Sequence and Genetic Organization of Adenovirus Type 35 Early Region 3

PHYLLIS R. FLOMENBERG,^{1*} MEI CHEN,² and MARSHALL S. HORWITZ^{2,3,4}

Departments of Cell Biology,² Medicine,¹ Microbiology-Immunology,³ and Pediatrics,⁴ Albert Einstein College of Medicine, Bronx, New York 10461

Received 18 April 1988/Accepted 19 July 1988

The early transcription region 3 (E3) of group B adenovirus type 35 (Ad35), a serotype isolated primarily from patients with acquired immunodeficiency syndrome and other immunodeficiency disorders, has been partially sequenced. We had previously identified an Ad35 29-kilodalton (kDa) early glycoprotein which, analogous to group C Ad2 E3-19K, associated with major histocompatibility complex class I antigens in the endoplasmic reticulum of infected cells. The open reading frame (ORF) of the Ad35 29-kDa protein has now been identified within a 2-kilobase-pair cloned Ad35 E3 fragment. The predicted amino acid sequence was very similar to that of group B Ad3 E3-19K. In contrast, homology between the Ad35 and Ad2 glycoproteins was limited to five cysteines in identical positions and a 20-amino-acid region proximal to the transmembrane domain. In addition, 20.3- and 20.6-kDa ORFs have been identified downstream from the ORF for the Ad35 glycoprotein. Analogous 20-kDa ORFs are present in the Ad3 E3 region but are not present in Ad2 and Ad5. In contrast, the region analogous to an Ad2 11.6-kDa ORF, which is 9 kDa in size in Ad3, was absent from the expected position within the Ad35 E3 region. Because the E3 region is likely to play an important role in the interaction between virus and host, analysis of the function of the Ad35 E3 proteins should further our understanding of adenovirus pathogenesis.

Adenovirus type 35 (Ad35) is a group B adenovirus that has been isolated primarily from patients with acquired immunodeficiency syndrome (AIDS) or other immunodeficiencies (8, 28, 30). We have shown that over 10% of patients with AIDS or AIDS-related complex surveyed at our institution had urine cultures positive for Ad35 (19). In contrast, this serotype was isolated rarely from normal individuals with febrile illnesses, and specific Ad35 antibody was present only at a very low titer in pooled gamma globulin (12). We have therefore proposed that Ad35 may be unique from other adenoviruses in its interaction with the immune system. Currently, there is no well-established animal model in which to test this hypothesis.

We have focused on the characterization of the Ad35 early transcription region 3 (E3) because this region is likely to play an important role in the interaction between the virus and the host. It has been demonstrated that the group C Ad2 and Ad5 E3 regions code for a glycoprotein, E3-19K, which associates with major histocompatibility complex (MHC) class I antigens (22). In addition, the E3 region is not essential for viral replication in vitro (20) yet appears to have been conserved among all adenovirus serotypes studied so far (6, 11, 16, 17, 26, 29). The Ad2 glycoprotein has been localized to the endoplasmic reticulum (ER), where it interferes with the transport of class I antigens to the cell surface (1, 3). Inhibition of the expression of class I antigens, which are needed for cytotoxic T-lymphocyte recognition of virusinfected cells (9), may enable adenovirus-infected cells to evade immune recognition and facilitate latency. Besides E3-19K, the Ad2 E3 region also contains eight other open reading frames (ORFs) which could code for proteins important for infection in vivo. Currently, only two other E3 proteins, coded by the 11.6-kilodalton (kDa) and 14.7-kDa ORFs, have been shown to be synthesized in adenovirusinfected cells (31, 33). (Another 11.6-kDa ORF is present in the Ad2 E3 region but is predicted to be noncoding.) Recently, the product of the 14.7-kDa ORF has been shown to protect infected cells against lysis by tumor necrosis factor (14).

We have previously identified and characterized a group B Ad35 29-kDa early glycoprotein which associated with MHC class I antigens in the ER, analogous to the Ad2 E3-19K (11). In contrast to the Ad2 glycoprotein, the Ad35 glycoprotein, E29, was synthesized in much smaller amounts, was more extensively glycosylated, and did not cross-react with polyclonal antibody against the Ad2 glycoprotein.

We now report the identification of the Ad35 E29 ORF within the Ad35 E3 region. A 2-kilobase (kb) Ad35 E3 fragment from map coordinates 79.5 to 83 was cloned and sequenced. Ad35 genomic DNA was extracted from purified virions grown in HeLa cells infected with Ad35 Holden (30), as previously described (18). The Ad35 BamHI A fragment, located between 60.5 to 85 map units (m.u.) (32), was cloned into pUC12 by standard recombinant DNA techniques. The right-hand end of the fragment, containing the E3 region, was subcloned by making use of one EcoRI restriction site at 79.5 m.u. and another within the multicloning site of the plasmid, adjacent to the BamHI site at 60.5 m.u. The construct was cut with EcoRI and religated, and the 60.5-to-79.5-m.u. EcoRI fragment was removed. A pUC12 plasmid containing the remaining Ad35 fragment between the 79.5m.u. EcoRI site and the 85-m.u. BamHI site was obtained. The Ad35 E3 fragment was cloned into both M13mp18 and M13mp19, between the EcoRI and BamHI sites in the multicloning regions of these vectors, and sequenced in both directions. A series of overlapping deletions within the insert were subcloned into both M13 vectors, through the use of the IBI Cyclone protocol, based on the method of Dale et al. (7). Sequencing reactions were performed by using a modification of the Sanger dideoxy method (27). The sequence of the entire 2-kb fragment was determined for both strands.

Translation of the L strand of the Ad35 1,978-base-pair E3

^{*} Corresponding author.

DNA sequence in all three reading frames identified four ORFs which have a capacity to code for proteins larger than 6 kDa. The predicted proteins had molecular masses of 18.5, 6.4, 20.3, and 20.6 kDa. The 2-kb Ad35 E3 nucleotide sequence exhibited a high degree of homology to a region within the group B Ad3 E3 sequence published previously (29). The three largest Ad35 ORFs, 18.5, 20.3, and 20.6 kDa, are analogous to Ad3 ORFs previously identified. The nucleotide sequence of the Ad35 E3 fragment and the predicted amino acid sequences of the four ORFs, together with the alignment of the Ad3 sequences, are shown in Fig. 1. Figure 2 shows a schematic representation and alignment of the Ad35 and Ad3 DNA sequences and ORFs. These data represent the first nucleotide sequence established for the unique group B Ad35.

Comparison of the Ad35 18.5-kDa ORF with the Ad3 E3-19K ORF revealed approximately 80% identity at both the DNA and protein levels. Analysis of the predicted amino acid sequence of the Ad35 ORF revealed that there are four canonical glycosylation sites (Asn-X-Thr/Ser), an observation which is consistent with our studies of the carbohydrate residues on the Ad35 E29 glycoprotein. The Ad35 glycoprotein has been shown by sodium dodecyl sulfate-polyacrylamide gel electrophoresis to migrate as a 29-kDa band because of the presence of four N-linked high-mannose oligosaccharides per molecule. The Ad3 glycoprotein has not been characterized, but its ORF also contains four putative glycosylation sites. In contrast, Ad2 E3-19K contains two high-mannose oligosaccharides per molecule and migrates as a 25-kDa protein (21). Hydrophobicity plots generated for both Ad35 and Ad3 glycoproteins reveal that both proteins have a typical NH₂-terminal signal sequence, a large hydrophobic transmembrane domain, and a short hydrophilic cytoplasmic tail, similar to Ad2 E3-19K (Fig. 3).

Both the Ad35 18.5-kDa and the Ad3 19.2-kDa ORFs exhibited much more limited homology to the Ad2 E3-19K ORF at the DNA (50%) and protein (30%) levels. This observation was consistent with our inability to demonstrate immunoreactivity between the Ad35 E29 glycoprotein and a polyclonal antibody to Ad2 E3-19K. However, comparison of the Ad35, Ad3, Ad2, and Ad5 MHC class I antigenbinding proteins has identified one prominent region of homology among all four proteins (Fig. 3 and 4). A conserved stretch of 20 amino acids is located adjacent to the transmembrane domain, within the lumenal globular domain of the protein. In this region, 15 amino acids are identical and 3 amino acids are related (L-I, F-Y, A-T), whereas only 2 amino acids are not conserved among all four proteins. Since group B and C glycoproteins all share the ability to bind to MHC class I antigens, one hypothesis is that the conserved domain represents a common binding site. This possibility is currently being addressed by site-directed mutagenesis experiments. In addition, there are five Cys residues which are absolutely conserved, located at identical positions within each protein. It is likely that these Cys residues are involved in maintaining a structure that is conserved by critical disulfide bonds.

Further comparison of the four class I antigen-binding proteins has revealed that the last three COOH-terminal amino acids, Lys-Met-Pro, are absolutely conserved. The cytoplasmic-tail amino acids have been shown to be important for localization of the Ad2 glycoprotein to the ER. Deletion of the last eight amino acids is sufficient to allow Ad2 E3-19K to move out of the ER (25). Our current data suggest that three residues may be sufficient for retention of the glycoproteins in the ER or that the three amino acids are an important conserved domain of the octapeptide ER retention signal.

In addition, we have identified Ad35 E3 ORFs of 20.3 and 20.6 kDa downstream from the 18.5-kDa ORF. Both of these ORFs are strongly predicted by TESTCODE analysis to be coding (10). Hydrophobicity plots of the predicted amino acid sequences indicate that both ORFs have features typical of transmembrane proteins (data not shown). Both proteins coded by the 20.3- and 20.6-kDa ORFs were predicted to be heavily glycosylated because the sequences contain four and five putative glycosylation sites, respectively. Despite these similarities to the E29 protein, neither ORF exhibited any homology to the 18.5-kDa ORFs may be distantly related to each other, but the alignment at either the protein (27%) or DNA (41%) level was not statistically significant.

Both the Ad35 E3 20.3- and 20.6-kDa ORFs are analogous to two ORFs in the group B Ad3 E3 unit. Specifically, the Ad35 20.3-kDa ORF showed 79% identity with the Ad3 20.1-kDa ORF at the DNA level and a 76% alignment at the protein level. The Ad35 20.6-kDa ORF was 70% homologous to the Ad3 20.5-kDa ORF at the DNA level, and the predicted proteins were 60% identical. Neither group C Ad2 nor Ad5 had an analogous 20-kDa ORF within its E3 region, which was consequently 1 kb smaller. The function of the unique group B E3 proteins remains unclear. In addition, a 6.4-kDa Ad35 ORF was present in another reading frame within the same DNA sequence as the 18.5-kDa ORF. It had no analog within the E3 units of the other serotypes sequenced, and TESTCODE analysis yielded no prediction in terms of probability of coding.

Unexpectedly, the analog of the Ad2 11.6-kDa and the Ad3 9-kDa ORFs was absent from its anticipated location in the Ad35 E3 region. Instead, directly downstream from the 20.6-kDa ORF, a DNA sequence closely related to the Ad3 10.2-kDa ORF was present (Fig. 1 and 2). The Ad35 E3 sequence data extend approximately two-thirds into this ORF. Since both complementary Ad35 DNA strands were sequenced at least once and the deletion observed corresponds specifically to a single complete ORF, the absence of the E3 ORF is unlikely to represent a sequencing or cloning artifact.

It is intriguing that the equivalent of the Ad2 11.6-kDa ORF was absent from the Ad35 E3 region. The Ad2 protein has previously been shown to be expressed in infected cells by analysis with antibody to synthetic peptides (33). An analogous 9-kDa ORF is present in Ad3, and a related 7.7-kDa ORF has been identified in group B Ad7 (J. Engler, personal communication). Perhaps the absence of an analogous ORF in Ad35 is related to the unique epidemiology of Ad35 infections. Ad35 has been isolated primarily from immunosuppressed hosts such as patients with AIDS, bone marrow transplant recipients, and children with severe combined immunodeficiency disease. Ad35 infection usually causes asymptomatic viruria but occasionally has resulted in fatal pneumonias. Perhaps the Ad2 11.6-kDa ORF codes for a virulence factor which is absent from Ad35. On the other hand, it may not be an essential protein in the adenovirus life cycle in vivo. Less likely, the ORF may be present in an aberrant location within the Ad35 E3 region and therefore was not included in the sequence data presented here.

In addition, the Ad35 E3 sequence does not contain the polyadenylation [poly(A)] consensus signal that would be predicted by comparison with the Ad2 E3 region. The Ad2 and Ad5 E3 units have been shown to encode multiple overlapping mRNAs; the mRNAs are separated into two

P CCT S S ATCC ATCC	и и	AGAGA :: : GGACA D 460	50 ACAGC C) 1 5ATA 5ATA 3ATA 3ATA	n on si
1170 CLCN11 CTCCAAT1 T 1330 T L 3::::::	P Y 1 300 1490 K E	$ \begin{array}{llllllllllllllllllllllllllllllllllll$	1650 CTTTAC	<pre>710 1770 1790 1790 1810 0 T I V H I T L G H <u>N H T</u> L V G P I T S E V I W T K L G S V D Y F D I CAAACATTGTCCATATTACTTAGGACCATAATTACATGAGGGCCCAAACTTGGAGGAGTGATTATTTGATA </pre>	135 sequence (bottom) is shown with its alignment to the Ad3 E3 sequence from nees of predicted ORFs are shown above the Ad3 nucleotide sequence; Ad35 amino -, Putative glycosylation sites (Asn-X-Thr/Ser). The computer programs DFASTP (D. J. Lipman and W. J. Wilbur, National Institutes of Health, Bethesda, Md.) were
A N L rgcraatct ri:::::: rgccaatt rgccaatt 150 150 150 150 150 150 150 150	T I 290 P S	CTCCCA ::::: CTCCTA 450	CCTTTT : CAGA 610	D Y TTGATTA' ::::::: TTGATTA' 760	seque ce; Ad grams esda, N
6 S / 1 566CTCT 566CTCT 566TTCT 566TTCT 566TTCT 140 140 140 140 140 140 140 140	о с д з	ratggc :::::: ratggc	CCTCATA : :: CTGTTTA 600	G S V GGAAGTGT :::::: GGAAGCGT 750	d3 E3 sequen er proj 1, Betho
1150 1150 GGGTTTGGGG GGGTTTGGGG GGAACCGGG GGAACCGGG E P 14(1310 1310 1310 14(1310 1310 1310 14(1310 1310 14(14) 14) 14(14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14(14) 14) 14) 14(14) 14) 14) 14(14) 14) 14) 14(14) 14) 14) 14) 14(14) 14) 14) 14) 14) 14) 14) 14) 14) 14)	I T K 280 1470 Q Y D L	ATGACC	9k 1630 :TTTTTC: :::::::::::::::::::::::::::::	1790 L G LACTTGG LACTTGG2 LACTGGGG	e the A comput Health
L 11 L 11 CCTTTTGG SCTTTTGG SCTTTTGG SCTTTTGG 130 H N N ACATAATA	СНТІТКР БТ 270 280 290 28 21470 S R Q Y D L W P P	ACAGT	end 19k 1630 * TAACCTTTT ::::::::: TTAACCTCTT * 590 60d 18.5k	M T K L TTGGACCAAACT ::::::::::: CTGGACCAAACT 740	nent to 13 nucl The c utes of
S L TTTCGC :::::::: TTCGC TTH TTAH ::::::: :TTACAC	L H 270 S R	TCATGAGCAG ^A ::::::::: TCATGAGCAA ^J TCATGAGCAA ^J S * 430 end 6.4k	K M P GAAAATGCCT ::::::: LAAAATGCCT 580	I W TATTTG :::: CATCTG	alignn the Ac nr/Ser). I Instit
L L L L IIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	P L K 260 1450 M F M	CATGTTCATG :::::::::: CATGTTCATG CATGTTCATG C S * 420 end 6	end 19k 1610 1630 K E M P * AAACATAACCTTAACCTTTTC :::::::::::::::::) E V AGAGGT :::::: AGAGGT 730	vith its above n-X-Th Vationa
1130 L A GGTTGC GGTTGT GGTTGT GGTTGT 1290 S V S V S V S V S C CGTCGGT GGTCCGT	G P L K L H T 260 270 1450 A S R (56CCATG 56CCATG 56CCATG P C	1610 : K E : AAAAGA :::::: AAAGA	1770 T S 1 FACTTCA(:::::: CACTTCA(iown w shown es (As ilbur, N
V V 1 AGTCGTGC :::::::::: AGTCTGC AGTCTTGC AGTCTGC 110 110 110 110	2250 A K D L	GATCTG :::::: GATCTG GATCTG 410 410	E E ATGAGGA :::: ATAAC 570	P I CCCAATT	n) is sh Fs are tion sit /. J. Wi
2k CCATACT. CCATACT. CCATACT. CCATACT. CCATACT. CCATACT. 100 5k M D 5k M C 5 K ATGGGAC ATGGGAC	M G N A G I 240 250 250 atg 6.4k 1430 L A 5 E M C D L A	TTTTCTGAAATGTGCGATCTC TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	S N AAAGCAA' :::: ACGCCAA' A 5	G P AGGGCCC ::::::: AGGACCC 710	bottom ed OR cosylal r and W
atg 19.2k 1110 M GG A 1 M A M GG A 1110 1100 10	M 240 atg 1430 S E M	rctgaa. :::::: rctgaa. L K 400	1590 R Q CCTAGACA TCCAAAAA TCCAAAAAA	1750 L V TCTTGTA :::::::: TCTCATA 7	Jence (predict ive gly Lipman
atg 11 13 6AAATG 5.::::: 2.:::::::: 90 90 90 90 81 90 81 12 12 12 12 12 12 12 12 12 12 12 12 12	£1.	ATTTT ATTTT F F	R P Treecco	I T ZACACT ::::: racact	35 sequences of the seq
GTTTAA GTTTAA GTTTAA GTTTAA GTTTAA T T T T	230 F	ACTTTCATT ::::::::: ACTTTCATT ACTTTCATT L S F 390	I V I R P R ATAGTTATTCGCCTAG ::::::::: CTAACCACTCGCATCAA V T T I K	H <u>N H T</u> ACATAATCACAC ACATAATTACAC ACATAATTACAC ACATAATTACAC Y 700	he Ad: sequer
$ \begin{array}{ccccccc} 70 & 1090 & \text{atg } 19.2 \text{k} & 1130 & 1130 & 1150 & 1150 & 110 \\ Grccatatitadrcatrategests testimassan segses can be considered as the construction of the constructin of$		ACAA CA ::::::: ACAA CA F T 30	V C I H L I I TGTGGGATACACTTGGTT ATAA ::::::::::::::::::::::::::::::::	0 G H TAGGAC :::::: TAGGAC	ient. T io acid quence id NU(
1090 1090 11250 1250 1250 1250 1250 1250 1250 12	220 1410 S N N	ATCCGCATTAGTAACAA :::::::::::::::: ATCCGCATTAGTAACAA ATCCGCATTAGTAACAA S A L V T T S A L V T T 370 380	1570 H L L ACACTTGC' ACACCTGC' ACACCTGC' 540	H I I G H I I G CCATATTACCTTAGG ::::::::::::::::::::::::::::::::::	I fragn . Amin tide se isco) ar
GTCATA ::: GTGGTG GTGGTG A 70 70 70 71 10 71 70 70 70 70 70 70 70 70 70 70 70 70 70	210 R I	CCCGCAT ::::: CCCGCAT 370 370	C I 5TGCAT 1:::: ATGCAT 530	H I FCCATAT FCCATAT I I I I FCTCTAT S 680	BamH a colon nucleo Franc
TATTTA TATTTA TTCCTA TTCCTA 60 D T 51::::: CGACAC	s s	TTCCAT :::::: TTCCAT	C V ITGTGTO :::::: STGCGTA	I V CAATTGT :::::: CAGTCGT 670	<i>Eco</i> RI- ted by Ad35 Ad35 t at San
070 IGTGCA : TTCACT TTCACT TTCACT 230 A P TGCACC TGCACC TGCACC	200 1390 D G	TGACGGTTCC/ ::::::::: TGACGGTTCC/ TGACGGTTCC/ 360	1550 I I ATCATT ATCATT ATCATT I I 520	1710 0 T 36CAAAC 36CAAAC 36CAAAC	35 E3 indicat indicat ow the ow the ilifornia
10. 10. 200716CTV 200716CTV 50 50 12. 12. 12. 12. 12. 12. 12. 12. 12. 12.	ы D	AGGTCC AGGTCC V L	0 I I T A 1 ccatgcatcatcactgcra ccatgcatcatcactgcra : :: : :::::: ccttgccttcttactgcr1 b L L 1 500 510	H G ETCACGG ETCACGG ETCACGG ETCACGG 660	air Ad les are ed bel / of Ca
acacao acacao acacao acacao acacao	190 V R	CGTCCCAG CGTCCCAG CGTCCCAG S E 350	T I T A MTCATCACTG : ::::: CTTCTTACTG	T V A CACTGTCGC ::::::: CACTGCCGC 550 A	base-pa icleotic indicat iversity ints, re
ACCTAT ACCTAT ACCTAT 40 40 ACCTAT ACTGC	80 V S V	CTGTCTCT :::::::: CTGTCTCT CTGTCTCT 340	ACATGCAT ACATGCAT ACATGCAT ACATGCAT A A 500 L	0 V T rGTCAC rGTCAC rGTCAC	1,978- tical nu ds are vry, Un nlignme
1050 1050 111105 1111111111111111111111	180 1370 Y T V	TATACTG :::::: TACACTG T L 34	1530 L V T SCTTGGTAACI SCTTGGTCACI SCTTGTGCGCC C A 500	1690 S I V AGCATTGT AGCATTGT ::::::::: AGCATTGT	of the). Iden ino aci aborato DNA a
TIGATCCT TTGATCCT TTATTCT TTTCTCT TTCGACC TTTCGACC	170 0 w Y	CCAGTGGTAT : ::::::: CGAGTGGTAC E E S 330 330	C L ATTGCT :::::: ATTGCT 490	I A S ATTGCCA(::: ::: ATTTGTCA(F V 640	uence op) (29 d3 am tion Lá in and
CTTGTGC CTTGTGA CTTGTGA 20 20 L D 1 L TTTAGAC TTTTAGAC	۵. ۲	F P	A Y TTGCTTA ::::::: TTGCTTA	570 A L I I A S AGCTCTAATTATTGCCAG : :::: ::::::::: ATCTCTCATATTTGCCAG S 630 640	ide seq 3341 (t from A omputa e prote
1030 GTGGGC7 GTGGGC7 GTGGGC7 GTGGGC7 1190 1190 1 CCATGT7 CCATGT7 CCATGT7	160 1350 v 5 G v	CCAGGAGT ::::::: CCAGGAGT : CCAGGAGT CCAGGAGT 320 320	1510 S I FTTCCAT' : ::::: FCTCCAT' 480	1670 T A ACAGCT ::::::: FACATCT	fucleoti 1018 to differ 1 atics Co uce the
1030 1050 1070 1090 atg 19.2k 1130 1150 1150 1170 GTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	یم ^۲ س	ATGGGAGC ::::::: ATGGGAGC :ATGGGAGC :ATGGGAGC :ATGGGAGC :310	A F rGGCATT : ::: rAACGTT 70	.1k 5 V 1 rcrGTT :::::: rcrcTT 620 .3k	FIG. 1. Nucleotide sequence of the 1,978-base-pair Ad35 E3 <i>Eco</i> RI- <i>Bam</i> HI fragment. The Ad35 sequence (bottom) is shown with its alignment to the Ad3 E3 sequence from nucleotides 1018 to 3341 (top) (29). Identical nucleotides are indicated by a colon. Amino acid sequences of predicted ORFs are shown above the Ad3 nucleotide sequence; Ad35 amino acids which differ from Ad3 amino acids are indicated below the Ad35 nucleotide sequence. —, Putative glycosylation sites (Asn-X-Thr/Ser). The computer programs DFASTP (Biomathematics Computation Laboratory, University of California at San Francisco) and NUCALN (D. J. Lipman and W. J. Wilbur, National Institutes of Health, Bethesda, Md.) were used to produce the protein and DNA alignments, respectively.
1170 1018 1030 1050 1070 1070 1090 atg 19.2k 1130 1150 1150 1170 1170 1170 a k 1 k 1 k 1 k 1 k 1 k 1 k 1 k 1 k 1 k	33 Fr	ACACATGGGAGGCCAGGAGTTCCCCAGGAGTCTTCTCTGACGAGGTCCTGAGGATTGTAACAA CACTTTTTTTTTT	1510151015301570157015701690161016101630ACATTGEGGATTATCATTGCTTACATGCTATACTTATGGGATACATGCTTATATTCGCCCTAGAGCAATGAGGAAAAAGGAAAAAGGAAAAAGGCAATATTCCTATACCTTATACAGC16301650ACATTGGGCATTTCCATTGCTTGCTTACATGCTACTGCTTATACTTATGGCCTAGGCAATAAGGAAAAAGGAAAAAGGCAATAACGCTTAACCTTATTCCTCATACCTTATCATACGCCTAGGCAATAAGGAAAAAGGGAAAAAGGGAAAAAGGCAAAAAAGGGAAAAAGGGCAATAACGCTTAACCTTATTCCTCATACGCTATACGCCTAGAGGAAAAAGGGCAAAAAAGGGCAAAAAAGGGCAAAAAA	atg 20.1k 1670 1670 1690 1770 1770 1770 1770 1770 1770 1810 M A S V T A L I I A S I V T V A H G Q T I V H I T L G H N H T L V G P P I T S E V I W T K L G S V D Y F D I ATGGCTTCTGTTACGCTCACTGTCGCTCACGGGGGCAACAATTGTCTTAGGAGCCCTCTTGTAGGGGCCCCCAATTACTTCGAGGGTTATTGGAACGTGATTATTGTATA IIIIIIIIIIIIIIIIIIIIIIIII	FIC nucle acids (Biom used 1

	14 30 14 30	. C	ç ç	× < · · <
0 LTA	210 CCC 112	TAC HACK	аў:: ЧС ПС	L L L
197 197 197 197 197 197	337 1	> 0 0	N 1444 1444 1444 10	
CCA.	M TA TA TA O	CAC) S N ATCAA/ AGCAA/ A 1400	AG0 - 60
P AGCCC :: AGTTG 920	C M ::::: TGCA 1080	2270 L L ACCTAN :::: ACCTCN 1240	130 130 130	590 FAA
CA::CA	TA.	CACAC	12/ TG/	1 21
	T A P T A P T A P	S TCA TCA	D V C	CGC M
TT ST	I SAA SAA SAA O		N] TTAAC/ :::: CCAACC	B B B B B B B B B B B B B B B B B B B
<pre>% I % I CGCAT CGCAT CGTGT % % % % % % % % % % % % % % % % % %</pre>	[10 I [AATA/ [:::::: [AATA/ 1070	6600 5123	гАТ :::т	
11C 1	ACT 21	ATGG TG :: OTC	AT' AT' AT'	
191 191 661	P P	Y G R S TTATGGCCGT : : : : : : : TTTCTGCTCGCT S A 1230	D T TTGACC : ::: CTAACC 1380	CAG CAG CAP 10 1
15::0	2 T T T T T T T T T T T T T T T T T T T	2250 V T GGTAACT : ::: CCTGACT L 1220	13 13	15 15
Y I TAC7 :::: TAC7 900	G M 566AATC 5:::::::: 565T6ATC V 1060	2250 V TGTAN : : TCTGN 1220	410 ACC	ZAU: TAV
N 1.1.1.1 1.1.1.1 1.1.1.1 1.1.	00:01	ACT 2 ACT 2 FOR	3AG 22 3AG 22 3AG 22	S LICC 5 S LICC 5
AAA aga	L CAT CAG CAG	CA CCA	G R GGGAAG :: :: GGCAG 1370	лт АТ 30
TATI TTATI	090 V A SGTGGG SGTGGG SGTGGG	LO LO L	13 CC CC C	T ACA 15
Y AATA' :::: AATA' 890	090 267 367 10	L A GCCTTGC 5 :: :: :: GTCTCGC 1210	N CAA CAA	AAT AGC
0000000		S STA(STA(STA(ACL
TAC		TAC) S C C T T T C T T C T T C T C C C C C C	AAG CTA 520
S S TCCAG : ::: TACAG 880	I A TTATTG TTATTG TATTG TGTTG 1040	CAT CAT SCT 200	00 111 111 1	L P CC P K
S TA 8 8		M M M	23 - A7	75.255 76.17
<pre>370 1930 1990 1910 1910 1930 1930 1950 1970 L I N U S K I Y N G Y Y G Y D R S S S Q Y K N Y L V R I T Q P K L T CGTTAATTAATGTTAGCAAAATTTATAAGGGTTACTAATTAGGGGGGTAAATTAGGTGGGGGGAAATTAAG SGTTAATTAATGTTAGCAAAATTTATAGGGTTACTAATTAGGGGGGGG</pre>	2030 2050 2070 2090 2110 2130 T S P T P N M I I C M L	end 20.1k atg 20.5k 2230 2250 2250 2270 V P N I * 2210 M I S T T I F I I T S L A V T Y G R S H L T V P TTAACATTTAATTTTTTATACAG-ATGATTTCATTATCATTACGGCGGCGGGGGGGGGG	GAA K	s cag cag cag
SAC	LTGA TTGA	E E	T AC T AC R AC	51(51(51(
Y I TATC :::: TATC 870	100 11 15 15 15 1030 1030	CCAN ::: CAT CAT 19(1)	, G CTGG1 ::::: CCTGG1 1350	S AGC AGC
0 11:12 0 11:12 0 0 11:12	207 M M AAT 1	L A A T	J CC J CC J CC	5:5
10 10 10 10 10 10 10 10 10 10 10 10 10 1	TTC S	T T		
	N AAA CAAT CAAT		SAT SAT	CTT CTT CTT CTT
Y :::: CTA' 860	P N CCCAAA ::::: CCCAGA	20.5k 2.5 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5	370 C D TGTG/ :::: TGTG/ 1340	T 30
TAC		20 1 0 1 0 1 20 20 20	23 23 23 23 23 23 23	25 S S AAC AAC
5555	I ACA ACA	atg 20.5k 2210 M I S T ATGATTTCCA :::::::: CATGGTTTCCA ::::::::: :::::::: ::::::::::::::	P BAC AAC	N L L L L L L L L L L L L L L L L L L L
	N SAAL		R TAC KA	149 149 149
ATA :: : OACA	50 50 50 50 50 50 50 50 50 50 50 50 50 5		TAC:CGC	
06 7.7.7 7.7.7 7.7.7	3GA : 20 3GA : 20	TAU TAU	LI.I.	CAC I
18 I AA V V	N AAT GAC	TT /	₹ <u>1</u> 20 8 <u>1</u> 20	
CAP CAP		T :TT: 60	0 2 GA 2 GA	
RAGC RAGC RAGC 840	r P ACACC ACACC ACACC 1000	AL - 21 C	2350 1 G MATGG MATGG 1320	CAC
61. 61.	CCH	end 20.1k 90 TTTAATTT :::::::::: TTTAATTTCTT TTAATTTCTT 0 1160 end 20.3k	ATH ACH	C LO
N AAT ::: AAT	C A A T	0 * TTA * * * *	ATG ATG:::	AGC AGC ATC 170
I AATT ::: GATT 830	30 55 P 11CAC 11CTC 990	enc 2190 1 1 , ACATTT :::::: ACATTT ACATTT 1150	Y TAT AT AT	
TA/ TG/	030 030 030 030 030 030 030 030	AAOCA AAOCA 11	I TAAT 1:::: 13	>5.51
870 CGT CAT CAT	T T 2 TAC	5 · · F · · · F ·	R GAG GAG	STA 5TA
TCA T TTA:	F	N F AAATT :: AAGGC ACGC	A L	460
ATC ATC ATC ATC	N AAAAT ::::: AAACT 980	L N ACTAA/ ::::: ACTAA/ R 1140	760 300 300	00 I T T T A T T T A T T I I I I I I I I I I I I I I I I I I
AAA :::: 8	9 5 A A 5 A A	L L L L L L	r 23 1 1 23	24 I ACA ACA
ACA ACA	TTA	J C C		N N N N N N N N N N N N N N N N N N N
R AG	CTC	6 J : 5 0	ATC	TTA 145
N F FAACA :::: CAACA 1 810	010 K A AAAG(:: AATT(970 970	2170 D P AGGATC :AGGATC : :::: :AGGATC L 1130	37C 37C 129	AGA AGA
	20 ACA: N	CACACA	D D D D D D D D D D D D D D D D D D D	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
185 CTT	N T TAP	4 - 4X	BGC AGC AGC	S AGT AGT 40
I TAA TAA TAA T	4000 0000 0000	Q H K TCAACATAA ::::::: TCATCCTAAAA H P H P 1120	S LCA	0 AAT AGT 14
V FTGTA F3:::: FAGTA 800	M A AATGGA : :: ATCCG, 960 960		P CCC 128	470 AA/ AA/
F AA7 I	R AA/ CG/	A AAO A TO F	000 ° ° ° ° ° ° ° ° ° ° ° ° ° ° ° ° ° °	ATP K ATP ATP
I AAT ::::		OCT:CE		Y ACT ACT ACT ACT
P CC P	1990 T I I IGACAATAA :::::: IGGCAAGAA A K 950	50 K F AAGTT' AAGTT' AAGTT' 1110	70 1	14 CGI
K P AAAC AAAC	990 1 1 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	150 K AAA 3AA 11	I CAC CAC CAC 12	AAC
830 ACT ACA	TGT 1 TGC 7 1 A	ZAA.2	T BUDE T	5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	M TAC	Y TA1 TAC	CTC CTC	Y TA1 TA1 420
N R AACA ::::: AACA 780	P T CCAAC CCAAC CCAAA 940	Y STTAC STCGC R 1100	90 260 260	50 7 TAT 1 ::: 1 1
GCA GCA GTA	4 00 :: 00 6	5::5 ⁻ -	22 27 27 27 21	CTC: 24
	M M M M	CT1 C	80 · · 80 10 · · 90	1 55 5 6 5 7 5 7 5 7 5 7 6 7 6
I C N K T K P I F V I C N R Q N L T L I N V S K I Y N G Y Y G Y D R S S Q Y K N Y L V R I T Q P K L T TAMTTECAACATAATTATTTAATTETAATCTEAACAGAAATTTAATGAGGTTACTAATTAACGGGTTACTAATTAAGGTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	V F T M T I I K M A N K A L E M F T S P T P N E K N I P N S M I A I I A A V A L G M A L I I I C M L L A L C A L I I C M L L B X I A I I A V A L G M A L I I I C M L L A L I I C M L I I C M L I I C M L I I C M L I I C M L U A L I I I C M L I I C M L I I I C M L I I C M L I I I C M L I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I I C M L I I I I C M L I I I I I C M L I I I I C M L I I I I C M L I I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I I C M L I I I C M L I I I I I C M L I I I I I I C M L I I I I I C M L I I I I I C M L I I I I I I C M L I I I I I I C M L I I I I I I C M L I I I I I I C M L I I I I I I I I I I I I I I C M L I I I I I I I I I I I I I I I I I I	Y A C Y Y K K F O H K O D P L L N F TATGCTTGTTATTATAAAAGTTTCAACATAAAGGGATCACTACTAATTT TATGCTTGTTATTATTTCAACATAAAGGGATCACTACTAATTT TATGCTTGTTGTTCATCAACAAAAAAAAAA	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	2450 2550 2470 2490 2590 2590 2590 2590 2590 2590 2590 25
H	~ U	ようふん	E E-	OU AX

E E	(i. Es 1	Ĺ Ĺ.		F1 ·· F1		
L L LTAC7	T T S Y A Y T N I Q P K Y A M Q L E I T I L I V I G I L I L S V I L Y F Ictaceteratecatacetrageatecetagaateceatecaateceateceateceate	end 9k 3030 8050 atg 10.2k 3070 2990 2010 8 NF 2990 3010 4 10.2k 3030 3050 3010 4 1 2 R N F CTATTCTCTAGATTAGGTGATCATCATCATCATCATCATCATCATCATCATCATCATCA		A T A S P D C I G P F 3170 A T A S P D C I G P F A S Y A L F A F V T C I C V C S I V C L V I N F AGCACTGCAAGGCCCAAAGGCCCTATGCTTCCTATGCTTTTGCCTTGGTTACTTGGCGGGGGGGG	50	
P 1		R 1 FAGAI	1790	I N Attaa :::::	1950	
750 D TGA1 TGA1 TGA1	2910 V I GTTA1	atg 10.2k 3070 M I P R CATGATCCTAG ::::::::::::::::::::::::::::::::::::		3230 v I GGTTAT :::::: GGTTAT		
2 6 8 8 8 8 8 8 8 1	S 2 TCTG	atg 10.2k 3070 M I P CATGATCC :::::::::	0	C L CTGCCTG CTGCCTG	1940	
CAAAA CAAAA CATAA CATA	L ACTA	at M ATCA ATCA	1780	сс с тстб тстб	1	
E FGAAG :: CAAAG FK 1 170	I FTAT	ATCA CCA		I V ATAG ATAG ATAG	0	
D D D D D D D D D D D D D D D D D D D	ATTC'	SAAC SAAC	1770	S LAGC LAGC	1930	
30 GAA7 GAA7	2890 	3050 ATGGT0 :::::: ATGGT0	.н	3210 V C S CGTGTGTAGC ::::::::::::::		
Y 27 TATA :::: TATA	26 I AATT	30 GTA1 5:::: GTA1	_	32 10 10 10 10 10 10 10 10 10 10 10 10 10	1920	
C 160 C 160 C 161	TTGT	TACA :::: TACA	1760	I C CATCTG ::::::	1	
C CCTG :::: CCTG 1680	L I CTAA	CTCT . : AT-T		TTGC TTGC TTGC		
LACG A	I ААТА	TTTT :::: TTTT	0	T TTAC :::: TTAC	1910	
T T T T T T T T T T T T T T T T T T T	70 TCAC	30 TTTT TTTT	1750	90 F V 1.1.CC 1.1.1 TTTC		
27: M M T T 16: 16:	2870 E I 7 SAAATC/	3030 ATCTTT :::: TTCTTT		3190 A F TGCCTT(:::::: TGCTTT	00	
	L I ACTAC	9k 19k 19		F F F F F F F F F F F F F F F F F F F	1900	
и 1116 11116 660 660	GCAI	end 9k I * MATCTAA		3170 Y A L F P F A S Y A L F i.accartrGctrcctArGcActrtr i.accartrGctrcctArGcActrtr i.accartrGctrcctArGcActrtr		-2k
F ATT1 ATT1 TGT1 V 1	A M GCTAT	E		Y Z	1890	end 10.2k I * TTATTAA
0 1171 11.: 11.: 0 1.1	Y TACG 	TGAA		0 5 117CC 111CC	-	end CTT2
269 ATGA ATGA ATGA 165	2850 P K Y GCCTAAAT	3010 A L GCTCT		3170 A TTGCT TTGCT		3330 a R L TAGGC
L ATTA ATTA: AATA: I	AGCC	MTATG		P F CCAT ::: GCAT		TTCT
6 6 6 6 6 6 7 7 7 6 7 6 7 7 7 7 7 7 7 7	AT TCAG	H CCCA'		6 6 6 7 8 7 8 7 8 7		C AC
TAGT STAGT STAGT STAGT	N I CAAACP	CGTCC		C I G CTCTATAGO :::::::: CTGTATAGO	1870	A A 2
>90.::0 .:::0 .:::	CAC	S F		ACTO ACTO	18	LTGT1
2670 A TGCC :::: TGCA 1630	2830 1 Y 5CATA(2990 I 'GATT		3150 P D CCCAGA CCCAGA	_	end 10.2 3330 3330 3330 3330 A L L R L I L I L I L A MATACGCAATCAAAATGTTGCGGGCACTTATTAGGCTTATTAA
TTTAC TTTAC TCGC TCGC	Y A TATG 	CTAT		S AAGC AACC	1860	AATC
I ATCA ATCA 20 20	s crcc	S P TCTC		CTGC CTGC :::: CAGC		A DOD
G TGGC :: CAGC S 16	T CTAC	Y CTAT		A T NGCCA NGCCA	50	AATA
I C AAT :::: C AAT	ACCA	I CCAT		V AGTA :::: AGTA	18	щü
2650 5 7 7 7 7 2 7 2 7 7 2 9 1610	2810 N AAAT	2970 R P NCGTC		3130 A CAGC CAGC		3290 H P ACCATC
T T	S AATC	AAGA		F T TTCA TTCA	1840	H GTCA
	ATG 9k 2810 90 I. F. Q. S. N. J. M. I. S. Q. S. N. J. S.	2970 SKRP TTCTAAAAGACGTCC		TACT		CTAC CTAC
I I ACAA' ACAA' ACAA'	ATTG	A T T S		3760 3760 3060	30	TAC
N CAATI CAATI TACAI	ATG 9k 90 I I M I I TATGATAT	R I TAGA		5777.C	1830	rt60
2630 P CTCC(E ::: CACA H L590	AT 2790 M MATTAT	2950 / H STTCA1		3110 N TAAT :::: TAAT		3270 R I CGAA
	raga!	AATG.		E CTTT CTTT	1820	L C L C L C L C L C L C L C L C L C L C
T ZACA(::: 10 10	.5k [TTA: 10 6k	P I VCCCI		3090 3110 3110 3110 3130 T I L I C A F N V C A T F T A V SECENTACTCATCTGTGCTTTTAATGTCTGTGGTGGTGGT ACCATACTCATCTGTGCTTTTAATGTTTGCGGGGGTGGTGGTGGTGGTGGGGGGGG		F CC FCC
ATTCC ATTCC ATTCC ATTCC	end 20.5k * TAATTTTT TAATTTT TAATTT * 1740 end 20.6k	LAAT/		I (CATC1 CATC1	10	rgGA1 rgGA1 rgGA1
AGTAJ	en¢ 111A/ 111A/ 111A/	2GTCL		L LACTC	1810	D V AGACT VGACT
2610 L 5 CTTA	2770 D I FGATAT FGATAT FGATAT FGATAT	2930 C F		3090 T I CACCAT CACCAT		3250 V 16677
26102630263026702670269027102730273027302750AAATGGCTTAAGTAATTCCTAAATGCAATTACTATTACTATTACTATTTTGTGCAATTAATT	end 20.5k ATG 9k 2810 N F D I * 2790 C 2810 AAATTTGATATTTAATTTTTAAGAATTAGAATAGTGATGA	2930 2950 2950 2970 I F C R Q I P N V H R N S K R R P I TCATATTCTGCGGTCAATACCCAATGTTCATAGAAATTCTAAAAGAGGGTCCCAT		3090 3100 3110 3110 3130 3130 3150 3150 3170 3170 3190 3190 3210 3210 3210 3210 3230 F F T L I C A F N V C A T F T A V A T A S P D C I G P F A L F A L F V T C I C V C S I V C L V I N F CTTCTTCACCATCTCTCTCTCTCTCTCTCTCCTCTCTCT	1800	F Q L V D W I F V R I A Y L R H H P E TTCCAACTGGTAACTGGGATCTTGGGGAATTGCCTACCTA
I AAA'I :: AACC R	N AAA1 : TAG7 R	TCAT		F CTTC CTTC		TTC(TTC(TTCC

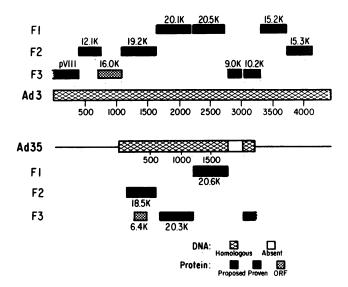


FIG. 2. Comparison of the E3 regions of Ad3 and Ad35. Alignment of the Ad35 sequence within the 4-kb Ad3 E3 *Xhol-Ball* fragment (from 76.5 to 89.2 m.u.) (29) is illustrated schematically. The gap in the Ad35 E3 region indicates the absence of approximately 250 nucleotides which code for a 9-kDa ORF in the Ad3 sequence. Boxes represent ORFs larger than 6 kDa. F1, F2, and F3 are the three reading frames; K, kilodaltons.

groups, E3A and E3B, on the basis of the use of two different poly(A) sites (4, 5). The major Ad2 E3 mRNA transcribed, which codes for E3-19K, uses the first poly(A) site downstream from the 11.6-kDa ORF; the poly(A) site is preceded by the hexanucleotide signal sequence ATTAAA. The consensus nucleotide sequence AATAAA is present at the second poly(A) site, downstream from the Ad2 5'-terminal ORF. In the Ad3 E3 region, the sequences AATAAA and J. VIROL.

	•		*		
3	MGAILVVLALLSLLGLGSANLNPLDHDPCLDFDPENCTLTFAP				
35	MGPILVLLVLLSLLEPGSANYDPCLDFDPENCTLTFAP				•••
00		UT SAT		(Dakonjo	
2	M-RY-MILGLLALAAVCSAAKKVEFKEPACNVTFKS	EANE-	CTTLIK	TTEHEK	LI
5	MIRY-MILGLLTLASAHGTTQKVDFKEPACNVTFAA	EANE-	CTTLIK	TTEHEK	LL
3	I THNNKTWNNTL STTWEPGVPQWYTVSVRGPDGS I R I SNNTF I	FSEMO	DLAMFM	ROYDLW	PP
				:	::
35	ITHNNKTWNNTLSTTWEPGVPEWYTVSVRGPDGSIR ISNNTFI	FSEMO	DLAMFM	SKQYSLW	PP
		: :::		:	
2	IRHKDKIGKYAVYAIWQPGDTNDYNVTVFQGENR-KTFMYKFP				
5	IRHKNKIGKYAVYAIWQPGDTTEYNVTVFQGKSH-KTFMYTFP	FYEM	DITMYM	SKQYKLW	PP
	•				
3	SKENIVAFSIAYCLVTCIITAIICVCIHLLIVIRPROSNEEKE	KMP			
		1			
35	SKDNIVTFSIAYCLCACLLTALLCVCIHLLVTTRIKNAN-NKE	КМР			
	: :::::::	:::			
2	QKCLENTGTFCSTALL ITALALVCTLLYLKYKSRRSFIDER				
_	· · · · · · · · · · · · · · · · · · ·				
5	QNCVENTGTFCCTAMLITVLALVCTLLYIKYKSRRSFIEE	UMP			
	FIG. 4. Alignment between the Ad35. Ad3.	Ad2	and A	d5 MI	HC

FIG. 4. Alignment between the Ad35, Ad3, Ad2, and Ad5 MHC class I antigen-binding proteins. Identical residues are indicated by a colon. Symbols: , conserved domains; *, Cys residues which are absolutely conserved among the four proteins.

ATTAAA are present close together within the 20.1-kDa ORF (at nucleotides 1996 and 2008), and an atypical sequence ATGAAA, is present at the end of the 9-kDa ORF (nucleotide 3016); one of these sequences may function as a poly(A) site for the Ad3 19.2-kDa ORF. In contrast, none of the above sequences is present in an analogous location in the Ad35 E3 sequence. It is possible that the absence of a poly(A) site adjacent to the Ad35 18.5-kDa ORF could result in less efficient transcription. This hypothesis is consistent with our data that show that the Ad35 E29 glycoprotein is produced in significantly smaller amounts compared with Ad2 E3-19K. For instance, studies of Ad2 E3 deletion

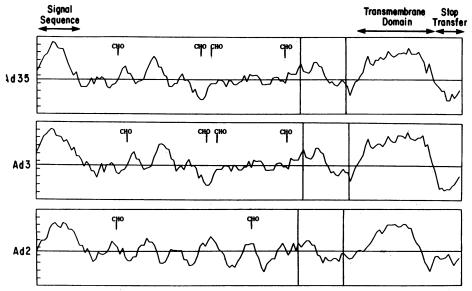


FIG. 3. Comparison of hydrophobicity plots for the MHC class I antigen-binding glycoproteins of Ad35, Ad3, and Ad2. Plots were generated by the computer program ANALYSEP (Roger Staden, Medical Council Center, Cambridge, England) with the hydrophobicity values of Kyte and Doolittle (23). A window span of seven was used. The y axis represents the hydropathy index from +50 (top) to -50 (bottom). Hydrophilic C-terminal stop transfer regions and domains which are typical for hydrophobic N-terminal signal sequences and transmembrane regions are indicated. CHO, Putative glycosylation sites. A 20-amino-acid domain that is conserved among all three glycoproteins is indicated within the vertical lines adjacent to the transmembrane domains on each plot.

mutants have demonstrated that the elimination of the first poly(A) site or other defined sites downstream from the 19-kDa ORF can decrease levels of Ad2 E3-19K mRNA (2). Alternatively, a novel sequence which can function as a poly(A) site for the 18.5-kDa mRNA may be present in the Ad35 E3 region.

The functions of the E3 proteins in adenovirus pathogenesis remain unclear. The MHC class I antigen-binding glycoprotein may effect virulence or facilitate latency, but direct evidence is lacking. Although recent studies of infection of the cotton rat with Ad5 mutants have indicated that deletion of E3 enhances virulence, similar experiments with hamsters have yielded contradictory results (13, 24). We are currently investigating an alternative model, using a mouse adenovirus in its natural host (15). Further studies will also provide important information in regard to the safety of using an adenovirus E3 deletion mutant as a live-virus vaccine vector in humans, as has been proposed elsewhere (24).

We thank E. Guetta for assistance in cloning the Ad35 BamHI A fragment, R. Korn and A. Weinreb for technical advice, G. DeFilippi for typing the manuscript, and J. Engler for critically reading the manuscript.

This work was supported by Public Health Service grants AI21105, AI20671, and CA13330 from the National Institutes of Health.

LITERATURE CITED

- 1. Andersson, M., S. Paabo, T. Nilsson, and P. A. Peterson. 1985. Impaired intracellular transport of class I MHC antigens as a possible means for adenoviruses to evade immune surveillance. Cell 43:215-222.
- Bhat, B. M., and W. S. M. Wold. 1987. A small deletion distant from a splice or polyadenylation site dramatically alters premRNA processing in region E3 of adenovirus. J. Virol. 61:3938– 3945.
- 3. Burgert, H., and S. Kvist. 1985. An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens. Cell **41**:987–997.
- 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.
- Cladaras, C., B. Bhat, and W. S. M. Wold. 1985. Mapping the 5' ends, 3' ends, and splice sites of mRNAs from the early E3 transcription unit of adenovirus 5. Virology 140:44-54.
- Cladaras, C., and W. S. M. Wold. 1985. DNA sequence of the early E3 transcription unit of adenovirus 5. Virology 140:28–43.
- 7. Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13:31–40.
- deJong, P. J., G. Valderrama, I. Spigland, and M. S. Horwitz. 1983. Adenovirus isolates from the urines of patients with the acquired immunodeficiency syndrome. Lancet ii:1293-1295.
- Doherty, P. C., and R. M. Zinkernagel. 1975. H-2 compatibility is required for T-cell mediated lysis of target cells infected with lymphocytic choriomeningitis virus. J. Exp. Med. 141:502-507.
- Fickett, J. W. 1982. Recognition of protein coding regions in DNA sequence. Nucleic Acids Res. 10:5303-5318.
- 11. Flomenberg, P. R., M. Chen, and M. S. Horwitz. 1987. Characterization of a major histocompatibility complex class I antigenbinding glycoprotein from adenovirus type 35, a type associated with immunocompromised hosts. J. Virol. 61:3665–3671.
- Flomenberg, P. R., M. Chen, G. Munk, and M. S. Horwitz. 1987. Molecular epidemiology of adenovirus type 35 infections in immunocompromised hosts. J. Infect. Dis. 155:1127–1134.
- Ginsberg, H. S., J. Valdesus, R. Horswood, R. M. Chanock, and G. Prince. 1987. Adenovirus gene products affecting pathogen-

esis, p. 322–326. *In* R. Chanock, R. A. Lerner, F. Brown, and H. Ginsberg (ed.), Vaccines 87. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

- Gooding, L. R., L. W. Elmore, A. E. Tollefson, H. A. Brody, and W. S. M. Wold. 1988. A 14,700 MW protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. Cell 53:341-346.
- 15. Hartley, J. W., and W. P. Rowe. 1960. A new mouse virus apparently related to the adenovirus group. Virology 11:645-647.
- Herisse, J., G. Courtois, and F. Galibert. 1980. Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome. Nucleic Acids Res. 8:2173-2192.
- Herisse, J., and F. Galibert. 1981. Nucleotide sequence of EcoRI E fragment of adenovirus 2 genome. Nucleic Acids Res. 9:1229–1240.
- 18. Horwitz, M. S. 1971. Intermediates in the synthesis of type 2 adenovirus deoxyribonucleic acid. J. Virol. 8:675-683.
- Horwitz, M. S., G. Valderrama, R. Korn, and I. Spigland. 1984. Adenovirus isolates from the urines of AIDS patients: characterization of group B recombinants, p. 187-207. In M. S. Gottlieb and J. E. Groopman (ed.), Acquired immune deficiency syndrome. Alan R. Liss, Inc., New York.
- Kelly, T. J., Jr., 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.
- 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.
- Kvist, S., L. Ostberg, H. Persson, and L. Philipson. 1978. Molecular association between transplantation antigens and a cell surface antigen in an adenovirus-transformed cell line. Proc. Natl. Acad. Sci. USA 75:5674–5678.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.
- 24. Morin, J. E., M. D. Lubeck, J. E. Barton, A. J. Conley, A. R. Davis, and P. P. Hung. 1987. Recombinant adenovirus induces antibody response to hepatitis B virus surface antigen in hamsters. Proc. Natl. Acad. Sci. USA 84:4626–4630.
- Paabo, S., B. M. Bhat, W. S. M. Wold, and P. A. Peterson. 1987. A short sequence in the COOH-terminus makes an adenovirus membrane glycoprotein a resident of the endoplasmic reticulum. Cell 50:311–317.
- Paabo, S. T., T. Nilsson, and P. A. Peterson. 1986. Adenoviruses of subgenera B, C, D, and E modulate cell-surface expression of major histocompatibility complex class I antigens. Proc. Natl. Acad. Sci. USA 83:9665-9669.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- Shields, A. F., R. C. Hackman, K. H. Fife, L. Corey, and J. D. Meyers. 1985. Adenovirus infections in patients undergoing bone-marrow transplantation. N. Engl. J. Med. 312:529–533.
- Signas, C., G. Akusjarvi, and U. Pettersson. 1986. Region E3 of human adenoviruses: differences between the oncogenic adenovirus-3 and the nononcogenic adenovirus-2. Gene 50:173-184.
- Stalder, H., J. C. Hierholzer, and M. N. Oxman. 1977. New human adenovirus (candidate adenovirus type 35) causing fatal disseminated infection in a renal transplant recipient. J. Clin. Microbiol. 6:257-265.
- Tollefson, A. E., and W. S. M. Wold. 1988. Identification and gene mapping of a 14,700-molecular-weight protein encoded by region E3 of group C adenoviruses. J. Virol. 62:33–39.
- Valderrama-Leon, G., P. Flomenberg, and M. S. Horwitz. 1985. Restriction endonuclease mapping of adenovirus 35, a type isolated from immunocompromised hosts. J. Virol. 56:647-650.
- Wold, W. S. M., C. Cladaras, S. C. Magie, and N. Yacoub. 1984. Mapping a new gene that encodes an 11,600-molecular-weight protein in the E3 transcription unit of adenovirus 2. J. Virol. 52: 307-313.