JOURNAL OF VIROLOGY, June 1986, p. 724–732
0022-538X/86/060724-09$02.00/0

Vol. 58, No. 3

Organization of Nonstructural Genes of the Autonomous Parvovirus Minute Virus of Mice

SUSAN F. COTMORE AND PETER TATTERSALL*

Departments of Laboratory Medicine and Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06510

Received 20 December 1985/Accepted 12 February 1986

Regions of open reading frame (ORF) from the genome of the autonomous parvovirus minute virus of mice (MVM) were cloned into a procaryotic expression vector, and bacterial fusion proteins containing MVM-specific amino acid sequences were isolated. Antibodies raised against these proteins were used to immunoprecipitate viral proteins synthesized in vitro in a rabbit reticulocyte lysate translation system programmed with mRNA isolated from cells infected with MVM and a number of different parvoviruses. These studies demonstrated that: the 83-kilodalton nonstructural protein NS-1 and the 25-kilodalton nonstructural protein NS-2 have a common amino-terminal sequence which is encoded by the single ORF located between nucleotides 225 and 534 in the viral genome; the ORF located between nucleotides 1110 and 1638 is only expressed in the NS-1 protein; and the sequence encoded in a small alternative ORF between nucleotides 2075 and 2291 is expressed exclusively in NS-2. These data confirm that NS-1 is the product of the 4.8-kilobase R1 viral transcript and demonstrate that NS-2 is synthesized from the 3.3-kilobase R2 transcript which arises from the left-hand promoter at map unit 4 on the viral genome. Antibodies against the MVM fusion proteins also cross-reacted with similar proteins encoded by the viruses H-1 and LuIII, but although antibodies against the carboxy-terminal half of NS-1 cross-reacted with a similar protein in CPV, we were unable to demonstrate an NS-2 protein encoded by this virus.

Minute virus of mice (MVM), an autonomous parvovirus, has a linear, nonpermuted, single-stranded DNA genome of some 5 kilobases (kb) contained within an icosahedral protein capsid approximately 20 nm in diameter (24, 26). Its coding region is confined to one DNA strand (13), and blocks of open reading frame (ORF) span most of the viral genome, with some regions having multiple ORFs (2) (Fig. 1). The viral genome encodes two overlapping transcription units which produce three major spliced cytoplasmic mRNA species of 4.8 kb (R1), 3.3 kb (R2), and 3.0 kb (R3) (13) (Fig. 1). Transcripts R1 and R2 are synthesized from a promoter near the left-hand end of the viral genome at map unit 4, while the R3 transcript, which is the major virally coded mRNA expressed in infected cells late in infection, arises from a promoter at map unit 38 (13). All these mRNA species coterminate close to the right-hand end of the genome (13). The R3 transcript programs the synthesis of two capsid proteins, VP-1 and VP-2, of 83 and 64 kilodaltons (kDa), respectively (5, 14). The third capsid protein (VP-3, 62 kDa) is not a primary translation product but is derived by proteolytic cleavage which removes the amino-terminal region of VP-2 and which occurs only after capsid assembly and packaging of the viral genome (23). All three transcripts have a short intron sequence between 46 and 48 map units removed, and the 3.3-kb R2 transcript also has a second major intron between map units 10 and 40 which removes a large region of ORF located in the left half of the viral genome. A minor 1.8-kb (R4) transcript has also been described, but it has not been mapped on the genome (13). We have previously shown that RNA from MVM-infected cells programs the synthesis in vitro of four viral polypeptides: the two capsid proteins VP-1 and VP-2 and two nonstructural proteins of 83 kDa (NS-1) and 24 kDa (NS-2) (5). RNA selected by hybridization to a bacterial plasmid which contained the MVM sequence between nucleotides 1084 and 1659 yielded only NS-1 when translated in vitro, showing that this protein is the product of a 4.8-kb R1 transcript, but the particular mRNA species encoding NS-2 was not identified (5). Our original study (5) exploited the finding that animals infected with a particular parvovirus make antibodies which, in addition to recognizing the nonstructural proteins of that parvovirus, will cross-react with similar polypeptides synthesized by parvoviruses of different serotype. The major disadvantages in using such antibody preparations are their polyspecificity with respect to individual polypeptides and the heterogeneity of their affinities for different domains of a single polypeptide. In the present paper we used chimeric proteins produced with a procaryotic expression system to obtain antibodies against the protein sequences encoded by particular regions of ORF in the left half of the MVM genome. These antibodies allowed us to define blocks of amino acid sequence used to specify the NS-2 protein of MVM and to examine the antigenic relatedness of individual domains of the nonstructural proteins of a number of different autonomous parvoviruses. The advantages of this approach are threefold. First, the high-affinity antibodies so produced are monospecific for a particular protein segment or domain. Second, by using predetermined fragments of viral DNA and confirming the size and reading frame of the viral DNA inserts by DNA sequencing, the polypeptide sequence against which the antibodies are raised is unambiguously mapped within the viral DNA. Finally, the approach provides invaluable tools for the precise analysis of those viral gene products in vivo, for example, in determining their intracellular location and the kinetics of their synthesis and processing.

MATERIALS AND METHODS

Materials. The procaryotic expression vectors pJS413, pHK412, and pHK414 7X were obtained under license from

* Corresponding author.

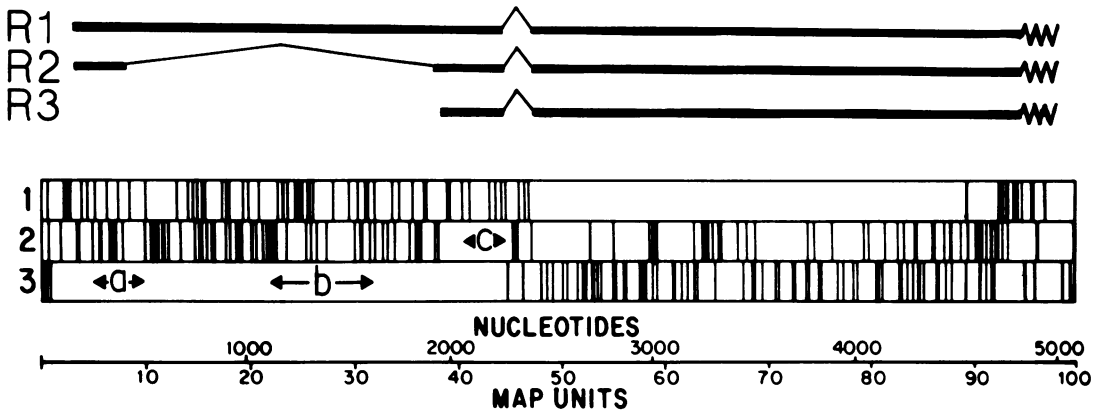

FIG. 1. Genetic map of MVM. The 5149-nucleotide-long viral genome is shown in a 3′ to 5′ orientation with a block diagram showing the translation termination codons in all three ORFs in the complementary strand. The three major cytoplasmic transcripts, R1 (4.8 kb), R2 (3.3 kb), and R3 (3 kb), are represented by thick black lines, and the thin lines indicate introns spliced out in the production of the mature message. Protein-coding regions are represented by open blocks. Sequences labeled a, b, and c on this diagram represent the nucleotide sequence and ORF expressed in the inserts of the expression plasmids pYT201, pYT202, and pYT203, respectively, constructed as described in the text. The boundaries of each MVM insert, determined by DNA sequencing, are nucleotides 225 to 534 for a, 1110 to 1638 for b, and 2075 to 2291 for c.

Molecular Genetics Inc., Minnetonka, Minn. Bacterial strains used were *Escherichia coli* NF1829 [*araD139* Δ(*araABC-leu*)*7679 galU galK* Δ(*lac*)*X74 rpsL thi* (F′ *lacI*^{q} *lacZ*::Tn*5* Y^+A^+)] and LE392F [*supE supF hsdR galK trpR metB lacY tonA* (F′ *lacI*^{q} *lacZ*::Tn*5* Y^+A^+)] and were also obtained from Molecular Genetics Inc. Restriction endonucleases and other DNA-modifying enzymes were obtained from New England BioLabs, Inc., Beverly, Mass. Reagents for DNA sequencing and radiochemicals were from Amersham Corp., Arlington Heights, Ill. pDR540, a plasmid containing the *trp-lac* hybrid or TAC promoter, was purchased from P-L Biochemicals, Inc., Milwaukee, Wis.

Construction of MVM expression constructs. (i) Procaryotic expression vector. pJS413 is one of a series of β-galactosidase-based expression vectors that have been described in detail elsewhere (21, 27–29). Sequences of interest are cloned into sites in a polylinker in the vector which links two procaryotic gene segments out of reading frame. Blocks of inserted ORF are expressed under control of the *lac* promoter as the middle part of a tripartite fusion protein as described in detail in the Results section.

(ii) MVM sequences. MVM sequences were obtained by restriction endonuclease digestion of a full-length infectious clone of MVM(p) contained in the plasmid pMM984 (9). The following sequences were then purified from agarose gels: (i) the *Hin*fI fragment between nucleotides 225 and 534; (ii) the *Pvu*II to *Xho*I fragment between nucleotides 763 and 2074; and (iii) the *Xho*I to *Nar*I fragment between nucleotides 2067 and 2291. These DNAs were then digested with exonuclease BAL 31 for various periods, and the extent of digestion was analyzed by gel electrophoresis. The Klenow fragment of *E. coli* polymerase I was then used in the presence of all four deoxynucleotides to blunt the ends of all termini, and the resulting mixtures were ligated into the *Sma*I site in the polylinker of pJS413. The position, reading frame assignment, and extent of the MVM insert in each construct analyzed in the present paper are outlined in Fig. 1.

(iii) Screening for expression constructs. Ligation mixtures were transformed into *E. coli* NF1829, a strain which lacks a functional β-galactosidase gene (29) and carries transposon *5* and *lacI*^{q} on an F factor. Bacteria carrying this F factor are kanamycin resistant (Tn*5*) and overproduce the *lac* repressor (*lacI*^{q}). Transformants were selected by growth on ampicillin-kanamycin and replica plated on to lactose-MacConkey agar indicator plates. Colonies containing an in-frame fusion protein with β-galactosidase activity are thus induced and can be identified by their red color. Figure 2A diagrams the construction and identification of one such expression clone, pYT202, containing a region of the MVM genome derived from the B fragment. Transformants giving red colonies on lactose-MacConkey agar were amplified as 1-ml cultures in yeast-tryptone broth for several hours, before being induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) overnight. Pellets were then obtained from 100-μl samples of induced cultures, boiled in sample buffer, analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) essentially according to the procedure of Laemmli (7), and stained with Coomassie brilliant blue R. Owing to their high molecular weight, the tripartite fusion proteins migrated in a region of the gel occupied by few other bacterial proteins (Fig. 2B), and this provides a ready means of identifying which bacterial clones accumulate the highest levels of fusion protein.

(iv) Introduction of amber termination codons. The construct pHK414 7X is similar to pJS413 but carries an amber termination codon (TAG) in the polylinker just downstream of the *Sma*I site (28). This termination codon can be aligned in-frame with fusion constructs engineered in the pJS413 vector by selecting a restriction endonuclease site (X) which cuts once in pJS413 outside of the polylinker and not at all in the insert and by using the *Bgl*II and *Bam*HI sites which are at the 5′ and 3′ borders, respectively, of the polylinker. Thus, ligating together the *Bgl*II-X fragment which contains the polylinker from pHK414 7X and the *Bam*HI-X fragment containing the polylinker from the expression construct produces a competent plasmid which has two in-frame copies of the polylinker linked together by a *Bam*HI-*Bgl*II fusion where the upstream polylinker brackets the sequence to be expressed and the downstream polylinker contains an in-frame amber termination codon.

(v) Expression of amber fragments. Constructs containing amber termination codons were transformed separately into NF1829 and into LE392F which carries the amber suppressor genes *supE* and *supF*. Colonies were screened for

overproduction of fusion proteins by gel electrophoresis as described above.

(vi) Replacing the *lac* promoter with the TAC promoter. pJS413 was digested with *Rsa*I, and an 864-base-pair fragment containing the Shine-Dalgarno sequence of *cro*, the polylinker, and part of *lacZ* was gel purified. *Bam*HI linkers were ligated to the termini of this fragment, and it was cleaved with *Bam*HI to yield two fragments. This mixture was then cloned into the *Bam*HI site of pDR540 (located just downstream of the TAC promoter), and constructs containing the pJS413 polylinker in the correct orientation with respect to the TAC promoter were identified. One such plasmid was then digested with *Pst*I and *Bgl*II, and the fragment containing the TAC promoter was used to replace a similar fragment containing the *lac* promoter in pJS413 (then called pJS413/TAC). The new promoter was introduced into the expression constructs by substituting a *Bgl*II-X fragment containing the TAC promoter for a similar fragment in each construct.

DNA sequencing. Constructs were sequenced by transferring the entire polylinker plus insert (a *Bgl*II-*Bam*HI fragment containing both of the junctions between the insert and the vector) into the *Bam*HI site of M13mp8 (10), selecting clones in both orientations, isolating single-stranded DNA, and using the universal primer to obtain the sequence by the dideoxynucleotide chain termination method (19). The MVM nucleotide numbers at both junctions in each construct determined in this way are detailed in the legend to Fig. 1 and in the Results section.

Purification of fusion proteins. Although fusion proteins are soluble when synthesized at lower levels, induction overnight with IPTG leads to such massive overproduction and accumulation of the fusion proteins that they become insoluble and effectively precipitate out in the bacterial cell, making purification exceedingly simple. Cultures (50 ml) of *E. coli* carrying each gene fusion plasmid were grown to an optical density at 600 nm of 1.0 before being induced with 1 mM IPTG overnight. Bacteria were then collected by centrifugation and incubated in 2 ml of 12.5% sucrose–0.15 M Tris hydrochloride (pH 8.0)–0.005 M EDTA containing 10 mg of lysozyme per ml for 30 min on ice. Samples were then freeze-thawed twice and mixed with 9 volumes of 0.05 M Tris hydrochloride (pH 8.0)–0.025 M EDTA–1% Triton X-100–1% sodium deoxycholate–1% Nonidet P-40–0.01% SDS and incubated at room temperature for 30 min. After sonication to reduce sample viscosity, the insoluble fusion proteins were collected by centrifugation at 8,000 × *g* for 20 min. Pellets were resuspended in SDS-sample buffer and further purified by preparative SDS-PAGE. Gels were briefly stained with Coomassie brilliant blue R, protein bands were excised and washed, the gels were crushed, and proteins were eluted overnight by agitation in 5 volumes of 0.005 M $NaHCO_3$ (pH 8.5) containing 0.1% SDS. After centrifugation to remove gel fragments, supernatants were concentrated with Centricon microconcentrators as described by the manufacturer (Amicon Corp., Lexington, Mass.).

Antisera. Antisera against all fusion proteins were raised in MVM-free female BALB/c mice (Jackson Laboratory, Bar Harbor, Maine) by repeated intraperitoneal injection of 100-μl samples of the fusion proteins (20 to 50 μg) emulsified with either Freund complete adjuvant (first injection) or Freund incomplete adjuvant (subsequent injections). Animals were immunized over the course of 1 to 2 months during which time they received three to four injections. Subsequently, antisera were raised against selected fusion proteins in male albino rabbits by the initial injection of approximately 200 μg of protein in Freund complete adjuvant at multiple intramuscular and subcutaneous sites, followed by repeated injections of similar amounts of protein emulsified in Freund incomplete adjuvant at multiple subcutaneous sites during the course of several months. Sera were collected 5 to 6 days after the last injection.

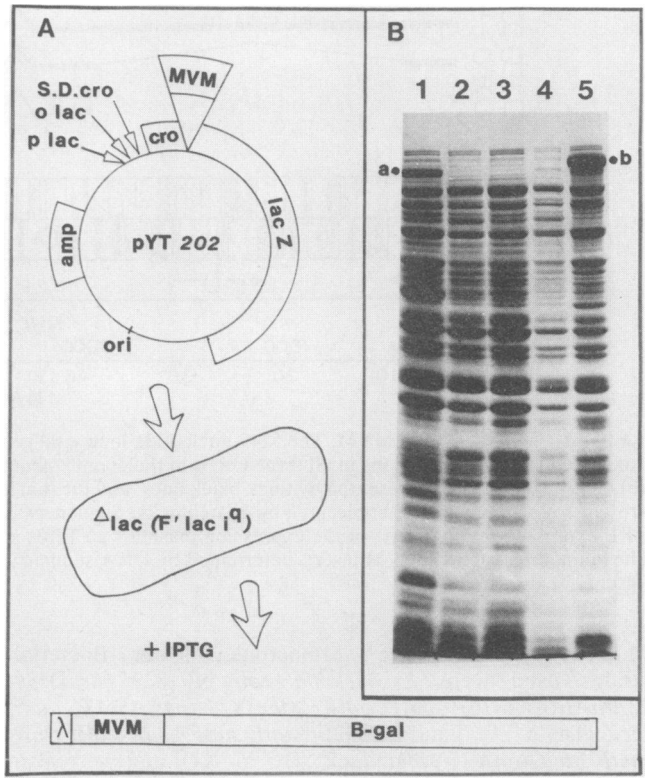

FIG. 2. Construction and identification of procaryotic expression clones. (A) The construction and analysis of pYT202 is shown as an example of a tripartite gene fusion. The region designated b (nucleotides 1110 to 1638) in Fig. 1, a BAL 31 digestion product of the *Pvu*II to *Xho*I fragment of MVM (nucleotides 763 to 2074), was inserted at the *Sma*I site of the expression vector pJS413. The resulting fusion gene (*cro*-MVM-*lacZ*) is expressed under the control of the UV5 *lac* promoter (p lac), the lac operator (o lac), and the Shine-Dalgarno sequence of the *cro* protein of phage lambda (S.D.cro). This construct was introduced into NF1829, a strain of *E. coli* which lacks an active β-galactosidase gene (Δ*lac*) and carries *lac*I^q on an F factor. Induction of the *lac* operon in this construct with IPTG gives rise to the synthesis of the tripartite fusion protein diagrammed below. The chimeric molecules contain the first 22 amino acids of λ *cro*, 176 amino acids of MVM between nucleotides 1110 and 1637 in ORF 3, and, at their carboxy termini, almost all of β-galactosidase (28). Such hybrid molecules have variable, but significant, β-galactosidase activity. (B) Coomassie-stained 10% SDS-polyacrylamide gel showing the proteins synthesized with (lanes 1, 3, and 5) and without (lanes 2 and 4) IPTG induction of bacterial clones carrying the following plasmids: lane 1, pHK412 (in which *cro* and *lacZ* are in frame and there is no insert); lanes 2 and 3, pJS413 (in which *cro* and *lacZ* are out of frame and there is no insert); and lanes 4 and 5, pYT202 (in which the *cro* and *lacZ* genes of pJS413 are joined in open reading register by insert b from MVM). The tripartite fusion protein (b) produced upon induction of pYT106 constitutes approximately 3 to 5% of the total bacterial protein and has a higher apparent molecular weight than the 116-kDa *cro-lacZ* fusion synthesized from pHK412 (a).

Cells and viruses. The prototype strain of MVM was grown

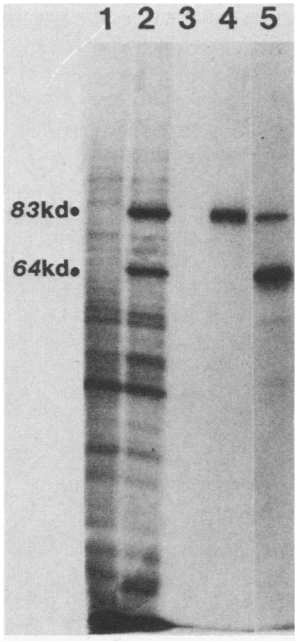

FIG. 3. Immunoprecipitation with antisera to tripartite and bipartite fusion polypeptides. The [^{35}S]methionine-labeled products of in vitro translation systems programmed with uninfected A9 cell RNA (lane 1) and MVM-infected A9 cell RNA (lanes 2 through 5) were prepared and processed as described in the Materials and Methods section. Lanes 1 and 2 show total translation products on a 10% SDS-polyacrylamide gel, and lanes 3 through 5 show proteins immunoprecipitated by antisera raised in mice against the tripartite fusion protein specified by pYT202 (lane 3), the bipartite amber fragment specified by pYT202Am-TAC (lane 4) (see Fig. 4 and text), and purified empty MVM capsids (lane 5). kd, Kilodalton.

in the mouse L-cell derivative A9 ouab^{r}11, and H-1 and LuIII were grown in the human simian virus 40-transformed fibroblast line 324K, as previously described (22). Canine parvovirus (CPV) was grown in Crandall feline kidney cells (CFK), and bovine parvovirus (BPV) was grown in EBTr cells. Cells were parasynchronized by a single thymidine block according to the method of Ward and Dadachanji (25), infected with 30 PFU of virus per cell, washed, and suspended in medium containing 10^{-5} M deoxycytidine. Cells were harvested 22 to 24 h postinfection.

RNA isolation, cell-free translation, and immunoprecipitation. Cytoplasmic RNA was isolated and translated in a rabbit reticulocyte lysate containing [^{35}S]methionine as previously described (5). Immunoprecipitation was performed essentially as described by Kessler (6), using Formalin-fixed, heat-killed *Staphylococcus aureus* (Cowan I strain obtained from Boehringer Mannheim Biochemicals, Indianapolis, Ind.) Immunoprecipitation and autoradiography were performed as described previously (5).

RESULTS

Procaryotic expression of MVM sequences. (i) Tripartite fusion proteins. The procaryotic expression vector pJS413 has been described in detail elsewhere (21, 27–29). Briefly, the construct contains a small polylinker into which coding sequences of interest are inserted and thus joined in phase to a *lacI-lacZ* gene fusion (Fig. 2A). Expression of the fusion protein is under control of the UV5 mutant *lac* promoter-operator region, and efficient translation of hybrid genes is facilitated by the presence of a short peptide leader (from the *cro* protein of bacteriophage lambda) which is situated at the correct distance from its own ribosome-binding site and provides a natural initiation codon and the first 22 amino acids at the amino terminus of the fusion. In pJS413 the *cro* leader is specifically out of translation phase with the *lacZ* gene owing to the polylinker and thus makes no β-galactosidase. However, insertion into the polylinker of DNA fragments $3n + 1$ nucleotides long and beginning at the first nucleotide of codon 1 in an ORF will correctly phase the *cro* leader and the *lacZ* gene, allowing the translation of a tripartite protein with demonstrable β-galactosidase activity. This enzyme activity is detected in bacteria carrying such recombinant plasmids since they grow as red colonies when replica plated on to lactose-MacConkey agar indicator plates. When BAL 31-digested fragments A (nucleotides 225 to 534) and C (nucleotides 2067 to 2291) from the MVM genome were inserted into pJS413, between 5 and 10% of the resultant colonies had β-galactosidase activity. This approximates the frequency expected for the cloning of randomly cut fragments containing one or two ORFs if the presence of the eucaryotic sequences does not substantially impair host viability. However, with BAL 31 derivatives of fragment B (nucleotides 763 to 2074) many fewer red colonies were obtained, and most of these gave low-level expression of fusion proteins when analyzed by SDS-PAGE. The most likely interpretation of this observation seems to be that, despite overproduction of the *lac* repressor in these cells, the *lac* promoter is slightly active under these conditions and allows the synthesis of small amounts of a toxic fusion protein. However, alternative explanations, such as the chimeric RNA or DNA themselves being toxic, have not been excluded. The bacterial clone in group B which produced by far the highest level of fusion protein contained the plasmid pYT202, diagrammed in Fig. 2A. DNA sequencing showed that this construct contained MVM sequences between nucleotides 1110 and 1638 in the viral genome, designated b in Fig. 1. Thus, *BAL* 31 removed around 350 bases from the left end and 430 bases from the right end of the B fragment. As such, this was certainly one of the shortest sequences presented to the vector and so supports the suggestion that longer sequences from this region impair bacterial viability. Interestingly, an overlapping, but much longer sequence, from this part of the genome (*Taq*I fragment nucleotides 227 to 2071) was somewhat tolerated by NF1829. However, prolonged IPTG induction of such clones did not lead to the production of very large amounts of the protein, since the cells ceased to replicate long before they reach the normal saturation density of this strain (unpublished observations).

SDS-PAGE showed that many of the clones containing BAL 31 derivatives of fragment A or fragment C overproduced fusion proteins. Constructs selected for further study were subcloned into M13, and the inserts were sequenced to determine both their exact sequence and the MVM reading frame they represented. The construct pYT201 contained a fragment A insert (designated a in Fig. 1) between nucleotides 225 and 534 expressed in ORF 3 of the MVM genome, while pYT203 contained nucleotides 2075 to 2291 (designated c) expressed in ORF 2. Bacteria expressing all three of the selected clones (pYT201, pYT202, and pYT203) accumulated fusion proteins to approximately 5% of the total cell mass when induced with IPTG.

The tripartite fusion proteins were then tested for their ability to elicit antibody responses against the MVM sequences they contained when injected into mice. In each case the mice became immune to the β-galactosidase portion

of the molecule, as assessed by immunoprecipitation of [^{35}S]methionine-labeled bacterial proteins (data not shown). However, none of the sera specifically immunoprecipitated any viral protein from the in vitro translation products of MVM-infected cell RNA, as shown in Fig. 3, lane 3, for antibodies against the b-region tripartite fusion. This suggested that the MVM sequences were masked in some way in such fusion polypeptides, and to overcome this we further engineered these plasmids to separate the MVM sequences from the β-galactosidase sequence, by inserting an in-frame amber codon between the two.

(ii) Bipartite fusion proteins—amber fragments. NF1829 bacteria carrying constructs with amber termination codons downstream of the MVM sequences did not accumulate either bipartite or tripartite fusion proteins after induction with IPTG (Fig. 4B, lanes 9 and 10). This result was obtained because the UAG codons effectively terminated translation, and the truncated amber fragments, which are detectable in short-term [^{35}S]methionine labeling experiments, were rapidly degraded in the bacterial cell and failed to accumulate to high levels. However, when such constructs are transformed into the bacterium LE392F, which carries the *supE* and *supF* genes, mutant tRNAs may insert glutamine or tyrosine, respectively, at the position of the amber codon (28). *supE* and *supF* are reported to give 14 and 55% suppression of UAG terminators, respectively (28), and thus in this bacterium both tripartite and bipartite proteins are synthesized. Overproduction and accumulation of the full-length fusion proteins appeared to protect the truncated from form degradation (Fig. 4B, lanes 4 and 6), and both proteins coprecipitated (Fig. 4B, lane 7). The amber fragment was then further purified by preparative gel electrophoresis to yield a protein which appeared essentially free of contamination as assessed by analytical SDS-PAGE (Fig. 4B, lane 8) and which was nonpyrogenic when injected into animals.

The 529-base MVM fragment in pYT202Am was expressed with a total of 92 nucleotides from the vector to yield a bipartite fusion protein with an apparent molecular size of around 20 kDa. This protein accumulated appreciably in LE392F, but smaller fusions, such as that derived from pYT203Am (approximately 10 kDa), were less well protected, and it was difficult, or impossible, to obtain useful amounts of such proteins at the levels of expression obtained. The TAC promoter (which contains the −35 region from the *trp* promoter and the −10 region from the *lac* UV5 promoter) is much stronger than the *lac* promoter although it is still regulated by the *lac* repressor (18). Substitution of this promoter into the amber constructs allowed considerable accumulation of even the smaller bipartite fusion proteins.

In contrast to the results obtained with the tripartite fusion proteins, the bipartite amber fragments proved highly immunogenic in both mice and rabbits when analyzed in the same way, as shown for the pYT202Am-TAC product in Fig. 3, lane 4.

Nucleotides 1110 to 1638 in ORF 3 are expressed in the NS-1 protein. Antisera raised against the amber fragment expressed by pYT202Am-TAC specifically precipitated the NS-1 polypeptides from the in vitro translation products of mRNA isolated from cells infected with MVM(p), H-1, LuIII, and CPV (Fig. 5). For all four viruses this protein had an apparent molecular size of around 83 kDa, although in the translation products of LuIII mRNA NS-1 was sometimes resolved into two bands of similar molecular weight. In contrast, no high-molecular-weight protein was precipitated from the translation products of BPV-infected cell mRNA, but a low-molecular-weight protein which comigrates with

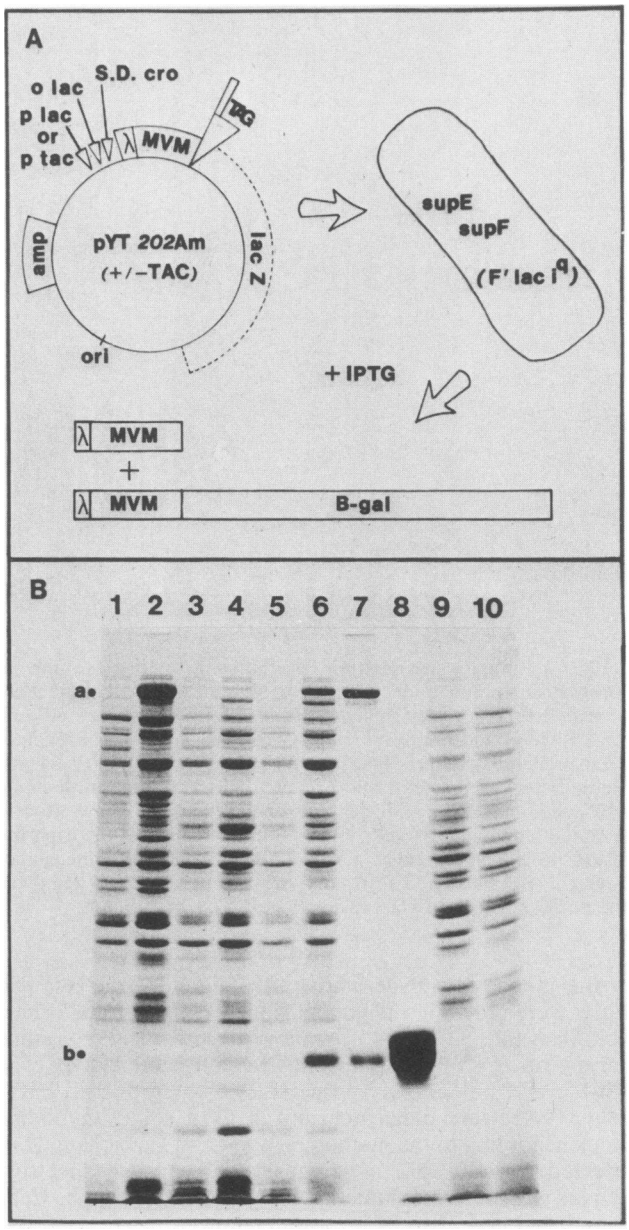

FIG. 4. Construction of amber fragments and introduction of the TAC promoter. (A) pYT202Am is derived from pTY202 by introduction of an in-frame amber termination codon just downstream of the MVM sequences. In most bacteria proteins synthesized from this construct (represented by open blocks) terminate at this codon. However, when this construct is introduced into LE392F, an *E. coli* strain carrying two amber suppressor genes (*supE* and *supF*), termination at this UAG is partially suppressed, and both truncated and full-length fusion proteins are synthesized upon induction with IPTG. (B) An 11% SDS-polyacrylamide gel showing the proteins synthesized by pYT202 (lanes 1 and 2), pYT202Am (lanes 3 and 4), and pYT202Am-TAC (lanes 5, 6, 9, and 10) before (lanes 1, 3, 5, and 9) and after (lanes 2, 4, 6, and 10) induction with IPTG. In Sup$^-$ bacteria such as NF1829 which lack amber suppressor tRNAs (lanes 9 and 10), the fusion proteins fail to accumulate to useful levels after induction, whereas in the Sup$^+$ bacterium LE392F (lanes 3 through 6), synthesis and accumulation of the full-length fusion protein (a) stabilizes and protects the truncated amber product (b), and both accumulate in the cell (lanes 4 and 6). Lane 7 shows the proteins harvested in the 8,000 × *g* pellet after detergent extraction and sonication of the bacteria. Lane 8 shows the amber fragment after gel purification.

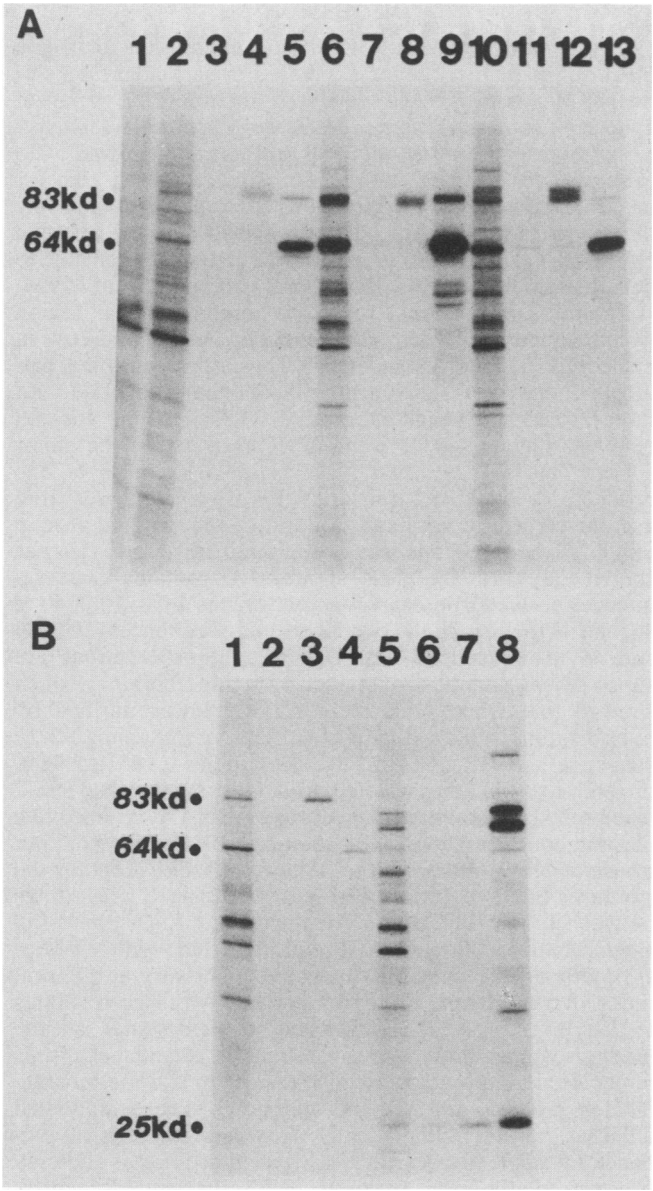

FIG. 5. Protein encoded in the b region of MVM. Immunoprecipitation of [^{35}S]methionine-labeled in vitro translation products of MVM, H-1, LuIII, CPV, and BPV with antisera against the bipartite amber fragment specified by pYT202Am-TAC, encoded in ORF 3 of MVM between nucleotides 1110 and 1638. (A) A 10% SDS-polyacrylamide gel showing the total in vitro translation products of mRNA obtained from A9 cells (lane 1), A9 cells infected with MVM(p) (lane 2), 324K cells infected with H-1 (lane 6), and 324K cells infected with LuIII (lane 10). Lanes 4, 8, and 12 show the NS-1 polypeptides of MVM, H-1, and LuIII, respectively, immunoprecipitated with rabbit antiserum raised against the bipartite fusion protein synthesized from pYT202Am-TAC. Lanes 3, 7, and 11 show proteins precipitated from the translation products of MVM, H-1, and LuIII RNA, respectively, with preimmunization serum. For comparison, in lane 5 the MVM capsid proteins VP-1 (83 kDa [kd]) and VP-2 (64 kDa) are precipitated with rabbit antiserum against purified MVM capsids, in lane 9 capsid polypeptides of H-1 are precipitated with guinea pig antiserum against H-1, and in lane 13 capsid polypeptides of LuIII are precipitated with hamster antiserum against LuIII. (B) A 10% SDS-polyacrylamide gel showing total in vitro translation products (lanes 1 and 5) and immunoprecipitates of the products of mRNA obtained from CFK cells infected with CPV (lanes 1 through 4) and EBTr cells infected with BPV (lanes 5 through 8). An 83-kDa NS-1 polypeptide is precipitated with rabbit antiserum against the fusion protein synthesized from pYT202Am-TAC from CPV-specified translation products (lane 3), but not from those of BPV (lane 7). Proteins precipitated by preimmunization serum from this rabbit are seen in lanes 2 for CPV and 6 for BPV. A 25-kDa protein which comigrates with NP-1 of BPV (lane 8) is weakly precipitated by both the immune and nonimmune sera from this rabbit (lanes 6 and 7) and may well be nonspecific. Capsid polypeptides VP-1 and VP-2 of CPV are weakly precipitated in lane 4 with a mixture of guinea pig anti-H-1 capsid and rabbit anti-porcine parvovirus capsid. Capsid polypeptides (VP-1, 80 kDa; VP-2, 72 kDa; and VP-3, 62 kDa) and the NP-1 protein (~27 kDa) of BPV are precipitated with rabbit anti-BPV capsid serum (lane 8). The figures at the left of each panel indicate apparent molecular sizes.

the NP-1 protein of BPV (8) was weakly precipitated by both immune and nonimmune sera from this rabbit (Fig. 5B, lanes 6 and 7) and is likely to be nonspecific.

Nucleotides 225 to 534 in ORF 3 are expressed in both NS-1 and NS-2. Antisera raised against the fusion protein specified by pYT201Am-TAC specifically recognized both an NS-1 protein (~83 kDa) and an NS-2 protein (~25 kDa) encoded by MVM, H-1, and LuIII (Fig. 6), demonstrating that mRNAs specified by all three of these viruses encode both an NS-1 and an NS-2 polypeptide. Additionally, the data show that in all three viruses this region of the genome (nucleotides 225 to 534 in MVM) must encode an amino-terminal peptide which is common to both of these proteins and that this amino-terminal region shows strong antigenic cross-reactions among all three viruses.

Both the NS-1 and NS-2 proteins specified by LuIII mRNA migrated as doublets (Fig. 6A, lane 11), suggesting that the heterogeneity between the two forms of each molecule is likely to reside in the common amino-terminal residues. We do not know the significance of this doublet, but suspect it may be trivial, for example, the result of transcription from both wild-type virus and a mutant with a relatively small in-frame deletion within the amino-terminal coding sequence of NS-1.

In contrast to the results obtained with H-1 and LuIII, antisera against the pYT201Am-TAC fusion totally failed to precipitate translation products specified by CPV-infected cell mRNA (Fig. 6B, lane 4). Thus, despite the high level of antigenic cross-reaction seen between the middle region of the NS-1 polypeptides in MVM and CPV demonstrated in Fig. 5, linear antigenic determinants expressed within the amino-terminal regions of these two NS-1 proteins appear to be completely unrelated. Similarly, this antiserum provides no evidence for an NS-2 molecule carrying MVM cross-reactive determinants encoded by CPV.

Mouse antisera against the pYT201Am-TAC fusion did precipitate a 27-kDa protein from the translation products of BPV-infected cell mRNA which we presume to be the NP-1 protein described by Lederman et al. (8) (Fig. 6B, lane 7). Preimmune sera from the same mice failed to precipitate this protein, but until this result is confirmed with specific rabbit antiserum (not currently available), we remain cautious about its interpretation.

Nucleotides 2075 to 2291 in ORF 2 are expressed in the NS-2 protein. Antibodies raised against the fusion protein specified by pYT203Am-TAC, which expresses the MVM sequence in ORF 2 between nucleotides 2075 and 2291, specifically precipitated the NS-2 polypeptide from the translation products of MVM, H-1, and LuIII (Fig. 6A) but did not react with proteins specified by CPV or BPV (Fig. 6B).

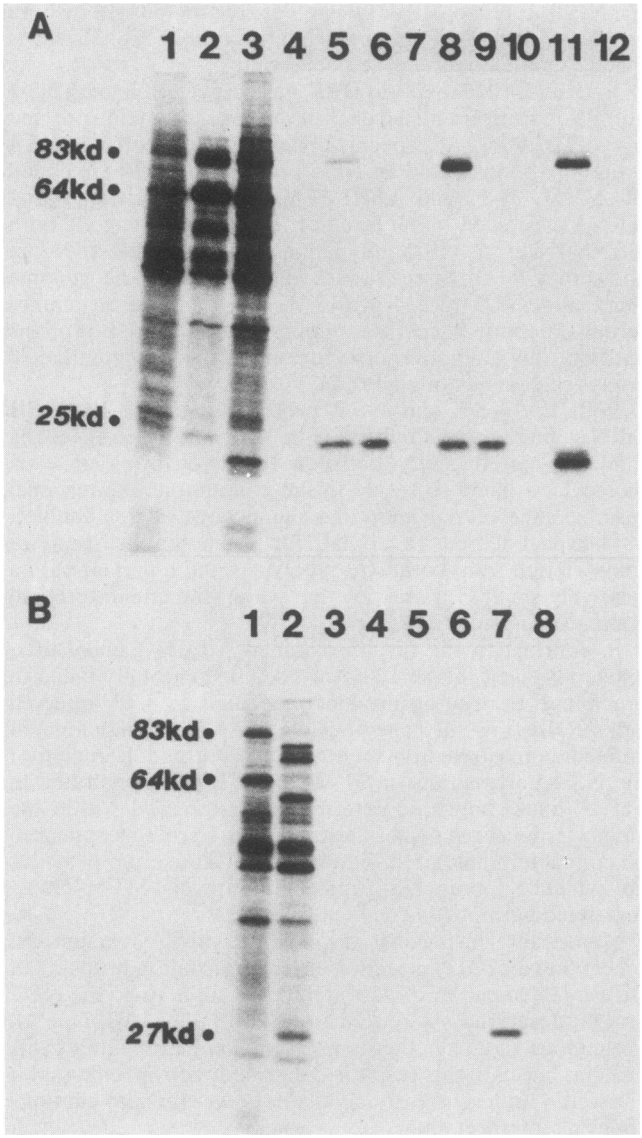

FIG. 6. Proteins encoded in the a and c regions of MVM. Immunoprecipitation of [^{35}S]methionine-labeled in vitro translation products of MVM, H-1, LuIII, CPV, and BPV mRNAs with antisera raised against the amber fragments specified by pYT201Am-TAC and pYT203Am-TAC (encoded in ORF 3, nucleotides 225 to 534, and ORF 2, nucleotides 2075 to 2291, of MVM, respectively). (A) A 10% SDS-polyacrylamide gel showing the total translation products of mRNA from cells infected with MVM (lane 1), H-1 (lane 2), and LuIII (lane 3). Lane 5, 8, and 11 show proteins immunoprecipitated from the total translation products of MVM, H-1, and LuIII, respectively, with mouse antiserum against the amber fragment synthesized from pYT201Am-TAC, while lanes 4, 7, and 10 show that preimmunization sera from these mice did not precipitate the nonstructural proteins from products of MVM, H-1, or LuIII RNA, respectively. A rabbit antiserum against the amber fragment synthesized from pYT203Am-TAC precipitates only the NS-2 polypeptide of MVM (lane 6), H-1 (lane 9), and LuIII (lane 12; faint band present in original autoradiograph not visible in figure). (B) A 10% SDS-polyacrylamide gel of immunoprecipitates of in vitro translation products of RNA from cells infected with CPV (lanes 3, 4, and 5) and BPV (lanes 6, 7, and 8), using preimmune mouse serum (lanes 3 and 6), mouse serum against the amber fragment of pYT201Am-TAC (lanes 4 and 7), and rabbit serum against the amber fragment specified by pYT203Am-TAC (lanes 5 and 8). All of these sera fail to precipitate CPV-specified proteins, while only the serum against the common amino-terminal region of NS-1 and NS-2 (nucleotides 225 to 534) effectively precipitates a low-molecular-weight protein (lane 7) which is presumed to be the NP-1 protein of BPV. Lanes 1 and 2 show the total translation products of the cytoplasmic poly(A)$^+$ RNA isolated from CPV- and BPV-infected cells, respectively. kd, Kilodalton.

DISCUSSION

In this study we used procaryotic expression to isolate blocks of protein sequence expressed in particular ORFs of the MVM genome. These sequences, incorporated as part of a bipartite bacterial fusion protein, were then relatively easy to purify from the other bacterial proteins and allowed us to raise antisera in mice and rabbits which were specific for particular domains of the viral nonstructural proteins. Using the in vitro translation products specified by virus-infected cell mRNA as a source of viral antigen, we were then able to demonstrate that the NS-1 and NS-2 polypeptides of MVM, H-1, and LuIII share a common amino-terminal region which contains the sequence encoded in MVM between nucleotides 225 and 534 in ORF 3. The carboxy-terminal half of NS-2 does not share protein sequence with NS-1, but rather utilizes a small alternative ORF (ORF 2) located between nucleotides 2075 and 2291, upstream of the minor splice. This suggests that NS-2 is most likely to be the product of a 3.3-kb R2 transcript (Fig. 1) which arises from a promoter at the left-hand (3′) end of the genome at map unit 4. S1 nuclease analysis of the viral transcripts (13) has shown that this size class of message contains exon sequences derived from map unit coordinates 4.0 to 10.0, 40 to 46, and 48 to 95, with the two intervening sequences (10 to 40 and 46 to 48) spliced out. Our present observations are supported by unpublished studies from this laboratory which showed that when virus-specific mRNA was purified by hybridization to individual plasmid DNAs containing MVM sequence 1 to 415, 2067 to 2204, 2290 to 2654, 2651 to 4000, or 3997 to 5148, the selected RNAs all synthesized NS-2 when translated in vitro. Conversely, mRNA hybridized to a plasmid containing the MVM sequence 1084 to 1659 did not translate NS-2, although this latter RNA did program the synthesis of NS-1. It has been estimated that R2 constitutes between 15 and 20% of the virus-specific mRNA present in asynchronous cultures of A9 cells infected with MVM(p) (13), although presumably this value might vary at different times after infection. We do not know how this corresponds to the level of NS-2 synthesized in vitro, since all the methods we might use to evaluate relative abundance of this molecule are indirect and involve the use of [^{35}S]methionine-labeled proteins and specific antibodies whose individual affinities can vary dramatically. However, there is no evidence available at present to suggest that NS-2 is the only product synthesized from an R2 transcript. Recently, Jongeneel and his colleagues obtained the sequence of a cDNA clone derived from MVM(i) mRNA which carries the exon regions characteristic of R2 (C. V. Jongeneel, G. McMaster, R. Sahli, and B. Hirt, Abstr. P11, p. 64, EMBO Workshop on Parvoviruses, Grangeneuve, Switzerland, September 1985). In this clone, MVM nucleotide 514 lies next to nucleotide 1990, suggesting a splice of 1,475 bases which would transfer any protein coding in ORF 3 to ORF 2. This is clearly compatible with the data presented here for NS-2. This clone also contains a minor splice which juxtaposes nucleotides 2280 and 2377, leaving the protein-coding sequence in ORF 2 for another six amino acids before terminating at an amber codon at nucleotide 2396. In the

absence of this second minor splice the protein would still terminate seven residues downstream of nucleotide 2280 at nucleotide 2299, a fact which is of interest because it has been suggested, although not yet proven, that while the great majority of R3 transcripts do use this 5′ splice site, a minority, which encode VP-1, do not (1, 4, 15). If the NS-2 transcripts use the splice sites described by Jongeneel et al., a protein starting at the AUG at position 261 would have a molecular weight of 25,009, which is in close agreement with the apparent molecular weight of NS-2 estimated from SDS-PAGE analysis.

A 76-kDa protein homologous to the NS-1 of MVM has been recognized previously in the in vitro translation products of mRNA from cells infected with H-1 (17), but this is the first clear demonstration that H-1 also encodes an NS-2 protein. Although the in vitro translation products of LuIII have not been examined previously, it is perhaps not surprising that this virus also encodes an NS-1 and NS-2 protein since heteroduplex mapping studies show that LuIII DNA shares considerable sequence homology with the left half of the genomes of both MVM and H-1 (3).

Similarly, the observation that an antibody against the middle region of the NS-1 molecule of MVM (nucleotides 1110 to 1638) cross-reacts with the NS-1 molecule of CPV is not surprising since CPV is a host-range mutant of feline panleukopenia virus (FPV), and the nucleotide sequence of FPV (4) could specify a protein which would share 156 of the 176 amino acids contained in the MVM insert of pYT202 (i.e., 87% homology). What is surprising is that the NS-1 of CPV, although having an apparent molecular weight very similar to that of MVM NS-1, entirely lacks linear antigenic determinants which cross-react with the amino terminal of MVM NS-1. Unfortunately, the nucleotide sequence of this region remains to be determined for both CPV and FPV.

Nucleotides 1110 to 1638 in the MVM genome cloned in pYT202 specified a protein sequence which is expressed exclusively in the NS-1 molecule. The only mRNA which incorporates this region of the genome is the 4.8-kb R1 transcript (13) (Fig. 1), and therefore this observation confirms our previous finding that NS-1 is encoded by the R1 transcript. All the parvoviruses which have been sequenced to date share a region of amino acid homology (MVM nucleotides 1428 to 1832) in the middle of the NS-1-coding region (20). The degree of homology through this region varies between viruses, but even such disparate viruses as MVM(p) and the dependovirus AAV-2 or MVM(p) and the human serum parvovirus B19 share around 50% homology at the amino acid level, while the more closely related viruses MVM and FPV show 96% homology. Seventy amino acids from the amino-terminal half of this region are contained in the fusion protein specified by pYT202Am-TAC, and the first 55 of these are known to show over 60% homology with a theoretical protein deduced from the nucleotide sequence of BPV (B. C. Shull, M. Lederman, K. C. Chen, E. S. Moses, E. R. Stout, and R. C. Bates, Abstr. S1/3, p. 17, EMBO Workshop on Parvoviruses, Grangeneuve, Switzerland, September 1985). However, we have not been able to identify a BPV-specified protein which cross-reacts with this region of the MVM NS-1.

Downstream of nucleotide 1832 in MVM the amino acid homology between the NS-1 polypeptides of MVM and CPV rapidly diminishes (1, 2, 4, 15), and although there are still clusters of similar residues, it is necessary to introduce gaps into the nucleotide sequence of CPV to get maximum alignment of encoded amino acids. This makes it essentially impossible to encode homologous proteins in alternative ORFs in this region. Although there is a dual block of ORF between nucleotides 288 and 566 in CPV, this sequence terminates upstream of the 5′ splice site proposed for the minor splice (1, 4, 15) (discussed above), and, when compared with the amino acid sequence expressed by pYT202, the alternative sequence in CPV shows only very weak homology (a maximum homology of 28% over a region of 72 amino acids). The minor transcripts of FPV and CPV have not been enumerated or mapped although the major (R3 equivalent) mRNA of FPV is known to comprise a 270-base exon spliced to a 2,500-base region complementary to the right-hand half of the genome (4). The possibility therefore exists that the transcription patterns of CPV and FPV and the organization of the left-hand ends of the genomes of these viruses may be significantly different from those determined for MVM and H-1.

The NS-1 proteins of MVM and H-1 are nuclear phosphoproteins (11; S. F. Cotmore and P. Tattersall, Virus Res., in press) which appear to have multiple functions in vivo, including *trans* activation of the middle promoter (16), an as yet unspecified effect on the later stages of viral DNA replication (M. Merchlinsky, Ph.D. thesis, Yale University, New Haven, Conn., 1984) and an inhibitory effect on cellular DNA replication (unpublished observations). The NS-2 protein has yet to be localized in vivo, and at present we do not have either biochemical or genetic information concerning its function. However, the demonstration that NS-1 and NS-2 share approximately 84 amino acids at their amino terminal suggests that at least one of the functions of both of these molecules is to interact with and coregulate the activity of a common element in the infected cell.

ACKNOWLEDGMENTS

We thank Solon Rhode III for guinea pig anti-H-1 serum and hamster anti-LuIII serum, Tom Molitor for rabbit anti-porcine parvovirus capsid serum, Jeffry Leary for rabbit anti-BPV capsid serum, and Gunter Siegl for supplying virus seed stocks. We thank Molecular Genetics Inc., Lynn Enquist, Roger Watson, and John Salstrom for help and advice with the expression vector system.

This work was supported by Public Health Service grants CA29303 and CA16038 from the National Cancer Institute.

LITERATURE CITED

1. **Astell, C. R., E. M. Gardiner, and P. Tattersall.** 1986. DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain. J. Virol. **57**:656–669.
2. **Astell, C. R., M. Thomas, M. Merchlinsky, and D. C. Ward.** 1983. The complete DNA sequence of minute virus of mice, an autonomous parvovirus. Nucleic Acids Res. **11**:999–1018.
3. **Banerjee, P. T., W. H. Olson, D. P. Allison, R. C. Bates, C. E. Snyder, and S. Mitra.** 1983. Electron microscopic comparison of the sequences of single-stranded genomes of mammalian parvoviruses by heteroduplex mapping. J. Mol. Biol. **166**: 257–272.
4. **Carlson, J., K. Rushlow, I. Maxwell, F. Maxwell, S. Winston, and W. Hanh.** 1985. Cloning and sequencing of DNA encoding structural proteins of the autonomous parvovirus feline panleukopenia virus. J. Virol. **55**:574–582.
5. **Cotmore, S. F., L. J. Sturzenbecker, and P. Tattersall.** 1983. The autonomous parvovirus MVM encodes two nonstructural proteins in addition to its capsid polypeptides. Virology **129**: 333–343.
6. **Kessler, S. W.** 1975. Rapid isolation of antigens from cells with staphylococcal protein A-antibody adsorbent: parameters of the interaction of antibody-antigen complexes with protein A. J. Immunol. **115**:1617–1624.
7. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the

assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
8. **Lederman, M., R. C. Bates, and E. R. Stout.** 1983. In vitro and in vivo studies of bovine parvovirus proteins. J. Virol. **48**: 10–17.
9. **Merchlinsky, M. J., P. Tattersall, J. J. Leary, S. R. Cotmore, E. M. Gardiner, and D. C. Ward.** 1983. Construction of an infectious molecular clone of the autonomous parvovirus MVM. J. Virol. **47**:227–232.
10. **Messing, J., and J. Vieira.** 1981. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
11. **Paradiso, P. R.** 1984. Identification of multiple forms of the noncapsid parvovirus protein NCVP1 in H-1 parvovirus-infected cells. J. Virol. **52**:82–87.
12. **Paradiso, P. R., K. R. Williams, and R. L. Constantino.** 1984. Mapping of the amino terminus of the H-1 parvovirus major capsid protein. J. Virol. **52**:77–81.
13. **Pintel, D., D. Dadachanji, C. R. Astell, and D. C. Ward.** 1983. The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. Nucleic Acids Res. **11**:1019–1038.
14. **Pintel, D., M. Merchlinsky, and D. C. Ward.** 1984. Expression of minute virus of mice structural proteins in murine cell lines transformed by bovine papillomavirus-minute virus of mice plasmid chimera. J. Virol. **52**:320–327.
15. **Rhode, S. L., III.** 1985. Nucleotide sequence of the coat protein gene of canine parvovirus. J. Virol. **54**:630–633.
16. **Rhode, S. L., III.** 1985. *Trans*-activation of parvovirus P_{38} promoter by the 76K noncapsid protein. J. Virol. **55**:886–889.
17. **Rhode, S. L., III, and P. R. Paradiso.** 1983. Parvovirus genome: nucleotide sequence of H1 and mapping of its genes by hybrid-arrested translation. J. Virol. **45**:173–184.
18. **Russell, D. R., and G. N. Bennett.** 1982. Construction and analysis of in vivo activity of *E. coli* promoter hybrids and promoter mutants that alter the −35 to −10 spacing. Gene **20**:231–243.
19. **Sanger, S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Shade, R. O., M. C. Blundell, S. F. Cotmore, P. Tattersall, and C. R. Astell.** 1986. Nucleotide sequence and genome organization of human parvovirus B19 isolated from the serum of a child during aplastic crisis. J. Virol. **58**:921–926.
21. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Tattersall, P., and J. Bratton.** 1983. Reciprocal productive and restrictive virus-cell interactions of immunosuppressive and prototype strains of minute virus of mice. J. Virol. **46**:944–955.
23. **Tattersall, P., A. J. Shatkin, and D. C. Ward.** 1977. Sequence homology between the structural polypeptides of minute virus of mice. J. Mol. Biol. **111**:375–394.
24. **Tattersall, P., and D. C. Ward.** 1978. The parvoviruses—an introduction, p. 3–12. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Ward, D. C., and D. K. Dadachanji.** 1978. Replication of minute virus of mice DNA, p. 297–313. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Ward, D. C., and P. Tattersall.** 1982. Minute virus of mice, p. 313–334. *In* H. L. Foster, J. D. Small, and J. G. Fox (ed.), The mouse in biomedical research, vol. 10. Academic Press, Inc., New York.
27. **Watson, R. J., J. H. Weis, J. S. Salstrom, and L. W. Enquist.** 1982. Herpes simplex virus type-1 glycoprotein D gene: nucleotide sequence and expression in *Escherichia coli*. Science **218**:381–384.
28. **Watson, R. J., J. H. Weis, J. S. Salstrom, and L. W. Enquist.** 1984. Bacterial synthesis of herpes simplex type 1 and 2 glycoprotein D antigens. J. Invest. Dermatol. **83**:102–111.
29. **Weis, J. H., L. W. Enquist, J. S. Salstrom, and R. J. Watson.** 1983. An immunologically active chimeric protein containing herpes simplex virus type 1 glycoprotein D. Nature (London) **302**:72–74.