

Identification of an Immediate-Early Gene in the Marek's Disease Virus Long Internal Repeat Region Which Encodes a Unique 14-Kilodalton Polypeptide

YU HONG AND PAUL M. COUSSENS*

*Molecular Virology Laboratory, Department of Animal Science, Michigan State University,
East Lansing, Michigan 48824*

Received 13 December 1993/Accepted 21 February 1994

Marek's disease virus (MDV) is an oncogenic avian herpesvirus whose genomic structure is similar to those of herpes simplex virus and varicella-zoster virus. Repeat regions of the MDV genome have been intensively investigated because of a potential relationship to MDV oncogenicity and abundant expression of immediate-early transcripts. In this study, a 1.6-kb immediate-early transcript was localized to the *Bam*HI-*I*₂ region by Northern (RNA) hybridization analysis. With cDNA cloning and sequencing, two cDNAs of 1.4 kb (C1) and 1.35 kb (C2) were identified. Both cDNAs are derived from spliced mRNAs spanning the *Bam*HI-H and -*I*₂ fragments. C1 and C2 use the same splice acceptors and 3' ends, but they differ at their 5' ends and utilize different splice donors. The upstream promoter-enhancer region of C1 cDNA has been defined as a bidirectional regulatory region shared by the MDV pp38 gene. Sequencing analysis shows two small open reading frames (ORFs) within each cDNA (ORF1a and ORF2 in C1, ORF1b and ORF2 in C2). Potential ORFs of the sequence have no significant homology with any known protein in the Swiss-Protein data base. DNA fragments encoding ORF1a and ORF1b were cloned into pGEX-3X vectors to produce glutathione S-transferase fusion proteins and induce antisera. In Western blot (immunoblot) analysis of MDV-infected-cell lysates, a 14-kDa polypeptide was identified by antisera against both ORF1a and ORF1b. This 14-kDa protein is expressed in cells which are lytically infected with MDV strains GA, Md11 passage 14 (oncogenic), and Md11 passage 83 (attenuated), as well as in the latently MDV-infected and transformed MSB-1 cell line.

Marek's disease virus (MDV) is a highly cell-associated avian herpesvirus which induces T-cell lymphomas and peripheral nerve demyelination (Marek's disease) in chickens (5, 11). Although MDV has been effectively controlled by vaccination with an antigenically related but apathogenic herpesvirus of turkeys, the mechanism of MDV tumorigenicity is still unknown. Recent efforts in this regard have been focused on the identification of the genes which may be responsible for, or related to, malignant transformation by MDV.

MDV was classified as a gamma-herpesvirus on the basis of its biological characterization (31). However, the overall genome structure of MDV is more similar to those of human alpha-herpesviruses (e.g., varicella-zoster virus and herpes simplex virus), consisting of unique long (*U*_L) and unique short (*U*_S) regions, each bounded by a set of inverted repeats (*TR*_L, *IR*_L, *IR*_S, and *TR*_S) (6, 15). As with other herpesviruses, MDV gene expression is coordinately regulated and sequentially ordered in a cascade fashion (26, 27, 35). Three major kinetic classes of genes are expressed as immediate-early (IE), early, and late genes. IE genes are expressed immediately upon infection and do not require de novo viral protein synthesis. Characterization of RNA transcripts isolated from MDV-infected cells treated with cycloheximide (CHX) indicates that transcripts from IE genes are clustered in repeat regions, similar to the locations of other herpesvirus IE genes (32, 35). In addition, MDV IE transcripts can be detected not only in lytically infected cells but also in transformed cell lines in which MDV infection is considered latent. To understand MDV gene expression and regulation during MDV tumor induction, we

are interested in examining IE transcripts in MDV repeat regions and investigating related IE gene products.

Repeat regions of the MDV genome have been intensively investigated because of speculation that transcripts within this region may be important in MDV-induced tumorigenicity. Several investigators have demonstrated that serial in vitro passage of virulent MDV in primary chicken embryo fibroblast (CEF) cells results in loss of MDV tumorigenicity. This attenuation was found to correlate with amplification of a specific 132-bp repeat sequence found within the MDV *TR*_L and *IR*_L (*Bam*HI-H and -D, respectively) (8, 16, 25, 36). Bradley et al. (3, 4) reported that a 1.8-kb gene family is transcribed rightwardly from the expanded region of the *Bam*HI-H fragment. According to Bradley et al., these transcripts are expressed only in oncogenic MDV and are absent or truncated in attenuated MDV. Antisense inhibition of proliferation of an MDV-derived lymphoblastoid cell line by using an oligonucleotide complementary to the putative splice donor sequence of the 1.8-kb gene family supports the hypothesis of an association between this gene family and the tumorigenic potential of MDV (21). Recently, Peng et al. (28) further developed a cDNA library from this *Bam*HI-H gene family and identified four cDNA clones. While two cDNAs of 1.9 and 2.2 kb were reported to be nonspliced transcripts, two other cDNAs of 1.5 and 1.9 kb were recognized as single-spliced transcripts spanning the *Bam*HI-H and *Bam*HI-*I*₂ fragments of MDV DNA. Protein products associated with these cDNAs or other transcripts from the *Bam*HI-H gene family have not been identified. Other transcripts that are initiated or terminated within or near the 132-bp repeat region have also been described (8). Recently, a 38-kDa phosphoprotein expressed in both lytically MDV-infected cells and MDV-transformed lymphoblastoid tumor cell lines was localized to the *Bam*HI-H

* Corresponding author.

region and found to be transcribed in a leftward direction (7, 13). A basic-leucine zipper gene, designated *meq* (identified as a homolog of the *fos/jun* oncogene family), has been mapped to the rightward region of the *Bam*HI-*I*₂ fragment, within the MDV IR_L. By using antiserum against a synthetic peptide deduced from the *meq* DNA sequence, a 40-kDa protein was detected in MDV-transformed lymphoblastoid cell lines but not in cells lytically infected with MDV strain GA (20).

In this article, we report the analysis of an IE gene localized within the MDV IR_L. By cDNA cloning and sequencing, we have determined that distinct transcripts are derived from this same gene by altered splicing patterns. In all cases examined, the 132-bp repeat region is not present in mature mRNA species. Computer sequence analysis revealed only small open reading frames (ORFs) within our cDNA transcripts, consistent with a previous report (28). Using antisera raised against fusion proteins, we demonstrated that a 14-kDa protein encoded by two small ORFs in these transcripts is expressed in cells lytically infected with both oncogenic and attenuated MDV, as well as in cells latently infected and transformed by MDV. Our results have important implications regarding the role of proteins encoded within the MDV IR_L in MDV-induced tumorigenicity.

MATERIALS AND METHODS

Cell culture and virus. Duck embryo fibroblast (DEF) and CEF cells were prepared, maintained, and infected with MDV as described previously by Glaubiger et al. (17). Cell-associated MDV strains GA passage 7, Md11 passage 16 (designated a low-passage strain), and Md11 passage 83 (designated a high-passage strain) were obtained from the Avian Disease and Oncology Laboratory (ADOL), U.S. Department of Agriculture, East Lansing, Mich., and have been described previously (42). To obtain IE RNA transcripts, 100 µg of CHX per ml was added at the time of infection. RNA was extracted 16 to 24 h after infection and CHX treatment. Early RNA was obtained by treating cells with 100 µg of phosphonoacetic acid (PAA) per ml at the time of infection. Fresh medium containing 100 µg of PAA per ml was added to infected and control cultures after 24 h. RNA was extracted after an additional 24 h of incubation in PAA. The MSB-1 cell line (a producer, expression cell line) (1a) was used as a representative of MDV-induced lymphoblastoid cell lines and was cultured in Leibovitz L15-McCoy 5A medium supplemented with 10% fetal calf serum at 41°C in a humidified atmosphere containing 5% CO₂.

RNA isolation and Northern (RNA) blot analysis. Total cellular RNA was isolated from uninfected and MDV-infected DEF cells by the guanidinium-phenol-chloroform method as described previously (9). Polyadenylated [poly(A)⁺] mRNAs were purified from total RNA by using the polyAtract mRNA kit (Promega, Madison, Wis.) according to the manufacturer's specifications.

For Northern blot analysis, 0.5 µg of poly(A)⁺ mRNA per well was loaded, electrophoresed through 1.2% formaldehyde-agarose gels, and transferred to Hybond-N nylon membranes (Amersham Corp., Arlington Heights, Ill.) essentially as described previously (33). α-³²P-labeled probes were generated by using a random-primed-labeling kit (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) as recommended by the manufacturer. Hybridization was performed under high stringency conditions (50% formamide, 3× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 5% dextran sulfate, 50 mM phosphate buffer, 5× Denhardt's solution, 0.1% sodium dodecyl sulfate [SDS], and 100 µg of salmon sperm DNA per ml at 42°C for 12 to 16 h). Transcript sizes were determined by

comparison with a 0.24- to 9.5-kb RNA ladder (Bethesda Research Laboratories, Gaithersburg, Md.) run on the same gel. RNA blots were stripped of probe DNA and rehybridized with a chicken β-actin gene probe (12). Intensities of RNA bands on autoradiographies were analyzed on an FB910 densitometer with Zeineh 1-D Videophoresis II software (FisherBiotech).

cDNA library construction, screening, and Southern blot analysis. To enrich for cDNAs derived from IE genes, poly(A)⁺ mRNAs isolated from MDV-infected DEF cells treated with CHX were utilized. Construction of a cDNA library was performed as described by Hu et al. (18). Briefly, plasmid pBluescript II KS⁺ was digested with the restriction enzyme *Eco*RV to generate blunt ends. Oligo(dT) tails were added to the blunt ends with T4 terminal transferase. After removal of the oligo(dT) tail at one end by digestion with *Sma*I, poly(A)⁺ mRNA was annealed to the oligo(dT) tail of the remaining end and cDNA was synthesized directly along the vector-primer. Following digestion of RNA and second-strand cDNA synthesis, blunt ends were generated by treatment with T4 polymerase and religated to generate recombinant cDNA clones. cDNA clones were transformed into *Escherichia coli* DH10B cells (Bethesda Research Laboratories), and the library was screened by in situ hybridization as described previously (33). Positive colonies were isolated and expanded for analysis of plasmid DNA by Southern hybridization.

Primer extension. Primer extension studies were performed essentially as described previously (2). Three oligonucleotide primers were designed on the basis of cDNA sequences described in Results. Oligonucleotide primer P1 is 5'-AGGAA ATATATCGGGGTACGGCCGT-3', oligonucleotide primer P2-A is 5'-ATGGAAAGTGGGTCCGCAGTCAATG-3', and primer P2-B is 5'-GTCAATGCATCCGGGGTTCGTTCCC A-3'. The primers were 5' end labeled with [γ-³²P]ATP and annealed to 30 µg of total RNA isolated from uninfected and MDV-infected DEF cells for 90 min at 65°C. After cooling at room temperature for 30 min, reverse transcription was carried out at 42°C for 1 h. Reactions were terminated by phenol-chloroform extraction and ethanol precipitation. Precipitated nucleic acids were resuspended and analyzed on 9% polyacrylamide-7 M urea sequencing gels.

Viral genomic DNA and cDNA clone sequencing. Viral genomic DNA sequencing was initially performed on a 2.3-kb *Bam*HI-*Xba*I subfragment of the *Bam*HI-*I*₂ fragment from an MDV strain GA *Bam*HI library (kindly provided by M. Nonoyama). This subfragment was subcloned into pUC18 and used to generate overlapping deletions by using *Exo*III and *S1* nucleases. At the same time, *Sau*3A and *Taq*I libraries of the 2.3-kb *Bam*HI-*Xba*I fragment were constructed in M13 vectors. Nucleotide sequences of both strands were obtained by the dideoxy-chain termination method (34), using the Sequenase enzyme (United States Biochemical Corp., Cleveland, Ohio). Both forward and reverse 17-mer universal primers were used to sequence the ends of subcloned DNA. For determining the junction sequence of *Bam*HI-*I*₂ and *Bam*HI-H, a *Hind*III-*Xba*I subfragment was isolated from the *Eco*RI-F fragment of an MDV strain GA *Eco*RI library (kindly provided by R. F. Silva, ADOL, East Lansing, Mich.). This subfragment was cloned into pUC18 and sequenced in both directions as described above. For cDNA clone sequencing, the dideoxy-chain termination procedure was conducted as described for viral genomic DNA clones. T7 forward and KS reverse primers were used for the end sequences of each cDNA.

Computer analysis of DNA sequence. Nucleotide sequences were analyzed by using the Genepro program (Riverside

Scientific Enterprises, Seattle, Wash.) and MacVector 3.5 (International Biotechnologies, Inc., New Haven, Conn.). Amino acid sequences of the putative polypeptides encoded by ORFs were searched against the protein sequence data deposited in the Swiss-Protein data bases by using the Genetics Computer Group program FASTA from the University of Wisconsin (14).

Expression of GST fusion proteins and antibody production. The vector system used to express C1 and C2 ORFs in *E. coli* is plasmid pGEX-3X (Pharmacia, Alameda, Calif.), which contains a glutathione S-transferase (GST) gene under the control of an isopropylthiogalactopyranoside (IPTG)-inducible *tac* promoter. DNA fragments containing ORF1a of C1 cDNA and ORF1b of C2 cDNA were ligated into the 3' end of the GST ORF as in-frame insertions. The respective GST fusion proteins were purified by using glutathione-Sepharose 4B (Pharmacia) as recommended by the manufacturer. Purified fusion proteins were applied to a 0.1% SDS–10% polyacrylamide gel. Positions of the fusion proteins were identified by Coomassie brilliant blue staining. Sizes of fusion proteins were estimated by comparison with protein molecular weight standards run on the same gel.

New Zealand White rabbits were initially immunized with 400 to 500 µg of purified fusion proteins in Titer-Max adjuvant (CytRx Co., Norcross, Ga.). The rabbits were boosted with the same amount of protein in Titer-Max every 4 weeks and bled 10 days after the final boost. The immune sera were preabsorbed with GST carrier protein to remove cross-reacting antibody. Antibody titers were determined by enzyme-linked immunosorbent assay (ELISA) as described previously (37).

Western immunoblot analysis. Cultured cells were lysed with triple-detergent lysis buffer (33) and separated on 12.5 or 15% polyacrylamide gels containing 0.1% SDS. Proteins were electrophoretically transferred to nitrocellulose filters. Immune detections were performed with an Amersham ECL Western blot kit according to the manufacturer's specifications. The filters were blocked with 5% nonfat milk and probed with anti-ORF1a and -ORF1b sera at a 1:100 dilution. Donkey anti-rabbit immunoglobulin conjugated with horseradish peroxidase was used as the second antibody. Proteins were detected by luminescence reagent and exposed to X-ray film. Protein sizes were estimated by reference to prestained protein molecular weight standards (Bio-Rad, Richmond, Calif.) run on each gel.

Nucleotide sequence accession number. The nucleotide sequences of the C1 and C2 cDNAs reported in this paper have been given GenBank accession numbers L26394 and L26395, respectively.

RESULTS

Detection of an IE transcript within the *Bam*HI-I₂ region. Previous studies have shown that MDV IE genes are located mainly in repeat regions, but there have been conflicting reports on the sizes and numbers of transcripts from these regions (3, 4, 8, 26, 28, 35). To avoid problems associated with unprocessed precursor transcripts and introns, poly(A)⁺ mRNAs isolated from uninfected DEF cells, MDV-infected DEF cells, and infected DEF cells treated with CHX or PAA were purified for Northern blot analysis. Initial studies revealed that the *Bam*HI-I₂ fragment hybridized to a 1.6-kb transcript. This 1.6-kb *Bam*HI-I₂ transcript was abundantly expressed in CHX-treated cells (enhanced 3.5-fold over that in untreated MDV-infected cells as normalized relative to a chicken actin probe) and was identified in PAA-treated cells after longer exposure (data not shown), suggesting that it is

expressed with IE kinetics (Fig. 1A). To verify the location of this 1.6-kb IE transcript, various subfragments of the 5.1-kb *Bam*HI-I₂ fragment were utilized as probes in Northern hybridizations. These studies confirmed localization of the 1.6-kb transcript to the leftward *Bam*HI-*Xba*I region of *Bam*HI-I₂, juxtaposed to the right region of the *Bam*HI-H fragment (Fig. 1B). Thus, this 1.6-kb transcript may be related or identical to the 1.8-kb transcripts which hybridize to the *Bam*HI-H fragment (3) and is also consistent with the 1.7-kb transcript of *Bam*HI-I₂ identified by Peng et al. (28). Smaller RNA species (1.0, 0.5, and 0.4 kb) expressed with IE kinetics were also observed in Northern hybridizations along with the 1.6-kb transcript (Fig. 1A and B). As described below, smaller RNA species are most likely derived from undegraded intron RNAs.

Viral genomic DNA sequence of the *Bam*HI-*Xba*I region from an MDV *Bam*HI-I₂ fragment. Data from Northern blot hybridizations indicated that an IE transcript was encoded within the leftward *Bam*HI-*Xba*I region of the *Bam*HI-I₂ fragment. This region of MDV DNA was subcloned and completely sequenced as described in Materials and Methods. Despite abundant transcription in infected cells, no large ORF was found in either direction in this region, although several smaller ORFs (<100 amino acids) were identified (data not shown). To confirm the accuracy of sequence data and determine the junction sequence between the *Bam*HI-H and *Bam*HI-I₂ fragments, a *Hind*III-*Xba*I subclone derived from the *Eco*RI-F fragment was also sequenced. As before, no continuous large ORFs were observed (data not shown). Two possibilities were considered to explain a lack of ORFs within an abundantly transcribed region: (i) transcripts from the *Bam*HI-I₂ region may function similarly to latency-associated transcripts in herpes simplex virus type 1-infected neurons (30, 38), or (ii) RNAs encoded within this region are extensively spliced. Data from reports by Peng et al. (28), Bradley et al. (3, 4), and Chen and Velicer (8) suggested that the latter possibility was more likely. However, a protein product from this region had not been identified by any of those groups.

Isolation of cDNA and cDNA sequence analysis. In an attempt to distinguish between the possibilities described above, a cDNA library was constructed by using poly(A)⁺ mRNA from DEF cells infected with MDV strain GA and treated with CHX. For the construction of this library, we considered both the reported sequence of *Bam*HI-H and our own sequence of the MDV *Bam*HI-I₂ fragment. Both sequences contain many stretches with high AT contents. The use of oligo(dT) to prime cDNA synthesis in this area could lead to multiple starts within transcripts. For this reason, we chose a modified Okayama-Berg procedure (18) as described in Materials and Methods. This procedure is less likely to support priming within transcripts. The resulting cDNA library was screened with a 5.1-kb *Bam*HI-I₂ fragment by in situ hybridization. Three cDNA clones (C1 through C3) were isolated. Insert sizes for these clones were estimated to be 1.4 kb for C1, 1.35 kb for C2, and 0.7 kb for C3 by Southern blot analysis. The three clones were sequenced on both strands, and the sequences were compared with MDV genomic sequence data. C1 and C2 were identified as being derived from different spliced RNAs, and C3 was a partial cDNA clone. The C1 and C2 transcripts both extend from *Bam*HI-H into *Bam*HI-I₂ in a rightward direction (Fig. 2). C1 and C2 start from different 5' initiators and have different splice donor sites within the *Bam*HI-H region, but they share the same splice acceptor site and 3' end within the *Bam*HI-I₂ region (Fig. 2).

The complete nucleotide sequences of C1 and C2 are presented in Fig. 3. The C1 cDNA clone has an insertion of 1,295 bp before the poly(A) tail. The 5' 38 bp of C1 is identical

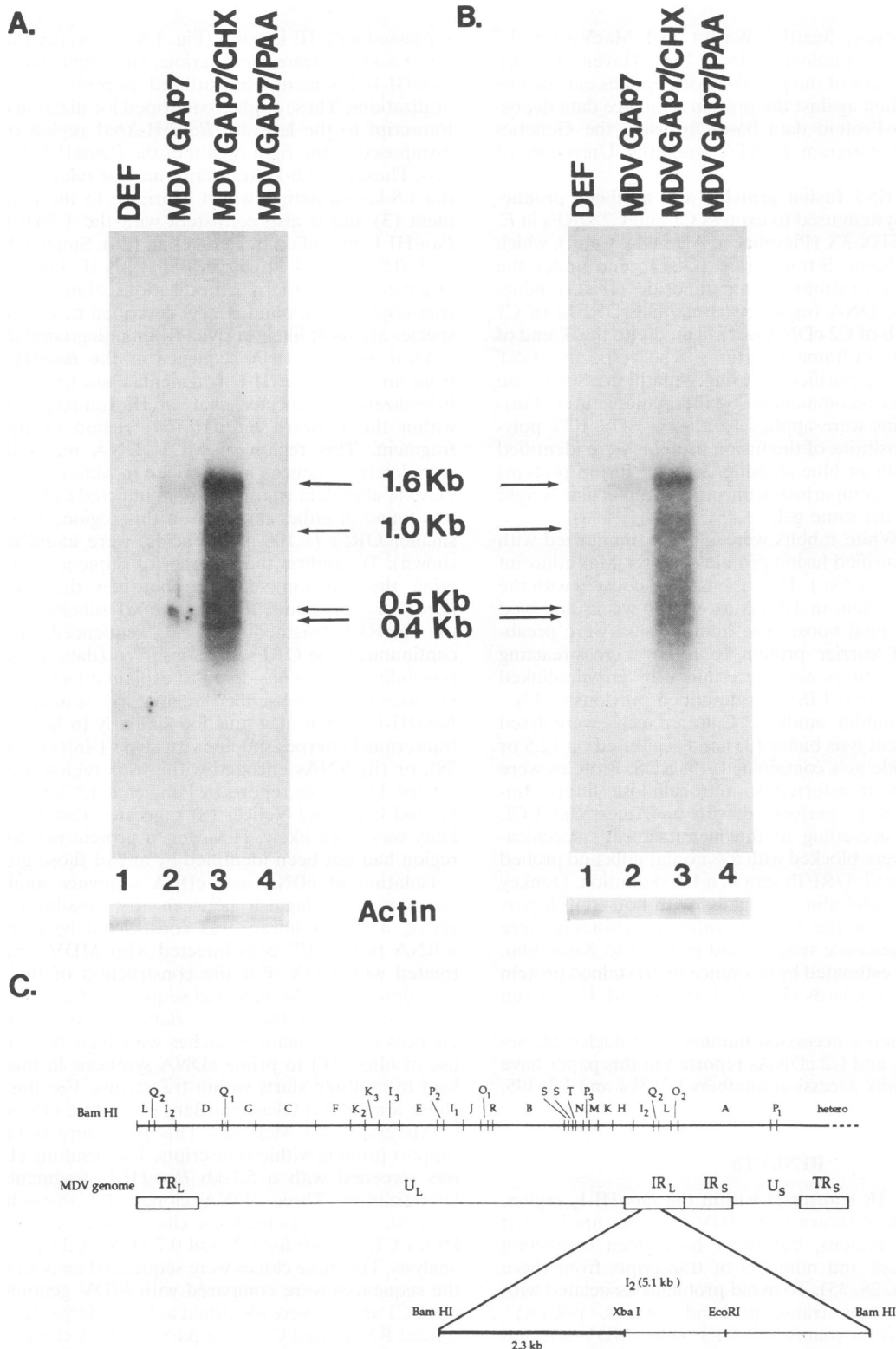


FIG. 1. Northern blot hybridization to identify IE gene transcripts in MDV *BamHI-I₂*. (A) MDV strain GA passage 7 (GAP7) was propagated in DEF cells. To obtain IE and early RNA transcripts, CHX or PAA was added at the time of infection and was present during the entire culture process. Poly(A)⁺ mRNAs were isolated and electrophoresed in 1.2% formaldehyde-agarose gels, transferred to Hybond-N nylon membranes, and hybridized with the 5.1-kb MDV *BamHI-I₂* fragment. Lane 1, uninfected DEF cells; lane 2, DEF cells infected with MDV GAP7; lane 3, DEF cells infected with GAP7 and treated with CHX; lane 4, DEF cells infected with GAP7 and treated with PAA. Sizes of various transcripts were determined by comparison with a 0.24- to 9.5-kb RNA ladder. (B) The same membrane used for panel A was stripped and reprobed with the 2.3-kb *BamHI-XbaI* subfragment of MDV *BamHI-I₂*. (C) Schematic representation of MDV genomic structure with its *BamHI* sites (16). The enlarged portion is a restriction map of the *BamHI-I₂* fragment, with the heavy line representing the 2.3-kb *BamHI-XbaI* subfragment probe.

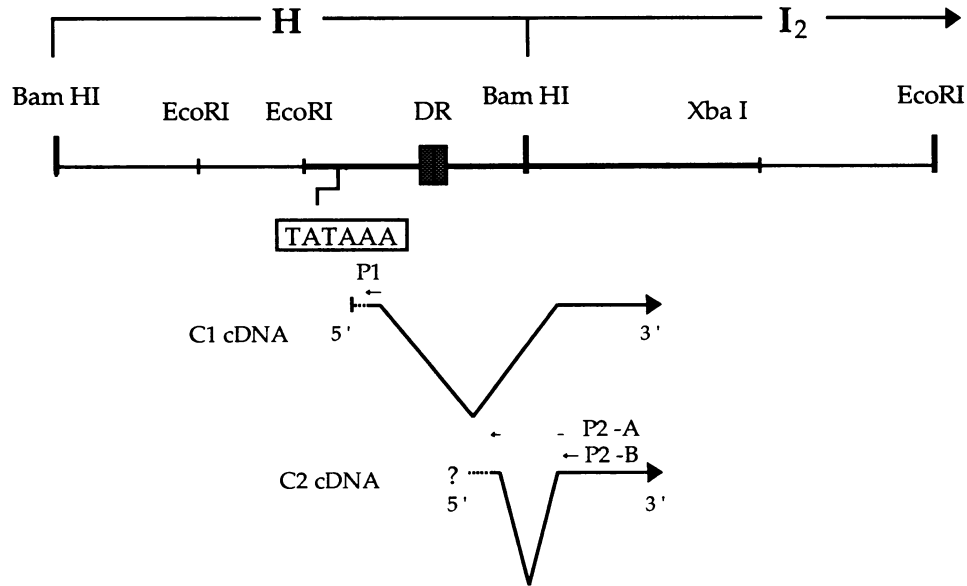


FIG. 2. Schematic representation of the locations and structures and primer extension analysis of C1 and C2 cDNAs. A partial restriction map of the MDV *Bam*HI-H and -I₂ fragments is presented, with a putative TATA sequence for the C1 transcript in a horizontal box. The dark line represents the region from which cDNAs C1 and C2 are derived. Shaded boxes represent the 132-bp direct repeat (DR) region. The locations of C1 and C2 cDNAs are shown with the approximate positions of their respective introns. Arrowheads indicate the directions of transcription for C1 and C2. Oligonucleotides used for primer extensions were designed on the basis of the cDNA sequence. P1 primer was complementary to C1 cDNA positions 14 to 39, just upstream of the intron. Two primers, P2-A and P2-B, were designed for the C2 transcript. P2-A is complementary to C2 cDNA sequence positions 29 to 53, 10 bp downstream of the splice acceptor. P2-B is complementary to nucleotides 11 to 35 of the C2 clone and crosses over to the spliced intron. Primers P1, P2-A, and P2-B are depicted with small arrows. Dashed lines indicate sequences extended by primer extension analysis.

to nucleotides 701 to 738 of a previously published *Bam*HI-H sequence (3). Beginning with nucleotide 39, the C1 sequence was aligned with the *Bam*HI-I₂ genomic sequence from nucleotide 224 to 1489. Sequence alignment revealed the presence of a 2.3-kb intron, which encompasses the 132-bp repeat region. A potential TATA sequence was found 42 bp upstream of the 5' end of C1, and a putative polyadenylation signal sequence (ATTAAA) is located 17 bp upstream of the 3' end of our C1 cDNA clone. This spliced transcript utilized a splice donor, GAAGGA, beginning at nucleotide 739 of the *Bam*HI-H sequence, and a splice acceptor, ATCGTTGCAG, at nucleotides 214 to 223 of the *Bam*HI-XbaI subfragment of *Bam*HI-I₂.

The C2 cDNA contains a 1,275-bp insert. The beginning 18 bp was identical to nucleotides 2071 to 2088 of the *Bam*HI-H sequence (3). The remaining sequence of C2 was aligned to the *Bam*HI-I₂ fragment, which is identical to C1 cDNA. C2 cDNA initiates 512 bp downstream of the 132-bp repeat region and ends at the same 3' site as C1 cDNA. The C2 transcript thus contains a 1.0-kb intron. This intron employs a splice donor, GTATGC, located at nucleotides 2088 to 2093 of the *Bam*HI-H fragment, and the same splice acceptor used by the C1 transcript. C2 cDNA therefore appears to be analogous to cDNA 3 identified by Peng et al. (28).

Despite significant attempts, we were unable to identify any cDNA clones which initiated within the 132-bp region of MDV *Bam*HI-H.

Transcriptional mapping of C1 and C2 cDNAs. In order to determine the precise 5' ends of our cDNA clones, primer extension experiments were conducted. Three oligonucleotide primers were designed on the basis of our cDNA sequence. The positions and directions of these primers are schematically

depicted in Fig. 2. The first primer (P1), 5'-AGGAAATATATCGGGGTACGGCCGT-3', was complementary to C1 cDNA sequence positions 14 to 38, which are just upstream of the intron. The 42-bp extension observed in reactions with P1 and RNA isolated from MDV GA-infected cells (data not shown) indicated that the C1 transcript starts 13 bp downstream of a putative TATA sequence within *Bam*HI-H. This location is consistent with the start site of the 1.8-kb gene family described by Bradley et al. (3). An ATG translation start site was found 37 bp downstream of the putative TATA sequence. However, the optimal context surrounding AUG (A at -3 position and G at +4 position) was not found (22). The TATA sequence, upstream of the C1 start site, belongs to a putative promoter-enhancer region containing a variety of potential transcription-regulatory elements (3, 13). This promoter-enhancer region has been cloned into a chloramphenicol acetyltransferase reporter plasmid in both orientations in our laboratory. Transient expression assays indicate that this region is, in fact, a bidirectional promoter activated by infection with MDV (1).

The second primer (P2-A), 5'-ATGGAAAGTGGGTCCGCAGTCAATG-3', is complementary to C2 cDNA sequence positions 29 to 53, which are 10 bp downstream of the splice acceptor. The third primer (P2-B), 5'-GTCAATGCATCCGGGTCGTTCCCA-3', is complementary to C2 cDNA nucleotides 9 to 33 and spans the exon-intron junction. Primer extension bands of 135, 99, 77, and 54 bp in length were observed by using P2-A and P2-B primers hybridized to RNAs isolated from MDV GA-infected cells (data not shown). The 77-bp band matched the C1 initiation site. It is difficult to determine which extension product represents the actual initiation site of C2 RNA, since we can not exclude the presence

A).

1 CTCTCAGAATGGCAGCGCCGATCCCGGATATAATTTCTACCCCGGATGCATGACTGCGGACCCACTTTCCATCTCGAAACGGATACTGCGACAACCTAGG
 S Q N G T A D P R Y I S Y P G C I D C G P T F H L E T D T A T T R
 (G) (C) (H)
 101 AACCGAACATCAAGTTTTGGTGATCGCATCAGAAGTTTTTGTGAGAGGGCTCGCTCCTTAATTTCCAACCTTTGTAACATGGCGTAGTAGAAAACGAGAGCT
 N G T S S F G D R I R S F C E R A R S L I S N F V T W R S R N E S C
 (G) (H) (H) (H) (H) (G) (C)
 201 GTGAGGTTCTGGCAGAGATTCACAAGAGAAAGAAGTGGAAATCGACCCCTCGAATCCAGTATAAATAGTAGCTAGGCGGGATAATGAGTCGCTGTTTGCA
 E V L A E I P Q E K E V E S T L *
 301 CATTATCAAAGCTACGCATTAGATAACTGCAGAAAGACGCTGCGTATAGTTATGTATTTAGAATACGCTCTGTATATACGCACGAACATATAAGTCTGT
 401 AAGAAATGTAATGCTTCGTACAGATCAGTGTTTATTTGAAGTTCAACCGTATGAAAATTTGAGTATACCTAGAATTAGCAGGATAATTGGCATCTCAAATTTCT
 501 CGAGGCTTTTTTTTTTTTGACACATTCACCTACCGAAAAGTGTGTATGCGATTCGCTTACCCTTCCCAACTTTCTCTGTCGGTCTGTGTATATTGAGG
 601 CCGGTGTATGTAGAGAGTCTACGATCTTCGATCTCCCTTCGGATCACATGGAGCGGAGATGTTGTAGGGTTCGAGAGGGGTGAGACCTAAACATGCGATC
 701 GCATGCATGGTGGAAACAGGATTTGGCCGTTGTAGCATACAAGCAGTACACATGGCGAAAAGTTTGGCGTCCGCTGTTTCGGTGTGACGTCATGTTTAGGTTTG
 M W N T I G R C S I Q A V H M A K V C R P P V R C D V M F R F E
 (H)
 801 AGCATGTAAGAAAAATGGAACCTGTTAACTCTAAAAAGAAGTATCTCGCCCATTTGTATCATTCGGCGGTGGGAAAATAGGTAATAGGAAAAATCATACT
 H V R K M E L L T L K R S I S P H L Y H S A V G N I G N R K N H T
 (H) (H) (H) (G)
 901 CTAGCTCAGGCTTCTTCGTCTCGCATCGCAGAAGTGTCTGGAGACGCGCAAAGAAGGCTGGCGCATTCGATATAGTTTGCAGCCAAATGCTTTGGTCCCG
 Y V R L L R L R S Q K C L E T R K E G L A H S D I V C S Q C L V R
 (H)
 1001 GGAAATCGGGAATCGGAGCCGATTAATCGGATAATACCGAAGCAAAGGGGATAACTTCCTTTGTTTACATAGAAATGTATGACCGGACACGCAACACCGAAAATCC
 G I G I G A D Y R *
 1101 TGTCCAAAGCTTTTGTGCGGAAATAGAACGGATATGACGCTTCTGTAACAGATTCGCGAAGTACGCGGTCGCAGACGAAACCGGACGTTGATAGCGGG
 PA
 1201 GACGTTCAATGTCCTTTGTTTTCATAATTTTGAATGTATGTATATTTTCAGCCCTCATCTGTAATTCGGTCGAGCAATTAAGAGGTTACGGAT

B).

1 CGGCTCTGGGTGGGAACGACCCCGGATGCATGACTGCGGACCCACTTTCCATCTCGAAACGGATACTGCGACAACCTAGGAAACGGAACATCAAGTTTTTGG
 A L G G N D P G C I D C G P T F H L E T D T A T T R N G T S S F G
 (G)
 101 TGATCGCATCAGAAGTTTTTGTGAGAGGGCTCGCTCCTTAATTTCCAACCTTTGTAACATGGCGTAGTAGAAAACGAGAGCTGTGAGGTTCTGGCAGAGATT
 D R I R S F C E R A R S L I S N F V T W R S R N E S C E V L A E I
 (H) (H) (H) (G) (C)
 201 CCACAAGAGAAAGAAGTGGAAATCGACCCCTCGAATCCAGTATAAATAGTAGCTAGGCGGGATAATGAGTCGCTGTTTGACACATTAACAAGCTACGCATT
 P Q E K E V E S T L *
 301 AGATAACTGCAGAAAGACGCTGCGTATAGTTATGTATTTCTTAGAATACGCTCTGTATATACGCACGAACATATAAGTCTGTAAGAATGTAATGCTTCGTAC
 401 AGATCACGTTTTATTTGAAGTTCAACCGTATGAAAATTTGAGTATACCTAGAATTAGCAGGATAATTGGCATCTCAAATTTCTCGAGGCTTTTTTTTTTTTGCA
 501 CAITGACATCTACCGAAAAGTGTGTATGCGATTCGCTTACCCTTCCCAACTTTCTCTGTCGGTCGTGTTATTGAGGCCGGTGTATGTAGAGAGTCT
 601 ACGATCTTCGATCTCCCTTCGGATCACATGGAGCGGAGATGTTGTAGGGTTCGAGAGGGGTGAGACCTAAACATGCAGTCGATGCATGGTGGAAACAGGAT
 M W N T I
 701 TGGCCGTTGTAGCATACAAGCAGTACACATGGCGAAAAGTTTGGCGTCCGCTGTTTCGGTGTGACGTCATGTTTAGGTTTGGAGCATGTAAGAAAAATGGAA
 G R C S I Q A V H M A K V C R P P V R C D V M F R F E H V R K M E
 (H)
 801 CTGTTAACTCTAAAAAGAAGTATCTCGCCCATTTGTATCATTCGGCGGTGGGAAAATAGGTAATAGGAAAAATCATACTACGTCAGGCTTCTTCGTC
 L L T L K R S I S P H L Y H S A V G N I G N R K N H T Y V R L L R L
 (H) (H) (H) (G)
 901 TGGCATCGCAGAAGTGTCTGGAGACGCGCAAAGAAGGCTGGCGCATTCGATATAGTTTGCAGCCAATGCTTGGTCCGCGAATCGGGATCGGAGCCGA
 R S Q K C L E T R K E G L A H S D I V C S Q C L V R G I G I G A D
 1001 TTATCGATAATACGGAAGCAAAGGGGATAACTTCTTGTGTTTACATAGAAATGTATGACCGGACACGCAACACGGAAAATCTGTCCAAAGCTTTTGTGTTG
 Y R *
 1101 CGGAAATAGAACGGATATGACGCTTCTGTAACAGATTCGCGAAGTACGCGGTCGCAGACGAAACCGGACGTTGATAGCGGGACGTTCAATGTCCTTTGTTT
 PA
 1201 TTCTAATTTATTTGAATGTATGTATATTTTCAGCCCTCATCTGTAATTCGGTCGAGCAATTAAGAGGTTACGGAT

FIG. 3. Complete nucleotide sequences of two cDNA clones. The overall length of the C1 cDNA insert (A) is 1,295 bp, and that of the C2 cDNA insert (B) is 1,275 bp. The amino acid sequences of potential ORFs are shown in single-letter code below the appropriate nucleotide sequence. ATG start codons and ATTAAG consensus sequences for polyadenylation (PA) are underlined. N-glycosylation sites (G), casein kinase consensus phosphorylation sites (C), and histone kinase consensus phosphorylation sites (H) are also noted.

of other spliced transcripts using the same splice acceptor as C1 and C2 or the presence of unspliced transcript precursors. To confirm the origin from which our cDNA clones were derived, Northern hybridization was conducted with the C1 cDNA insert as a probe. A 1.6-kb transcript was detected in poly(A)⁺ mRNA isolated from MDV strain GA-infected cells and was enhanced in mRNA from infected cells treated with CHX (Fig. 4, lanes 2 and 3). No hybridization to uninfected-cell mRNA was observed (Fig. 4, lane 1). This result indicates that the cDNA was derived from the 1.6-kb IE RNA previously

detected with a *Bam*HI-I₂ fragment probe. Consistent with sequencing and primer extension data, the C1 cDNA is 1.3 kb in length, without a poly(A) tail. Interestingly, minor transcripts of 1.0, 0.5, and 0.4 kb observed in Northern hybridizations probed with the *Bam*HI-*Xba*I subfragment of *Bam*HI-I₂ were not detected when the C1 cDNA insert was used as a probe (Fig. 4). This result suggests that these smaller RNA species are most likely derived from undegraded intron RNAs. In support of this observation, unspliced RNA precursors (3.3 and 2.6 kb) as well as the 2.3-kb intron are readily detected in

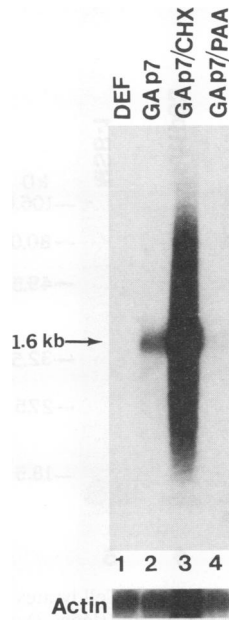


FIG. 4. Northern blot hybridization to confirm the origin of cDNA sequence. Poly(A)⁺ mRNAs from uninfected DEF cells (lane 1), DEF cells infected with MDV GA passage 7 (GAp7) (lane 2), and DEF cells infected with GAP7 and treated with CHX (lane 3) or PAA (lane 4) were hybridized with the 1.3-kb C1 cDNA insert. Transcript sizes were determined by comparison with a 0.24- to 9.5-kb RNA ladder.

Northern blots of total RNA from MDV-infected cells (data not shown).

Analysis of potential ORFs in MDV *Bam*HI-I₂-related cDNAs. Nucleotide sequences of C1 and C2 cDNA were analyzed by using the Genepro, Macvector, and Genetics Computer Group programs. Translation of the cDNAs in all six reading frames identified two small potential ORFs in both C1 and C2. In C1 cDNA, ORF1a could encode 83 amino acid residues, resulting in a putative protein of 9,307 Da. ORF1a has two potential casein kinase consensus phosphorylation sequences (amino acids 5 and 66), four potential histone kinase consensus phosphorylation sequences (amino acids 9, 42, 49, and 63), and three N-glycosylation sites (amino acids 3, 34, and 64). ORF2 could encode a 107-amino-acid protein of 12,332 Da. ORF2 has four sites of histone kinase consensus phosphorylation sequence (positions 7, 43, 44, and 71) and one potential site of N glycosylation (position 63). C2 cDNA could encode two ORFs of 76 (ORF1b) and 107 (ORF2) amino acids. The predicted size of an ORF1b-encoded protein is 8,381 Da with one potential casein kinase phosphorylation consensus sequence (position 59), three potential histone kinase phosphorylation sequences (positions 35, 42, and 56), and two potential N-glycosylation sites (positions 27 and 57). ORF2 of C2 cDNA is the same as ORF2 in C1 (Fig. 5) and corresponds to ORF-F reported by Peng et al. (28). All potential ORFs were compared with protein sequence data deposited in the Swiss-Protein data base without highly significant homology being found. However, limited homologies were found with mouse zinc finger protein ZFP-27 (mrk4) and *myc* proto-oncogene protein. ORF1a has 29% identity to mouse zinc finger protein ZFP-27 (mrk4) (21 of 78 amino acids overlapping), and ORF1b has 28% identity to ZFP-27 (14 of 50 amino acids overlapping) (10). The N terminus of ORF2 has 30% homology with the *myc* proto-oncogene exon 3 (19 of 66 amino acids overlapping) (40) (data not shown). A lack of

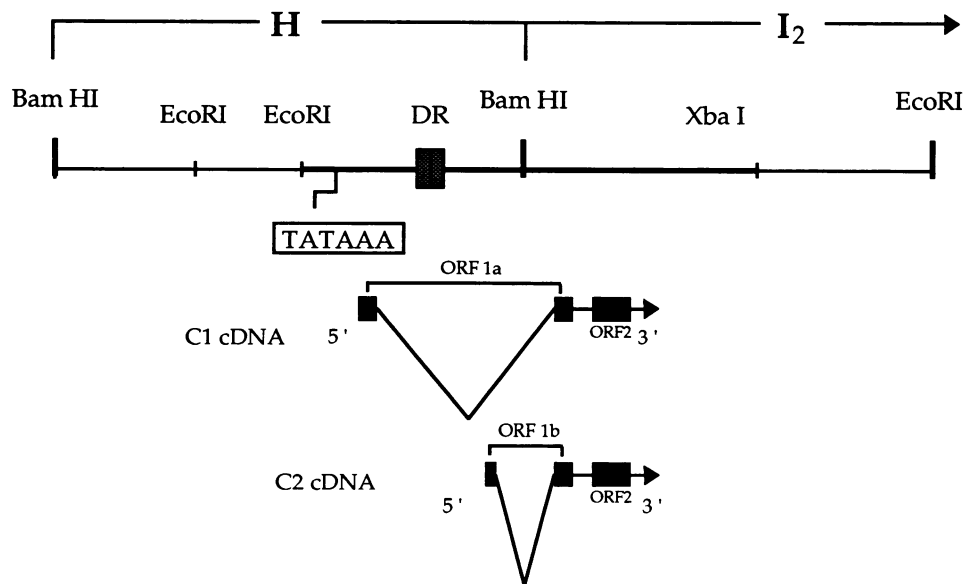


FIG. 5. Schematic representation of the positions of ORFs in C1 and C2 cDNA clones. Nucleotide sequences of C1 and C2 cDNAs were analyzed by using the Genepro and Genetics Computer Group programs. A partial restriction map of the *Bam*HI-H and -I₂ fragments is shown, with the putative TATA sequence upstream of the C1 clone highlighted. Shaded boxes represent the 132-bp direct repeat (DR) region. ORF1a of the C1 clone, which is interrupted by a 2.3-kb intron, encodes a putative 83-amino-acid polypeptide, while ORF1b of C2 (which crosses over a 1.0-kb intron) encodes a putative 76-amino-acid polypeptide. ORF2, which is located completely in exon 2, encodes 107 amino acid residues.

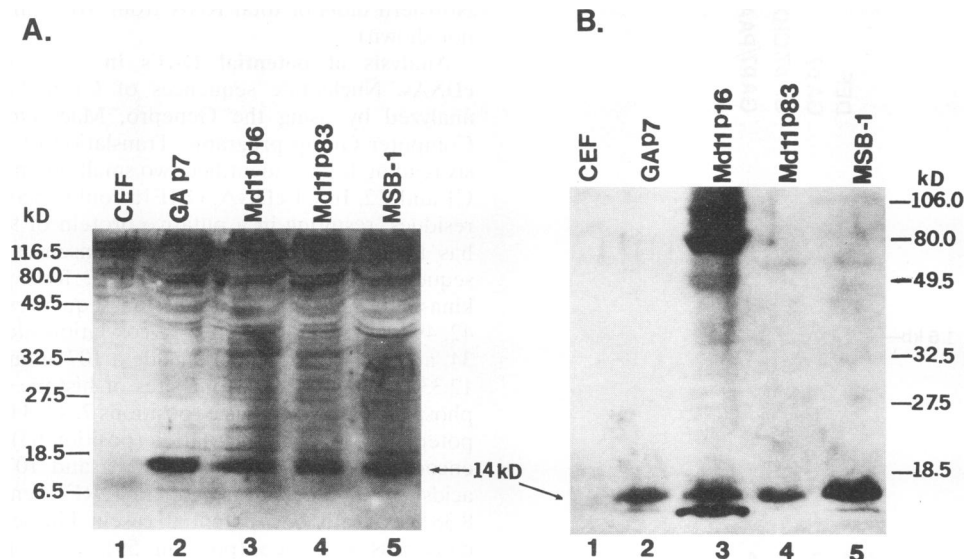


FIG. 6. Western blot analysis detection of MDV-specific proteins encoded by ORF1a (A) and ORF1b (B). Cell lysates from uninfected CEF cells (lanes 1), CEF from cells infected with MDV GA passage 7 (GAp7) (lanes 2), Md11 passage 16 (Md11p16) (lanes 3), or Md11p83 (lanes 4), and from the MDV-induced lymphoblastoid cell line MSB-1 (lanes 5) were resolved in SDS-12.5% polyacrylamide gels and transferred to nitrocellulose membranes, which was followed by immunodetection using fusion protein antiserum against ORF1a.

continuous ORFs within our cDNA sequences is consistent with the results of Peng et al. (28) and raises questions regarding the function of these abundantly expressed spliced transcripts.

Identification of a 14-kDa protein of ORF1a and ORF1b gene products. Previous analyses of transcripts encoded by the *Bam*HI-H and -I₂ regions of MDV DNA have failed to determine if any protein product is produced by these extensively spliced RNAs. Given the potential importance of this region in viral oncogenicity, identification of a protein product associated with *Bam*HI-H and -I₂ transcripts would be of considerable interest. To generate adequate amounts of protein for antibody production, DNA fragments which contain ORF1a (83 amino acids) of C1 and ORF1b (76 amino acids) of C2 were cloned into the 3' end of the GST gene of pGEX-3X vectors (in-frame insertions) and induced by IPTG to express fusion proteins in *E. coli* cells. Two fusion proteins, GST-ORF1a and GST-ORF1b, were generated and purified by glutathione-affinity chromatography. Antibodies against these proteins were raised in rabbits as described in Materials and Methods. The immune antisera were preabsorbed with purified vector protein (GST) to remove cross-reacting antibodies. After the second boost, the antiserum titer reached a 1:1,000 dilution as determined by ELISA (data not shown).

Western blot analysis was conducted to identify putative protein products of the C1 and C2 clones. Both ORF1a and ORF1b antisera detected a 14-kDa protein in CEF cells infected with MDV strain GA but not in control CEF cells (Fig. 6A and B, lanes 1 and 2). As described above, ORF1a spans a 2.3-kb intron which includes the 132-bp repeat region amplified in the genome of attenuated MDV. To determine whether the ORF1a and ORF1b gene products, identified in this study, incur any structural changes when the 132-bp region expands in attenuated MDV, we employed Md11 passage 16 as a representative low-passage (oncogenic) virus and Md11 passage 83 as a high-passage (attenuated) virus. Both fusion protein antisera detected a 14-kDa protein in lysates of cells infected with low- and high-passage Md11 (Fig. 6A and B,

lanes 3 and 4). Higher-molecular-weight proteins detected by antiserum to ORF1b (Fig. 6B, lane 3) were not consistently observed and thus likely represent nonspecific interactions. Since these gene products originate from the repeat region, which is speculated to be associated with MDV tumorigenicity, we further explored the gene products by using the MDV lymphoblastoid cell line MSB-1. A 14-kDa protein was readily detected by both ORF1a and ORF1b antisera in the MSB-1 tumor cell line without iododeoxyuridine induction (Fig. 6A and B, lanes 5). These results suggest that ORF1a and ORF1b indeed encode polypeptides in MDV-infected cells. Furthermore, oncogenic and attenuated strains of MDV produce similar amounts of this polypeptide.

DISCUSSION

Reports on the sizes, numbers, and directions of transcripts derived from MDV repeat segments (IR_L and TR_L), especially around the 132-bp expansion region, have been well documented (3, 4, 8, 26, 35). Many reports offer confusing and often conflicting data regarding initiation, splicing, and termination. Transcripts derived from the *Bam*HI-H fragment may constitute a 1.8-kb gene family, including three transcripts of 1.8, 3.0, and 3.8 kb as reported by Bradley et al. (3, 4). Four groups of transcripts are postulated to initiate or terminate within or near the 132-bp repeats (8). Three transcripts of 4.1, 3.0, and 1.9 kb in *Bam*HI-H and six transcripts of 1.8 to 5.9 kb in *Bam*HI-I₂ were reported by Schat et al. (35). Two transcripts (5 and 2.5 kb) in *Bam*HI-H and similar-sized transcripts in *Bam*HI-I₂ were reported by Maray et al. (26). Among these transcripts, the 1.9-kb transcript in *Bam*HI-H and the 1.8-kb transcript in *Bam*HI-I₂ identified by Schat et al. (35), as well as the 5- and 2.5-kb transcripts in *Bam*HI-H reported by Maray et al. (26), were cataloged as IE gene transcripts.

Here, we report that a 1.6-kb major transcript which hybridized to *Bam*HI-I₂ is highly expressed in MDV strain GA-infected and CHX-treated DEF cells. Accumulation of this 1.6-kb transcript in CHX-treated cells suggests that it is

expressed with IE gene kinetics. The decrease of this transcript in PAA-treated cells is expected, because inhibition of viral replication by PAA can limit new viral infection and subsequently reduce accumulation of viral transcripts in infected-cell cultures. Though CHX would also prevent virus spread, IE gene transcription is not limited by feedback inhibition as it would be in PAA-treated or untreated cells. Complete nucleotide sequence analysis of the 2.3-kb *Bam*HI-*Xba*I fragment of MDV *Bam*HI-I₂ and its positioning relative to previously published MDV *Bam*HI-H sequences (3) revealed no large ORFs in over 5 kb of continual viral genome. Several smaller ORFs (<100 amino acids) which could comprise exons of a spliced gene were identified within the *Bam*HI-I₂ sequence. In support of this possibility, spliced IE mRNAs have been identified in other herpesviruses. ICP0 of human herpes simplex virus, IE1 and IE2 of cytomegalovirus, BZLF1 and BRLF1 of Epstein-Barr virus, and IE RNA 1 and IE RNA 2 of bovine herpesvirus 4 (23, 24, 39, 41) all result from spliced transcripts. In cytomegalovirus, exons 1, 2, and 3 of IE1 can be ligated onto IE2 by alternate splicing to produce different mRNAs (39). A variant form of splicing which omits the middle exon of BZLF1 to produce a shortened protein has also been described for Epstein-Barr virus (23). It was therefore considered likely that the 1.6-kb transcript identified in our Northern blots was spliced.

Given the confusing and conflicting reports regarding transcription in the IR₁ region of MDV, we analyzed the available sequence data for regions which may interfere with cDNA synthesis and S1 nuclease mapping. The sequence of *Bam*HI-H from an internal *Eco*RI site to the right-end *Bam*HI site has been reported (3). Within this sequence, multiple stretches of high-AT-content DNA exist. For example, nucleotides 1701 to 1900 contain two tracts, one with 19 A residues and another with 13 T residues. These sections could dramatically affect S1 mapping and act as sites for oligo(dT) initiation during cDNA synthesis. In our poly(A)⁺ mRNA Northern blots, smaller species of RNA (1.0, 0.5, and 0.4 kb) hybridize to the *Bam*HI-*Xba*I fragment but are not observed when a cDNA insert is used as probe. These results suggest that the small RNAs may be derived from undegraded or partially degraded intron RNA which can contaminate poly(A)⁺ mRNA, perhaps because of specific AU-rich sequences. In Northern blots of total infected-cell RNA, additional bands of 3.3 and 2.6 kb are visible when the RNA is probed with an intact *Bam*HI-I₂ fragment. Interestingly, the sizes of these larger species correspond to the predicted sizes of unspliced RNAs from which our 1.6-kb transcript is likely derived. Given these facts, we chose to utilize an oligo(dT)-tailed plasmid for initiation of cDNA synthesis. In this system, the poly(dT) primer extends from a double-stranded plasmid end. Thus, annealing to internal poly(A) tracts is thermodynamically less favored than annealing to terminal poly(A) tracts (e.g., internal annealing could occur only if the RNA bent to accommodate the complementary plasmid strand). Poly(A)⁺ mRNA was also used to reduce the contributions of unspliced RNA precursors and spliced introns. In the analysis of over 3,000 clones from this library, we detected no transcripts which initiated within the 132-bp repeat region.

As this research was in process, Peng et al. (28) reported a cDNA library constructed from MDV strain RBIB-infected cells. Three classes of cDNA from the long inverted repeat region were detected, and all were transcribed rightwardly. Among them, two cDNAs are derived from spliced mRNAs. The class II 1.9-kb cDNA (no. 3 cDNA) has the same splice donor and acceptor sites as our C2 cDNA but is 663 bp longer at the 5' end and about 30 bp shorter at the 3' end. The class

I 1.5-kb cDNA (no. 4 cDNA) has the same splice acceptor as the class II 1.9-kb cDNA, as well as our C1 and C2 cDNAs, but uses a different splice donor site which is 250 bp downstream of our C1 cDNA splice donor site. Thus, there are at least three cDNAs, which utilize different splice donors but share a common splice acceptor site, present in the *Bam*HI-H and -I₂ fragments of MDV.

Two unspliced mRNAs (1.69-kb cDNA 1 and 2.2-kb cDNA 6) from the *Bam*HI-H and -I₂ regions were also reported by Peng et al. (28). Previously, Iwata et al. (19) reported the construction of a cDNA library derived from MDV strain Md5-infected cells. Twelve clones localized to the *Bam*HI-H fragment contained 1.1- to 1.8-kb insertions, and all were derived from unspliced mRNAs transcribed in a rightward direction. Iwata et al. therefore concluded that the majority of mRNA from the *Bam*HI-H region is composed of unspliced transcripts. Chen and Velicer (8) also reported four groups of unspliced transcripts which initiated or terminated in the *Bam*HI-H 132-bp repeat region in a bidirectional manner. On the basis of the results presented in this report and extensive analysis of cDNA clones, we believe that many of these cDNAs result from priming within transcripts and within relatively stable introns containing the 132-bp repeats.

The 5' end of C1 has been defined by a combination of primer extension and sequence analysis. The 5' upstream promoter-enhancer has been characterized as a bidirectional regulatory region shared with the pp38 gene. The structural arrangement of this bidirectional regulatory region shared by C1 and pp38 is similar to that of the HSV ICP4 and ICP22/47 genes (29, 43). Whether C1 and pp38 gene expression is coordinately regulated or mutually exclusive or whether their gene products are functionally synergistic will be an important issue in extending our knowledge of MDV gene regulation.

Protein similarity searches with putative C1- and C2-encoded polypeptides against the Swiss-Protein data base revealed only limited homology with zinc finger proteins and the *myc* proto-oncogene. However, these similarities were not in conservative or functional regions. The regions of similarity between ORF1a, ORF1b, and the mouse zinc finger protein ZFP-27 are not within the zinc finger motif region (10). Similarly, ORF2 has a limited homology with exon 3 of the *myc* proto-oncogene, a region outside the conservative "myc-box" (40). Therefore, we do not believe that these similarities have significance regarding potential gene functions. We did not find similarity between ORF1a or ORF1b and the mouse TLM oncogene as reported by Peng et al. (28), because the region in question was not present in our cDNA sequence as a result of altered splicing. We also did not find any similarity between ORF-F (corresponding to our ORF2) and the *fes/fps* family of kinase-related transforming proteins as reported by Peng et al. (28). It is therefore likely that any protein product encoded by these cDNAs is MDV specific.

By using predicted amino acid sequences deduced from our cDNA nucleotide sequence, fusion proteins and specific antibodies have been prepared. A 14-kDa protein was detected in Western blots of MDV-infected-cell lysates by antisera against both ORF1a and ORF1b. It is possible that these two ORFs encode proteins similar in size which differ slightly at their amino-terminal ends. ORF1a is predicted to encode 83 amino acids. An additional 2 amino acids would be added if the 5' sequence determined by primer extension is included. ORF1b could encode only 76 amino acids and is 18 amino acids shorter at the amino-terminal end than ORF-D described by Peng et al. (28). If we assume that the 5' end of ORF-D is the same 5' end for our ORF1b, ORF1b could encode 93 amino acids. Thus, the calculated sizes of polypeptides encoded by ORF1a

and ORF1b would be proximately 9.6 and 10.3 kDa, respectively. The predicted sizes of polypeptides encoded by ORF1a and ORF1b, therefore, are smaller than those of the polypeptides identified by ORF1a and ORF1b fusion protein antisera in SDS-polyacrylamide gels. According to computer data analysis, ORF1a and ORF1b both could be heavily phosphorylated (six sites for ORF1a and four sites for ORF1b) and glycosylated (three sites for ORF1a and two sites for ORF1b). Thus, posttranslational modification, such as phosphorylation or glycosylation, may be an important factor contributing to the discrepancy between predicted and apparent protein sizes. Alternatively, ORF1a and ORF1b may be exons encoding only part of a larger polypeptide. We believe this latter possibility is remote, since our primer extension data indicate that the C1 and C2 transcripts end very near the 5' ends contained in our cDNA clones and since we did not detect any larger polypeptide in Western blots.

Our data suggest that the 132-bp repeat region is removed by splicing and does not affect the protein encoded by transcripts C1 and C2 from this region of MDV DNA. These results also suggest that expansion of the 132-bp repeat may not be as critical in MDV attenuation as was previously thought. However, our data cannot rule out the possibility of subtle changes in amino acid content coincident with attenuation and expansion of the 132-bp repeat region.

Expression of the 14-kDa C1 and C2 gene products in an MDV-induced lymphoma cell line indicates that these proteins may play at least some role in maintenance of the transformed state or latent infection. Detection of the C1 and C2 14-kDa proteins in MDV-induced lymphoblastoid cells represents the third such identification of MDV antigen in latently infected transformed cells, after the identification of pp38 (7, 13) and *meq* (20). All of the proteins identified in MSB-1 cells are encoded within viral repeat regions. Whether these proteins have coordinated functions in MDV gene regulation and cell transformation remains to be investigated. The 14-kDa proteins identified in this study are derived from an IE gene. Investigation of whether these proteins can activate or inhibit expression from other MDV genes and whether they execute different roles mediated by altered splicing patterns will provide important clues regarding the nature of MDV-induced tumorigenicity and pathogenicity.

ACKNOWLEDGMENTS

This project was supported by the Michigan Agriculture Experiment Station, Research Excellence Fund, State of Michigan, and grants 90-34116-5329 and 92-8420-7430 awarded to P. M. Coussens under the Competitive Research Grants Program administered by the U.S. Department of Agriculture.

We thank Melinda Wilson and Amin Abujoub for review of the manuscript and helpful discussions. We also thank R. Southwick for his expertise with computers and assistance with preparation of graphics.

REFERENCES

- Abujoub, A., and P. M. Coussens. Unpublished observations.
- Akiyama, Y., and S. Kato. 1974. Two cell lines from lymphomas of Marek's disease. *Biken J.* **17**:105-116.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1991. *Current protocols in molecular biology*. Greene Publishing Associates, New York.
- Bradley, G., M. Hayashi, G. Lancz, A. Tanaka, and M. Nonoyama. 1989. Structure of Marek's disease virus *Bam*HI-H gene family: genes of putative importance for tumor induction. *J. Virol.* **63**:2534-2542.
- Bradley, G., G. Lancz, A. Tanaka, and M. Nonoyama. 1989. Loss of Marek's disease virus tumorigenicity is associated with truncation of RNAs transcribed within *Bam*HI-H. *J. Virol.* **63**:4129-4135.
- Calnek, B. K. 1985. Marek's disease—a model for herpesvirus oncology. *Crit. Rev. Microbiol.* **12**:293-320.
- Cebrian, J., C. K. Dierich, N. Berthelot, and P. Sheldrick. 1982. Inverted repeat nucleotide sequences in the genome of Marek's disease virus and the herpesvirus of the turkey. *Proc. Natl. Acad. Sci. USA* **79**:555-558.
- Chen, X., P. J. A. Sondermeijer, and L. F. Velicer. 1992. Identification of a unique Marek's disease virus gene which encodes a 38-kilodalton phosphoprotein and is expressed in both lytically infected and latently infected lymphoblastoid tumor cells. *J. Virol.* **66**:85-94.
- Chen, X., and L. F. Velicer. 1991. Multiple bidirectional initiations and terminations of transcription in Marek's disease virus long repeat regions. *J. Virol.* **65**:2445-2451.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156-159.
- Chowdhury, K., H. Rohdewohld, and P. Gruss. 1988. Specific and ubiquitous expression of different Zn finger protein genes in the mouse. *Nucleic Acids Res.* **21**:9995-10011.
- Churchill, A. E., and P. M. Biggs. 1967. Agent of Marek's disease in tissue culture. *Nature (London)* **215**:528-530.
- Cleveland, D. W., M. A. Lopata, R. J. MacDonald, N. J. Cowan, W. J. Rutter, and M. W. Kirschner. 1980. Number and evolutionary conservation of α - and β -tubulin and cytoplasmic β - and γ -actin genes using specific cloned cDNA probes. *Cell* **20**:95-105.
- Cui, Z., F. Lee, J.-L. Liu, and H.-J. Kung. 1991. Structural analysis and transcriptional mapping of the Marek's disease virus gene encoding pp38, an antigen associated with transformed cells. *J. Virol.* **65**:6509-6515.
- Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387-395.
- Fukuchi, K., M. Sudo, Y. S. Lee, A. Tanaka, and M. Nonoyama. 1984. Structure of Marek's disease virus DNA: detailed restriction enzyme map. *J. Virol.* **51**:102-109.
- Fukuchi, K., A. Tanaka, L. W. Shierman, R. L. Witter, and M. Nonoyama. 1985. The structure of Marek's disease virus DNA: the presence of unique expansion in nonpathogenic viral DNA. *Proc. Natl. Acad. Sci. USA* **82**:751-754.
- Glaubiger, C., K. Nazerian, and L. F. Velicer. 1983. Marek's disease herpesvirus. IV. Molecular characterization of Marek's disease herpesvirus A antigen. *J. Virol.* **45**:1228-1234.
- Hu, W., W. Kopachik, and R. N. Band. 1992. A simple, effective method to create a cDNA library. *Biotechniques* **13**:862-864.
- Iwata, A., S. Ueda, A. Ishihama, and K. Hirai. 1992. Sequence determination of cDNA clones of transcripts from the tumor-associated region of Marek's disease virus genome. *Virology* **187**:805-808.
- Jones, D., L. Lee, J.-L. Liu, H.-J. Kung, and J. K. Tillotson. 1992. Marek's disease virus encodes a basic-leucine zipper gene resembling the *fos/jun* oncogenes that is highly expressed in lymphoblastoid tumors. *Biochemistry* **89**:4042-4046.
- Kawamura, M., M. Hayashi, T. Furuichi, M. Nonoyama, E. Isogai, and S. Namioka. 1991. The inhibitory effect of oligonucleotides, complementary to Marek's disease virus mRNA transcribed from the *Bam*HI-H region, on the proliferation of transformed lymphoblastoid cells, MDCC-MSB1. *J. Gen. Virol.* **72**:1105-1111.
- Kozak, M. 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**:19867-19870.
- Lau, R., G. Packham, and P. J. Farrell. 1992. Differential splicing of Epstein-Barr virus immediate-early RNA. *J. Virol.* **66**:6233-6236.
- Leib, D. A., D. M. Coen, C. L. Bogard, K. A. Hichs, D. R. Yager, D. M. Knipe, K. L. Tyler, and R. A. Schaffer. 1989. Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. *J. Virol.* **63**:759-768.
- Maotani, K., A. Kanamori, K. Ikuta, S. Ueda, S. Kato, and K. Hirai. 1986. Amplification of a tandem direct repeat within

- inverted repeats of Marek's disease virus DNA during serial in vitro passage. *J. Virol.* **58**:657-660.
26. **Maray, T., M. Malkinson, and Y. Becker.** 1988. RNA transcripts of Marek's disease virus (MDV) serotype-1 in infected and transformed cells. *Virus Genes* **2**:49-68.
 27. **Nazerian, K., and L. F. Lee.** 1976. Selective inhibition by phosphonoacetic acid of MDV DNA replication in a lymphoblastoid cell line. *Virology* **74**:188-193.
 28. **Peng, F., G. Bradley, A. Tanaka, G. Lancz, and M. Nonoyama.** 1992. Isolation and characterization of cDNAs from *BamHI-H* gene family RNAs associated with the tumorigenicity of Marek's disease virus. *J. Virol.* **66**:7389-7396.
 29. **Preston, C. M., M. C. Frame, and M. E. M. Campbell.** 1988. A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early regulatory DNA sequence. *Cell* **52**:425-434.
 30. **Rock, D. L., A. B. Nesburn, H. Ghiasi, J. Ong, T. L. Lewis, J. R. Lokensgard, and S. Wechsler.** 1987. Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1. *J. Virol.* **61**:3820-3826.
 31. **Roizman, B., L. E. Carmichael, F. Deinhardt, G. de-The, A. J. Nahmias, W. Plowright, F. Rapp, P. Sheldrick, M. Takahashi, and K. Wolf.** 1981. Herpesviridae: definition, provisional nomenclature, and taxonomy. *Intervirology* **16**:201-217.
 32. **Roizman, B., and A. Sears.** 1990. Herpes simplex viruses and their replication, p. 849-895. *In* B. Fields and D. M. Knipe (ed.), *Fundamental virology*, 2nd ed. Raven Press, New York.
 33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
 35. **Schat, K. A., A. Buckmaster, and L. J. N. Ross.** 1989. Partial transcription map of Marek's disease herpesvirus in lytically infected cells and lymphoblastoid cell lines. *Int. J. Cancer* **44**:101-109.
 36. **Silva, R. F., and R. L. Witter.** 1985. Genomic expansion of Marek's disease virus DNA is associated with serial in vitro passage. *J. Virol.* **54**:690-696.
 37. **Sjogren-Jansson, E., and S. Jeansson.** 1990. Growing hybridomas in dialysis tubing: optimization of technique, p. 41-60. *In* H. Zola (ed.), *Laboratory methods in immunology*, vol. 1. CRC Press Inc., Boca Raton, Fla.
 38. **Spivack, J. G., and N. W. Fraser.** 1987. Detection of herpes simplex virus type 1 transcripts during latent infection in mice. *J. Virol.* **61**:3841-3847.
 39. **Stinski, M. F., C. L. Malone, T. W. Hermiston, and B. Liu.** 1991. Regulation of human cytomegalovirus transcription, p. 245-260. *In* E. K. Wagner (ed.), *Herpesvirus transcription and its regulation*. CRC Press Inc., Boca Raton, Fla.
 40. **van Beneden, R. J., D. K. Watson, T. T. Chen, J. A. Lautenberger, and T. S. Papas.** 1986. Cellular *myc* (*c-myc*) in fish (rainbow trout): its relationship to other vertebrate *myc* genes and to the transforming genes of the MC29 family of viruses. *Proc. Natl. Acad. Sci. USA* **83**:3698-3702.
 41. **van Santen, V. L.** 1993. Characterization of bovine herpesvirus 4 immediate-early RNA encoding a homolog of the Epstein-Barr virus R transactivator. *J. Virol.* **67**:773-784.
 42. **Wilson, M. R., J. T. Pulaski, R. A. Southwick, V. L. Tieber, Y. Hong, and P. M. Coussens.** 1994. Molecular analysis of the glycoprotein C-negative phenotype of attenuated Marek's disease virus. *Virology* **199**:393-402.
 43. **Wong, S. W., and P. A. Schaffer.** 1991. Elements in the transcriptional regulatory flanking herpes simplex virus type 1 oriS stimulate origin function. *J. Virol.* **65**:2601-2611.