Sequence of the large double-stranded RNA segment of the N1 strain of infectious pancreatic necrosis virus: a comparison with other Birnaviridae

L. S. Håvarstein,1* K. H. Kalland,3 K. E. Christie2 and C. Endresen1

1Laboratory of Biotechnology and 2Norbio A/S, University of Bergen, P.O. Box 3152, Årstad, N-5029 Bergen, Norway and 3Dana Farber Cancer Institute, Boston, Massachusetts 02115, U.S.A.

The cDNA sequence of the large dsRNA segment (segment A) of the N1 strain of infectious pancreatic necrosis virus (IPNV) has been determined. The nucleotide and deduced amino acid sequences were compared to the sequences of segment A of the Jasper strain of IPNV and to the sequences of segments A and B (5' and 3' flanking regions) of the 002-73 strain of infectious bursal disease virus (IBDV). The comparison demonstrated that the precursor protein of the major structural polypeptide, pVP2, is highly conserved at the N and C termini, whereas the amino acid sequence of an internal segment shows greater diversity between the strains. This internal segment probably carries the serotype-specific epitopes of birnaviruses. An alternative open reading frame (ORF) (444 bp) partly overlapping with the large ORF (2916 bp) of segment A was found to be conserved among the IPNV strains and is probably also present in the 002-73 strain of IBDV. This small ORF may encode a novel birnavirus polypeptide with an Mr of 17K. SDS–PAGE of radiolabelled purified IPNV particles revealed a band corresponding to the possible novel 17K polypeptide. Short terminal inverted repeats are found in segment A of the N1 and Jasper strains of IPNV and in segment B of the 002-73 strain of IBDV. Segment A of IPNV and segment B of IBDV also contain adjacent inverted repeats at their 5'-terminal flanking regions.

Introduction

Infectious pancreatic necrosis virus (IPNV) is the prototype of a family of viruses called the Birnaviridae (Dobos et al., 1979). Birnaviruses also include infectious bursal disease virus (IBDV) of domestic fowl (Müller et al., 1979) and drosophila X virus of Drosophila melanogaster (Teninges et al., 1979). IPNV causes an acute, contagious disease in juvenile salmonids, e.g. brook trout (Salvelinus fontinalis) and rainbow trout (Salmo gairdneri) (Hill, 1982). IBD is a highly contagious disease of young chickens, characterized by the destruction of the lymphoid cells in the bursa of Fabricius (Chevill, 1967).

Several strains of IPNV and IBDV with different immunological and/or biochemical properties have been described (Becht et al., 1988; Caswell-Reno et al., 1986; Kibenge et al., 1988). In this article only the birnavirus strains that have been characterized by cDNA cloning and nucleotide sequence analysis will be considered. In addition to the N1 strain, these include the Jasper strain of IPNV (Duncan & Dobos, 1986) and the Australian IBDV strain 002-73 (Azad et al., 1985; Hudson et al., 1986; Morgan et al., 1988).

Birnaviruses possess a bisegmented, dsRNA genome contained within an unenveloped, icosahedral capsid (Dobos & Roberts, 1983). The larger genome segment A (approx. 3100 bp) of IPNV (N1 and Jasper strains) encodes three proteins in a single large open reading frame (large ORF): the 60K to 62K precursor (pVP2) of the 52K to 54K major structural protein VP2, the 29K non-structural protein (NS) and the 31K minor structural protein VP3 (Duncan et al., 1987; Nagy et al., 1987). The corresponding M, of the segment A proteins of IBDV (strain 002-73) are 50K to 60K (precursor to VP2), 41K to 37K (VP2), 29K (VP4) and 32K (VP3) (Fahey et al., 1985; Kibenge et al., 1988). The smaller B segment (approx. 2900 bp) of birnaviruses encodes a single gene product (VP1) with an M, of approximately 90K, presumed to be the viral RNA polymerase (Gorbalenya & Koonin, 1988; Morgan et al., 1988). Nucleotide and peptide sequence analyses have shown that the large ORF of the A segment of IPNV is monocistronic and encodes a polyprotein in which three viral polypeptides are arranged in the order N-pVP2-NS-VP3-C (Duncan & Dobos, 1986; Hudson et al., 1986; Nagy et al., 1987). However, the precise borders of the three coding regions have not yet been defined (Duncan et al., 1987). The same applies to IBDV, apart from the fact that the protein equivalent to pVP2 is called VPX and the protein equivalent to NS is called VP4 (Kibenge et al., 1988).
VP4 is involved in the processing of the precursor polyprotein (large ORF gene product), in cleaving between VPX and VP4 and between VP4 and VP3 (Azad et al., 1987; Jagadish et al., 1988).

The Jasper strain of IPNV was originally isolated from rainbow trout of the Jasper river in Alberta, Canada. Recently we isolated a new serotype of IPNV from Atlantic salmon (Salmo salar) (Christie et al., 1988). This serotype, called the N1 strain, was detected in young hatchery-reared salmon from western Norway. We report here the cDNA cloning of the N1 strain and sequencing of the A segment of the N1 genome. The nucleotide and deduced amino acid sequences were compared to the corresponding sequences of the Jasper and the 002-73 strains. We also present results indicating that a conserved ORF, overlapping with the N-terminal part of VP2, encodes a fifth birnavirus polypeptide (VP5).

**Methods**

**Cells and viruses.** The culturing of the chinook salmon embryo cells (CHSE-214) and the propagation and purification of IPNV have been described previously (Christie et al., 1988). The Ab and Sp reference strains were obtained from P. E. V. Jørgensen (Denmark) and the Jasper strain was obtained from P. Dobos (Canada).

**Construction of a lambda gt10 cDNA library.** Synthesis of double-stranded cDNA from N1 genomic dsRNA was performed using random primers as described by Azad et al. (1985). A lambda gt10 cloning system (Amersham) was used to construct a recombinant lambda gt10 library from the cDNA. Briefly, the cDNA was methylated with EcoRI methylase, EcoRI linkers were ligated to the cDNA and the cDNA was ligated to EcoRI-digested and phosphatase-treated lambda gt10 arms. The products of the ligation were packaged using packaging extracts from Amersham.

**Isolation of cDNA clones.** The lambda library was initially screened with two 5' end-labelled synthetic oligonucleotides (40-mers). Their nucleotide sequences were taken from the 5' and 3' ends of segment A of the Jasper strain [nucleotides (nt) 31 to 70 and 3050 to 3089, Fig. 2]. Standard hybridization and washing conditions were used. Lambda phage DNA was isolated by using LambdaSorb phage adsorbent (Promega).

**cDNA sequencing.** Isolated virus clones were subcloned in both orientations into the pGEM-7ZF(+) vector (Promega). When necessary, DNA subclones of varying lengths were generated using the Erase-a-Base system (Promega). ssDNA, prepared by infection with the M13 K07 helper phage, was sequenced according to the dideoxynucleotide chain termination method (Sanger et al., 1977), using the Sequenase system (United States Biochemical Corporation). Both strands of the cDNA were sequenced completely. The sequence was compiled from independent, overlapping clones. In the sequence presented in Fig. 1 each base has been confirmed from at least two overlapping independent clones, except for the terminal 87 bp at the 5' end and the terminal 25 bp at the 3' end. All ambiguous or compressed sequence regions were resolved by substitution of dITP for dGTP in the sequencing reactions. Sequences were assembled and analysed using Staden-Plus software (Amersham). The hydrophobicity and charge plots were made using the method of Kyte & Doolittle (1982) with a window size of 13, as implemented in the Staden-Plus software. Sequence alignment was performed using the Align program (Scientific & Educational Software). Mismatch penalty, open gap penalty and extend gap penalty were the recommended values of 2, 4 and 1, respectively.

**SDS-PAGE of radiolabelled virus particles.** A monolayer of CHSE-214 cells was infected with 1 p.f.u. per cell (25 cm² flasks, Costar). The medium was exchanged 22 h post-infection with methionine-free medium containing 16 μCi [35S]methionine (Amersham) per ml. The infection was allowed to proceed to complete c.p.e. and the virus was harvested and purified as described earlier (Christie et al., 1988). SDS-PAGE was performed as originally described by Laemmli (1970). The Mr values of the viral polypeptides were estimated from their electrophoretic mobilities relative to 14C-labelled Mr markers (Amersham) run in parallel.

**Results**

**Nucleotide sequences of segment A**

The cDNA sequence of genome segment A of the N1 strain of IPNV is given in Fig. 1. The sequence is 3104 bp long and contains a large ORF of 2916 bp. The large ORF of the N1 strain is identical in length to the one already described for the Jasper strain (Duncan & Dobos, 1986). We found an identity of 79.5% between the nucleotide sequences of the A segments of the N1 and Jasper strains. A comparison with segment A of the 002-73 strain resulted in 54% identity (Table 1).

To look for possible new virus-encoded proteins we searched for conserved alternative ORFs on the A segments of the IPN viruses. Only one such perfectly conserved ORF was found, which is 444 bp long and overlaps with that encoding the N-terminal part of VP2 (Fig. 1). This small ORF has the potential of encoding a 17K peptide containing 148 amino acids. The initiation codon of the small ORF is located 79 bp from the 5' terminus of the A segment. It is closer to the 5' terminus than the ATG codon of the large ORF, which is located 131 bp from the 5' end of segment A (Fig. 1). Presumably, the 002-73 strain of IBDV also contains this small ORF. Unfortunately, much of the 5' flanking region is missing from the sequence published by Hudson et al. (1986). Therefore, 15 bp of the small ORF is not available and it is not possible to tell whether there is a correctly positioned initiation codon at the N-terminal end of the small ORF of the 002-73 strain.

The optimal sequence for initiation by eukaryotic ribosomes has been determined to be ACCATGG (Kozak's rule) (Kozak, 1986). At the least a functional initiation codon should be flanked by a neighbouring purine (usually A) at position -3 and/or a G at position +4. Neither the large nor the small ORF of segment A contains a perfect Kozak consensus sequence at their start region (Fig. 2). The large ORF start region contains
Table 1. Comparison between corresponding nucleic acid and amino acid sequences of the N1 and Jasper strains of lPNV and the 002-73 strain of lBDV*

<table>
<thead>
<tr>
<th>Segment A</th>
<th>ORF A</th>
<th>pVP2/VPX†</th>
<th>NS/VP4</th>
<th>VP3</th>
<th>VP5</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1/Jasper</td>
<td>NA 79-5</td>
<td>–</td>
<td>80-9</td>
<td>76-6</td>
<td>77-3</td>
</tr>
<tr>
<td>N1/IBDV</td>
<td>NA 53-8</td>
<td>–</td>
<td>88-8</td>
<td>79-9</td>
<td>78-2</td>
</tr>
<tr>
<td>Jasper/IBDV</td>
<td>NA 54-2</td>
<td>–</td>
<td>56-6</td>
<td>49-0</td>
<td>51-2</td>
</tr>
</tbody>
</table>

* Results are given as percentage of identity.
† The large ORFs were divided into three segments corresponding to the gene products pVP2/VPX, NS/VP4 and VP3; the putative novel peptide VP5 was also included in the analysis.
‡ NA, nucleic acid.
§ AA, amino acid.

the important purine in the −3 position, but lacks a G in position +4. For the small ORF it is the other way around.

The large ORF of segment A of the N1 strain is bracketed by 5′- and 3′-terminal flanking regions consisting of 130 bp and 55 bp, respectively (Fig. 1). These regions are highly conserved within the IPNV strains (Fig. 2). A comparison between the corresponding flanking regions of the A segments of IPNV and IBDV reveal no significant sequence similarity (data not shown).

Terminal inverted repeats were found 10 bp from the 5′ terminus (nt 11 to 32) and 37 bp from the 3′ terminus (nt 3047 to 3067) of segment A of the N1 strain (Fig. 3a). Segment A of the Jasper strain has identical sequences at its 5′ and 3′ termini. It is not possible to determine whether segment A of IBDV has equivalent inverted repeats, because sequence information for the 5′ terminus is missing. On the other hand, segment B of the 002-73 strain of IBDV contains terminal inverted repeats (nt 67 to 78 and nt 2746 to 2757; Fig. 3b). The terminal inverted repeats of the two segments contain similar sequences and have the motif 5′ AAGAG 3′ in common, but the repeats on segment B (5′ TCTTCTTCTT 3′/3′ AGGAGAAGAA 5′) are inverted relative to the repeats on segment A (5′ AGAAAGAGAG 3′/3′ TCTTCTTCTC 5′) (Fig. 3a and b; 5′ ends). In addition, adjacent inverted repeats are also found at the 5′-terminal flanking regions of segment A of IPNV (nt 41 to 82; open arrows, Fig. 3a) and segment B of IBDV (nt 17 to 61; open arrows, Fig. 3b).

Another puzzling sequence structure of IPN viruses is found approximately half way between the initiation codons of the large and the small ORF (Fig. 2). In the N1 strain this sequence can be read forwards and backwards to give the same sense (5′ TCTAACAAAAACAAAACAAATCCT 3′). The corresponding Jasper sequence has four mismatches compared to the N1 sequence (Fig. 2). All mismatches are positioned to the right side of the axis of symmetry when the sequence is written in the 5′ to 3′ direction.

Deduced amino acid sequences

The alignment of the deduced amino acid sequences of the large and small ORFs of the N1, Jasper and 002-73 strains is presented in Fig. 4 and Fig. 5, respectively. When the amino acid sequences of the N1 and Jasper polypeptides are compared, pVP2 proves to be the most conserved viral protein (88.8% identity), whereas the possible new gene product VP5 is least conserved (63.5% identity). On the other hand, if the IPNV gene products are compared to those of the 002-73 strain of IBDV, pVP2/VPX is still the most conserved polypeptide (approx. 44% identity), but the least conserved one has now changed to NS/VP4 (approx. 20% identity, Table 1).

The pVP2/VPX protein consists of two strongly conserved segments (the N and C termini) and a less conserved internal segment (I). The amino acid positions (Fig. 4) for these three segments are for the N1 and Jasper strains of IPNV 1 to 182 (N), 183 to 338 (I) and 339 to 470 (C), and for the 002-73 strain of IBDV 1 to 185 (N), 186 to 330 (I) and 331 to 469 (C). A comparison between the segments of the two strains of IPNV shows approximately 94% identity for the end segments and 78% identity for the internal segments. Similarly, a comparison between the IPNV strains and the 002-73 strain of IBDV resulted in approximately 50% and 25% identity for the end segments and the internal segments, respectively. The internal segments of the IPNV strains in particular have many amino acid replacements at positions 234 to 264 (Fig. 4). The same region also contains several possible N-glycosylation sites (N-X-S/T).

Charge plots of the putative VP5 polypeptides
Fig. 1. cDNA plus-strand sequence of segment A of the N1 strain and deduced amino acid sequences of the small ORF (nt 79 to 522) and the large ORF (nt 131 to 3046).
Sequence of segment A of the N1 strain of IPNV

Fig. 2. Alignment of the 5' and 3' flanking regions of segment A of the N1 and Jasper strains of IPNV. The sequence for the Jasper strain of IPNV is from Duncan & Dobos (1986).

| N1                  | 5'-GTGTGGTTGACGAAAGAGAGTTTCAACGGTTAGGGTCACCCACGACGAGCTCTTA |
| Jasper              | 5'-GGAAAGAGAGTTTCAACGTTAGGTAACCCACGACGAGCTCTTA |
|                     | 5'-GGAAAGAGAGTTTCAACGTTAGGTAACCCACGAGCGGAGAGCTCTTA |

| N1                  | CGGAGGAGCTCTCGTGCATGGCGAAAGCCCTTTCTAACAACAAACCAACAAATCTATAT |
| Jasper              | CGGAGGAGCTCTCGTGCATGGCGAAAGCCCTTTCTAACAACAAACCAACAAATCTATAT |
|                     | CGGAGGAGCTCTCCGTCGATGGCGAAAGCCCTTTCTAACAACAAACCAACAAATCTATAT |

| N1                  | TCAATACAAGATGAAC- Large ORF |
| Jasper              | TACATGAATCATGAGC- Large ORF |

| N1                  | TAACAGCTACTCTTCTTGCTGACTGATCCCCTGCCGAAAAACCCCGGCCGGGGGGGGG |
| Jasper              | TAACAGCTACTCTTCTTGCTGACTGATCCCCTGCCGAAAAACCCCGGCCGGGGGGGGG |

* demonstrates that the patterns of the N1, Jasper and 002-73 strains are quite similar. A comparison between the corresponding VP5 hydrophobicity plots shows that the IPNV strains have very similar plots, whereas the hydrophobicity plot of VP5 of the 002-73 strain of IBDV shows hardly any resemblance to the IPNV plots (Fig. 6).

**SDS–PAGE of [35S]methionine-labelled viral proteins**

SDS–PAGE of radiolabelled purified IPNV particles (Ab, Sp, Jasper and N1 strains) revealed the polypeptide pattern shown in Fig. 7. In addition to bands corresponding to VP1, pVP2, VP2 and VP3, there are bands at the bottom of the gel with Mr of approximately 25K and 17K. The 17K bands of the Jasper and N1 lanes (Fig. 7) are poorly visible in the photograph.

**Discussion**

Results obtained by Azad et al. (1985) upon in vitro translation of IBDV RNA (strain 002-73) indicate that segment A may code for a 16K polypeptide in addition to PVX, PV3 and PV4 (Azad et al., 1985). A second small ORF in segment A of IPNV, capable of encoding a 17K polypeptide, has also been mentioned by Duncan et al. (1987). However, these observations have not been discussed further, as the existence of a 17K polypeptide has never been corroborated. We consider it unlikely that the positions of the start and stop codons of the small ORF should be so perfectly conserved between the strains unless it encodes a functional gene product. No other ORFs on segment A are conserved in that way. Alignment of the amino acid sequences of the putative VP5 polypeptides of the N1, Jasper and 002-73 strains revealed some remarkably well conserved cysteine residues and demonstrated that many of the non-identical amino acids are conservative amino acid replacements (Fig. 5). Likewise, the comparison of hydrophobicity and charge shown in Fig. 6 demonstrated a greater similarity in physicochemical properties than would be expected from just considering matching amino acids. Furthermore, the autoradiograph shown in Fig. 7 revealed a band which corresponded to the theoretically calculated Mr of VP5. Since the SDS–PAGE gel was loaded with [35S]methionine-labelled purified virus, the 17K band is probably of viral origin. Based on this new information we now think it is reasonable to believe that the small ORF on segment A encodes a fifth birnavirus protein (VP5). However, further investigation is required to prove the existence of VP5 and to understand its biological function. For that reason we are now preparing synthetic oligopeptides in order to make monospecific antisera against VP5. These antisera will be used to study the production of VP5 in vivo.
The mechanism of replication has not been well studied in the Birnaviridae group, but evidence suggests that birnavirus replication is initiated independently at the ends of the segments and proceeds by strand displacement (Bernard, 1980; Mertens et al., 1982; Spies et al., 1987). All birnaviruses probably use identical, or at least very similar, mechanisms for replicating and packaging their genomes. It is also to be expected that segment A is replicated and packaged in the same way as segment B. The finding of terminal inverted repeats in both segment A of IPNV and segment B of IBDV indicates that these sequences are somehow essential to birnaviruses and raises a number of questions pertaining to function. We suggest that the terminal inverted repeats play important roles in birnavirus dsRNA replication and/or packaging.

Stem-and-loop structures might be expected to form at the 5' untranslated regions of the mRNAs transcribed from segment A of IPNV and segment B of IBDV (Fig. 3a and 3b, adjacent inverted repeats). Several studies have demonstrated that a stem-and-loop structure 5' to the start codon of the mRNA significantly decreases
translational efficiency in eukaryotes (Pelletier & Sonenberg, 1985; Kozak, 1988). Thus, it seems logical to predict that the stem-and-loop structures of the 5' untranslated regions of segments A and B regulate viral protein synthesis by decreasing or blocking translation.
Fig. 5. Alignment of the amino acid sequences of the small ORFs (VP5) of the N1 and Jasper strains of IPNV and the 002-73 strain of IBDV. Symbols: X, conserved cysteine residues; :, identical amino acids; ., conserved amino acid replacements [scores greater than zero in the mutation probability data matrix (Dayhoff, 1978)]. The amino acid sequences of the small ORFs of the Jasper strain of IPNV and 002-73 strain of IBDV were deduced from their respective nucleic acid sequences (Duncan & Dobos, 1986; Hudson et al., 1986).

Fig. 6. Plots of hydrophobicity and charge of VP5 by the method of Kyte & Doolittle (1982). Comparison between the N1 and Jasper strains of IPNV and the 002-73 strain of IBDV. The amino acid sequences of the small ORFs (VP5) of the Jasper strain of IPNV and the 002-73 strain of IBDV used in these analyses were deduced from their respective nucleic acid sequences (Duncan & Dobos, 1986; Hudson et al., 1986).
Sequence of segment A of the N1 strain of IPNV

Acids differ between the two serotypes of IPNV. It turns out that the hypervariable internal segment identified by amino acid comparison analysis is identical to the segment recognized by MAb 17/82. On the basis of these two independent observations we suggest that the approximately 150 amino acid long hypervariable internal segment of VP2 carries serotype-specific epitopes of both IPNV and IBDV.

As far as we know it has not yet been determined whether VP2 of IPNV is glycosylated. Müller & Becht (1982) made attempts to demonstrate carbohydrates in IBDV-specific polypeptides and found that no appreciable amounts of common carbohydrates are present in IBDV. However, the conservation in all three birnavirus strains of several potential N-glycosylation sites in VP2 (Fig. 4), especially in the heterogeneous region mentioned above, suggests that it should not be completely ruled out that oligosaccharides may form part of VP2 of birnaviruses.

Azad et al. (1987) have shown that expression in *Escherichia coli* of the large ORF of segment A of IBDV results in autocatalytic cleavage of the polyprotein. Furthermore, Jagadish et al. (1988) demonstrated by mutagenesis studies that the only mutants which affected processing at the VPX/VP4 and VP4/VP3 junctions were those in which pieces of DNA were either inserted into or deleted from VP4. We have expressed a region of the large ORF of the N1 strain (positions 452 to 972; Fig. 4) in *E. coli* as a β-galactosidase fusion protein. Subsequent Western blot analysis using specific antisera revealed bands with *M*<sub>r</sub> values corresponding to those of NS and VP3, indicating autocatalytic cleavage of the fusion protein (not shown). It is puzzling that the NS polypeptide of IPNV and VP4 polypeptide of IBDV, which carry out equivalent cleavage reactions, have diverged to such an extent that there is very little homology left between their amino acid sequences (Table 1). However, the NS protein is comparatively well conserved between the two strains of IPNV. The adaptation of IPN and IBD viruses to different hosts appears to have increased the rate of evolution of the NS/VP4 proteins relative to the other gene products of segment A.

We wish to thank Rein Aasland and Dr Johan Glette for advice and helpful discussions. We also thank Dr P. E. V. Jørgensen and Dr P. Dobos for providing the Ab, Sp and Jasper strains of IPNV. This work was supported by the Royal Norwegian Council for Science and Industrial Research and the Research Council of Fisheries.

**References**


(Received 17 July 1989; Accepted 17 October 1989)