JOURNAL OF VIROLOGY, Jan. 1983, p. 173–184
0022-538X/83/010173-12$02.00/0

Vol. 45, No. 1

Parvovirus Genome: Nucleotide Sequence of H-1 and Mapping of Its Genes by Hybrid-Arrested Translation

SOLON L. RHODE III* AND PETER R. PARADISO

Institute for Medical Research of Bennington, Bennington, Vermont 05201

Received 27 July 1982/Accepted 5 October 1982

The nucleotide sequence of the parvovirus H-1 has been determined by the chain-terminating method of Sanger. The sequence is 5,176 nucleotides long. Two large open reading frames (1 and 2) and two smaller open reading frames (3 and 4) of potential importance were identified in the plus-strand sequence. Promoter sequences are located at map positions 4 and 38 when map positions are expressed as percent of genome length from the 3′ end of the virion minus strand. The locations for the genes for the parvovirus capsid proteins and a 76,000-dalton noncapsid protein (NCVP1) were mapped by hybrid-arrested translation. The gene for the capsid proteins VP1 and VP2′ is located in the 5′ half of the virus genome. The gene for NCVP1 is located in the 3′ half of the viral DNA.

The autonomously replicating parvoviruses are animal viruses with linear single-stranded DNA genomes and icosahedral capsids (13, 26, 28, 31). H-1 is one of the group of parvoviruses of rodents that includes Kilham rat virus, minute virus of mice (MVM), and H-3 (28). These viruses have a high degree of nucleotide sequence homology and can complement each other for many functions (3, 14, 21, 24). Other parvoviruses less closely related to the rodent group, such as canine parvovirus, feline panleukopenia virus, porcine parvovirus, and bovine parvovirus are important pathogens of their respective hosts (2, 28).

The method of replication of parvoviruses involves the synthesis and replication of a duplex replicative form DNA (RF DNA), but many of the details of the replication process are unknown (31). Analysis of several temperature-sensitive mutants of H-1 has indicated that one or both of the capsid proteins, VP1 and VP2′, are required for progeny single-stranded DNA synthesis (19). The phenotypes of a deletion mutant of H-1, *dl*1, as well as defective interfering particles, suggest that a viral protein (RF rep) is also required for replicative form DNA replication (22–24). In addition, a protein has been found covalently bound to the 5′-termini of both the plus and minus strands of RF DNA (19). It has been suggested that the terminal protein is a site-specific nickase required for RF DNA replication (25). Except for the capsid proteins VP1 and VP2′, there has been no evidence that parvoviruses contain the genes for one or more of these other proteins implicated in their replication process.

The H-1 parvovirus is known to encode two capsid proteins designated VP1 (88,000 daltons) and the major capsid protein VP2′ (68,000 daltons) (9, 10). Two other capsid proteins which have been reported are VP2, which is a post-translational cleavage product of VP2′, and VP3, which appears to be a breakdown product of VP2′ and perhaps VP1 (16). Nearly all of the amino acid sequence of VP2′ appears to be contained within VP1 for the parvovirus MVM (30), and these two proteins are highly homologous in H-1 as well (P. R. Paradiso, this paper and unpublished data). Studies on H-1 mRNA (7, 12) have identified two promoters in the H-1 genome mapping at positions 4 and 38, and three major cytoplasmic mRNAs of 4,700, 3,000, and 2,800 bases, each of which is spliced RNAs with a 2.6-kilobase (kb) common, unspliced region. The 2.8- and 3.0-kb mRNAs are present in a ratio of approximately five to one (12), which is the same as the ratio of VP2′ to VP1 in the capsid (8). This ratio and the large regions of homology in both the RNAs and proteins leads to the speculation that the 2.8-kb mRNA encodes VP2′ and the 3.0-kb mRNA encodes VP1 (7). Furthermore, Tal et al. (29) have identified an MVM message in which 30% of the sequences, those between map positions (m.p.'s) 9 and 39, have been spliced out.

To derive information on the structure of parvovirus capsid proteins and the possible existence of genes for noncapsid proteins, we have determined the nucleotide sequence of H-1. As an adjunct to this analysis, we have used hybrid-arrested translation (HART) (17) with M13 recombinant clones containing H-1 minus-strand DNA of defined sequence to assign parvovirus gene products to the major open reading frames

in the nucleotide sequence. These experiments have identified a new noncapsid virus protein (NCVP1), which maps in the left half of the genome (between m.p.'s 4 and 49). VP1 and VP2', on the other hand, map in the right half of the genome (between m.p.'s 38 and 88).

MATERIALS AND METHODS

Materials. Restriction enzymes, molecular linkers, M13 specific primers, nucleotides, and dideoxynucleotides were obtained from commercial sources as described previously (25). The 17-mer synthetic M13 primer was superior to the 12-mer for some templates. The Klenow fragment of *Escherichia coli* DNA polymerase I came from Boehringer-Mannheim Corp., New York; Bethesda Research Laboratories, Inc., Bethesda, Md.; or New England Nuclear Corp., Boston, Mass.; all gave satisfactory results. T4 DNA polymerase was purchased from Bethesda Research Laboratories.

Molecular cloning of parvovirus DNA. H-1 RF DNA was used as the source of H-1 DNA as described previously (25). In addition, the M13 vector mp7-01, kindly provided by David Bentley, Oxford University, was used in some cases. The cloning of *Hha*I fragments was carried out by first making flush ends with T4 DNA polymerase treatment in the presence of each dNTP at 100 μM for 30 min at 37°C before the addition of *Eco*RI linkers and cloning in M13 mp5. Fragments generated by *Xba*I, *Ava*II, and *Eco*RII were similarly made to have blunt ends before cloning.

DNA sequencing. Sequencing reactions and gel conditions were as previously described (25, 27).

Isolation and translation of H-1 mRNA. NB cells were infected with H-1 virus at a multiplicity of infection of 25 (or no virus in the case of the mock infections), and the total cell RNA was harvested at 16 h postinfection with 8 M guanidine-hydrochloride as described by Adams et al. (1). The RNA (1 to 2 μg) was then translated in a cell-free rabbit reticulocyte lysate system (Bethesda Research Laboratories) incorporating [^{35}S]methionine (New England Nuclear; 1,000 Ci/mmol; 5 μCi per reaction) into the proteins.

Immunoprecipitation of H-1 translation products. Precipitation of H-1 proteins was achieved by adding 2 μl of serum directly to the reticulocyte lysate reaction mixture at 4°C for 4 h, followed by 100 μl of a 10% solution of Formalin-fixed *Staphylococcus aureus* in phosphate-buffered saline (Pansorbin; Calbiochem-Behring, Inc., La Jolla, Calif.) at 4°C for 15 min. The pellets were washed twice with 0.5 ml of phosphate-buffered saline, taken up in 75 μl of gel sample buffer (11), and incubated at 100°C for 5 min before electrophoresis.

HART. Total infected-cell RNA (2 μg) and H-1 viral-strand DNA cloned into an M13 vector (25) (1 μg) was made to 0.1 M KCl in a total volume of 5 μl, sealed into a siliconized glass micropipette, and incubated at 55°C for 30 min. The hybrids were then translated, and the proteins were immunoprecipitated with serum from hamsters that survived a neonatal H-1 infection.

Polyacrylamide gel electrophoresis. Polyacrylamide gels were run by the procedure of Laemmli (11) and analyzed by fluorography as described by Bonner and Lasky (5). The peptide mapping gel was done according to the procedure of Cleveland et al. (6). H-1 mRNA was translated, cold carrier capsid protein was added, and the proteins were separated on a 7.5% polyacrylamide gel and stained with Coomassie brilliant blue. ^{3}H-labeled capsid proteins were also included as a control. The appropriate bands were sliced out and analyzed on a 15% polyacrylamide gel in the second dimension in the presence of 10 μl of *S. aureus* V8 protease (0.1 absorbance unit at 280 nm per ml).

Purification and amino acid analysis of viral proteins. VP1 and VP2' were purified out of a 7.5% acrylamide preparative slab gel with a single 10-cm sample lane after staining with Coomassie brilliant blue. Proteins were eluted from stained gel bands by soaking in 70% formic acid, and the stain was removed by passing the concentrated eluate (100 μl) over a 1-ml Bio-Gel P-6 column in 70% formic acid and collecting everything that eluted ahead of the dye. Aliquots were dried and hydrolyzed in 6 N HCl–0.2% phenol for 24 h at 110°C with subsequent amino acid analysis performed on a Beckman 121M analyzer equipped with an Autolab computing integrator.

RESULTS

Nucleotide sequence of H-1. The nucleotide sequences of the terminal 126 nucleotides from the 3' end of H-1 minus-strand DNA and the terminal 644 nucleotides from the minus-strand 5'-terminus which includes the origin of replication have been reported (3, 25). In this study, we determined the sequence beginning with the *Hha*I site at nucleotide 59 (numbering from the 5'-terminus of the plus or sense strand) (Fig. 1). Both DNA strands were independently analyzed. The sequencing strategy used is outlined in Fig. 2.

We obtained a slightly different sequence within the 5'-hairpin (plus strand) than that previously reported, which consisted of an addition of a T at nucleotide 62, an interchange of nucleotides 71 and 72, and a deletion of a T that would have been nucleotide 86 (3). This change in the sequence results in a slightly more perfect hairpin, but it would seem to have no other profound consequences. All restriction sites representing fragment boundaries were sequenced as internal sites in other overlapping fragments, except for the terminal *Hha*I site at nucleotide 59, the *Eco*RI site at 1,086, and the terminal *Hin*dII site at 5,169 (25).

Location of the open reading frames. The sequence of the plus strand was divided into the three reference reading frames, and the positions of stop codons and AUG codons in each of the reference reading frames were determined (Fig. 3). In the first reference reading frame, there is an open reading frame (frame 1) of 2,139 nucleotides or 713 codons from nucleotide 141 to 2,279. However, the only TATA sequence indicative of a promoter for this reading frame is at nucleotide 180, and the first available AUG start

codon is at nucleotide 264. If frame 1 is translated without interruption through RNA splicing to the terminator at nucleotide 2,280, a protein of 672 amino acids or a molecular weight of 76,000 would be produced.

In the second reference reading frame, another large open reading frame (frame 2) is located between nucleotides 2,386 and 4,572. This represents 729 codons, which would generate a protein of about 79,300 daltons. This reading frame has promoter sequences preceding it at m.p. 38 (nucleotide 1.979; Fig. 1).

Codon usage. Although the exact arrangements of the coding sequences into the parvovirus capsid proteins VP1 and VP2′ and the NCVP1 have yet to be determined, the codon usage for frame 1 and frame 2 was tabulated (data not shown). The codon frequencies for frame 1 are similar to those for frame 2, except for the Val codon GUG, the Glu codon GAU, the Arg codon CGA, the Gly codon GGG, and the Cys codon UGU, which differ by more than factors of 3. The arginine codons beginning with AG are much more frequent than those with CG in both open frames. Codons with CG in the second and third positions are also infrequent. This is similar to the reported bias against CG pairs in other eucaryotic DNA viruses (13, 15).

Minus strand. The only RNAs transcribed from parvovirus DNA that have been reported to date are complementary to the minus strand (7, 29). In the present study, analysis of the stop codons in the minus-strand sequence revealed an open reading frame of 483 nucleotides at nucleotides 1,334 to 1,817 and another of 459 nucleotides at nucleotides 2,401 to 2,859. Since there are no appropriate promoter or AUG sequences at these sites, nor any other sizable open reading frames on this strand, it is unlikely that there are any structural genes represented in the minus strand.

Translation of H-1 mRNA and identification of a noncapsid virus protein. Total cell RNA was extracted from both H-1-infected NB (simian virus 40-transformed human newborn kidney) cells and mock-infected NB cells and was translated in a rabbit reticulocyte lysate system, incorporating [^{35}S]methionine into the proteins. Figure 4 shows the positions of the translated capsid proteins VP2′ and VP1, which are absent (or at least greatly reduced) in the translated mock-infected cell RNA. Furthermore, the infected-cell RNA encoded large amounts of a protein with an apparent molecular weight of 84,000 (P. R. Paradiso, Fed. Proc. **41**:1278, 1982). This protein is clearly absent in the mock-infected RNA translation. To look directly at the virus-specific proteins, immunoprecipitations were done with two different anti-H-1 sera: hamster serum from animals that survived neonatal infection, and guinea pig antiserum raised against purified capsid proteins. The hamster H-1 antiserum was found to precipitate the two capsid proteins VP1 and VP2′ as well as the 84,000-dalton (84K) protein (Fig. 5, lane 2). The guinea pig H-1 antiserum, on the other hand, precipitated only VP1 and VP2′ but not the 84K protein (Fig. 5, lane 3). Interestingly, the serum from hamsters surviving a neonatal infection with the parvovirus MVM(i) also reacted with all three viral proteins (Fig. 5, lane 3). When mock-infected cell RNA was used, none of these products was precipitated (data not shown).

The fact that the 84K protein did not react with the guinea pig antiserum against the capsid proteins suggested that it was not a degradation product of VP1 or a precursor to VP2′. This conclusion was confirmed by peptide mapping, using the method of Cleveland et al. (6) (data not shown). Thus, the 84K protein is virus specific but noncapsid in nature, and will be referred to from here on as noncapsid protein 1 (NCVP1).

Mapping the virus proteins by HART. To map the virus proteins to the open reading frames predicted from the DNA sequence, we took advantage of our library of minus-strand (viral-strand) clones of H-1 DNA in M13 vectors to study the HART of H-1 mRNA. Figure 6 shows a schematic representation of the predicted open reading frames of H-1 and the nine different probes we hybridized to the mRNA before translation. Figure 7 shows the effect of the various DNAs on the proteins synthesized by using the hamster anti-H-1 serum to precipitate the virus proteins. The translation of NCVP1 mRNA was arrested by all of the probes within m.p.'s 1.1 to 39.2, which includes the 1.1–9.5, 9.5–19.2, 21.0–31.1, 35.0–39.2, 35.8–45.9, and 39.2–44.1 H-1 DNA sequences. However, it was not arrested by the 49.5–55.2 fragment or by anything to the right of m.p. 49.5. VP1 and VP2′ mRNA translation was arrested by all of the probes within map positions 35.8 and 81.0, which includes 35.8–45.9, 39.2–44.1, 49.5–55.2, 55.2–72.6, and 81.0–87.6 H-1 DNA sequences. Thus, NCVP1 is encoded predominantly in the left (3′) half of the viral genome, whereas VP1 and VP2′ share the right (5′) half of the genome. The 35.8–45.9 and the 39.2–44.1 H-1 sequences arrest all three of the virus proteins, indicating a potential overlap in the coding regions of these proteins. As indicated by the arrowheads in the left margin of Fig. 7, three truncated proteins with approximate molecular weights of 65,000, 68,000, and 72,000 (lanes 4 to 6, respectively) arose as a result of the hybridizations with promoter proximal boundaries of 35.0, 35.8, and 39.2; they may represent shortened forms of NCVP1 (see Discussion).

Determination of the precise structure of the

0001 CATTTTTAGA ACTGACCAAC CATGTTCACG CAAGTGACGT GATGACGCGC

0051 GCTGCGCGCG CTGCCTTCGG CAGTCACACG TCACTAGCGT TTCACATGGT

0101 TGGTCAGTTC TAAAAATGAT AAGCGGTTCA GAGAGTTTGA AACCAAGGCG

0151 GGAAACGGAA GTGGGCGTGG CTAACTGTAT ATAAGCAGTC ACTCTGGTCG

0201 GTTACTCACT CTGCTTTCAT TTCTGAGTTT GTGAGACACA GGAGCGAGAC

0251 TAACCAACTA ACC ATG GCT GGA AAC GCT TAC TCC GAT GAG GTT TTG GGA

0300 GTA ACA AAC TGG CTG AAG GAC AAA AGT AGC CAG GAG GTG TTC TCA TTT

0348 GTT TTT AAA AAT GAA AAC GTC CAA CTA AAT GGA AAG GAC ATC GGT TGG

0396 AAT AGT TAC AGA AAG GAG CTA CAA GAT GAC GAG CTG AAG TCT CTA CAA

0444 CGA GGG GCG GAG ACC ACT TGG GAC CAA AGC GAG GAC ATG GAA TGG GAG

0492 AGC GCA GTG GAT GAC ATG ACC AAA AAG CAA GTA TTT ATT TTT GAT TCT

0540 TTG GTT AAG AAG TGT TTG TTT GAA GTG CTC AGC ACA AAG AAC ATA GCT

0588 CCT AGT AAT GTT ACT TGG TTC GTG CAG CAT GAA TGG GGA AAG GAC CCA

0636 GGC TGG CAC TGT CAT GTG CTG ATT GGA GGC AAG GAC TTT AGT CAA CCT

0684 CAA GGA AAA TGG TGG AGA AGG CAG CTA AAT GTG TAC TGG AGT AGA TGG

0732 TTG GTG ACT GCC TGT AAT GTT CAA CTA ACA CCA GCT GAA AGA ATT AAA

0780 CTG AGA GAA ATA GCA GAG GAC AGT GAA TGG GTC ACT TTG CTT ACC TAT

0828 AAG CAT AAG CAC ACC AAG AAG GAC TAT ACC AAG TGT GTT CTT TTT GGA

0876 AAC ATG ATT GCT TAT TAC TTT TTA AGC AAA AAG AAA ATA TGT ACC AGT

0924 CCA CCA AGG GAC GGA GGC TAT TTT CTT AGC AGT GAC TCT GGC TGG AAA

0972 ACT AAC TTT TTG AAA GAG GGC GAG CGC CAT CTA GTG AGC AAA CTG TAT

1020 ACT GAT GAG ATG AAA CCA GAA ACG GTC GAG ACC ACA GTG ACC ACT GCA

1068 CAG GAA GCT AAG CGC GGC AGA ATT CAA ACT AGA GAG GAG GTC TCG ATT

1116 AAA ACC ACA CTC AAA GAG TTG GTA CAT AAA AGA GTA ACC TCA CCA GAA

1164 GAC TGG ATG ATG ATG CAG CCA GAC AGT TAC ATT GAA ATG ATG GCT CAA

1212 CCA GGT GGA GAA AAC TTG CTT AAA AAT ACA CTA GAG ATC TGT ACA CTG

FIG. 1. The nucleotide sequence for the plus strand of the parvovirus H-1. The sequences for nucleotides 1 to 58 are from reference 2, and those for 4,535 to 5,176 are from reference 25. The 5′ terminal nucleotide has not been definitively identified. The large open reading frames, 1 and 2, are shown with the nucleotides organized as codons. The sequencing reactions and gel electrophoresis conditions were the same as those previously described (25). Sites referred to in the text are underlined.

1260 ACT CTA GCA AGA ACC AAA ACA GCC TTT GAC TTG ATT CTG GAA AAA GCT

1308 GAA ACC AGC AAA CTA GCC AAC TTT TCC ATG GCT AGC ACC AGA ACC TGT

1356 AGA ATC TTT GCT GAG CAT GGC TGG AAC TAT ATT AAA GTC TGC CAT GCC

1404 ATC TGT TGT GTG CTG AAT AGA CAA GGA GGC AAA AGG AAC ACT GTG CTC

1452 TTT CAC GGA CCA GCC AGC ACA GGC AAA TCT ATT ATT GCA CAA GCC ATA

1500 GCA CAA GCA GTT GGT AAT GTT GGT TGT TAC AAT GCT GCC AAT GTG AAC

1548 TTT CCA TTT AAT GAC TGT ACC AAC AAA AAC TTG ATT TGG GTG GAA GAA

1596 GCT GGT AAC TTT GGC CAG CAA GTA AAC CAA TTC AAA GCT ATT TGT TCT

1644 GGC CAA ACC ATA CGC ATT GAT CAA AAA GGA AAA GGC AGC AAA CAG ATT

1692 GAA CCA ACA CCA GTT ATT ATG ACC ACC AAC GAG AAC ATT ACC GTG GTT

1740 AGA ATA GGC TGT GAG GAA AGA CCA GAA CAC ACT CAA CCA ATC AGA GAC

1788 AGA ATG CTC AAC ATT CAC CTG ACA CGT ACA CTA CCT GGT GAC TTT GGT

1836 TTG GTG GAT AAG CAC GAA TGG CCT CTG ATC TGT GCT TGG TTG GTG AAG

1884 AAT GGT TAC CAA TCT ACC ATG GCT TGT TAC TGT GCT AAA TGG GGC AAA

1932 GTT CCT GAT TGG TCA GAG GAC TGG GCG GAG CCG AAG CTA GAC ACT CCT

1980 ATA AAT TCG CTA GGT TCA ATG CGC TCA CCA TCT CTG ACT CCG AGA AGT

2028 ACG CCT CTC AGC CAA AAC TAC GCT CTT ACT CCA CTT GCA TCG GAC CTT

2076 GCG GAC CTA GCT CTA GAG CCT TGG AGC ACA CCA AAT ACT CCT GTT GCG

2124 GGC ACT GCA GCA AGC CAA AAC ACT GGG GAG GCT GGT TCC ACA GCC TGC

2172 CAA GGT GCT CAA CGG AGC CCA ACC TGG TCC GAG ATC GAG GCG GAT TTG

2220 AGA GCT TGC TTC AGT CAA GAA CAG TTG GAG AGC GAC TTC AAC GAG GAG

2268 CTG ACC TTG GAC TAA GGTACA ATG GCA CCT CCA GCT AAA AGA GCT AAA

2316 AGA GGT AAG GGG CTA AGG GAT GGT TGG TTG GTG GGG

2352 TAC TAA TGTATGACTACCTGTTTTACAGGCCTGA frame 2: AAT

2389 CAC TTG GTT CTA GGT TGG GTG CCT CCT GGC TAC AAG TAC CTG GGA CCA

2437 GGG AAC AGC CTT GAC CAA GGA GAA CCA ACC AAC CCT TCT GAC GCC GCT

2485 GCC AAA GAA CAC GAC GAA GCC TAC GAC CAA TAC ATC AAA TCT GGA AAA

2533 AAT CCT TAC CTG TAC TTC TCT CCT GCT GAT CAA CGC TTC ATT GAC CAA

FIG. 1. *Continued*

2581 ACC AAA GAC GCC AAG GAC TGG GGC GGC AAG GTT GGT CAC TAC TTT TTT

2629 AGA ACC AAG CGA GCT TTT GCA CCT AAG CTT TCT ACT GAC TCT GAA CCT

2677 GGC ACT TCT GGT GTG AGC AGA CCT GGT AAA CGA ACT AAA CCA CCT GCT

2725 CAC ATT TTT GTA AAT CAA GCC AGA GCT AAA AAA AAA CGC GCT TCT CTT

2773 GCT GCA CAG CAG AGG ACT CTG ACA ATG AGT GAT GGC ACC GAA ACA AAC

2821 CAA CCA GAC ACT GGA ATC GCT AAT GCT AGA GTT GAG CGA TCA GCT GAC

2869 GGA GGT GGA AGC TCT GGG GGT GGG GGC TCT GGC GGG GGT GGG ATT GGT

2917 GTT TCT ACT GGG ACT TAT GAT AAT CAA ACG ACT TAT AAG TTT TTG GGA

2965 GAT GGA TGG GTA GAA ATA ACT GCA CAT GCT TCT AGA CTT TTG CAC TTG

3013 GGA ATG CCT CCT TCA GAA AAC TAC TGC CGC GTC ACC GTT CAC AAT AAT

3061 CAA ACA ACA GGA CAC GGA ACT AAG GTA AAG GGA AAC ATG GCC TAT GAC

3109 ACA CAT CAA CAA ATT TGG ACA CCA TGG AGC TTG GTA GAT GCT AAT GCT

3157 TGG GGA GTT TGG TTC CAA CCA AGT GAC TGG CAG TTC ATT CAA AAC AGC

3205 ATG GAA TCG CTG AAT CTT GAC TCA TTG AGC CAA GAA CTA TTT AAT GTA

3253 GTA GTC AAA ACA GTC ACT GAA CAA CAA GGA GCT GGC CAA GAT GCC ATT

3301 AAA GTC TAT AAT AAT GAC TTG ACG GCC TGT ATG ATG GTT GCT CTG GAT

3349 AGT AAC AAC ATA CTG CCT TAC ACA CCT GCA GCT CAA ACA TCA GAA ACA

3397 CTT GGT TTC TAC CCA TGG AAA CCA ACC GCA CCA GCT CCT TAC AGA TAC

3445 TAC TTT TTC ATG CCT AGA CAA CTC AGT GTA ACC TCT AGC AAC TCT GCT

3493 GAA GGA ACT CAA ATC ACA GAC ACC ATT GGA GAG CCA CAG GCA CTA AAC

3541 TCT CAA TTT TTT ACT ATT GAG AAC ACC TTG CCT ATT ACT CTC CTG CGC

3589 ACA GGT GAT GAG TTT ACA ACT GGC ACC TAC ATC TTT AAC ACT GAC CCA

3637 CTT AAA CTT ACT CAC ACA TGG CAA ACC AAC AGA CAC TTG GCA TGC CTC

3685 CAA GGA ATA ACT GAC CTA CCA ACA TCA GAT ACA GCA ACA GCA TCA CTA

3733 ACT GCA AAT GGA GAC AGA TTT GGA TCA ACA CAA ACA CAG AAT GTG AAC

3781 TAT GTC ACA GAG GCT TTG CGC ACC AGG CCT GCT CAG ATT GGC TTC ATG

3829 CAA CCT CAT GAC AAC TTT GAA GCA AAC AGA GGT GGC CCA TTT AAG GTT

FIG. 1. *Continued*

3878 CCA GTG GTA CCG CTA GAC ATA ACA GCT GGC GAG GAC CAT GAT GCA AAC

3925 GGA GCC ATA CGA TTT AAC TAT GGC AAA CAA CAT GGC GAA GAT TGG GCC

3973 AAA CAA GGA GCA GCA CCA GAA AGG TAC ACA TGG GAT GCA ATT GAT AGT

4021 GCA GCT GGG AGG GAC ACA GCT AGA TGC TTT GTA CAA AGT GCA CCA ATA

4069 TCT ATT CCA CCA AAC CAA AAC CAG ATC TTG CAG CGA GAA GAC GCC ATA

4118 GCT GGC AGA ACT AAC ATG CAT TAT ACT AAT GTT TTT AAC AGC TAT GGT

4165 CCA CTT AGT GCA TTT CCT CAT CCA GAT CCC ATT TAT CCA AAT GGA CAA

4213 ATT TGG GAC AAA GAA TTG GAC CTG GAA CAC AAA CCT AGA CTA CAC GTA

4261 ACT GCA CCA TTT GTT TGT AAA AAC AAC CCA CCA GGT CAA CTA TTT GTT

4309 CAC TTG GGG CCT AAT CTG ACT GAC CAA TTT GAC CCA AAC AGC ACA ACT

4357 GTT TCT CGC ATT GTT ACA TAT AGC ACT TTT TAC TGG AAG GGT ATT TTG

4405 AAA TTC AAA GCC AAA CTA AGA CCA AAT CTG ACC TGG AAT CCT GTA TAC

4453 CAA GCA ACC ACA GAC TCT GTT GCC AAT TCT TAC ATG AAT GTT AAG AAA

4501 TGG CTC CCA TCT GCA ACT GGC AAC ATG CAC TCT GAT CCA TTG ATT TGT

4549 AGA CCT GTG CCT CAC ATG ACA TAC TAA CCAAC CAACTATGTT TCTCTGTTTG

4601 CTTCACATAA TACTTAAACT AACTAGACTA CAACATAAAA ATATACACTT

4651 AATAATAGAT TATTAAAAAT AACATAATAT GGTAGGTTAA CTGTAAAAAA

4701 TAATAGAACT TTTGGAATAA ATATAGTTAG TTGGTTAATG TTAGATAGAA

4751 TATAAAAAGA TTTTGTATTT TAAAATAAAT ATAGTTAGTT GGTTAATGTT

4801 AGATAGAATA TAAAAAGATT TTGTATTTGG GAAATAAAAA GGGTGGTTGG

4851 GTGGTTGGTT GGTACTCCCT TAGACTGAAT GTTAGGGACC AAAAAAATAA

4901 TAAAATAATT AAAATGAACA AGGACTACTG TCTATTCAGT TGACCAACTG

4951 AACCTATAGT ATCACTATGT TTTTAGGGTG GGGGGGTGGG AGATACATAC

5001 GTTCGCTATG GACCAAGTGG TACCGGTTGG TTGCTAAGCT CGAACAAGAC

5051 GGCTAAGCCG GTCCGGTTGG TTGAGCGCAA CCAACCGGTA CCACTTGGTC

5101 CATAGCGAAC GTATGTATCT CCCACCCCCC CACCCTAAAA ACATAGTGAT

5151 ACTATAGGTT CAGTTGGTCA ACTGAA(T)-5176-3' TERMINUS

FIG. 1. *Continued*

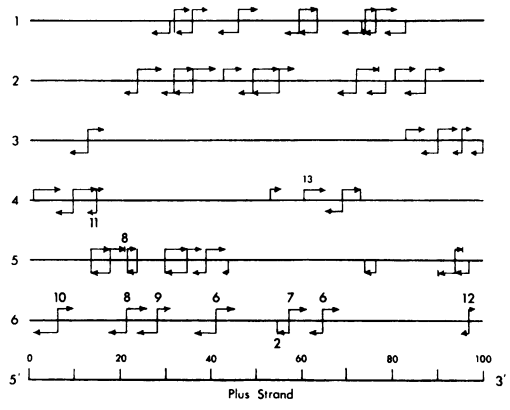

FIG. 2. The sequencing strategy used to determine the nucleotide sequence of the parvovirus H-1. Sequencing in the direction of the plus strand is represented as arrows above the reference lines and in the minus-strand direction below the reference lines. Map position 0 is at the plus-strand 5′ terminus, and 100 is at the 3′ terminus. It should be noted that the positions of the *Hha*I fragments at m.p. 53.4–69.2 and 69.2–73.3 were mistakenly reported in the reverse order (25). The restriction endonuclease sites referred to are: 1. *Hae*III, 2. *Sau*3A, 3. *Hin*dII, 4. *Hha*I, 5. *Rsa*I, 6. *Pst*I, 7. *Xba*I, 8. *Eco*RI, 9. *Ava*II, 10. *Eco*RII, 11. *Pvu*II, 12. *Taq*I, and 13. *Alu*I.

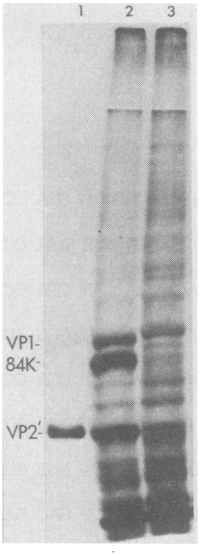

FIG. 4. Translation of H-1 mRNA. mRNA (1 to 2 μg) from H-1-infected and mock-infected cells was translated as described in Materials and Methods. The total translation products were separated on a 7.5% polyacrylamide gel (11) and analyzed by fluorography (5). Lane 1, ^{3}H-labeled, purified empty capsid; lane 2, translation of infected-cell RNA; lane 3, translation of mock-infected-cell RNA.

H-1 gene products will require analysis of the transcription products and their splicing arrangements. We also intend to analyze H-1 protein peptide sequences to confirm the complete accuracy of the DNA sequence. As a preliminary step to this confirmation, we have compiled the amino acid composition of the hypothetical translation products of frame 1, frame 2, and the 86,933-dalton VP1 (see Discussion)—which utilizes part of frame 1 and most of frame 2—and have compared them with the compositions of H-1 capsid proteins (Table 1). The amino acid compositions predicted by frame 2 and the 87K VP1 are reasonably close matches to those for the values measured for the capsid proteins VP1 and VP2′. This is in agreement

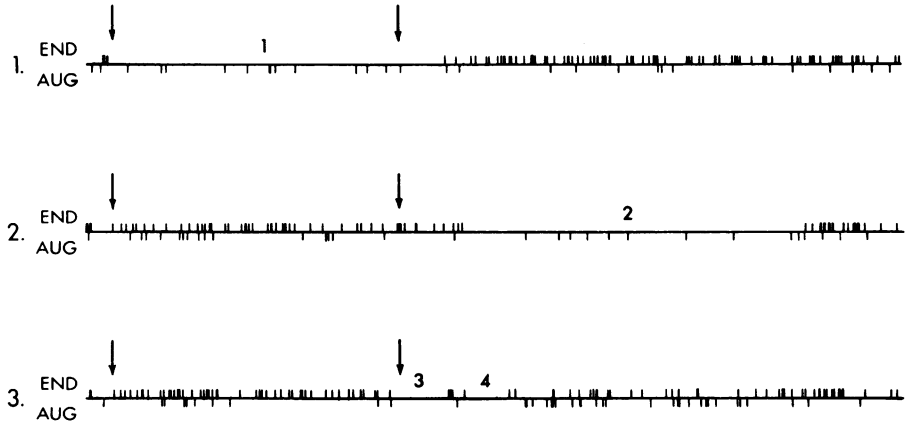

FIG. 3. A diagram of the stop codons and AUG codons in each of the three possible reading frames of the plus-strand sequence. Reading frame 1 was empirically chosen to be in frame with open reading frame 1; it begins with nucleotide 3. The major open reading frames 1 through 4 are indicated, and the arrows are positioned at the sites of the TATAA promoter sequences at m.p.'s 4 and 38.

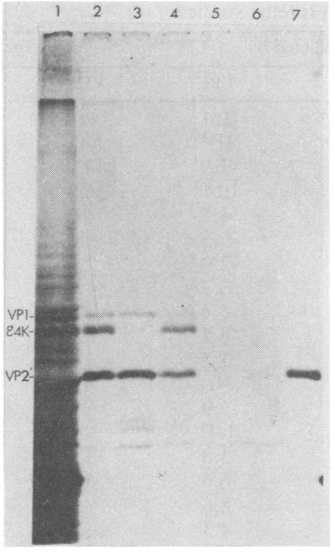

FIG. 5. Immunoprecipitation of H-1 translation products. Infected cell mRNA was translated, immunoprecipitated as described in Materials and Methods, and analyzed on a 7.5% polyacrylamide gel as described in the legend to Fig. 4. Lane 1, unprecipitated translation products; lane 2, serum from a hamster which survived a neonatal H-1 infection; lane 3, guinea pig antiserum raised against purified capsid protein; lane 4, serum from a hamster which survived a neonatal MVM(i) infection; lane 5, normal hamster serum; lane 6, guinea pig pre-bleed serum; lane 7, ^{3}H-labeled empty capsid.

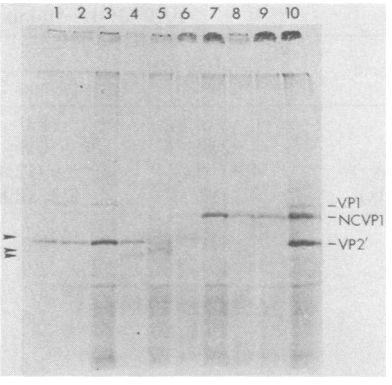

FIG. 7. Hybrid-arrested translation of H-1 mRNA. Total infected-cell RNA (2 μg) and H-1 viral strand DNA cloned into an M13 vector (1) (1 μg) were made to 0.1 M KCl in a total volume of 5 μl, sealed into a siliconized glass micropipette, and incubated at 55°C for 30 min. The hybrids were then translated, and the proteins were immunoprecipitated with serum from hamsters surviving neonatal infection as described in the legend to Fig. 5. The precipitates were run on a 7.5% polyacrylamide gel and fluorographed. The DNA sequences used are described in the legend to Fig. 6. Lane 1, *Hha*I 1.1–9.5; lane 2, *Hha*I 9.5–19.2; lane 3, *Eco*RI–*Hae*III 21.0–31.1; lane 4, *Rsa*I 35.0–39.2; lane 5, *Hae*III 35.8–45.9; lane 6, *Rsa*I 39.2–44.1; lane 7, *Sau*3A 49.5–55.2; lane 8, *Sau*3A 55.2–72.6; lane 9, *Sau*3A 81.0–87.6; lane 10, unhybridized. The arrows point to apparently truncated forms of NCVP1 seen in lanes 4, 5, and 6.

with our assignment of frame 2 to these structural proteins.

DISCUSSION

The parvovirus H-1 has a genome of 5,176 nucleotides that contains two large open reading frames (1 and 2) of 672 and 729 codons, respectively, that are either headed by an AUG start codon or preceded by one. There are two promoter sequences at m.p.'s 4 and 38. HART experiments have identified the open reading frame 1 (m.p.'s 5 to 44) as the gene for NCVP1 protein of 76,000 daltons. We speculate that NCVP1 is the RF rep gene product (24) and is

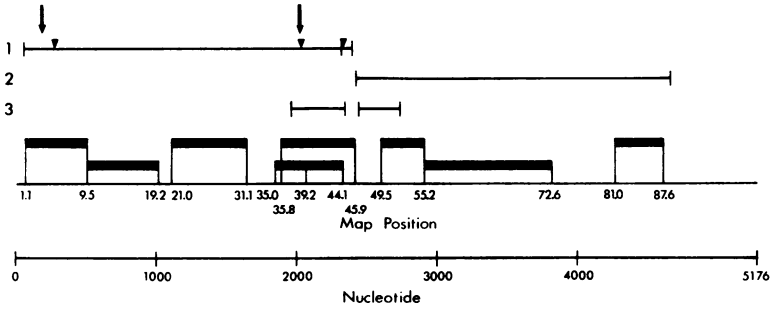

FIG. 6. Strategy for hybrid-arrested translation. The top three lines indicate the major open regions in the three reading frames of the H-1 genome, with potential initiation AUG codons at nucleotides 264, 1,998, and 2,289 indicated with arrowheads. The heavy vertical bars correspond to the H-1 DNA restriction fragments cloned into M13 vectors that were used to arrest the H-1 mRNA translation. The restriction sites used were: *Hha*I, 1.1–9.5 and 9.5–19.1; *Eco*RI–*Hae*III, 21.0–31.1; *Rsa*I, 35.0–39.2; 39.2–44.1; *Hae*III, 35.8–45.9; *Sau*3A, 49.5–55.2, 55.2–72.6, and 81.0–87.6. The arrows represent putative promoter sites at m.p.'s 4 and 38. The nucleotide 1 and m.p. 0 are at the 5' terminus of the sense strand.

TABLE 1. Amino acid composition of proteins studied

Amino acid	Composition predicted by[a]:					
	Frame 1	Frame 2	87K VP1	VP1	VP2′	Capsid
Ala	6.6	8.0	9.7	10.1	8.8	8.0
Cys	2.8	0.8	1.0	ND[b]	ND	Tr
Asp (Asx)	4.9 (10.1)	5.9 (12.4)	5.7 (11.8)	12.9	14.0	10.8
Glu (Glx)	7.3 (12.2)	3.3 (8.8)	3.5 (8.8)	10.1	10.3	11.5
Phe	3.4	4.1	4.0	4.5	4.9	3.1
Gly	6.0	8.2	7.8	9.5	9.1	13.0
His	1.9	2.9	2.5	2.9	2.9	3.0
Ileu	4.8	4.1	3.7	3.7	4.1	3.8
Lys	7.1	4.5	4.3	5.4	3.5	6.5
Leu	7.9	6.4	6.6	6.2	6.2	7.8
Met	2.2	1.6	1.4	1.5	1.7	2.1
Asn	5.4	6.4	6.1	ND	ND	ND
Pro	4.2	7.4	7.6	6.9	6.4	7.7
Gln	4.9	5.5	5.5	ND	ND	ND
Arg	4.2	4.1	2.9	3.0	3.7	3.2
Ser	6.9	6.2	7.1	6.6	6.5	7.0
Thr	8.2	9.4	9.4	8.6	9.8	6.8
Val	5.7	4.9	4.5	4.3	4.1	4.8
Tryp	3.6	2.2	2.2	ND	ND	
Tyr	2.2	4.0	3.7	3.7	4.0	1.0

[a] Amino acid composition of predicted proteins coded by frame 1, frame 2, and a hypothetical 87K VP1 as proposed in the Discussion are compared with capsid protein amino acid compositions. The compositions are expressed as mole percents, and the amino acid composition of whole capsids is from reference 19. The amino acid compositions of VP1 and VP2′ do not include Trp and Cys in the 100%. Frame 2 codes for an 80,000-dalton protein and therefore cannot reflect the exact composition of either VP1 (88,000) or VP2′ (68,000).

[b] ND, Not determined.

either the postulated site-specific nickase (25), the terminal bound protein (19), or both. The open reading frame 2 (m.p.'s 46 to 88.3) has been assigned to the capsid proteins VP1 and VP2′, but the detailed arrangements of the coding sequences for these two proteins are unknown.

The noncapsid nature of the NCVP1 protein is indicated by the fact that (i) it does not react with antiserum to purified capsids and (ii) it has never been identified as a protein component of purified H-1 virus (9, 18). A protein with the same characteristics as NCVP1 has been immunoprecipitated from extracts of H-1-infected cells (P. R. Paradiso, unpublished data), and a similar protein has recently been reported for the parvovirus Aleutian disease virus (4).

The HART data indicate clearly that NCVP1 utilizes reading frame 1 (see Fig. 6) and terminates to the left of map position 49, since the 49.5–55.2 H-1 DNA did not arrest the NCVP1 mRNA translation. The protein presumably initiates at nucleotide 264, which is the first AUG start signal downstream of the promoter at nucleotide 180, and terminates either at nucleotide 2,280 in reading frame 1 or at nucleotide 2,303 in frame 3 via a splicing mechanism. In either case, the DNA sequence predicts a molecular weight of ca. 76,000 for NCVP1, which is 10% smaller than predicted from gel electrophoresis. Lanes 4, 5, and 6 of Fig. 7 also show apparent truncated proteins with approximate molecular weights of 65,000, 68,000, and 72,000, respectively. The DNA sequence predicts that truncation of NCVP1 at m.p.'s 35, 36, and 39 would give proteins of 58,000, 59,900, and 65,700 daltons, respectively. Like the intact protein, these truncated proteins are about 10% larger on polyacrylamide gels than predicted by the DNA sequence.

The HART results also show that VP1 and VP2′ are predominantly encoded in reading frame 2 (Fig. 4) in the right half of the genome, and they probably terminate at nucleotide 4,573. The leftward boundary of these proteins is still not known. However, the HART data suggest that both are at or to the right of map position 39.2 (nucleotide 2,026). If the promoter sequence at nucleotide 1,979 signals the start of a message at the expected 30 nucleotides downstream, then translation would begin at the AUG at nucleotide 2,289 or perhaps 2,336 or 2,357. This would provide more than enough coding sequences for the 68K VP2′. The most likely position would be nucleotide 2,289 (in reading frame 1), since the probe DNA sequence representing 39.2–44.1 has its distal end at nucleotide 2,284 and arrests the translation of VP2′. This message would splice into reading frame 2 be-

fore the terminator at nucleotide 2,356 and continue to the polyadenylate signal at nucleotide 4,918. VP1, on the other hand, cannot use the promoter at nucleotide 1,979 because the AUG at nucleotide 1,998 is too close to the TATA signal and the next available AUG would not produce a large enough protein. Thus, it appears that VP1 mRNA uses the promoter at nucleotide 180 as does NCVP1, but not the AUG at nucleotide 264, since the translation was not arrested by the 1.1–9.5 H-1 DNA probe. An mRNA species, similar to that described by Tal et al. (29) for MVM, in which a major portion of the left half of the genome has been spliced out, would explain the use of the leftward promoter with coding sequences in the right half of the genome. If this is the case, the HART data indicate that the AUG at nucleotide 1,998 may be used, since sequences to the left of this codon did not arrest translation of VP1. If, for example, the codons in frame 1 from the AUG at nucleotide 1,998 to a splice at 2,231 and in frame 2 from 2,401 to the end were translated into protein, an 86,933-dalton product would result. The predicted amino acid composition of this protein matches closely that of VP1 (Table 1). On the other hand, determination of the DNA sequence of canine parvovirus shows it to be 80% homologous to H-1 in this area and to have the promoter sequences at m.p. 38, but there is no AUG near the position analogous to the H-1 AUG at nucleotide 1,998 (S. L. Rhode, unpublished data).

Figure 8 shows three potential mRNAs based on the previous mRNA analyses (7, 12, 29) and the data presented in this paper. The predicted mRNA species corresponding to NCVP1, VP1, and VP2′ are approximately 4.5, 2.8, and 2.6 kb, respectively. Splice sites have been assigned to accommodate the predicted initiation codons (see above), the known terminator codons, the previously predicted common splice junctions (7), and the sizes of the proteins. The previously predicted 4.78-kb mRNA (7), which we have assumed to correspond to NCVP1, was shown to have a common downstream splice junction with the two major mRNA species (7). We have depicted the NCVP1 mRNA here with splice junctions similar to those of the VP2′ mRNA. The splices in the VP1 mRNA permit the use of the promoter at m.p. 4, initiation of protein synthesis in reading frame 1 (see Fig. 6), and coding of the bulk of the protein in reading frame 2. Similarly, the VP2′ mRNA uses the promoter at m.p. 38, initiates the protein synthesis in reading frame 1, and is spliced into reading frame 2 to encode the remainder of the protein. The polyadenylate site for all three messages is at m.p. 95. Although these mRNA species are somewhat smaller than predicted by Green et al. (7), the splicing pattern and the relative sizes are similar.

The discovery of a new parvoviral gene product may help to answer several questions regarding parvovirus replication, since this protein may be the protein covalently linked to the 5′ terminus of H-1 RF DNA (19) or the predicted site-specific nickase (25), or both. It is of interest that the approximately 300-base pair deletions found in defective interfering genomes (20) and *dl*1 (21) that include m.p. 38 render these genomes fully defective, with no detectable viral proteins synthesized (P. R. Paradiso, unpublished data). This is also the region which arrested the translation of all three virus proteins. The fact that the H-1 NCVP1 was precipitated by the hamster anti-MVM serum suggests that this protein is not unique to H-1 and in fact may be a common product of the autonomous parvoviruses. Parvoviruses appear to make efficient use of their limited coding capacity for the synthesis of viral proteins.

ACKNOWLEDGMENTS

We thank Joachim Messing and David Bentley for kindly providing the M13 vectors, Helene Toolan and John Kongsvik for providing the antiserum used in this study, and Kenneth Williams for the amino acid analyses. We also thank William Burhans, Jeanne Helft, and M. Suzanne Hopkins for their

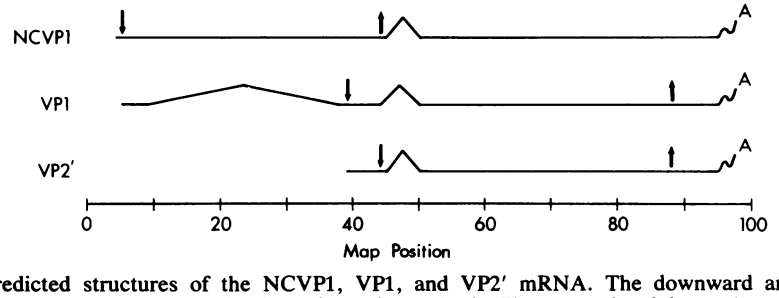

FIG. 8. Predicted structures of the NCVP1, VP1, and VP2′ mRNA. The downward arrow represents initiation, and the upward arrow termination, of protein synthesis. The untranslated (introns) are indicated by the breaks in the straight lines. Although the exact details of the structure of these mRNAs are not shown, these predictions are compatible with our results and those described previously (7, 12, 29).

excellent technical assistance, and Virginia Haas for typing this manuscript.

This work was supported in part by Public Health Service grants CA-25866 and CA-26801 from the National Institutes of Health.

LITERATURE CITED

1. **Adams, S. L., M. E. Sobel, B. H. Howard, K. Olden, K. Yamada, B. DeChrombrugghe, and I. Pastan.** 1977. Levels of translatable mRNAs for cell surface protein, collagen precursors, and two membrane proteins are altered in Rous sarcoma virus-transformed chick embryo fibroblasts. Proc. Natl. Acad. Sci. U.S.A. **74**:3399–3403.
2. **Appel, M. J. G., B. J. Cooper, H. Greisen, and L. E. Carmichael.** 1978. Status report: canine viral enteritis. J. Am. Vet. Med. Assoc. **173**:1516–1518.
3. **Astell, C., M. Smith, M. Chow, and D. Ward.** 1979. Structure of the 3′ hairpin termini of four rodent parvovirus genomes: nucleotide sequence homology at origins of DNA replication. Cell **17**:691–703.
4. **Bloom, M. E., R. E. Race, and J. B. Wolfinbarger.** 1982. Identification of a nonvirion protein of Aleutian disease virus: mink with Aleutian disease have antibody to both virion and nonvirion proteins. J. Virol. **43**:608–616.
5. **Bonner, W., and R. Laskey.** 1974. A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
6. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1101–1106.
7. **Green, M. R., R. M. Lebovitz, and R. G. Roeder.** 1979. Expression of the autonomous parvovirus H1 genome: evidence for a single transcriptional unit and multiple spliced polyadenylated transcripts. Cell **17**:967–977.
8. **Kongsvik, J. R., J. F. Gierthy, and S. L. Rhode III.** 1974. Replication process of the parvovirus H-1. IV. H-1-specific proteins synthesized in synchronized human NB kidney cells. J. Virol. **14**:1600–1603.
9. **Kongsvik, J. R., M. S. Hopkins, and K. A. O. Ellem.** 1978. Two populations of infectious virus produced during H-1 infection of synchronized transformed cells, p. 505–520. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Kongsvik, J. R., and H. W. Toolan.** 1972. Capsid components of the parvovirus, H-1. Proc. Soc. Exp. Biol. Med. **139**:1202–1205.
11. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
12. **Liebovitz, R., and R. Roeder.** 1981. The autonomous parvovirus H-1 consists of two overlapping transcription units with independent promoter sites. J. Supramolec. Struct. Cell Biochem. **5**(Suppl):431.
13. **McGeoch, D. J., L. V. Crawford, and E. A. C. Follett.** 1970. The DNAs of three parvoviruses. J. Gen. Virol. **6**:33–40.
14. **McMaster, G. K., P. Beard, H. D. Engers, and B. Hirt.** 1981. Characterization of an immunosuppressive parvovirus related to the minute virus of mice. J. Virol. **38**:317–326.
15. **Morrison, J. M., H. M. Keir, H. Subak-Sharp, and L. V. Crawford.** 1967. Nearest neighbour base sequence analysis of the deoxyribonucleic acids of a further three mammalian viruses: simian virus 40, human papilloma virus and adenovirus type 2. J. Gen. Virol. **1**:101–108.
16. **Paradiso, P. R.** 1981. Infectious process of the parvovirus H-1: correlation of protein content, particle density, and viral infectivity. J. Virol. **39**:800–807.
17. **Paterson, B. M.** 1977. Structural gene identification and mapping by DNA-mRNA hybrid-arrested cell-free translation. Proc. Natl. Acad. Sci. U.S.A. **74**:4370–4374.
18. **Peterson, J., R. Dale, R. Karess, D. Leonard, and D. Ward.** 1978. Comparison of parvovirus structural proteins—evidence for post-translational modification, p. 431–446. *In* D. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Revie, D., B. Tseng, R. Grafstrom, and M. Goulian.** 1979. Covalent association of protein with replicative form DNA of parvovirus H-1. Proc. Natl. Acad. Sci. U.S.A. **76**:5539–5545.
20. **Rhode, S. L. III.** 1976. Replication process of the parvovirus H-1. V. Isolation and characterization of temperature-sensitive H-1 mutants defective in progeny DNA synthesis. J. Virol. **17**:659–667.
21. **Rhode, S. L. III.** 1978. Replication process of the parvovirus H-1. X. Isolation of a mutant defective in replicative-form DNA replication. J. Virol. **25**:215–223.
22. **Rhode, S. L. III.** 1978. H-1 DNA synthesis, p. 279–296. *In* D. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Rhode, S. L. III.** 1978. Defective interfering particles of parvovirus H-1. J. Virol. **27**:347–356.
24. **Rhode, S. L. III.** 1982. Complementation for replicative form DNA replication of a deletion mutant of H-1 by various parvoviruses. J. Virol. **42**:1118–1122.
25. **Rhode, S. L. III, and B. Klaassen.** 1982. DNA sequence of the 5′ terminus containing the replication origin of parvovirus replicative form DNA. J. Virol. **41**:990–999.
26. **Rose, J. A.** 1974. Parvovirus reproduction, p. 1–61. *In* H. Fraenkel-Conrat and R. Wagner (ed.), Comprehensive virology, vol. 4. Plenum Publishing Corp., New York.
27. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
28. **Siegl, G.** 1976. The parvoviruses. *In* S. Goard and C. Hallauer (ed.), Virology monographs. Springer-Verlag, New York.
29. **Tal, J., D. Ron, P. Tattersall, S. Bratosin, and Y. Aloni.** 1979. About 30% of minute virus of mice RNA is spliced out following polyadenylation. Nature (London) **279**:649–651.
30. **Tattersall, P., A. Shatkin, and D. Ward.** 1977. Sequence homology between the structural polypeptides of minute virus of mice. J. Mol. Biol. **111**:375–394.
31. **Ward, D., and P. Tattersall (ed.).** 1978. Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.