Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 181, 541–552 (1991)

Sequence Analysis of Infectious Pancreatic Necrosis Virus Genome Segment B and Its Encoded VP1 Protein: A Putative RNA-Dependent RNA Polymerase Lacking the Gly–Asp–Asp Motif

ROY DUNCAN,*,[1] CARLA L. MASON,† EVA NAGY,‡ JO-ANN LEONG,† AND PETER DOBOS§

*Department of Microbiology and Infectious Diseases, University of Calgary, Calgary, Alberta, Canada T2N 4N1; †Department of Microbiology, Oregon State University, Corvallis, Oregon 97331-3804; ‡Department of Veterinary Microbiology and Immunology, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada N1G 2W1; and §Department of Microbiology, College of Biological Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Received July 30, 1990; accepted December 18, 1990

The genome segment B sequence of infectious pancreatic necrosis virus was determined for both the Jasper and Sp serotypes. The sequences are 2784 and 2630 bp long, respectively, and contain a single large open reading frame encoding the VP1 protein, the putative RNA-dependent RNA polymerase (RdRp) of IPNV. The proteins exhibit an 88% homology with each other, but only 41% with infectious bursal disease virus (IBDV) VP1, another member of the Birnaviridae. Despite the low overall homology between the IPNV and IBDV VP1 proteins, homologous regions were detected within the central portion of the proteins. The carboxy-proximal regions of the VP1, which contain very low amino acid homology, displayed evidence of conservation in structural features such as a hydrophilic, highly basic domain. Consensus sequences associated with GTP-binding proteins and RdRps were also detected in VP1. However, unlike the RdRps associated with single-stranded plus RNA viruses, the birnavirus RdRp lacks the Gly–Asp–Asp motif characteristic of this enzyme family.

INTRODUCTION

Members of the family Birnaviridae possess a bisegmented double-stranded RNA (dsRNA) genome consisting of two segments (A and B) within an unenveloped, single-shelled icosahedral capsid of 60 nm (Dobos *et al.*, 1979). Members of the family include infectious pancreatic necrosis virus (IPNV), the causative agent of an acute, contagious disease of young salmonid fishes; infectious bursal disease virus (IBDV) which causes a highly contagious disease of chickens, characterized by the destruction of lymphoid cells in the bursa of Fabricius; and *Drosophila* X virus (DXV) of *Drosophila melanogaster* (Dobos *et al.*, 1979; Brown, 1986).

Birnavirus gene expression involves transcription from the dsRNA parental genome segments by a virion-associated RNA-dependent RNA polymerase (RdRp) (Mertens *et al.*, 1982; Spies *et al.*, 1987). The enzyme directs the synthesis of nonpolyadenylated mRNA species corresponding in size to the A and B segments of the viral genome. No spliced messenger RNA species have been detected (Somogyi and Dobos, 1980; Bernard, 1980; Mertens and Dobos, 1982; Azad *et al.*, 1985).

[1] To whom reprint requests should be addressed.

The sequence of the viral A segment (approximately 3.1 kbp) has been determined for two IPNV Jasper strains (Duncan and Dobos, 1986; Haverstein *et al.*, 1990) and for three IBDV strains (Hudson *et al.*, 1986; Spies *et al.*, 1989; Kibenge *et al.*, 1990). This genome segment A encodes an approximately 100K polyprotein which is cleaved to produce, in order from the amino terminal end, the major virion structural protein, VP2, and the minor structural proteins, VP4 and VP3 (Nagy and Dobos, 1984; Huang *et al.*, 1986; Duncan *et al.*, 1987; Jagadish *et al.*, 1988). Autocatalytic protease activity has been associated with the VP4 protein (Duncan *et al.*, 1987; Jagadish *et al.*, 1988).

The viral B segment encodes VP1, the presumptive virion-associated RdRp. This protein is approximately 90 kDa in size as estimated by gel migration (MacDonald and Dobos, 1981; Nagy and Dobos, 1984; Azad *et al.*, 1985). IPNV, IBDV, and DXV all possess genome-linked proteins (VPg) tightly associated with the ends of the genomic RNA (Persson and MacDonald, 1982; Revet and Delain, 1982; Muller and Nitschke, 1987). In the cases of IPNV and IBDV, this VPg has been shown to be VP1. The VPg of DXV is a 67K protein of undetermined origin (Revet and Delain, 1982). The RdRp may also contain guanylyl and methyl transferase activities (Spies and Muller, 1990).

Morgan *et al.* (1988) reported the sequence of IBDV genome segment B (Australian strain 002-73). They

0042-6822/91 $3.00

reported no homology between the predicted VP1 sequence and the sequences of putative ssRNA-dependent RNA polymerases (ssRdRp). Subsequently, Gorbalenya and Koonin (1988) reanalyzed the IBDV sequence and detected homology with consensus sequence elements found in ssRdRps (Kamer and Argos, 1984; Argos, 1988). We present in this report the genome segment B nucleotide sequences of the IPNV-Jasper and -Sp strains and the predicted sequences of their encoded VP1 proteins. Sequence analysis revealed the presence of extensive homology between IPNV and IBDV VP1 in the central regions of the proteins and the presence of several conserved domains associated with RdRps and GTP binding proteins. Contrary to other putative RdRps, however, IPNV VP1 lacks the conserved Gly–Asp–Asp motif, the proposed catalytic site of this enzyme family.

MATERIALS AND METHODS

Viral strains and cDNA production

The IPNV-Jasper and -Sp strains were grown in CHSE-214 cells and genomic viral RNA was isolated from purified virions as previously described (Huang *et al.*, 1986; Duncan *et al.*, 1987). The viral RNA was used as templates for cDNA production and recombinant plasmids were screened for viral inserts by specific hybridization to B segment RNA in Northern blots (Huang *et al.*, 1986; Duncan *et al.*, 1987). A recombinant plasmid with a 2.8-kbp cDNA insert of the B segment of IPNV-Jasper was isolated and subcloned into M13 mp18 and mp19 for DNA sequencing. A recombinant plasmid containing a 2.6-kbp insert of the B segment for IPNV-Sp was constructed in pT7 from two overlapping cDNA clones (Huang *et al.*, 1986) and selected restriction fragments from this recombinant plasmid were recloned into M13 mp18 and mp19 for sequence analysis. Subsequently, the large cDNA inserts were subcloned into Gemini vectors (Promega Biotec) for *in vitro* transcription and translation.

DNA sequencing and analysis

The sequence of the IPNV genome segment B for both virus strains was determined by the dideoxy chain termination procedure (Sanger *et al.*, 1977) using standard 6% polyacrylamide, 8 *M* urea gels or buffer gradient gels (Biggin *et al.*, 1983). The sequences were determined by sequencing overlapping subclones in both directions. The entire Jasper sequence was determined from both strands. Unambiguous sequence was determined in one direction for several small regions of the Sp sequence from several independently isolated cDNA clones. In addition, oligonucleotide primers complementary to Sp nucleotide positions 96–117, 681–697, 805–821, 1746–1763, 2079–2096, and 2371–2388 were used to confirm these sequences.

The cDNA sequence was analyzed on a VAX780 using the DNA sequence analysis package from the University of Wisconsin Genetics Computer Group (Devereux *et al.*, 1984).

RESULTS AND DISCUSSION

Nucleotide sequence comparison of IPNV segment B Jasper and Sp serotypes

Recombinant plasmids containing cDNA copies of the B segment of IPNV were detected by Northern blot analysis of cDNA libraries prepared from total genomic RNA of the Jasper and Sp serotypes of IPNV (Huang *et al.*, 1986; Duncan *et al.*, 1987). In the case of the IPNV-Jasper library, a recombinant plasmid containing an insert of approximately 2.8 kbp was identified and used for sequence analysis. The IPNV-Sp library contained no insert larger than 2 kbp. Thus, two of the IPNV-Sp recombinant plasmids were physically mapped using a series of restriction endonucleases and a large continuous insert was produced by ligation of insert fragments at overlapping restriction sites (Huang *et al.*, 1986). The reconstruction of the complete coding regions for both B segment RNAs was verified by *in vitro* translation of RNA transcripts of segment B and radioimmune precipitation of the VP1 protein (Huang *et al.*, 1986; R. Duncan, data not shown).

The nucleotide sequences of both serotypes of IPNV genome segment B were determined using the sequencing strategy outlined in Fig. 1 and the plus strand cDNA sequences are shown in Fig. 2. The sequences were aligned using the algorithm of Needleman and Wunsch (1970). The genome segment B sequence summary is outlined in Table 1.

The Sp sequence contained 2630 bp while the Jasper sequence consisted of 2784 bp. Both sequences were colinear with no insertions or deletions outside of the terminal sequences and with the exception of an additional codon present in the Jasper sequence at the end of the VP1 coding region (nucleotides 2633–2635). There were a total of 521 nucleotide mismatches representing an overall homology of 80.7%. In view of the high degree of homology between the Jasper and Sp segment B sequences, it seems likely that the short 3′-nontranslated region of the Sp cDNA indicates that the clone is truncated. Each sequence contained a single large ORF encoding

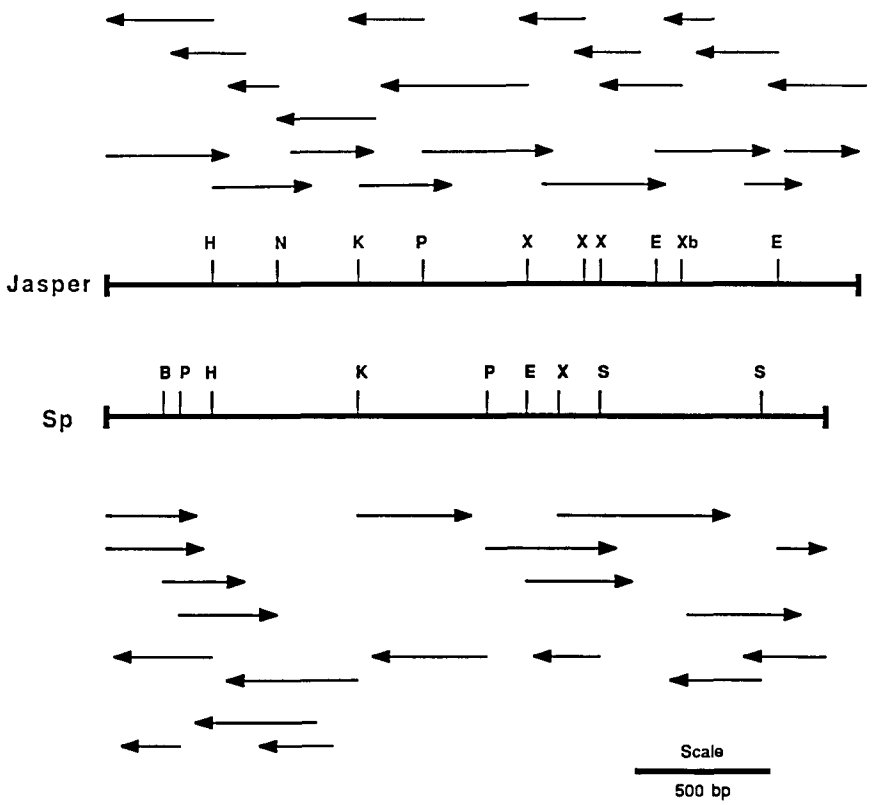

FIG. 1. Sequencing strategy of IPNV segment B Jasper and Sp serotypes. The partial physical maps of Jasper and Sp segment B cDNA are diagrammed. Arrows indicate the extent and direction of nucleotide sequence obtained from several overlapping clones. Restriction endonuclease sites: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Nhe*I; P, *Pst*I; S, *Sma*I; X, *Xho*I; Xb, *Xba*I.

the 845-amino-acid Jasper or 844-amino-acid Sp VP1 proteins. The 2532-nucleotide Sp ORF begins at nucleotide 94 and ends with a single TAA termination codon at nucleotide 2626, while the 2535-nucleotide Jasper ORF begins at nucleotide 101 and terminates at a single TAA termination codon at nucleotide 2636 (Fig. 2 and Table 1).

There are no published birnavirus terminal RNA sequences available; therefore, it is not possible to state that the sequences presented represent the entire genome segment B sequence. However, the importance of the terminal sequences in transcription and translation initiation and possibly genome packaging prompted an analysis of these regions to identify consensus sequence elements. Figure 3 shows a comparison of the 5′- and 3′-terminal sequences of IPNV-Jasper segment B with the previously published Jasper and N1 serotype segment A sequences (Duncan and Dobos, 1986; Haverstein *et al.*, 1990).

The Jasper segment A and B cDNA sequences both start with a conserved pentanucleotide sequence (GGAAA) and terminate with a conserved tetranucleotide sequence (CCCC). The NI serotype segment A sequence contains an additional 11 nucleotides at the 5′-end and the first G of the GGAAA pentanucleotide was replaced by an A. In addition, the reported N1 sequence ended immediately prior to the conserved 3′-terminal CCCC sequence present in the Jasper segment A and B sequences. There was extensive homology between the three sequences near the 5′- and 3′-termini while there was little homology between the segment A and B coding regions. When the 5′-terminal sequences of the A and B segments of IPNV-Jasper were optimally aligned, 32 of the 50 nucleotides were conserved. Similarly, 29 of 50 nucleotides at the 3′-terminus of B were conserved in the segment A sequence (Fig. 3). These extensive nucleotide sequence homologies presumably reflect regions important for genome replication and expression. In addition, as previously reported for the N1 segment A sequence (Haverstein *et al.*, 1990), the Jasper B sequence contained inverted terminal repeats of 14 nucleotides (Fig. 3). Unlike the segment A sequences, however, there was no indication of extensive *adjacent* inverted repeats in the segment B termini. The function of these small inverted repeats has not been determined. The consensus IPNV terminal sequences derived from the available IPNV sequence data are presented in Fig. 3. Refine-

```
Ja  GGAAACAGTG GGTCAACGTT GGTGGCACCC GACATACCAC GACTGTTTAC GTATGCACGC AAGTGCCCTT TAACAAAACC CTATACACAC AACTCATGAT          100
Sp  .......                                           T          G          T A  CCT  TACA T T      T                   93

Ja  ATG TCG GAC ATC TTC AAC TCA CCA CAA AAC AAG GCA TCC ATC TTG ACT GCT CTG ATG AAG AGC ACA ACA GGA GAC GTA GAG GAC GTG CTG  190
Sp                              T             T  T         A   A   C             G CAG          G       T  T  A    123

Ja  ATA CCA AAA CGC TTC AGG CCA GCC AAG GAC CCC CTC GAC AGC CCG CAA GCA GCA GCA CAG TTC CTG AAG GAC AAC AAG TAC CGG ATA CTT  280
Sp        C  G  G  T  A  C  A      T  G T A  T      C  G  T      C GC          A  A C        T                          273

Ja  AGG CCG CGA GCC ATC CCG ACC ATG GTC GAA CTA GAG ACA GAT GCC GCT CTG CCT CGA CTG CGA CAA ATG GTG GAA GAC GGC AAG CTT AAG  370
Sp                  A  C          T    A        G                          A GCG GCC          C  T                        363

Ja  GAC ACG GTA AGC GTC CCA GAA GGA ACC ACT GCG TTC TAC CCA AAG TAC TAC CCA TTC CAC AAG CCA GAC CAT GAC GAA GTG GGG ACG TTC  460
Sp    A  T   C  AT  T  C          A  C                 A                    GA  C             C      A      T    453

Ja  GGG GCT CCG GAC ATC ACA CTT CTG AAG CAA CTC ACC TTC TTC CTG TTG GAG AAT GAC TTC CCA ACA GGA CCG GAG ACG CTC AGA CAA GTC  550
Sp           A             A  C  A      G  G           C              T      T  T  A      C  A                          543

Ja  CGA GAA GCC ATA GCC ACA CTT CAG TAC GGC TCA GGC AGC TAC TCA GGA CAG CTG AAC AGG CTC CTA GCA ATG AAG GGC GTC GCC ACT GGC  640
Sp  A        A  C  A  C  G  A      G  C                    A  C         A  G                          G  G                 633

Ja  AGG AAT CCA AAC AAG ACT CCA AAA ACA GTG GGC TAC ACG AAC GAA CAG CTA GCA AAA CTG CTG GAG CAA ACA CTA CCG ATC AAC ACA CCA  730
Sp           C               CTG G C  T          C      G    A G     G     A           C T G  T        C T        723

Ja  AAG CAT GAG GAC CCC GAC CTC CGG TGG GCC CCA AGC TGG CTG ATC AAC TAC ACC GGA GAC CTG AGC ACA GAC AAG TCA TAT CTG CCA CAT  820
Sp      A           A          A                   T    A C G              GCA TCA  T                C  T               813

Ja  GTG ACC ATA AAG TCC TCA GCC GGC CTA CCA TAC ATA GGC AAA ACC AAA GGA GAC ACG ACG GCA GAA GCG CTC GTA CTG GCT GAC TCC TTC  910
Sp        A GCT      A  T            C                                C      C      C  G  G      A                      903

Ja  ATA CGT GAC CTC GGA AGA GCC GCC ACA TCA GCA GAC CCA GAA GCA GGT GTG AAG AAA ACC ATC ACC GAC TTC TGG TAC CTG AGC TGT GGG  1000
Sp      A G            A       A          C         GG  G  CA  CT        GTA C G T                    C      C  T    993

Ja  CTG CTC TTC CCA AAG GGC GAA AGA TAC ACA CAA GTG GAC TGG GAC AAG AAG ACC AGA AAC ATT TGG AGC GCG CCC TAC CCA ACA CAC CTA  1090
Sp                  A  G  G             G AAA           TT          G      C      T  C      T      G                      1083

Ja  CTA CTA TCA ATG GTG TCA ACC CCG GTG ATG AAC GAG TCC AAA CTC AAC ATC ACC AAC ACC CAG ACC CCG TCT CTG TAC GGG TTC TCA CCA  1180
Sp                      G T A           G T                                T          T                                  1173

Ja  TTC CAC GGA GGA ATG GAC AGA ATC ATG ACA ATC ATA AGA GAC AGC CTG GAC AAT GAC GAG GAC CTA GTC ATG ATC TAT GCG GAC AAC ATC  1270
Sp    T       T  G   C A              C      C      G CAT      T C A  G C                   A      C          A          1263

Ja  TAC ATA CTG CAG GAC AAC ACA TGG TAC TCA ATT GAC CTG GAG AAA GGC GAG GCC AAC TGC ACT CCA CAG CAC ATG CAG GCC ATG ATG TAC  1360
Sp          A             C          C  C  T  A  A  G  A  A  A          A      A             A                          1353

Ja  TAC CTC CTG ACG AGG GGA TGG ACG AAC GAG GAC GGC TCA CCG CGA TAC AAC CCG ACA TGG GCC ACA TTC GCA ATG AAC GTC GCT CCA TCC  1450
Sp       GT   C   A C C  A       A                     A  G             G      A          C    T    T  GG   C   A        1443

Ja  ATG GTG GTG GAC TCA TCG TGT CTT TTG ATG AAC CTT CAA CTG AAG ACC TAC GGC CAG GGC AGT GGG AAT GCC TTT ACA TTC CTG AAC AAC  1540
Sp        A            A C  C  G C        T  G  G          T ACG      A                     C  C       C                1533

Ja  CAC CTC ATG TCC ACA ATA GTC GTG GCC GAG TGG GTA AAA GCA GGG AAA CCA AAC CCC ATG ACC AAA GAG TTC ATG GAC CTC GAG GAG AAA  1630
Sp    T                 T  G  C  G        CAC  G      A GG      T        T        A                A  CA  G    1623
```

FIG. 2. Nucleotide sequence of IPNV genome segment B. The cDNA nucleotide sequence of the plus-strand B segment of IPNV Jasper (Ja) and Sp strains. Nucleotide mismatches in the Sp sequence are shown. (–) Locations of insertions/deletions; (●) sequences not present in the Sp cDNA. The sequence is grouped in triplets beginning at the start of the VP1 coding region. These sequences have been deposited with the GenBank nucleotide sequence database and have been assigned Accession Nos. M58756 (Jasper) and M58757 (Sp).

```
Ja  ACG GGC ATC AAC TTC AAG ATA GAG CGC GAG CTG AAA AAC CTT AGA GAG ACC ATC GTT GAG GCC GTC GAG ACG GCA CCC CAG GAT GGA TAC   1720
Sp        G              A  C                  G    A  G TC  T     A G      G  A  C  C  C  A TC  C  C  T   1713

Ja  CTC GCC GAT GGG TCC GAC CTA CCC CCC ATC AGA CCA GGA AAA GCA GTA GAG CTT GAC CTG CTC GGG TGG TCG GCC ATC TAC AGC CGC CAA   1810
Sp   A      C             G  A     GG GTG      G  G  G  G      C      T  A  A      C  A G G         A      1803

Ja  ATG GAG ATG TTC GTT CCA GTC CTC GAG AAC GAG AGA CTA ATT GCC TCA GCA GCC TAC CCA AAG GGG CTC GAG AAC AAA GCC CTG GCC AGA   1900
Sp  C C             C  C      T  A      A             A     TC             A      A           T    A  T C    1893

Ja  AAA CCC GGA GCC GAG ATT GCG TAC CAG ATA GTG CGG TAC GAG GCA ATC AGG CTG GTA GGC GGC TGG AAC AAT CCG CTA CTA GAG ACC GCG   1990
Sp          G          C  A      A      A A    T  A  G  T C    C A C  A                A  C A C  A  A  A   1983

Ja  GCC AAA CAC ATG TCC CTC GAC AAG AGG AAG AGA CTG GAA GTG AAG GGG ATA GAC GTC ACC GGA TTC CTC GAT GAC TGG AAC AAC ATG TCC   2080
Sp   A                 G      A                 G  A  A  C  C                    A  C           C       G   2073

Ja  GAA TTC GGA GGA GAC CTA GAA GGA ATA ACG CTA TCA GAA CCC CTA ACA AAC CAA ACA CTT GTT GAC ATC AAC ACG CCG CTG GAC AGC TTT   2170
Sp   G          C  T  G  G  T  C T A    A  CT      C         G  C  C C A            A  A GA  AC  GAG  C   2163

Ja  GAC CCC AAG GCA AGA CCC CAA ACA CCC AGG TCT CCA AAG AAG ACT CTA GAC GAG GTG ACT ACG GCC ATC ACC TCG GGC ACA TAC AAG GAC   2260
Sp     GT   A AC         CC  G  G C C  C    GGC  A  C  C C       A  C G A  G  A  A  A  G  C      A      2253

Ja  CCC AAA AGC GCT GTG TGG CGA CTG CTA GAC CAG AGG ACC AAA CTC CGA GTC AGC ACA CTA CGA GAT CAA GCA CTA GCG CTA AAA CCC GCG   2350
Sp          T  A        A G  C  C                   A  C  G      C      C      C  G AC               A   2343

Ja  TCG TCC TCG GTG GAC AAC TGG GCC GAA GCC ACT GAG GAA CTG GCA CAA CAG CAA CAG CTA CTC ATG AAG GCC AAC AAT CTG CTG AAG AGC   2440
Sp  G    A A A TCC    TT      GG  C      A  A      C  C G    A      A  G  G      A  G      C      A          2433

Ja  AGC CTG ACG GAG ACA AGA GAA GCA CTG GAG ACC ATC CAG TCA GAC AAA ATC ATC GCA GGG AAA TCC AAC CCC GAG AAG AAC CCG GGG ACC   2530
Sp       C      A G G  G      C  C  A    G G         T      A    T    C    A  TCT  A         T  C          2523

Ja  GCA GCC AAC CCA GTC GTT GGC TAC GGG GAA TTC AGT GAG AAG ATT CCT CTG ACC CCC ACG CAG AAA AAG AAT GCC AAG CGG AGG GAA AAG   2620
Sp   C  A      C  G  G CG  T  A  G      C      A         C  T         A  G      C         T C    G       2613

Ja  CAG AGA AGA AAC CAG TAA GAAGACCCA AACCGGGAAG AATCCGAAAT GAATCAGCTG GACTCATATG AAAGCTCCGC GCCGCACGGC AAGCTGGACA AAAGTAGTGA  2727
Sp                ---      ..... ..... ..... ..... ..... ..... ..... ..... .....  2630

Ja  CCCGACAACG TGCCACCAAC ATGACCCCTG AAAACATCCG GTTCCGCCAG GGACCCC   2784
```

FIG. 2—*Continued*

ment of these consensus sequences awaits further cDNA sequence analysis of other serotypes and direct genomic RNA sequence determinations.

The presence of perfectly conserved terminal nucleotides in the Jasper segment A and B cDNA sequences suggested that the cloned cDNAs represented complete genomic sequences. In fact, two large independent segment A cDNA clones were isolated and both contained identical terminal sequences (Duncan *et al.*, 1987). In addition, an identical sequence was obtained by direct RNA sequencing of the 3′-end of the minus-strand genome segment A and B RNAs (E. Nagy and P. Dobos, unpublished data). While it is not possible to state unequivocally that the cDNA sequences represent complete genomic RNA sequences, it is reasonable to assume that extensive terminal sequences are not missing, and that a complete sequence determination of the viral genome of IPNV-Jasper has been made.

Analysis of the segment B-encoded VP1 proteins

The predicted amino acid sequences of the IPNV-Jasper and -Sp serotype VP1 proteins are shown in Fig. 4 and were aligned with that of the IBDV VP1 (Morgan *et al.*, 1988). Translation commencing at the first methionine codon of the B segment ORF produced a polypeptide of 845 amino acids for IPNV-Jasper and 844 amino acids for IPNV-Sp. The amino acid number corresponded to the estimated molecular weight of VP1, which has ranged from 95 kDa for the Jasper isolate (Duncan *et al.*, 1987) to 89 kDa for the Sp and Ab isolates of IPNV (Hedrick *et al.*, 1983). The amino acid

TABLE 1

IPNV GENOME SEGMENT B SEQUENCE SUMMARY

	Serotype	
	Jasper	Sp
Segment B	2784	2630
5′-noncoding	100	93
3′-noncoding[a]	149	8
VP1 ORF (nucleotides[b])	101–2635	94–2625
VP1 ORF (codons)	845	844
Predicted molecular weight	94,441	94,064
$_p$I	6.60	6.60

[a] Including the termination codon.
[b] Excluding the termination codon.

compositions of the three VP1 proteins were similar to each other; the only unusual feature was their very low cysteine content (three cysteines in IPNV, VP1, eight cysteines in IBDV).

The Jasper sequence contained eight potential N-linked glycosylation sites, five of which were conserved in the Sp sequence and two in the IBDV sequence (Table 2). The Sp and IBDV sequences each contained one additional potential N-linked glycosylation site not found in the Jasper sequence. It is not known whether any of these sites are glycosylated in the birnavirus VP1 proteins.

Recent evidence indicates that Jasper VP1 in IPNV-infected cells exists in multiple phosphorylated forms (P. Dobos, unpublished observation). A similar situation has been recently reported for poliovirus RdRp, where multiple species of 3Dpol were detected which differed in their phosphorylation level of serine residues. (Ransone and Dasgupta, 1989).

While there is no recognition sequence for phosphorylation, preferred flanking sequences have been identified (Hunter, 1982; Bramson *et al.*, 1984). The phosphorylation of Ser residues by cAMP-dependent protein kinases exhibits a marked preference for one or two basic amino acids located one or two residues upstream of the Ser or Thr target site. Similarly, Tyr-specific kinases like pp60^{Src} preferentially phosphorylate Tyr residues that are preceded by an acidic residue four amino acids upstream. The VP1 sequences were scanned for such motifs and the results are summarized in Table 2. Jasper VP1 contained six potential Ser phosphorylation sites, four of which were found at the same sites in Sp VP1, and one was conserved in IBDV VP1. In addition, Sp VP1 contained one, and IBDV VP1 contained five, unique potential Ser phosphorylation sites. There was one potential Tyr phosphorylation site in Jasper which was conserved in Sp, while IBDV VP1

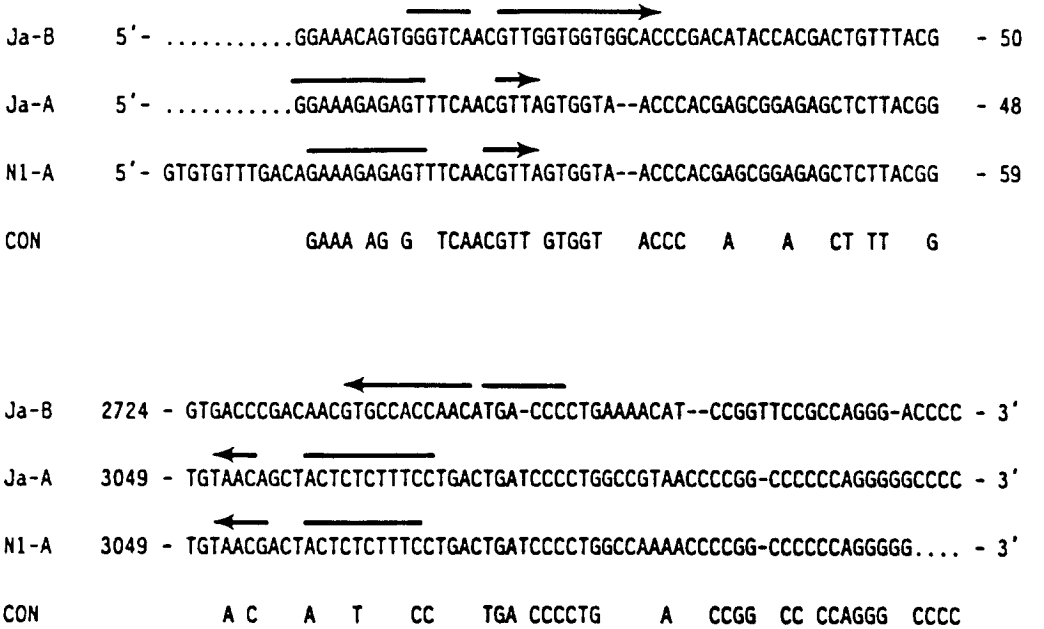

FIG. 3. IPNV terminal cDNA sequences. The 5′-terminal (top) and 3′-terminal (bottom) plus-strand cDNA sequences of IPNV Jasper segment B (Ja-B) and A (Ja-A) and IPNV N1 segment A (NI-A) are listed. Insertions (–) were introduced to optimize the alignment. (●) Sequences not present in the cDNA. The numbers refer to the last nucleotide listed of the 5′-terminal sequence or the first nucleotide listed of the 3′-terminal sequence. Arrows above the sequence indicate the location of the inverted terminal repeats which appear in the 5′- and 3′-terminal regions. A consensus sequence (con) for the IPNV terminal sequences is shown.

```
Ja  MSDIFNSPQNKASILTALMKSTTG--DVEDVLIPKRFRPAKDPLDSPQAAAQFLKDNKYRILRPRAIPTMVELETDAALPRLRQM--VEDGKLKDTVSVPEGT  99
Sp  MSDIFNSPQNKASILNALMKSTQG--DVEDVLIPKRFRPAKDPLDSPQAAAAFLKEHKYRILRPRAIPTMVEIETDAALPRLAAM--VDDGKLKEMVNVPEGT  99
IB  MSDVFNSPQARTKISAAFGIKPTAGQDVEELLIPKVWVPPEDPLASPSRLAKFLRENGYKILQPRSLPENEEYETDQILPDLAWMRQIEGAVLKPTLSLHWGP 103
    *** *****   * *        ***  ****  *  *** **  * **   * ** **  *  * ***  ** * *      **      *

Ja  TAFYPKYYPFHKPDHDEVGTFGAPDITLLKQLTFFL-LENDFPTGPETLRQVREAIATLQYGSGSYSGQLNRLLAMKGVATGRNPNKTPKTVGYTNEQLAKLL 201
Sp  TAFYPKYYPFHRPDHDDVGTFGAPDITLLKQLTFFL-LENDFPTGPETLRQVREAIATLQYGSGSYSGQLNRLLAMKGVATGRNPNKTPKLAGYTNEQMARLM 201
IB  RVL-PKVLLNSPPEQGKAQCVPTRHCTTQADIYLFLQVPEATESLKDEVTLLTQNIRDKAYGSGTYMGQATRLVAMKEVATGRNPNKDPLKLGYTFESIAQLM 205
        **      *            *      **                  *    **** * **  ** *** ********* *   *** *  * *

                                                    GTP
Ja  EQTLPINTPKHED---PDLRWAPSWLINYTGDLSTD---KSYLPHVTIKSSAGLPYIGKTKGDTTAEALVLADSFIRDLG-----RAATSADPEAGVKKTITD 293
Sp  EQTLPINPPKNED---PDLRWAPSWLIQYTGDASTD---KSYLPHVTAKSSAGLPYIGKTKGDTTAEALVLADSFIRDLG-----KAATSADPGAAAKKVLSD 293
IB  DITLPVGPPGEDDKPWVPLTRVPSRMLVLTGDVDGDFEVEDYLPKINLKSSSGLPYVGRTKGETIGEMIAISNQFLRELSALLKQGAGTKGSNKKKLLSMLSD 308
      ***   *   *     *   **     ***   *     ***    *** **** * *** *  *       * * *       * *           *

Ja  FWYLSCGLLFPKGERYTQVDWDKKTRNIWSAPYPTHLLLSMVSTPVMNESKLNITNTQT-PSLYGFSPFHGGMDRIMTIIRDSLDNDEDLVMIYADNIYILQN 395
Sp  FWYLSCGLLFPKGERYTQKDWDLKTRNIWSAPYPTHLLLSMVSSPVMDESKLNITNTQT-PSLYGFSPFHGGINRIMTIIREHLDQEQDLVMIYADNIYILQD 395
IB  YWYLSCGLLFPKAERYDKSTWLTKTRNIWSAPSPTHLMISMITWPVMSNSPNNVLNIEGCPSLYKFNPFRGGLNRIVEWI---LAPEEPKALVYADNIYIVHS 408
     *********** ***    *  ********* ****  **   ***  *  *  *    **** * ** **  **   *   *         *******

          1                                                                                  2
Ja  NTWYSIDLEKGEANCTPQHMQAMMYYLLTRGWTNEDGSPRYNPTWATFAMNVAPSMVVDSSCLLMNLQLKTYGQGSGNAFTFLNNHLMSTIVVAEWVKAGKPN 498
Sp  NTWYSIDLEKGEANCTPQHMQAMMYYRLTREWTNEDGSPRYNPTWATFAMYVGPSMVVDSTCLLMNLQLKTTGQGSGNAFTFLNNHLMSTIVVAEWHKAGRPN 498
IB  NTWYSIDLEKGEANCTRQHMQAAMYYILTRGW-SDNGDPMFNQTWASFAMNIAPALVVDSSCLIMNLQIKSYGQGSGNAATFINNHLLSTLVLDQWNLMKQPN 510
    ***************** ***** *** *** *    * *  * *** ***   *  **** ** **** *  ******* ** **** ** *   *     **

                  3
Ja  PMTKEFMDLEEKTGINFKIERELKNLRETIVEAVETAPQDGYLADGSDLPPIRPGKAVELDLLGWSAIYSRQMEMFVPVLENERLIASAAYPKGLENKALARK 601
Sp  PMSKEFMDLEAKTGINFKIERELKDLRSIIMEAVDTAPLDGYLADGSDLPPRVPGKAVELDLLGWSAVYSRQLEMFVPVLENERLIASVAYPKGLENKSLARK 601
IB  PDSEEFKSIEDKLGINFKIERSIDDIRGKLRQLVPLA-QPGYLSGG--VEPEQSSPTVELDLLGWSATYSKDLGIYVPVLDKERLFCSAAYPKGVENKSLKSK 610
    *  **   * * ********     *         *   ***  *    *      ********** **      ****  ***  * ***** *** * *

Ja  PGAEIAYQIVRYEAIRLVGGWNNPLLETAAKHMSLDKRKRLEVKGIDVTGFLDDWNNMSEFGGDLEGITLSEPLTNQTLVDINTPLDSFDPKARPQTPRSPKK 704
Sp  PGAEIAYQIVRYEAIRLIGGWNNPLIETAAKHMSLDKRKRLEVKGIDVTGFLDDWNTMSEFGGDLEGISLTAPLTNQTLLDINTPETEFDVKDRPPTPRSPGK 704
IB  VGIEQAYKVVRYEALRLVGGWNYPLLNKACKNNASAARRHLEAKGFPLDEFLAEWSELSEFGETFEGFNIKLTVTRENLAELNKPV---PPKPPNVNRPVNTG 713
     * * **  ***** ** **** **   * *      *  ** **     **  *   ****   **       *   *   * *      *

Ja  TLDEVTTAITSGTYKDPKSAVWRLLDQRTKLRV-STLRDQALALKPASSSVD----NWAEATEELAQQQQLLMKANNLLKSSLTETREALETIQSDKIIAGKS 802
Sp  TLAEVTAAITSGTYKDPKSAVWRLLDQRTKLRV-STLRDHAHALKPAASTSD----FWGDATEELAEQQQLLMKANNLLKSSLTEAREALETVQSDKIISGKT 802
IB  GLKAVSNALKTGRYRNEAGLSGLVLLATARSRLQDAVKAKAEAEKLHKSKPDDPDADWFERSETLSDLLEKADVASKVAHSALVETSDALEAVQSSSVYTPKY 813
     * * *   * *          *      *        * * *   * *     *    * *         *     * * *   ***  **      *

Ja  NPEKNPGTAANPVVGY-----------------GEFSEKIPLTPTQKKNAKRREKQRRNQ-----  845
Sp  SPEKNPGTAANPVVAY-----------------GEFSEKIPLTPTQKKNAKRREKQRRN------  844
IB  PEVKNPQTASNPVVGLHLPAKRATGVQAALLGAGTSRPMGMEAPTRSKNAVKMAKRAQRQKESRQ  878
      *** ** ****                  *         **  ***    *
```

FIG. 4. Amino acid sequence comparisons of birnavirus VP1 proteins. The predicted amino acid sequences of the IPNV Jasper (Ja) and Sp VP1 proteins were aligned with the sequence of IBDV (IB) VP1 using the algorithm of Needleman and Wunsch (1970) and insertions (–) were introduced to optimize the alignment. The amino acids are numbered on the right. Residues conserved in all three proteins are indicated (*). The conserved GTP-binding motif (GTP) and RdRp consensus sequence elements 1, 2, and 3 are overlined (see text).

TABLE 2

POTENTIAL PHOSPHORYLATION AND GLYCOSYLATION SITES IN VP1

	Jasper	Sp	IBDV
N-linked glycosylation sites	N-184[a]	N-184[a]	N-422[b]
	N-226	N-344[a]	N-449[b]
	N-339	N-409[b]	N-643
	N-344[a]	N-437[b]	
	N-409[b]	N-595	
	N-437[b]	N-677[a]	
	N-658		
	N-677[a]		
Ser phosphorylation sites	S-13[a]	S-13[a]	S-15
	S-245[b]	S-245[b]	S-256[b]
	S-375	S-292	S-297
	S-738[a]	S-738[a]	S-304
	S-751	S-781[a]	S-441
	S-781[a]		S-716
Tyr phosphorylation sites	Y-399[a]	Y-399[a]	Y-73
			Y-247

Note. N, Asn; S, Ser; Y, Tyr. Numbers refer to the amino acid.
[a] Conserved between Ja and Sp.
[b] Conserved between Ja, Sp, and IBD.

contained two unique potential Tyr phosphorylation sites. The role of these residues on VP1 phosphorylation and the role of phosphorylation on VP1 function are presently unclear.

The 507 nucleotide substitutions in the VP1 coding region between the two viral serotypes resulted in only 95 amino acid changes in the VP1 proteins. The majority of nucleotide changes (75%) occurred in the third position of a codon, only 17% of which resulted in amino acid substitutions. This led to an amino acid homology of 88.6%.

The comparison of the IPNV VP1 sequences with the 877 amino acid IBDV sequence revealed 365 amino acids conserved between all three proteins. This represented an overall homology of only 41%, extremely low compared to the homology that exists between the RdRps of the enteroviruses (74% between Coxsackie virus B3 and poliovirus, or 64% between Coxsackie virus B3 and human rhinovirus 14; Lindberg *et al.*, 1987). While the overall homology between IPNV and IBDV VP1 was quite low, several large regions of extensive homology particularly through the central regions of the proteins were observed (Fig. 4). The first 293 amino acids of IPNV were approximately 38% homologous with IBDV, but one region of 37 amino acids (positions 159–195) exhibited a 70% homology. The region spanning residues 294–487 of IPNV contained 62% homologous amino acids and represented the most extensively conserved segment of the three proteins; the last 102 amino acids in this region (positions 386–487) were 71% homologous. The next 140 residues (positions 487–626) were 51% homologous and once again, a stretch of 41 amino acids (positions 556–590) in this region contains 70% homology with IBDV. The carboxy-terminal 220 amino acids represented the most divergent region of the three proteins; only 28% of the residues were homologous. This region also contained a major insertion of 17 amino acids near the carboxy terminus of IBDV. The low level of amino acid homology over the last quarter of VP1 and the large insertion in the IBDV sequence suggested that the function of this portion of the protein was dependent on the presence of similar, rather than identical, amino acids that impart a particular secondary or tertiary structure on the protein.

A comparison of the predicted secondary structures of the three VP1 proteins was made. The proteins were predicted to contain approximately 30% of their residues in an α helix conformation and in β sheet structures. The remaining 40% of the amino acids were predicted to exist in turns and random coil or loop structures characteristic of globular proteins. Hydropathy plots revealed a protein with continuously alternating hydrophobic and hydrophilic regions also suggestive of a large globular protein (data not shown). Although the amino acid sequences at the carboxy termini of these proteins showed evolutionary divergence, there were several conserved structural motifs. The carboxy termini of all three proteins were highly hydrophilic as reflected by their basic nature. Approximately 30% of the carboxy-terminal 30 amino acids were basic. The presumptive role of VP1 in RNA replication suggests a possible role for this region in RNA association.

Interestingly, two of the most extensively conserved secondary structure motifs resided in the carboxy-terminal 250 amino acids. A large region of predicted β sheet-turn-α helix (residues 605–642 of IPNV) and an approximately 37-amino-acid α helix (residues 757–793) were predicted for all three proteins, by both the Garnier *et al.* (1978) and Chou and Fassman (1978) algorithms (data not shown). The low amino acid homology and the conserved predicted structural motifs support the notion that this region of VP1 may not be directly involved in the putative catalytic activity of this protein, but may play a structural role in facilitating VP1 function.

The VP1 protein is the presumptive birnavirus RdRp. Comparative analysis of the amino acid sequences of nucleotide binding proteins and putative RdRps has revealed the presence of conserved sequence motifs

which may be involved in the enzymatic functions of this group of proteins. The sequence GXXXXGKS/T is a constant motif present in *ras*-type GTP binding proteins (Argos and Leberman, 1985; Moller and Amons, 1985) and is found in several viral proteins with a tentative role in RNA replication (Kaariainen *et al.,* 1987). The same sequence is present in IPNV VP1 between residues 248 and 255 (GLPYIGKT). The corresponding region of IBDV VP1 is GLPYVGRT, a closely related sequence motif also found in the VSV N protein (Gallione *et al.,* 1981) and similar to the sequence GFIIKGRS found in the influenza A PA protein (Fields and Winter, 1982). As such, this region represents a potential GTP binding site in the VP1 of birnaviruses. This sequence may relate to the recently reported guanylyl transferase activity associated with IBDV VP1 (Spies and Muller, 1990), and IPNV VP1 (P. Dobos, unpublished data).

Kamer and Argos (1984) were the first to report the presence of conserved sequence motifs in the aligned sequences of the picornavirus polymerase, 3Dpol and putative RdRps from other plus-strand RNA viruses of plants and animals. Three regions of similarity which were grouped in the central region of the proteins were detected. The most distinctive pattern detected was a GDD motif flanked by predominantly hydrophobic residues. This motif is highly conserved in putative RdRps and has been proposed to represent the functional site of RdRps (Kamer and Argos, 1984; Gorbalenya *et al.,* 1989).

We scanned the IPNV VP1 sequence for the conserved RdRp motifs and aligned these regions with the corresponding regions of 38 putative RdRps from several plus-strand RNA and dsRNA virus families of animals and plants. Representative sequences from each of the virus families examined are listed in Fig. 5. The first two conserved regions, appropriately spaced, were clearly present in IPNV VP1 in the highly homologous central portion of the protein. The generalized consensus sequence for these regions is DXXXXD/E$(X)_{45-65}$S/TGXXXTXXXN, where X represents any amino acid. In fact, the amino acids flanking the absolutely conserved residues show a decidedly nonrandom distribution leading to a more extensive consensus sequence as outlined at the top of Fig. 5. The only exceptions to the absolutely conserved amino acids in the consensus sequence were found in maize chlorotic mottle virus (MCMV), which substitutes Ile for Thr in region 2, and reovirus, which substitutes His for Asn.

As previously reported (Kamer and Argos, 1984; Argos, 1988; Gorbalenya *et al.,* 1989), the GDD sequence in region 3 is a highly conserved motif that is present in almost all putative RdRps. The IPNV VP1 proteins represent the only putative RdRps that deviate from the conserved GDD motif so typical of RdRps. While IBDV has conserved the Asp–Asp sequence, IPNV does not contain this motif in the corresponding region of its VP1. The Sp sequence contained a Lys–Asp while the Jasper sequence showed even less conservation with a Lys–Asn sequence at the corresponding region of VP1 (Fig. 5). This represents a significant change in the proposed catalytic site with acidic residues being replaced by a basic residue of very different size and a polar residue of similar size.

Although IBDV contained the Asp–Asp sequence in region 3, the flanking residues show a high degree of variation from the consensus sequence. As evident in Fig. 5, the flanking residues are hydrophobic or neutral, while the birnavirus flanking residues are for the most part charged or polar hydrophilic residues. Argos (1988) has recently analyzed the GDD motif and found that the Asp–Asp or Asp–Thr–Asp sequence is present in numerous RNA-dependent and DNA-dependent polymerases, as well as several reverse transcriptases. An analysis of these sequences led to the development of a series of rules governing the composition of the residues flanking the Asp–Asp sequence present in this 15-amino-acid consensus sequence (Argos, 1988).

Interestingly, none of the putative RdRps from dsRNA viruses in Fig. 5 fulfil the Argos rules for a core polymerase sequence, even those that contain the GDD motif (reovirus matches the consensus sequence at the top of Fig. 5, but the Gln in position 6 does not agree with the Argos core polymerase sequence). All of the plus-strand RNA viruses fulfill all of the rules except for infectious bronchitis virus, which has a Ser at position 7. It is possible that there is something intrinsically different about RNA polymerases from dsRNA viruses, such that the regions flanking the "core" are structurally distinct from RdRps of plus-strand RNA viruses. Alternatively, the Asp–Asp sequence may represent a highly conserved structural motif, and not the active site, which is more diverged in the case of RdRps from the dsRNA virus families. The fact that IPNV VP1 lacks the GDD sequence motif would seem to argue in favor of the latter possibility. Regions 1 and 2 do not appear in DNA-dependent RNA polymerases or reverse transcriptases (Argos, 1988) and, as such, represent conserved sequence motifs of RdRps.

The absence of the GDD motif in IPNV VP1 should serve as a caution for anyone analyzing the sequence of a putative RdRp. The GDD sequence has become the hallmark of RdRps and may mislead others to

	VIRUS	REGION 1			REGION 2			REGION 3	
			* *		* *	* *		* * *	
	CONSENSUS	J J Z B D J O Z U D Z O (E)			S G - B B T B J B N B J J O B J (T)			- - J J B B G D D B J J - -	
1	Polio	229 - F A - F * Y T G Y * A S	-47-		* * C S G * S I F * S M I N N L		-17-	L K M I A Y * * * V I A S Y	3D-pol
2	CB3	229 - L I A F * Y S G Y * A S	-48-		* * C S G * S I F * S M I N N I		-17-	F R M I A Y * * * V I A S Y	3D-pol
3	EMC	230 - V Y D V * Y S N F * S T	-50-		* * C A A * S M L * T I M N N I		-17-	V K V L S Y * * * L L V V A	E52
4	FMDV	235 - V W D V * Y S A F * A N	-50-		* * C S A * S I I * T I L N N I		-17-	Y T M I S Y * * * I V V A S	p56a
5	SBV	365 - V L E T * I A S F * K S	-51-		* * M F L * L F V * T V L N V V		-13-	R C A A F I * * * N I I H G	nsP4
6	MBV	365 - V L E T * I A S F * K S	-51-		* * M F L * L F V * T M L N M t		-13-	K C A A F I * * * N I V H G	ns72
7	YFV	531 - F Y A d * t A G W * T R	-60-		* * Q V V * Y A L * T I t N L k		-41-	K R M A V S * * * C V V R P	NS5
8	IBV	572 - L M G W * Y P K c * R A	-56-		* * D A T * A Y A * S V F N I I		-55-	F S L M I L s * * G V V C Y	F2 ORF
9	REO	580 - s I N I * I S A c * A S	-89-		* * S T A * S t e h T A n N S T		-29-	R N Y V C Q * * * G L M I I	lambda 3
10	ROTA	515 - V L y T * V S Q W * S S	-63-		* * E k Q * k A A * S I A N L A		-17-	K I I r V d * * * N Y A V L	VP1
11	BTV-10	664 - Y I R L * e S E r * K G	-45-		* * E N S * L I A * S M H N M A		-19-	L S e Q Y V * * * T L F Y T	VP1
12	IBDV	411 - W Y S I * L e K g E A N	-56-		* * N A A * F I N * H L I S T L		-27-	F K I e r S i * * I r g K L	VP1
13	IPNV-Ja	398 - W Y S I * L e K g E A N	-57-		* * N A F * F L N * H L M S T I		-27-	F K I e r e l k n L r e T I	VP1
14	IPNV-Sp	398 - W Y S I * L e K g E A N	-57-		* * N A F * F L N * H L M S T I		-27-	F K I e r e l k * L r s T I	VP1
15	CPMV	274 - V L c C * Y S S F * G l	-52-		* * F P M * V I V * S I F N e I		-27-	I G L V T Y * * * N L I S V	87K
16	TMV	163 - V L E L * I S K Y * K S	-51-		* * D V T * F I G * T V I i A A		-11-	I K g A F C * * * S L L Y F	183K
17	BMV	267 - F L E A * L S K F * K S	-52-		T * D A F * Y F G * T L V T M A		-11-	D C A I F S * * * S L I I S	RNA 2
18	TEV	243 - Y C D A * g S Q F * S S	-54-		* * Q P S * V V d * T L M v I I		-15-	I V Y Y V N * * * L L I A I	54K
19	CNV	524 - A I G L * A S R F * Q h	-53-		* * D I N * S L G * Y L L m C A		-13-	Y S L A N C * * * C V L I V	92K
20	MCMV	663 - A I G F * A S R F * Q N	-53-		* * D M N i S L G * C I L A T A		-12-	A R L I N N * * * N V L I C	111K

FIG. 5. Conserved RdRp consensus sequences. The amino acid sequences of conserved domains present in putative RdRps from the members of several plus-strand (1–8) and double-stranded RNA (9–14) virus families of animals and plus-strand RNA plant virus families (15–20) are aligned with the corresponding regions of birnavirus VP1 proteins. Numbers refer to the numbers of amino acids between the domains and from the predicted amino terminus of the proteins. The right-hand column lists the names of the proteins or genes where assigned, or the approximate size (in mol wt × 1000) of the gene product. The asterisks in the body of the table indicate direct amino acid conservation with the consensus sequence listed at the top of the figure. A double asterisk above the consensus sequence indicates an absolutely conserved position, a single asterisk denotes a highly conserved position. Upper case letters indicate that the amino acid conforms to the generalized consensus sequence with the following groups: J, hydrophobic (LIVMFYWCHA); O, neutral (GASTNQP); B, hydrophobic or neutral (J + O); Z, neutral or charged (O + KDER); U, aromatic (FYW); (–), any amino acid. Viruses and source of sequence—picornaviruses: polio, poliovirus type 1; EMC, encephalomyocarditis virus; FMDV, foot and mouth disease virus (Kamer and Argos, 1984, and references therein); CB3, Coxsackie virus B3 (Lindberg *et al.*, 1987). Togaviruses: SBV, Sindbis virus; MBV, middleberg virus (Strauss *et al.*, 1983). Flaviviruses: YFV, yellow fever virus (Rice *et al.*, 1985). Coronavirus: IBV, infectious bronchitis virus (Boursnell *et al.*, 1987). Reoviruses: REO, mammalian reoviruses serotypes 1, 2, and 3 (Weiner and Joklik, 1989); ROTA, bovine rotavirus (Cohen *et al.*, 1989); BTV, blue tongue virus serotype 10 (Roy *et al.*, 1988). Birnaviruses: IBDV, infectious bursal disease virus (Morgan *et al.*, 1988); IPNV-Ja and Sp—this report. Plant viruses: CPMV, cowpea mosaic virus; TMV, tobacco mosaic virus; BMV, brome mosaic virus (Kamer and Argos, 1984, and references therein); TEV, tobacco etch virus (Allison *et al.*, 1986); CNV, cucumber necrosis virus (Rochon and Tremaine, 1989); MCMV, maize chlorotic mottle virus (Nutter *et al.*, 1989).

merely scan their sequence for the presence of this motif. Indeed, outside of IPNV the only exceptions to the GDD motif are IBDV (IDD) and IBV (JDD) which were not identified as RdRps until more recently analyzed (Gorbalenya and Koonin, 1988; Gorbalenya *et al.*, 1989). Our data indicate that even the Asp–Asp sequence may not be present in some putative RdRps.

Proof of the role of birnavirus VP1 as a RdRp and the regions involved in polymerase function awaits direct experimental analysis and the development of a functional *in vitro* polymerase assay.

ACKNOWLEDGMENTS

Work performed at the University of Guelph was supported by the Natural Sciences and Engineering Research Council of Canada. R.D. is an Alberta Heritage Foundation for Medical Research Postdoctoral Fellow. Portions of this work were performed in Dr. Patrick Lee's laboratory at the University of Calgary. The work performed at Oregon State University was supported by Oregon Sea Grant with funds from the National Oceanic and Atmospheric Administration, Office of Sea Grant, Department of Commerce, under Grants NA81AA-D-00086 (Project R/FSD-9) and NA85AA-D-SG095 (Project R/FSD-11 and R/FSD-13) and from appropriations made by the Oregon State Legislature.

REFERENCES

ALLISON, R., JOHNSTON, R. E., and DOUGHERTY, W. G. (1986). The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: Evidence for the synthesis of a single polyprotein. *Virology* **154,** 9–20.

ARGOS, P. (1988). A sequence motif in many polymerases. *Nucleic Acids Res.* **16,** 9909–9915.

ARGOS, P., and LEBERMAN, R. (1985). Homologies and anomalies in primary structural patterns of nucleotide binding proteins. *Eur. J. Biochem.* **152,** 651–656.

AZAD, A. A., BARRETT, A., and FAHEY, KH. (1985). The characterization and molecular cloning of the double-stranded RNA genome of an Australian strain of infectious bursal disease virus. *Virology* **143,** 35–44.

BERNARD, J. (1980). *Drosophila* X RNA polymerase: Tentative model for *in vitro* replication of the double-stranded virion RNA. *J. Virol.* **33,** 717–723.

BIGGIN, M. D., GIBSON, T. J., and HONG, G. F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.

BOURSNELL, M. E. G., BROWN, T. D. K., FOULDS, I. J., GREEN, P. F., TOMLEY, F. M., and BINNS, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* **68,** 57–77.

BRAMSON, H. N., MILDVAN, A. S., KAISER, E. T., and CORBIN, J. D. (1984). Mechanistic studies of cAMP-dependent protein kinase action. *CRC Crit. Rev. Biochem.* **15**(2), 93–124.

BROWN, F. (1986). The classification and nomenclature of viruses: Summary of results of meetings of the International Committee on Taxonomy of Viruses in Sendai, September 1984. *Intervirology* **25,** 141–143.

CHOU, P. Y., and FASMAN, G. D. (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* **47,** 251–276.

COHEN, J., CHARPILIENNE, A., CHILMONCZYK, S., and ESTES, M. K. (1989). Nucleotide sequence of bovine rotavirus gene 1 and expression of the gene product in baculovirus. *Virology* **171,** 131–140.

DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395.

DOBOS, P., HILL, B. J., HALLETT, R., KELLS, D. T. C., BECHT, H., and TENINGES, D. (1979). Biophysical and biochemical characterization of five animal viruses with bisegmented double-stranded RNA genomes. *J. Virol.* **32,** 593–605.

DUNCAN, R., and DOBOS, P. (1986). The nucleotide sequence of infectious pancreatic necrosis virus (IPNV) dsRNA segment A reveals one large ORF encoding a precursor polyprotein. *Nucleic Acids Res.* **14,** 5934.

DUNCAN, R., NAGY, E., KRELL, P. J., and DOBOS, P. (1987). Synthesis of the infectious pancreatic necrosis virus polyprotein, detection of a virus-encoded protease, and fine structure mapping of genome segment A coding regions. *J. Virol.* **61,** 3655–3664.

FIELDS, S., and WINTER, G. (1982). Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment. *Cell* **28,** 303–313.

GALLIONE, C. J., GREENE, J. R., IVERSON, L. E., and ROSE, J. K. (1981). Nucleotide sequences of the mRNAs encoding the vesicular stomatitis virus N and NS proteins. *J. Virol.* **39,** 529–535.

GARNIER, J., OSGUTHORPE, D. J., and ROBSON, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120,** 97–120.

GORBALENYA, A. E., and KOONIN, E. V. (1988). Birnavirus RNA polymerase is related to polymerases of positive RNA viruses. *Nucleic Acids Res.* **16,** 7735.

GORBALENYA, A. E., KOONIN, E. V., DONCHENKO, A. P., and BLINOV, V. M. (1989). Coronavirus genome: Prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. *Nucleic Acid Res.* **17,** 4847–4861.

HAVERSTEIN, L. S., KALLAND, K. H., CHRISTIE, K. E., and ENDRESEN, C. (1990). Sequence of the large double-stranded RNA segment of the N1 strain of infectious pancreatic necrosis virus: A comparison with other Birnaviridae. *J. Gen. Virol.* **71,** 299–308.

HEDRICK, R. P., OKAMOTO, M., JANO, T., and FRYER, J. L. (1983). Biochemical characterization of eel virus European. *J. Gen. Virol.* **64**(6), 1421–1426.

HUANG, M. T. F., MANNING, D. S., WARNER, M., STEPHENS, E. B., and LEONG, J. C. (1986). A physical map of the viral genome for infectious pancreatic necrosis virus sp: Analysis of cell-free translation products derived from viral cDNA clones. *J. Virol.* **60,** 1002–1011.

HUDSON, P. J., MCKERN, N. M., POWER, B. E., and AZAD, A. A. (1986). Genomic structure of the large RNA segment of infectious bursal disease virus. *Nucleic Acids Res.* **14,** 5001–5012.

HUNTER, T. (1982). Synthetic peptide substrates for a tyrosine protein kinase. *J. Biol. Chem.* **257,** 4843–4848.

JAGADISH, M. N., STATON, V. J., HUDSON, P. J., and AZAD, A. A. (1988). Birnavirus precursor polyprotein is processed in *Escherichia coli* by its own virus-encoded polypeptide. *J. Virol.* **62,** 1084–1087.

KAARIAINEN, L., TAKKINEN, K., KERANEN, S., and SODERLUND, H. (1987). Replication of the genome of alphaviruses. *J. Cell Sci.* **7,** 231–250.

KAMER, G., and ARGOS, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12,** 7269–7283.

KIBENGE, F. S. B., JACKWOOD, D. J., and MERCADO, C. C. (1990). Nucleotide sequence analysis of genome segment A of infectious bursal disease virus. *J. Gen. Virol.* **71,** 569–577.

LINDBERG, A. M., STALHANDSKE, P. O. K., and PETTERSSON, U. (1987). Genome of coxsackievirus B3. *Virology* **156,** 50–63.

MACDONALD, R. D., and DOBOS, P. (1981). Identification of the proteins encoded by each genome segment of infectious pancreatic necrosis virus. *Virology* **114,** 414–422.

MERTENS, P. P. C., and DOBOS, P. (1982). Messenger RNA of infectious pancreatic necrosis virus is polycistronic. *Nature* **297,** 243–246.

MERTENS, P. P. C., JAMIESON, P. B., and DOBOS, P. (1982). *In vitro* RNA synthesis by infectious pancreatic necrosis virus-associated RNA polymerase. *J. Gen. Virol.* **59,** 47–56.

MOLLER, W., and AMONS, R. (1985). Phosphate-binding sequences in nucleotide-binding proteins. *FEBS Lett.* **186,** 1–7.

MORGAN, M. M., MACREADIE, I. G., HARLEY, V. R., HUDSON, P. J., and AZAD, A. A. (1988). Sequence of the small double-stranded RNA genomic segment of infectious bursal disease virus and its deduced 90-kDa product. *Virology* **163,** 240–242.

MULLER, H., and NITSCHKE, R. (1987). The two segments of the infectious bursal disease virus genome are circularized by a 90,000-Da protein. *Virology* **159,** 174–177.

NAGY, E., and DOBOS, P. (1984). Coding assignments of drosophila X virus genome segments: *In vitro* translation of native and denatured virion dsRNA. *Virology* **137,** 58–66.

NEEDLEMAN, S. B., and WUNSCH, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48,** 443–453.

NUTTER, R. C., SCHEETS, K., PANGANILAN, L. C., and LOMMEL, S. A.

(1989). The complete nucleotide sequence of the maize chlorotic mottle virus genome. *Nucleic Acids Res.* **17**(8), 3163.

PERSSON, R. H., and MACDONALD, R. D. (1982). Evidence that infectious pancreatic necrosis virus has a genome-linked protein. *J. Virol.* **44,** 437–443.

RANSONE, L. J., and DASGUPTA, A. (1989). Multiple isoelectric forms of poliovirus RNA-dependent RNA polymerase: Evidence for phosphorylation. *J. Virol.* **63,** 4563–4568.

REVET, B., and DELAIN, E. (1982). The *Drosophila* X virus contains a 1-μM double-stranded RNA circularized by a 67-kd terminal protein: High-resolution denaturation mapping of its genome. *Virology* **123,** 29–44.

RICE, C. M., LENCHES, E. M., EDDY, S. R., SHIN, S. J., SHEETS, R. L., and STRAUSS, J. H. (1985). Nucleotide sequence of yellow fever virus: Implications for flavivirus gene expression and evolution. *Science* **229,** 726–733.

ROCHON, D'A. M., and TREMAINE, J. H. (1989). Complete nucleotide sequence of the cucumber necrosis virus genome. *Virology* **169,** 251–259.

ROY, P., FUKUSHO, A., RITTER, G. D., and LYON, D. (1988). Evidence for genetic relationship between RNA and DNA viruses from the sequence homology of a putative polymerase gene of Bluetongue virus with that of vaccinia virus: Conservation of RNA polymerase genes from diverse species. *Nucleic Acids Res.* **16,** 11,759–11,767.

SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5453–5467.

SOMOGYI, P., and DOBOS, P. (1980). Virus-specific RNA synthesis in cells infected by infectious pancreatic necrosis virus. *J. Virol.* **33**(1), 129–134.

SPIES, U., and MULLER, H. (1990). Demonstration of enzyme activities required for cap structure formation in infectious bursal disease virus, a member of the birnavirus group. *J. Gen. Virol.* **71,** 977–981.

SPIES, U., MULLER, H., and BECHT, H. (1987). Properties of RNA polymerase activity associated with infectious bursal disease virus and characterization of its reaction products. *Virus Res.* **8,** 127–140.

SPIES, U., MULLER, H., and BECHT, H. (1989). Nucleotide sequence of infectious bursal disease virus genome segment A delineates two major open reading frames. *Nucleic Acids Res.* **17,** 7982.

STRAUSS, E. G., RICE, C. M., and STRAUSS, J. H. (1983). Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proc. Natl. Acad. Sci. USA* **80,** 5271–5275.

WEINER, J. R., and JOKLIK, W. K. (1989). The sequences of the reovirus serotype 1, 2 and 3 L1 genome segments and analysis of the mode of divergence of the reovirus serotypes. *Virology* **169,** 194–203.