

Coding sequences of both genome segments of a European 'very virulent' infectious bursal disease virus

Michael D. Brown, Michael A. Skinner *

Institute for Animal Health, Compton Laboratory, Newbury, Berks., RG20 7NN, UK

Received 7 December 1994; revised 10 March 1995; accepted 10 April 1995

Abstract

The sequences of segment A (encoding the VP2-VP4-VP3 polyprotein and VP5) and segment B (encoding VP1) of a recent, 'very virulent' (VV) European isolate (UK661) of infectious bursal disease virus (IBDV), a birnavirus, have been determined. There are 26 to 36 amino acid substitutions compared to any other type I IBDV within the segment A polyprotein (of these, 15 are unique) and about 50 substitutions within VP1 (of which 16 are unique). There is more variation compared to classical and antigenic variant viruses, of both virulent and attenuated phenotype, in VP1, VP3 and VP4 than in VP2, even though the latter has previously been identified as the most variable protein between different strains of type I IBDV. In VP3 and VP4, UK661 is the most diverged type I IBDV. Thus the origin of the virus is unclear. It is possible that strong functional constraints have preferentially maintained the primary structure of VP2, though the possibility of recombination cannot be excluded. There are no clear candidate mutations to account for the enhanced virulence of the VV IBDV. Polymerase motifs are well conserved in VP1 but there is an amino acid substitution next to the predicted active-site serine of the viral protease (VP4). In addition, there is a conservative substitution close to the postulated VP2-VP4 cleavage site. It is also now apparent that sequences of IBDV segment B (the segment encoding the RNA polymerase) do not group according to serotype (specified by the capsid proteins encoded on segment A), indicating that segment reassortment has occurred.

Keywords: Birnavirus; IBDV; Gumboro disease; Serine protease motif

1. Introduction

Infectious bursal disease (IBD), initially described in the 1960's (Edgar and Cho, 1965; Snedeker et al., 1967), is a highly contagious disease of young chickens caused by a birnavirus, infectious bursal disease virus (IBDV). Although the disease caused some mortality in older birds (of 3 to 6 weeks of age), the major problem

* Corresponding author. Tel. +1635 577270; fax: +1635 577263; E-mail: SKINNER@BBSRC.AC.UK.

The nucleotide sequence data reported in this paper will appear in the EMBL database under the accession numbers: X 92760, X 92761.

associated with IBD was that of immunosuppression following infection of younger chicks. Until the late 1980's, the disease was well controlled by the use of a range of live attenuated and killed commercial vaccines but, in 1987, a pathotypic variant of IBDV arose in the Benelux countries and spread rapidly throughout Europe, including the UK (Chettle et al., 1989; Van den Berg et al., 1991). The variant was called 'very virulent' (VV) as it caused high flock mortalities (up to 70% for pullets, 25% for broilers and 100% mortality for specific pathogen-free birds). In contrast to antigenic variants isolated in the US in the mid 1980's (Snyder, 1990; Van der Marel et al., 1990; Öppling et al., 1991), the pathotypic variant retains very similar antigenicity to previously circulating strains of IBDV (referred to as 'classical' strains). VV IBDV can break through maternal antibody one week earlier than so-called 'intermediate' live vaccine strains. Based on pathology data, VV IBDV appears to have reached Poland by 1991 (Minta and Daniel, 1994) and South Africa by 1989 (Horner et al., 1994). A VV virus was also isolated and characterised in Japan in mid-1990 (Nunoya et al., 1992). Anecdotal reports indicate that the pathotypic variant may also be established in the Middle and Far East and in Central Africa. There are currently no indications that it has reached Australasia or the Americas.

IBDV has a bisegmented, double-stranded RNA genome contained within a non-enveloped icosahedral capsid. Genome segment A (3.3 kb) encodes a VP2–VP4–VP3 polyprotein which is auto-cleaved to the individual proteins (Müller and Becht, 1982; Azad et al., 1985; Hudson et al., 1986; Azad et al., 1987). VP2, the major structural protein, which carries an immunodominant, conformational, antigenic site (capable of inducing neutralizing antibodies), displays the greatest amount of amino acid sequence variation between classical strains (Bayliss et al., 1990; Kibenge et al., 1990). VP4 is believed to be a viral protease (Azad et al., 1987; Jagadish et al., 1988) and VP3 is a structural peptide which bears a minor neutralizing site (Jagadish and Azad, 1991). Segment A also encodes a small protein (M_r 17K) known as VP5, from sequences upstream of and overlap-

ping the start of the polyprotein coding sequence (Spies et al., 1989; Mundt et al., 1995). The other segment (B; 2.8 kb) encodes VP1, a single polypeptide, believed to be the viral RNA-dependent RNA polymerase (Morgan et al., 1988; Koonin, 1992).

We recently reported the partial VP2 sequences of several VV UK isolates of IBDV spanning the major neutralization site (Brown et al., 1994). We showed that the VV isolates formed a closely related group clearly distinguishable from other classical strains of IBDV. Moreover, we showed that the UK isolates were very similar to the earliest VV isolate from mainland Europe and, surprisingly, that their sequence was almost identical to that of a Japanese VV isolate (Lin et al., 1993). These results have allowed us to attribute new outbreaks of IBDV in the UK to the VV strains (Brown and Skinner, unpublished data) using restriction enzyme analysis of amplified cDNAs (RT-PCR) as suggested by Lin et al., 1993.

Determination of the complete sequence of VV IBDV is a prerequisite to identifying the determinants responsible for its enhanced virulence and should facilitate the production of specific sero-diagnostic tools for the discrimination of VV IBDV from other pathotypes. Here we report the complete coding sequences for both genome segments of a strain (UK661) representative of the UK VV IBDV isolates. Comparison with other strains of IBDV reinforces our previous conclusions that VV IBDV represents a distinct strain of the virus. Paradoxically, VV IBDV appears more closely related to other classical strains in VP2, hitherto the most variable viral protein, than in the other viral proteins.

2. Methods

VV IBDV UK661 was isolated by J. C. Stuart from 10-day-old broilers in the UK in May 1989. Propagation of the virus *in vivo* was described previously (Brown et al., 1994). Extraction and purification of virus from homogenised bursae, the extraction of viral RNA, reverse transcription and polymerase amplification were as described (Brown et al., 1994).

2.1. Amplification and cloning of Segment A cDNA

Segment A cDNA (covering the region 120 to 3185 as defined by Bayliss et al., 1990) was amplified in three fragments of 1204 bp (120 to 1323), 1302 bp (660 to 1961) and 1300 bp (1886 to 3185). Three independently-amplified copies of each fragment were purified, cleaved and cloned into pBlue-script II KS- (Stratagene), as described previously (Brown et al., 1994), for sequencing. The 5' end was amplified as a 186 bp fragment (1 to 186), purified by isopropanol precipitation and directly sequenced using the amplification primers.

2.2. Amplification and cloning of Segment B cDNA

Segment B cDNA (covering the region 49 to 2763, according to the sequence for 002-73 of Morgan et al., 1988) was amplified in two fragments of 1420 bp (49 to 1468) and 1449 bp (1335 to 2763). Three independently-amplified copies of each fragment were cloned into pBluescript KS- (Stratagene) then plasmid DNA was prepared and purified through caesium chloride for automated sequencing.

2.3. Cycle sequencing using an automated sequencing machine

Cycle sequencing reactions were carried out using the Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems) using the manufacturer's recommended conditions. Sequence determination was performed on an ABI 373A automated sequencer.

2.4. Sequence assembly and analysis

Sequence data, checked using SEQED 1.03 (Applied Biosystems, Inc.), was assembled on a Dec Vax mainframe (VMS) using the Staden Sequencing package (Staden, 1982). Database searches were performed on the Vax using the FASTA program and sequence analysis and comparisons were carried out using the GCG package (Sequence Analysis Software Package, 7.2, 1991, Ge-

netics Computer Group, Madison, Wisconsin, USA).

All the phenogramatic analysis was carried out on sequence alignments compiled by the GCG PILEUP program using the PHYLIP package (Phylogeny Inference Package v 3.5c; Joseph Felsenstein, Department of Genetics, University of Washington, Seattle, Washington, USA, 1993). Distance matrixes were first constructed, from the multiple sequence alignments, using either DNADIST (using the model of Jukes and Cantor, 1969) or PROTDIST (using the tables of Eck and Dayhoff, 1966). Phylogenetic trees were then constructed from these matrixes using the FITCH program, with the 'global' option and shuffling of input order (Fitch and Margoliash, 1967). Trees were then drawn and displayed by the DRAWTREE program.

Sequences used for comparison were from the following references and EMBL database accessions. Segment A, serotype I: 52/70 (Bayliss et al., 1990. Accession D00869), STC (Kibenge et al., 1990. Accession D00499), GIs (Vakharia et al., 1992. Accession M97346), PBG 98 (Bayliss et al., 1990. Accession D00868), CU-1 (Spies et al., 1989. Accession X16107), 002-73 (Hudson et al., 1986. Accession X03993); segment A, serotype II: OH (Kibenge et al., 1991. Accession M66722), 23/82 (Chettle et al., 1985; Bernstein, 1993. Accession Z21971); segment B, serotype I: Cu-1 (Spies, 1990. Supplied by H. Becht, Giessen); 002-73 (Morgan et al., 1988. Accession M19336) and segment B, serotype II: 23/82 (Chettle et al., 1985; Bernstein, 1993. Supplied by H. Becht, Giessen) and OH (Avirsegb; Dybing, 1992. Accession L19502).

3. Results

3.1. Sequence analysis of segment A

Sequences of segment A, encoding VP5 and the polyprotein, and of segment B, encoding VP1, have been determined for VV strain UK661. The sequences were derived from three independently-amplified clones for each cDNA fragment to remove the risk of errors or artefacts. No differences were observed between the clones. In addition, the

sequence reported here spans the partial VP2 sequence that we previously reported for UK661 (Brown et al., 1994) and the relevant sequence is identical to that reported previously, even though both were derived from cDNA amplified independently using different primers.

There are 136 to 307 nucleotide differences between segment A of UK661 and any other serotype I strain of IBDV, though UK661 has only 57 unique substitutions. The polyprotein encoded by UK661 shows 26 to 35 amino acid differences from those of any other serotype I strain, with 15 substitutions unique to UK661 (Fig. 1).

We showed previously that, between residues 191 to 384, the VP2 of four U.K. VV IBDV isolates showed only three unique amino acid substitutions (P-A 222, V-I 256 and L-I 294). In the rest of VP2 of UK661 (residues 1–454), there is only one more unique mutation (I-L 452), just before the postulated dibasic residue protease cleavage site. Thus there are a total of 4 amino acid substitutions unique to UK661 (as a representative VV isolate) within the 454 residues of VP2 (representing 0.88% variation).

Within the 289 residues of VP3 (polyprotein residues 724–1012), there are 6 amino acid substitutions (representing 2.1% variation) unique to UK661: H-D 751, E-G 849, E-G 919, I-V 951, A-V 990 and T-A 1005. The E-G substitution at 919 is at the C-terminus of a region mapped as the binding site for neutralizing monoclonal antibody 17/80 (Jagadish and Azad, 1991) but is after a substitution (at 918) in 52/70.

Within the 269 residues of VP4 (polyprotein residues 455–723), there are 5 amino acid substitutions (representing 1.86% variation) unique to UK661: D-V 624, D-G 632, N-S 651, K-N 685 and P-S 715. The substitution at 651 (N-S) is adjacent to what we believe to be the active site serine of the VP4 serine protease (i.e. S 652; see Discussion and Fig. 7), though this has not yet been demonstrated experimentally.

Within the 145 residues of VP5 there are 2 amino acid substitutions (representing 1.4% variation) unique to UK661: G-R 45, I-L 74 (Fig. 2). These two substitutions are encoded by a region that overlaps the polyprotein open reading frame. The mutations causing the substitutions in VP5 are silent in the polyprotein.

3.2. Sequence analysis of segment B

There are currently only 2 full IBDV VP1 coding sequences from segment B in the databases: those for the attenuated Australian serotype I virus, 002-73, and for an avirulent, serotype II virus, OH. Additional full VP1 coding sequences have been reported for Cu-1 (Spies, 1990), a mutant of Cu-1 (Cu-1M Bernstein, 1993) and another serotype II strain, 23/82 (Bernstein, 1993). In addition there is a partial sequence (from residue 326) for 52/70 (PIR accession: RRXSI5). There are 274 to 306 nucleotide differences between the VP1 coding sequence of UK661 and any other strain of IBDV of either serotype. Of these, 153 are unique to UK661.

Within the 879 residues of VP1, there are 23 to 59 amino acid differences between UK661 and any other strain, of either serotype. Of these, 16 are unique to UK661 (Fig. 3). They are: T-S 12, A-P 91, E-D 146, D-E 242, W-C 330, L-C 331, I-K 337, P-Q 341, P-S 343, L-M 390, E-D 393, S-P 562, S-P 687, K-R 694, K-M 751, N-H 825. None of the substitutions are located in the 8 RNA-dependent RNA polymerase (RdRp) motifs identified by Bruenn (1991). Mutations L-C 331 and I-K 337 are, however, located in an additional RdRp motif, I (Koonin, 1991; Koonin, 1992). The change to a lysine at 337 brings the sequence of UK661 in this motif in line with the consensus for positive strand RNA viruses (Koonin, 1991), whereas other IBDVs and also infectious pancreatic necrosis viruses (IPNV; a birnavirus of salmonids) have an isoleucine at this position.

Fig. 1. Alignment of the predicted amino acid sequence of the segment A polyprotein (VP2-VP4-VP3) of the UK VV isolate UK661 to the corresponding sequences of published IBDV serotype I strains 52/70, GIs, 002–73 and serotype II strain 23/82. The alignment is shown as a PRETTYPLOT in which residues identical to a consensus are indicated by a dash and non-identical residues are shown in lower case. The start sites of VP4 and VP2 (assuming cleavage at the dibasic residues) are indicated by asterisks at residues 455 and 725). The A-x-A-A-S repeats, postulated as alternative cleavage sites (Hudson et al., 1986), are underlined. Note the presence of an extra residue in 23-82 (S 249).



Fig. 2. Multiple sequence alignment of the predicted amino acid sequences of VP5 encoded by segment A of IBDV strains: UK661 (17k661), 52/70 (17k5270), G1s (17kg1s), 002-73 (17k002-73), OH (17koh) and 23/82 (17k23-82). The alignment is displayed as a PRETTYBOX plot in which identical residues are shaded black, similar residues are lightly shaded and non-identical residues are unshaded.

Fig. 5. Amino acid sequence comparison of residues 61 to 177 of segment B of IBDV (strains 002-73, UK661, Cu-1 and two serotype II strains, 23/82 and OH, "Avirsegb") and IPNV strains Jasper (Ipnvja), drt (Ipnvdrt) and Sp (Ipnvsp). The alignment is displayed as a PRETTYBOX plot in which identical residues are shaded black, similar residues are lightly shaded and non-identical residues are unshaded.

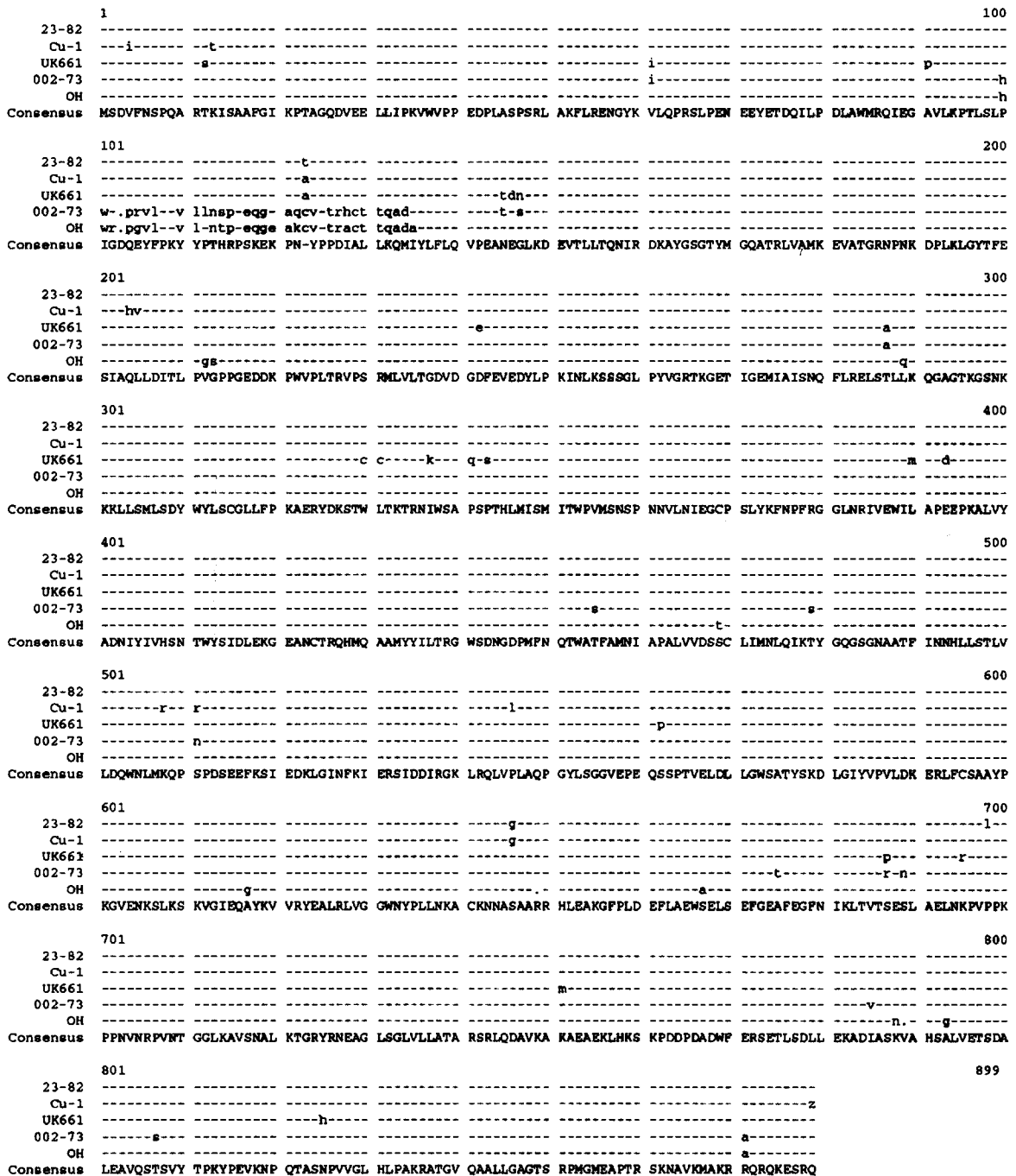


Fig. 3. Alignment of the predicted amino acid sequence of VP1 of UK661 to all the published IBDV sequences: 002-73, Cu-1 and two serotype II strains OH and 23/82. The alignment is shown as a PRETTYPLOT in which residues identical to a consensus are indicated by a dash and non-identical residues are shown in lower case. Note that deleted residues in 002-73 and OH means that their coordinates will differ from the consensus.

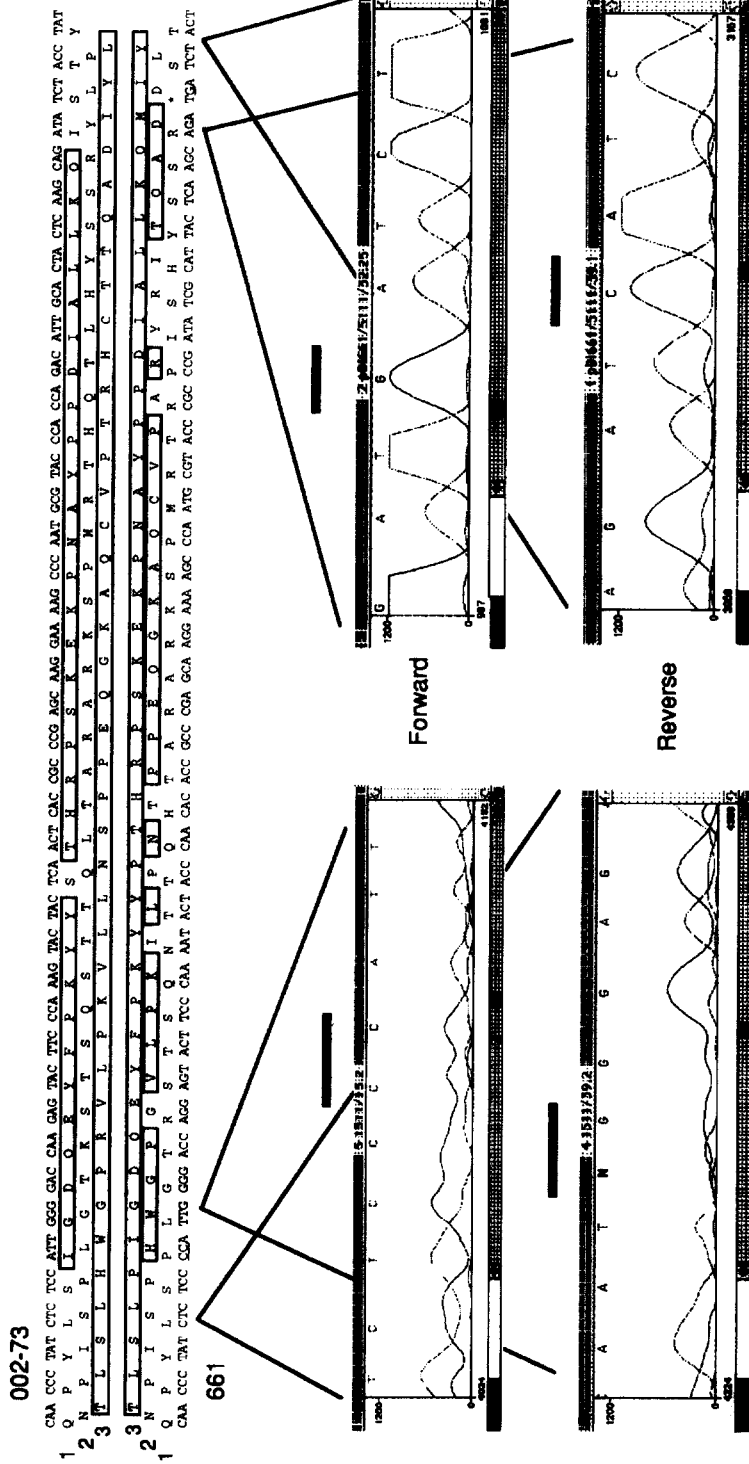


Fig. 4. Diagram showing the nucleic acid sequences and the corresponding three frame translations of UK661 and 002-73 encompassing the region of VP1 between amino acid residues 96 and 135 (in UK661). The sequence in the figure starts at base 327 of the UK661 segment B sequence. The predicted open reading frame (ORF) for VP1 of UK661 is boxed and labelled 3. The published, predicted ORF for 002-73 is also boxed and labelled 3. Also highlighted by boxes are the homologous sequences from each strain which are found in different reading frames (that in reading frame 2 of UK661 shows homology to the published 002-73 ORF and that in reading frame 1 of 002-73 is homologous to the ORF predicted for UK661). Bases underlined in the UK661 nucleic acid sequence are insertions relative to 002-73. The chromatograms shown are the forward (top) and reverse (bottom) sequence data around the insertion points in UK661. The bars above each chromatogram highlight the bases inserted into the sequence of UK661 to create the frame-shift in comparison to the 002-73 strain.

3.2.1. A frame-shift between the VP1 of different IBDVs?

The amino acid sequence of the VP1 of UK661 shows high divergence in the region between residues 100 and 135 when compared with those of 002-73 and OH (with no more than 6 out of 34 residues identical; see Fig. 3). An alternative read-002-73 has 32 (out of 33) residues identical to that in frame of UK661 has 26 (out of 34) residues identical to the predicted product from 002-73 (Fig. 4). An alternative open reading frame of predicted for UK661. The apparent frame-shifts in UK661, compared to 002-73 and OH, are due to insertions in UK661 of 2 bases (CC 389-390) and one base (G 445), respectively (Fig. 4).

A similar discrepancy with 002-73 and OH was also described by Bernstein in his thesis (Bernstein, 1993). The amino acid sequences observed for Cu-1 (Spies, 1990) and Cu-1M (Bernstein, 1993), in the region between residues 100 and 135, are identical to that for UK661. Serotype II strain 23/82 (Bernstein, 1993) has only one substitution (A-T 123) within the same region. Moreover, sequences of IPNV show much higher identity over this region to those of UK661, Cu-1, Cu-1M and 23/82 than to 002-73 and OH (Fig. 5).

We re-examined our sequence data for UK661 and are confident that the extra bases, relative to 23/82 and OH, are actually present (Fig. 4). Strains 002-73 and OH were not available to us so we cannot comment on whether the differences are real or due to sequence artefacts. The published sequence of 002-73 results in 8 amino acid substitutions relative to OH in this region but the alternative reading frame (equivalent to that reported here for UK661) would result in only 4 substitutions compared to the product of the alternative reading frame of OH. Thus there would be more silent mutations in the alternative reading frames of 002-73 and OH than in the published open reading frames.

We note, however, that in the sequences of 002-73 and OH, there is conservation of the motif PK-(6)-P-(8)-P found in UK661 and the other sequences. This raises the possibility that the alternative sequences have both been ex-

pressed ancestrally and have been subject to the same selection and constraints. It is possible, therefore, that a switch to the alternative sequence might occur in a particular strain of IBDV, though the consequences of such a switch are as yet unknown.

3.3. Phylogenetic relationship of UK661 to other IBDVs

We previously showed (Brown et al., 1994), by comparison of sequences encoding the major antigenic site on VP2, that the VV isolates formed a closely related group clearly distinguishable from other classical strains of IBDV. Comparison of the amino acid sequences of the VP2 antigenic site showed that the VV isolates were more closely related to the virulent strains (52/70 and STC) than to other pathotypes.

Comparison of the complete peptide sequence of the polyprotein of UK661, reported here, with those of other strains, also shows that it is most closely related to 52/70 (not shown). It is, however, diverged from 52/70, and the other strains, by approximately the same as the hitherto most highly diverged serotype I strain, 002-73 (a highly attenuated Australian strain).

Analysis of the distribution of the UK661-specific amino acid substitutions shows that they are twice as frequent in VP3 (2.1%) and VP4 (1.9%) as in VP2 (0.9%). It was, therefore, necessary to perform cluster analysis for each protein. This revealed that, in VP2, UK661 is most closely related to 52/70 and vice versa. In VP3 and VP4, however, the nearest neighbours to UK661 are PBG98 and Cu-1 but, of the serotype I viruses, UK661 is the most diverged (Fig. 6b-d). In VP5, the frequency of UK661-specific amino acid substitutions is 1.4%, considerably higher than in the overlapping VP2, but the nearest neighbour to UK661 is 52/70 (Fig. 6e), as is the case for VP2.

The frequency of UK661-specific amino acid substitutions in VP1 is 1.8%, similar to that observed for VP4. Cluster analysis of VP1 protein sequence is problematic, given the potential frame-

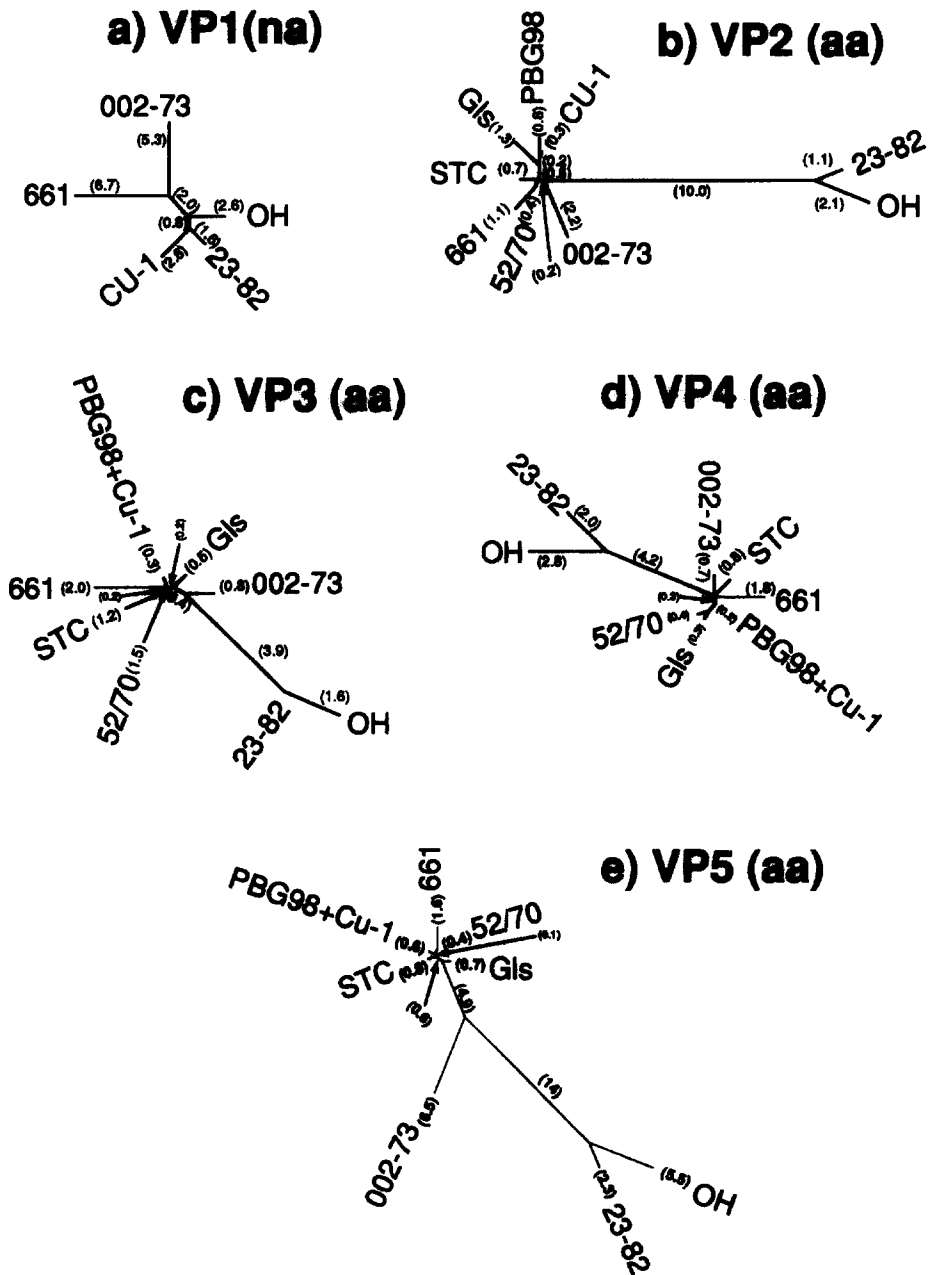


Fig. 6. Phenograms, produced by the programs in the PHYLIP package 3.5c (Felsenstein, Department of Genetics, University of Washington, Seattle, Washington, USA.), displaying the relationship of the nucleic acid sequences of the VP1 open reading frame (a) and of the amino acid sequences of VP2 (b), VP3 (c), VP4 (d) and VP5 (e) of the UK VV isolate UK661 to those of serotype I strains 002-73, Cu-1, PBG 98, STC, Gls and 52/70 as well as serotype II strains OH (avir.) and 23/82. The branch lengths are proportional to the expected percentage (shown in brackets) of nucleic acid or amino acid substitutions based on the observed data.

	I	II	III	IV
	a* e (11)	*	(35) g e* * g (0)	* * i
TRP:	34 QWV SA A HCYKS (33)	SLTINNDIMLIKLS (76)	GKDS C MGDSGGPVVC.SG (0)	KLQGI V SWG SG
661	539 **I* ***** (32)	***** (46)	*****S***** (35)	***** **
IBDV:	539 PGVL RGA H NLDC (32)	IEGVREDLQPPSQRG (46)	PIPPIV G NSGNLAIAYMD (35)	HRL Q LKLAG PG
TBE:	47 KGV L HTM W EVTRG (13)	WADVREDVVCYGGAW (43)	PIDL V K G T S GSPIINAQG (0)	VV V GLYNG LK
IPNV:	539 KGVIFPVV H TVES (27)	LSYFKNDTGC V WGIG (31)	FAGPAM G PSAQLALSLLV (36)	HEH L PLI C QCPG
	bc1d3 2 (37)	4ddd b b f (78)	dd f h ffb (4)	f c f

a: TRP + IBD	f: IBD + IPN	1: V in IBDV CU-1, PBG98, 002-73, 52/70 and Stc; I in 661, Gls, OH and 23-82.
b: TBE + IPN	g: TRP + TBE	2: K in IPNV SN1
c: TBE + IPN + IBD	h: TRP + TBE + IBD	3: V in IPNV SN1
d: TBE + IBD	i: TRP + IBD + IPN	4: H in IPNV SN1
e: TRP + IPN	*: All	

Fig. 7. Alignment of VP4 sequences of IBDV and IPNV with serine protease motifs from trypsin (TRP) and the NS3 protease of the flavivirus, tick-borne encephalitis virus (TBE). The alignment consists of 4 domains, separated for each sequence by the number of residues in brackets (the numbers in brackets at the top and bottom represent the minimum and maximum separation, respectively, seen in viral and cellular proteases, Chambers et al., 1990). Numbers at the start of the sequences represent the position of the first residue (relative to the polyprotein for the birnavirus sequences). Domains I, II and III represent sequences flanking each of the residues comprising the catalytic triad: H 546, D 589 and S 652 in IBDV. Domain IV contains the substrate-binding-pocket. *Indicates residues conserved across all the sequences shown. Letters above or below the aligned sequences show the patterns of homology, as indicated in the key. For UK661, only differences relative to 52/70 are indicated.

shift, described above. Analysis of the nucleic acid sequence, however, shows that UK661 is the most diverged of all the viruses, serotype I or II (Fig. 6a). The strain showing the least number of differences from UK661 is serotype II strain 23/82 and that showing the most is 002-73. However, UK661 clusters with 002-73, indicating shared substitutions, but is more distantly related to Cu-1, the other serotype I strain, than to serotype II strain OH. Whether VP1 sequences are compared as coding nucleic acid sequences or as derived amino acid sequences, they do not group according to serotype.

4. Discussion

Our main interest in determining the full sequence of VV IBDV isolate UK661 was as a basis for the identification of the mutation, or mutations, which is, or are, responsible for the enhanced virulence of this pathotype. We are aware that in a virus of this size, major changes in virulence can be the result of just one or two

base changes in coding or non-coding sequences (Skinner et al., 1989; Westrop et al., 1989). It is likely that a reverse genetics approach will be necessary to permit the direct identification of virulence determinants, as performed for poliovirus (Westrop et al., 1989). At the moment there is no true 'reverse genetics' system for IBDV (nor indeed for any other segmented, double-strand RNA virus) to allow the in vitro construction and recovery of mutants or recombinants. For this reason we have not extended the sequence of UK661 to the ends of the non-coding regions, as there is no way yet of examining the function of these regions. Clearly, knowledge of these sequences will be required for a 'reverse genetics' approach. The terminal sequences of P2, Cu-1, Cu-1M and 23/82 have been reported recently (Mundt and Müller, 1995).

Of the UK661-specific amino acid substitutions that we have observed, some are in locations which may affect the phenotype of the virus, possibly affecting virulence. These are mutations in the VP4 viral protease, in VP1 polymerase, near the VP2-VP4 cleavage site and in the antigenic sites of VP2 and VP3.

No significant relationships have been demonstrated between the birnavirus VP4 sequences and any known type of protease (Koonin, 1992; Dougherty and Semler, 1993). Koonin (1992) reported finding 'marginal similarity to serine protease active centres' but no data was presented. The active site of the protease in IBDV has not been mapped but we propose that IBDV polyprotein residues H 546, D 589 and S 652 are candidates to form the catalytic triad of a serine protease (Fig. 7). The spacing between the residues is consistent with that observed in viral and cellular serine proteases (Chambers et al., 1990). The sequence around residue D 589 (VREDL, or AREDL in strain OH) is particularly well conserved with the equivalent region in the NS3 protease of flaviviruses (Chambers et al., 1990). The sequence around the postulated catalytic site at 652 (GSSGNL in UK661, GNSGNL in other IBDVs) shows conservation of residues (namely G × SG) that are invariant in the catalytic sites of viral and cellular serine proteases (Chambers et al., 1990). Alignment of the IBDV and IPNV sequences is difficult, as the homology is very low. The postulated active site serine of IBDV is, however, found in a similar context in IPNV (IPNV polyprotein residue S 633), as is the first G, but the second G is not conserved. The other components of the catalytic triad can be found in IPNV at positions similar to those observed in IBDV. A histidine (IPNV polyprotein residue H 547) is found in IPNV at position 93 of VP4 (assuming the same dibasic residue cleavage site), the position occupied by residue H 546 of the IBDV polyprotein. Candidate aspartate residues exist in IPNV at polyprotein residues 573 and 585 (the latter being in a better context, as compared with other serine proteases, but closer to the active site serine than usual; Fig. 7). A motif in VP4 of IBDV (G₇₀₀XXXXG) and IPNV, downstream of the putative active site, shows conservation with the substrate-binding site of serine proteases (Fig. 7), particularly with those of flaviviruses (Chambers et al., 1990). This motif is, however, much further removed from the putative active site in VP4 than in other serine proteases (separated from the active site by 48 or 49 residues in VP4 compared to 12 to 17 residues in

other proteases). It will be interesting to see if this motif does contribute to substrate specificity in VP4.

The protease cleavage sites, like the protease active site, have not been established for IBDV (nor for IPNV). Dibasic residues (at residues 453 and 723 in IBDV), which are conserved in IBDV and IPNV, were proposed as likely cleavage sites (Hudson et al., 1986). Repeats of the sequence A-x-A-A-S were also suggested as possible cleavage sites in IBDV (Hudson et al., 1986). Three such repeats are found between polyprotein residues 485 and 504 and there is one at residue 754. The triple repeat in IBDV can be aligned (Heppell et al., 1993) with an IPNV sequence which spans the C-terminus of the VP2 of IPNV (Manning and Leong, 1990). There is no A-x-A-A-S motif in IPNV at this position but two copies of an A-A-G-G-R-Y repeat are well conserved (Heppell et al., 1993).

In UK661, two mutations could potentially affect the specificity or efficiency of cleavage of the viral polyprotein by VP4 protease. These are the substitution in the polyprotein at 651 (N-S), adjacent to the serine residue we predict to be the active site of the VP4 serine protease, and the I-L substitution at 452, just before the postulated dibasic residue protease cleavage site between VP2 and VP3. There is also a substitution (H-D 752) just upstream of the A-x-A-A-S postulated as the alternative VP4-VP3 cleavage site. By affecting polyprotein processing, such mutations might alter the virus replication rate. Cellular proteins could also be targets for any modified protease, so such mutations could also affect virulence by diverse effects on cell function. Mutations in the NS2B and NS3 proteins of Japanese encephalitis flavivirus, both of which are important to the activity of the virus-encoded serine protease, have recently been implicated in the attenuation of vaccine strains (Ni et al., 1995).

Of the mutations in VP1, those (L-C 331 and I-K 337) in RdRp motif I (Koonin, 1991) are perhaps most likely to affect polymerase function and thus alter the replication rate, possibly influencing virulence.

Although no clear antigenic changes have been reported for VV IBDV (Snyder, 1990; Van der

Marel et al., 1990; Öppling et al., 1991), the observation that it can cause disease in the face of higher levels of maternal antibody raises the possibility that subtle antigenic changes might play a role in the increased virulence. As we reported previously (Brown et al., 1994), the VP2 mutations V-I 256 and L-I 294 are conservative changes in a hydrophobic part of the antigenic site and, as such, are unlikely to make a direct contribution to antibody specificity. More subtle conformational changes cannot be excluded, however.

The remaining VP2 antigenic site mutation (P-A 222) is in one of the two hydrophilic domains, at a position where US antigenic variants have P-T, P-Q or P-S substitutions (Lana et al., 1992; Vakharia et al., 1994). VV IBDV have no changes in the second hydrophilic domain, unlike the US antigenic variants. Heine et al. (1991) analysed the binding of neutralizing monoclonal antibodies to wild-type, variant and recombinant VP2 expressed in *E. coli*. This showed that the second hydrophilic domain was crucial to the binding of the monoclonal antibodies (by dot blot immunoassay) but that the first hydrophilic domain played a secondary, possibly conformational role (as assayed by Western blotting). The P-A 222 substitution found in VV IBDV was frequently observed in antibody escape mutants of PBG98 or D78 (Van den Berg et al., 1994). Unlike in VV IBDV, however, the escape mutants, like the US antigenic variants, always had additional mutations in the second hydrophilic domain. It may be that mutations in the first hydrophilic domain are necessary to stabilise the protein structure after mutation in the second hydrophilic domain. Whether mutations in the first hydrophilic domain are, by themselves, destabilising is not clear. If it were so, it is possible that the mutations in VV IBDV in the region between the hydrophilic domains would compensate. Clearly there is a need for more detailed epitope mapping of this region in all pathotypes.

Monoclonal antibodies against VP3 are generally group-specific and poorly neutralizing (Becht et al., 1988). The immunodominant site of VP3, represented by 002-73 MAb 17/80, was mapped

to between residues 860 to 923 (Jagadish and Azad, 1991). Mahardika and Becht (1995) have recently shown that group-specific epitopes are spread throughout VP3 but that serotype-specific epitopes are located towards the C-terminus (from residue 890). Whether any of the UK661 substitutions in VP3 affect group- or serotype-specific responses and play a role in reduced *in vivo* protection is not known. The substitution at 919 (E-G) is within the region mapped to contain the immunodominant epitope on VP3 but it lies between sites of substitutions in 52/70 (E-D 918) and STC (R-K 923). Any of the other three C-terminal substitutions (I-V 951, A-V 990 and T-A 1005) might affect serotype-specific interactions but this will be clarified by finer epitope mapping.

The origins of VV IBDV are still unclear. In VP1, VP3 and VP4, UK661 is the most diverged of the serotype I viruses. In VP1 it is the most diverged of viruses of both serotypes. We assume that the source of VV IBDV is likely to be some as yet unrecognised reservoir. It is not clear whether that reservoir would be a geographically isolated population of poultry or a different avian species. The higher relative homology of UK661 with other serotype I viruses in VP2 and VP5 means that we cannot exclude the possibility of a recombination event having taken place. Sequences encoding the VP2 of VV IBDV are, however, still clearly distinct from those of other serotype I viruses. It is possible that structural or functional constraints account for the higher sequence conservation in VP2 relative to the other viral proteins.

The observation that the IBDV VP1 coding sequences of serotype I strains UK661 and 002-73 are more closely related to those of two serotype II strains (23/82 and OH) than to the other serotype I strain (Cu-1), seems to indicate that VP1 sequences do not correlate with serotype. The most obvious explanation for this would be that segment B is able to reassort with segment A. Should the determination of more segment B sequences confirm the observation, it will become necessary to introduce a classification for the segment B sequences that is independent of virus serotype.

Acknowledgements

We wish to thank Professor Becht (Giessen, Germany) for making available to us the thesis of Dr. F. Bernstein and for providing us with the sequences for 23/82 and Cu-1M in a computer-readable form. We thank J. McCauley, D. Cavanagh and P. Britton for critical reading of the manuscript and for constructive criticism. We are also grateful for the support of the British Egg Marketing Board Research and Education Trust in funding M.D.B. with a pre-doctoral fellowship.

References

- Azad, A.A., Barrett, S.A. and Fahey, K.J. (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.
- Azad, A.A., Jagadish, M.N., Brown, M.A. and Hudson, P.J. (1987) Deletion mapping and expression in *Escherichia coli* of the large genomic segment of a birnavirus. *Virology* 161, 145–152.
- Bayliss, C.D., Spies, U., Shaw, K., Peters, R.W., Papageorgiou, A., Muller, H. and Boursnell, M.E. (1990) A comparison of the sequences of segment A of four infectious bursal disease virus strains and identification of a variable region in VP2. *J. Gen. Virol.* 71, 1303–1312.
- Becht, H., Muller, H. and Muller, H.K. (1988) Comparative studies on structural and antigenic properties of two serotypes of infectious bursal disease virus. *J. Gen. Virol.* 69, 631–640.
- Bernstein, F. (1993) Analyse und Vergleich der Genome pathogener und apathogener Stämme des Virus der infektiösen Bursitis. Ph.D. Thesis, Justus-Liebig-Universität Giessen.
- Brown, M.D., Green, P. and Skinner, M.A. (1994) VP2 sequences of recent European 'very virulent' isolates of infectious bursal disease virus are closely related to each other but are distinct from those of 'classical' strains. *J. Gen. Virol.* 75, 675–680.
- Bruenn, J.A. (1991) Relationship among the positive strand and double-strand RNA viruses as viewed through their RNA dependent RNA polymerases. *Nucleic Acids Res.* 19, 217–226.
- Chambers, T.J., Hahn, C.S., Galler, R. and Rice, C.M. (1990) Flavivirus genome organization, expression and replication. *Annu. Rev. Microbiol.* 44, 649–688.
- Chettle, N.J., Eddy, R.K. and Wyeth, P.J. (1985) The isolation of infectious bursal disease virus from turkeys in England. *Br. Vet. J.* 141, 141–145.
- Chettle, N., Stuart, J.C. and Wyeth, P.J. (1989) Outbreak of virulent infectious bursal disease in East Anglia. *Vet. Rec.* 125, 271–272.
- Dougherty, W.G. and Semler, B.L. (1993) Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.* 57, 781–822.
- Dybing, J.K. (1992) Sequence analysis of IBDV of serotype 2 and expression of the VP5 cDNA. M.Sc., University of Prince Edward Island, Charlottetown.
- Eck, R.V. and Dayhoff, M.O. (1966) Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Spring, Maryland.
- Edgar, S.A. and Cho, Y. (1965) Avian nephrosis (Gumboro disease) and its control by immunisation. *Poultry Science* 44, 1366.
- Fitch, W.M. and Margoliash, E. (1967) Construction of phylogenetic trees. *Science* 155, 279–284.
- Heine, H.G., Haritou, M., Failla, P., Fahey, K. and Azad, A. (1991) Sequence analysis and expression of the host-protective immunogen VP2 of a variant strain of infectious bursal disease virus which can circumvent vaccination with standard type I strains. *J. Gen. Virol.* 72, 1835–1843.
- Heppl, J., Berthiaume, L., Corbin, F., Tarrab, E., Lecomte, J. and Arella, M. (1993) Comparison of amino acid sequences deduced from a cDNA fragment obtained from infectious pancreatic necrosis virus (IPNV) strains of different serotypes. *Virology* 195, 840–844.
- Horner, R.F., Parker, M.E. and Pike, R.N. (1994) Vaccination of maternally immune commercial broilers provides limited protection against virulent IBD. International Symposium on Infectious Bursal Disease and Chicken Infectious Anaemia, pp. 312–335.
- 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.
- 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.
- Jagadish, M.N. and Azad, A.A. (1991) Localization of a VP3 epitope of infectious bursal disease virus. *Virology* 184, 805–807.
- Jukes, T.H. and Cantor, C.R. (1969) Evolution of protein molecules. In: H.N. Munro (Ed), *Mammalian Protein Metabolism*, pp. 21–132. Academic Press, New York.
- Kibenge, F.S., 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.
- Kibenge, F.S., McKenna, P.K. and Dybing, J.K. (1991) Genome cloning and analysis of the large RNA segment (segment A) of a naturally avirulent serotype 2 infectious bursal disease virus. *Virology* 184, 437–440.
- Koonin, E.V. (1991) The Phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* 72, 2197–2206.
- Koonin, E.V. (1992) Evolution of double-stranded RNA viruses: a case for polyphyletic origin from different groups of positive-stranded RNA viruses. *Sem. Virol.* 3, 327–339.

- Lana, D.P., Beisel, C.E. and Silva, R.F. (1992) Genetic mechanisms of antigenic variation in infectious bursal disease virus: analysis of a naturally occurring variant virus. *Virus Genes* 3, 247–259.
- Lin, Z., Kato, A., Otaki, Y., Nakamura, T., Sasmaz, E. and Ueda, S. (1993) Sequence comparisons of a highly virulent infectious bursal disease virus prevalent in Japan. *Avian Dis.* 37, 315–323.
- Mahardika, G.N.K. and Becht, H. (1995) Mapping of cross-reacting and serotype-specific epitopes on the VP3 structural protein of the infectious bursal disease virus (IBDV). *Arch. Virol.* 140, 765–774.
- Manning, D.S. and Leong, J.C. (1990) Expression in *Escherichia coli* of the large genomic segment of infectious pancreatic necrosis virus. *Virology* 179, 16–25.
- Minta, Z. and Daniel, A. (1994) Infectious bursal disease in Poland: current situation and vaccinal control. International Symposium on Infectious Bursal Disease and Chicken Infectious Anaemia, pp. 208–214.
- 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.
- Müller, H. and Becht, H. (1982) Biosynthesis of virus-specific proteins in cells infected with infectious bursal disease virus and their significance as structural elements for infectious virus and incomplete particles. *J. Virol.* 44, 384–392.
- Mundt, E., Beyer, J. and Muller, H. (1995) Identification of a novel viral protein in infectious bursal disease virus-infected cells. *J. Gen. Virol.* 76, 437–443.
- Mundt, E. and Müller, H. (1995) Complete nucleotide sequences of 5'- and 3'-noncoding regions of both genome segments of different strains of infectious bursal disease virus. *Virology* 209, 10–18.
- Ni, H., Chang, G.J., Xie, H., Trent, D. and Barrett, A.J. (1995) Molecular basis of attenuation of neurovirulence of wild-type Japanese encephalitis virus strain SA14. *J. Gen. Virol.* 76, 409–413.
- Nunoya, T., Otaki, Y., Tajima, M., Hiraga, M. and Saito, T. (1992) Occurrence of acute infectious bursal disease with high mortality in Japan and pathogenicity of field isolates in specific-pathogen-free chickens. *Avian Dis.* 36, 597–609.
- Öppling, V., Muller, H. and Becht, H. (1991) Heterogeneity of the antigenic site responsible for the induction of neutralizing antibodies in infectious bursal disease virus. *Arch. Virol.* 119, 211–223.
- Skinner, M.A., Racaniello, V.R., Dunn, G., Cooper, J., Minor, P.D. and Almond, J.W. (1989) New model for the secondary structure of the 5'-non-coding RNA of poliovirus is supported by biochemical and genetic data that also show that RNA secondary structure is important in neurovirulence. *J. Mol. Biol.* 207, 379–392.
- Snedeker, C., Wills, E. and Moulthrop, I.M. (1967) Some studies on the infectious bursal agent. *Avian Dis.* 11, 519–528.
- Snyder, D.B. (1990) Changes in the field status of infectious bursal disease virus. *Avian Pathol.* 19, 419–423.
- Spies, U. (1990) Bestimmung der Nukleotidsequenz des Genoms von IBDV und Zuordnung von Antigenstruktur und funktionellen Bereichen. Ph.D. thesis, Justus-Liebig-Universität, Giessen.
- 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, 7982–7982.
- Staden, R. (1982) Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.* 10, 4731–4751.
- Vakharia, V.N., Ahamed, B. and He, J. (1992) Use of polymerase chain reaction for efficient cloning of dsRNA segments of infectious bursal disease virus. *Avian Dis.* 36, 736–742.
- Vakharia, V.N., He, J., Ahamed, B. and Snyder, D.B. (1994) Molecular basis of antigenic variation in infectious bursal disease virus. *Virus Res.* 31, 265–273.
- Van den Berg, T.P., Gonze, M. and Meulemans, G. (1991) Acute infectious bursal disease in poultry: isolation and characterisation of a highly virulent strain. *Avian Pathol.* 20, 133–143.
- Van den Berg, T.P., Morales, D., Gonze, M. and Meulemans, G. (1994) Relevance of antigenic variation for protection in infectious bursal disease. International Symposium on Infectious Bursal Disease and Chicken Infectious Anaemia, pp. 22–36.
- Van der Marel, P., Snyder, D. and Luticken, D. (1990) Antigenic characterisation of IBDV field isolates by their reactivity with a panel of monoclonal antibodies. *Dtsch. tierärztl. Wschr.* 97, 81–83.
- Westrop, G.D., Wareham, K.A., Evans, D.M.A., Dunn, G., Minor, P.D., Magrath, D.I., Taffs, F., Marsden, S., Skinner, M.A., Schild, G.C. and Almond, J.W. (1989) Genetic basis of attenuation of the Sabin type 3 oral Poliovirus vaccine. *J. Virol.* 63, 1338–1344.