JOURNAL OF VIROLOGY, Nov. 1984, p. 307–313
0022-538X/84/110307-07$02.00/0

Vol. 52, No. 2

Mapping a New Gene That Encodes an 11,600-Molecular-Weight Protein in the E3 Transcription Unit of Adenovirus 2

WILLIAM S. M. WOLD,* CHRISTOS CLADARAS, SANDRA C. MAGIE, AND NUHA YACOUB
Institute for Molecular Virology, St. Louis University School of Medicine, St. Louis, Missouri 63110

Received 30 April 1984/Accepted 10 July 1984

The DNA sequence of the early E3 transcription unit of adenovirus 2 (Ad2) (J. Hérissé et al., Nucleic Acids Res. 8:2173–2192, 1980), indicates that an open reading frame exists between nucleotides 1860 and 2163 that could encode a protein of M_r 11,600 (11.6K). We have determined the DNA sequence of the corresponding region in Ad5 (closely related to Ad2) and have established that this putative gene is conserved in Ad5 (a 10.5K protein). To determine whether this protein is expressed, we prepared an antiserum in rabbits against a synthetic peptide corresponding to amino acids 66 to 74 in the 11.6K protein of Ad2. The peptide antiserum immunoprecipitated a ca. 13K-14K protein doublet, as estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, from [^{35}S]methionine-labeled Ad2- or Ad5-early-infected KB cells. The antiserum also immunoprecipitated a 13K-14K protein doublet translated in vitro from Ad2 or Ad5 early E3-specific mRNA purified by hybridization to Ad2 *Eco*RI-D (nucleotides −236 to 2437). The synthetic peptide successfully competed with the 13K-14K protein doublet in immunoprecipitation experiments, thereby confirming the specificity of the antiserum. As deduced from the DNA sequence, the 11.6K protein (and the corresponding 10.5K Ad5 protein) has a conserved 22-amino-acid hydrophobic domain, suggesting that the protein may be associated with membranes. We conclude that a gene located at nucleotides 1860 to 2143 in the Ad2 E3 transcription unit (nucleotides 1924 to 2203) in the Ad5 E3 transcription unit) encodes an 11.6K protein (10.5K in Ad5).

We are interested in the E3 region of group C adenoviruses (e.g., adenoviruses 2 [Ad2] and 5 [Ad5]) as a model for understanding gene organization and expression in a eucaryotic complex transcription unit (12). E3 is located at map positions (mp) 76 to 86 on the adenovirus genome and is transcribed off the *r*-strand (8, 11, 15, 16, 29). A schematic representation of the E3 transcription unit of Ad2 is presented in Fig. 1. A single promoter (4, 7, 13, 32) controls transcription of ca. 10 overlapping mRNAs which form two major families (referred to herein as E3A and E3B) of 3′ coterminal mRNAs (8, 11, 20, 29, 36). The mRNAs in each family have different spliced structures. Nucleotide (nt) sequences have been identified that encode the 5′ ends, the E3A and E3B 3′ ends, and the major 5′ and 3′ splice sites of the E3 mRNAs of Ad5 (C. Cladaras, B. Bhat, and W. S. M. Wold, manuscript in preparation) and Ad2 (1, 2, 4, 18, 32, 44).

The E3 DNA sequence has been determined for both Ad2 (23, 24) and Ad5 (C. Cladaras and W. S. M. Wold, manuscript in preparation). The Ad2 E3 transcription unit consists of 3,255 base pairs (3,308 in Ad5) between nt +1, the transcription initiation site, and nt 3255, the major E3B 3′ end site (44). (nt +1 corresponds to nt 236 in the DNA sequence of the Ad2 *Eco*RI D fragment as described by Hérissé et al. [23].) The major 3′ end site for the E3A mRNAs of Ad2 is at nt 2183. Four open reading frames (ORFs) that contain an ATG and that could encode a protein of M_r 6,000 (6K) or larger are conserved in each of the E3A and E3B regions of Ad2 and Ad5. However, only one E3 gene has been identified with certainty, that for a well characterized glycoprotein, E3A-gp19K (25, 30, 35, 36, 38). This protein was initially mapped to E3 by cell-free translation of *r*-strand and E3-specific mRNA (35, 36), by analysis of E3 deletion mutants (27, 35, 41, 45), and by immunoprecipitation with antisera directed against adenovirus-transformed cells that express E3 (26, 38, 40, 46). The N-terminus of E3A-gp19K of Ad2 has been mapped to ATG_{1204} in the Ad2 E3 transcription unit by N-terminal sequence analysis (36). Therefore, E3A-gp19K would be coded by the nt 1204 to 1681 ORF in Ad2 (23). The N-terminal sequence of mature E3A-gp19K (i.e, after cleavage of the N-terminal signal sequence) of Ad5 has been mapped to nt 1219 (W. S. M. Wold, C. Cladaras, S. L. Deutscher, and Q. Kapoor, manuscript in preparation), indicating that the Ad5 version of E3A-gp19K is coded at nt 1168 to 1648 in the E3 transcription unit of Ad5. A 14K protein has been mapped to the E3B region of Ad2 by translation of mRNA specific to *Eco*RI-E (mp 83.4 to 89.7) and by analysis of E3 deletion mutants (37). Proteins of 15.5K, 14K, and 13K have been translated from Ad2 E3-specific mRNA (22); whether any of these correspond to the E3B-specific 14K or to E3A-gp19K is not known.

As an approach to identifying E3 genes, we prepared synthetic peptides corresponding to putative proteins coded by ORFs in E3 of Ad2 or Ad5, or both. Here we report that nt 1860 to 2163 in Ad2 (nt 1924 to 2203 in Ad5) encode an 11.6K protein (10.5K in Ad5).

MATERIALS AND METHODS

Materials. Divinylbenzene resin (1% cross-linked) substituted with 0.56 meq of N^{α}-*t*-butyloxycarbonyl(Boc)-*S*-*p*-methoxybenzyl-L-cysteine per g was purchased from Peninsula Laboratories, Inc. Boc-L-amino acids were purchased from the following: isoleucine, N^{G}-tosyl arginine, and *O*-benzyl tyrosine, Chemical Dynamics; alanine and proline, Sigma Chemical Co. Dichloromethane, trifluoroacetic acid, *N*,*N*′-dimethylformamide, and tertiary butanol were purchased from Mallinckrodt, Inc. Anisole and triethylamine were from Fisher Scientific Co., *N*,*N*′-dicyclohexylcarbodiimide was from Sigma, *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester was from Pierce Chemical Co., hydrogen

* Corresponding author.

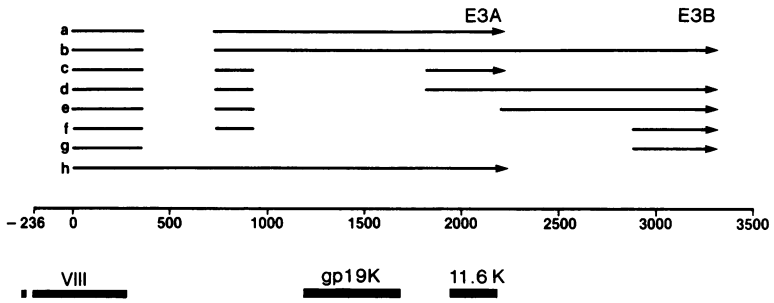

FIG. 1. Schematic representation of the E3 transcription unit of Ad2. nt +1 indicates the transcription initiation site (4, 7, 13, 32) and nt −236 indicates the *Eco*RI site at mp 79.6 (23). The arrows at the top represent spliced mRNA molecules as determined by electron microscopy (11, 29), nuclease gel mapping (8), and DNA sequence analysis of cDNA clones (1, 2, 44). The black bars at the bottom indicate the coding sequences for E3A-gp19K, E3A-11.6K (see text), and the C-terminus of virion protein VIII.

fluoride was from Union Carbide Corp. (Linde Division), and keyhole limpet hemocyanin (KLH) was from Calbiochem-Behring.

Peptide synthesis, peptide coupling to KLH, immunizations, and ELISA. Procedures for solid-phase peptide synthesis (6), peptide coupling to KLH (21, 33), rabbit immunizations, and enzyme-linked immunosorbent assay (ELISA) were essentially as described in detail by Green et al. (19). Briefly, synthesis of peptide 1 (see below) was initiated on 0.75 g of Boc-*S*-*p*-methoxybenzyl-L-cysteine esterified to divinylbenzene resin (0.4 mmol of cysteine) in a 25-ml reaction vessel equipped with a medium porous fritted glass filter for evacuation of reactants and solvents. Agitation was in a Burrell wrist action shaker in a reaction or solvent volume of 15 ml. For each coupling, the previous reactant amino acid was washed out, the resin was shrunk and then expanded by washes in 95% tertiary butanol and then in dichloromethane, the Boc group was removed from the amino group with 25% trifluoroacetic acid in dichloromethane, the resin was neutralized with 10% triethylamine in dichloromethane, and the next Boc-amino acid was coupled in a fivefold molar excess of amino acid and *N*,*N*′-dicyclohexylcarbodiimide overnight in 50% *N*,*N*′-dimethylformamide in dichloromethane. The efficiency of coupling was assayed by the ninhydrin reaction (43). Anhydrous HF with *ca.* 5% anisole was used at 0°C to cleave the completed peptide from the resin and to remove the protective side groups from the amino acids. After removal of HF under vacuum and removal of anisole by ether extraction, the peptide was solubilized in 50% acetic acid and lyophilized. The yield for peptide 1 was 42%. The peptide was soluble at pH 4 to 9. Amino acid analysis gave the expected molar ratios of amino acids.

Peptide 1 (8 mg) was coupled to 8 mg of KLH with *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (21, 33). Two 4-lb (ca. 1.8-kg) New Zealand rabbits were inoculated subcutaneously at several sites in the neck with peptide-KLH conjugate (1.3 mg of peptide) in complete Freund adjuvant. This was repeated 1 week later with incomplete Freund adjuvant; 1 week later, each rabbit received an intraperitoneal inoculation of conjugate (1.3 mg of peptide) in alum (aluminum hydroxide gel). The fourth and subsequent inoculations were every 3 weeks by the intraperitoneal route with conjugate (0.7 mg of peptide) in alum. Rabbits were bled from the ear ca. 10 days after the fourth and subsequent inoculations.

Sera were analyzed for peptide 1 antibodies by ELISA, essentially as described (19); 5 to 500 pmol of peptide 1 was adsorbed to 96-well Immulon 2 plates (Dynatech Laboratories, Inc.) and reacted with rabbit sera, alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin G (Sigma), and, finally, alkaline phosphatase substrate (*p*-nitrophenyl phosphate; Sigma).

Preparation of [^{35}S]Met-labeled cell extracts and immunoprecipitation analysis. Procedures for the preparation and immunoprecipitation analysis of cell extracts labeled with [^{35}S]methionine (Met) have been described previously (10, 46). Briefly, suspension cultures of KB cells in minimal essential medium were infected with 250 PFU of Ad2 or Ad5 per cell, or were mock infected. Cycloheximide was added to 25 μg/ml at 3 h postinfection. Cells were labeled at 7×10^5 cells per ml from 6 to 11 h postinfection in Met-free minimal essential medium containing 5% horse serum, 20 μg of 1-β-D-arabinofuranosylcytosine per ml, and 20 μCi of [^{35}S]Met (1,058 Ci/mmol; New England Nuclear Corp.) per ml. Cells were collected, rinsed twice in cold phosphate-buffered saline, and lysed on ice with 10 times the pellet volume of iso-high-pH buffer (0.14 M NaCl, 1 mM $MgCl_2$, 10 mM Tris-chloride, pH 8.5) containing 0.5% Nonidet P-40 and 1 mM phenylmethylsulfonyl fluoride. Nuclei were removed, and the cytoplasm was clarified by centrifugation at 35,000 rpm for 1 h at 4°C in a Ti50 rotor; portions were stored at −85°C.

Extract portions (3 to 5×10^6 cpm) were analyzed by immunoprecipitation with 5 to 10 μl of peptide antiserum and Formalin-fixed *Staphylococcus aureus* containing protein A (10). For peptide competition experiments, 0.1, 1, or 10 μg of peptide was included in the immunoprecipitation reaction (volume, 160 μl). Immunoprecipitates were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 15% polyacrylamide gels (0.75 mm by 15 cm; acrylamide:*N*,*N*′-methylene-bisacrylamide, 29.8:0:2), and the gels were fluorographed (9).

Preparation of RNA extracts and cell-free translation. Suspension cultures of KB cells were infected with 200 PFU of Ad2 or Ad5 per cell (47). Cycloheximide was added to 25 μg/ml at 3 h postinfection, and at 13 h postinfection cells were collected, rinsed, and lysed (47). After removal of nuclei, cytoplasmic RNA was extracted (14), and polyadenylated RNA was purified by two passages through oligodeoxythymidylate-cellulose (Collaborative Research, Inc.) (3, 5).

E3-specific mRNA was purified by hybridization to Ad2 *Eco*RI-D (cloned at the *Eco*RI site in plasmid pBR322) immobilized on nitrocellulose filters (39; Jim Lewis, personal communication) by using 10 μg of RNA per 5 μg of DNA per filter.

E3-specific mRNA purified from four filters was translated in a 25-μl volume in the rabbit reticulocyte system (New

England Nuclear) by using procedures suggested by the manufacturer. A 1-μl portion was removed for SDS-PAGE analysis, and the remainder was assayed by immunoprecipitation with the peptide 1 antiserum by to the following procedure (Margarita H. Cladaras, personal communication). The translation mixture was mixed with an equal volume of 10% SDS, heated for 2 min at 100°C, and then diluted with 10 volumes of immunoprecipitation buffer containing 40 mM Met. The mixture was preabsorbed for 1 h at 20°C with *S. aureus* (final concentration, 1%) and immunoprecipitated as described previously (10). For the competition experiment, 0.1 μg of peptide 1 was included in the immunoprecipitation reaction (volume, 530 μl). Before and after immunoprecipitation, proteins were electrophoresed on an SDS 15% polyacrylamide gel (0.75 mm by 30 cm), and the gel was fluorographed.

DNA sequencing. The Ad5 *Hin*dIII B fragment (mp 73.6 to 89.1), cloned in plasmid pBR322, was used as the source of DNA for sequencing. The *Bgl*II-*Eco*RI fragment, mp 78.0 to 83.4, nt 566 to 2482 in the Ad5 E3 transcription unit, was purified on a 4% polyacrylamide gel. This fragment was subcleaved with appropriate restriction enzymes, and the resulting fragments were labeled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase (34). The individual strands of these fragments were purified on 5% polyacrylamide gels, and were sequenced by the Maxam-Gilbert procedure (34). The regions sequenced on the *l*-strand included (i) nt 1793 (*Hin*fI site) to 2100, (ii) nt 2098 (*Fnu*DII site) to beyond 2250, and (iii) nt 2224 (*Xho*I site) to beyond 2250. Regions sequenced on the *r*-strand included (i) nt 2224 (*Xho*I site) to 2000, and (ii) nt 1879 (*Fnu*DII site) to beyond 1800. Other fragments in these regions were also sequenced. Both strands of the region shown in Fig. 5 were sequenced at least once with unambiguous results. Sequence analyses were performed across all restriction sites. Computer programs (31) were used to record and analyze the sequence data.

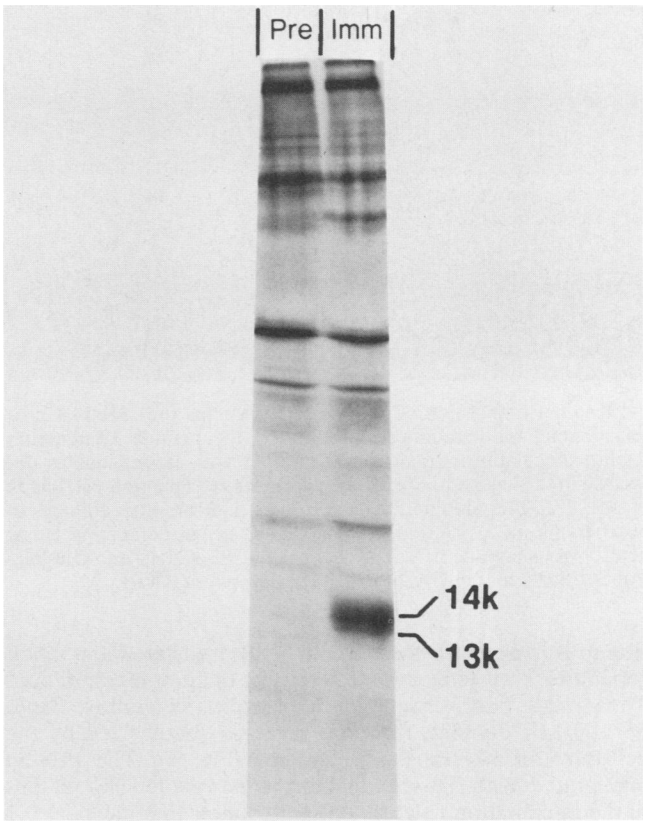

FIG. 2. Immunoprecipitation by peptide 1 antiserum of a 13K-14K protein doublet from Ad2-early-infected cells. The proteins were labeled with [^{35}S]Met, immunoprecipitated with antiserum to peptide 1 (Imm, eighth inoculation), electrophoresed on an SDS 15% polyacrylamide gel, and fluorographed. Pre, Preimmune serum.

RESULTS

Induction of antibodies to peptide 1. Two rabbits were immunized with peptide 1 (NH_2-Arg-Ala-Arg-Pro-Pro-Ile-Tyr-Arg-Pro-COOH) coupled to KLH through a C-terminal cysteine (the cysteine is not shown). Peptide 1 corresponds to amino acids 66 to 74 in the putative Ad2 protein of 101 amino acids. The rabbits were bled after the fourth and most of the subsequent inoculations, and the sera were examined for peptide-specific antibodies by ELISA. A strong reaction was observed in sera from both rabbits after the fourth inoculation. The titers peaked after the fifth inoculation (9 weeks after the initial inoculation) and then gradually declined with each successive inoculation (data not shown). Both rabbits gave about the same response. Preimmune sera from each rabbit, as well as sera from rabbits immunized with other peptides, were negative in the ELISA.

Immunoprecipitation of a 13K-14K protein doublet from Ad2-infected cells. The sera from each blood sample were assayed for their ability to immunoprecipitate a protein from Ad2 early-infected cell extracts labeled with [^{35}S]Met. The immune serum from the eighth inoculation immunoprecipitated what appeared to be two proteins of ca. 13K and 14K (or one protein that gave a diffuse protein band) (Fig. 2). The preimmune serum did not immunoprecipitate those proteins (Fig. 2), nor did immune sera prepared against unrelated synthetic peptides coupled to KLH (not shown). A protein(s) of about the same apparent M_r was also immunoprecipitated from Ad5-early-infected cells (not shown). However, the antisera were much less potent against the Ad5 protein; this is not surprising because the Ad2 and Ad5 sequences have a mismatch of one amino acid in the area corresponding to peptide 1 (see Fig. 6). The ability of the sera from both rabbits to immunoprecipitate the Ad2 13K-14K protein doublet increased from the fourth to the eighth inoculation (data not shown), which contrasts with the ELISA results discussed above. Apparently, the antibodies most active in the ELISA are not so effective as other antibodies present in the antisera in immunoprecipitating the native protein.

Competition experiments with peptide 1 were done to ascertain the specificity of the antisera. A 0.1-μg sample (in a volume of 160 μl) of peptide 1 completely prevented the antiserum from immunoprecipitating the 13K-14K protein doublet (Fig. 3). However, competition was not observed with 10 μg of an unrelated peptide, P2 (see legend to Fig. 3 for the P2 sequence). Similar results were obtained with antisera from both rabbits, and for all blood samples analyzed. The diffuse nature of the immunoprecipitated protein(s) is characteristic, and the immunoprecipitate appeared to resolve into two bands (Fig. 3). Unfortunately, a doublet of host bands (Fig. 3) was coprecipitated, which made the visualization of the viral protein(s) more difficult.

Immunoprecipitation of a 13K-14K protein doublet translated in vitro from E3-specific mRNA. To verify that the peptide 1 antisera were indeed immunoprecipitating an E3-coded protein, E3-specific mRNA was purified by hybridiza-

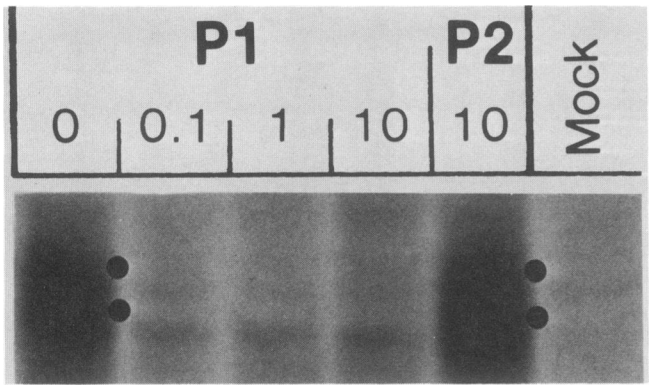

FIG. 3. Competition of peptide 1 (P1) versus the 13K-14K protein doublet for immunoprecipitation by the peptide 1 antiserum (eighth inoculation). Immunoprecipitation was carried out as described in the legend to Fig. 2; the dots mark the apparent two bands of the 13K-14K doublet. The numbers indicate the amount in micrograms (in a reaction volume of 160 μl) of competing P1 or heterologous peptide (P2). The sequence of P2 is NH_2-Asp-Glu-Ser-Val-Ser-Ser-Ser-Leu-Gly-Leu-Arg-Pro-Asp-Cys-COOH.

tion to Ad2 *Eco*RI-D (mp 75.9 to 83.4), translated in a rabbit reticulocyte system, and assayed by immunoprecipitation. The peptide 1 antiserum immunoprecipitated a diffuse band, or 13K-14K doublet, from the proteins synthesized by the cell-free mRNA translation system (Fig. 4). The protein comigrated with E3-specified proteins visible without immunoprecipitation (Fig. 4, E3 RNA lane). Peptide 1, at 0.1 μg in a volume of 530 μl, completely outcompeted the doublet protein. Similar results were observed with E3-specific mRNA from Ad5, although the quantity of doublet immunoprecipitated was much less than that with Ad2 (data not shown).

DNA sequence of the Ad5 gene that corresponds to the Ad2 gene encoding the 13K-14K protein doublet. The results given above clearly indicate that the ORF at nt 1860 to 2163 in the Ad2 transcription unit represents a gene that encodes an 11.6K protein, although this protein has an apparent M_r of 13,000 to 14,000, as measured by SDS-PAGE. It was of interest to sequence the corresponding gene of Ad5 to determine the extent to which the protein sequence was conserved. The Ad5 DNA and protein sequences and their comparisons with those of Ad2 are shown in Fig. 5. The sequence shown includes (i) the 3′ splice site for the Ad5 mRNA that probably specifies this protein, (ii) the protein coding sequences, and (iii) the putative ATTAAA cleavage/polyadenylation signal and the major 3′ end site for the E3A family of mRNAs of Ad5 (Cladaras et al., manuscript in preparation) and Ad2 (1, 2, 44).

A comparison of the hydrophilic and hydrophobic domains of the Ad2 and Ad5 proteins is given in Fig. 6. An interesting feature is a domain of 22 amino acids (amino acids 41 to 62 in Ad2) that is very hydrophobic and is completely conserved in both viruses. This domain is flanked by polar regions, very basic on the C-terminal side and acidic on the N-terminal side. The region between the hydrophobic domain and amino acid 85 in the Ad2 protein (amino acids 63 to 92 in the Ad5 protein) is also quite conserved. The proteins diverge considerably at both the N- and C-termini. Near the N-terminus, seven amino acids are present in the Ad2 protein that are not found in the Ad5 protein. The mismatch in the last eight or nine amino acids at the C-terminus of the Ad2 and Ad5 proteins is caused by an extra nt at position 2180 in Ad5, which changes the reading frame (Fig. 5). We have confirmed the sequence in this region of Ad2 as described by Hérissé et al. (23); therefore, this change in reading frame is genuine.

DISCUSSION

We prepared antisera directed against a 9-amino-acid synthetic peptide whose sequence is predicted by nt 2055 to 2081 in the Ad2 E3 transcription unit. These antisera immunoprecipitated a [^{35}S]Met-labeled ca. 13K-14K protein doublet, as estimated by SDS-PAGE, from Ad2-early-infected KB cells, and from cell-free translates of E3-specific mRNA. We conclude that the protein we have detected is coded by the ORF at nt 1860 to 2163. Based on the predicted amino acid sequence, the true M_r of this protein is 11,600.

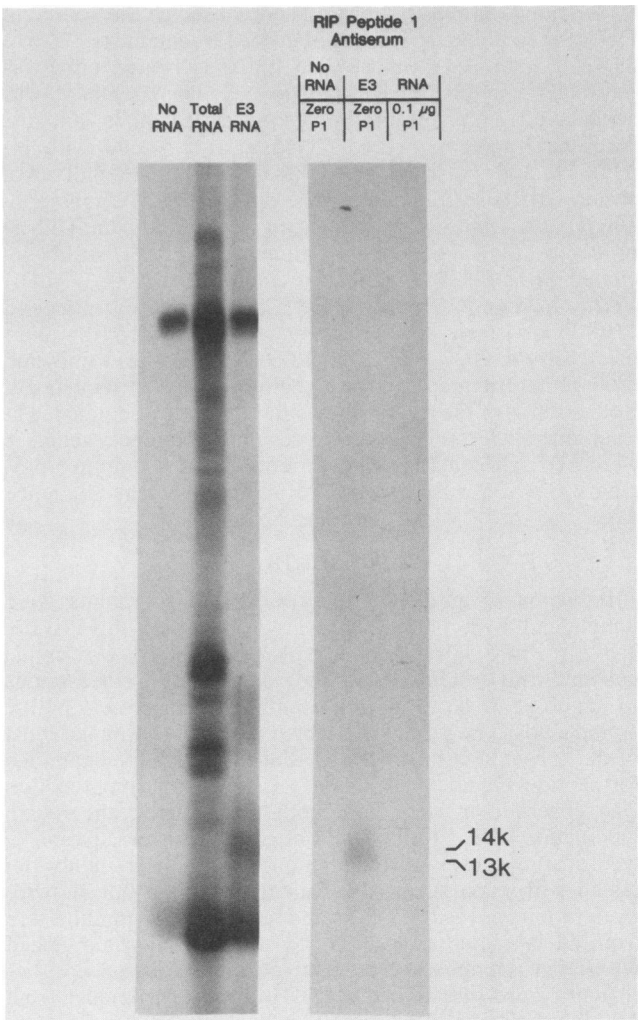

FIG. 4. Cell-free translation of Ad2 E3-specific mRNA and immunoprecipitation of the 13K-14K protein doublet by the peptide 1 antiserum (eighth inoculation). E3-specific mRNA was purified from Ad2-early-infected KB cells by hybridization to cloned Ad2 *Eco*RI-D and translated in the rabbit reticulocyte system. Translated proteins in lanes marked "no RNA" or "E3 RNA" were assayed by immunoprecipitation with peptide 1 antiserum. Peptide 1 (P1), at 0.1 μg in a 530-μl reaction volume, was used to compete with the 13K-14K protein doublet. Proteins were electrophoresed on an SDS–15% polyacrylamide gel, and the gel was fluorographed. RIP, Radioimmunoprecipitation.

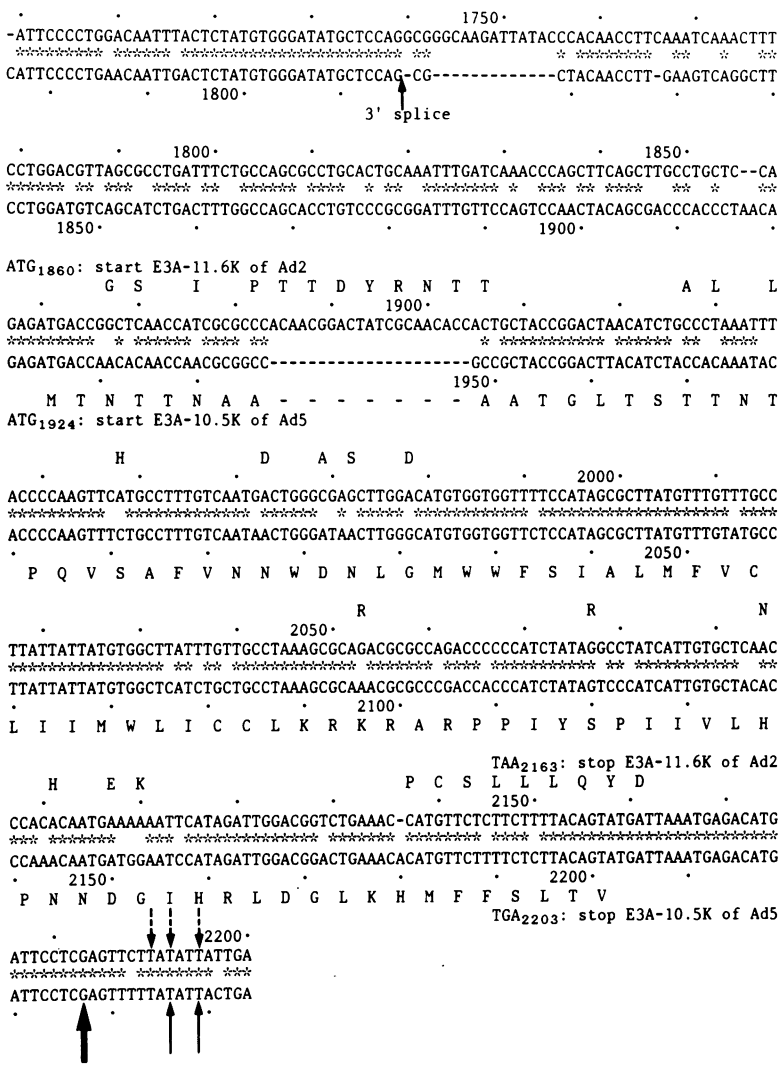

FIG. 5. DNA sequence of the Ad5 E3A-10.5K gene, the predicted amino acid sequence of the Ad5 E3A-10.5K protein, and comparison with the corresponding Ad2 sequence. The Ad2 and Ad5 sequences are given at the top and bottom, respectively. Only amino acids in the Ad2 protein that mismatch the Ad5 protein are presented. The Ad2 DNA sequence was determined by Hérissé et al. (23) and was confirmed at nt ca. 2050 to 2200 in our laboratory. The nt numbers refer to the E3 transcription unit of Ad2 or Ad5, in which nt +1 represents the first transcription initiation site (4; Cladaras et al., in preparation). nt +1 is nt 236 in Ad2 *Eco*RI-D (23). The arrow at nt 1817 in the Ad5 sequence indicates the 3′ splice site for the mRNA that presumably codes the E3A-10.5K protein (see text). Although there is no direct evidence, we presume that a 3′ splice occurs at the corresponding position in the Ad2 sequence. The major (thick solid arrows) and minor (thin solid arrows) 3′ ends of the E3A family of mRNAs of Ad5 were determined by nuclease gel mapping (Cladaras et al., in preparation). The 3′ ends of Ad2 E3A mRNAs (dashed arrows) were determined by DNA sequencing of cDNA clones (1, 2, 44).

We therefore refer to this protein as E3A-11.6K. We have sequenced the equivalent gene in Ad5, located at nt 1924 to 2203 in the E3 transcription unit; the Ad5 gene would encode a 10.5K protein.

The Ad2 11.6K protein appeared to migrate as a single diffuse band or to resolve into two bands in SDS-PAGE. Perhaps the unusual structure of this protein, i.e., a relatively long hydrophobic domain flanked by polar domains, causes the protein to assume alternate conformations with differing mobilities in SDS-PAGE.

We have not proven that Ad2 E3A-11.6K initiates at ATG_{1860} (this would require NH_2-terminal sequencing), but this is almost certain. The next ATG in the ORF is 40 codons downstream, and initiation at that ATG would yield a 6.7K protein; this is much smaller than the 13K-14K protein we observed by SDS-PAGE. In Ad5 there is a 3′ splice at nt 1817 (Cladaras et al., in preparation), and the corresponding splice in Ad2 is probably at nt 1740 (Fig. 1 and 5). This splice probably joins the main body of the E3c and E3d mRNAs of Ad2 (11) to two leaders coded at nt 1 to 372 and 766 to 952 (2). E3A-11.6K is probably coded for by one or both of these mRNAs. If so, these mRNAs contain only one ATG 5′ to ATG_{1860}, i.e., ATG_{291} (23). However, ATG_{291} is followed by an in-phase termination signal 34 codons downstream (2, 23), and so it cannot initiate E3A-11.6K. In Ad5 there is an ATG at nt 1846 which is downstream of the nt 1817 splice and which is in phase with ATG_{1924} (Fig. 5). If ATG_{1846} is used, the Ad5 protein would be 2,200 daltons larger than the Ad2 protein. The Ad5 protein translated in vitro and immunoprecipitated by peptide 1 antiserum appeared to comigrate with

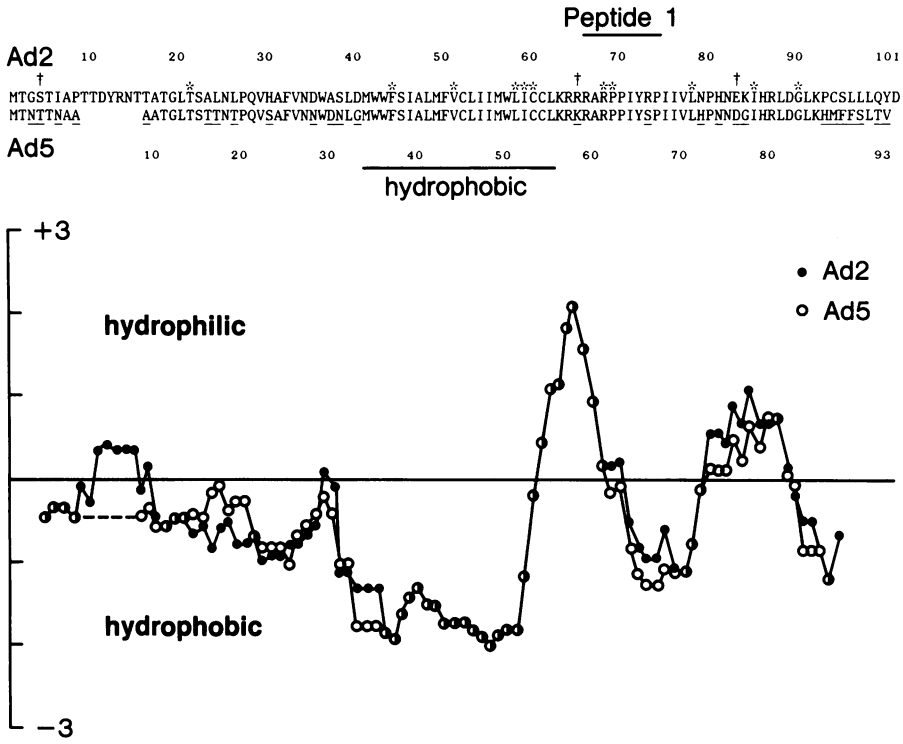

FIG. 6. Comparison of the amino acid sequences and the hydrophobicity-hydrophilicity profiles of the Ad2 E3A-11.6K protein and the Ad5 E3A-10.5K protein. Amino acids that mismatch are underscored. †, Conservative substitution; *, nucleotide differences that do not affect the amino acid coded. The hydrophobic-hydrophilic profiles were generated by a computer program (31); the values represent a running average of seven amino acids. Polar amino acids (R, K, D, E) have values of +3. Values for some of the more hydrophobic amino acids are −3.4 for W, −2.5 for F, −2.3 for Y, and −1.8 for I. The International Union of Biochemistry single-letter amino acid code is used.

the Ad2 protein by SDS-PAGE, so that the Ad5 protein probably initiates mainly at ATG_{1924}.

A comparison of the amino acid sequences and hydrophobic-hydrophilic profiles of the Ad2 and Ad5 proteins reveals some interesting features. The only highly conserved areas are a 22-amino-acid hydrophobic domain (zero substitutions) followed by a 30-amino-acid domain (six substitutions, two of which are conservative) with both polar (basic) and hydrophobic features. This conservation suggests that these domains are important in the function of this protein. In contrast, the proteins diverge considerably at the N- and C-terminal domains, suggesting that these domains are under less evolutionary pressure. A 22-amino-acid domain is of sufficient length to traverse a membrane, and this invites speculation. That is, this protein may be oriented in the membrane with the hydrophobic domain traversing the membrane, and with the N- and C- portions of the protein on either side of the membrane. The protein would be stabilized in the membrane by the polar domains that flank the hydrophobic domain, very basic on the C-terminal side and somewhat acidic on the N-terminal side. This type of structure and orientation would be similar to that of many transmembrane glycoproteins, especially the highly basic domain on the C-terminal side (reviewed in reference 42).

We know nothing about the function of this protein except that it is nonessential for adenovirus multiplication in exponentially growing cultured human cells. That is, this gene is deleted in a number of nondefective adenovirus-simian virus 40 viruses (17, 28) and in viable mutants of Ad2 (unpublished results). Nevertheless, the conservation of this gene between Ad2 and Ad5 argues that it is beneficial to the virus in natural adenovirus infections of humans.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants CA24710 and GM31276 from the National Institutes of Health. W.S.M.W. was supported by Research Career Development award CA00650, and S.C.M. was supported by Postdoctoral Training grant CA09222, both from the National Institutes of Health.

We thank Tammo A. Kramer, Janey S. Symington, Maurice Green, and the staff at Peninsula Laboratories for kindly giving advice on the synthesis of peptides and the generation of peptide antibodies. We thank C. E. Mulhall for typing the manuscript.

LITERATURE CITED

1. **Ahmed, C. M. I., R. Chanda, N. Stow, and B. S. Zain.** 1982. The sequence of 3′-termini of mRNAs from early region III of adenovirus 2. Gene **19**:297–301.
2. **Ahmed, C. M. I., R. S. Chanda, N. D. Stow, and B. S. Zain.** 1982. The nucleotide sequence of mRNA for the M_r 19,000 glycoprotein from early gene block III of adenovirus 2. Gene **20**:337–344.
3. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. U.S.A. **69**:1408–1412.
4. **Baker, C. C., and E. B. Ziff.** 1981. Promoters and heterogeneous 5′ termini of the messenger RNAs of adenovirus serotype 2. J. Mol. Biol. **149**:189–221.
5. **Bantle, J. A., I. H. Maxwell, and W. E. Hahn.** 1976. Specificity of oligo(dT)-cellulose chromatography in the isolation of polyadenylated RNA. Anal. Biochem. **72**:413–427.

6. **Barany, G., and R. B. Merrifield.** 1980. Solid-phase peptide synthesis, p. 1–284. *In* E. Gross and J. Meienhofer (ed.), The peptides, vol. 2. Academic Press, Inc., New York.
7. **Berk, A. J., and P. A. Sharp.** 1977. Ultraviolet mapping of the adenovirus 2 early promoters. Cell **12**:45–55.
8. **Berk, A. J., and P. A. Sharp.** 1978. Structure of the adenovirus 2 early mRNAs. Cell **14**:695–711.
9. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
10. **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.
11. **Chow, L. T., T. R. Broker, and J. B. Lewis.** 1979. Complex splicing patterns of RNAs from the early regions of adenovirus-2. J. Mol. Biol. **134**:265–303.
12. **Darnell, J. E., Jr.** 1982. Variety in the level of gene control in eukaryotic cells. Nature (London) **297**:365–371.
13. **Evans, R. M., N. Fraser, E. Ziff, J. Weber, M. Wilson, and J. E. Darnell.** 1977. The initiation sites for RNA transcription in Ad2 DNA. Cell **12**:733–739.
14. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
15. **Flint, S. J., S. M. Berget, and P. A. Sharp.** 1976. Adenovirus transcription. III. Mapping of viral RNA sequences in cells productively infected by adenovirus type 5. Virology **72**:443–455.
16. **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.
17. **Flint, S. J., Y. Wewerka-Lutz, A. S. Levine, J. Sambrook, and P. A. Sharp.** 1975. Adenovirus transcription. II. RNA sequences complementary to simian virus 40 and adenovirus 2 DNA in $Ad2^{+}ND_1$- and $Ad2^{+}ND_3$-infected cells. J. Virol. **16**:662–673.
18. **Frazer, N. W., C. C. Baker, M. A. Moore, and E. B. Ziff.** 1982. Poly(A) sites of adenovirus serotype 2 transcription units. J. Mol. Biol. **155**:207–233.
19. **Green, M., K. H. Brackmann, L. A. Lucher, J. S. Symington, and T. A. Kramer.** 1983. Human adenovirus 2 E1B-19K and E1B-53K tumor antigens: antipeptide antibodies targeted to the NH_2 and COOH termini. J. Virol. **48**:604–615.
20. **Green, M., W. S. M. Wold, and W. Büttner.** 1981. Integration and transcription of group C human adenovirus sequences in the DNA of five lines of transformed rat cells. J. Mol. Biol. **151**:337–366.
21. **Green, N., H. Alexander, A. Olson, S. Alexander. T. M. Shinnick, J. G. Sutcliffe, and R. A. Lerner.** 1982. Immunogenic structure of the influenza virus hemagglutinin. Cell **28**:477–487.
22. **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.
23. **Hérissé, J., G. Courtois, and F. Galibert.** 1980. Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome. Nucleic Acids Res. **8**:2173–2192.
24. **Hérissé, J., and F. Galibert.** 1981. Nucleotide sequence of the EcoRI E fragment of adenovirus 2 genome. Nucleic Acids Res. **9**:1229–1240.
25. **Jeng, Y.-H., W. S. M. Wold, and M. Green.** 1978. Evidence for an adenovirus type 2-coded early glycoprotein. J. Virol. **28**:314–323.
26. **Johansson, K., H. Persson, A. M. Lewis, U. Pettersson, C. Tibbetts, and L. Philipson.** 1978. Viral DNA sequences and gene products in hamster cells transformed by adenovirus type 2. J. Virol. **27**:628–639.
27. **Kapoor, Q. S., W. S. M. Wold, and G. Chinnadurai.** 1981. A nonessential glycoprotein is coded by early region E3 of adenovirus type 7. Virology **112**:780–784.
28. **Kelly, T. J., and A. M. Lewis, Jr.** 1973. Use of nondefective adenovirus-simian virus 40 hybrids for mapping the simian virus 40 genome. J. Virol. **12**:643–652.
29. **Kitchingman, G. R., and H. Westphal.** 1980. The structure of adenovirus 2 early nuclear and cytoplasmic RNAs. J. Mol. Biol. **137**:23–48.
30. **Kornfeld, R., and W. S. M. Wold.** 1981. Structures of the oligosaccharides of the glycoprotein coded by early region E3 of adenovirus 2. J. Virol. **40**:440–449.
31. **Larson, R., and J. Messing.** 1983. Apple II computer software for DNA and protein sequence data. DNA **2**:31–35.
32. **Lee, D. C., R. G. Roeder, and W. S. M. Wold.** 1982. DNA sequences affecting specific initiation of transcription *in vitro* from the EIII promoter of adenovirus 2. Proc. Natl. Acad. Sci. U.S.A. **79**:41–45.
33. **Liu, F.-T., M. Zinnecker, T. Hamaoka, and D. H. Katz.** 1979. New procedures for preparation and isolation of conjugates of proteins and a synthetic copolymer of D-amino acids and immunochemical characterization of such conjugates. Biochemistry **18**:690–697.
34. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
35. **Persson, H., M. Jansson, and L. Philipson.** 1980. Synthesis and genomic site for an adenovirus type 2 early glycoprotein. J. Mol. Biol. **136**:375–394.
36. **Persson, H., H. Jörnvall, 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.
37. **Persson, H., B. Öberg, and L. Philipson.** 1978. Purification and characterization of an early protein (E14K) from adenovirus type 2-infected cells. J. Virol. **28**:119–139.
38. **Persson, H., C. Signäs, and L. Philipson.** 1979. Purification and characterization of an early glycoprotein from adenovirus type 2-infected cells. J. Virol. **29**:938–948.
39. **Ricciardi, R. P., R. L. Jones, C. L. Cepko, P. A. Sharp, and B. E. Roberts.** 1981. Expression of early adenovirus genes requires a viral coded acidic polypeptide. Proc. Natl. Acad. Sci. U.S.A. **78**:6121–6125.
40. **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.
41. **Ross, S., and A. J. Levine.** 1979. The genomic map position of the adenovirus type 2 glycoprotein. Virology **99**:427–430.
42. **Sabatini, D. D., G. Kreibich, T. Morimoto, and M. Adesnik.** 1982. Mechanisms for the incorporation of proteins in membranes and organelles. J. Cell Biol. **92**:1–22.
43. **Sarin, V. K., S. B. H. Kent, J. P. Tam, and R. B. Merrifield.** 1981. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. Anal. Biochem. **117**:147–157.
44. **Stalhandske, P., H. Persson, M. Perricaudet, L. Philipson, and U. Pettersson.** 1983. Structure of three spliced mRNAs from region E3 of adenovirus type 2. Gene **22**:157–165.
45. **Storch, T. G., and J. V. Maizel, Jr.** 1980. The early proteins of the non-defective Ad2-SV40 hybrid viruses: the 19K glycoprotein is coded by Ad2 early region 3. Virology **103**:54–67.
46. **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.
47. **Wold, W. S. M., M. Green, K. H. Brackmann, M. A. Cartas, and C. Devine.** 1976. Genome expression and mRNA maturation at late stages of productive adenovirus type 2 infection. J. Virol. **20**:465–477.