JOURNAL OF VIROLOGY, Dec. 1989, p. 5455–5458
0022-538X/89/125455-04$02.00/0

Vol. 63, No. 12

The Mouse Adenovirus Type 1 Contains an Unusual E3 Region

KANAKATTE S. RAVIPRAKASH,[1] AVRAHAM GRUNHAUS,[1] MOHAMED A. EL KHOLY,[1] AND MARSHALL S. HORWITZ[1,2,3]*

Departments of Microbiology-Immunology,[1] Cell Biology,[2] and Pediatrics,[3] Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received 22 May 1989/Accepted 2 August 1989

Since the E3 region of human adenoviruses codes for a series of proteins that are probably involved in viral pathogenesis, the nucleotide sequence for a 3.6-kilobase DNA fragment in the corresponding region (map units 77 through 89) of the mouse adenovirus type 1 genome has been determined. Analysis of the sequence revealed that the genes for the fiber and for the precursor to the hexon-associated protein, pVIII, that usually flank the E3 region, are well conserved. However, many of the open reading frames contained in the E3 region of human adenoviruses between the pVIII and the fiber genes were absent from the mouse adenovirus type 1 genome.

Adenovirus genomes are organized into complex transcription units. Early regions E1 and E2 of human adenoviruses encode products that are important in cell transformation, transcriptional activation, and viral DNA replication. Early region E3 directs the synthesis of several proteins that appear to be involved in the immune response and, thereby, in viral pathogenesis. A 19-kilodalton (kDa) glycoprotein product of E3 (gp19), which binds the heavy chain of the class I major histocompatibility complex (MHC) in the endoplasmic reticulum membrane and blocks the appearance of the MHC on the cell surface, has been extensively investigated (1, 3, 4, 15, 16). Two other protein products of E3 have interesting properties: a 14.7-kDa protein from adenovirus type 5 (Ad5) E3 has recently been shown to inhibit lysis of infected cells by tumor necrosis factor (10), and a 10.4-kDa polypeptide downregulates the epidermal growth factor receptor on the cell surface (5). Because the E3 region is not essential for growth of adenoviruses in tissue culture but has been preserved among most human adenoviruses studied, it has been postulated that E3 genes control important viral pathogenesis functions in vivo. However, because human adenoviruses are relatively limited in their abilities to infect cells or animals of other species, molecular studies on in vivo pathogenesis have been difficult. Therefore, we have been characterizing the mouse adenovirus for use in murine pathogenesis studies in vivo. Detailed molecular genetic information on mouse adenovirus type 1 (MAV-1 [or FL]) has only recently begun to be available. On the basis of regions of DNA homology with the human adenovirus Ad2 and the polarity of encapsidation, Larsen et al. (14) have oriented the MAV-1 genome with its *Eco*RI fragment C at the left terminus. Ball et al. (2) have recently studied the MAV-1 E1 transcription unit by way of genomic and cDNA sequence analysis and have shown considerable differences from the human E1A and E1B transcripts.

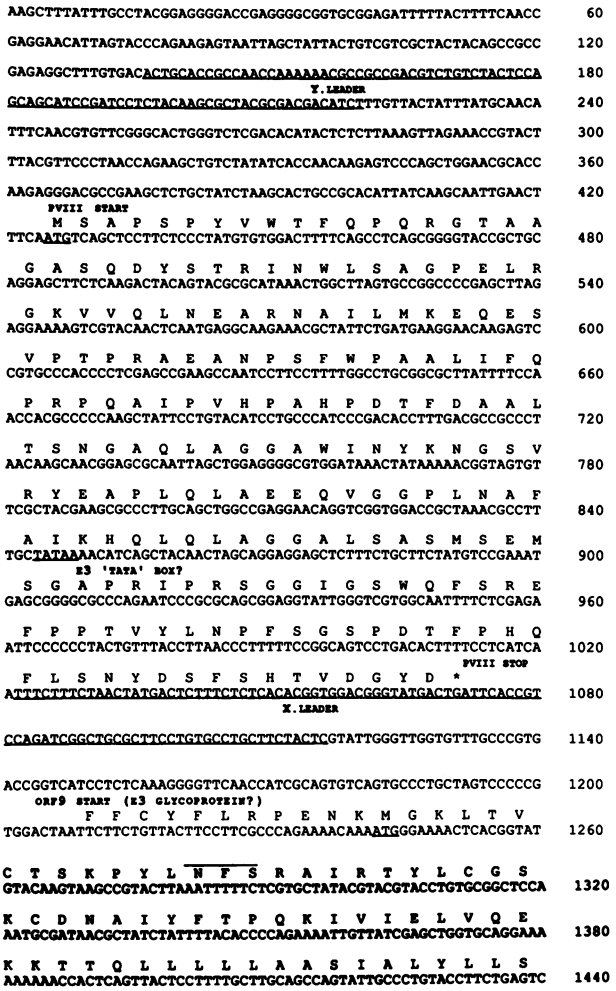
In this report, we describe the sequence of the *Hin*dIII C fragment of MAV-1 which extends between map units 77 and 89 (14). Corresponding regions of human adenovirus genomes contain multiple overlapping open reading frames (ORFs) of the E3 transcription unit flanked on the left by the gene for the hexon-associated protein precursor pVIII and on the right by the fiber gene (7). Analysis of the deduced MAV-1 amino acid sequence for homology and preservation of motif sequences has revealed that two of the ORFs within the *Hin*dIII C fragment code for pVIII and the fiber. However, in contrast to the structure of the human adenoviruses sequenced thus far, our analysis of MAV-1 DNA sequence (77 to 89 map units) shows that many of the ORFs that could code for the proteins homologous to those of the human adenovirus E3 region are absent.

The analysis of the E3 ORFs was approached by cloning and sequencing the 3.6-kilobase *Hin*dIII C fragment of MAV-1 DNA which extends from 77 to 89 map units. MAV-1, obtained from S. Larsen, Indiana University School of Medicine, Indianapolis, was propagated in mouse L cells in suspension, and the virions were purified by standard techniques used for human-adenovirus purification (11). The *Hin*dIII C fragment isolated from purified virion DNA was cloned into the *Hin*dIII site of M13mp18. A nested population of overlapping deletions of the clone was generated by the method of Dale and Arrow (8), using the IBI cyclone system. Using these deletions and selected synthetic oligonucleotide primers, both strands were sequenced in their entirety by the dideoxy-chain termination method. DNA sequence analysis was performed by using the sequence analysis package software developed by the Biomathematics Computation Laboratory at the University of California in San Francisco.

The nucleotide sequence for the *Hin*dIII C fragment of MAV-1 DNA is shown in Fig. 1. Analysis of this sequence revealed several interesting features. Initially, it resolved the positioning of the short *Xho*I L and *Hpa*I G fragments which could not be determined by the restriction endonuclease analysis of MAV-1 DNA (14). The 342-base-pair *Xho*I L fragment was located between the *Xho*I A and B fragments, and the 501-base-pair *Hpa*I G fragment was located between the *Hpa*I B and C fragments (14; Fig. 2). Further analysis of the sequence showed 9 ORFs which could code for polypeptides of 60 residues or more (Fig. 2). Only ORFs 6, 7, and 9 contained an initiating ATG codon from which a polypeptide of 6 kDa or more could be synthesized. A homology search of the predicted amino acid sequences for each of the 9 ORFs against the National Biomedical Research Foundation protein database sequence revealed that only ORFs 6 and 7 showed significant homology with any known adenovirus protein product. ORF 6 shared homologous sequences with the Ad2 hexon-associated protein precursor pVIII, and ORF 7 showed structural similarities to the Ad2 fiber protein.

The fiber protein is encoded in late transcription region L5 of the adenovirus genome and is one of the well-conserved proteins among different adenoviruses (6, 17). Although the most-conserved region within the fiber is the amino-terminal (tail) portion associated with the virion capsid, a repeating structural motif of about 15 residues in the shaft regions of Ad2, Ad3, Ad5, and Ad7 fibers has been recognized (6, 12, 13, 17). Such a repeating motif, consisting of periodically occuring

* Corresponding author.



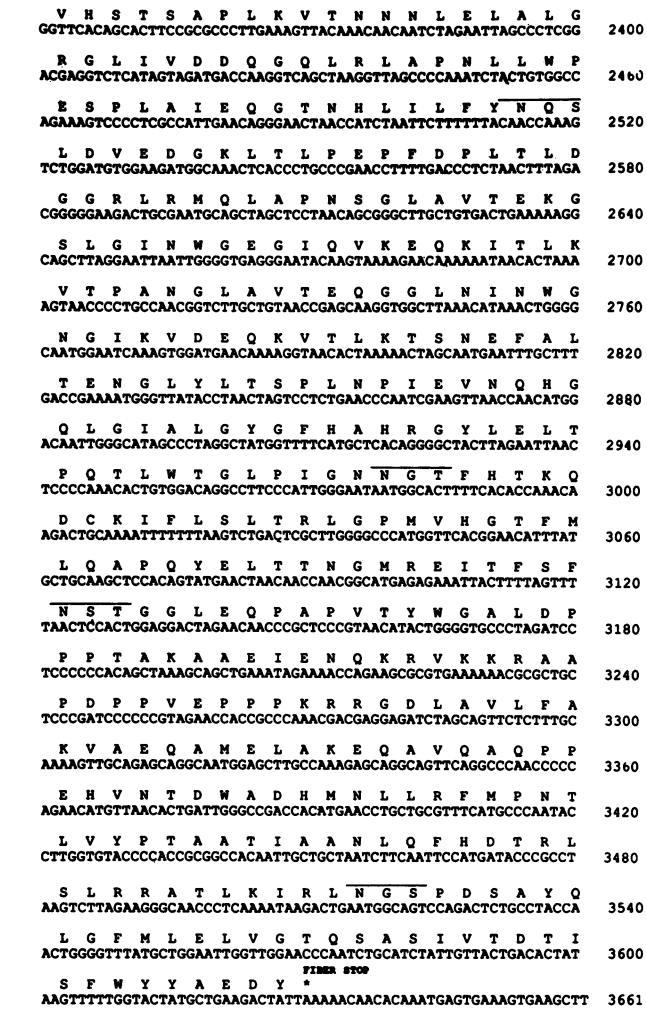


FIG. 1. Nucleotide sequence of the *Hin*dIII C fragment of the MAV-1 genome. The ORFs for the fiber, the pVIII, and the putative glycoprotein gene (ORF 9) are shown. Salient features such as the probable translational start sites, poly(A) signal, and y-, x-, and z-leader-like sequences are marked. These leaders are designated on the basis of their homology with the corresponding Ad2 sequences. Potential glycosylation sites are designated by a line above the sequence.

prolines and hydrophobic residues in the shaft region, is also preserved in the deduced amino acid sequence for the MAV-1 fiber (Fig. 3), which further emphasizes the functional importance of the proposed structure for the adenovirus fiber (12). The MAV-1 fiber is 612 residues long, resembling, in size, the fibers of Ad2 and Ad5 (6), which are longer than those of Ad3 and Ad7 (13). Although there has been no evidence for N-linked glycosylation of any adenovirus fiber, Hong et al. (13) have identified three potential sites for N-glycosylation (N-X-T/S) in the deduced amino acid sequence for the Ad7 fiber. Four such canonical sequences are also found in the MAV-1 fiber sequence (Fig. 1). In addition to the tripartite leader sequences which are contained in most of the adenovirus late mRNAs, the Ad2 fiber mRNA can also contain additional leader sequences (20). These x-, y-, and z-leader sequences overlapped region E3, which is just upstream of the fiber gene in the Ad2 genome. y-, x-, and z-leader-like sequences, with 54 to 65% homologies over approximately 100 base pairs, were also found upstream of the fiber gene in MAV-1 DNA (Fig. 1). However, the MAV-1 x, y, and z leaders are homologous to the Ad2 y, x, and z leaders, respectively.

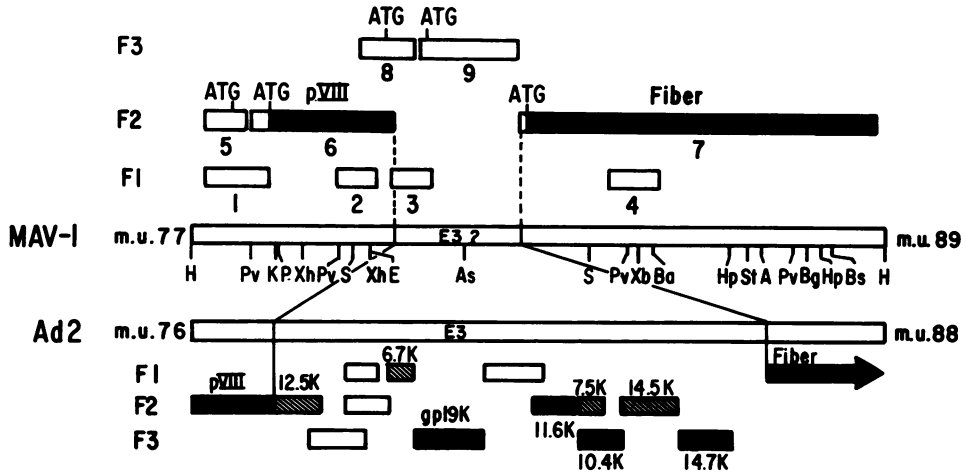


FIG. 2. Comparison of the MAV-1 ORFs between map units 77 and 89 with those of Ad2 in the corresponding region. F1, F2, and F3 are the three reading frames; the first ATG in each of the MAV-1 ORFs is shown. For Ad2, the solid bars represent proven proteins, and hatched bars represent proteins that are proposed to exist (adopted from Cladaras et al. [7]). Also shown are the following restriction sites for the MAV-1 genome: *Apa*I (A), *Asu*II (As), *Bal*I (Ba), *Bgl*II (Bg), *Bst*XI (Bs), *Eco*RI (E), *Hin*dIII (H), *Hpa*I (Hp), *Kpn*I (K), *Pst*I (P), *Pvu* II (Pv), *Sst*I (S), *Stu*I (St), *Xba*I (Xb), *Xho*I (Xh). m.u., Map unit.

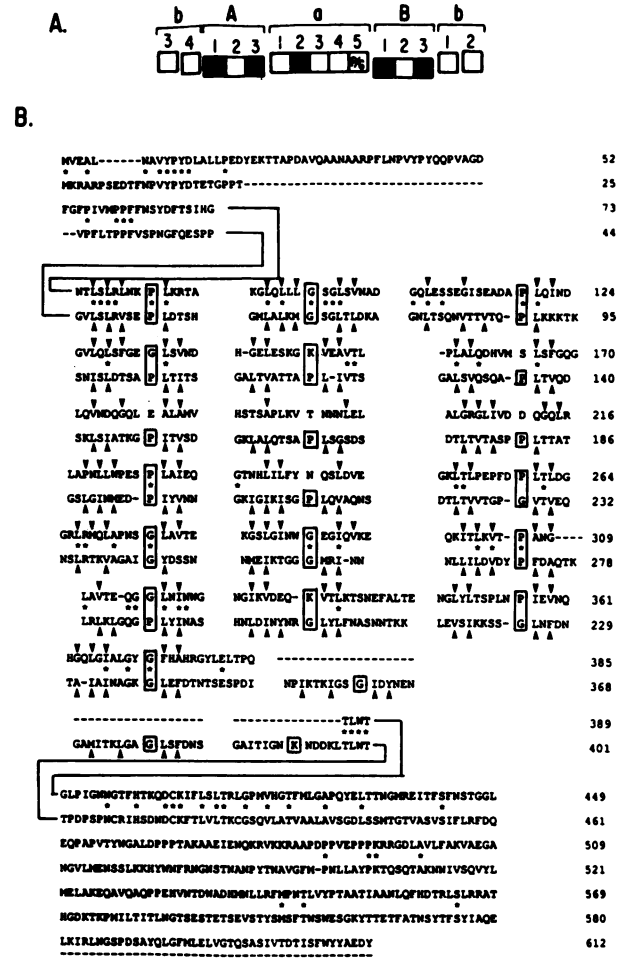


FIG. 3. Structure of fiber. (A) Repeating motif sequence based on the amino acid sequences in the shaft regions of Ad2, Ad3, and Ad5 fibers. Position a5 is almost always a proline or a glycine, although occasionally a lysine residue has been found in this position. The closed boxes are occupied by hydrophobic residues. (B) Comparison of the deduced amino acid sequence for the MAV-1 fiber with that of Ad2. The amino-terminal tail, the shaft, and the carboxy-terminal knob regions are shown separately. The shaft is represented as 22 motif sequences suggested for Ad2 fiber (12). The sequences in the motif which are aligned on the almost-invariant a5 position (P/G/K) are shown in boxes. Hydrophobic residues in the motif sequence are indicated by arrowheads. Asterisks indicate amino acid identities between the two fibers.

Not much is known about the structural and functional aspects of adenoviral pVIII. The deduced amino acid sequence for the MAV-1 pVIII exhibited an overall homology of about 40% compared with that of the Ad2 pVIII. By sequence alignment it was possible to recognize three distinct domains in the pVIII polypeptide (Fig. 4). Two highly conserved domains (55 to 60% identity plus an additional 10% conservative substitutions) at each end of the polypeptide were held together by a domain that is poorly conserved (22% homologous). It is interesting that the previously determined partial sequence for Ad3 pVIII (18) also exhibits the same degree of homology with the C-terminal domain of MAV-1.

The E3 regions of Ad2 and Ad5 were located between the L4 pVIII gene and the L5 fiber gene (9). The E3 region for most of the human subgroups (B through E) is known to encode several mRNAs, which fall into two families depending on which poly (A) site is used (7). There were at least 10 ORFs between the genes for pVIII and fiber in Ad2 and Ad5 (Fig. 2); protein products for at least 3 of these ORFs have been identified (5, 7, 19). However, ORF 9 of the MAV-1 DNA (Fig. 2) was the only ORF between the MAV-1 pVIII and fiber genes which was capable of coding for a peptide larger than 6 kDa from an initiating ATG codon located at nucleotide 1242 (Fig. 1). This ORF encoded a polypeptide of approximately 17 kDa, whose deduced amino acid sequence contains two canonical N-glycosylation sites of the form N-X-T/S (Fig. 1) and a transmembranelike stretch of hydrophobic residues near the carboxyl end. Although these features suggest that this hypothetical MAV-1 protein may be a homolog of the human adenovirus E3 glycoprotein, the predicted sequence differs from the human adenovirus E3 glycoprotein in that the MAV-1 ORF 7 does not contain an amino-terminal signal sequence required for insertion into the endoplasmic reticulum. Such a putative sequence could be spliced from DNA sequences upstream of the ORF. However, the five conserved cysteines and the sequences just

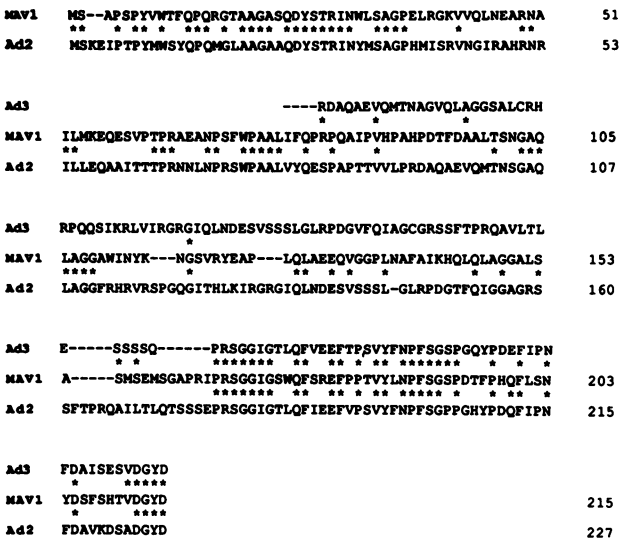


FIG. 4. Alignment of MAV-1 pVIII sequence with those of Ad2 and Ad3 pVIII. For Ad3 pVIII, the complete amino-terminal sequences have not been reported (18).

upstream of the transmembrane domain which are conserved in four of the human adenovirus glycoproteins (9) are absent from the sequences of MAV-1 ORF 9. Although there is a TATA box at nucleotide 844 and a polyadenylation signal (AA UAAA) at nucleotide 1723 which overlaps the termination codon for this ORF at 1725, it is unclear if ORF 9 codes for a polypeptide product.

These results raise further questions about the E3 region of MAV-1. The hybridization experiments of Larsen et al. (14), the E1 organization data of Ball et al. (2), and our results on pVIII and fiber presented above strongly argue that the genome of MAV-1 is colinear with that of other adenoviruses. However, the absence of all or most of the MAV-1 E3 between the L4 pVIII and the L5 fiber genes suggests either critical and complex splicing in the region to generate the many homologs of the human adenovirus E3 ORFs or deletion of functional coding sequences in this region. It is also interesting to note that the MAV-1 genome is smaller (31 kilobases) than the 36 kilobases of human adenoviruses. Although the nonessential regions of the MAV-1 genome may have been inadvertently deleted during some of the earlier passages in tissue culture, this possibility is less likely since DNA prepared from an earlier passage (obtained from Janet Hartley; passage 22 of October 1960) showed an identical *Hin*dIII pattern of digestion (data not shown). A previous report of an E1A splice downstream of MAV-1 E1B (2) suggests that novel splicing may be possible for the MAV-1 E3 transcription unit. If the MAV-1 contains no functional E3, it should be possible to construct recombinant MAV-1 containing the well-characterized Ad2, Ad3, Ad5, or Ad35 E3 glycoprotein gene to study its function in the murine system in vivo. By taking advantage of the well-characterized genetics of the murine class I MHC system and the fatal adenovirus FL disease in many mouse strains, it should be possible to modulate the levels of MHC on the cell surface and to determine the in vivo effect on pathogenesis. In addition, comparable effects of the Ad2 10.4- and 14.7-kDa E3 proteins, which interact with other host molecules potentially affecting pathogenesis, can be studied in the murine system to assess their functional importance in an in vivo system.

This work was supported by Public Health Service grant RO1-AI 27199 and Core Cancer grants PO1-CA 13330 (M.S.H.) and 5T32-CA 09060 (A.G.) from the National Institutes of Health. K.S.R. is a fellow of the Cancer Research Institute, New York, N.Y.

We thank Clayton Beard, Hannah Wooley, and Katherine Spindler for sharing some of their MAV-1 sequence data before publication. We also thank Gina DeFilippi for assistance in preparing the manuscript and Jeffrey Engler for critical appraisal of the work and helpful suggestions.

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 adenovirus to evade immune surveillance. Cell **43**:215–222.
2. **Ball, A. O., C. W. Beard, S. D. Redick, and K. R. Spindler.** 1989. Genome organization of mouse adenovirus type 1 early region 1: a novel transcription map. Virology **170**:523–526.
3. **Burgert, H.-G., and S. Kvist.** 1985. An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens. Cell **41**:987–997.
4. **Burgert, H.-G., and S. Kvist.** 1987. The E3/19K protein of adenovirus type 2 binds to the domains of histocompatibility antigens required for CTL recognition. EMBO J. **6**:2019–2026.
5. **Carlin, C. R., A. E. Tollefson, H. A. Bradley, B. L. Hoffman, and W. S. M. Wold.** 1989. Epidermal growth factor receptor is downregulated by a 10,400 MW protein encoded by the E3 region of adenovirus. Cell **57**:135–144.
6. **Chroboczek, J., and B. Jarcot.** 1987. The sequence of adenovirus fiber: similarities and differences between serotypes 2 and 5. Virology **161**:549–554.
7. **Cladaras, C., and W. S. M. Wold.** 1985. DNA sequence of the early E3 transcription unit of adenovirus 5. Virology **140**:28–43.
8. **Dale, R. M. K., and A. Arrow.** 1987. A rapid single stranded cloning, sequencing, insertion and deletion strategy. Methods Enzymol. **155**:204–214.
9. **Flomenberg, P. R., M. Chen, and M. S. Horwitz.** 1988. Sequence and genetic organization of adenovirus type 35 early region 3. J. Virol. **62**:4431–4437.
10. **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.
11. **Green, M., and M. Pina.** 1963. Biochemical studies on adenoviral replication. IV. Isolation, purification and chemical analysis of adenovirus. Virology **20**:562–565.
12. **Green, N. M., N. G. Wrigley, W. C. Russel, S. R. Martin, and A. D. McLachlan.** 1983. Evidence for a repeating cross-sheet structure in adenovirus fiber. EMBO J. **2**:1357–1365.
13. **Hong, J. S., K. G. Mullis, and J. A. Engler.** 1988. Characterization of the early region 3 and fiber genes of Ad7. Virology **167**:545–553.
14. **Larsen, S. H., R. F. Margolskee, and D. Nathans.** 1979. Alignment of the restriction map of mouse adenovirus FL with that of human adenovirus 2. Virology **97**:406–414.
15. **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.
16. **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.
17. **Signas, C., G. Akusjarvi, and U. Petterson.** 1985. Adenovirus 3 fiber polypeptide gene: implications for the structure of the fiber protein. J. Virol. **53**:672–678.
18. **Signas, C., G. Akusjarvi, and U. Petterson.** 1986. Region E3 of human adenoviruses: differences between the oncogenic adenovirus-3 and the non-oncogenic adenovirus-2. Gene **50**:173–184.
19. **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.
20. **Uhlen, M., C. Svensson, S. Josephson, P. Alestrom, J. B. Chattapadhyaya, U. Petterson, and L. Phillipson.** 1982. Leader arrangement in the adenovirus fiber mRNA. EMBO J. **i**:249–254.