

## Early Region 1B of Adenovirus 2 Encodes Two Coterminal Proteins of 495 and 155 Amino Acid Residues

CARL W. ANDERSON,<sup>1</sup> ROBERT C. SCHMITT,<sup>2</sup> JOHN E. SMART,<sup>3†</sup> AND JAMES B. LEWIS<sup>2\*</sup>

*Biology Department, Brookhaven National Laboratory, Upton, New York 11973<sup>1</sup>; Fred Hutchinson Cancer Research Center, Seattle, Washington 98104<sup>2</sup>; and Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724<sup>3</sup>*

Received 17 October 1983/Accepted 7 February 1984

Partial sequence analysis of tryptic peptides has identified the E1B-495R (E1b-57K) (early transcription region 1B of 495 amino acid residues, with an approximate molecular weight of 57,000) protein of adenovirus 2 as encoded by the 495 amino acid open reading frame located in the adenovirus 2 DNA sequence between nucleotides 2016 and 3500. Additional proteins of 16,000  $M_r$  and 18,000  $M_r$  that are related to the E1B-495R protein were identified by cell-free translation of hybridization-selected mRNA. Analysis of [<sup>35</sup>S]methionine-containing amino terminal tryptic peptides by thin-layer chromatography showed that the E1B-495R, E1B-18K, and E1B-16K proteins all begin at the same initiation codon. The E1B-495R protein from 293 cells also has the same initial tryptic peptide, acetyl-methionyl-glutamyl-arginine. Sequence analysis of E1B-18K tryptic peptides indicated that this protein also has the same carboxy terminus as the E1B-495R protein and that it is derived from an mRNA that is spliced to remove sequences between nucleotides 2250 and 3269, resulting in a protein product of 155 amino acid residues. Analysis of E1B-16K tryptic peptides has not yet revealed the carboxy terminal structure of this protein. Both the E1B-495R and the E1B-155R (E1B-18K) proteins, as well as the E1B-16K protein, were precipitated from cell-free translations and from extracts of infected cells by antiserum against an amino terminal nonapeptide common to these proteins.

Early transcription region 1B (E1B) of adenovirus, located between 4.5 and 11 map units (m.u.) on the viral genome, is necessary for oncogenic transformation by adenovirus and for one or more unidentified functions during productive growth of the virus in cultured cells (9, 12, 17, 20, 22, 24, 49). DNA comprising E1B is not by itself capable of transforming rodent cells, but together with E1A DNA it can induce morphologically complete transformation (23). In addition, one or more E1B functions is responsible for the greater tumorigenicity in nude mice of adenoviruses of subgroup A (adenovirus 12 [Ad12]) compared with subgroup C (Ad5) (5, 6, 51). Careful analysis of how much of region E1B is required for transformation suggests that several E1B gene products are involved in a complex process that differs for transformation via DNA transfection and transformation via virus infection (6, 9, 45).

To study these differences further, it is important to identify and characterize each of the E1B-encoded proteins and to determine which DNA sequences are necessary to encode each protein. The DNA sequences of Ad5 (7) and Ad2 (15) show two open reading frames that could correspond to E1B proteins: the first AUG from the mRNA start would initiate synthesis of a protein of 175 amino acid residues; the second AUG from the start, in a different reading frame from the first, would begin a protein of 495 residues (the number of residues refers to the Ad2 sequence). Proteins of approximate apparent molecular weights of 15,000 and 57,000 (E1B-15K, E1B-57K) have been translated *in vitro* from E1B-specific mRNAs (19, 32). The E1B-57K protein corresponds to a prominent tumor antigen immunoprecipitated from infected cells (14, 28, 30, 44, 48, 53). Variable immunoprecipitation of the E1B-15K protein has been reported, but this product has recently been

purified (43). A third E1B protein, E1B-18K, encoded by an mRNA distinct in size from those that encode the E1B-15K and 57K proteins, has been noted (10). This E1B-18K protein is probably the same as the E1B-20K protein that has been purified and shown to be related to the E1B-57K protein in tryptic peptide composition and distinct from the E1B-15K protein (8, 18, 36). Also encoded within E1B is virion component IX, whose mRNA has a different 5' end but is 3'-coterminal with the E1B mRNAs (1).

Partial amino acid sequence analysis of the amino terminus of the E1B-15K protein has confirmed its identification with the E1B-175 residue open reading frame (4). Here, we report confirmation by protein sequence studies of the correlation of the E1B-57K protein with the 495 residue open reading frame and the identification of the E1B-18K protein as comprising the amino- and carboxy terminal segments of the E1B-57K protein, with 340 internal residues deleted. The relationships of the various E1B proteins and mRNAs to the E1B DNA sequence, including the results presented here, are summarized in Fig. 1.

The literature of adenovirus proteins has been confused by the use of apparent molecular weight to designate viral proteins, with different laboratories assigning different numbers to the same protein. Some of the alternative designations of the E1B proteins are noted in the legend to Fig. 1. We suggest that viral proteins be identified in terms of apparent molecular weight only when their DNA coding sequences are in doubt, and thereafter by the number of amino acid residues comprising the protein. Accordingly, the E1B-15K, E1B-57K, and E1B-18K proteins are referred to here as E1B-175R, E1B-495R, and E1B-155R proteins, respectively.

### MATERIALS AND METHODS

**Cells and viruses.** Human adenovirus 2 was propagated on HeLa cells as previously described (4).

**Preparation of E1B proteins by translation *in vitro*.** E1B

\* Corresponding author.

† Present address: Biogen, Inc., Cambridge, MA 02142.

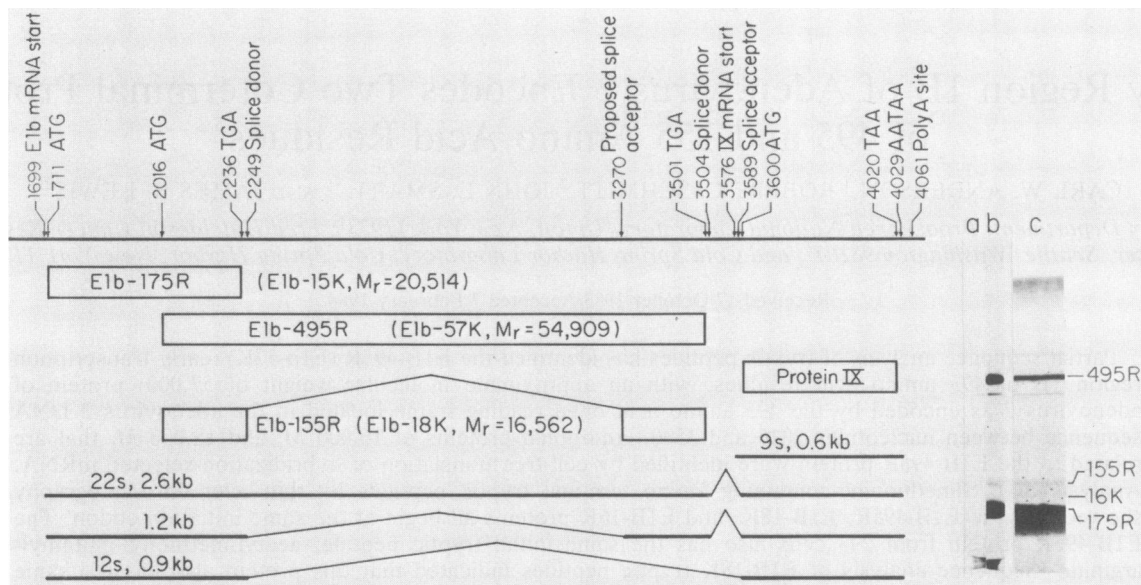


FIG. 1. Map of adenovirus 2 early region 1B, showing the protein encoding regions and mRNA structures. The positions (nucleotide number) on the Ad2 DNA sequence (15) of various E1B landmarks are shown along the top, with the mRNA structures and protein encoding locations drawn below. The structure of the E1B-155R protein and mRNA are deduced from the peptide and sequence information presented here. The new nomenclature for adenovirus proteins suggested in the text is used for the E1B proteins, with our older designations in parentheses, along with the theoretical molecular weight calculated from the protein sequence. The E1B-175R protein corresponds to the E1B-19K protein of Matsuo et al. (36), and the E1B-155R protein corresponds to their E1B-20K protein. On the right is shown the products of translation in vitro of E1B mRNA. Lane a is a reaction with the addition of no RNA; lane b is a reaction with the addition of E1B mRNA from early productive infection, increased in concentration by inhibition of protein synthesis, as described in the text. Lanes a and b were from a fluorogram exposed for 16 h. Lane c is a sector of one gel used for the preparation of in vitro-synthesized E1B proteins. To facilitate elution of the proteins, the preparative gel was not fixed but instead was washed twice with water for 20 min each time, dried, and autoradiographed for 3 days.

mRNA was prepared from total cytoplasmic RNA by hybridization to DNA immobilized on nitrocellulose filters and then translated in vitro as previously described (31). RNA was prepared at 8 h postinfection (p.i.) from cells that had been treated with 25  $\mu$ g of cycloheximide per ml or 10  $\mu$ M anisomycin (33) from 3 to 8 h p.i. The DNA segment used to select E1B mRNA was either the region 4.30 to 7.67 m.u. (*Hpa*I to *Hind*III) cloned into the plasmid pBR322 or a *Pst*I fragment of this segment (5.01 to 6.85 m.u.) subcloned into the single-strand phage M13 in the orientation complementary to E1B mRNA.

**Protein and peptide analysis.** Electrophoretic separation of proteins and trypsin digestion were done as described previously (31), except that in some cases (e.g., see Fig. 5) the intact proteins were eluted from the gel slice before oxidation and digestion. Reverse-phase chromatography was performed as described by Smart et al. (50) (see Fig. 2) or by Oosterom-Drögen and Anderson (39) (see Fig. 5). Sequence analysis was done as described previously (2, 50), as was thin-layer chromatography (31).

**Preparation of peptide-specific antibodies.** The peptide acetyl-Met-Glu-Arg-Arg-Asn-Pro-Ser-Glu-Arg-Cys-Lys was synthesized by using the Merrifield technique as previously described (31). The first nine residues correspond to the nine amino-terminal residues of the E1B-495 codon open reading frame; the last two residues were included to facilitate conjugation to carrier protein for immunization.

The synthetic peptide (5 mg) was coupled to carrier protein (10 mg of bovine serum albumin [BSA]; Sigma Chemical Co.) by using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (Sigma) (16). Sephadex G-50 chromatography in phosphate-buffered saline (PBS) was

used to separate unlinked from coupled peptide. Approximately 10% of the peptide was linked to BSA.

Antiserum was obtained by immunization of adult female New Zealand white rabbits with 200  $\mu$ g of peptide coupled to BSA. The conjugate was emulsified with an equal volume of Freund complete adjuvant (GIBCO Diagnostics) and injected intradermally at several sites. Rabbits were boosted at 2 weeks with 100  $\mu$ g of conjugated peptide emulsified with Freund incomplete adjuvant. Subsequent boosts were administered at monthly intervals. Antiserum was collected 1 week after boosting.

A protein A-Sepharose column was used to isolate the immunoglobulin G fraction from the crude antiserum after ammonium sulfate precipitation (11). The sample in 10 mM sodium phosphate (pH 8.0) was loaded on the column, washed with loading buffer, and eluted with 0.1 M citric acid (pH 3.0). Fractions were collected into 1.5 M Tris-hydrochloride (pH 8.9) to neutralize the acid.

Antibodies specific for the synthetic peptide were purified, using affinity chromatography by the procedure of Gentry et al. (13). The immunoglobulin G fraction in PBS was first passed over a BSA-Sepharose column to eliminate antibodies to the BSA carrier and then loaded onto a peptide-Sepharose column. After washing the final column with PBS and 2 M KCl in PBS, the bound specific antibody was eluted with 5 M NaI-1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>. The peak fractions were dialyzed against PBS, made 2 mg per 100 ml in BSA, and stored at -70°C. The BSA-Sepharose and peptide-Sepharose columns were prepared by using CNBr-activated Sepharose 4B (35).

**Immunoprecipitation.** HeLa cell monolayers at 80% confluency were infected with Ad2 (100 PFU per cell), essential-

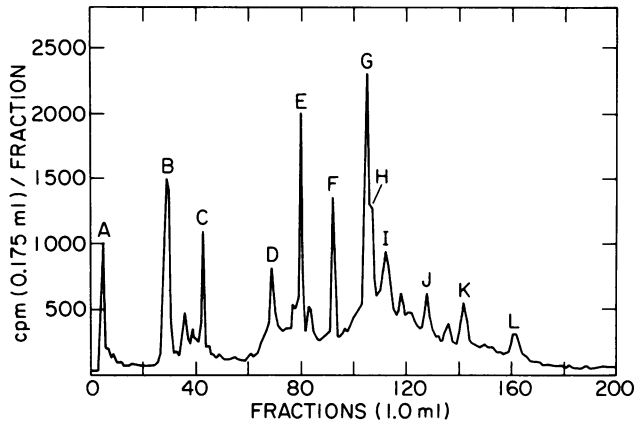


FIG. 2. Reverse-phase high-pressure chromatography of in vitro-synthesized, [<sup>35</sup>S]methionine-labeled E1B-495R tryptic peptides. E1B-495R protein was synthesized in a cell-free system programmed with hybridization-selected E1B mRNA as described in the text. After SDS-polyacrylamide gel electrophoresis of the translation mixture, the region of the gel containing the E1B-495R protein was excised, performic acid-oxidized, and incubated with tolylsulfonyl phenylalanyl chloromethyl ketone-treated trypsin. Extracted peptides were dissolved in 20% formic acid and applied to a C-18 reverse-phase column equilibrated with 5% formic acid. Peptides A through L were eluted by application of a linear 0 to 62.5% ethanol gradient in 5% formic acid as previously described (50). A portion of each fraction (175  $\mu$ l) was mixed with scintillation fluid and counted for radioactivity. Appropriate fractions representing peaks of radioactivity were pooled and lyophilized for subsequent sequence analysis.

ly following the protocol of Harter and Lewis (21). Virus was added in 0.5 ml of serum-free medium per 10-cm plate. At 1 h p.i., 10 ml of methionine-free medium containing 5% fetal calf serum and cycloheximide was added. Cytosine arabinoside was added 4 h p.i. At 5 h p.i., cells were washed with methionine-free medium containing cytosine arabinoside, and 100  $\mu$ Ci of [<sup>35</sup>S]methionine was added in 3 ml of

methionine-free medium with 5% fetal calf serum and cytosine arabinoside. Labeling was continued for 3 h. Labeled cells were washed two times with 5 ml of PBS at 4°C, scraped into a solution containing 1 ml of 0.01 M Tris-hydrochloride (pH 7.4), 0.01 M NaCl, 0.0015 M MgCl<sub>2</sub>, 0.5% Nonidet P-40, and 5  $\mu$ g of aprotinin per ml, vortexed, and held on ice for 5 min. After centrifugation for 5 min in an Eppendorf centrifuge at 4°C, the supernatant was removed, and a volume containing 2.5  $\times 10^7$  cpm was made 0.1% in sodium dodecyl sulfate (SDS). The cell extract was pre-cleared by incubating with 5  $\mu$ l of crude preimmune serum for 30 min at 4°C. Immune complexes were removed by incubating for 30 min at 4°C with 70  $\mu$ l of a 10% suspension of *Staphylococcus aureus* in PBS containing 1% BSA and 0.01% NaN<sub>3</sub> (25). After centrifugation, the supernatant was incubated with 20  $\mu$ g of purified antibody or blocked antibody (antibody incubated with 20  $\mu$ g of peptide for 30 min at 4°C) for 2 h at 4°C. Immune complexes were removed by the addition of 70  $\mu$ l of a 10% suspension of *S. aureus*. The resulting pellet was washed two times with a solution containing 10 mM Tris-hydrochloride (pH 7.4), 0.15 M NaCl, 1% Nonidet P-40, 1% sodium desoxycholate, 0.1% SDS, and 5  $\mu$ g of aprotinin per ml and one time with each of the following: high salt (2 M NaCl, 10 mM Tris-hydrochloride [pH 7.4], 1% Nonidet P-40, 0.5% sodium desoxycholate), low salt (0.5% Nonidet P-40, 0.1% SDS in PBS), 1 M MgCl<sub>2</sub>, and 0.01 M Tris-hydrochloride (pH 7.5). The final pellet was suspended in 50  $\mu$ l of sample buffer (27) and heated in a boiling water bath for 3 min. The sample was centrifuged, and the supernatant was electrophoresed on a 15% SDS-polyacrylamide gel (27). Products of translation in vitro were processed for immunoprecipitation immediately after translation, as freezing the translation mixtures often resulted in nonspecific precipitation.

## RESULTS

Cell-free translation of E1B-mRNA produces the E1B-495R, E1B-175R (10, 19), and E1B-155R proteins (10, 36). The results of a typical translation are shown in Fig. 1, along

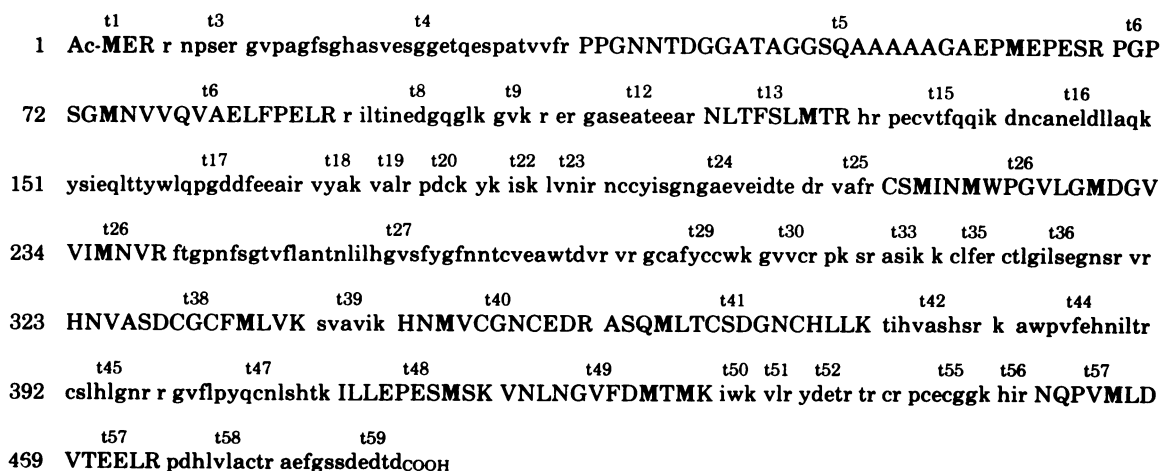


FIG. 3. Amino acid sequence and predicted tryptic peptides of the Ad2 E1B-495R protein. The amino acid sequence deduced from the 495 codon open reading frame in the Ad2 DNA sequence between nucleotides 2016 and 3500 (15) is shown in the single-letter code. The sequence has been broken (by a space) after each arginine (R) and lysine (K) codon to reflect the specificity of trypsin; note, however, that trypsin cleaves slowly when a proline (P) immediately follows either residue. Predicted tryptic peptides have been numbered sequentially from the amino terminus. Capitalized sequence indicates methionine-containing tryptic peptides; methionine (M) residues are in bold type. Numbers at the beginning of each line indicate the position of the first amino acid in the line from the amino terminus of the protein. The protein is shown in this work to have its initiating methionine blocked by acetylation (Ac-).

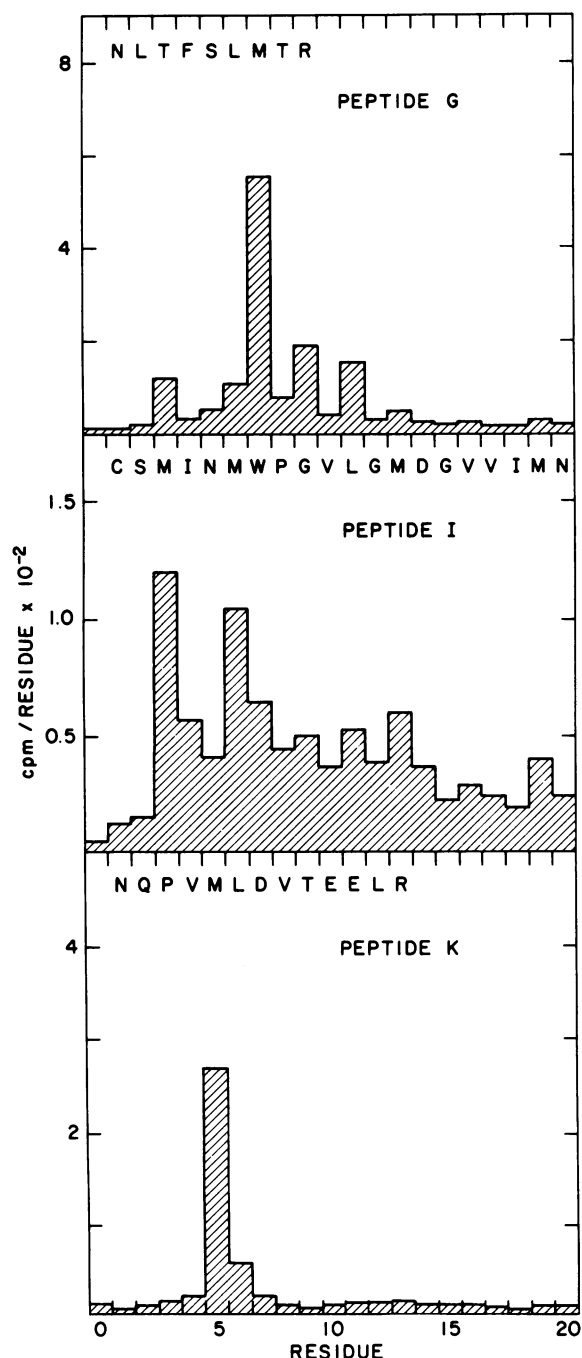


FIG. 4. Amino acid sequence analysis of selected Ad2 E1B-495R [<sup>35</sup>S]methionine-containing tryptic peptides. Individual peaks of radioactivity from the experiment shown in Fig. 2 were pooled, lyophilized, dissolved in 20% formic acid, and applied to a Beckman 890C protein sequencer as previously described (2, 50). The total radioactivity found in the amino acid residue after each Edman degradation cycle has been plotted for three selected peptides: G (2,450 cpm applied to sequencer), I (2,000 cpm applied), and K (620 cpm applied). The deduced sequence of each peptide, based on the nucleotide sequence of Ad2 DNA (15), is shown at the top of each panel. Similar experiments confirmed the identity of most of the other methionine-containing tryptic peptides from the chromatographic separation shown in Fig. 2 (Table 1).

with a sector of an electropherogram used to prepare the E1B-495R and E1B-155R proteins and a minor component of the 16K protein for analysis of peptide composition and amino acid sequence. The major product of translation *in vitro*, the E1B-175R protein, has already been subjected to amino terminal sequence analysis (4) and will not be considered here.

The tryptic peptides of the E1B-495R protein that contain methionine were resolved by reverse-phase chromatography (Fig. 2) and then subjected to sequence analysis to determine the position of the methionine residue(s) in each peptide. The sequence of the Ad2 E1B-495R protein, predicted from the DNA sequence (15), is shown in Fig. 3 divided into 59 tryptic peptides, with the 11 tryptic peptides that contain methionine residues shown in upper-case letters. As shown in Fig. 4 and Table 1, 8 of the 12 peaks resolved in Fig. 2 have been identified with 7 of the 11 tryptic peptides expected from the E1B-495R protein. Peaks J and K both have the amino terminus of peptide t57. The existence of two peptides with the amino terminal sequence of t57 might arise from inefficient cleavage of the Arg-Pro peptide bond predicted to occur at the carboxy terminus or from posttranslational modification, such as phosphorylation of a threonine residue. The seven peptides identified among the resolved peaks total 93 residues and span the region from residues 117 to 474 of the 495 codon open reading frame. Four expected peptides were not identified by sequence analysis, including the initial peptide t1, so that these experiments, although showing the structure of the large E1B protein to be consistent with the 495 codon open reading frame, do not establish the amino terminus of the protein. Subsequent analysis (see below) has shown that peptide t1, the amino terminal peptide, is blocked. Peptide A corresponds to the expected position of t1. The amino termini of peptides t5 and t6 are produced by cleavage at Arg-Pro bonds. This bond is only cleaved slowly by trypsin. The positions of methionines in peptides B and F (Fig. 4) were not identified by sequence analyses of these peptides (data not shown). Peptides B and F may correspond to partial trypsin cleavage products containing combinations of t4, t5, and t6.

To identify the amino terminus of the E1B-495R protein and to determine the relationship of this protein to the smaller products identified in Fig. 1 as 155R (E1B-18K) and 16K, the methionine-containing tryptic peptides from the three proteins were compared by reverse-phase chromatography (Fig. 5). An elution buffer and program different from that used to produce Fig. 2 was used for this comparison. We have previously determined (31) with this second system the elution position of acetyl-Met-Glu-Arg, the expected E1B-495R amino terminal tryptic peptide. Also included in this comparison was E1B-495R protein prepared by immunoprecipitation of methionine-labeled 293 cells. 293 is a human cell line that constitutively produces Ad5 E1B-495R protein (34). Ad5 E1B-495R protein differs from the predicted Ad2 E1B-495R sequence at only 11 residues, only one of which (a Thr to Ala change in t13; Fig. 3) is in a methionine-containing tryptic peptide that we have identified. Ad5 E1B-495R from 293 cells provided protein for identification of any amino terminal modification produced *in vivo*.

Peak M2 had the expected position of acetyl-Met-Glu-Arg, and all four proteins yielded methionine-containing peptides with this elution position. Peptide M2 from each of the four proteins and acetyl-Met-Glu-Arg comigrated during thin-layer chromatographic analysis (Fig. 6). Thus, each of these proteins has the same amino terminus. This amino terminus corresponds to the 5' end of the 495 codon open reading

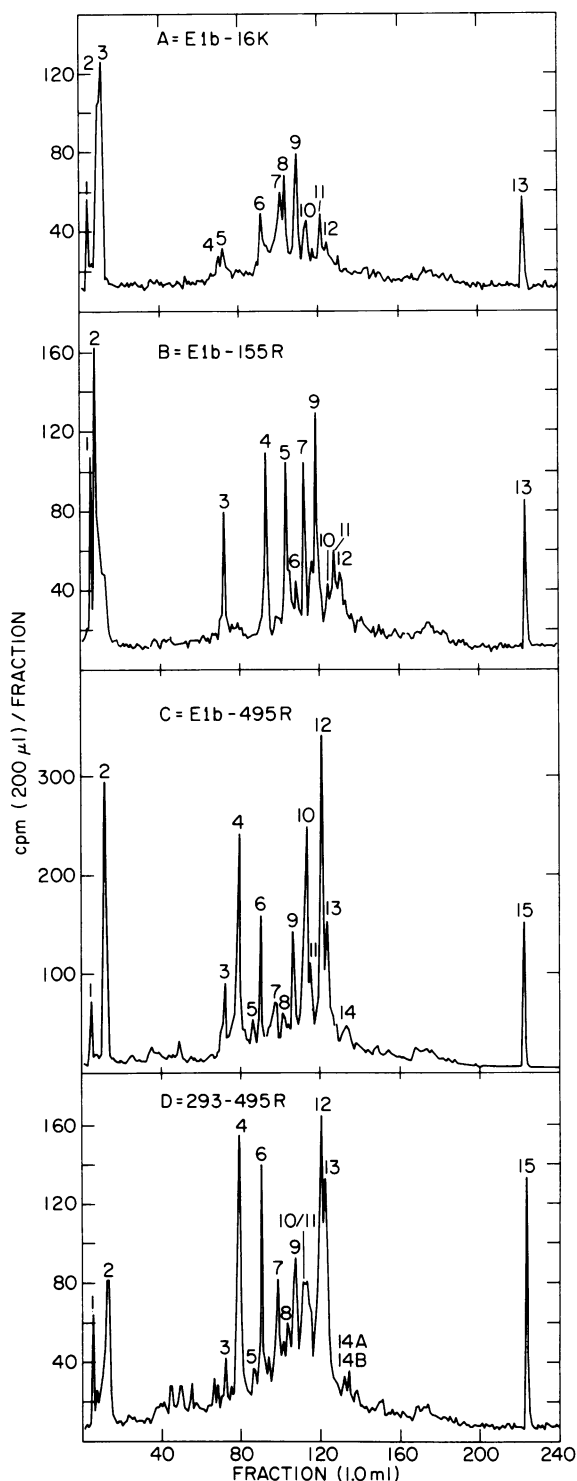


FIG. 5. Reverse-phase high-pressure chromatography of the tryptic peptides from the E1B-related E1B-16K, E1B-155R (E1B-18K), and E1B-495R proteins. The E1B-16K, E1B-155R, and E1B-495R proteins, labeled with [ $^{35}$ S]methionine, were prepared by cell-free translation of hybridization-selected mRNA as described in the text and the legend to Fig. 1. Ad5 E1B-495R protein was obtained by immunoprecipitation of [ $^{35}$ S]methionine-labeled 293 cells as described by Sarnow et al. (47). The translation products and the immunoprecipitate were fractionated by SDS-polyacrylamide gel electrophoresis, appropriate regions of the gels were excised, and

frame at N2016, and the E1B-495R product is acetylated at its amino terminus both in the rabbit reticulocyte cell-free protein synthesizing system and in vivo.

Sequence analysis of several E1B-495R peptides confirmed the above-described correspondence of this polypeptide with the 495 codon open reading frame, although the elution order of several peptides differed in the two buffer systems used (Table 1). The 293 cell (Ad5) 495R protein, synthesized in vivo, and the Ad2 495R protein, synthesized in vitro, gave essentially identical chromatographic profiles, despite the amino acid difference cited above. The major tryptic peptides from the E1B-155R protein eluted at positions similar to peptides from the E1B-495R protein. Sequence analysis of 155R peptides M7 and M9, corresponding in position to 495R peptides M9 and M10, respectively, indicated that the methionines matched positions in peptides t49 and t57, respectively, as predicted. Thus, the E1B-155R and 495R products share carboxy terminal sequences as well as their amino termini. This apparent conflict of common termini for two proteins of very different sizes was resolved by sequence analysis of E1B-155R peptide M3, which unlike its E1B-495R counterpart, was found to have methionine as its 6th and 14th residues (Fig. 7). No peptide with methionines spaced 8 residues apart can be directly encoded by Ad2 E1B, and only peptide t6 has a first methionine as its sixth residue. Inspection of the E1B nucleotide sequence revealed that a peptide with precisely this sequence of methionine residues would be produced by an mRNA splice after the glutamine codon 78 (from N2249) to the proline codon 419 (to N3270). Nucleotide N2249 is the donor site for a splice that produces the 12S or 0.9-kilobase mRNA that encodes the E1B-175R (E1B-15K) product. The normal acceptor site for this splice is N3589, just before the initiation site for polypeptide IX. Inspection of the sequence around N3270 reveals that this sequence also resembles a consensus of the known splice acceptor sites (Fig. 8). A splice from N2249 to a novel acceptor site at N3270 would produce an mRNA 319 nucleotides larger than the E1B-12S (0.9-kilobase) mRNA, in good agreement with a size estimate of 1.2 kilobases for the E1B-18K mRNA (10). An mRNA thus spliced would be expected to produce a protein of 155 amino acids, 78 of which correspond to the amino terminus and 77 to the carboxy terminus of the E1B-495R protein. We propose this to be the structure of the E1B-18K protein.

We have not yet identified any of the E1B-16K tryptic peptides, with the exception of M2, the amino terminus. Indeed, many of the peaks in the E1B-16K profile had considerably less radioactivity than M2 and may result from contaminants present in lower molar amounts. The E1B-16K

proteins were eluted from gel slices electrophoretically. After removal of SDS, performic acid oxidation, and digestion with trypsin, the resulting peptides were dissolved in 20% formic acid and applied to a C-18 reverse-phase column equilibrated in 10 mM ammonium acetate (pH 5.4)-1.0% acetonitrile. The column was developed with a series of discontinuous linear acetonitrile gradients as previously described (39). A 200- $\mu$ l portion of each fraction was mixed with scintillation fluid for determination of radioactivity, and the counts obtained have been plotted without background subtraction. Peaks of radioactivity have been numbered according to their elution order and are referred to in the text as M1, M2, etc., for methionine-containing peak 1, peak 2, etc. Appropriate peak fractions were pooled, mixed with apomyoglobin, and lyophilized for subsequent sequence analysis, which permitted the identification of many of these peptides (see Fig. 7 and Table 1).

TABLE 1. Identification of HPLC<sup>a</sup> peaks with the 495 codon E1B open reading frame of Ad2

Methionine tryptic peptides	Residues	Methionine positions	HPLC system 1, Ad2 495R	HPLC system 2			293 495R
				Ad2 495R	Ad2 155R	Ad2 16K	
t1	1-3	1 <sup>b</sup>	A	M2 <sup>c</sup>	M2 <sup>c</sup>	M2 <sup>c</sup>	M2 <sup>c</sup>
t5 <sup>d,e</sup>	38-68	26					
t6 <sup>d</sup>	69-87	6			M3 <sup>f</sup>		
t13	117-125	7	G	M12			
t26	218-239	3, 6, 13, 19	I				
t38	323-336	11					
t40	343-353	3	C				
t41	354-369	4	D	M3			
t48	415-424	8	E	M4			
t49	425-436	9, 11	H	M7, M9	M5, M7		
t57 <sup>e</sup>	462-474	5	J, K	M10, M13	M9		

<sup>a</sup> HPLC, High-pressure liquid chromatography.

<sup>b</sup> Amino terminus blocked by acetylation.

<sup>c</sup> Identity deduced from chromatographic behavior of synthetic acetyl-methionyl-glutamyl-arginine (see text).

<sup>d</sup> Amino terminus produced by cleavage at Arg-Pro bond, a site only cleaved slowly by trypsin.

<sup>e</sup> Carboxy terminus produced by cleavage at Arg-Pro bond, a site only cleaved slowly by trypsin.

<sup>f</sup> E1B-155R peptide M3 was shown to have methionine at residues 6 and 14 (Fig. 6); this sequence results from a splice which joins t6 sequences with t48 sequences.

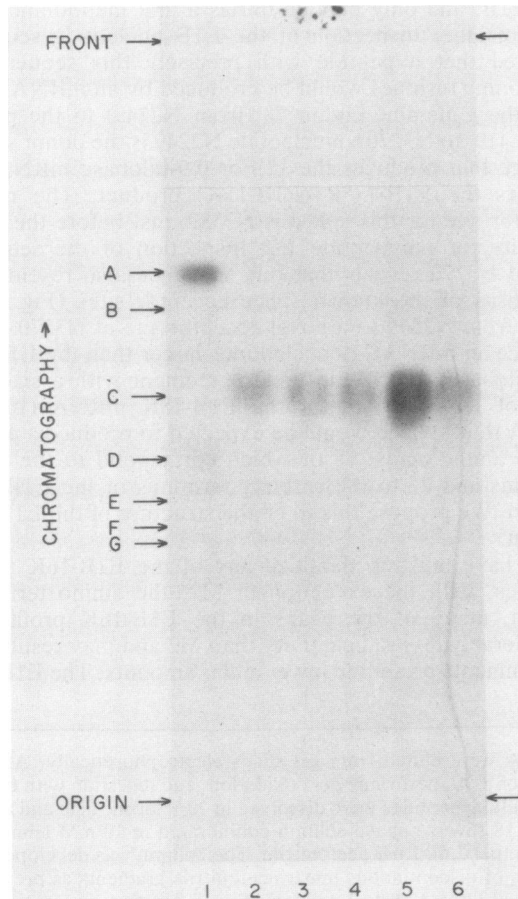


FIG. 6. Thin-layer chromatographic identification of the amino terminal tryptic peptide from the adenovirus E1B-related E1B-495R, E1B-155R, and E1B-16K proteins. Preliminary experiments indicated that peptide M2 from each of the E1B proteins was likely to correspond to the predicted amino terminal peptide methionyl-glutamyl-arginine. This peptide, from each of the experiments shown in Fig. 5, was lyophilized, dissolved in water, and applied to a cellulose thin-layer chromatography sheet together with appropriate

protein could represent the product of initiation of translation at the E1B-495R initiation site on the 0.9-kilobase mRNA, but this would be difficult to show with methionine-labeled material. Only peptide t5, which has not been identified, and a modified t6, in addition to peptide t1, would contain label.

To determine whether the E1B-155R and E1B-16K proteins were present in Ad2-infected cells, an antiserum prepared against the acetylated amino terminal nonapeptide common to the E1B-495R, 155R, and 16K polypeptides was used to immunoprecipitate methionine-labeled cell extracts. The E1B-495R, E1B-175R, E1B-155R, and E1B-16K proteins were all precipitated by the antipeptide antiserum from both Ad2-infected cell extracts and cell-free translations by using E1B-mRNA (Fig. 9). The precipitation of the E1B-175R protein was quite surprising since this protein has an entirely different amino terminus (4) and shares no sequence homology with the E1B-495R product. That this coprecipitation is specific is shown by the fact that it was blocked by pretreatment of the antiserum with the nonapeptide, that none of the E1B-related proteins precipitated with the preim-

synthetic peptide standards (31). After development in *n*-butanol-pyridine-acetic acid-water (300:200:60:240), the chromatogram was stained with cadmium-ninhydrin and then exposed to X-ray film at -70°C. Shown is a portion of a 2-week autoradiogram that contained radioactive samples: (1) synthetic <sup>14</sup>C-labeled acetyl-methionyl-glutamyl-arginine (250 cpm applied); (2) synthetic <sup>14</sup>C-labeled acetyl-methionyl-glutamyl-arginine after performic acid oxidation (330 cpm applied); (3) peptide M2 (Fig. 5) from the E1B-16K protein synthesized in vitro (350 cpm applied); (4) peptide M2 from the E1B-155R protein synthesized in vitro (220 cpm applied); (5) peptide M2 from the E1B-495R protein synthesized in vitro (1,300 cpm applied); and (6) peptide M2 from the E1B-495R protein isolated from 293 cells (450 cpm applied). Letters to the side of the autoradiogram indicate the positions of synthetic peptide standards chromatographed on the same thin-layer sheet and detected by ninhydrin staining or by autoradiography: (A) acetyl-methionyl-glutamyl-arginine; (B) methionine; (C) performic acid-oxidized acetyl-methionyl-glutamyl-arginine; (D) methionyl-glutamyl-arginine; (E) performic acid-oxidized methionyl-arginine; (F) methionine sulfone; and (G) performic acid-oxidized methionyl-glutamyl-arginine.

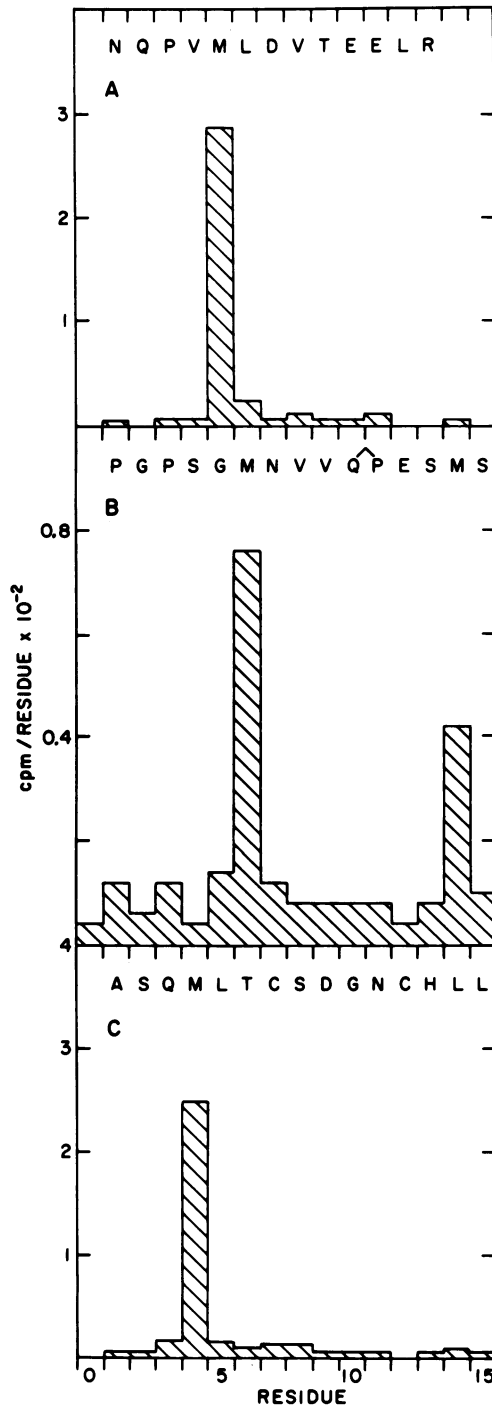
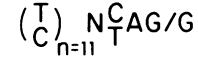


FIG. 7. Sequence analysis of [<sup>35</sup>S]methionine-containing tryptic peptides from adenovirus E1B-related proteins. [<sup>35</sup>S]methionine-labeled peptides from the experiment shown in Fig. 5 were subjected to automated protein sequence analysis as described in the text. The total radioactivity obtained in the amino acid fraction has been plotted for three selected peptides: (A) M9 from the E1B-155R protein (400 cpm applied to sequencer); (B) M3 from the E1B-155R protein (560 cpm applied); and (C) M3 from the E1B-495R protein made in vitro (570 cpm applied). Near the top of each panel, the predicted amino acid sequence, derived from the DNA sequence of Ad2 (15), is given in the single letter code. Peptide M3 from the E1B-155R protein spans the splice junction in the E1B-155R protein mRNA (see text); the position of the splice is indicated by a  $\blacktriangle$ .

Acceptor Consensus :



E1a Acceptor : GTGATTTTTTAAAAG/G n=7

E1b (9.83 m.u.) : GTATCTGTTTTGCAG/C n=8

E1b (8.96 m.u.) : AAGATATTGCTTGAG/C n=5

FIG. 8. Comparison of splice acceptor sequences. The novel splice junction used in the mRNA for the E1B-155R protein (E1B at 8.96 m.u.) is compared with the other splice acceptor site in E1B mRNAs (9.83 m.u.), the splice acceptor common to the E1A mRNAs, and the consensus sequence catalogued by Mount (38).

mune serum, and that proteins endogenous to the translation system and non-E1B proteins in the lysate were not precipitated.

### DISCUSSION

The peptide sequencing experiments reported here identify the amino terminus of the E1B-495R (57K) protein and confirm that this product is translated from the 495 codon open reading frame in E1B of the Ad2 DNA sequence. The amino terminus of this protein is acetylated in vivo and as synthesized in the rabbit reticulocyte cell-free system. Two related proteins of 16,562  $M_r$  (E1B-155R) and ca. 16,000  $M_r$  are also encoded by this reading frame, have the same amino termini as the 495 protein, and also have their amino termini acetylated by the reticulocyte lysate. Most adenovirus virion components, except those subjected to amino terminal proteolytic processing during virion assembly (3), have blocked, presumably acetylated, amino termini (29), but nonstructural proteins may have either blocked or free amino termini. The E4-116R (11K) protein isolated from infected cell extracts has nonacetylated Met-Ile-Arg as its amino terminal sequence (46). The proteins of the E1A region, synthesized in vivo (50), and the E3-16K protein, synthesized in vitro (42), both have nonacetylated Met-Arg as amino termini. In contrast, the Met-Arg amino terminus of the L0-13.6K (late transcription region) protein, encoded by the major late "i"-leader, is acetylated, as is the Met-Glu-Arg terminus of the L0-13.5K protein, encoded between the first and second late leader segments; both proteins are synthesized in vitro (31). Although some rules are emerging, we still cannot accurately predict whether a given amino terminus will be acetylated.

The E1B-155R (18K) protein was shown by Matsuo et al. (36) to be related in peptide composition to the E1B-495R product. Using partial amino acid sequence analysis, we have been able to predict the complete sequence of the E1B-155R product and to determine its precise relationship to the E1B-495R protein. The E1B-155R protein consists of the amino terminal 78 and carboxy terminal 77 residues of the 495R sequence. It has a calculated molecular weight of 16,604 (including amino terminal acetate) and, except for its moderately high content of glutamic acid (15 residues), proline (13 residues), and methionine (7 residues), is otherwise unremarkable in amino acid composition. Although we have not confirmed the actual carboxy terminus of either

Other E1B-155R and E1B-495R peptides have been identified by sequence analysis (data not shown). These identifications are summarized in Table 1.

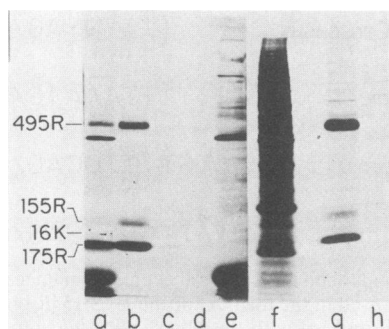


FIG. 9. Immunoprecipitation of E1B protein from a cell-free translation mixture and from extracts of infected cells by using antiserum against the E1B-495R amino terminal nonapeptide. The peptide acetyl-Met-Glu-Arg-Arg-Asn-Pro-Ser-Glu-Arg-(Cys-Lys-BSA) was used to raise rabbit antiserum, which was subsequently affinity purified as described in the text. For the experiments whose results are shown in lanes a through d, cell-free translation was programmed by E1B mRNA as described in the text and the legend to Fig. 1. To lane a was applied 1  $\mu$ l of the total translation mix. Lanes b through d show the immunoprecipitates obtained from 40  $\mu$ l of translation mix, using 20  $\mu$ g of affinity-purified antipeptide serum (b), the same amount of antipeptide serum preblocked with 35  $\mu$ g of the peptide (c), and 10  $\mu$ l of preimmune serum (d). No RNA was added to the 5  $\mu$ l of translation mix analyzed in lane e. Immunoprecipitates of extracts of Ad2-infected HeLa cells are shown in lanes g and h. The HeLa cell extracts were precleared with preimmune serum before incubation with affinity-purified antipeptide serum. Extract (380  $\mu$ l) was precipitated with 20  $\mu$ g of antiserum (g) or antiserum preblocked with 20  $\mu$ g of peptide (h). One microliter of extract before immunoprecipitation is analyzed in lane f. The peptide used to block the antiserum in lane c was not the peptide used for immunization but is the similar peptide acetyl-Met-Glu-Arg-Arg-Asn-Pro-Ser-Glu-Arg-Gly-Tyr, which is the E1B-495R amino terminal decapeptide with a Tyr at the C-terminus to permit, in other experiments, linkage to carrier protein via diazotization. Immunoprecipitates were analyzed on a 15% SDS-polyacrylamide gel, which was exposed for fluorography for 8 days (lanes a through e) or 3.5 days (lanes f through h).

protein, we have identified a methionine residue in a tryptic peptide only 29 residues from the predicted carboxy terminus. Molecular weight estimates of both proteins by gel electrophoresis are consistent with their predicted sequences.

The acceptor site for the splice that creates the E1B-155R protein mRNA was deduced from a partial sequence of the peptide that spans the splice junction. The 155R splice junction peptide fortuitously eluted at the same position as an internal 495R methionine-containing peptide (t41) and would not have been recognized without partial sequence analysis. The proposed acceptor site sequence at 8.96 m.u. (N3270) is compared in Fig. 8 with previously determined E1A and E1B acceptor sites (40, 41) and with a consensus of known splice acceptor sequences (38). The 8.96-m.u. site conforms to the terminal AG dinucleotide rule but fits the remainder of the consensus sequence less well than do other Ad2 early region acceptor sequences. Perhaps the 8.96-m.u. acceptor site is used so rarely that the E1B-155R protein mRNA is present in only small amounts, thus explaining why this structure has not been previously reported. Molecules with structures similar to those expected for the E1B-155R mRNA have occasionally been observed by electron microscopy (L. Chow, personal communication). An E1B mRNA qualitatively similar in structure to the E1B-155R mRNA,

i.e., having a second exon that would permit synthesis of a protein comprised of the amino- and carboxy terminal sequences of the large E1B protein, has been observed among Ad12 mRNAs (52). The Ad12 acceptor lies 102 nucleotides farther downstream than does the Ad2 8.96-m.u. site.

Formally, the Ad2 E1B-495R and 155R proteins bear the same relationship to each other as do the E1A-289R and 243R proteins; each pair shares amino- and carboxy terminal sequences. In the case of the E1A proteins, Montell et al. (37) used site-specific mutagenesis to show that the 289R protein was sufficient for virus growth in culture. However, recent genetic studies have shown that both the E1A-289R and 243R proteins, in addition to the E1B region products, are required for full transformation of rat embryo cultures (C. Montell, E. F. Fisher, M. H. Caruthers, and A. J. Berk, personal communication). Experiments similar to those performed by Montell et al. (37) will be required to determine whether the E1B-155R protein has a function independent of the 495R protein or can itself provide the transformation function currently ascribed to the 495R protein.

The E1B transcription region differs from the E1A region in at least one interesting aspect. All known E1B transcripts begin at N1699 (counting from the left end of the genome; Fig. 1). The first AUG initiation sequence, that for the E1B-175R protein, occurs at N1711. Thus the AUG initiation sequence for the E1B-495R, 155R, and 16K proteins is actually the second AUG from the 5' end of all E1B mRNAs. Montell and colleagues (personal communication) have recently shown that an Ad5 mutant that is defective in the splice donor site at N2249 makes both E1B-175R and E1B-495R proteins. We have shown in this report that the second AUG (at N2016) is, in fact, the initiation site for the E1B-495R protein. Thus, the 22S E1B mRNA must be competent to initiate protein synthesis at both sites. Both AUG sequences belong to an infrequently used class of initiation sites (26). Although both have the favored G in the +4 position with respect to the A of the AUG, the N1711 site has a C in the -3 position, whereas the N2016 site has a T in the -3 position.

At the present time, we do not know whether the E1B-495R and E1B-155R proteins are present at the same intercellular locations in productively infected or in transformed cells. The immunoprecipitation experiment presented in Fig. 9, using antibody raised against the amino terminal nonapeptide, suggests that either or both of these proteins may form a complex with the E1B-175R protein. Alternatively, yet another E1B-495R-related product, which we would not have detected by amino terminal sequence analysis, is masked by migration with the 175R product. Further experiments with other specific antipeptide antisera or additional peptide and protein sequence analysis will be required to distinguish between these possibilities.

After we had submitted this manuscript for publication, we learned that U. Pettersson and A. Virtanen (personal communication) had discovered a new Ad2 E1B mRNA containing a splice between the donor site at N2249 and an acceptor site near N3265, in agreement with the results presented here.

#### ACKNOWLEDGMENTS

We thank Medora Hardy, Richard Feldman, Sherri Osborn, and Ruth LaFlamme for expert technical assistance. The 2A6 monoclonal antibody directed against the E1B-495R protein (47) was generously provided by P. Sarnow.

This research was supported by Public Health Service grant CA29600 from the National Cancer Institute to J.B.L., a core grant



from the National Cancer Institute to the Fred Hutchinson Cancer Research Center, and the U.S. Department of Energy.

#### LITERATURE CITED

- Aleström, P., G. Aküsjarvi, M. Perricaudet, M. B. Mathews, D. F. Klessig, and U. Pettersson. 1980. The gene for polypeptide IX of adenovirus type 2 and its unspliced messenger RNA. *Cell* 19:671-681.
- Anderson, C. W. 1982. Partial sequence determination of metabolically labeled radioactive proteins and peptides, p. 147-167. *In* J. K. Setlow and A. Hollaender (ed.), *Genetic engineering: principles and methods*, vol. 4. Plenum Publishing Corp., New York.
- Anderson, C. W., P. R. Baum, and R. F. Gesteland. 1973. Processing of adenovirus 2-induced proteins. *J. Virol.* 12:241-252.
- Anderson, C. W., and J. B. Lewis. 1980. Aminoterminal sequence of adenovirus type 2 proteins: hexon, fiber, component IX, and early protein 1B-15K. *Virology* 104:27-41.
- Bernards, R., A. Houweling, P. I. Schrier, J. L. Bos, and A. J. van der Eb. 1982. Characterization of cells transformed by Ad5/Ad12 hybrid early region 1 plasmids. *Virology* 120:422-432.
- Bernards, R., P. I. Schrier, J. L. Bos, and A. J. van der Eb. 1983. Role of adenovirus types 5 and 12 early region 1b tumor antigens in oncogenic transformation. *Virology* 127:45-53.
- Bos, J., L. J. Polder, R. Bernards, P. I. Schrier, P. J. van den Elsen, A. J. van der Eb, and H. van Ormondt. 1981. The 2.2 kb E1b mRNA of human Ad12 and Ad5 codes for two tumor antigens starting at different AUG triplets. *Cell* 27:121-131.
- Brackmann, K. H., M. Green, W. S. M. Wold, M. Cartas, T. Matsuo, and S. Hashimoto. 1980. Identification and peptide mapping of human adenovirus type 2-induced early polypeptides isolated by two-dimensional gel electrophoresis and immunoprecipitation. *J. Biol. Chem.* 255:6772-6779.
- Chinnadurai, G. 1983. Adenovirus 2 1p<sup>+</sup> locus codes for a 19 kd tumor antigen that plays an essential role in cell transformation. *Cell* 33:759-766.
- Esche, H., M. B. Mathews, and J. B. Lewis. 1980. Proteins and messenger RNAs of the transforming region of wild-type and mutant adenoviruses. *J. Mol. Biol.* 142:399-417.
- Ey, P. L., S. J. Prowse, and C. R. Jenkin. 1978. Isolation of pure IgG<sub>1</sub>, IgG<sub>2a</sub> and IgG<sub>2b</sub> immunoglobulins from mouse serum using protein A-Sepharose. *Immunochemistry* 15:429-436.
- Gallimore, P. H., P. A. Sharp, and J. Sambrook. 1974. Viral DNA in transformed cells. II. A study of the sequences of adenovirus 2 DNA in nine lines of transformed rat cells using specific fragments of the viral genome. *J. Mol. Biol.* 89:49-72.
- Gentry, L. E., L. R. Rohrschneider, J. E. Casnellie, and E. G. Krebs. 1983. Antibodies to a defined region of *pp60<sup>src</sup>* neutralizes d tyrosine specific kinase activity. *J. Biol. Chem.* 258:11219-11228.
- Gilead, Z., Y. Jeng, W. S. M. Wold, K. Sugawara, H. M. Rho, M. L. Harter, and M. Green. 1976. Immunological identification of two adenovirus 2-induced early proteins possibly involved in cell transformation. *Nature (London)* 264:263-266.
- Gingeras, T. R., D. Sciaky, R. E. Gelinis, J. Bing-Dong, C. E. Yen, M. M. Kelly, P. A. Bullock, B. L. Parsons, K. E. O'Neill, and R. J. Roberts. 1982. Nucleotide sequences from the adenovirus-2 genome. *J. Biol. Chem.* 257:13475-13491.
- Goodfriend, T. L., L. Levine, and G. D. Fasman. 1964. Antibodies to bradykinin and angiotensin: a use of carbodiimides in immunology. *Science* 144:1344-1346.
- Graham, F. L., T. Harrison, and J. Williams. 1978. Defective transforming capacity of adenovirus type 5 host range mutants. *Virology* 86:10-21.
- Green, M., K. H. Brackmann, M. A. Cartas, and T. Matsuo. 1982. Identification and purification of a protein encoded by the human adenovirus type 2 transforming region. *J. Virol.* 42:30-41.
- Halbert, D. N., D. J. Spector, and H. J. Raskas. 1979. In vitro translation products specified by the transforming region of adenovirus type 2. *J. Virol.* 31:621-629.
- Harrison, T., F. Graham, and J. Williams. 1977. Host range mutants of adenovirus type 5 defective for growth in HeLa cells. *Virology* 77:319-329.
- Harter, M. L., and J. B. Lewis. 1978. Adenovirus type 2 early proteins synthesized in vitro and in vivo: identification in infected cells of the 38,000- to 50,000-molecular-weight protein encoded by the left end of the adenovirus type 2 genome. *J. Virol.* 26:736-749.
- Ho, Y.-S., R. Galos, and J. Williams. 1982. Isolation of type 5 adenovirus mutants with a cold-sensitive host range phenotype: genetic evidence of an adenovirus transformation maintenance function. *Virology* 122:109-124.
- Houweling, A., P. J. van den Elsen, and A. J. van der Eb. 1980. Partial transformation of primary rat cells by the leftmost 4.5% fragment of adenovirus 5 DNA. *Virology* 105:537-550.
- Jones, N., and T. Shenk. 1979. Isolation of adenovirus type 5 host range deletion mutants defective for transformation in rat embryo cells. *Cell* 17:583-589.
- Kessler, S. W. 1975. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody adsorbent: parameters of the interaction of antibody-antigen complexes with protein A. *J. Immunol.* 115:1617-1624.
- Kozak, M. 1983. Comparison of initiation of protein synthesis in prokaryotes, eucaryotes, and organelles. *Microbiol. Rev.* 47:1-45.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
- Lassam, N. J., S. T. Bayley, and F. L. Graham. 1979. Tumor antigens of human Ad5 in transformed cells and in cells infected with transformation defective host-range mutants. *Cell* 18:781-791.
- Laver, W. G., J. R. Suriano, and M. Green. 1967. Adenovirus proteins. II. N-Terminal amino acid analysis. *J. Virol.* 1:723-728.
- Levinson, A., and A. J. Levine. 1977. The isolation and identification of the adenovirus group C tumor antigens. *Virology* 76:1-11.
- Lewis, J. B., and C. W. Anderson. 1983. Proteins encoded near the adenovirus late messenger RNA leader segments. *Virology* 127:112-123.
- Lewis, J. B., J. F. Atkins, P. R. Baum, R. Salom, R. F. Gesteland, and C. W. Anderson. 1976. Location and identification of the genes for adenovirus type 2 early polypeptides. *Cell* 7:141-151.
- Lewis, J. B., and M. B. Mathews. 1980. Control of adenovirus early gene expression: a class of immediate early products. *Cell* 21:303-313.
- Lewis, J. B., and M. B. Mathews. 1981. Viral messenger RNAs in six lines of adenovirus-transformed cells. *Virology* 115:345-360.
- Malinowski, K., and W. Manski. 1981. Microprocedures for quantitative immunochemical analysis of antigenic molecules and antigenic determinants, p. 418-436. *In* J. J. Langone and H. Van Vunakis (ed.), *Methods in enzymology*, vol. 73. Academic Press, Inc., New York.
- Matsuo, T., W. S. M. Wold, S. Hashimoto, A. Rankin, J. Symington, and M. Green. 1982. Polypeptides encoded by transforming region E1b of human adenovirus 2: immunoprecipitation from transformed and infected cells and cell-free translation of E1b-specific mRNA. *Virology* 118:456-465.
- Montell, C., E. F. Fisher, M. H. Caruthers, and A. J. Berk. 1982. Resolving the functions of overlapping genes by site-specific mutagenesis at a mRNA splice site. *Nature (London)* 295:380-384.
- Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10:459-472.
- Oosterom-Dragon, E. A., and C. W. Anderson. 1983. Polypeptide structure and encoding location of the adenovirus serotype 2 late, nonstructural 33K protein. *J. Virol.* 45:251-263.
- Perricaudet, M., G. Aküsjarvi, A. Virtanen, and U. Pettersson. 1979. Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. *Nature (London)* 281:694-696.

41. Perricaudet, M., J. M. Le Moullec, and U. Pettersson. 1980. Predicted structure of two adenovirus tumor antigens. *Proc. Natl. Acad. Sci. U.S.A.* **77**:3778-3782.
42. Persson, H., H. Jornvall, and J. Zabielski. 1980. Multiple mRNA species for the precursor to an adenovirus-encoded glycoprotein: identification and structure of the signal sequence. *Proc. Natl. Acad. Sci. U.S.A.* **77**:6349-6353.
43. Persson, H., M. G. Katze, and L. Philipson. 1982. Purification of a native membrane-associated adenovirus tumor antigen. *J. Virol.* **42**:905-917.
44. Ross, S. R., S. J. Flint, and A. J. Levine. 1980. Identification of the adenovirus early proteins and their genomic map positions. *Virology* **100**:419-432.
45. Rowe, D. T., and F. L. Graham. 1983. Transformation of rodent cells by DNA extracted from transformation-defective adenovirus mutants. *J. Virol.* **46**:1039-1044.
46. Sarnow, P., P. Hearing, C. W. Anderson, N. Reich, and A. Levine. 1982. Identification and characterization of an immunologically conserved adenovirus early region 11,000 M<sub>r</sub> protein and its association with the nuclear matrix. *J. Mol. Biol.* **162**:565-583.
47. Sarnow, P., C. A. Sullivan, and A. J. Levine. 1982. A monoclonal antibody detecting the adenovirus type 5 E1b-58Kd tumor antigen: characterization of the E1b-58Kd tumor antigen in adenovirus-infected and -transformed cells. *Virology* **120**:510-517.
48. Schrier, P. I., P. J. van den Elsen, J. J. L. Hertoghs, and A. J. van der Eb. 1979. Characterization of tumor antigens in cells transformed by fragments of adenovirus type 5 DNA. *Virology* **99**:372-385.
49. Shiroki, K., K. Maruyama, I. Saito, Y. Fukui, and H. Shimojo. 1981. Incomplete transformation of rat cells by a deletion mutant of adenovirus type 5. *J. Virol.* **38**:1048-1054.
50. Smart, J. E., J. B. Lewis, M. B. Mathews, M. L. Harter, and C. W. Anderson. 1981. Adenovirus type 2 early proteins: assignment of the early region 1A proteins synthesized *in vivo* and *in vitro* to specific mRNAs. *Virology* **112**:703-713.
51. van den Elsen, P., S. de Pater, A. Houweling, J. van der Veer, and A. van der Eb. 1982. The relationship between region E1a and E1b of human adenoviruses in cell transformation. *Gene* **18**:175-185.
52. Virtanen, A., U. Pettersson, J. M. Le Moullec, P. Tiollais, and M. Perricaudet. 1982. Different mRNAs from the transforming region of highly oncogenic and non-oncogenic human adenoviruses. *Nature (London)* **295**:705-707.
53. Wold, W. S. M., and M. Green. 1979. Adenovirus type 2 early polypeptides immunoprecipitated by antisera to five lines of adenovirus-transformed rat cells. *J. Virol.* **30**:297-310.