components formed several distinct groups. These molecules are presumably intermediates in the splicing processes that generate mature messages.

Adenovirus serotype 2 (Ad2) (class C) mRNA's have been mapped by a variety of techniques, including hybridization to separated strands of DNA restriction fragments (12, 29), fractionation by gel electrophoresis followed by hybridization to restriction fragments (5, 9, 11, 13, 23, 26), and electron microscopic RNA loop analysis (8, 24, 25). Surprisingly, its mRNA's have a very complex sequence organization. Most early (2- to 8-h) (3, 18; L. T. Chow, T. R. Broker, and J. B. Lewis, submitted for publication) and late (18- to 38-h) (2, 4, 6, 7, 19) cytoplasmic RNAs consist of sequences derived from two to five or more separate locations on the chromosome which become spliced together by the deletion of intervening sequences.

The majority of late RNA of Ad2 is transcribed from the *r*-strand from a promoter located near map coordinate 16.4 (41) and is synthesized as a continuous polymer of up to 26,000 nucleotides to a position near the right end of the chromosome (37). Each of about a dozen different cytoplasmic mRNA's (6, 7, 23, 26) is then processed from the nuclear precursor transcript (16). Only one message can be generated from a precursor because each of the mRNA's has the same three-part "leader" sequence derived from Ad2 coordinates 16.6, 19.6, and 26.6 spliced to its 5′ terminus (6, 7). The total length of the leader is about 200 bases (2, 7, 19, 40).

One means of deciphering the biological role and the mechanisms of splicing of the leaders present on most late Ad2 mRNA's is to find adenoviruses that exhibit variations. We therefore decided to compare RNAs from different serotypes of adenoviruses to search for similarities and differences in leader sequences and splicing patterns. Adenovirus serotypes 3 and 7 (class B) were selected for such studies because they differ substantially in DNA sequences from class C adenoviruses (14, 17) but their early and late RNA transcription patterns are organized much like those of Ad2 (36). In this report we describe our experiments that map the leader sequences for the *r*-strand late mRNA's of class B serotype adenoviruses as well as experiments that compare the sequence relationships among the late RNA leaders of class B and class C serotypes.

MATERIALS AND METHODS

Cells and viruses. Class C adenovirus serotypes (Ad1, Ad2, Ad5, and Ad6) were grown in HeLa cell suspension cultures as described by Pettersson and Sambrook (28) at an input multiplicity of about 20 PFU/cell for viral DNA preparation and at 100 PFU for late viral mRNA production. Class B adenovirus serotypes (Ad3 and Ad7) were grown in HeLa cell monolayer cultures with Eagle minimal media containing 2% A-γ calf serum (Flow Labs, Inc.) at similar multiplicities of infection.

Preparation and characterization of viral DNA. DNA was isolated from purified virions of all adenovirus serotypes as described by Pettersson and Sambrook (28). DNA stock preparations were 100 μg/

ml in 10 mM Tris-hydrochloride–1 mM EDTA, pH 7.2. The serotype identity of DNA preparations was verified by digestion with *Hpa*I, *Bam*HI, *Xho*I, and *Sal*I restriction enzymes (36; M. Zabeau, cited in reference 4) followed by electrophoresis in 1% agarose gels (34). DNA preparations were also examined by electron microscopy for sequence homogeneity and for intactness of single strands by denaturing and self-annealing at room temperature. Any preparation in which fewer than 80% of the DNA strands were intact was not used. Very few heteroduplexes containing deletions or substitutions were observed in the DNA preparations used; some molecules with long, instead of the normally short, inverted terminal duplications were found. Ad2:3, -2:7, and -3:7 heteroduplexes were also prepared, and the general patterns of homology were similar to those reported between class C and class B serotypes (14).

The *r*-strand of Ad2 *Bam*HI B (0 to 29.1) restriction fragment was prepared as described previously (7).

Ad3 and Ad7 DNAs were digested with *Bam*HI restriction endonuclease in 6 mM Tris-hydrochloride (pH 7.9), 6 mM $MgCl_2$, and 6 mM β-mercaptoethanol at 37°C for 3 h. Digestion was judged to be 100% complete, as analyzed by gel electrophoresis and ethidium bromide staining. The total digests were extracted successively with phenol and chloroform–isoamyl alcohol (24:1), precipitated with ethanol, and then redissolved in 0.01 M Tris-hydrochloride (pH 7.2)–0.01 M EDTA.

Preparation of late mRNAs. Class C viral mRNA was extracted from cells 24 h after infection, and that of class B was extracted 14, 26, and 48 h after infection. At 48 h after infection, each of the class B serotypes had induced approximately equivalent late stages of cytopathic change in the host cells as had the class C serotypes at 24 h. Late viral mRNA was isolated and purified from polysomes by a modification of the magnesium precipitation method (27) as described previously (15). Some data were derived from polysomal RNA prepared without detergent disruption of infected cells (1, 32). RNA stock preparations were 6 mg/ml in 10 mM Tris-hydrochloride, pH 7.2, and contained an estimated 5 to 10% adenovirus-specific mRNA and about 95 to 90% rRNA.

To confirm that a variety of viral mRNA species was contained in each RNA preparation, R loops were formed with the homologous DNA of each serotype and examined by electron microscopy. The R-loop patterns (data not shown) were similar to those of Ad2 reported by Chow et al. (8) and Meyer et al. (24).

R loops were formed as described by Thomas et al. (35), White and Hogness (38), and Chow et al. (8). Adenoviral DNA and adenoviral RNA samples were mixed at final concentrations of 5 to 10 μg/ml and at 50 to 100 μg/ml, respectively, in 0.1 M HEPES (*N*-2-hydroxyethyl piperazine-*N'*-2-ethanesulfonic acid, pH 7.8), 0.4 M NaCl, 0.01 M EDTA, and 70% formamide in a total volume of 10 μl. The mixtures were incubated at 52°C for 16 to 20 h, and portions were removed, diluted into hyperphase solution, and immediately prepared for electron microscopy.

Hybridization of Ad3 and -7 and Ad-1, -2, -5, and -6 mRNA's with Ad2 *Bam*HI-B fragment *r*-strand. Class B (Ad3, Ad7) or class C (Ad1, Ad2, Ad5, Ad6) mRNA was mixed with purified *r*-strands of Ad2 *Bam*HI B restriction fragment (map coordinates 0 to 29.1) under conditions that readily allowed hybridization of this DNA strand with Ad2 mRNA leader components, as described by Chow et al. (7). Specifically, Ad2 *Bam*HI-B DNA fragment was dissolved at 4 to 5 μg/ml, and the respective mRNA was dissolved at 40 to 50 μg/ml in 0.1 M HEPES (pH 7.8), 0.4 M NaCl, 0.01 M EDTA, and 70% formamide. The mixture was placed at 52°C and gradually cooled to 35°C over a 4-h period, or to 0°C over a 6-h period. Portions from the reaction were prepared for electron microscopy by spreading from 40 to 45% formamide at 22°C.

Hybridization of Ad3 or Ad7 mRNA's with Ad3 and Ad7 single-stranded DNA. Annealing of Ad3 or Ad7 mRNA's and homologous DNA was performed with intact DNA or DNA digested with restriction enzymes. Intact DNA (0.05 μg in 2.5 μl of water) was fully denatured by adding 1 μl of 1 N NaOH at room temperature for 10 min. Denaturation was stopped by chilling on ice and the addition of 3 μl of 2.0 M Tris-hydrochloride (pH 7.0), 2.5 μl of 0.2 M Na-EDTA (pH 8.5), and 1.9 μl of distilled water. The solution was adjusted to 70% formamide and 0.3 M Na^+ by adding 35 μl of redistilled formamide and 3.8 μl of 4.0 M NaCl. After adding 1.8 μg of RNA stock in 0.3 μl, the mixture had a final volume of 50 μl. In reactions involving mixtures of Ad3 and Ad7 RNA, one-half of this amount of RNA was contributed by each serotype. Hybridization was carried out in sealed tubes at 52°C for 30 to 90 min.

Alternative conditions were used to effect more efficient DNA:RNA hybridizations when starting with double-stranded DNA restriction fragments, for which the rate of the competing DNA:DNA reannealing is more rapid than for intact DNA. DNA was dissolved at 4 μg/ml and RNA was dissolved at 200 μg/ml in 0.1 M HEPES (pH 7.8), 0.4 M NaCl, 0.01 M EDTA, and 80% formamide in a total volume of 10 μl. The mixtures were heated in sealed tubes at 67°C for 10 min to denature the DNA, and then immediately transferred to a 58°C water bath to allow formation of DNA:RNA heteroduplexes. At least 20% of the DNA single strands formed RNA:DNA heteroduplexes within 30 to 60 min of incubation. All hybridizations were terminated by 5- or 10-fold dilutions with ice-chilled water. RNA:DNA heteroduplexes were stable for at least 4 days after dilution and storage on ice or at 4°C.

Electron microscopy. Portions of the hybridization mixtures were diluted into 0.1 M Tris-hydrochloride (pH 8.5), 0.01 M EDTA, 50 μg of cytochrome *c* per ml, and 40 or 45% formamide, and these hyperphase solutions were spread over hypophases of 0.01 M Tris-hydrochloride (pH 8.5), 0.001 M EDTA, and 10 or 15% formamide (10). Length standards of single-stranded ϕX174 DNA (5,386 bases) and double-stranded colicin E1 DNA (6,300 base pairs) were included in each sample. Cytochrome films were picked up on Parlodion-coated 200-mesh grids, stained with 10^{-4} M uranyl acetate in 90% ethanol, rotary shadowed with evaporated Pt-Pd (80:20), and examined with a Philips 201C electron microscope.

Data analysis. Micrographs were taken at 10,000× and were projected with an enlarger to yield a final

magnification of about 80,000×. Contour lengths were measured with a Numonics Corp. electronic planimeter. Average values of several double-stranded and single-stranded length standards were determined for each micrograph, and these values were used to normalize double-stranded and single-stranded regions of the DNA:RNA heteroduplexes. Map coordinates are expressed as percentages of the adenoviral DNA (35,000 base pairs) (17). When hybridized to single-stranded DNA, spliced mRNA molecules constrain the DNA into one or more deletion loops intervening between the conserved RNA segments, and the splice coordinates are readily measurable.

RESULTS

To map the leader sequences of mRNA's from different serotypes of adenoviruses, we annealed mRNA to single-stranded DNA. If the message were spliced, the main body would pair with its coding sequence and the leader sequences or other spliced RNA components would hybridize back to each region of DNA from which they were derived, thereby drawing the DNA strand into a series of loops which would correspond to the sequences deleted during splicing (2, 6, 18).

Hybridizations of the *r*-strand of Ad2 *Bam*HI-B restriction fragment with late mRNA's of class C and class B adenovirus serotypes. The *r*-strand of the Ad2 *Bam*HI-B restriction fragment spans coordinates 0 to 29.1 and encodes the three leader segments common to the late Ad2 mRNA's transcribed downstream from the *r*-strand. The fragment will hybridize with the 5′ ends of these mRNA's displayed in R loops, or with the 5′ ends of otherwise unhybridized mRNA's, at three separate sites and be constrained into deletion loops representing the intervening sequences (7).

The Ad2 *Bam*HI-B *r*-strand was used as a probe to test for the presence of leader sequences homologous to those of Ad2 in polysomal late mRNA synthesized by various adenoviruses of class C and B serotypes. Characteristic double-loop structures indicative of a three-part leader were formed by each of the class C serotypes tested: Ad1, Ad2, Ad5, and Ad6 (Fig. 1A and B). The leader coordinates and lengths found for the intervening DNA segments of Ad1, Ad5, and Ad6 are summarized in Table 1. These values are almost identical to measurements of heteroduplexes prepared with Ad2 RNA and DNA (2, 6, 7) and indicate that all class C adenoviruses encode tripartite leaders for the late mRNA's with the same or very similar nucleotide sequences.

In contrast, when Ad2 *Bam*HI *r*-strand DNA and mRNA from the class B (Ad3 and Ad7) serotypes were mixed, no such double-loop structures were formed under the same conditions, whether the mRNA was free in solution or present in R loops. To maximize the opportunity for hybridization of sequences with partial homology, the hybridization temperature was dropped slowly over a greater temperature range, from 60 to 0°C. Still, no double-loop structures or any other heteroduplexes were found. Furthermore, no heteroduplex structures with deletion loops were found when intact Ad2, Ad3, or Ad7 DNA was mixed with adenovirus mRNA's derived from serotypes outside its respective class.

Splicing of adenovirus class B late mRNA. To assess whether class B late mRNA was simply devoid of leaders or contained leaders with nucleotide sequences different from those of Ad2 mRNA, Ad3 and Ad7 mRNA's were hybridized with intact single-stranded DNA from the homologous serotype. In this experiment the class B serotypes formed DNA:RNA heteroduplex structures with characteristic triple loops. The first two loops were reproducible in sizes and positions and were identical to those formed between the DNA and RNA of the class C serotypes (2, 6, 7; this study). The third DNA loop was variable in length and corresponded to the intervals between the third common leader segment and various gene transcripts (Fig. 2). As shown in Table 1, the first DNA loop of Ad7 began at coordinates 16.5 ± 0.4 (24 measurements) and had a length of 3.1 ± 0.3% (82 measurements). A duplex of about 0.2% corresponded to the second leader segment, with a midpoint of 19.8 ± 0.4 (82 measurements). The second loop had a length of 6.9 ± 0.7% (80 measurements). A duplex of about 0.3% corresponded to the third leader, with its midpoint at 26.8 ± 0.6 (74 measurements). As also shown in Table 1, essentially identical values were found for Ad3.

To confirm these coordinates, Ad3 late message transcripts were hybridized to the Ad3 *Bam*HI-A fragment (coordinates 9.6 to 36.7) (present in a total digest). Similarly, Ad7 mRNA was annealed to the Ad7 *Bam*HI-B fragment (coordinates 15.9 to 36.7). In both cases, double- and triple-loop structures were formed (Fig. 1C and D), depending on the mRNA hybridized (cf. Table 2). The first loop in the Ad3 fragment was located 6.7 ± 0.4% (28 measurements) from the coordinate 9.6 restriction site, confirming the first leader segment to be encoded near 16.3 in the genome. Similarly, two or three loops were observed with Ad7 mRNA and Ad7 *Bam*HI-B restriction fragment. The single-stranded DNA segment to the left of the first loop in most molecules was barely resolvable. Since the practical limit of resolution under the sample prep-

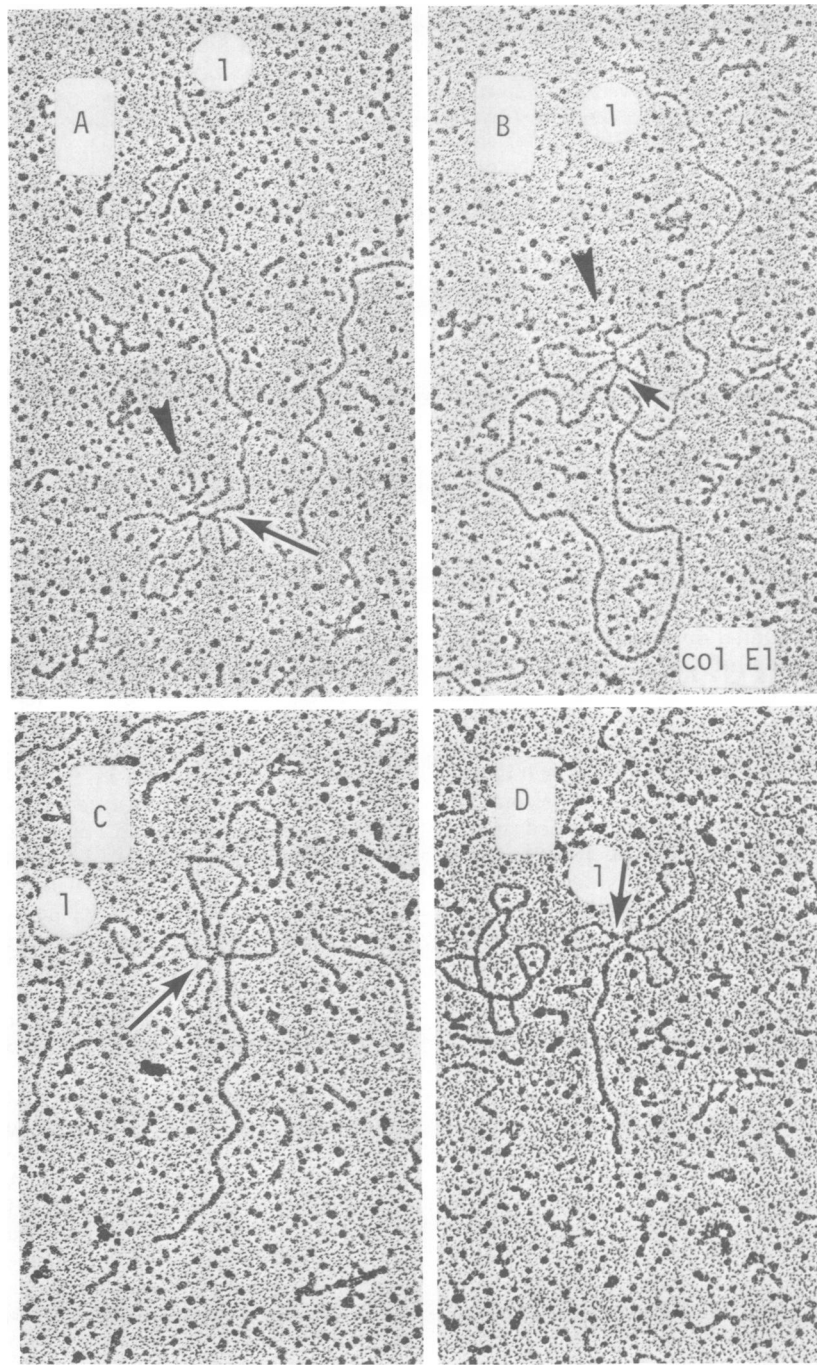

FIG. 1. *Heteroduplexes between r-strands of BamHI restriction fragments of adenovirus class B or C DNA and polysomal late mRNA. Ad2 BamHI B DNA (coordinates 0 to 29.1) is shown cross-hybridized to Ad1 mRNA (A) and to Ad5 mRNA (B). Ad3 BamHI A DNA (coordinates 9.6 to 36.7) is hybridized to Ad3 mRNA (C) and Ad7 BamHI B (coordinates 15.9 to 36.7) is hybridized to Ad7 mRNA (D). l indicates the left end of the restriction fragment. In (A) and (B) only the leader sequences have hybridized; arrowheads point to unhybridized message bodies and arrows point to the 5′ ends of the transcripts. In (C) and (D), both the leaders and the main bodies (5′ ends at 30.7) of the RNAs have hybridized to the fragments. Because the 3′ end of the mRNA extends beyond the right restriction site in (C), it has cohybridized to a neighboring DNA fragment.*

TABLE 1. *Genome coordinates of the leaders and lengths of intervening DNA of the class B and class C serotypes*

Hybridization mixtures	First leader coordinate	First loop length	Midpoint second leader coordinate[a]	Second loop length	Midpoint third leader coordinate[b]
Ad7 intact DNA vs Ad7 RNA	16.5 ± 0.4 (24)[c]	3.1 ± 0.3 (82)	19.8 ± 0.4 (82)	6.9 ± 0.7 (80)	26.8 ± 0.6 (74)
Ad3 intact DNA vs Ad3 RNA	16.6 ± 0.4 (21)	3.1 ± 0.4 (59)	19.9 ± 0.4 (60)	6.7 ± 0.2 (58)	26.9 ± 0.8 (60)
Ad7 intact DNA vs Ad3 RNA	16.6 ± 0.4 (8)	3.1 ± 0.3 (32)	19.8 ± 0.3 (32)	6.6 ± 0.6 (32)	26.7 ± 0.8 (32)
Ad7 *Bam*HI B vs Ad7 RNA	16.5[d]	3.2 ± 0.4 (27)	19.9 ± 0.4 (27)	7.1 ± 0.8 (21)	26.9 ± 0.6 (19)
Ad3 *Bam*HI A vs Ad3 RNA	16.4 ± 0.4[e] (28)	3.1 ± 0.4 (31)	19.6 ± 0.7 (28)	7.0 ± 0.7 (27)	27.0 ± 0.9 (27)
Ad2 *Bam*HI B vs Ad1 RNA	16.4 ± 0.8 (13)	2.8 ± 0.2 (14)	19.4 ± 0.2 (14)	6.8 ± 0.4 (14)	26.4 ± 0.4 (14)
Ad2 *Bam*HI B vs Ad5 RNA	16.6 ± 0.8 (8)	2.9 ± 0.2 (12)	19.6 ± 0.2 (12)	6.7 ± 0.6 (13)	26.5 ± 0.6 (13)
Ad2 *Bam*HI B vs Ad6 RNA	16.6 ± 0.6 (8)	2.8 ± 0.2 (6)	19.6 ± 0.2 (6)	6.9 ± 0.5 (9)	26.7 ± 0.5 (9)

[a] Average length of second leader was 0.2%.
[b] Average length of third leader was 0.3%.
[c] Numbers in parentheses are the number of measurements.
[d] Since the DNA segment between the 15.9 *Bam*HI restriction site and the first leader was too small to measure accurately, its length was estimated as described in the text as 0.6 to 0.7%.
[e] Evaluated by adding the measured distance of 6.7 ± 0.4% to the restriction site at coordinate 9.6.

aration conditions is about 200 bases (0.6% of the genome), this positioned the first leader close to 16.5%, in good agreement with data obtained with intact DNA strands. The lengths of the DNA deletion loops and the map coordinates of the other leader segments also agreed with measurements made on intact DNA strands, as summarized in Table 1.

Although the nucleotide sequences composing the tripartite leaders of Ad3 and Ad7 were nonhomologous under our hybridization conditions with the class C serotypes, they were very similar or identical to one another. This was established by cross-hybridizations of Ad3 mRNA with intact Ad7 DNA, and vice versa. In both reactions, typical heteroduplexes with three deletion loops were formed for most mRNA's (Fig. 3 and Tables 1 and 2). The map coordinates and lengths of intervening DNA loops were the same as those found for homologous RNA:DNA hybridization reactions of the class B serotypes, strongly suggesting that the Ad3 and Ad7 leaders were derived from the same chromosome positions. This was confirmed by hybridizing mixtures of Ad3 and Ad7 mRNA with either Ad3 or Ad7 DNA. Only the usual, well-layed-out, triple-looped heteroduplexes were found. This ruled out the possibility that the Ad3 and Ad7 leaders were different and were encoded at slightly different, nonoverlapping sites. Had this been the case, sequences remaining exposed in the intervening DNA segments of the RNA:DNA heteroduplexes would still be available for pairing with leaders from the other serotype RNA. Hybridization, then, of the second RNA would further constrain the heteroduplexes into more complex looped structures. No such RNA:DNA:RNA structures were found after mixed cross-hybridization reactions. These results confirm that Ad3 and Ad7 encode very similar, if not identical, tripartite leaders. They mapped at coordinates 16.5, 19.8, and 26.8, essentially identical to those of Ad2.

Mapping Ad3 and Ad7 spliced late mRNA's. In none of the heteroduplexes formed with intact DNA or restriction fragments was there evidence of any systematic difference in positions or lengths of leader components with respect to the particular mRNA hybridized. To map the 5′ ends of the main coding region for each mRNA, the lengths of the third deletion loop between the third leader segment and the main body of the transcript were evaluated. The 3′ end of all third leaders was standardized as 26.9 to minimize compounding measurement errors and to utilize data from heteroduplexes in which the left end of the DNA was broken. As shown in Table 2, essentially the same distributions of 5′ ends were found with both Ad3 and Ad7. The mean positions of 5′ ends were at 30.7, 33.9, 37.8, 39.1, 42.9, 45.4, 49.1, 51.3, 65.6, 68.3, 74.2, and 86.3. These coordinates are almost

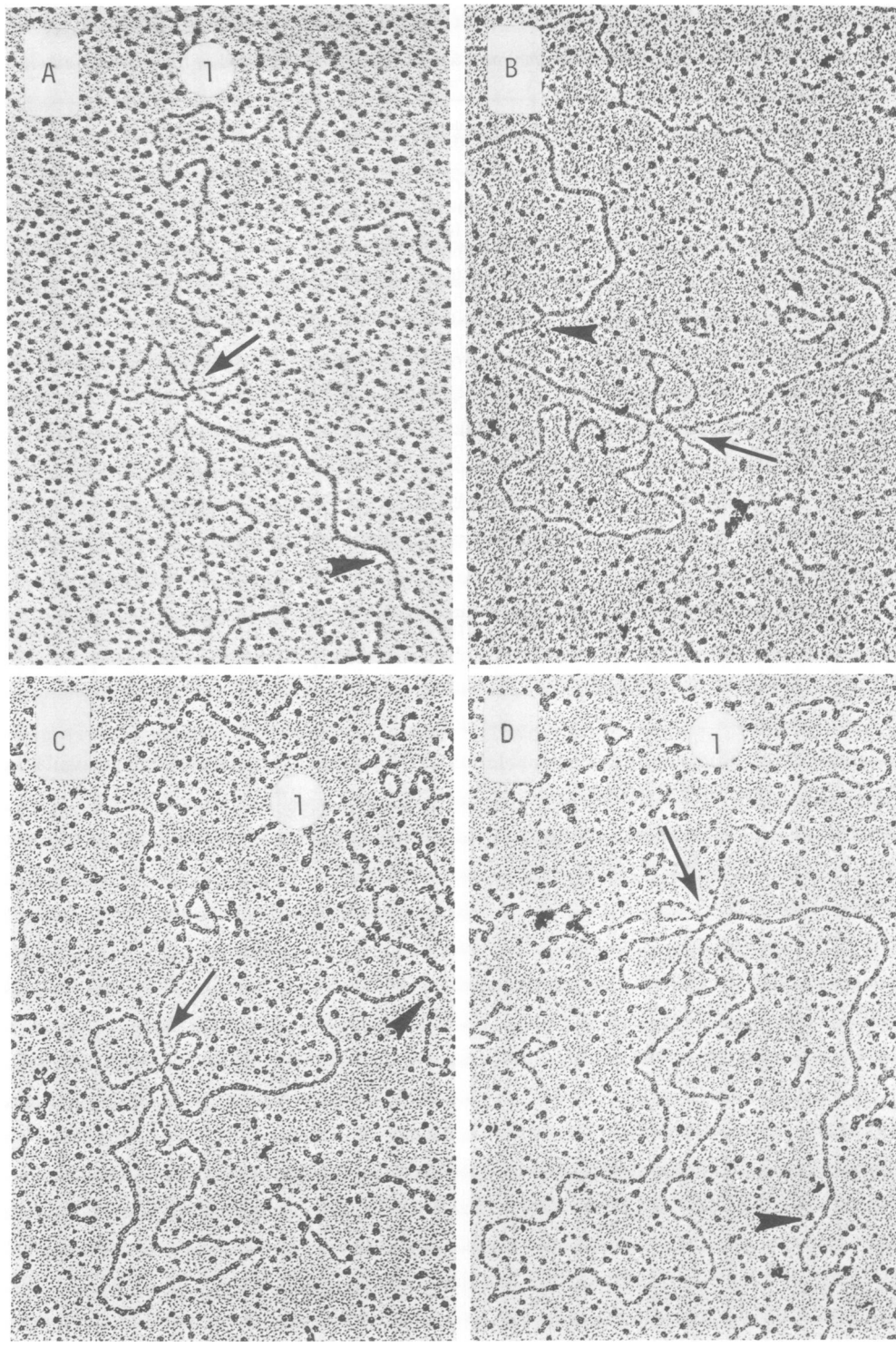

FIG. 2. *Heteroduplexes between intact r-strands of class B adenovirus DNAs and polysomal late mRNA of the homologous serotype. (A) and (B): Ad3; (C) and (D): Ad7. The 5′ and 3′ coordinates of the message bodies are (A) 39.2 to 45.1; (B) 44.8 to 49.6; (C) 51.4 to 61.2; and (D) 68.4 to 79.6. Molecule (B) is in a circular form due to renaturation of the terminal inverted duplications. In this and subsequent figures, l (and r) indicate the left (and right) orientation(s) of the DNA according to the conventional genetic map; arrows point to the 5′ ends of transcripts, and arrowheads point to the 3′ ends. In some cases, additional RNAs have annealed to DNA sequences in the third deletion loop.*

identical to those of Ad2 late mRNA's (6, 7), and the transcripts are tentatively identified in Table 2 according to the mRNA:protein assignments achieved with Ad2 (4, 22).

Mapping other spliced mRNA's of Ad3 and Ad7. The similarity of mRNA splicing among class C (Ad2) and class B (Ad3 and Ad7) serotypes was further evaluated by examining polysomal mRNA of Ad3 and Ad7 for composite mRNA species encoded on the *l*-strand and on the *r*-strand to the left of coordinate 16.0. The *l*-strand of Ad3 and Ad7 gives rise to mRNA's similar to the Ad2 IVa2 mRNA (6). The main body of this mRNA from Ad3 mapped from 14.4 ± 0.2 to 10.0 ± 0.1 (five measurements) (Table 3 and Fig. 4A). It has a single-spliced leader component derived from coordinate 15.4 ± 0.3 (five measurements). Only one example of the IVa2 message was found in Ad7 mRNA. Another message encoded by the *l*-strand is for the single-stranded DNA binding protein (DBP) (36). The DBP message from Ad2 is unusual in that the splicing pattern changes gradually during the transition from early to late transcription (Chow et al., submitted for publication). At early times the 5′ proximal leader is derived from coordinate 75.0 (3, 18; Chow et al., submitted for publication). At late times, two alternative 5′ leaders are derived from either coordinate 72 (major form) or 86 (minor species) (Chow et al., submitted for publication). When we examined the DBP-encoding region of Ad3 and Ad7 heteroduplexes with 26-h RNA, spliced RNA identical to that seen for early Ad2 transcripts was found. The DBP mRNA from Ad3 has the 5′ leader at 74.9 ± 0.4 (7 measurements), with a second segment derived from coordinates 68.5 ± 0.5 (11 measurements), and the main body of the message complementary to coordinates 66.3 ± 0.6 (11 measurements) through 61.6 ± 0.9 (9 measurements) (Fig. 4B and Table 3). Similar molecules have also been found for Ad7 (Table 3). Examples were infrequently observed with the 5′ leader from coordinate 72 or from about 87, whereas the other two segments were the same as in the early mRNA species from Ad2 and Ad3 or Ad7 (Fig. 4C).

As summarized in Table 3, Ad3 has a unique late *r*-strand transcript that was found in both unspliced and spliced versions from coordinates 4.3 ± 0.5 through 10.2 ± 0.5 or from 4.2 ± 0.5 through 5.7 ± 0.6 spliced to 9.0 ± 0.3 through 10.7 ± 0.1 (Fig. 3B, 4D, 5A). Similar mRNA's have been seen in Ad7 RNA preparations (data not shown). These species correspond closely to the Ad2 early region 1 B mRNA's (6, 8, 18).

Extra common leader segment. In the course of this work, infrequent examples (less

TABLE 3. *Map coordinates of Ad3 and Ad7 messages: l-strand and other r-strand messages*

Message	Coordinates[a]
Ad3 region 1B spliced	4.2 ± 0.5 (8)–5.7 ± 0.6 (8)/9.0 ± 0.3 (8)–10.7 ± 0.1 (8)
Ad3 region 1B unspliced	4.3 ± 0.5 (3)–10.2 ± 0.5 (3)
Ad3 "IVa2"	15.4 ± 0.3 (5)/14.4 ± 0.2 (5)–10.0 ± 0.1 (5)
Ad3 DBP	74.9 ± 0.4 (7)/68.5 ± 0.5 (11)/66.3 ± 0.6 (11)–61.6 ± 0.9 (9)
Ad7 DBP	75.0 ± 0.1 (3)/68.0 ± 0.3 (5)/65.8 ± 0.3 (5)–61.0 ± 0.8 (5)

[a] Numbers in parentheses are the number of measurements, dashes indicate continuous RNA, and diagonal lines indicate splice junctions.

TABLE 2. *Map coordinates of Ad3 and Ad7 messages: 5′ coordinates of r-strand messages spliced to the common tripartite leaders*

Ad3 5′ ends	Ad7 5′ ends	Ad2[a] 5′ ends	Tentative protein assignment[b]
		29.1 ± 0.4 (13)	
30.7 ± 0.4 (36)[c]	30.8 ± 0.5 (5)	30.5 ± 0.3 (42)	Peripentonal hexon-associated (IIIa)
33.9 ± 0.4 (12)	34.0 ± 0.5 (8)	33.9 ± 0.3 (15)	
37.8 ± 0.9 (9)		37.0 ± 0.6 (10)	Penton (III)?
39.8 ± 0.4 (4)	39.1 ± 0.8 (10)	38.8 ± 0.6 (51)	
42.9 ± 0.5 (15)	42.9 ± 0.8 (8)	42.8 ± 0.7 (41)	Precursor, major core (pVII)
45.4 ± 0.8 (13)	45.5 ± 1.0 (9)	45.1 ± 0.7 (25)	Minor core (V)
49.1 ± 0.9 (24)	49.1 ± 0.9 (12)	49.5 ± 0.6 (32)	Precursor, hexon associated (pVI)
51.3 ± 0.4 (10)	51.4 ± 0.8 (15)	51.2 ± 0.5 (24)	Hexon (II)
65.6 ± 0.9 (12)	65.0 ± 0.6 (5)	66.1 ± 0.7 (12)	Nonvirion 100K
68.3 ± 0.8 (7)	68.2 ± 1.0 (4)	68.0 ± 1.0 (16)	
74.2 ± 0.2 (2)		74.0 ± 0.7 (5)	Precursor, hexon associated (pVIII)
86.3 ± 2.7 (4)		86.2 ± 0.4 (52)	Fiber (IV)

[a] From reference 6.
[b] From references 8 and 22.
[c] Numbers in parentheses indicate the number of measurements.

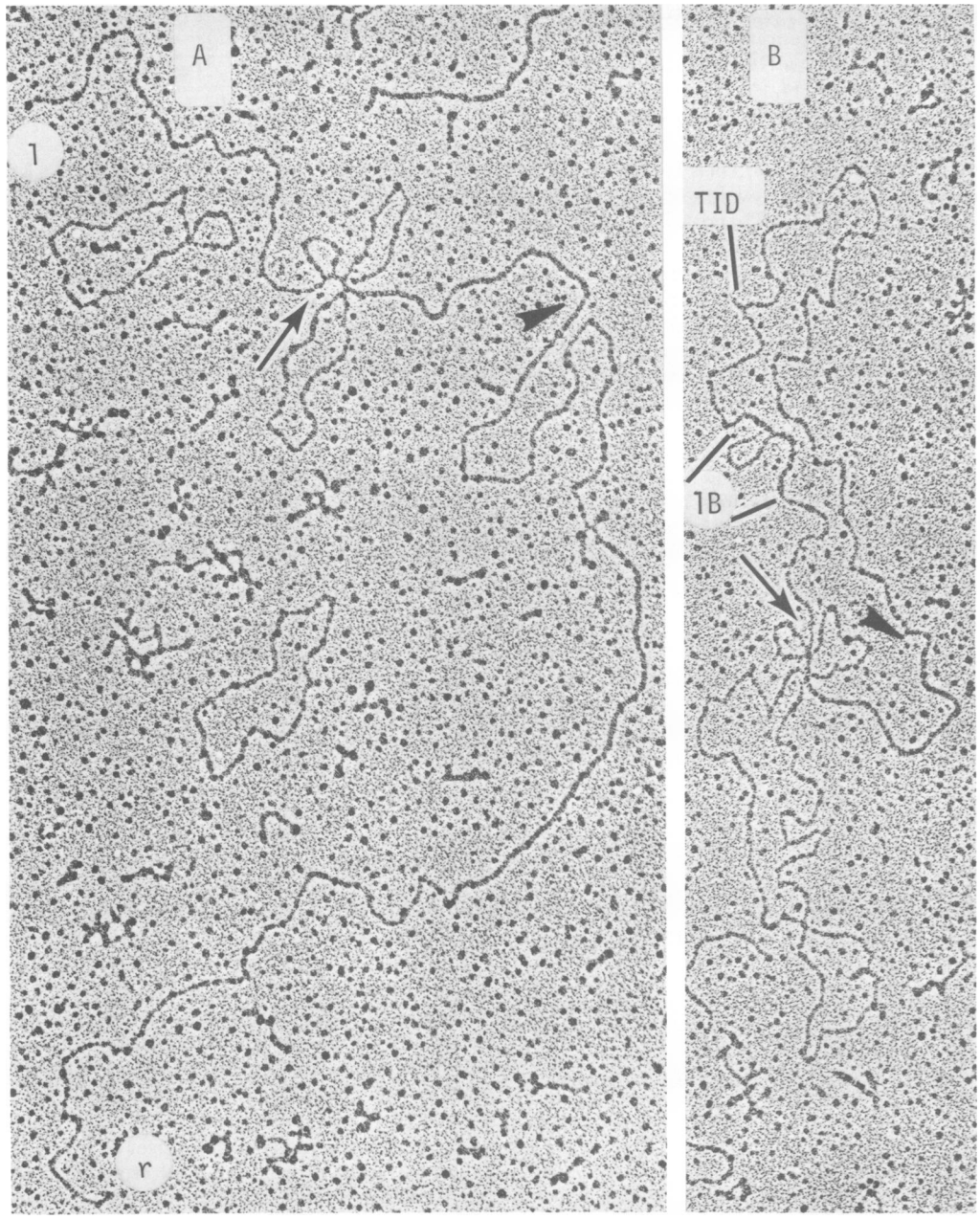

FIG. 3. *Heteroduplexes between intact r-strands of Ad7 DNA and Ad3 polysomal late mRNA. The 5′ and 3′ coordinates of the message body are (A) 38.5 to 45.5; (B) 65.8 to 72.5. The spliced form of region 1 B RNA has also annealed to molecule (B) and the terminal inverted duplications (TID) of the DNA have paired to form a panhandle and circularize the DNA.*

than 5%) of various *r*-strand messages containing either a longer third leader or four leader segments were discovered in Ad3 and Ad7 polysomal late mRNA's extracted at 26 and 48 h postinfection. These were revealed by heteroduplexes in which the second loop (intervening between the second and third leaders) had a contour length shorter than normal due to a longer third leader segment or by RNA:DNA heteroduplexes in which the DNA was con-

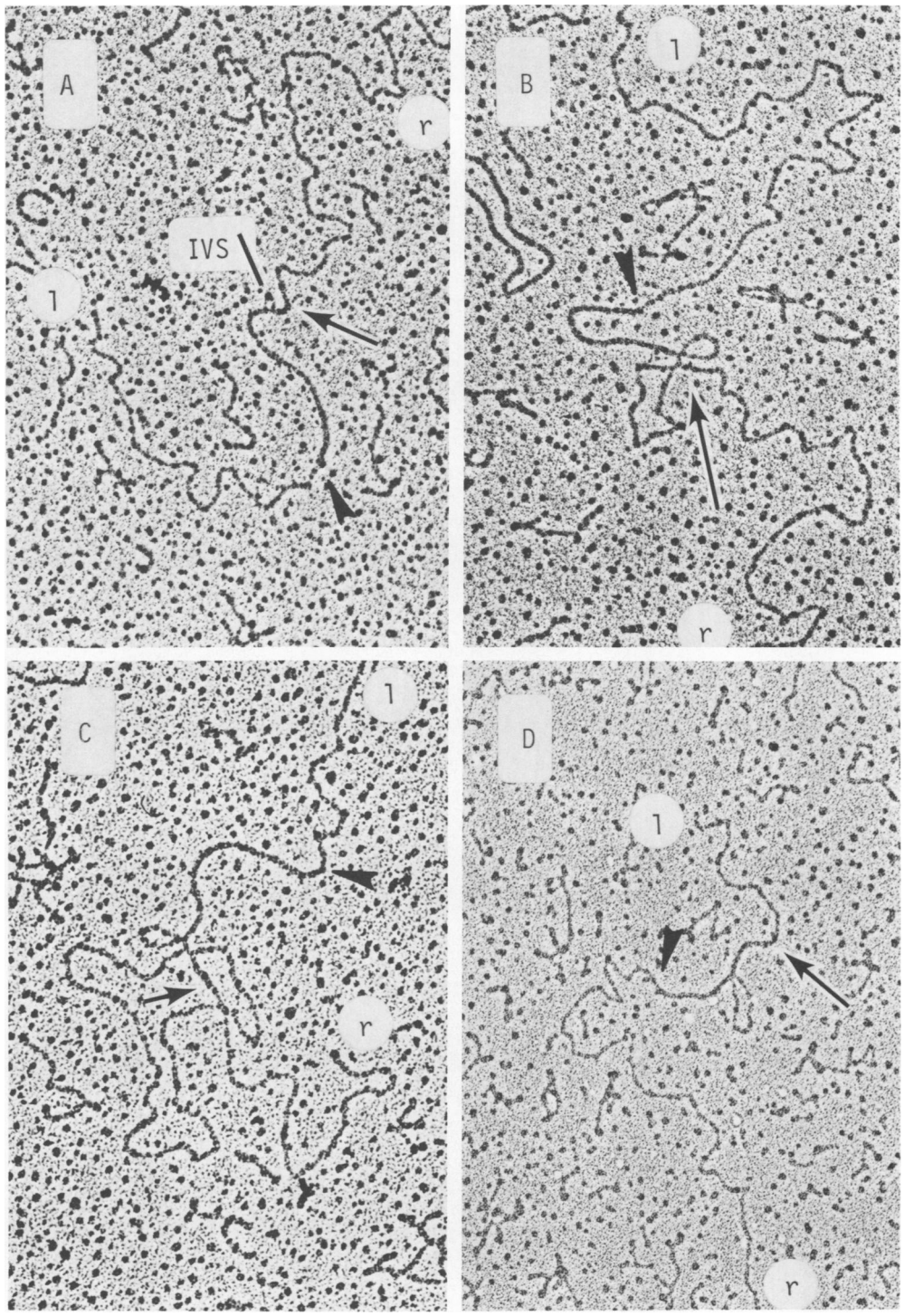

FIG. 4. *Heteroduplexes between intact single-stranded DNA of class B serotypes and polysomal mRNA from the homologous serotype. (A) Ad3 "IVa2" mRNA hybridized to Ad3 l-strand (10.1 to 14.5/15.3). IVS points to the short intervening sequence between the leader and the main body of the RNA; (B) Ad3 DBP mRNA (analogous to Ad2 early form) hybridized to Ad3 l-strand (62.3 to 66.7/69/75); (C) Ad7 DBP mRNA (analogous to Ad2 minor late form) hybridized to Ad7 l-strand (60.6 to 65.6/68/87.0 to 87.8); (D) Ad3 early region 1B spliced mRNA hybridized to Ad3 r-strand (4.6 to 6.1/9.2 to 10.9).*

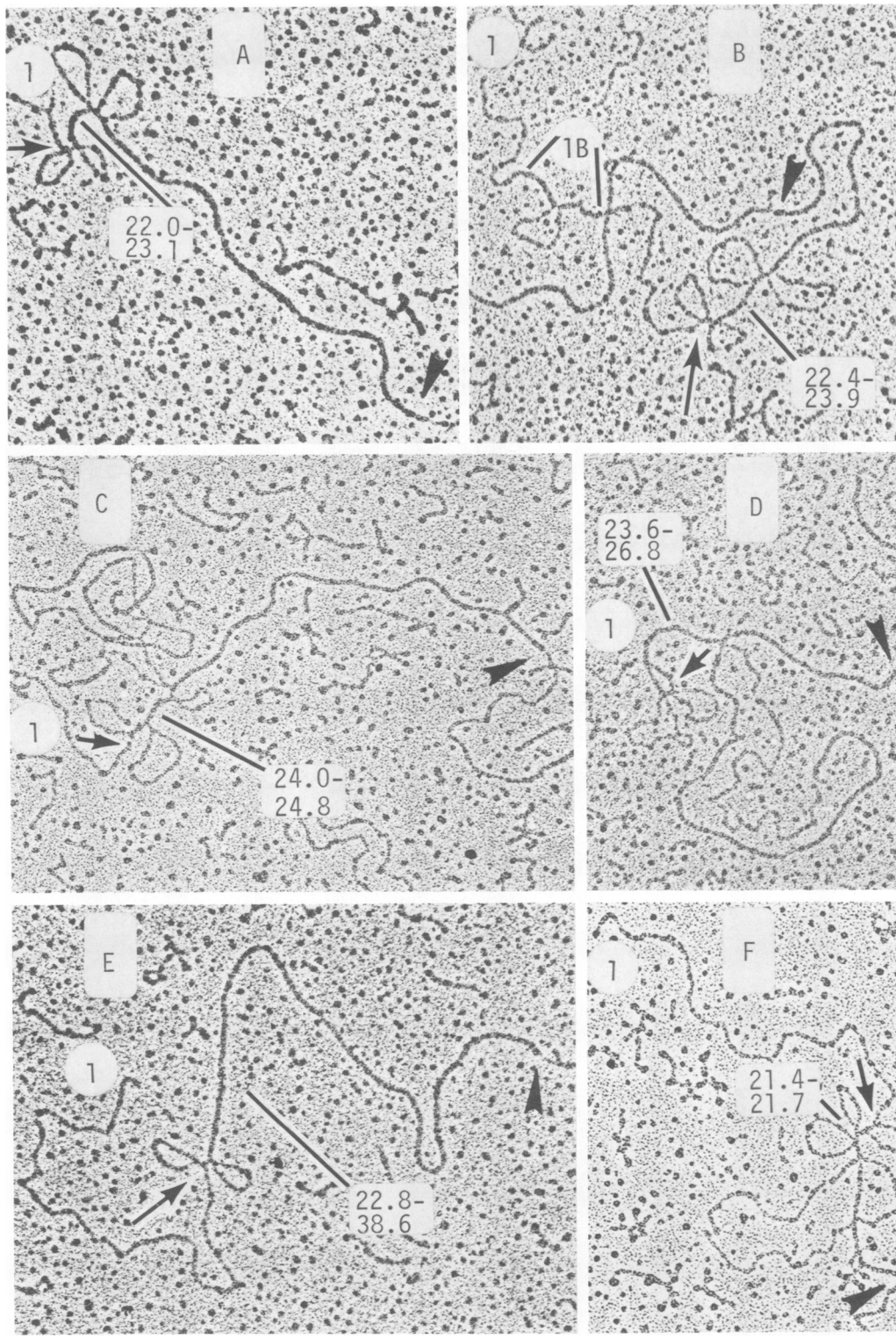

FIG. 5. *Ad3 and Ad7 r-strands hybridized to transcripts containing extra RNA segments. All transcripts in this figure contain the common three leaders mapping at coordinates 16.5, 19.8, and 26.8 (cf. Table 1). In addition each transcript contains an extra RNA segment located between the second and third leaders as indicated by the bar. Coordinates are given for the extra segments. (A) through (E) are from Ad3; (F) is from Ad7. The map coordinates for the message bodies are: (A) 30.2 to 40.8; (B) 31.2 to 39.2; (C) 52.8 to 62.9; (D) 52.1 to 56.8; (E) leader sequences continuous through the message body; (F) 44.7 to 49.1.*

strained into four loops rather than three. In addition to the standard tripartite leaders, these transcripts contained one of several different extra sequences (Table 4). In the most frequently observed type, the extra segment extended from 22.0 ± 0.4 (29 measurements) to 22.9 ± 0.4 (29 measurements) (Fig. 5A). A second version mapped from 22.9 ± 0.4 (four measurements) to 24.0 ± 0.3 (four measurements) (Fig. 5B). In a third type, the extra segment mapped from 23.8 ± 0.4 (three measurements) to 24.3 ± 0.6 (three measurements) (Fig. 5C). The longest of these extra leader region segments extended from 22.4 ± 0.5 (four measurements) continuously through the third leader at 26.9 ± 0.2 (three measurements). A third leader of intermediate length extended from 26.1 ± 0.1 (four measurements) to 26.9 ± 0.1 (four measurements). Other rare partially spliced molecules are shown in Fig. 5D, E, and F. In general, only one extra segment was found per molecule. Figure 6 illustrates the arrangements of RNA components in the leader region of these molecules. The presence of these extra components did not appear to be restricted to any given mRNA species, and no alterations in coding positions or lengths of the other spliced segments were found.

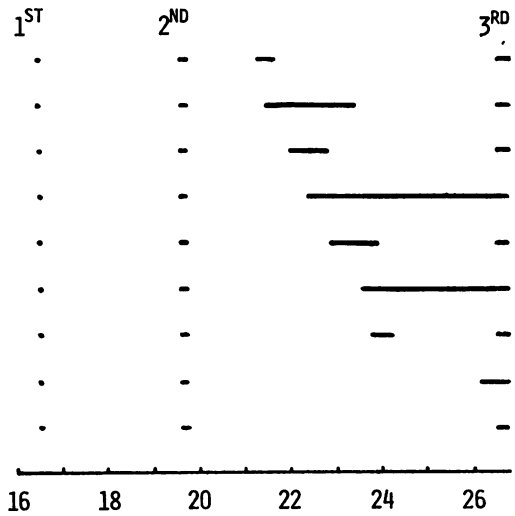

FIG. 6. *Leader components of the presumptive processing intermediates in the leader region of both Ad3 and Ad7 transcripts. For comparison, the leaders on the mature mRNA's are also presented at the bottom of the figure. Map coordinates of these leaders are listed in Table 4.*

DISCUSSION

The experiments described in this report reveal that the phenomena and the patterns of mRNA splicing are highly conserved among several adenovirus serotypes regardless of their genetic or sequence relatedness. Late mRNA of class C serotypes (Ad1, Ad2, Ad5, and Ad6) contain tripartite leaders spliced to the 5′ end of the message bodies. Each of the three leader components is highly homologous in nucleotide sequence. Similarly, late mRNA from the class B serotypes (Ad3 and Ad7) contains tripartite leaders at the 5′ end of message bodies. The Ad3 and Ad7 leaders appear to be identical in coding position and nucleotide sequence and are encoded at the same map positions as the tripartite leader segments of class C serotypes. However, the class B serotype leaders have substantially diverged in nucleotide sequence from the class C serotype leaders to the point where they no longer cross-anneal under relaxed hybridization conditions.

TABLE 4. *5′ and 3′ ends of extra RNA segments in the common leader region of class B serotypes*

5′ ends		3′ ends	
21.4	(1)[a]	21.7	(1)
21.6	(1)	23.5	(1)
22.0 ± 0.4	(29)	22.9 ± 0.4	(29)
22.4 ± 0.5	(4)	26.6 ± 0.2	(3)[b] [third leader]
22.9 ± 0.4	(4)	24.0 ± 0.3	(4)
23.6	(1)	26.8	(1) [third leader]
23.8 ± 0.4	(3)	24.3 ± 0.6	(3)
26.1 ± 0.1	(4)	26.8 ± 0.1	(4) [third leader]

[a] Numbers in parentheses are the numbers of measurements.

[b] One molecule has coordinates 16.6/19.9/22.8–38.6.

The common tripartite leaders of both Ad3 and Ad7 are present on a variety of late message species with 5′ ends mapping at essentially the same coordinates as in Ad2 (6). In the same region as the Ad2 message for the IVa2 protein, the class B serotypes encode a small *l*-strand mRNA (14.4 to 10.0) with a single-spliced 5′ leader (15.4). Thus, as with Ad2, the class B serotypes encode two transcription promotors on opposite strands within the genome interval 15.4 to 16.5 (about 385 base pairs).

Early mRNA of the class B serotypes is spliced similarly to Ad2 mRNA's in at least the early region 1B (*r*-strand, coordinates 4.5 to 11.1) and region 2 (DNA binding protein) (*l*-strand, coordinates 75 to 61). We expect that a more extensive characterization of spliced early mRNA's will reveal splicing patterns identical with Ad2 for other versions of mRNA in these regions and in early regions 1A, 3 and 4 (3, 18; Chow et al., submitted for publication). Although homology of nucleotide sequences in early mRNA was not directly examined between class B and class C serotypes, at least some of

these sequences have significantly diverged since no region 1B spliced mRNA was found in cross-hybridization mixtures of mRNA and heterologous DNA. This is compatible with the nonhomology in this region as determined by DNA:DNA heteroduplex analysis (14). Very few examples of transcripts from the genome position expected for fiber (86.3 to 91.3), IVa2, and single-stranded DNA binding proteins were found under a variety of hybridization conditions, suggesting a reduced production of these messages in the class B serotypes at the times of RNA harvest. The results of in vitro protein synthesis programmed with Ad3 and Ad7 polysomal mRNA isolated at 48 h were consistent with these lower relative abundances (B. A. Kilpatrick and J. B. Lewis, unpublished data).

Filter hybridization (17, 20) and heteroduplex mapping studies (14) have revealed extensive DNA homology among adenovirus serotypes belonging to the same serological class. These studies have reported 85 to 90% homology among class C serotypes and among class B serotypes. However, comparative cross-hybridizations between class B and class C serotypes revealed only 10 to 25% DNA homology. Thus, it is intriguing that the genome organization (36; this study) and splicing patterns of mRNA species are faithfully maintained. This may also be true for class A adenoviruses. Our preliminary studies of Ad12 (class A), which has 10 to 25% homology with class C serotypes, have shown that its late polysomal mRNA contains tripartite leaders that are not homologous with the class B and class C serotype leaders (data not shown). Yet, saturation hybridization experiments of Ad12 have revealed major blocks of early and late transcripts similar to those described for Ad2, Ad3, and Ad7 (31, 33).

Adenovirus serotypes belonging to the same class recombine readily and produce progeny (30). However, it has not been possible to obtain viable recombinants between serotypes belonging to different classes. Interclass recombination experiments among a variety of Ad5 (class C) temperature-sensitive mutants and wild-type Ad12 viruses (class A), for instance, resulted in production of some progeny virions possessing antigenic determinants from both serotypes (39). Analysis revealed that these progeny were the result of phenotypic mixing of certain proteins, and not of genetic recombination. A general nonhomology of DNA sequences appears to pose a primary barrier to genetic recombination between classes. The likelihood that the inverted terminal duplications must match one another for DNA replication to proceed (21) may establish another barrier. Nevertheless, a low frequency of recombinants might still be anticipated as a result of double crossover. However, we propose that another, possibly more formidable requirement for recombination between classes exists; it is likely that leader sequences must be matched to message sequences at the splice junctions or that separated regions of the primary transcripts must establish secondary structures to draw the remote conserved segments together for RNA splicing to occur. It should be possible to test this hypothesis by marker rescue studies with DNA restriction fragments as donors of wild-type alleles to mutant viruses from different classes.

The divergence of nucleotide sequences in leader components does not rule out the possibility that a short nucleotide sequence, homologous secondary structure, and common host-encoded RNA splicing enzyme are primarily responsible for RNA splicing. Work is in progress to determine nucleotide sequences at the splice junctions in Ad3 and Ad7 mRNA's. By comparing the nucleotide sequences and the secondary structures of class B and class C viral RNAs, it should be possible to determine the salient features of these transcripts that signal and possibly control splicing.

Several polysomal mRNA species were found in which the common second and third leaders were not directly spliced together but were, instead, separated by an additional RNA segment. The extra RNA segments formed more than one distinct group, but often shared one of several common 5′ or 3′ ends; one group with a 5′ end at 22.0 and a 3′ end at 22.9 occurred most frequently. These molecules were found in polysomal RNA prepared with or without detergent disruption of infected cells. They are presumptive intermediates of mRNA splicing in which the extra RNA components represent residual, unspliced portions of the primary transcript encoded between the second and third leader. The structures are significant since they likely reflect features in the general scheme by which adenovirus messages are processed from the primary transcript to mature, fully spliced mRNA's. Since these species with extra leaders were found at a low frequency in the polysomal RNA preparations, we cannot determine whether they originate from nuclear leakage. However, some of the same species have been found in high abundance in cytoplasmic RNA preparations from Ad2 and Ad2-simian virus 40 hybrid-infected cells (L. Chow, J. Lewis, T. Grodzicker, J. Sambrook, and T. Broker, unpublished data). This suggests that at least some of the rare Ad3 and Ad7 polysomal transcripts are not derived from nuclear leakage. As will be described in detail in this latter study, these presumptive intermediates indicate that splicing of adenovi-

rus messages occurs via numerous intermediate steps in which small, defined portions of RNA are successively removed (see also reference 6). Since it is not yet clear that these mRNA species are indeed processing intermediates, as opposed to aberrant molecules, it will be necessary to follow their appearance and ultimate fate during infection to sort out their role in RNA maturation.

ACKNOWLEDGMENTS

We thank Merilyn Sleigh and James McDougall for providing initial virus seed stocks. We also thank John Scott for maintaining the electron microscope facilities and Marie Moschitta for excellent assistance with manuscript preparation.

This research was sponsored by Public Health Service grant CA13106 from the National Cancer Institute to Cold Spring Harbor Laboratory. B.A.K. is the recipient of a Damon Runyon-Walter Winchell Cancer Fund postdoctoral fellowship, DRG-219.

LITERATURE CITED

1. **Anderson, C. W., J. B. Lewis, J. F. Atkins, and R. F. Gesteland.** 1974. Cell-free synthesis of adenovirus-2 proteins programmed by fractionated mRNA: a comparison of polypeptide products and mRNA lengths. Proc. Natl. Acad. Sci. U.S.A. **71**:2756–2760.
2. **Berget, S. M., C. Moore, and P. A. Sharp.** 1977. Spliced segments at the 5′ terminus of adenovirus-2 late mRNA. Proc. Natl. Acad. Sci. U.S.A. **74**:3171–3175.
3. **Berk, A. J., and P. A. Sharp.** 1978. Structure of the adenovirus-2 early mRNAs. Cell **14**:695–711.
4. **Broker, T. R., L. T. Chow, A. R. Dunn, R. E. Gelinas, J. A. Hassell, D. F. Klessig, J. B. Lewis, R. J. Roberts, and B. S. Zain.** 1978. Adenovirus-2 messengers—an example of baroque molecular architecture. Cold Spring Harbor Symp. Quant. Biol. **42**:531–553.
5. **Büttner, W., Z. Veres-Molnár, and M. Green.** 1976. Preparative isolation and mapping of adenovirus-2 early messenger RNA species. J. Mol. Biol. **107**:93–114.
6. **Chow, L. T., and T. R. Broker.** 1978. The spliced structures of adenovirus-2 fiber message and the other late mRNAs. Cell **15**:497–510.
7. **Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts.** 1977. An amazing sequence arrangement at the 5′ ends of adenovirus-2 messenger RNA. Cell **12**:1–8.
8. **Chow, L. T., J. M. Roberts, J. B. Lewis, and T. R. Broker.** 1977. A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. Cell **11**:819–836.
9. **Craig, E. A., M. Sayavedra, and H. J. Raskas.** 1977. Strand assignment of polyadenylated nuclear RNAs synthesized early in infection with adenovirus-2. Virology **77**:545–555.
10. **Davis, R. W., M. Simon, and N. Davidson.** 1971. Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. Methods Enzymol. **21**:413–428.
11. **Flint, S. J.** 1977. Two "early" mRNA species in adenovirus type 2-transformed rat cells. J. Virol. **23**:44–52.
12. **Flint, S. J., P. H. Gallimore, and P. A. Sharp.** 1975. Comparison of viral RNA sequences in adenovirus 2 transformed and lytically-infected cells. J. Mol. Biol. **96**:47–68.
13. **Fraser, N., and E. Ziff.** 1978. RNA structures near poly(A) of adenovirus-2 late messenger RNAs. J. Mol. Biol. **124**:27–51.
14. **Garon, C. F., K. W. Berry, J. C. Hierholzer, and J. A. Rose.** 1973. Mapping of base sequence heterologies between genomes from different adenovirus serotypes. Virology **54**:414–426.
15. **Gelinas, R. E., and R. J. Roberts.** 1977. One predominant 5′ undecanucleotide in adenovirus 2 late messenger RNAs. Cell **11**:533–544.
16. **Goldberg, S., D. Nevins, and J. E. Darnell.** 1978. Evidence from UV transcription mapping that late adenovirus type 2 mRNA is derived from a large precursor molecule. J. Virol. **25**:806–810.
17. **Green, M., J. T. Parsons, M. Piña, K. Fujinaga, H. Caffier, and I. Landgraf-Leurs.** 1970. Transcription of adenovirus genes in productively infected and in transformed cells. Cold Spring Harbor Symp. Quant. Biol. **35**:803–818.
18. **Kitchingman, G. R., S.-P. Lai, and H. Westphal.** 1977. Loop structures in hybrids of early RNA and the separated strands of adenovirus DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:4392–4395.
19. **Klessig, D. F.** 1977. Two adenovirus mRNAs have a common 5′ terminal leader sequence encoded at least 10 kb upstream from their main coding regions. Cell **12**:9–21.
20. **Lacy, S., and M. Green.** 1967. The mechanism of viral carcinogenesis by DNA mammalian viruses: DNA-DNA homology relationships among the 'weakly' oncogenic human adenoviruses. J. Gen. Virol. **1**:413–418.
21. **Lechner, R. L. and T. J. Kelly, Jr.** 1977. The structure of replicating adenovirus-2 DNA molecules. Cell **12**:1007–1020.
22. **Lewis, J. B., C. W. Anderson, and J. F. Atkins.** 1977. Further mapping of late adenovirus genes by cell-free translation of RNA selected by hybridization to specific DNA fragments. Cell **12**:37–44.
23. **McGrogan, M., and A. J. Raskas.** 1978. Two regions of the adenovirus-2 genome specify families of late polysomal RNAs containing common sequences. Proc. Natl. Acad. Sci. U.S.A. **75**:625–629.
24. **Meyer, J., P. D. Neuwald, S.-P. Lai, J. V. Maizel, Jr., and H. Westphal.** 1977. Electron microscopy of late adenovirus type 2 mRNA hybridized to double-stranded viral DNA. J. Virol. **21**:1010–1018.
25. **Neuwald, P. D., J. Meyer, J. V. Maizel, Jr., and H. Westphal.** 1977. Early gene expression of adenovirus type 2: R-loop mapping of mRNA and time course of viral DNA, mRNA, and protein synthesis. J. Virol. **21**:1019–1030.
26. **Nevins, J. R., and J. E. Darnell.** 1978. Groups of adenovirus type 2 mRNA's derived from a large primary transcript: probable nuclear origin and possible common 3′ ends. J. Virol. **25**:811–823.
27. **Palmiter, R. D.** 1974. Magnesium precipitation of ribonucleoprotein complexes. Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. Biochemistry **13**:3606–3615.
28. **Pettersson, U., and J. Sambrook.** 1973. Amount of viral DNA in the genome of cells transformed by adenovirus type 2. J. Mol. Biol. **73**:125–130.
29. **Pettersson, U., C. Tibbetts, and L. Philipson.** 1976. Hybridization maps of early and late mRNA sequences on the adenovirus type 2 genome. J. Mol. Biol. **101**:479–502.
30. **Sambrook, J., J. Williams, P. A. Sharp, and T. Grodzicker.** 1975. Physical mapping of temperature-sensitive mutations of adenoviruses. J. Mol. Biol. **97**:369–390.
31. **Scheidtmann, K.-H., and W. Doerfler.** 1977. Transcription of the genome of adenovirus type 12. IV. Maps of stable late RNA from productively infected human cells. J. Virol. **22**:585–590.
32. **Schreier, M., and T. Staehelin.** 1973. Initiation of mammalian protein synthesis: the importance of ribosome and initiation factor quality for in vitro synthesis. J. Mol. Biol. **73**:329–349.

33. **Smiley, D. R., and S. Mak.** 1978. Transcription map for adenovirus type 12 DNA. J. Virol. **28**:227–239.
34. **Sugden, B., B. DeTroy, R. J. Roberts, and J. Sambrook.** 1975. Agarose slab-gel electrophoresis equipment. Anal. Biochem. **68**:36–46.
35. **Thomas, M., R. L. White, and R. W. Davis.** 1976. Hybridization of RNA to double stranded DNA: formation of R-loops. Proc. Natl. Acad. Sci. U.S.A. **73**: 2294–2298.
36. **Tibbetts, C.** 1977. Physical organization of subgroup B human adenovirus genomes. J. Virol. **24**:564–579.
37. **Weber, J., W. Jelinik, and J. E. Darnell, Jr.** 1977. The definition of a large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. Cell **10**:611–616.
38. **White, R. L., and D. S. Hogness.** 1977. R-loop mapping of the 18s and 28s sequences in the long and short repeating units of *Drosophila melanogaster* rDNA. Cell **10**:177–192.
39. **Williams, J., H. Young, and P. Austin.** 1975. Complementation of human adenovirus type 5 *ts* mutants by human adenovirus type 12. J. Virol. **15**:675–675.
40. **Zain, S., J. Sambrook, R. J. Roberts, W. Keller, M. Fried, and A. R. Dunn.** 1979. Nucleotide sequence analysis of the leader segments in a cloned copy of adenovirus-2 fiber mRNA. Cell **16**:851–861.
41. **Ziff, E. B., and R. M. Evans.** 1978. The promoter and capped 5′ terminus of RNA from the adenovirus-2 major late transcription unit. Cell **15**:1463–1475.