CHARACTERISATION OF INFECTIOUS BURSAL DISEASE VIRUS (IBDV) POLYPROTEIN PROCESSING

by

PHILLIA RIXONGILE VUKEA
B.Sc. (Hons) Genetics (UKZN)

Submitted in fulfillment of the academic requirements for the degree of Doctor of Philosophy in the Discipline of Biochemistry, School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal

Pietermaritzburg

November 2011

DECLARATION - PLAGIARISM

I, Phillia Rixongile Vukea, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a. Their words have been re-written but the general information attributed to them has been referenced
b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed……

PREFACE

The experimental work described in this thesis was carried out at the School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal, Pietermaritzburg, South Africa from February 2007 to November 2011, under the supervision of Prof. Theresa H.T. Coetzer and co-supervision of Dr Alain Boulangé.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any University. Where use has been made of the work of others it is duly acknowledged in the text.

_____

Candidate: Phillia Rixongile Vukea

_____

Supervisor: Prof. Theresa H.T. Coetzer

ABSTRACT

Infectious bursal disease virus (IBDV) is a birnavirus that infects the B-cells in the bursa of Fabricius of young chickens, causing Gumboro disease. The IBDV 114 kDa polyprotein (NH_2-pVP2-VP4-VP3-COOH) is thought to be processed at ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} through the proteolytic activity of VP4, a serine protease which uses a Ser/Lys catalytic dyad, to release pVP2, VP4 and VP3. Precursor VP2 (pVP2) is further processed at its C-terminus to generate VP2 and structural peptides through the cleavage of the ^{441}Ala-Phe^{442}, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502} peptide bonds to release VP2 and four structural peptides, pep46, pep7a, pep7b and pep11. While the processing at the ^{441}Ala-Phe^{442} site was shown to be mediated by the endopeptidase activity of VP2, the processing at the other two sites is not well understood.

The products resulting from the processing of the IBDV polyprotein were previously identified by anti-VP2 and anti-VP3 antibodies. The present study used anti-VP4 peptide antibodies to identify products resulting from the IBDV polyprotein processing. It was hypothesised that VP4 exists in two forms, the embedded form which exists as an integral part of the polyprotein and a mature form which is released after the processing. In order to characterise the two forms of VP4, six different fragments i.e. full-length polyprotein (Met^{1}-Glu^{1012}), truncated polyprotein (Ile^{227}-Trp^{891}), VP4-RA (Arg^{453}-Ala^{755}), VP4-RK (Arg^{453}-Lys^{722}), VP4-ΔVP3 (Ala^{513}-Trp^{891}, called VP4-AW for the sake of simplicity) and VP4-AA (Ala^{513}-Ala^{755}) were amplified from the IBDV dsRNA, cloned into a T-vector and sub-cloned into several expression vectors. The constructs were sequenced prior to expression.

The sequence of the polyprotein coding region was used to determine the pathotype of the isolate used for viral dsRNA isolation. This isolate was from IBDV-infected bursae harvested from commercial chickens during an IBD outbreak in KwaZulu-Natal, South Africa in 1995, thus naming the isolate SA-KZN95. The comparison of the deduced amino acid sequence of SA-KZN95 polyprotein with 52 sequences of other IBDV strains highlighted 21 residues which could be molecular markers of different IBDV pathotypes. The residues of SA-KZN95 were identical to those of the Malaysian very virulent UPM94/273 strain.

The constructs representing the embedded and mature forms of VP4 were recombinantly expressed. Processing was observed from the expression of the full-length polyprotein, truncated polyprotein, VP4-RA, VP4-RK and VP4-AW, but not from VP4-AA expression.

The mutation of the Ser/Lys catalytic dyad in the full-length polyprotein, truncated polyprotein, VP4-RA, VP4-RK and VP4-AW, prevented processing thus verifying that the proteolytic activity was due to VP4. Anti-VP4 peptide antibodies were raised in chickens for the identification of the polyprotein cleavage products. The anti-VP4 peptide antibodies detected more cleavage products than expected from the polyprotein, suggesting that additional or different cleavage sites may be used. The characterisation of the cleavage products suggested that the processing for the release of VP4 occurs either at the ^{487}Ala-Ala^{488} or the ^{512}Ala-Ala^{513} site in a single polyprotein molecule. Ultimately, an IBDV polyprotein processing strategy that would explain the release of the additional products was proposed in the present study.

The present study also illustrated the importance of Pro^{377} in the processing of the polyprotein where its replacement with Leu induced a prominent change in polyprotein processing. The mutation seemed to induce structural changes that may possibly affect the cleavage sites.

Although no autocatalytic activity was observed during the expression of VP4-AA (mature form), it cleaved mutant VP4-RK in *trans*. It seemed to be active as a dimer on a gelatine gel but no activity was observed against a dialanyl fluorogenic peptide substrate. It also appeared to form peptidase-inhibitor complexes with anti-thrombin III.

The present study also describes attempts to detect native VP4 in IBDV-infected bursa homogenates by anti-VP4 peptide antibodies on a western blot and by proteolytic activity determination on gelatine-containing SDS-PAGE gels.

The findings of the study provide new information that may contribute to the development of anti-viral agents. These anti-viral agents may target polyprotein processing, capsid assembly and thus prevent virus replication during IBDV infection.

ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to God, the Father of my Lord and Savior Jesus Christ. The work was not easy but “being confident of this, that he who began a good work in me will carry it on to completion (Philippians1:6) is the Word that became truth in this work.

I would also like to express my deep and sincere gratitude to the following people and institutions:

My supervisor, Prof. Theresa Coetzer for her supervision with patience and knowledge whilst allowing me the room to work in my own way, for her detailed and constructive comments during the writing stages of the thesis and for her support and mentoring while I was in LEAP (Leadership Equity Advancement Program).

I am grateful to Prof Alain Boulangé for his guidance and assistance.

Prof. Dean Goldring, Dr Edith Elliott and Dr Carola Niesler, who later became colleagues for their help; Charmaine Ahrens and Robyn Hillebrand for their diligent help with administration; Jess Subramami and Yegan Pillay for assisting with finding reagents and helping with technical problems. I also would like to thank Megan Brunkhorst for help with incubators.

N-terminal sequencing was conducted at the Centre of Protein Engineering (CIP) (University of Liege, Belgium), kindly facilitated by Dr Paulette Charlier, who also hosted me during my three-week visit.
Dr Rodger Horner and Raina Maharaj of Allerton Regional Veterinary laboratory (Pietermaritzburg, South Africa), for kindly providing the infected bursa samples.

My laboratory mates, the Tyrps/IBDV team (Cara Bartlett, Davita Pillay, Hlumani Ndlovu, Kayleen Brien, Laura Huson, Laurelle Jackson, Lorelle Bizaaré, Perina Vather, Richard Kangethe and Sabelo Hadebe) and Malaria team (Bridgette Cumming, David Choveaux, Ikechukwu Achilonu, Jackie Viljoen, Ramona Hurdayal, Robert Krause) for their assistance, laughter and for providing a free working environment.

My parents (Leah and Daniel), especially my mom for her love, encouraging words, support and always giving me a shoulder to cry on when it was difficult; my siblings; Sbuti, Munene, Nyeleti and Boyana for their love, support and laughter. I love you guys.

To my spiritual parents Pastors James and Emma White and Full Counsel Church members for their prayers and love; Aunt Nomfundo Molefe for her motherly love and encouraging words all the time and Mama Sheila for her weekly encouraging words and hugs.

Uncle Stan, Aunt Saroj and family for their love, support and prayers.

To all my friends for their moral support and all the fun we had.

The Oppenheimer family and Career Wise for funding during my first 4 years of university and NRF for funding this work through the Thuthuka Programme.

To God (The Father, the Son and the Holy Spirit) be the Glory !!!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2xYT	yeast tryptone media
2xYT-Amp	yeast tryptone media containing ampicillin
3C, 3CL	chymotrypsin-like protease
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
AMC	aminomethyl coumarin
AMV	avian myeloblastosis virus
BCA	bicinchoninic acid
BMGY	buffered media glycerol yeast extract
BMM	buffered minimal media
BLAST	basic local alignment search tool
BSA	bovine serum albumin
BSNV	blotched snakehead virus
cDNA	complementary deoxyribonucleic acid
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
CEF	chicken embryo fibroblast
DEPC	diethylpyrocarbonate
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
dNTP	deoxy nucleotide triphosphate
ds	double stranded
DTT	dithiothreiotol
DXV	drosophila X virus
E-64	*trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino)butane
EDTA	ethylene diamine tetra-acetic acid
ELISA	enzyme linked immunosorbent substrate assay
FMDV	foot and mouth disease virus
GSH	reduced glutathione
GSSG	oxidised glutathione

GST	glutathione S transferase
HAV	hepatitis A virus
HCV	hepatitis C virus
HC-Pro	helper component-protease
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRV	human rhinovirus
IBD	infectious bursal disease
IBDV	infectious bursal disease virus
ICTV	International Committee on Taxonomy of Viruses
ILRI	International Livestock Research Institute
IPNV	infectious pancreatic necrosis virus
IPTG	isopropyl-β-D-thiogalactopyranoside
LSCC-BK3	a bursal-derived lymphoblastoid cell line
MBS	*m*-maleimidobenzoyl-N-hydroxysuccinimide ester
MBP	maltose binding protein
M-MLV	moloney murine leukemia virus
NS	non-structural
nsP	non-structural polyprotein
OpdB	oligopeptidase B
ORF	open reading frame
OVA	ovalbumin
PBS	phosphate buffered saline
PDB	protein data bank
PCR	polymerase chain reaction
PEG	polyethylene glycol
PL1, PL2	papain-like proteases
POP	prolyl oligopeptidase
PVDF	polymer of vinylidene fluoride
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction

SAP	shrimp alkaline phosphatase
SARS	severe acute respiratory syndrome
SBTI	soybean trypsin inhibitor
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sIg	surface immunoglobulin
SVP	subviral particle
TAE	tris acetate buffer
Trx	thioredoxin
TV-1	Tellina virus-1
VLP	virus-like particle
VP	viral protein
VPg	genome-linked viral protein
VP4-AA	IBDV polyprotein sequence from Ala^{513} to Ala^{755}
VP4-AW	IBDV polyprotein sequence from Ala^{513} toTrp^{891}
VP4-RA, VP4-R^1A	IBDV polyprotein sequence from Arg^{453} to Ala^{755}
VP4-RK	IBDV polyprotein sequence from $Arg^{453}$to Lys^{722}
VP4-R^1W	IBDV polyprotein sequence from Arg^{453} to Trp^{891}
VP4-R^2W	IBDV polyprotein sequence from Arg^{505} to Trp^{891}
VP4-R^2A	IBDV polyprotein sequence from Arg^{505} to Ala^{755}
vvIBDV	very virulent IBDV strains
YNB	yeast nitrogen base
YPD	yeast peptone dextrose media
^nX-X^n	cleavage site (X = amino acid residue; n = residue number)
X^n-X	protein or polypeptide sequence (X = amino acid residue; n = residue number)

CHAPTER 1

Literature review

1.1 Introduction

Infectious bursal disease (IBD) or Gumboro disease is an immunosuppressive disease of young chickens, caused by infectious bursal disease virus (IBDV) infection with tropism for the bursa of Fabricius. The disease became known in 1957 in the United States of America (USA) and now occurs in most parts of the world. The disease is highly contagious and spreads quickly through flocks, thus resulting in major economic losses for the poultry industry. It is distributed through contaminated feed and water and is commonly diagnosed in 3-6 week old chickens (Muller *et al.*, 2003).

The symptoms of the disease are diarrhoea, trembling, weight loss, paleness, depression, lameness, ruffled feathers and ultimately death. The course usually takes 5-7 days during which mortality rises rapidly. The primary target organ, the bursa of Fabricius, shows lesions within 24 hours after infection (Kaufer and Weiss, 1980). Chickens surviving the disease have permanent immunosuppression (van den Berg *et al.*, 2000), which results in increased susceptibility to a variety of infectious diseases and poor or no response to vaccination (Liu *et al.*, 1994). Chickens are important in the production of meat and eggs. Being faced with the challenge of maintaining high quality production on the one hand and the problem of infectious diseases on the other, there is a need for effective prevention and control of infectious diseases in regions where outbreaks can cause food shortages and economic hardships.

IBDV is highly infectious and very resistant to inactivation. It can persist in poultry houses even after thorough cleaning and disinfection. Vaccination is the only available control system for IBDV, however, the efficacy is challenged by several factors. Firstly, in order to induce high titres of maternally derived antibodies that persist over the entire laying period, layers are vaccinated with inactivated vaccines. After hatching, chickens are immunised with live vaccines. However, the time-point of vaccination is crucial as persisting maternally derived antibodies might neutralise the vaccine. Therefore, the efficacy of live IBDV vaccines decreases in the presence of maternal antibodies, which are essential for the protection of young chickens during the first few weeks. Secondly, live IBDV vaccines also cause some degree of bursal atrophy and may contribute to the emergence of antigenic variant virus strains. Thirdly, the emergence of very virulent strains, as they can break through immunity provided by vaccine strains (Muller *et al.*, 2003; Wang *et al.*, 2007).

In view of these problems, there is still a need for new measures for disease control. A promising strategy is to develop anti-viral agents that target crucial processes in the virus life cycle such as polyprotein processing. Polyprotein processing is a common mechanism used by many viruses to produce proteins required for the assembly of infectious viral particles. Thus, polyprotein processing is a potential target for the development of antiviral drugs. Processing of the IBDV polyprotein by its encoded protease is the focus of the present study.

1.2 Infectious bursal disease virus (IBDV)

1.2.1 Taxonomy and structure

IBDV is a member of the genus *Avibirnavirus* of the family *Birnaviridae*. The family Birnaviridae comprises two other genera: *Aquabirnavirus* including infectious pancreatic necrosis virus (IPNV) of salmonid fish and *Entemobirnavirus* which includes drosophila X virus of the fruit fly (Dobos, 2002). Another aquatic birnavirus, blotched snakehead virus (BSNV), has been characterised and is considered a representative of a new genus rather than belonging to genus *Aquabirnavirus* (Da Costa *et al.*, 2002). Recently, tellina virus 1 (TV-1), a birnavirus isolated from the marine bivalve mollusk *Tellina tenuis* has also been characterised and genome sequence analysis established that it is representative of a viral cluster distant from other birnaviruses (Nobiron *et al.*, 2008).

The IBDV double stranded (ds) RNA genome has two segments, A and B (Fig. 1.1). The larger segment A (3.2 kb) contains two partially overlapping open reading frames (ORF) A1 and A2. The short ORF A1 overlaps with the 5' end of the long ORF A2 (Yamaguchi *et al.*, 1997). The short ORF A1 encodes VP5, a 17 kDa highly basic, cysteine rich protein which is conserved among IBDV pathogenic strains (Yao *et al.*, 1998). It is a membrane pore-forming protein which plays a role in host cell penetration and the release of the virus into the host cell (Yao *et al.*, 1998; Lombardo *et al.*, 2000). It is expressed in IBDV-infected cells but it is not incorporated into the virion (Mundt *et al.*, 1995).

The long ORF A2 encodes a 114 kDa precursor polyprotein (NH_2-pVP2-VP4-VP3-COOH) which is processed into pVP2 (residues 1 to 512), VP4 (residues 513 to 755) and VP3 (residues 756 to 1012), through the autoproteolytic activity of the integral VP4 protease (Lejal *et al.*, 2000). The precursor VP2 (pVP2) undergoes secondary processing, which starts at the C-terminus and progresses towards the N-terminus to release VP2 (residues 1 to 441), and four structural peptides (residues 442 to 487, 488 to 494, 495 to 501 and 502 to 512) (Da Costa *et al.*, 2002). The smaller RNA segment B (2.9 kb) encodes VP1, a

90 kDa RNA-dependent RNA polymerase that also functions as a protein primer for RNA synthesis. The initial step in IBDV protein priming is the self-guanylation of VP1 through covalent binding of a guanosine monophosphate moiety to the hydroxyl group of a serine residue through the α-phosphate group via a phosphodiester bond.

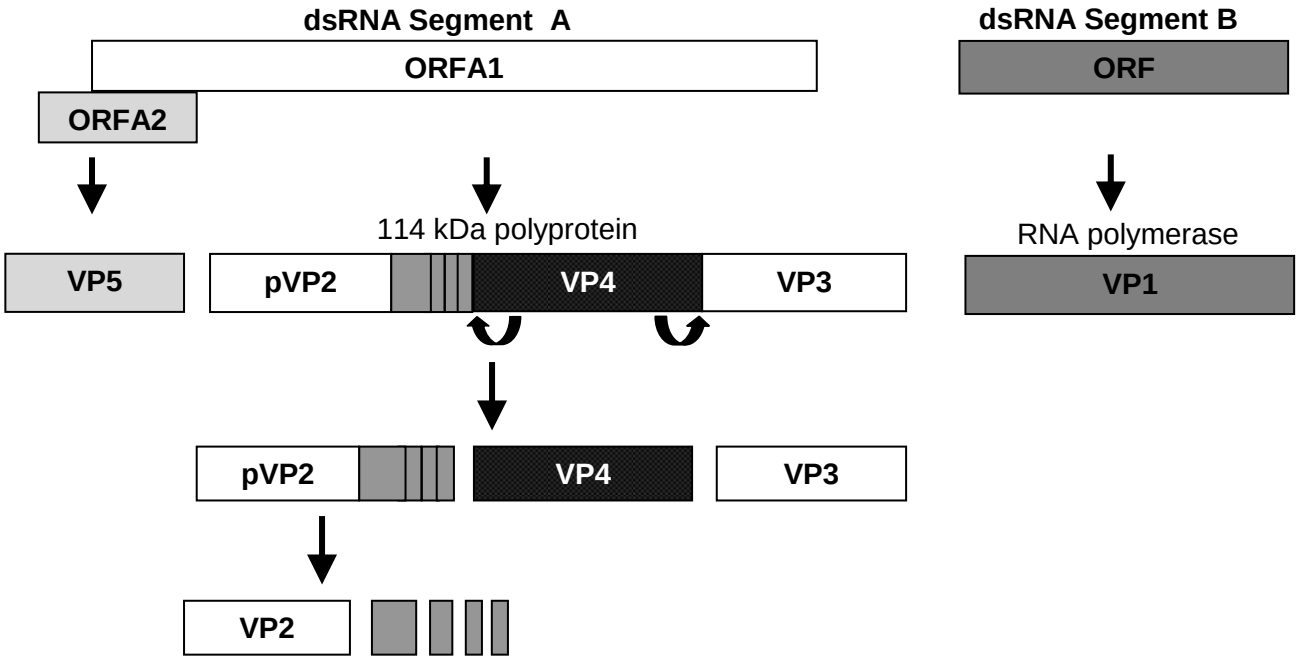


Figure 1.1 The IBDV genome organisation and expression strategy. The boxes represent the coding sequences as open reading frames (ORF). The large arrows show the translation steps and polyprotein processing steps. The small curved arrows show the major polyprotein processing sites.

VP2 and VP3 are the major structural proteins of the virion. VP2 is the major host-protective immunogen and carries all the neutralising epitopes responsible for antigenic variation (Brown *et al.*, 1994). It is also the receptor binding protein (Yip *et al.*, 2007) that contains tissue culture adaptation determinants (Lim *et al.*, 1999; Mundt, 1999; van Loon *et al.*, 2002). VP3 is a multifunctional protein (Casanas *et al.*, 2008) that plays a role in virus assembly, providing a scaffold for the assembly of the capsid (Maraver *et al.*, 2003b). It interacts directly with VP1 through its C-terminal tail and by so doing incorporates it into the virion (Lombardo *et al.*, 1999; Maraver *et al.*, 2003a). The VP3-VP1 interaction removes the innate structural obstruction of the polymerase active site (Luque *et al.*, 2009). Structural studies on the central region of VP3 showed significant similarities with different transcription regulation factors, suggesting that VP3 may also act as a transcriptional activator (Casanas *et al.*, 2008). Its basic C-terminal region also extends inward to interact with the dsRNA allowing for the encapsidation of the viral genome (Tacken *et al.*, 2000).

Birnaviruses are the only dsRNA viruses that initiate RNA synthesis using protein primers (Pan *et al.*, 2009). VP1 is present in the virion in two forms, as a free protein or as a genome-linked viral protein (VPg) which is covalently linked to the 5' ends of the two genomic segments (Tacken *et al.*, 2000; Pan *et al.*, 2009).

1.2.2 Capsid assembly

IBDV is a non-enveloped icosahedron (T=13 Laevo symmetry) with a diameter of 60 nm (Fig. 1.2A). Cryoelectron microscopy and image processing analysis showed that the capsid is formed by a single shell with a thickness of approximately 9 nm. The exterior of the capsid is formed by 260 trimers of VP2 while VP3 is present in the inner surface (Bottcher *et al.*, 1997). It has also been shown that the trimeric subunits contain a variable pVP2/VP2 ratio (Chevalier *et al.*, 2002). Preparations of purified IBDV contain full and empty icosahedrons, type I tubules with a diameter of 60 nm and type II tubules of 24 ± 26 nm in diameter (Fig. 1.2B and C). Type I tubules have a hexagonal arrangement of trimeric units, similar to those observed in IBDV virions (Martinez-Torrecuadrada *et al.*, 2000), with their surfaces formed by hexamers of pVP2 trimers. Type II tubules seem to contain VP4 and are only detected in IBDV-infected cells (Granzow *et al.*, 1997).

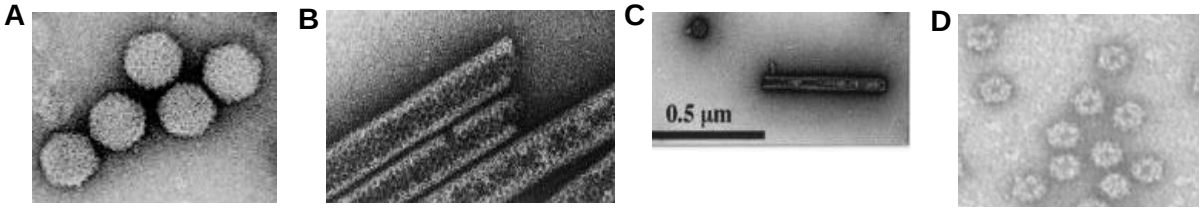


Figure 1.2 Different assemblies of IBDV particles. A) IBDV virions**, B)** Type I tubules (Chevalier *et al.*, 2004), **C)** Type II tubule (Chevalier *et al.*, 2002) and **D)** SVPs (Caston *et al.*, 2001).

The expression of VP2 in baculovirus (Caston *et al.*, 2001; Lee *et al.*, 2004) and yeast cells (Dey *et al.*, 2009) leads to the assembly of icosahedral T=1 subviral particles (SVPs) of 23 nm in diameter (Fig. 1.2D). The SVPs showed the same immunogenicity as the intact IBDV virion (Caston *et al.*, 2001). SVPs are made up of VP2 alone and the inner surface is negatively charged, as is the outer surface, therefore it does not support dsRNA encapsidation (Lee *et al.*, 2006a).

VP2 trimers are the common building blocks of the single-shelled capsid between T=13 and T=1 particles. Each VP2 subunit is folded into a helical base (B) domain, a shell (S) domain and a protrusion (P) domain (Fig. 1.3A and B). The B-domain consists mainly of

α-helices from the N- and C-termini of VP2 (Garriga *et al.*, 2006). VP2 is homologous to the capsid proteins of T=3 positive sense RNA nodaviruses and the T=13 dsRNA reoviruses (Coulibaly *et al.*, 2005). All three domains participate in inter- and intra-trimer interactions, which stabilise the trimers. In the T=13 IBDV virion or the T=1 SVP, each trimer interacts with other trimers mainly through the S domain via the sides of the equilateral triangles (Fig. 1.3D and E) (Coulibaly *et al.*, 2005).

The VP2 trimers have been extensively studied (Coulibaly *et al.*, 2005; Garriga *et al.*, 2006; Lee *et al.*, 2006a). The study of VP2 trimers revealed a calcium binding site on the internal surface and a chloride ion on the external surface of the SVP. The calcium ion binds to three pairs of symmetry-related Asp^{31} and Asp^{174} in domain S (Fig. 1.3B). The calcium-binding residues are conserved among four birnaviruses, IBDV, IPNV DXV and BSNV, however, in DXV the Asp^{174} is substituted by Asn but it still retains the calcium binding capacity. The Ca^{2+} coordinates the two Asp residues from three VP2 subunits in an octahedral geometry thus acting as a sealing element of the VP2 trimers (Lee *et al.*, 2006a).

The structure of VP2 from Garriga *et al.* (2006) shows an additional C-terminal helix α4 (Fig. 1.3C), which is not shown in the VP2 structure by Coulibaly *et al.* (2005) and Lee *et al.* (2006a). The C-terminal helix α4 is projected away from the protein core but is orientated towards the 3-fold axis of the neighbouring VP2 trimer thus mediating interaction between trimers. This C-terminal helix α4 and the bound Ca^{2+} are the stabilising elements of the viral capsid. The rotavirus major capsid protein VP6 also forms trimers and has a Zn^{2+} bound at the threefold axis of each trimer (Lee *et al.*, 2006a).

The interactions between α4 helices between S domains of neighbouring trimers lead to a pentameric grouping of trimers resulting in a T=1 icosahedral capsid. In the T=13 capsid of IBDV virions, the VP2 trimers are grouped not only in pentamers but also in hexamers (Lee *et al.*, 2006a). Luque *et al.* (2007) elucidated the molecular mechanism that defines the multimeric state of VP2 as hexamers and pentamers where it was shown that the pVP2 C-terminal region determines whether a pentamer or hexamer will be formed.

The expression of VP2 results in SVPs, but expression of VP2 fused with an N-terminal His tag led to efficient assembly of VLPs (Saugar *et al.*, 2005). VP3 C-terminal residues (DEDLE) show some similarities to the acidic region of the His tag (DYDIPTTE), therefore the His tag mimicked the VP3 oligomerisation function during VLP assembly (Maraver *et al.*, 2003b; Luque *et al.*, 2007). VP3 is responsible for the formation and/or stabilisation of

the five-fold vertex, which is crucial for the bending of the pVP2/VP2 lattice into icosahedral capsids (Martinez-Torrecuadrada *et al.*, 2000).

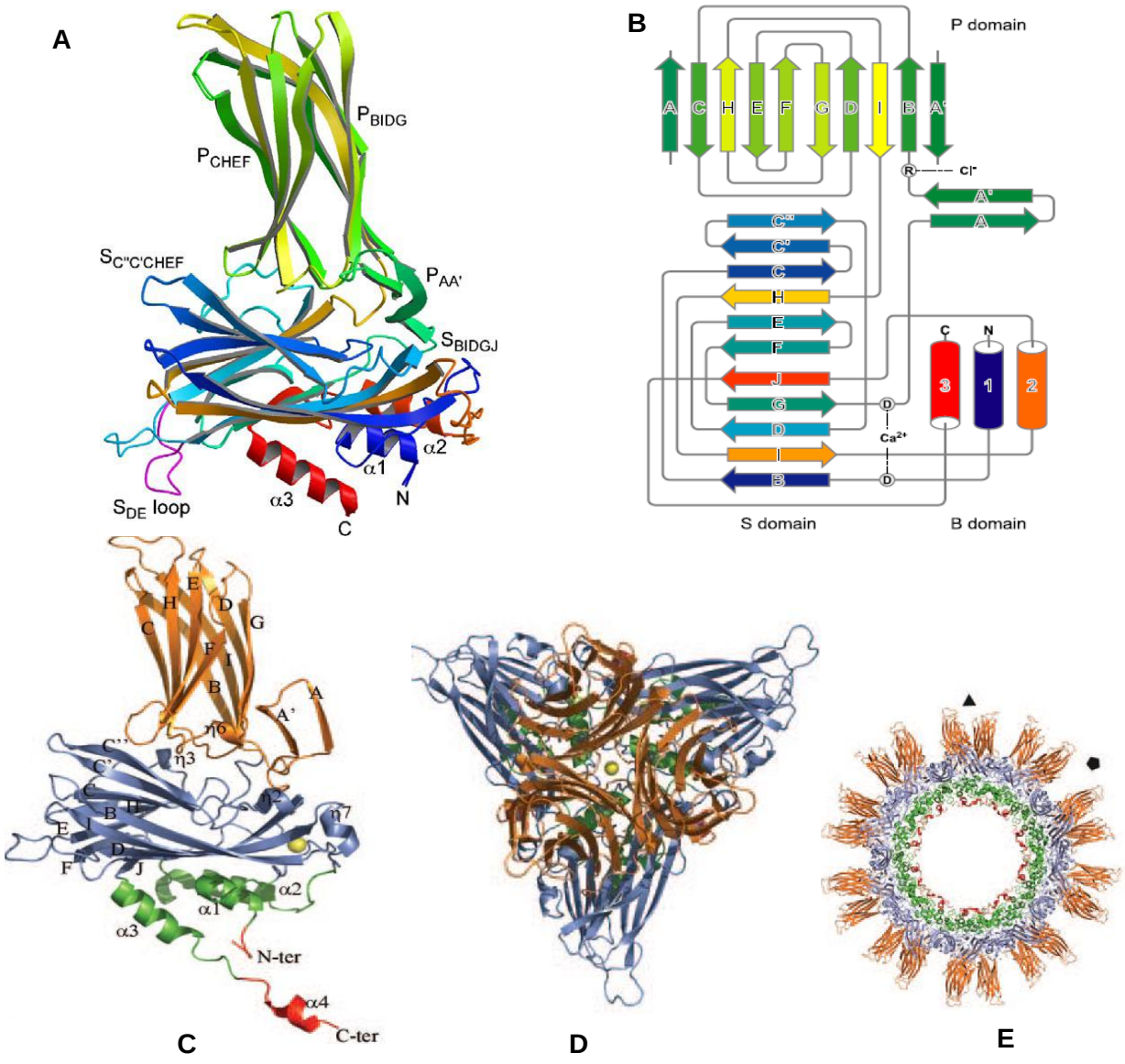


Figure 1.3 The structure of IBDV VP2 and its trimers. A) The ribbon structure of VP2. Major secondary structural elements are labeled, including the β-sheets of strands BIDGJ and C"C'CHEF in the S (shell) domain, the AA' β-ribbon, the BIDG and CHEF sheets of the P (protrusion) domain, as well as the three major helices α1, α2 and α3 of the B (base) domain. **B)** The strands and helices are shown in a topology diagram with arrows and cylinders, colored as in A (Lee *et al.*, 2006a). **C)** The ribbon structure of VP2 showing the additional C-terminal α4 helix. **D)** The top view of a VP2 trimer. The colors used to show the different domains are the same as in C. The Ca^{2+} ion at the centre of the trimer is shown as a yellow sphere (Garriga *et al.*, 2006). **E)** A ribbon structure of T=1 SVP, the colors used are the same as in C and D (Garriga *et al.*, 2006).

While SVPs are associated with VP2, different assemblies have been reported for the expression of pVP2. Caston *et al.* (2001) reported formation of type I tubules in Sf9 cells,

while expression in Hi-5 cells resulted in the formation of irregular particles and occasionally twisted tubular structures (Lee *et al.*, 2004). The observed discrepancies seem to be associated with the ability of cells to process pVP2. It was shown that pVP2 was processed into VP2 in Hi-5 insect cells in the absence of VP4 but not in Sf9 cells (another insect cell line) and inefficient processing generated irregular particles (Lee *et al.*, 2006b).

Expression of the polyprotein results mainly in the formation of IBD virus-like particles (VLPs) but other observations have also been made. Although VLPs do not contain genomic RNA, they have the same T=13 laevo icosahedral geometry as the virions, therefore capsid assembly does not depend upon the presence of RNA (Pous *et al.*, 2005). VLP formation was initially shown in African green monkey kidney epithelial BSC-1 cells in the absence of VP1 (Fernandez-Arias *et al.*, 1998) and in its presence (Lombardo *et al.*, 1999; Chevalier *et al.*, 2004). The expression of the polyprotein in Sf9 and Hi-5 cells resulted in a mixed population of particles consisting of VLPs, type I and type II tubules (Lee *et al.*, 2006b). Chevalier *et al.* (2004) studied capsid assembly in Sf9 cells and they showed that the polyprotein alone assembled into type I tubules. However, when an exogenous sequence [green fluorescence protein (288aa) or truncated ovalbumin (259 aa)] was fused with VP3, VLPs were formed.

Expression of polyprotein sequences in which the last five residues (DEDLE) or the C-terminal Glu^{1012} residue were deleted gave rise to VLPs. In contrast, expression of a polyprotein in which the last ten residues were deleted, abolished VLP formation, but resulted in type I tubules. The study showed that Glu^{1012} negatively controls capsid assembly, thus for VLP formation it has to be deleted, or hidden through interaction with an exogenous protein. The removal of the ten C-terminal residues completely abolishes VP3-VP1 complex formation which in turn, abolishes VLP formation (Chevalier *et al.*, 2004).

VP3 is proteolytically trimmed at the C-terminal end to release 13 amino acid residues. The accumulation of the cleaved or trimmed VP3 results in the formation of type I tubules, thus showing that the cleaved form is unable to trigger the bending of the capsomere lattice. The formation of VP3-VP1 complexes blocks the proteolytic cleavage of the C-terminal end of VP3 thus allowing the proper functioning of the protein. The protection of the VP3 C-terminus by VP1 binding is required on a limited number of VP3 molecules, just sufficient for capsid assembly (Maraver *et al.*, 2003a). The VP3-VP1 complexes are

formed immediately or shortly after translation in the cytoplasm of IBDV-infected cells (Tacken *et al.*, 2000).

The analysis of purified IBDV particles reveals the presence of VP1, VP2 and VP3 when analysed by SDS-PAGE. Mass spectrometry studies identified the presence of the structural peptides in the virions and VLPs (Da Costa *et al.*, 2002). Therefore, the assembly process is associated with pVP2 processing into VP2 (Chevalier *et al.*, 2002). Four peptides are released during pVP2 processing, namely pep46 (residues 442 to 487), pep7a (residues 488 to 494), pep7b (residues 495 to 501) and pep11 (residues 502 to 512) (Da Costa *et al.*, 2002). A polyprotein sequence with a pep46 deletion resulted in the assembly of SVPs instead of VLPs, whereas pep11 deletion led to no assembly (Chevalier *et al.*, 2005). This further confirms the relationship between capsid formation and pVP2 maturation. Pep46 contains a positively charged N-terminus which forms an amphipathic α-helix, hydrophobic central domain which suggests a role of membrane interaction and a negatively charged C-terminus (Chevalier *et al.*, 2005; Galloux *et al.*, 2007). A similar peptide of about 40 residues that can destabilise the cell membrane is present in noda- and tetraviruses. The N-terminal 22 residues of pep46 were shown to be more efficient in membrane permealisation than membrane active domain of nodavirus γ peptide (Galloux *et al.*, 2007).

The T=13 capsids of IBDV virions contain VP3, which is not present in T=1 SVPs. The C-terminus of VP3 contains oligomerisation domains that interact with both VP2 and VP1 (Maraver *et al.*, 2003b). The correct functioning of VP3 is also important for the processing of pVP2 to VP2 (Chevalier *et al.*, 2002). It is clear that correct IBDV capsid assembly is dependent upon complete and efficient processing of the polyprotein, thus showing the importance of VP4 protease. The role of pep46 in membrane permealisation shows a link between polyprotein processing and virus entry.

1.2.3 IBDV classification and virulence

Two serotypes of IBDV, termed serotype I and II, have been recognised. Serotype I viruses are pathogenic to chickens, whereas serotype II viruses isolated from turkeys are apathogenic to chickens (Vakharia *et al.*, 1994). Serotype I viruses cause severe infections in chickens while serotype II only cause sub-acute infections in turkeys (Snyder *et al.*, 1992). The serotype I viruses differ in their virulence and can further be divided into four major groups: classical virulent, attenuated, antigenic variant and very virulent (vvIBDV) strains (Lim *et al.*, 1999). *In vivo* studies categorise the subtypes into classic and

variant strains (Ismail and Saif, 1991) whereas *in vitro* studies conducted using monoclonal antibodies and nucleotide sequencing classify the serotype I subtypes into the four distinct groups (Vakharia *et al.*, 1994; Eterradossi *et al.*, 1998).

Classical virulent strains, isolated from flocks in the USA in the early 1960's, induce severe lymphoid necrosis and bursal inflammation in infected chickens, thus causing 30 – 60% mortality. In the late 1980's, antigenic variant strains, which are highly cytolytic and cause rapid bursal atrophy without inflammation, were isolated from flocks in the Delmarva Peninsula. In the mid 1990's, vvIBDV strains causing severe clinical symptoms and high mortality emerged in Asia and Europe. Even though these strains showed enhanced virulence, they have the same antigenic structure as classical strains (Boot *et al.*, 2000a). Lastly, there are attenuated strains which are adapted to primary cell culture such as chicken bursal lymphoid cells, chicken embryo kidney cells and chicken embryo fibroblast (CEF) cells through serial passages. They have reduced virulence and do not cause the disease in chickens and are thus used as vaccines (Snyder *et al.*, 1992; Lim *et al.*, 1999; Brandt *et al.*, 2001).

VP2 has been shown to be a determinant of IBDV virulence and serotype specificity. It contains the variable region or the major conformational antigenic domain at positions 206 to 350 in which mutations in IBDV strains occur (Bayliss *et al.*, 1990). In the hydrophobicity plot of the VP2, Etteradossi *et al.* (1998) showed that the domain consists of a hydrophobic region flanked by two hydrophilic peaks, A and B, which cover residues 210 to 225 and 312 to 324, respectively. Residues at positions 253, 279 and 284 were shown to be responsible for the attenuation of the vvIBDV strains (Yamaguchi *et al.*, 1996b; Lim *et al.*, 1999), whereas residues at positions 222, 249, 254, 270 and 330 were shown to play a role in the variant phenotype (Brandt *et al.*, 2001).

Vaccination with one pathotype does not ensure protection against another, but cross-protection can occur between many of the pathotypes (Ismail and Saif, 1991). It was shown that a variant strain vaccine protected chickens against a classic strain, suggesting that all the pathotypes contain shared epitopes that induce cross-protection. It was also suggested that the epitopes may not be expressed equally (Oppling *et al.*, 1991). It was then later suggested that the cross-protection between antigenic pathotypes may be caused by the existence of quasispecies expressing different antigens in the vaccine preparation (Domingo *et al.*, 1997). Jackwood and Sommer (2002) demonstrated that some IBDV strains indeed contain quasispecies at the hydrophilic peak B. This region has

relatively high mutation rates and thus contributes to the antigenic diversity of the quasispecies.

1.2.4 Pathogenesis and the bursa of Fabricius

The target organ for IBDV is the bursa of Fabricius, the primary site of B-cell development and generation of antibody diversity in avian species. The bursa starts developing from an epithelial bud which is first detected on the fourth day of the embryonic development in the cloacal region of a chicken. The B-cell precursors are first observed in the bursa on the eighth day and they subsequently migrate across the bursal epithelial basement membrane and begin to proliferate in epithelial buds to form the bursal follicles (Sayegh *et al.*, 1999). Each follicle consists of three zones, the cortex, a thin epithelial basement membrane and the medulla. Movement of cells is within and not between follicles, hence each follicle contains the progeny of a single donor cell (Ratcliffe *et al.*, 1996).

Only B-cell precursors that undergo productive immunoglobulin (Ig) gene rearrangement and consequently express surface immunoglobulin (sIg), can migrate across the basement membrane (Ratcliffe *et al.*, 1996). Productive Ig gene arrangements result in the production of the heavy or light chain and the cells can therefore express sIg. The cells that do not express sIg undergo apoptosis (Neiman *et al.*, 1991; Ratcliffe *et al.*, 1996). Apoptosis is associated with agglutinin binding to a cell surface glycoprotein through an α(1-6)-linked fucose moiety (Funk and Thompson, 1998), whereas, the survival of B-cells is associated with the expression of the cell surface chL12 antigen (Lampisuo *et al.*, 1998). The developing B-cells start expressing the chL12 antigen as they become mature and are ready to migrate (Lampisuo *et al.*, 1998).

The cells which express sIg continue to divide in the bursal cortex. As a result, the cortex contains densely packed cells while the medulla contains loosely packed cells. It is during this cell division that Ig genes are diversified by somatic gene rearrangement. Shortly before hatching, diversified B-cells move from the bursa across the bursal epithelial basement membrane to the secondary lymphoid organs and blood and are therefore referred to as post bursal B-cells (Nieminen *et al.*, 2002).

While mammals generate antibody diversity throughout their lives, chickens only do so during a short period before hatching (Tizard, 2002). As a result, the chicken survives throughout adult life with these postbursal B-cell as no new B-cells are produced after hatching (Ratcliffe *et al.*, 1996). The bursa reaches its maximal stage of development at

the age of three to six weeks (van den Berg *et al.*, 2000). Therefore, by the time sexual maturity is reached, the bursa reduces in size and the post-bursal cells in the secondary lymphoid organs are the only source of B-cell renewal throughout the lifetime of the chicken (Nieminen *et al.*, 2002). When chickens finally lay eggs, they pass the maternal antibodies on to their offspring to provide early immunity to the newly hatched chicken for the first four or five weeks. However, these antibodies interfere with vaccination in the young chicken until they disappear between 10 to 20 days after hatching (van den Berg and Meulemans, 1991).

IBDV infection causes a depletion of the dividing B-cells thus resulting in severe immunosuppression that is accompanied by little or no inflammatory response (Tham and Moon, 1996). IBDV virus replication was associated with apoptosis of chicken peripheral blood lymphocytes (Vasconcelos and Lam, 1994), bursal lymphoid cells (Vasconcelos and Lam, 1995), CEFs, Vero cells (a kidney epithelial cell line) (Tham and Moon, 1996) and thymus (Tanimura and Sharma, 1998), as well as necrosis and haemorrhage (Kim *et al.*, 2000). Apoptosis is characterised by nuclear fragmentation and cellular breakdown into apoptotic vesicles, but unlike necrosis there is no release of cellular contents and consequently no inflammation surrounding the dead cells (Vasconcelos and Lam, 1994).

The number of apoptotic cells in the bursa is correlated with viral replication. VP2 and VP5 are the only two viral proteins that have been associated with apoptosis induction (Fernandez-Arias *et al.*, 1997; Yao *et al.*, 1998). Apoptosis occurs at the late stage of viral replication; about 12 to 14 h post infection. Apoptosis was also suggested to be the anti-viral mechanism of the host to prevent virus spread (Jungmann *et al.*, 2001). VP5 deficient viruses caused earlier and greater apoptotic effects than the wild-type strain. The apoptosis was demonstrated to be caspase-dependent (Liu and Vakharia, 2006). This early apoptosis was accompanied by diminished production of the virus. It was suggested that VP5 acts as an anti-apoptotic regulatory protein at early stages of infection to prevent infected cells from dying before the virus completes its infection cycle and thereafter targets the plasma membranes to induce lysis (Liu and Vakharia, 2006).

IBDV infection was also shown to cause changes in the potassium current, as a result of virus attachment or membrane penetration (Repp *et al.*, 1998). Potassium channels play important roles in cellular signaling processes and as transport proteins for passive potassium ion movement across membranes (Shieh *et al.*, 2000). The effect of IBDV infection on cellular responses was studied by Zheng *et al.* (2008). IBDV infection was shown to turn off the host translational machinery for initiating its viral translation in

infected cells. The infection also caused cytoskeleton disruption, which was suggested to be a mechanism by which IBDV particles are released from infected cells. The infection also suppressed the expression of proteins involved in signal transduction, ubiquitin-mediated protein degradation, stress response, RNA processing, biosynthesis and energy metabolism.

1.3 Polyprotein processing

Polyprotein processing is a crucial step in the life cycle of the majority of viruses that infect eukaryotic cells. It plays a role in the liberation of the structural and non-structural proteins prior to capsid assembly and genome replication (Babe and Craik, 1997). This processing is co- or/and post-translationally mediated by proteases which are encoded as integral parts of the viral polyprotein. Up to three proteases may be encoded in a virus. The processing may also require both host and viral proteases, as in the case of hepatitis C virus (HCV) (Schultheiss *et al.*, 1995). Viral proteases may have host cellular proteins as targets as a mechanism to modulate cell processes in order to promote viral replication. Accurate and precise processing is required for the production of fully infectious viral particles. Mutations that produce imprecise cleavage or alter the order in which sites are cleaved may result in the assembly of less infectious virions (Chevalier *et al.*, 2004).

1.3.1 Genomic organisation and polyprotein processing

There are different ways of classifying viruses. Viral classification starts at the level of order (-*virales*), followed by family (-*viridae*), subfamily (-*virinae*), genus (-*virus*) and species. So far the International Committee on Taxonomy of Viruses (ICTV) has established six orders which include viruses from different hosts: the *Caudovirales, Herpesvirales, Mononegavirales, Nidovirales, Picornavirales, and Tymovirales. Caudovirales* are tailed dsDNA bacteriophages, *Herpesvirales* includes large eukaryotic dsDNA viruses, *Mononegavirales,* non-segmented negative sense ssRNA plant and animal viruses, *Nidovirales*, positive sense ssRNA vertebrate viruses, *Picornavirales*, small positive sense ssRNA viruses that infect plant, insect and animal hosts, and finally *Tymovirales* includes monopartite ssRNA viruses that infect plants. Family *Birnaviridae* is among 65 viral families that have not yet been assigned to an order *(http://www.ictvonline.org/virusTaxonomy.asp*; accessed 19 November 2011).

The Baltimore classification classifies viruses into groups depending on their genome in terms of the nucleic acid type (DNA or RNA), strandedness (single or double-stranded), sense (positive or negative) and method of replication (http://www.web-

books.com/MoBio/Free/Ch1E2.htm, accessed 14 March 2011). Viruses are classified into seven groups: dsDNA viruses (I), positive sense DNA viruses (II), dsRNA viruses (III), positive sense RNA viruses (IV), negative sense RNA viruses (V), ssRNA reverse transcribing viruses (VI) and dsDNA reverse transcribing viruses (VII).

The families *Potyviridae*, *Picornaviridae* and *Flaviviridae*, can be included in the *Picornavirales* order, based on their genome organisation and expression strategy. The genome consists of a 5' covalently linked VPg, a 3' poly (A) tail and a single ORF that is translated into a large polyprotein that is proteolytically processed into individual proteins (Fig. 1.4). The *Potyviridae* 340 - 370 kDa polyprotein is cleaved by three virus-encoded proteases: P1 protease and helper component-protease (HC-Pro) cleave only at their respective C-termini autocatalytically; and NIa protease which cleaves the remaining proteolytic sites (Urcuqui-Inchima *et al.*, 2001; Merits *et al.*, 2002) (Fig. 1.4A). Proteolytic processing of the picornavirus polyprotein is mediated by two proteases, 2A and 3C (Fig. 1.4B). The initial cleavage of the precursor occurs at the VP1-2A junction in *cis* by the 2A protease and the other proteolytic cleavages are largely carried out by the 3C protease (Peters *et al.*, 2005).

Some *Picornaviridae* genera like aphtovirus (foot and mouth disease virus, FMDV) and cardiovirus (encephalomyocarditis virus) also encode another protease, known as the leader (L) protease, located at the N-terminus of the polyprotein (Guarne *et al.*, 1998). In the intramolecular cleavage, the protease cleaves itself at the L-VP1 junction (Fig. 1.4B) (Skern *et al.*, 1998) and subsequently during virus replication, it intermolecularly cleaves a host cell protein, the eukaryotic initiation factor 4G (Hinton *et al.*, 2002; Gradi *et al.*, 2004). The final stage of picornavirus assembly is characterised by an autocatalytic cleavage of the capsid precursor protein VP0 into VP2 and VP4 (Fig. 1.4B) (Hindiyeh *et al.*, 1999).

Viruses within the *Flaviviridae* family (flavi-, pesti- and hepaciviruses) also encode a polyprotein that is co- and post translationally cleaved into structural (C, prM, and E) and non-structural proteins (P7, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). Unlike picornaviruses, the polyprotein of viruses within the *Flaviviridae* family is cleaved by a combination of viral and host proteases (Fig. 1.4C). The structural proteins are cleaved by the host cell proteases furin and secretase, whereas the structural protein boundaries (NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/NS4B, and NS4B/NS5) are cleaved by NS3 protease (Bouffard *et al.*, 1995; Ryan *et al.*, 1998). Pestiviruses also contain an autocatalytic Npro domain that cleaves the N-terminal cleavage site (Ryan *et al.*, 1998).

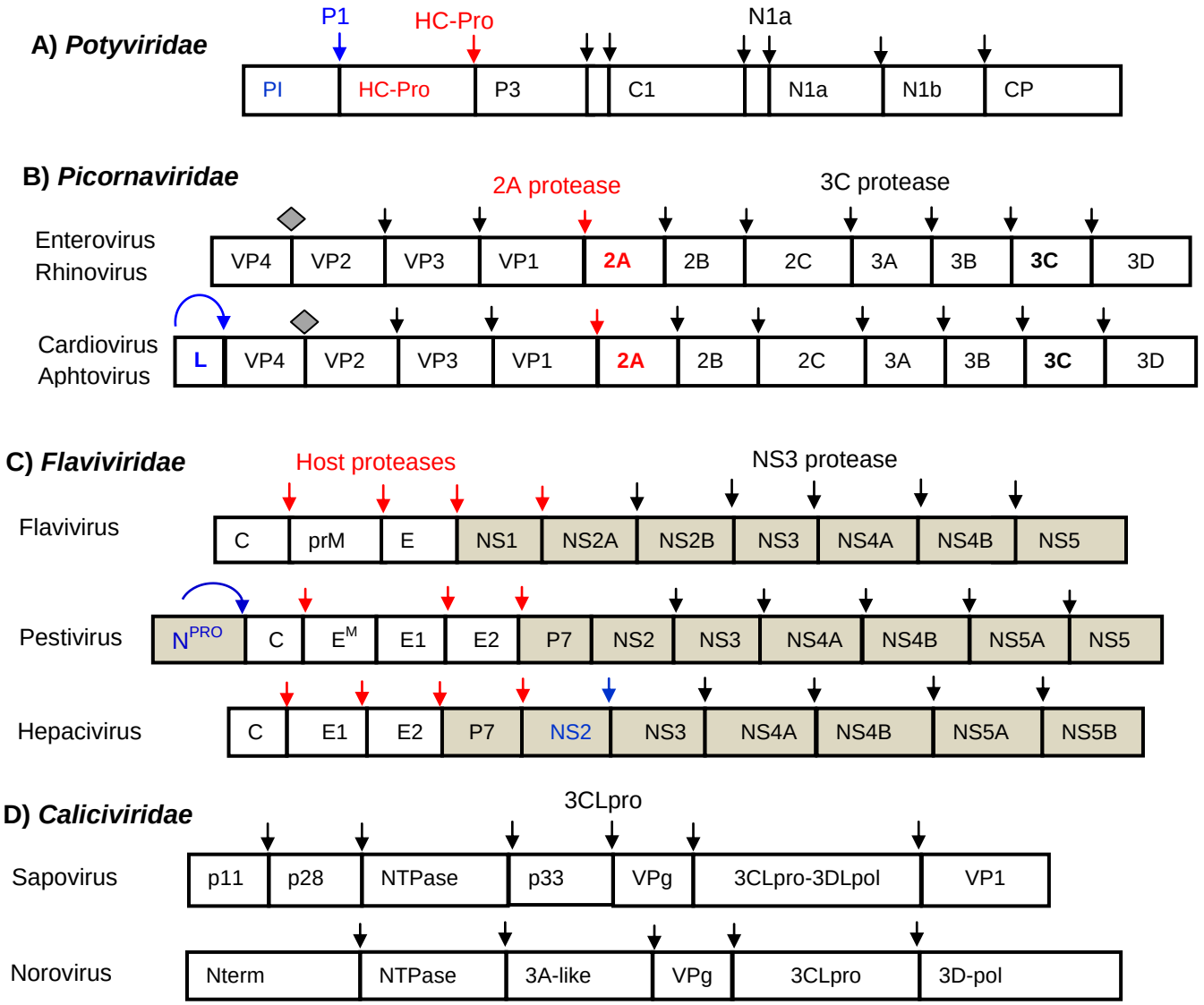


Figure 1.4 The genomic organisation and polyprotein processing maps of *Potyviridae, Picornaviridae, Flaviviridae* and *Caliciviridae* polyproteins. Each polyprotein is encoded by a 5' VPg covalently linked and 3' poly (A) tailed genome containing a single ORF. For the *Flaviviridae* family, the non-structural proteins are shaded. The positions at which cleavage occurs are indicated by vertical lines within the polyprotein. The viral proteases and their cleavage sites in the polyprotein, presented with short arrows are indicated by corresponding colours. The diamond indicates an autocatalytic processing of capsid precursor protein VP0 into VP4 and VP2 in picornaviruses [Figures were adapted from Urcuqui-Inchima *et al.* (2001); Peters *et al.* (2005); Ryan *et al.* (1998) and Matsuura *et al.* (2002)].

In the genus hepacivirus, the C/E1, E1/E2, E2/p7 and p7/NS2 junctions are processed by host signal proteases. The non-structural proteins NS2 to NS5B are released by two virus-encoded proteases: NS2 protease catalyses the cleavage of the NS2-NS3 junction while NS3 processes the remainder of the C-terminal junctions in the polyprotein (Fig. 1.4B). NS2 is a cysteine protease with a catalytic triad that consists of Cys, His and Glu (Table 1.1) (Suzuki *et al.*, 2007; Schregel *et al.*, 2009). Efficient cleavage at the NS2-NS3 site by NS2 requires the first 180 amino acids residues of NS3 (Suzuki *et al.*, 2007). Even though

the *Caliciviridae* genome also has a 5' covalently linked VPg and a 3' poly (A) tail, it is organised into two or three ORFs depending on the genus. ORF1 encodes a polyprotein precursor that is co-translationally processed by a chymotrypsin-like cysteine protease (3CLpro) (Oka *et al.*, 2007) (Fig. 1.4D).

Families *Arteriviridae* and *Coronaviridae* are classified in the *Nidovirales* order. The genome of these viruses is a single stranded positive sense RNA that contains a 5' RNA 7-methylguanosine cap, a 3' poly (A) tail and untranslated regions at both termini (Fig. 1.5). It is also polycistronic, containing a large 5' terminal replicase coding gene, and a downstream set of structural protein coding genes. The structural proteins are expressed via transcription of subgenomic mRNAs from the 3'-terminal region of the genome. The structural proteins of arteriviruses are different from those of coronaviruses. The common structural proteins are envelope protein (E), nucleocapsid protein (N) and membrane protein (M). While arteriviruses have the glycoproteins GP2-5, coronaviruses have a spike protein (de Vries *et al.*, 1997; Gorbalenya *et al.*, 2006).

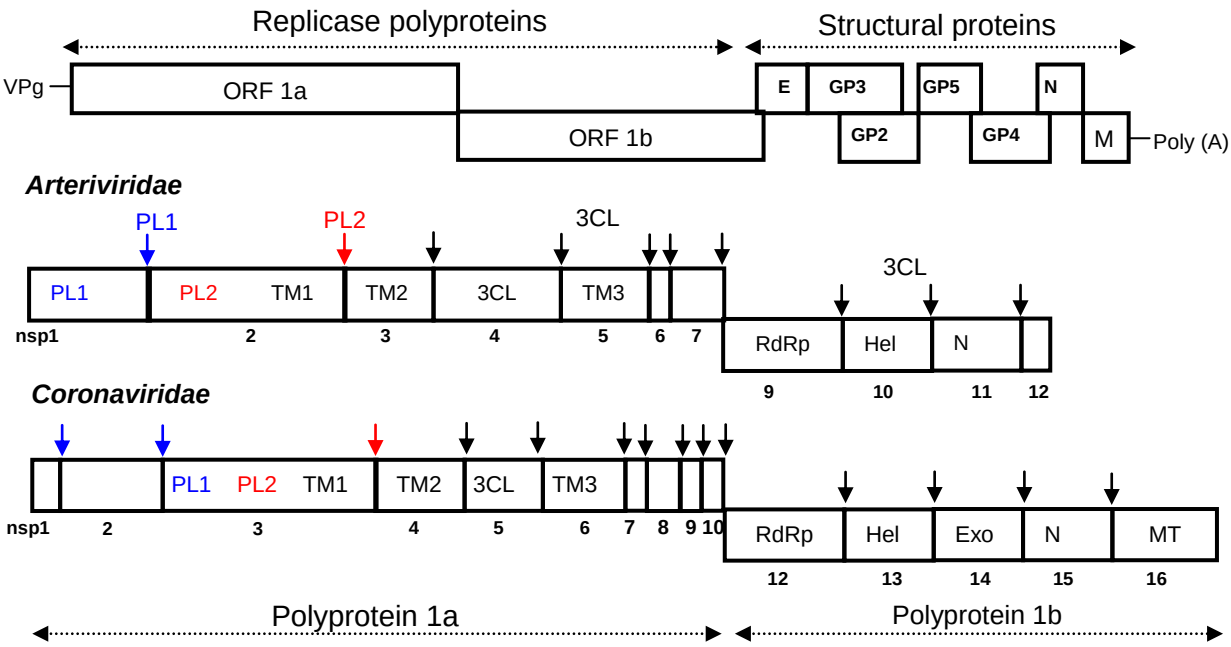


Figure 1.5 Genome organisation of families *Coronaviridae* and *Arteriviridae* of the order *Nidovirale*. The genome indicates the ORFs encoding the replicase polyproteins. Downstream of ORF1b, there are genes that encode structural proteins: envelope protein (E), glycoproteins GP2-5, nucleocapsid protein (N) and membrane protein (M). The gene organisation of structural proteins is specific for arteriviruses. The positions at which cleavage occurs are indicated by vertical lines within the polyprotein. The mapped domains in the nidovirus replicase polyproteins include papain-like cysteine proteases (PL1 and PL2), chymotrypsin-like protease (3CL) transmembrane (TM) domains, RNA-dependent RNA polymerase (RdRp), helicase (Hel), uridylate-specific endoribonuclease (N), exonuclease (Exo) and methyl transferase (MT). The viral proteases and the positions at which they cleave the polyprotein are presented with short arrows: PL1 (blue), PL2 (red) and 3CL (black) [adapted from Gorbalenya *et al*. (2006)]. The replicase proteins are cleaved into 12 to 16 non-structural proteins (nsps) which is denoted by the numbering.

The replicase gene is translated into ORF1a and ORF1b polyproteins, with expression of ORF1b involving a ribosomal frameshift just upstream of the ORF1a termination codon (Pasternak *et al.*, 2001) (Fig. 1.5). For arteriviruses, polyproteins 1a and 1b are cleaved by three different ORF1a-encoded proteases: papain-like cysteine proteases (nsp1 and nsp2 or PL1 and PL2) and a chymotrypsin-like serine protease, 3CL or nsp4. Following the autocatalytic release of PL1 and PL2, 3CL processes the remaining eight sites, five in the ORF1a polyprotein and three in ORF1b encoded polyproteins (nsp9-12) (Barrette-Ng *et al.*, 2002). The latter cleavages release flanking transmembrane (TM) domains as well as domains associated with RNA-dependent RNA polymerase (RdRp), helicase (Hel) and uridylate-specific endoribonuclease (N) activities (Fig. 1.5) (Gorbalenya *et al.*, 2006).

For coronaviruses, three proteases are also involved in the processing of the polyproteins. While the papain-like cysteine proteases, PL1 and PL2 cleave the N-terminus of polyprotein 1a, the 3CL serine protease is involved in the processing of the C-terminus of both polyproteins 1a and 1b (de Vries *et al.*, 1997). In addition to the release of the RdRp, Hel and N domains, the processing of polyprotein 1b also releases the exonuclease (Exo) and methyl transferase (MT) domains (Gorbalenya *et al.*, 2006).

The genomic organisation of family *Togaviridae* is similar to that of nidoviruses (Fig 1.6). The genome has a 5’ cap and a 3’ poly (A) tail. The 5’ terminal segment of the genomic RNA is directly translated into a polyprotein (nsP1-3) containing non-structural proteins. Through a translational readthrough of a UGA stop codon between nsP3 and nsP4 genes, an amino acid is introduced at the stop codon to produce the nsP1-4 polyprotein. NsP1 contains methyltransferase and guanylyltransferase activities that function in capping of viral RNA. The N-terminal region of nsP2 contains RNA 5’-triphosphatase, NTPase and helicase activities while the C-terminal domain contains a cysteine protease that is responsible for processing the non-structural polyprotein. While nsP3 is a phosphoprotein, nsP4 is the viral RNA-dependent RNA polymerase (De *et al.*, 2003).

Another feature of the replication of togaviruses is the synthesis of a single sub-genomic mRNA corresponding to the 3’ region of the genome due to the presence of stop codons at the end of the non-structural polyprotein (nsP) coding region (Fig. 1.6). The translation of the subgenomic mRNA also produces a polyprotein that comprises structural proteins, capsid protein (C), and envelope proteins (E1-3 and 6K) (Lemm *et al.*, 1994; Merits *et al.*, 2001). The N-terminally located capsid protein also functions as a protease that cleaves itself from the structural polyprotein (Nicola *et al.*, 1999).

The sobemovirus genome contains a 5' covalently linked VPg and a non-polyadenylated 3' terminus. The genome organisation consists of four ORFs. ORF1 codes for a movement protein and ORF2a codes for a polyprotein 2a (Protease-VPg-P) (Fig. 1.6). Polyprotein 2a is organised in two different ways among sobemoviruses, namely Protease-VPg-P3 and Protease-VPg-P10-P8, although the P domain has not yet been identified in all someboviruses. In some viruses, a ribosomal frameshift occurs upstream of ORF2b/3 to produce polyprotein 2ab (protease-VPg-RdRp). ORF3 is expressed from a sub-genomic RNA and codes for the viral coat protein (Satheshkumar *et al.*, 2004; Nair and Savithri, 2010). The processing of the polyprotein is mediated by the N-terminal protease domain coded by ORF2a (Fig. 1.6) (Gayathri *et al.*, 2006).

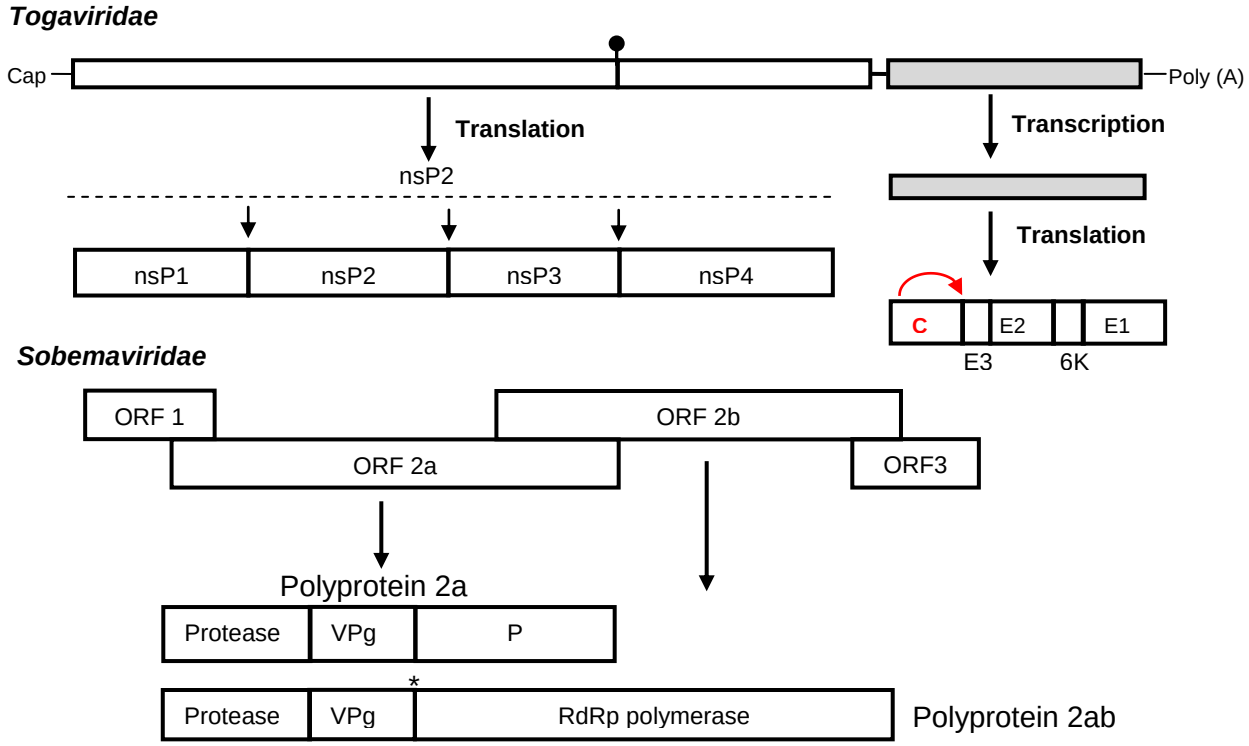


Figure 1.6 Genome organisations of family *Togaviridae* (Alphaviruses) and *Sobemaviridae* (Sesbania mosaic virus). For *Togaviridae*, the (•) indicates a termination codon that has to be read through in order to produce the nsP1234 precursor polyprotein. The non-structural proteins are processed by the protease activity of nsP2 as indicated by the small arrows (black). The structural proteins are expressed from subgenomic 26S mRNA as a polyprotein consisting of capsid protein (C) and envelope proteins (E1, E2, E3 and 6K). The capsid protein (C) cleaves itself from the polyprotein (indicated by the curved red arrow). In the *Sobemoviridae* genome, (*) corresponds to the region where the −1 ribosomal frameshift would occur. [Adapted from ten Dam *et al.* (1999) and Nair and Savithri (2010)]

The ribosomal frame shifting mechanism in ORFs of astroviruses is similar to that observed in retroviruses. The astrovirus genome is organised into three overlapping ORFs, ORF1a, ORF1b, and ORF2 (Fig. 1.7). ORF1a and ORF1b are linked by a

translational ribosomal frame shifting and they encode the viral protease and polymerase, respectively, whereas ORF2 encodes the capsid proteins (Mendez *et al.*, 2003). The processing of the non-structural polyproteins is not completely understood (Speroni *et al.*, 2009).

Astroviridae

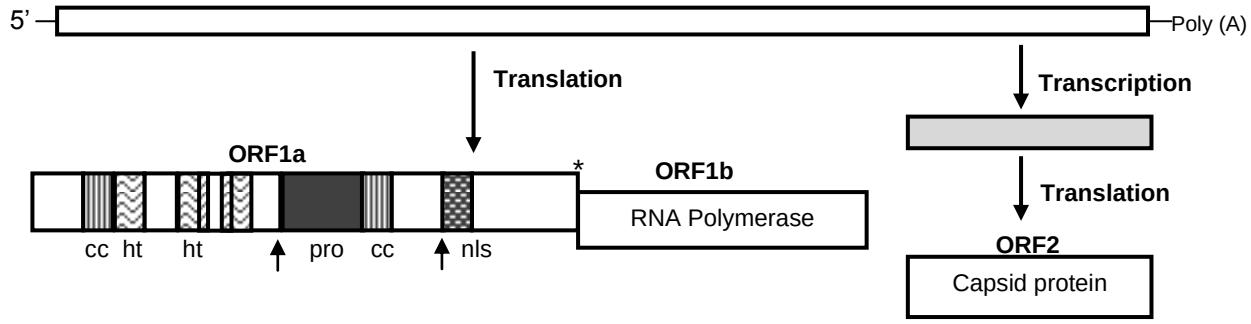


Figure 1.7 Genome organisation of family *Astroviridae*. Long open boxes represent the non-structural polyproteins 1a and 1b, and the small boxes inside represent the motifs for the predicted hydrophobic transmembrane (ht), protease (pro), nuclear localisation signal (nls) and coiled-coil structures (cc). The short arrows indicate the cleavage points for the protease. [Adapted from (Jiang *et al.*, 1993; Mendez *et al.*, 2003)].

Herpesviruses are characterised by long double stranded DNA genomes, ranging from about 125 to 240 kb, with unique regions bounded by direct or inverted repeats (Tyler *et al.,* 2011). The genomes can be grouped into classes based on the organisation of the repeats (Fig. 1.8). The class A genome, described in betaherpesviruses, consists of a unique sequence flanked by a direct repeat. Class B genomes have tandem repeat sequences at the termini and this organisation characterises most gammaherpesviruses. The class C genomes also found in gammaherpesviruses, have both terminal and internal repeats. Class D genomes, characteristic of alphaherpesviruses, contain two unique regions: long (U_L) and short (U_S), each flanked by inverted repeats (TR_S/IR_S). Class E of alphaherpesviruses is similar to class D, except that U_L and U_S sequences can invert resulting in isomers ($TR_L/IR_L$and TR_S/IR_S). Lastly, the Class F genome, also found in betaherpesviruses lacks the types of inverted and direct repeats found in other herpesvirus genomes (Davison, 2007).

An ORF within the U or U_L region encodes a serine protease that is expressed as part of a precursor protein that consists of an N-terminal protease domain, linker domain and a structural scaffold protein. During capsid maturation, the precursor protein undergoes processing at two sites, the release (R) site to release the N-terminal protease domain and the maturation (M) site to release a 25 amino acid peptide that joins the linker-scaffold protein to the capsid shell (Fig. 1.8) (Ertl *et al.,* 2000).

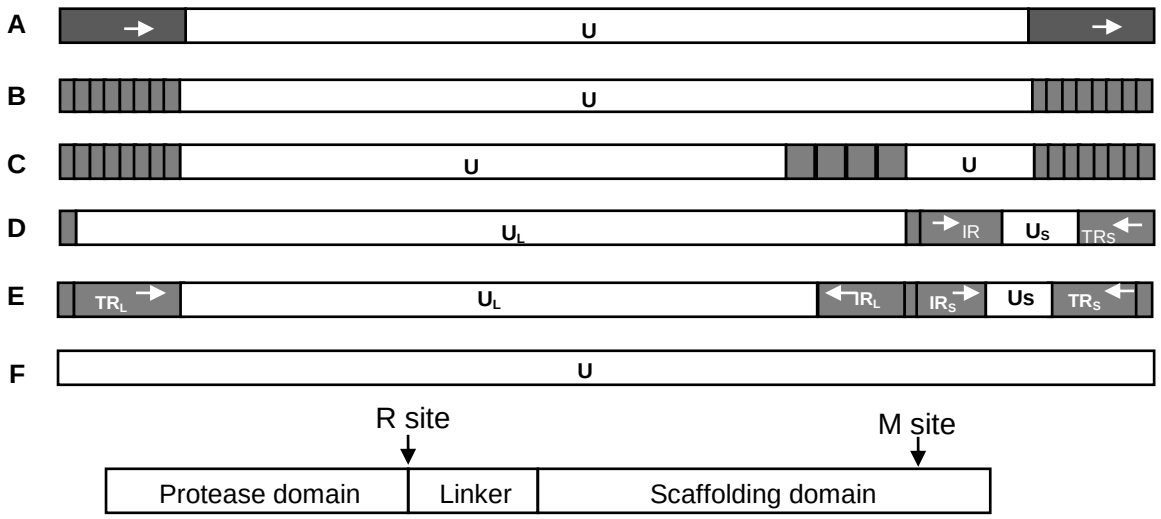


Figure 1.8 Genomic organisation of the *Herpesviridae* family. A to F indicate the different classes of genome structures (not to scale). Unique (U) and repeat regions are shown as white and grey rectangles, respectively. Classes D and E genomes contain two unique regions, long (U_L) and short (U_S), each flanked by inverted repeats (TR_L/IR_L and TR_S/IR_S). The orientations of repeats are shown by white arrows. The precursor protein is autocatalytically cleaved at the release (R) and maturation (M) sites by the N-terminal serine protease domain. Adapted from Gibson (2001) and Davison (2007).

The *Orthomyxoviridae* sense ssRNA genome comprises eight segments (Fig. 1.9) (Tsai and Chen, 2011). Segments 1 to 3 encode RNA polymerase complex subunits, PB2, PB1 and PA, respectively. Segments 4 to 6 encode surface glycoprotein hemagglutinin (HA), nucleoprotein (NP) and surface glycoprotein neuraminidase (NA), respectively. Segment 7 encodes two matrix proteins: M1, which is the major component of the virus and M2, an ion channel protein. Lastly, segment 8 encodes two non-structural proteins, NS1 and a nuclear export protein NEP (also known as NS2) (Tsai and Chen, 2011). The RNA polymerase subunit PA has a serine protease domain (Hara *et al.*, 2001a; Yuan *et al.*, 2009).

The *Birnaviridae* genome consists of a bisegmented dsRNA, where segment A contains two overlapping ORFs: a long ORF encoding the polyprotein and a short ORF (Fig. 1.10). The position of the short ORF is genus specific. The BSNV and TV-1 have a 71 and 106 amino acid long polypeptides, respectively, located between pVP2 and VP4 sequence indicated as X. However, these X polypetides do not show any sequence homology (Da Costa *et al.*, 2003; Nobiron *et al.*, 2008). The polyprotein is processed and cleaved at the pVP2-VP4 and the VP4-VP3 junctions by VP4 protease. Secondary processing of pVP2 generates mature VP2 and small structural peptides (Da Costa *et al.*, 2003).

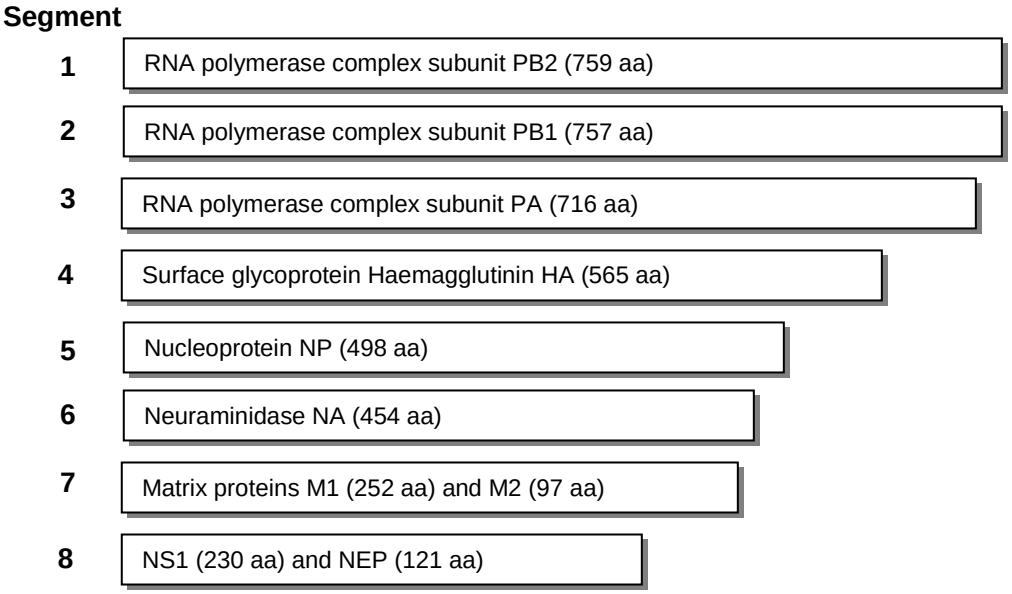


Figure 1.9 The eight genomic RNA segments of the influenza virus belonging to the *Orthomyxoviridae* family. These eight segments encode the viral RNA polymerase complex subunits PB2, PB1, and PA, haemagglutinin HA, nucleoprotein NP, surface glycoprotein neuraminidase NA, matrix proteins, non-structural protein NS1 and a nuclear export protein NEP. The number of amino acid residues for each protein is indicated in brackets (Data obtained from (Tsai and Chen, 2011).

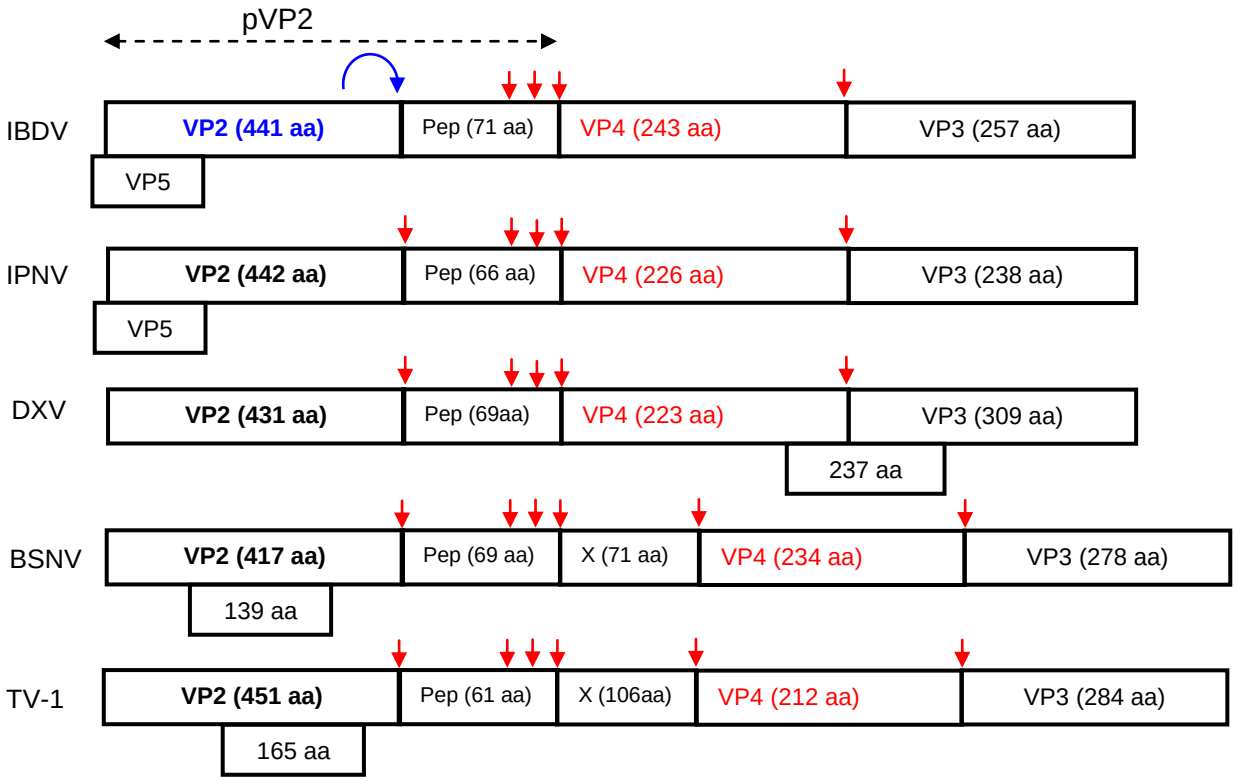


Figure 1.10 Gene rearrangement of genome segment A of the five birnaviruses. The cleavage sites of the protease are indicated by vertical lines, and their locations are given by the P_1-P'_1 amino acid number. The locations and sizes of the small ORFs coding for VP5 are indicated below the polyprotein coding ORFs. Numbers in parentheses at the 5' ends indicate the first nucleotide involved in the initiation codon (Nobiron *et al.*, 2008).

1.3.2 Viral proteases

Proteases also termed proteinases or peptidases catalyse the hydrolysis of peptide bonds and they are classified based on their mode of action. The hydrolysis of the peptide bond can occur from the N- or C-terminus (exopeptidase) or internally (endopeptidase). Based on their catalytic mechanisms, proteolytic enzymes are also classified into the following seven classes: aspartic, cysteine, glutamic, metallo, serine, threonine and asparagine (Rawlings *et al.*, 2004). According to the MEROPS peptidase database (http://merops.sanger.ac.uk) (Rawlings *et al.*, 2004; Rawlings *et al.*, 2011).The mechanism for peptide bond cleavage involves a nucleophilic attack of an amino acid residue (asparagine, cysteine, serine and threonine proteases) or a water molecule (aspartic, metallo and glutamic proteases) on the carbonyl group of the peptide bond (Rawlings *et al.*, 2008).

Proteases are further grouped into families and clans. A family groups protease sequences according to the alignment score of their catalytic domain or the whole sequence, whereas clans comprise a group of families with the same order of catalytic residues and show the same three-dimensional fold of the catalytic domain (Rawlings and Barrett, 1993). So far, only five protease classes have been identified in viruses, namely aspartic, asparagine, cysteine, metallo- and serine proteases (Table 1).

Aspartic proteases have only been found in viruses from the families *Retroviridae* and *Caulimoviridae.* These aspartic proteases are active as homodimers and the dimerisation results in the formation of the active site cleft containing two catalytic aspartic acid residues, one contributed by each monomer (Torruella *et al.*, 1989; Wlodawer and Gustchina, 2000; Ingr *et al.*, 2003). There are ten clans of cysteine proteases (CA, CD, CE, CF, CH, CL, CM, CN, CO and C- consisting of proteases not assigned to a particular clan) according to the MEROPS peptidase database. So far, only five clans have been recognised in viruses, namely CA, CE, CM, CN and C-. Clan CA has a catalytic triad consisting of Cys, His and Asp found in papain. Clans CE and CM have a catalytic triad, with Glu as the third active site residues. The catalytic residues of clans CN and C- are Cys and His which form a catalytic dyad (Table 1).

The best studied classes of proteases are serine proteases (Hedstrom, 2002), the most abundant being the chymotrypsin-like proteases of family S1 of clan PA (formerly SA) (Polgar, 2005). The active site of these serine proteases contains a Ser-His-Asp catalytic triad with Ser acting as the nucleophile. There are cysteine proteases that have a

chymotrypsin-like fold, and as a result, they are assigned to a separate clan PA. These chymotrypsin-like cysteine proteases are found in viruses from the *Nidovirales* order.

Metalloproteases can be divided into two subtypes depending on the number of metal ions required for catalysis. In some of the metalloproteases, only one metal ion is required, whereas others may require two metal ions that act co-catalytically. The known metal ligands in metalloproteases are His, Glu, Asp or Lys residues (Byrd *et al.*, 2004). Metalloproteases have fifteen clans: MA, MC, MD, ME, MJ, MK, ML, MM, MO and MP with one catalytic metal ion; while MF, MG, MH, MN and MQ have two metal ions. In addition, clan M- contains metalloprotease families not yet assigned to a clan (Auld, 2004; Manefeld, 2007). Metalloproteases are encoded by viruses from family *Poxyviridae* and are included in clan ME (Hedengren-Olcott *et al.*, 2004).

Capsid proteins from several viruses including those from members of the families *Nodaviridae, Tetraviridae* and *Picornaviridae* autocatalytically cleave themselves during the late stages of virus assembly (Zlotnick *et al.*, 1994; Hindiyeh *et al.*, 1999; Bothner *et al.*, 2005; Duquerroy *et al.*, 2009). The catalytic mechanism involves an acidic side chain from Asp that activates the catalytic Asn, resulting in a nucleophilic attack of its side chain amide nitrogen on its own carbonyl carbon, cleaving the peptide bond. The nucleophilic Asn is also the P_1 residue in the cleavage site (Duquerroy *et al.*, 2009). These capsid proteases that cleave the asparaginyl bonds are classified as asparagine endopeptidases (Reddy *et al.*, 2004).

Serine proteases are named after the catalytic serine residue. In clans PA, SB and SC the catalytic triad consists of Ser-Asp-His, where Ser is the nucleophile, His being a proton donor and Asp is required for the orientation of the His imidazolium ring (Polgar, 2004). There are some variations to the classical serine protease catalytic Ser-Asp-His triad. Proteases from clans SB (S53 family) and SS have a catalytic triad that contains Glu in place of His (Polgar, 2005). In clan SH, the Asp residue is replaced by another His. On the other hand, clans SE, SF, SJ, SK and SR consist of proteases that use a catalytic dyad with Ser as the nucleophile and a Lys residue as the proton donor. In clans SF and SM, the Lys residue in the dyad is replaced by His (Polgar, 2005). Viral serine proteases belong to clans PA, SH and SJ. The serine protease from the *Orthomyxoviridae* family has not yet been assigned to a particular clan (Table 1). The viral proteases for several families have not yet been characterised therefore it cannot be ruled out that other classes may also exist. IBDV encodes a serine protease (Lejal *et al.*, 2000) therefore the focus of the polyprotein processing will be on serine and chymotrypsin-like cysteine proteases.

Table 1.1 Viral families and their virus-encoded proteases

Class Clan/Family	Viral family and examples	Protease and /or catalytic residues[a]	References
Aspartic			
AA/A2, A9	*Retroviridae (Group VI: ssRNA RT)* HIV, murine leukemia virus spumaretrovirus, foamy virus	PR (Asp ,Asp)	Ingr *et al.* (2003); Wlodawer and Gustchina (2000)
AA/A3	*Caulimoviridae (Group VII: dsDNA-RT)* Cauliflower mosaic virus	pV (Asp, Asp)	Torruella *et al.* (1989)
Cysteine			
CA/C6	*Potyviridae (Group IV: (+) ssRNA)* Potato Y virus, tobacco etch virus	HC-Pro (Cys, His)	Anindya and Savithri, (2004)
CA/C16	*Coronaviridae (Group IV: (+) ssRNA)* Infectious bronchitis virus, SARS	PLpro (Cys, His, Asp)	Ratia *et al. (2006)*
CA/C28	*Picornaviridae (Group IV: (+) ssRNA)* Foot and mouth disease virus (FMDV)	L protease (Cys, His, Asp)	Guarne *et al*. (1998); Guarne et al. (2000)
CE/C5	*Adenoviridae (Group I: dsDNA)* Adenovirus	Adenovirus protease (Cys, His)	Ding *et al.* (1996); McGrath *et al.* (2003)
CE/C57	*Poxyviridae (Group I: dsDNA)* Vaccinia virus	Vaccinia I7L (Cys, His, Asp)	Hedengren-Olcott *et* *al.* (2004)
CM/C18	*Flaviviridae (Group IV: (+) ssRNA)* Hepatitis C virus	HCV NS2 (Cys, His, Glu)	Suzuki *et al.* (2007); Schregel *et al.* (2009)
CN/C9	*Togaviridae (Group IV: (+) ssRNA)* Sindbis virus, Semliki forest virus	nsP2 proteinase (Cys, His)	Rudenskaya and Pupov, (2008)
CA/C6	*Potyviridae (Group IV: (+) ssRNA)* Potato Y virus	HC-Pro (Cys, His)	Guo *et al.* (2011)
C-/ C7, C8	*Hypoviridae (Group III: dsRNA)* Chestnut blight fungus virus p29	P29 (Cys, His)	Choi *et al.* (1991)
C-/C21	*Tymoviridae (Group IV: (+) ssRNA)* Turnip yellow mosaic virus	Papain-like protease (Cys, His)	Bransom *et al.* (1996)
C-/C23	*Flexiviridae (Group IV: (+) ssRNA)* Blueberry scorch carlavirus	Papain-like protease (Cys, His)	Lawrence *et al.* (1995)
C-/C27	*Togaviridae (Group IV: (+) ssRNA)* Rubella virus, semliki forest virus	nsP2 proteinase (Cys, His)	Golubtsov *et al*. (2006)
C-/C32, C33	*Arteriviridae (Group IV: (+) ssRNA)* Equine arteritis virus	PL1 and PL2 (Cys, His)	den Boon *et al.* (1995)
C-/C36	*Genus Benyvirus (No family)* Beet necrotic yellow vein virus	Papain-like protease (Cys, His)	Hehn *et al.* (1997)
C-/C42	*Closteroviridae (+) ssRNA* Beet yellows virus	Papain-like protease (Cys, His)	Zinovkin *et al.* (2003)
C-/C53	*Flaviviridae (Group IV: (+) ssRNA)* Pestivirus	Npro (Cys, His)	Rumenapf *et al.* (1998)
PA/C3	*Picornaviridae (Group IV: (+) ssRNA)* Enterovirus, rhinovirus, cardiovirus, aphthovirus, poliovirus, FMDV	3Cpro (Cys, His, Asp / Glu) 2A protease (Cys, His, Glu)	Birtley *et al.* (2005); Sousa *et al.* (2006)
PA/C4	*Potyviridae (Group IV: (+) ssRNA)* Potato Y virus, tobacco etch virus	N1a (Cys, His, Asp)	Maia *et al.* (1996)
PA/C24, C37	*Caliciviridae (Group IV: (+) ssRNA)* Rabbit haemorrhagic disease virus Norwalk virus, hepatitis E virus	3CL protease (Cys, His, Glu or Asp)	Wirblich *et al.* (1995); Zeitler *et al*. (2006); Oka *et al.* (2007)
PA/C30	*Coronaviridae (Group IV: (+) ssRNA)* Infectious bronchitis virus, porcine transmissible gastoenteris virus	3CL protease (Cys, His, Glu)	Lim *et al.* (2000); Anand *et al.* (2002)

Table 1.1 Continued

Class Clan/Family	Viral family and examples	Protease and /or catalytic residues[a]	References
Metalloprotease			
ME/M44	*Poxyviridae (Group I: dsDNA)* Smallpox virus, vaccinia virus	Vaccinia G1L (His, His, Glu)	Hedengren-Olcott *et al.* (2004)
Asparagine			
NA/N1	*Nodaviridae (Group IV: (+) ssRNA)*	Nodavirus capsid protein (Asp, Asn)	Zlotnick *et al.* (1994)
NA/N2	*Tetraviridae (Group IV: (+) ssRNA)*	Tetravirus capsid protein (Asp, Asn)	Bothner *et al.* (2005)
NA/N8	*Picornaviridae (Group IV: (+) ssRNA)* Poliovirus	Poliovirus capsid protein	Hindiyeh *et al.* (1999)
NE/N5	*Picobirnaviridae (Group III: dsRNA)*	Capsid protein Asn, Asp/Glu)	Duquerroy *et al.* (2009)
Serine			
PA/S1	*Astroviridae (Group IV: (+) ssRNA)* Astrovirus	(Ser, Asp, His)	Speroni *et al.* (2009)
PA/S3	*Togaviridae (Group IV: (+) ssRNA)* Rubella virus, Sindbis virus Semliki forest virus	Capsid protease (Ser, Asp, His)	ten Dam *et al.* (1999)
	Flaviviridae (Group IV: (+) ssRNA)		Murthy *et al.* (1999)
PA/S7	Flavivirus (Yellow fever virus)	NS3 protease	
PA/S29	Hepacivirus (Hepatitis C virus)	(Ser, Asp, His)	
PA/S31	Pestivirus (Swine fever virus)		
PA/S32	*Arteriviridae (Group IV: (+) ssRNA)* Equine Arteritis virus	Nsp4 (Ser, Asp, His)	Barrette-Ng *et al.* (2002)
PA/S30	*Potyviridae (Group IV: (+) ssRNA)* Potato Y virus	P1 protease (Ser, Asp, His)	Verchot and Carrington (1995)
PA/S39	*Sobemoviridae (Group IV:(+) ssRNA)* Sesbania mosaic virus	(Ser, Asp, His)	Gayathri *et al.* (2006)
PA/S75	*Nidovirales – No family* *White bream virus*	3CL (Ser, Asp, His)	Schutze *et al.* (2006)
	Herpesviridae (Group I: dsDNA)	Herpesvirus protease	Buisson *et al.* (2002)
SH/S21	Cytomegalovirus	(Ser, His, His)	
SH/S26	Herpes simplex virus, Epstein-Barr virus,		
	Birnaviridae (Group III: dsRNA)	*VP4*	Lejal *et al.* (2000);
SJ/S50	IBDV, IPNV, DXV, BSNV,	(Ser, Lys)	Petit *et al.* (2000); Da
SJ/S69	Tellina-1 virus		Costa *et al.* (2003); Nobiron *et al.* (2008)
S-S62	*Orthomyxoviridae (Group V: (-) ssRNA)* - Influenza virus	RNA polymerase PA (Ser, Asp, Glu)	Hara *et al.* (2001a); Yuan *et al.* (2009)

[a] The catalytic residues are not given in protein sequence order

1.3.2.1 Viral serine proteases

Classical serine proteases are characterised by a Ser-Asp-His catalytic triad. Serine proteases catalyse peptide bond hydrolysis in two steps (Fig. 1.11). The first step is the acylation reaction whereby the nucleophilic Ser attacks the substrate scissile peptide bond, forming first a tetrahedral intermediate and then a covalent acyl-enzyme complex

with the release of the C-terminal portion of the substrate. During the second (deacylation) step, a water molecule attacks the acyl-enzyme, leading to the formation of a second tetrahedral intermediate followed by the release of the N-terminal fragment of the substrate. The His imidazole acts as a general base in accepting a proton to activate Ser as a nucleophile, and subsequently acts as a general acid, donating the proton to the nitrogen of the peptide leaving group (Hedstrom, 2002; Polgar, 2004). Proteases with this classical triad have been found in several viral families for example potyvirus P1, togavirus capsid protein, flavivirus NS3, arterivirus nsp4, astrovirus and sobemovirus proteases (Table 1.1). These proteases have been crystallised and their 3D structures determined, except for potyvirus P1.

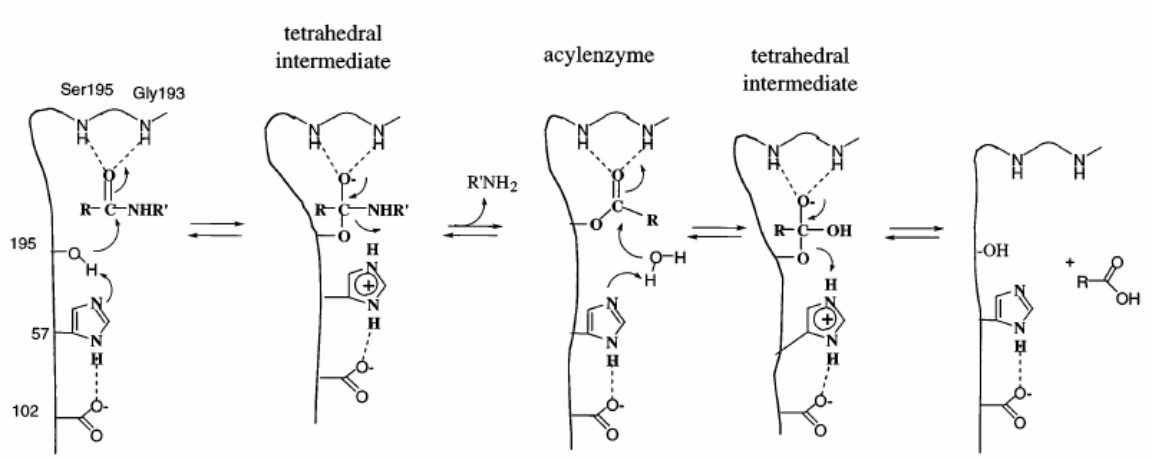


Figure 1.11 The catalytic mechanism of chymotrypsin-like serine protease. In the acylation reaction, Ser^{195} attacks the carbonyl of the peptide substrate, assisted by His^{57} acting as a general base, to yield a tetrahedral intermediate. The resulting positive charge formed on His^{57} is stabilised by the negative charge on the Asp^{102} through a hydrogen bond. When the double bond between the carbon and oxygen in the peptide bond reforms, the bond between the carbon and the nitrogen in the peptide bond is broken. The leaving part of the substrate is stabilised by the formation of a bond to a hydrogen atom from His^{57} to yield acylenzyme intermediate. During the deacylation reaction water attacks the acylenzyme, assisted by His^{57}, yielding a second tetrahedral intermediate. When the double bond between the oxygen and carbon atom in the remaining part of the substrate is reformed, the bond between carbon and the oxygen of Ser^{195} is broken. The hydroxyl group on Ser^{195} is restored by transfer of an H^+ ion from His^{57} (Hedstrom, 2002).

The potyvirus P1 protease cleaves itself from the polyprotein at the P1/HC-Pro junction (Fig. 1.4A) (Urcuqui-Inchima *et al.,* 2001). Its Ser-Asp-His catalytic triad and the Gly-X-Ser-Gly motif around the active site Ser are strictly conserved. In the P1 protease, the conserved His and Asp residues are present N-terminal to the active site Ser residue. The Tyr-Ser and Phe-Ser have been identified as cleavage site residues in several viruses. The P_1 residues are always Phe or Tyr whereas the P_1' residue is a conserved Ser. The P_4 residue is occupied by highly hydrophobic residues such Ile, Leu, Val or Met and the P_2 aromatic residue (His, Phe, Trp, Tyr) is also conserved (Table 1.2) [(Adams *et al.*, 2005); using the nomenclature of Schechter and Berger (1967) shown in Fig. 1.12].

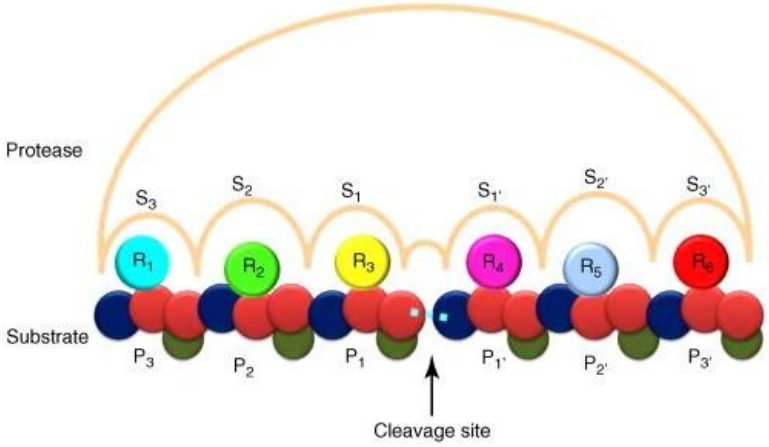


Figure 1.12 Schematic representation of the active site of proteases annotated according to Schechter and Berger notation (Smooker *et al.*, 2010) . The amino acid residues of the substrate bind in enzyme subsites of the active site. The N-terminal amino acid residues of the scissile bond are numbered P_3, P_2, and P_1, and the C-terminal residues are numbered P_1', P_2', P_3', etc. The subsites on the protease that complement the substrate residues are numbered S_3, S_2, S_1, S_1', S_2', S_3', etc.

Table 1.2 Comparison of cleavage site preferences for serine and chymotrypsin-like cysteine proteases

Protease	Class	Catalytic residues	P_2	P_1 ↓	P_1'
Potyvirus P1	Serine	His^{214} Asp^{223} Ser^{256}		F / Y	S
Togavirus capsid protease	Serine	His^{145} Asp^{167} Ser^{219}		W	S
Sobemovirus protease	Serine	His^{181} Asp^{216} Ser^{284}		E	T / S
Astrovirus protease	Serine	His^{461} Asp^{489} Ser^{551}		E / D	
Arterivirus nsp4	Serine	His^{29} Asp^{65} $Ser^{118,120}$		E	G / S
Hepacivirus NS3pro	Serine	His^{1083} Asp^{1107} Ser^{1165}		C, T	S / A
Flavivirus NS3pro	Serine	His^{51} Asp^{75} Ser^{135}		K / R	G / S / A
Pestivirus NS3pro	Serine	His^{1658} Asp^{1686} Ser^{1752}		L	S / A / N
Picornavirus 3Cpro	Cysteine	$His^{44,46}$ Asp^{84} $Cys^{163, 172}$		Q / E	G, S
Picornavirus 3Cpro -HRV	Cysteine	His^{40} Glu^{71} Cys^{147}		Q	G
Picornavirus 2A pro	Cysteine	His^{18} Asp^{35} Cys^{106}	T	X / Y	G
Coronavirus 3CL	Cysteine	Cys^{51} His^{148} Glu^{163}	L	E	G / A
Calicivirus 3CL	Cysteine	His^{27},^{30} $Cys^{104, 139}$ (Asp^{44})		G / A	E
Herpesvirus protease	Serine	$Ser^{120, 132}$ $His^{52, 63}$ His^{157}		A	S
Birnavirus VP4 - IBDV	Serine	Ser^{652} Lys^{692}		A	A
Birnavirus VP4 - BSNV	Serine	Ser^{692} Lys^{729}		A	A / S
Birnavirus VP4- IPNV	Serine	Ser^{633} Lys^{674}		A	S
Birnavirus VP4 - DXV	Serine	Ser^{627} Lys^{670}		S	A
Birnavirus VP4 - TV-1	Serine	Ser^{738} Lys^{777}		A	S

The three-dimensional (3-D) structures of the viral serine proteases reveal the chymotrypsin-like barrel domains (Fig. 1.13). About 140 residues of the alphavirus capsid protein shows a chymotrypsin-like fold that is smaller and more compact than chymotrypsin, with much shorter connecting loops and some of the β-strands are truncated (Fig. 1.13A) (Skoging and Liljestrom, 1998). The protease is expected to have a broad substrate specificity similar to chymotrypsin based on the size of the substrate binding pocket (Tong, 2002). The alphavirus capsid protein (264 amino acid residues) is located at the N-terminus of the structural polyprotein where it cleaves itself from the polyprotein at the Trp-Ser junction (Fig. 1.6). After its release from the polyprotein, it functions as a capsid protein. After the autoproteolytic cleavage, the free carboxylic group of Trp^{267} interacts with the catalytic triad leading to self inactivation (Skoging and Liljestrom, 1998; Nicola *et al.*, 1999).

The overall fold of the sobemovirus protease shows similar features to the chymotrypsin fold, consisting of two β barrels (domains I and II) connected by a long inter-domain loop (Fig. 1.13B). The active site and the substrate-binding cleft occur in between the two domains. Within domain II, there is a disulfide bond that stabilises the walls of the S1 specificity pocket. Its structure was found to be closer to that of cellular serine proteases than to that of other viral proteases (Gayanthri *et al.,* 2006). The protease cleaves the ORF2 polyprotein at three different positions (Fig. 1.6) with Glu-Thr or Glu-Ser sequences (Table 1.2) (Satheshkumar *et al.*, 2004).

The catalytic properties and the 3-D structure of astrovirus protease (Fig. 1.7) were reported by Speroni *et al.* (2009). The structure also adopts a chymotrypsin fold (Fig. 1.13C). However, the catalytic Asp is orientated away from His and does not form the His-Asp hydrogen bonds that characterise the classical triad of serine proteases. The β-ribbons are relatively shorter thus resulting in a highly solvent exposed active site. The protease crystallised as a hexamer or a trimer of dimers. The protease was shown to have a preference for Glu or Asp at P_1 (Table 1.2) and this was confirmed by the fact that the S_1 site is characterised by a positive electrostatic potential (Speroni *et al.*, 2009).

Arterivirus nsp4 protease adopts a structure that is slightly different from other chymotrypsin-like proteases. In addition to the double β-barrel structure, it has a C-terminal domain not found in most chymotrypsin-like proteases (Fig. 1.13D) (Barrette-Ng *et al.*, 2002). This C-terminal domain is not required for proteolytic activity (van Aken *et al.*, 2006a). It also adopts the smallest chymotrypsin-like fold with a canonical catalytic triad. The oxyanion hole was shown to adopt two conformations, either a collapsed inactive or

the standard active conformations, which seems to be a novel way of regulating proteolytic activity (Barrette-Ng *et al.*, 2002; Tian *et al.*, 2009). It shows a preference for Glu in the P_1 position and a small amino acid residue (Gly, Ser or Ala) at the P_1' position (Table 1.2), a specificity typical of picornavirus 3C cysteine proteases (van Aken *et al.*, 2006b).

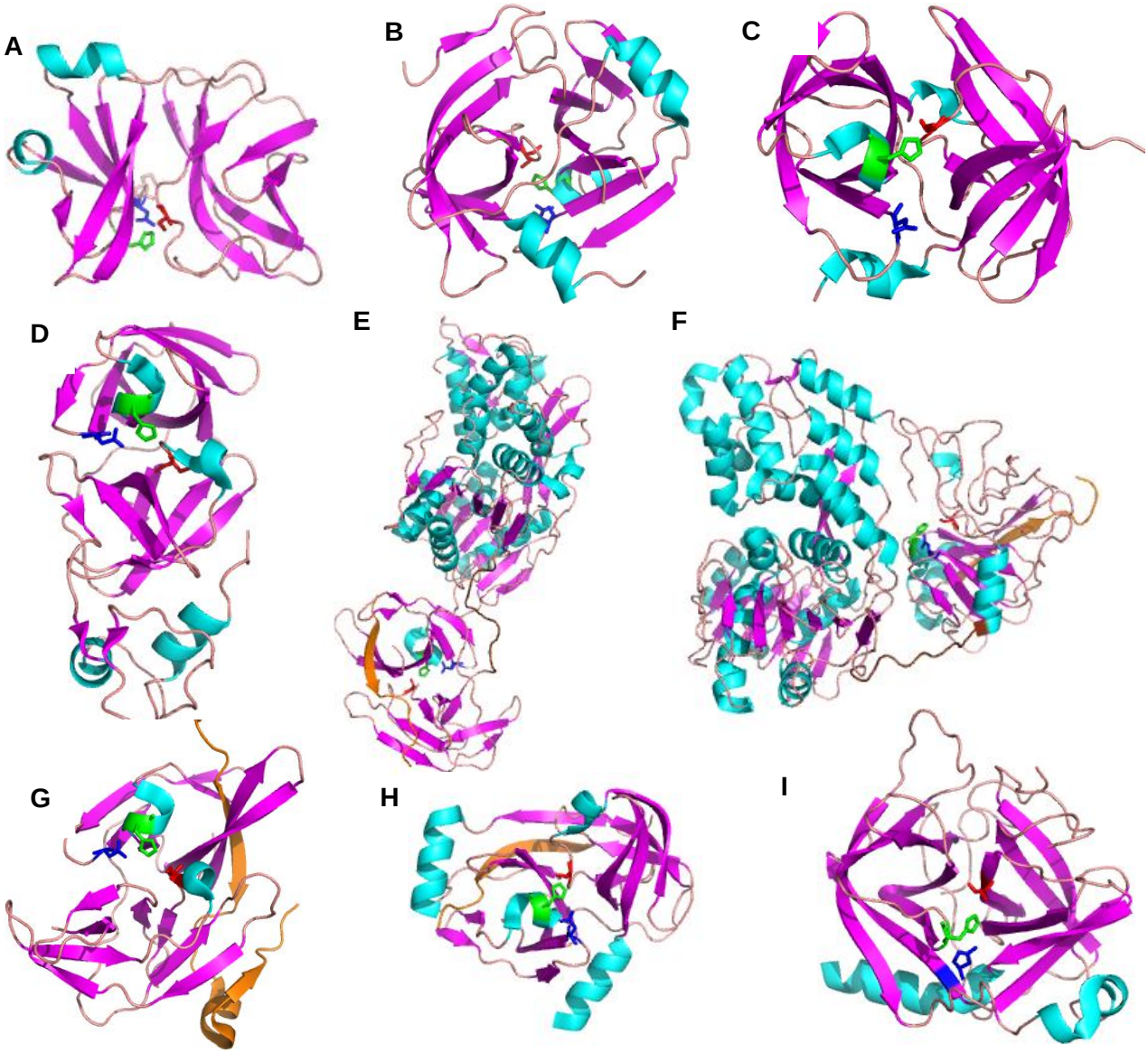


Figure 1.13 Crystal structures of chymotrypsin-like serine proteases. A) Togavirus capsid protein [PDB code 1VCQ, (Choi *et al.*, 1997)]; **B)** Sobemovirus protease [PDB code 1ZYO, (Gayathri *et al.*, 2006)]; **C)** Astrovirus protease [PDB code 2W5E, (Speroni *et al.*, 2009)]; **D)** Arterivirus nsp4 protease [PDB code 1MBM, (Barrette-Ng *et al.*, 2002)]; **E)** Flavivirus Dengue virus NS3 protein [PDB code 2VBC, (Luo *et al.*, 2008)]; **F)** Hepacivirus NS3 protein [PDB code 1CU1, (Yao *et al.*, 1999)]; **G)** Hepacivirus NS3pro with NS4A cofactor [PDB code 3P8N, (Lemke *et al.*, 2011)]; **H)** Flavivirus NS3pro with NS2B cofactor [PDB code 2FOM, (Erbel *et al.*, 2006)] and **I)** Bovine alpha chymotrypsin [PDB code 5CHA, (Blevins and Tulinsky, 1985)]. Secondary structure elements are coloured in cyan (α-helix), magenta (β-strand) and salmon (loops). For the NS3 protein, the cofactors (NS2Band NS4A) are orange, and the interdomain linkers are brown. The catalytic residues Ser (red), His (green) and Asp (blue) are shown as stick models.

The *Flaviviridae* NS3 is a multifunctional protein with an N-terminal serine protease domain (NS3pro) and a C-terminal region containing RNA helicase (NS3hel) and nucleoside triphosphate activities (Fig. 1.13E and F) (Tian *et al.*, 2009). The viral replication cycle requires that all these functions are regulated in a coordinated fashion. There is a flexible hinge between NS3pro and NS3hel which allows for twisting and closure movements (van Aken *et al.*, 2006b). The twisting movements reorientate NS3pro to expose its active site during the processing while the closure movements reorientate the subdomain of NS3hel for the unwinding of RNA duplexes (Bouffard *et al.*, 1995). The three dimensional structure of the 185 amino acid NS3pro domain forms six stranded anti-parallel β-barrels that are packed like those in chymotrypsin-like serine proteases (Fig. 1.13G and H) (Rosales-Leon *et al.*, 2007). However, several loops found in other chymotrypsin-like serine proteases that form the active site are missing from NS3pro (Rosales-Leon *et al.*, 2007).

NS3pro requires the presence of polypeptide cofactors for efficient proteolytic processing. NS3 protease domain with or without a cofactor has an overall similar topology (Murthy *et al.*, 1999). Flavivirus NS3pro requires a C-terminal 40 amino acid stretch of NS2B protein for activity (Fig. 1.13G). The NS2B functions as a cofactor and promotes the folding and functional activity of NS3pro (Tsumoto *et al.*, 2002). NS2B consists mostly of hydrophilic domains and three hydrophobic domains. During substrate binding, the hydrophilic region of the NS2B cofactor refolds to form an important component of the catalytic site. This refolding of NS2B serves to enhance the catalytic action of the protease (Erbel *et al.*, 2006). The N-terminus of the NS2B cofactor forms a β-strand that is inserted into the β-barrel of the protease (orange in Fig. 1.13G), hindering the exposure of the NS2B hydrophobic residues from the solvent thus providing stabilisation of the catalytic domain (Rosales-Leon *et al.*, 2007). The C-terminal end of NS2B cofactor is disordered in the absence of the substrate or inhibitor but upon binding of the substrate it becomes ordered (Lescar *et al.*, 2008).

NS4A serves as a cofactor to the hepacivirus NS3pro and is relatively shorter than NS2B (Fig. 1.13H). Only the N-terminal β-barrel of the protease domain in HCV is involved in binding the NS4A activation peptide cofactor (Rosales-Leon *et al.*, 2007). The interactions of the NS4A with NS3pro induce conformational changes in NS3pro that involve a structural reorganisation of both the N-terminal domain and the catalytic site (Lescar *et al.*, 2008). The reorientation of the catalytic triad ensures that the Asp residue is close enough to stabilise the His residue, after the imidazole ring of the His residue deprotonated the OH group of the Ser nucleophile. This indicates that in the absence of NS4A, NS3pro will

be inactive since the His^{57} is located too far away to be able to deprotonate the OH of the nucleophilic Ser (Barbato *et al.*, 1999; Rosales-Leon *et al.*, 2007). However, it was shown that the NS3-4A junction is cleaved in *cis* by NS3pro alone, whereas cleavages at NS4A-4B, NS4B-5A and NS5A-5B are mediated in *trans* by NS3pro with the NS4A cofactor (Barbato *et al.*, 1999). NS2B and NS4A also serve as membrane anchors attaching the cytoplasmic NS3 proteolytic processing to the membrane where the polyprotein resides (Rosales-Leon *et al.*, 2007).

The hepacivirus and flavivirus NS3pro are not only different in their cofactors but also in the cleavage site preferences. Flavivirus NS2B/NS3 recognises sites with dibasic residues, Lys or Arg at the P_1 and P_2 positions. Dibasic residues have to be followed by a short chain of residues such as Gly, Ala and Ser at the P_1' position to attain full activity of the protease (Bessaud *et al.*, 2006). The cleavage site specificities for pestivirus NS3pro consists of Leu at P_1 and either Ser, Ala or Asn at P_1' (Table 1.2). For hepacivirus NS3pro, the specificities include an acidic residue at P_6, Cys at P_1 for *trans* cleavage or Thr at P_1 for *cis* cleavage and either Ser or Ala at P_1' (Table 1.2) (Ryan *et al.*, 1998).

Taken together, these serine proteases not only share the catalytic triad residues with chymotrypsin-like proteases, but also adopt their structural fold. It is interesting that besides their similarities in the structural fold and the catalytic residues, their substrate specificities are mostly different.

1.3.2.2 Chymotrypsin-like cysteine proteases

The main proteases of picornaviruses, coronaviruses and caliciviruses, form a unique class of cysteine proteases with a three-dimensional structure showing an active site geometry and structural fold similar to chymotrypsin-like serine proteases. They have a conserved Cys-His-Asp/Glu catalytic triad at the active site, similar to the Ser-Asp-His triad found in serine proteases (Ryan *et al.*, 1998; Bessaud *et al.*, 2006) (Fig. 1.14). The shallow active site cavity uses the thiol group of Cys as the nucleophile and are thus classified as cysteine proteases. His and Asp/Glu residues are used as acid/base catalysts (Bolognesi *et al.*, 2001).

The substitution of the catalytic nucleophilic Cys of FMDV and poliovirus 3C protease with a Ser residue did not result in a completely inactive enzyme. In contrast, the replacement with Gly or Ala resulted in a completely inactive enzyme (Hedstrom, 2002). The incomplete inactivation suggests that the Ser hydroxyl group can substitute the Cys sulfhydryl group as a nucleophile (Grubman *et al.*, 1995; Marcotte *et al.*, 2007). In serine

proteases the His residue primes the Ser side chain for the nucleophilic attack on the scissile bond, while the Asp stabilises the resulting positive charge on the His (Birtley *et al.*, 2005). The Ser-His-Asp/Glu configuration is demonstrated in Fig. 1.14A using several picornaviral 3C protease structures. The Asp of the HAV 3C is directed away from the active site and has been proposed not to play a role in catalysis, the function is possibly replaced by the negatively charged Tyr^{147} residue. The HRV and poliovirus 3C catalytic Asp position is occupied by a Glu residue which is rare in serine proteases, however, it is unclear why some 3C proteases have Glu as the third member of the triad (Grubman *et al.*, 1995).

The 3C-like cysteine proteases have a chymotrypsin fold except for the coronavirus protease which contains an extra domain with an α-helical fold (Fig. 1.14B-F) (Yang *et al.*, 2003; Oka *et al.*, 2005; Zeitler *et al.*, 2006; Sweeney *et al.*, 2007). Even though potyvirus N1a protease has the same catalytic triad, the structure has not yet been crystallised therefore it is not known whether it has the same chymotrypsin fold. The second picornaviral 2A protease also possesses a chymotrypsin-like fold that is related to smaller serine proteases (Petersen *et al.*, 1999) (Fig. 1.14G).

The 3C-like proteases have well-defined substrate specificities. Picornavirus 3C protease specifically cleaves between a Gln-Gly sequence or less commonly between Gln-Ser, Gln-Ala, Glu-Ser and Glu-Gly sequences with different cleavage efficiencies (Table 1.2). For efficient cleavage, residues at P_4 and P_2 must have a small hydrophobic and a large aromatic residue, respectively (Zoll *et al.*, 1998; Birtley *et al.*, 2005). The 2A consensus recognition site consists of Thr-X-Gly where the Leu in the P_4 position and a Thr at P_2 are important cleavage determinants. The substrate preference for 2A is more stringent in poliovirus protease than in other picornaviral proteases in that it only cleaves at Tyr-Gly bonds (Blognesi *et al.*, 2001; Curry *et al.*, 2007). For coronaviruses, all known cleavage sites have a hydrophobic residue (mainly Leu) at P_2, Glu at P_1 and a small aliphatic residue at P_1' (Table 1.2) (Anand *et al.*, 2003). Cleavage preferences for 3C-like protease in caliciviruses are poorly understood but so far, the protease has been shown to cleave sites with Gly, Ala, Ser or Thr at P_1' and Glu at the P_1 position (Table 1.2) (Sosnovtseva *et al.*, 1999).

In addition to their role in polyprotein processing, picornavirus proteases have other functions or even target cellular substrates. They have a unique ability to specifically bind viral RNA thus playing a role in the regulation of protein and RNA synthesis during the viral life cycle (Oka *et al.*, 2007). Both 3C and 2A proteases have been shown to cleave the eukaryotic translation initiation factor 4G (eIF4G) (Li *et al.*, 2001; Strong and Belsham,

2004), amongst other cellular substrates. The 3C proteases from the other two families, *Coronaviridae* and *Caliciviridae* seem not to have any additional functions other than polyprotein processing.

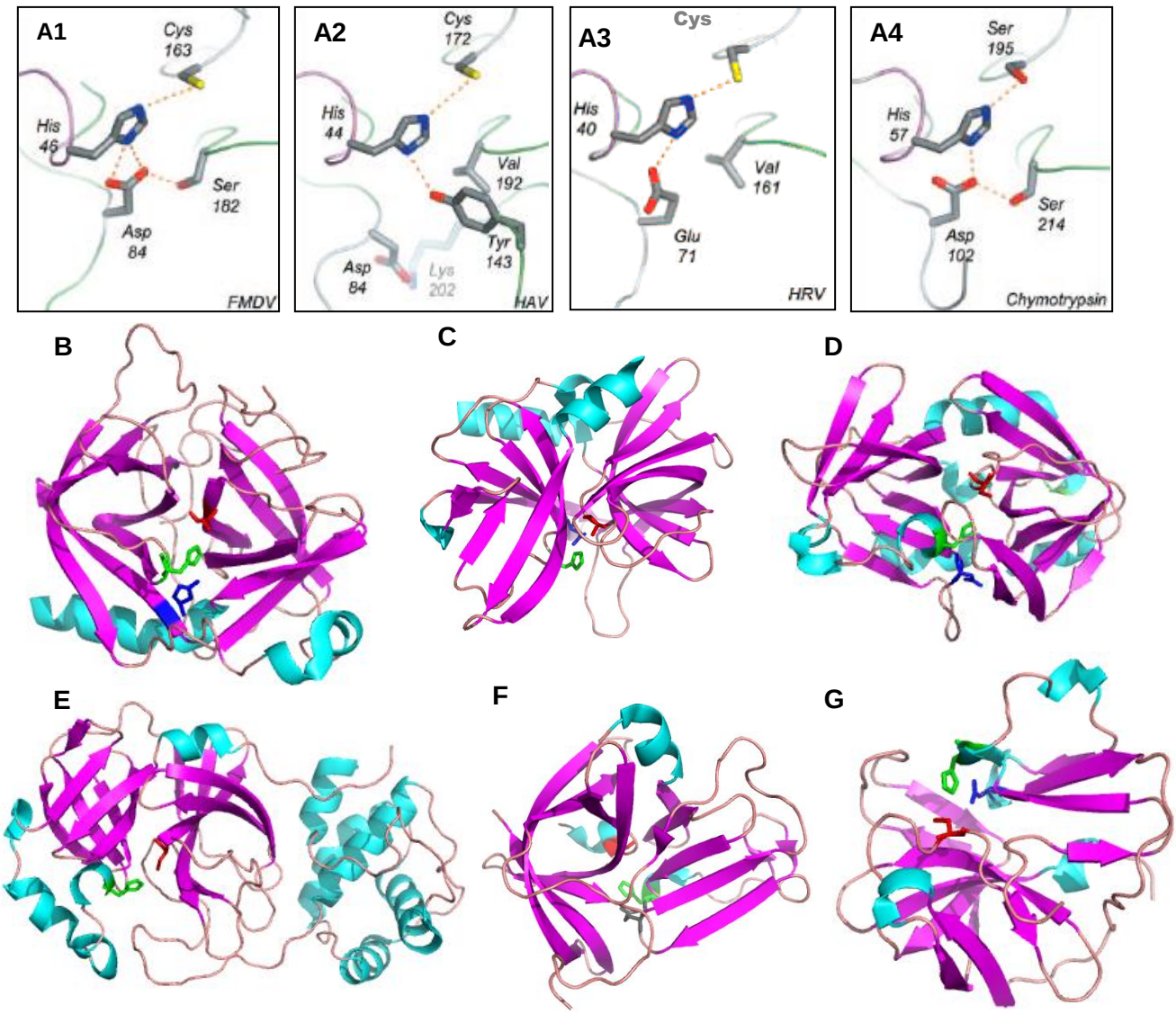


Figure 1.14 The active site and the structural fold of chymotrypsin-like cysteine proteases. **A)** Comparison of the active sites of picornavirus 3C proteases with chymotrypsin. The Cys/His/Asp catalytic triad of FMDV (A1) and HAV (A2) as well as Cys-His-Glu of HRV (A3) are compared to the serine protease Ser-Asp-His catalytic triad of chymotrypsin (A4) (Birtley *et al.*, 2005). **B)** The structure of bovine alpha chymotrypsin [PDB code 5CHA, (Blevins and Tulinsky, 1985)]. **C-F)** Chymotrypsin-like cysteine protease structures of C) FMDV of the *Picornaviridae* family [PDB code 2J92, (Sweeney *et al.*, 2007)]; D) HAV of the *Picornaviridae* family [PDB code 1HAV, (Bergmann *et al.*, 1997)]; E) SARS coronavirus [PDB code 1UK2, (Yang *et al.*, 2003)]; **F)** Norwalk virus of the family *Caliciviridae* [PDB code 2FQY, (Zeitler *et al.*, 2006)]; **G)** Structure of picornaviral HRV 2A protease [PDB code 2HRV, (Petersen *et al.*, 1999)]. Secondary structure elements are coloured in cyan (α-helix), magenta (β-strand) and salmon (loops). The catalytic residues Cys (red), His (green) and Asp (blue) or Glu (gray) are shown as stick models.

1.3.2.3 Novel serine proteases

Three viral families were identified to have novel serine proteases, namely *Orthomyxoviridae*, *Herpesviridae and Birnaviridae.* The influenza virus multifunctional RNA polymerase was reported to show a novel serine protease activity, with a Ser-Asp-Glu catalytic triad (Hara *et al.*, 2001a; Hara *et al.*, 2001b; Yuan *et al.*, 2009). The structure of the protease domain has an α/β architecture with five mixed β-strands forming a twisted plane surrounded by seven α-helices (Fig. 1.15A). Its substrate specificities have not yet been elucidated. The herpesvirus protease contains a central seven anti-parallel β-barrel surrounded by eight helices. The protease exists in a monomer-dimer equilibrium in solution (Fig. 1.15B-C). The active site containing Ser-His-His is a very shallow pocket located on top of two strands of the central β-barrel and it is complete in each monomer (Buisson *et al.*, 2002).

The comparison of the herpesvirus protease catalytic triads with the classical triad indicates that the second His^{63} is within hydrogen-bonding distance to the His^{157} (chymotrypsin numbering), thus it plays the same role that Asp plays in the stabilisation of the charge on the His (Fig. 1.15D) (Qiu *et al.*, 1997).

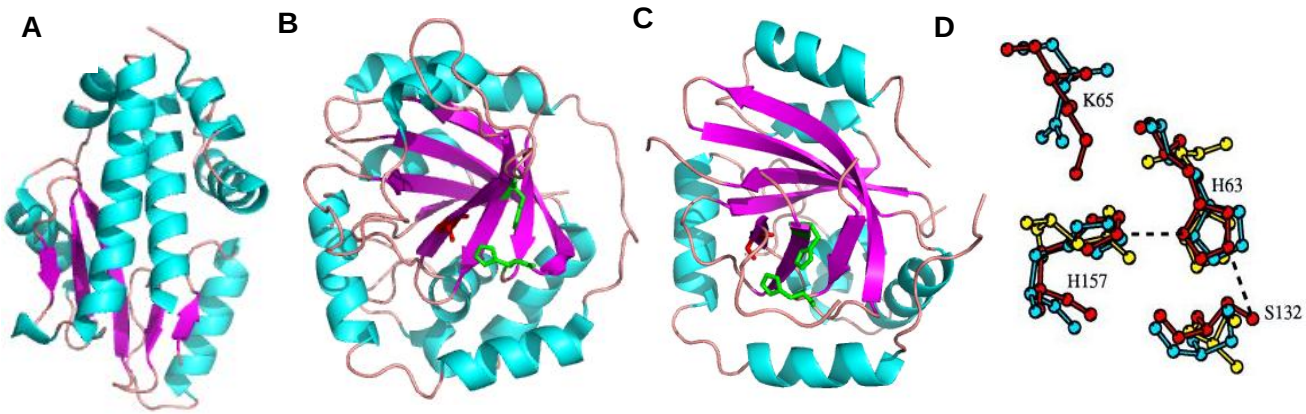


Figure 1.15 Crystal structures of novel viral serine proteases. A) Influenza RNA polymerase protease domain [PDB code 3EBJ, (Yuan *et al.*, 2009)]; **B)** Gamma-herpesvirus Epstein–Barr virus protease [PDB code 1O6E, (Buisson *et al.*, 2002)]; **C)** Beta-herpesvirus cytomegalovirus [PDB code 1LAY, (Qiu *et al.*, 1996)]. Secondary structure elements are coloured in cyan (α-helix), magenta (β-strand) and salmon (loops). The catalytic residues His (green) and Ser (red) are shown as stick models. **D)** The comparison of His-Ser-His catalytic residues of varicella-zoster virus (red); cytomegalovirus (blue) and chymotrypsin (yellow) (Qiu *et al.*, 1997).

1.3.2.4 Birnavirus VP4 protease

The birnavirus VP4 has been classified as a novel type of serine protease that utilises a Ser/Lys catalytic dyad instead of the canonical Ser-His-Asp triad (Lejal *et al.*, 2000; Petit

et al., 2000; Da Costa *et al.*, 2003; Nobiron *et al.*, 2008). The conserved Ser and Lys residues carry polarised side chains essential for catalysis therefore the substitution of either Ser or Lys abolishes polyprotein processing (Lejal *et al.*, 2000). The ε-amino group of Lys plays the role of a general base and activates the hydroxyl of Ser for catalysis; while the deprotonated OH group of Ser acts as a nucleophile which attacks the substrate (Nobiron *et al.*, 2008). However, the replacement of Ser with Cys in IBDV VP4 did not abolish processing but slowed it down, suggesting that Cys can replace the catalytic Ser efficiently (Birghan *et al.*, 2000).

This catalytic dyad is characteristic of the photosystem II D1 protein processing protease (CtpA) (Yamamoto *et al.*, 2001), bacterial Lon protease (Botos *et al.*, 2004), Lex A, leader protease of *E. coli* (Paetzel and Dalbey, 1997) and penicillin binding proteins (Sauvage *et al.*, 2008). Birnavirus VP4 forms a unique branch of the Lon family of serine proteases, that lack the ATPase domain, and are therefore called non-canonical Lon proteases. The approximately 80 amino acid C-terminal Lon^{pro} domain is conserved across the Lon/VP4 proteases (Birghan *et al.*, 2000). The structure of the Lon^{pro} domain has a unique fold consisting of six α helices and ten β strands (Fig. 1.16A). The six domains form a ring shaped hexamer that mimic the oligomerisation state of a holoenzyme (Botos *et al.*, 2004).

The 3-D structures of BSNV, IPNV and TV-1 VP4 proteases have been determined (Feldman *et al.*, 2006; Lee *et al.*, 2007; Chung and Paetzel, 2011a; Chung and Paetzel, 2011b) and although there is low sequence identity between birnavirus VP4 proteases, the three-dimensional structures adopt a similar topology (Fig. 1.16B-D). VP4 protease has an α/β fold, with the N-terminal anti-parallel β-sheet found in domain I and the C-terminal α/β structure forming domain II.

BSNV VP4 structure (Fig. 1.16B) contains thirteen β-strands, three α-helices and one 3_{10}-helix (shown in green). Eight long β-strands assemble in an anti-parallel fashion, three short β-strands form a parallel β-sheet and the remaining two strands (shown in lemon) form a short β-hairpin. While the anti-parallel β-sheet forms domain I, the α-helices pack against the parallel three-stranded β-sheet and the β-hairpin in domain II. The nucleophilic Ser^{692} is located at the N-terminal end of α1 and the general base Lys^{729} is within α2 (Fig. 1.16B) (Feldman *et al.*, 2006). IPNV VP4 structure also consists of thirteen β-strands but four α-helices and two 3_{10}-helices (shown in green). Seven long β-strands assemble in an anti-parallel fashion to provide the substrate-binding groove. While the anti-parallel β-sheet forms domain I, the three α-helices pack against a parallel three-stranded β-sheet and a short β-hairpin (shown in lemon) formed by the remaining two strands in domain II.

The active site nucleophilic Ser^{633} is located at the N-terminal end of α2 and Lys^{674} general base is part of α3 (Fig. 1.16C) (Lee *et al.*, 2007). IPNV VP4 protease consists of amino acid residues 509 to 734 while the structure shown in Fig. 1.16B represents of residues 518 to 716 (Lee *et al.*, 2007).

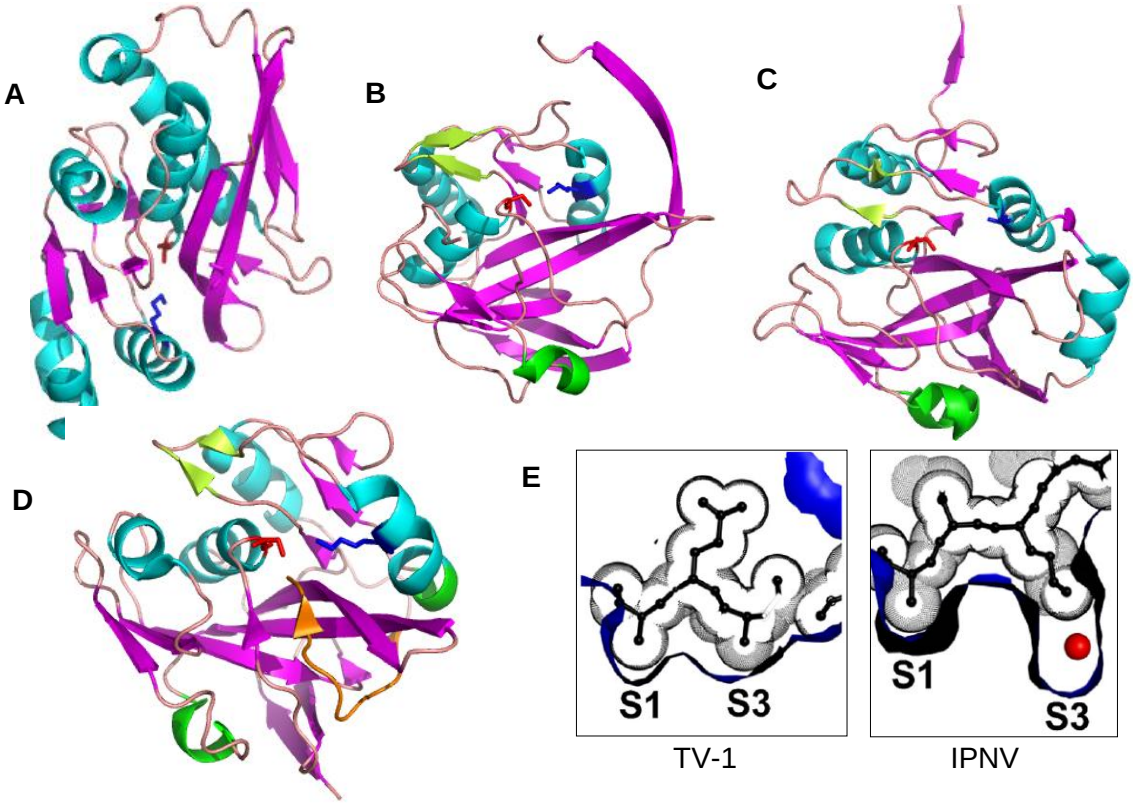


Figure 1.16 The ribbon structures of the Lon protease domain and the VP4 proteases. A) Lon protease [PDB code 1RRE, (Botos *et al.*, 2004)]; **B)** BSNV VP4 protease [PDB code 2GEF, (Feldman et al., 2006)]. **C)** IPNV VP4 protease [PDB code 2PNL, (Lee et al., 2007)]; **D)** TV-1 VP4 protease [PDB code 3PO6, (Chung and Paetzel, 2011b)]. The last five residues (Val^{826}–Ala^{830}) at the C terminus of VP4 are shown in orange with the last residue Ala^{830} forming an intramolecular (cis) acyl-enzyme with the nucleophile, Ser^{738} (Chung and Paetzel, 2011b). Secondary structure elements are coloured in cyan (α-helix), green (3_{10} helix), magenta (β-strand), lemon (short hairpin) and salmon (loops). The catalytic residues Ser (red) and Lys (blue) are shown in stick format. For IPNV, the Lys is mutated to Ala. **E)** Comparison of TV-1 and IPNV VP4 substrate specificity binding pockets. The amino acid residues within the bound substrate are shown as ball-and-stick (black) and the binding pockets are shown as blue surfaces (Chung and Paetzel, 2011b).

TV-1 VP4 protease structure comprises sixteen β-strands, four α-helices, and two 3_{10}-helices (shown in green). The anti-parallel form β-sheet that forms the substrate-binding groove is made up of eleven β-Strands. Three α-helices and a short β-hairpin formed by the remaining two strands (shown in green) surround a parallel three stranded β-sheet. The nucleophilic Ser^{738} is located at the N-terminal end of α2 and the Lys^{777} general base is found within α3 (Fig. 1.16D) (Chung and Paetzel, 2011b). The structure of IPNV VP4 shows the intermolecular (*trans*) acyl-enzyme intermediate where the C-terminus is

extended away from the structure and bound to an adjacent active site in a neighbouring VP4 molecule (Lee *et al.*, 2007). In contrast, the TV-1 VP4 structure shows an intramolecular (*cis*) acyl-enzyme intermediate between the nucleophilic Ser^{738} and the last residue Ala^{830} (Chung and Paetzel, 2011b).

The VP4 cleavage sites have been identified in all the birnaviruses. The BSNV VP4 cleaves multiple Pro-X-Ala↓(Ala/Ser) motifs except for the VP4-VP3 junction which has the sequence Cys-Gly-Ala↓Ala. The cleavage sites of IBDV, IPNV, DXV and TV-1 are defined by the (Thr/Ala)-X-Ala↓Ala, (Ser/Thr)-X-Ala↓(Ser/Ala)-Gly, (Ala/Gly)-X-Ser↓Ala and Ala-X-Ala↓Ala motifs, respectively (Da Costa *et al.*, 2003; Nobiron *et al.*, 2008). However the P_1' position in TV-1 can accommodate Ser, Gly or Trp (Nobiron *et al.*, 2008). For IBDV VP4, the P_1 and the P_1' residues are conserved as alanines, whereas the P_3 Ala can be substituted for Thr (Lejal *et al.*, 2000).

The molecular surfaces of the S_1 and S_3 binding pockets were analysed and compared between TV-1 and IPNV VP4 protease (Fig. 1.16E) (Chung and Paetzel, 2011b). Cleavage sites for all birnavirus VP4 proteases have Ala at P_1 but residues at P_3 position are variable. TV-1 cleavage sites have Ala at both P_1 and P_3 positions, whereas IPNV VP4 cleavage sites have Ser, Gln or Thr at P_3 position. The TV-1 pockets are shallow and hydrophobic thus complementary to the hydrophobic methyl group of Ala at P_1 and P_3 position. In IPNV VP4, the S_1 pocket is also shallow and hydrophobic, but the S_3 pocket is deep and hydrophilic allowing it to accommodate the variable residues at P_3 position (Chung and Paetzel, 2011b). In comparison to the Lon proteases, VP4 does not form a hexameric complex and has a larger and a wider substrate binding pocket (Feldman *et al.*, 2006).

The first VP4 structure to be solved was from BSNV (Feldman *et al.*, 2006), followed by IPNV (Lee et al., 2007) and TV-1(Chung and Paetzel, 2011b). The structures of IBDV and DXV VP4 proteases have not yet been crystallised but it is likely that they will adopt a similar α/β fold. This α/β fold seems to be conserved in this class of proteases. The structural insights provided by protein crystallisation are important for understanding polyprotein processing as well as drug design.

In conclusion, polyprotein processing is a major strategy used by viruses to release the structural and non-structural proteins. The viral protease domains present within the polyprotein play a major role in this maturation process. Therefore, viral proteases are attractive targets for anti-viral drug development (Feldman *et al.*, 2006). Protease inhibitor

based anti-viral drugs, saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir and atazanavir have been approved for the treatment of HIV infections (Lee *et al.*, 2007).

1.4 Objectives of the study

The overall aim of the current study was to obtain an understanding of the IBDV polyprotein processing by VP4 protease. The present study characterised polyprotein processing using anti-VP4 peptide antibodies and studied the two forms of VP4, namely the embedded form that exists as an integral part of the polyprotein and the mature form, which is released after the processing.

Several constructs of IBDV polyprotein and VP4 corresponding to embedded and mature forms of VP4 were designed and cloned into the eukaryotic *Pichia pastoris* and *Escherichia coli* expression systems. The virus isolate used in the present study was from bursae harvested from chickens killed in an IBDV outbreak in the KwaZulu-Natal province of South Africa. The full-length polyprotein clone was sequenced and compared to polyprotein sequences of other IBDV strains to establish the genetic relatedness of the isolate (Chapter 2). Anti-VP4 peptide antibodies were produced in chickens and used to characterise the expression of the polyprotein and VP4 constructs. The mature form of VP4 was purified for enzymatic characterisation using fluorogenic substrates, natural polypeptide substrate and gelatine-SDS-PAGE. Detection of native VP4 in infected bursa homogenates by western blotting using anti-VP4 peptide antibodies and determination of proteolytic activity by gelatine-SDS-PAGE analysis are also described (Chapter 3).

Chapter 4 describes the mutagenesis of the VP4 catalytic dyad in all the constructs and the expression of the mutants. In addition, a P377L mutation was also unexpectently introduced into the polyprotein in a different mutagenesis reaction and its effect on polyprotein processing was hence analysed. After the identification of two Arg-Ile dipeptides in the polyprotein sequence on the N-terminus of two cleavage sites, other VP4 constructs were designed commencing with either the first or the second dipeptide. Following their expression, the cleavage products were analysed to determine if these dipeptides may have any influence in the polyprotein processing (Chapter 5). The findings of the present study are summarised and discussed in chapter 6 (General discussion).

CHAPTER 2
Amplification, cloning and sequencing of different polyprotein and VP4 coding regions

2.1 Introduction

Infectious bursal disease virus (IBDV) is classified in the Avibirnavirus genus of the family *Birnaviridae,* consisting of a bisegmented dsRNA contained within a non-enveloped icosahedral capsid. The larger RNA segment (3.2 kb) contains two partially overlapping ORFs, A1 and A2. The short ORF A1 encodes a non-structural protein VP5. The long ORF A2 encodes a 114 kDa precursor polyprotein (NH_2-pVP2-VP4-VP3-COOH), which is processed into two capsid proteins pVP2 and VP3 by the viral protease VP4. The smaller RNA segment B (2.9 kb) encodes VP1, a 90 kDa RNA-dependent RNA polymerase (Fernandez-Arias *et al.*, 1998).

Polyprotein processing is a major strategy used by viruses to release structural and non-structural proteins. Despite the critical role that polyprotein processing plays in capsid assembly and virus replication, the events surrounding cleavages by the VP4 protease are not fully understood. The dibasic residues ^{452}Arg-Arg^{453} and ^{722}Lys-Arg^{723} were initially proposed by Hudson *et al.* (1986) as the polyprotein processing sites based on the sizes of proteins observed after processing and the fact that dibasic residues are frequent targets for proteolytic cleavage by serine proteases. In contrast, site-directed mutagenesis identified vicinal dialanines, ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} as processing sites at the pVP2-VP4 and VP4-VP3 junctions, respectively (Sanchez and Rodriguez, 1999). It has also been reported that the processing is an autocatalytic reaction and is co-translationally mediated by VP4 protease (Hudson *et al.*, 1986; Jagadish *et al.*, 1988).

From these observations, VP4 therefore exists in two forms, the embedded and mature form. The embedded form exists as an integral part of the polyprotein and it autocatalytically cleaves the ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} peptide bonds to produce the mature form (Ala^{513}-Ala^{755}). It was previously suggested that the mature form mediates the processing of pVP2 C-terminus to release structural peptides based on the presence of dialanine peptides (Sanchez and Rodriguez, 1999). In an attempt to determine whether the embedded and mature VP4 form behave in a similar fashion, different constructs coding for the IBDV polyprotein as well as the VP4 coding region were prepared and cloned.

There are two distinct serotypes of IBDV, I and II. Serotype I viruses are pathogenic to chickens causing Gumboro disease, while serotype II viruses isolated from turkeys are apathogenic to chickens (van den Berg *et al.*, 1991). Based on the degree of virulence, serotype I strains can further be divided into four major groups: classical, antigenic variant, very virulent (vvIBDV) and attenuated strains (Lim *et al.*, 1999). The classic strain was first identified and isolated from chickens in the Gumboro area of Delaware in USA in the early 1960's. In the late 1980s, highly cytolytic variant strains were isolated from flocks in the USA (Sreedevi *et al.*, 2007). The vvIBDV strains which cause severe clinical symptoms and high mortality, emerged in Europe in the early 1990s (Chettle *et al.*, 1989; van den Berg *et al.*, 1991).

Infectious bursal disease virus occurs in most parts of the world and is widespread in young commercial chickens. The virulent strain reached South Africa in 1989 and very quickly spread throughout the country affecting both the broiler and egg layer components of the poultry industry, despite the common use of commercially available IBD vaccines (Horner, 1994). However, the genetic identity of the strain was not determined and to date, sequence analysis involving South African isolates has been limited to the hypervariable region of VP2 (Jackwood *et al.*, 2008).

The present study thus presents the primer design, amplification, cloning and sequencing of different constructs of the polyprotein coding region. The sequence of the full-length polyprotein coding region was used for the determination of the genetic relatedness of the isolate with other IBDV strains, whereby the deduced amino acid sequence was compared to polyprotein sequences of classic, variant, very virulent and attenuated strains as well as non-pathogenic serotype II viruses. There are a number of residues in the polyprotein that are unique to the different pathotypes of IBDV and as a result they may be used for virus identification (Xia *et al*., 2008). These residues, together with phylogenetic trees were used to determine the relatedness of the isolate to other known IBDV isolates.

2.2 Materials

Virus isolates: The IBDV strain used in this study was isolated from bursa samples collected by Dr Roger Horner, Allerton Regional Veterinary laboratory, Pietermaritzburg, South Africa, from young commercial broilers collected during a South African IBD outbreak in 1995. The name of the isolate SA-KZN95 is based on the country (South Africa), the province (KwaZulu-Natal) and the year (1995).

Molecular biology: The following molecular biology reagents were obtained from Fermentas (Vilnius, Lithuania): BgII [Nomenclature according to Roberts *et al.* (2003)], EcoRI, NotI, shrimp alkaline phosphatase (SAP), T4 DNA ligase, 10 mM dNTP mix, X-gal, MassRuler® DNA Ladder Mix; MassRuler® high range DNA ladder; FastRuler® middle range DNA ladder, GeneRuler 1 kb DNA ladder mix, Taq DNA polymerase, Long PCR enzyme mix, GeneJET™ plasmid miniprep kit, pTZ57R/T and TransformAid™ bacterial transformation kit. SuperScript™ III first-strand synthesis system was purchased from Invitrogen Life technologies (California, USA). The peqGOLD gel extraction kit was obtained from PEQLAB Biotechnologie (Erlangen, Germany) and DNA clean and concentration kit™ was obtained from Zymo Research. The pGEM-T® Easy cloning kit, ImProm-II reverse transcription system and AMV reverse transcriptase/Tfl DNA polymerase Access RT-PCR System were purchased from Promega *(*Madison, USA*)*. All the expression vectors were provided as glycerol stocks by Dr A. Boulangé (University of KwaZulu-Natal).

2.3 Methods

2.3.1 Isolation of IBDV particles from infected bursa samples and preparation for EM

IBDV virus particles were isolated from infected bursae as described (Fernandez-Arias *et al.*, 1998) with some modifications. Briefly, an equal volume of homogenisation buffer (0.02 M Tris-HCl buffer, pH 7.8) was added to the infected bursae and homogenised using a Potter S homogeniser (Braun Biotech International GmbH, Germany). After centrifugation (17000 x g, 15 min, 4°C), the supernatant was loaded on top of a 2 ml cushion of 40% (w/v) sucrose in homogenisation buffer and centrifuged (86000 x g, 2.5 h, 4°C). The pellet was resuspended in homogenisation buffer and layered over half its volume of a 1.37 g/cm^3 CsCl solution. The preparation was centrifuged (86000 x g, 6 h, 4°C) and the viral band was collected by aspiration, resuspended in homogenisation buffer and dialysed overnight against dialysis buffer (0.01 M Tris-HCl buffer, pH 7.8) for removal of CsCl. The dialysed suspension was layered again over half its volume of a 1.27 g/cm^3 CsCl solution and centrifuged (86000 x g, 6 h at 4^{o}C). The band containing IBDV particles was collected and attached to formvar-coated copper grids. These grids were placed on drops of the virus particle suspension, and incubated for 2 min. Adsorbed particles were negatively stained for 5 min with 2% (w/v) uranyl acetate and viewed with an EM 120 Brotwin transmission electron microscope.

2.3.2 Isolation of viral dsRNA

The dsRNA was isolated from infected bursal homogenates using the proteinase K method (Akin *et al.*, 1998). Briefly, the bursal homogenates were treated with proteinase K (100 µg/ml) and 1% (w/v) sodium dodecyl sulfate (SDS) for 2 h at 60°C before 0.2% (v/v) diethylpyrocarbonate (DEPC) was added and incubation was continued an additional 30 min. The homogenate was treated with 1 M potassium acetate (pH 5) at 4°C for 30 min. The mixture was centrifuged (20000 x g, 30 min, 4°C), total nucleic acids were precipitated overnight at -20°C with 3 volumes of absolute ethanol, followed by centrifugation (14 000 x g, 30 min, 4°C) and phenol-chloroform extraction. Two volumes of absolute ethanol were added to the top aqueous phase and incubated at -20°C for 2 h and centrifuged (14000 x g, 30 min, 4°C). The pellet was resuspended in 500 µl DEPC treated water. The viral dsRNA was purified from the total nucleic acids by a two-step differential lithium chloride (2 and 4 M, respectively) precipitation method (Diaz-Ruiz and Kaper, 1978). The viral dsRNA pellet was finally dissolved in DEPC-treated water and stored at -20°C. The integrity and purity of the viral RNA was analysed on a 1% (w/v) agarose gel (Section 2.3.3).

2.3.3 Agarose gel electrophoresis

Agarose [1% (w/v)] was dissolved in 1x TAE buffer (200 mM Tris-acetate buffer, pH 8, 5 mM Na_2-EDTA) by heating, cooled to ±55°C and ethidium bromide added to a final concentration of 0.5 µg/ml. Samples were mixed with the loading buffer [10 mM Tris-HCl buffer pH 8, 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene cyanol FF, 60 mM EDTA, 60% (v/v) glycerol] in a (5:1 sample-sample loading buffer ratio) before loading into the wells. The images were captured with a VersaDoc™ imaging system (BioRad, California, USA) under UV light.

2.3.4 Primer design for the polyprotein and VP4 constructs

Six different constructs, i.e. full-length polyprotein (Met^{1}-Glu^{1012}), truncated polyprotein (Ile^{227}-Trp^{891}), VP4-RA (Arg^{453}-Ala^{755}), VP4-RK (Arg^{453}-Lys^{722}), VP4-AW (Ala^{513}-Trp^{891}) and VP4-AA (Ala^{513}-Ala^{755}), representing the embedded and mature forms of VP4, were designed (Fig. 2.1).

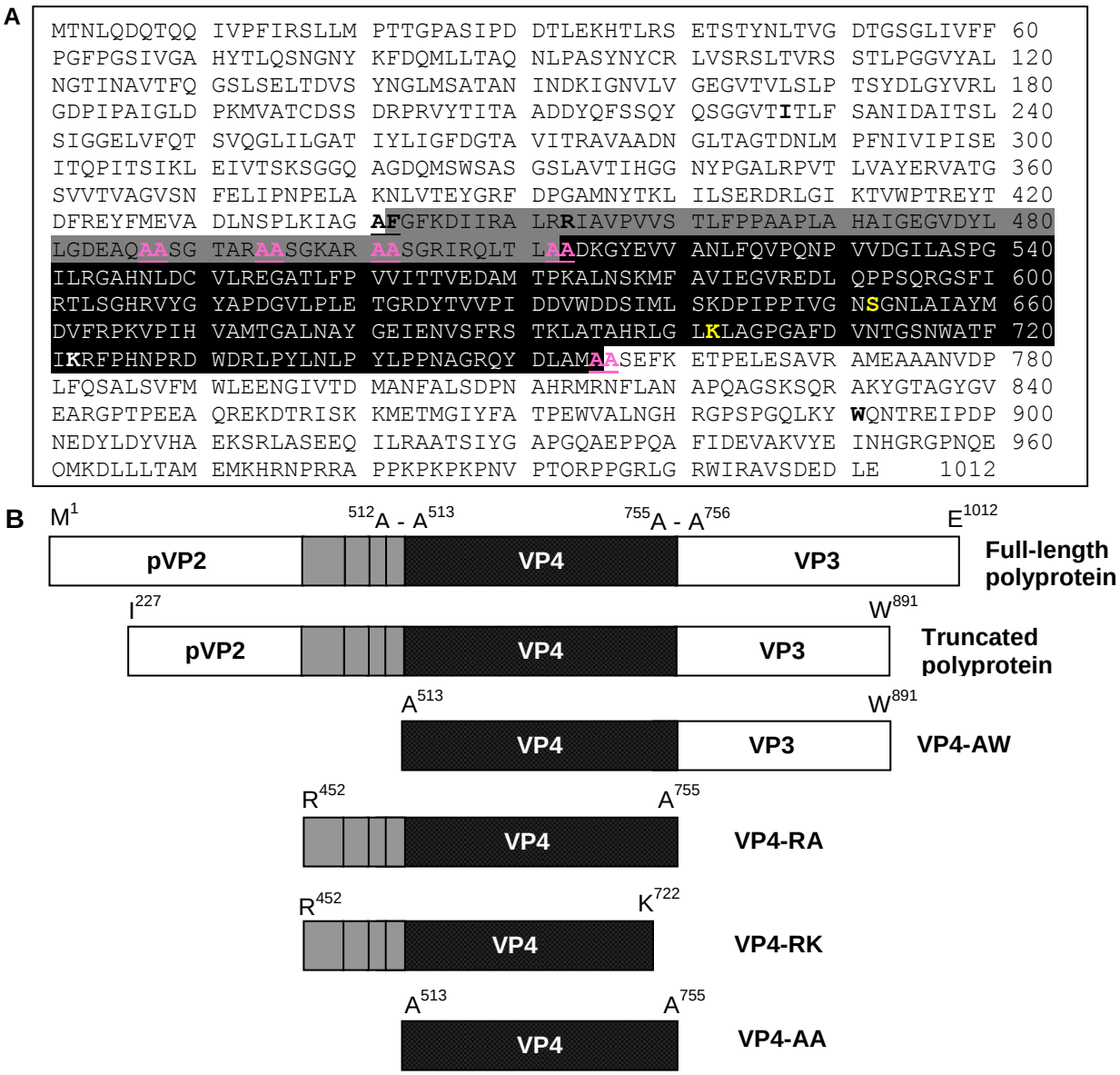


Figure 2.1 The IBDV polyprotein constructs used in the present study. A) Amino acid sequence of IBDV polyprotein (Accession number AF322444). The amino acid residue numbers are indicated on the right. The cleavage sites are in magenta and the catalytic residues are shown in yellow. The N- and C-terminal residues of the constructs are in bold. **B)** Schematic representation of the constructs. The constructs represent the embedded and mature forms of VP4.

There was no genetic information available for the isolates that were used in the present study, therefore primers were designed from the sequences available in the National Centre for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov) database. Sequence alignment of approximately 52 different segment A sequences in the database was carried out in order to design primers that would anneal to any possible IBDV sequence. IBDV strains used for the alignment are shown in Table 2.1 with their database accession numbers and geographic origin.

Table 2.1 Reference IBDV strains used for sequence analysis

Virulence	Strain	Accession No.	Origin	Virulence	Strain	Accession No.	Origin
Classic	002-73	X03993	Australia	**Attenuated**	B87	DQ906921	China
	52/70	D00869	Britain		CEF94	AF194428	Netherlands
	IM	AY029166	Germany		Cu-1	X16107	Germany
	STC	D00499	USA		CT	AJ310185	France
Variant	Gls	AY368653	USA		D78	AF499929	USA
	Variant E	AF133904	USA		GZ29112	AF051837	China
vvIBDVs	BD 3/99	AF362776	Bangladesh		Harbin	AF092171	China
	D6948	AF240686	Netherlands		HZ2	AF321054	China
	GHUT-12	DQ778035	Vietnam		JD1	AF321055	China
	Gx	AY444873	China		Lukert	AY918948	USA
	GZ/96	AY598356	China		NB	AY319768	China
	Harbin-1	A454945	China		P3009	AF109154	Taiwan
	HK46	AF092943	Hong Kong		TL2004	DQ088175	China
	Korean	AF508176	Korea		ZJ2000	AF321056	China
	ks	DQ927042	Israel	**ND**	02015.1	AJ879932	Venezuela
	OKYM	D49706	Japan		9109	AY462027	USA
	PO7	AY665672	Tunisia		CS-2-35	EF418033	USA
	SDH1	AY323952	Iran		GA-1	EF418034	USA
	SH/92	AF533670	Korea		H-30	EF418035	USA
	SH95	AY134874	China		HPR-2	EF418036	USA
	TASIK	AF322444	Indonesia		K310	AF165149	Korea
	TO9	AY099456	Nigeria		KK1	AF165150	Korea
	UK661	X92760	UK		KSH	AF165151	Korea
	UPM94/273	AF527039	Malaysia		KT1/99	AJ427340	India
	UPM97/61	AF247006	Malaysia		Soroa	AFA40705	Spain
Serotype II	23/82	AF362773	Germany				
	OH	M66722	USA				

ND means there was not sufficient information provided on the virulence

The sequence data was analysed with ClustalW in the BioEdit™ program (Hall, 1999) and Sequencher 4.7 (Gene Codes Corporation, USA). The sequences varied in length ranging from 3023 to 3245 bp. The resulting consensus sequence from the alignment was used

for primer design. When designing primers with restriction site extensions, it is essential to select enzymes that do not have restriction sites within the target sequence. Introducing restriction sites in the primers depended on how the gene would be ligated to the expression vectors. In the present study, the RT-PCR products were first cloned into a T-vector and sub-cloned into four different expression vectors, pET-28b, pET-32b, pGEX-4T-1 and pPIC9 (Appendix 1). Restriction sites unique to the polylinker of the expression vectors when compared to the IBDV polyprotein encoding sequences were identified by performing a restriction site search using Sequencher on all sequences (Table 2.2).

Table 2.2 Restriction sites in the polylinker region of expression vectors and IBDV segment A

Restriction Endonuclease[a]	Presence of site on the polylinker of the vector[b]				Number of sites in the segment A sequences[b]
	pET-28b	pET-32b	pPIC9	pGEX-4T-1	
AvaI		+			++++
AvaII			+		
BamHI	+	+		+	++
EagI	+	+			++
EcoRI	+	+	+	+	
EcoRV		+			
HindIII	+	+			++
NdeI	+				+
NheI	+				
NcoI	+	+			+++++
NotI	+	+	+	+	
SacI	+	+			+
SalI	+	+		+	+
SmaI				+	+++
SnaBI			+		
XhoI	+	+	+	+	+

[a] Nomenclature according to Roberts *et al.* (2003)

[b] Each + sign indicates a restriction site

Table 2.2 indicates that only EcoRI and NotI restriction sites were present in the polylinker region of the four expression vectors, but absent in the IBDV polyprotein coding region. This means that the target genes would be cloned into all the vectors with a single preparation of inserts. However, it is worth mentioning that some of the sequences had mutations that resulted in an EcoRI site (GAATTC) within the polyprotein coding sequence

(Table 2.3). These mutations could have resulted from the virus itself through the process of antigenic variation or been introduced by the reverse transcriptase during RT-PCR.

Table 2.3 Sequences showing the nucleotides that mutated to an EcoRI site and their positions

Accession number	Nucleotide residue[a]	Consensus sequence[b]	Amino acid number[a]	Position of the mutation
AY918948	565-570	GAATT**G**	143-144	VP2
AY029166	1752-1757	G**G**ATTC	535-536	VP4
D00499	1753-1758	G**G**ATTC	535-536	VP4
EF418036	1753-1759	G**G**ATTC	535-536	VP4
X03993	1642-1647	G**G**ATTC	535-536	VP4

[a] The numbers correspond to the numbering in the sequence of the strain
[b] The nucleotides in bold are those that have mutated and led to the EcoRI site GAATTC

The primers were designed with EcoRI/NotI restriction site extensions. Nested primers without restriction sites were also designed for reverse transcription. The sequences of all primers are listed in Table 2.4 and the forward and reverse primers used for each fragment as well as the expected sizes are summarised in Table 2.5. All primers were synthesised by Sylvean biotech (Pretoria, South Africa), except for the IBDA-F and IBDA-R primers which were synthesised by Sigma Proligo (France).

Table 2. 4 Description of primers used for segment A cDNA synthesis, amplification and sequencing of the polyprotein and VP4 coding regions

Primer	Application	Sequence (5'→ 3')[a]	Length
P1-18F	Reverse transcription	aggatacgatcggtctga	18
P3220-3241R	Reverse transcription	ggtgtccacacctgcgcgggt	21
P1820-R	Sequencing	cttaacaagcagtcgaggt	19
IBDA-F	PCR and sequencing	tc**gaattc**atgacaaacctgcaagatcaaacccaacag	38
IBDA-R	PCR and sequencing	cag**gcggccgc**ctcaaggtcctcatcagagacggtcct	38
VP2I-3W-F	PCR and sequencing	cc**gaattc**atcacactgttctcagccaacattgatgcc	38
VP2I-3W-R	PCR and sequencing	cgg**gcggccgc**ccagtactttagctggccggggcttgg	38
VP4-RK-F	PCR and sequencing	cc**gaattc**aggatagctgtgccggtggtctcc	32
VP4-RK-R	PCR and sequencing	ga**gcggccgc**tttgatgaacgttgcccagttggg	34
VP4-AA-F	PCR and sequencing	ct**gaattc**gccgacaaggggtacgaggtagtc	32
VP4-AA-R	PCR and sequencing	ac**gcggccgc**agccatggcaaggt ggtactggc	33

[a] The restriction site extensions are in bold

Table 2.5 Forward and reverse primers for the designed IBDV polyprotein constructs

Fragment	Forward Primers	Reverse Primer	Expected size (bp)
Full-length polyprotein	IBDA-F	IBDA-R	3036
Truncated polyprotein	VP2I-3W-F	VP2I-3W-R	1995
VP4-AW	VP4-AA-F	VP2I-3W-R	1137
VP4-RA	VP4-RK-F	VP4-AA-R	909
VP4-RK	VP4-RK-F	VP4-RK-R	836
VP4-AA	VP4-AA-F	VP4-AA-R	745

2.3.5 Reverse transcription and PCR

RT-PCR was initially attempted using the ImProm-II reverse transcription system and AMV reverse transcriptase/*Tfl* DNA polymerase Access RT-PCR System according to the manufacturer's instructions, but was unsuccessful. Later, a two-step RT-PCR was carried out using Superscript III reverse transcriptase and Long PCR enzyme mix. For segment A cDNA synthesis, primers P1-18F and P3220-3241R were used in the reverse transcription reaction. As template, 1 µg of IBDV dsRNA (10 µl) was mixed with 10 µl dimethyl sulfoxide (DMSO) and denatured at 100°C for 2 min. The denatured RNA (5 µl or 125 ng) was combined with 0.25 µM of the forward and reverse primers in a total volume of 9 µl. The dsRNA/primer mixture was incubated at 100°C for 5 min and immediately transferred onto ice for 1 min before mixing with the revere transcription mix (11 µl) containing 1x RT buffer, 5 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTP mix, 40 U of RNase inhibitor and 200 U of Superscript III reverse transcriptase. The 20 µl reaction mixture was incubated at 50°C for 1 h. The reverse transcriptase was inactivated by heating at 85°C for 5 min after which RNase H (2 U) was added and incubation was continued at 37°C for 20 min.

The reverse transcription reaction mix (2 µl) was used in a total volume of 50 µl reaction mix containing 1x Long PCR buffer without $MgCl_2$, 0.2 mM of dNTP mix, 200 nM each of forward and reverse primers, 2 mM $MgCl_2$ and 1.25 U of the Long PCR enzyme mix. Amplification was carried out in a GeneAmp 9700 thermocycler (Applied Biosystems, California, USA) programmed with the following temperature profile: 94°C (2 min) for initial denaturation, 10 cycles of denaturation at 94°C (30 s), annealing at 55°C (30 s), extension at 68°C (3 min), 25 cycles of denaturation at 94°C (30 s), annealing at 55°C (30 s), extension at 68°C (3 min + 10 s/cycle) and the final elongation step at 68°C (10 min). The PCR products were analysed by agarose gel electrophoresis (Section 2.3.3).

2.3.6 T-vector cloning of the polyprotein and VP4 amplicons

The RT-PCR products were purified from the agarose gel using the peqGOLD gel extraction kit. All purified RT-PCR products were ligated into the pGEM-T Easy vector (Fig. 2.2A) except for VP4-RA and VP4-AW which were ligated into pTZ57R/T (Fig. 2.2B) because they were cloned at a later stage in the study. The ligation was carried out using T4 DNA ligase at an insert:vector ratio of 4:1, and the reaction was incubated overnight at 4°C. The amount of insert DNA to be used for each ligation was calculated using the following formula (T4 ligase product information, Promega, USA):

The ligation mix was transformed into competent *E. coli* JM109 cells prepared using a TransformAid™ bacterial transformation kit. Transformed cells were grown overnight on 2xYT agar plates [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 86 mM NaCl, 1.5% (w/v) bacteriological agar] containing 100 µg/ml ampicillin, 0.1 mM IPTG and 0.005% (v/v) X-gal. Recombinant colonies were screened by colony PCR using *Taq* DNA polymerase and vector primers (T7 forward primer and SP6 reverse primer for pGEM-T Easy; M13 forward primer and T7 reverse primer for pTZ57R/T; see Appendix 2 for primer sequences). Plasmid DNA was isolated by the GeneJET™ plasmid miniprep kit and used for further screening by sequencing at the International Livestock Research Institute (ILRI), Nairobi, Kenya. For the full-length polyprotein coding region, the sequencing was carried out using multiple primers as shown in Fig. 2.3.

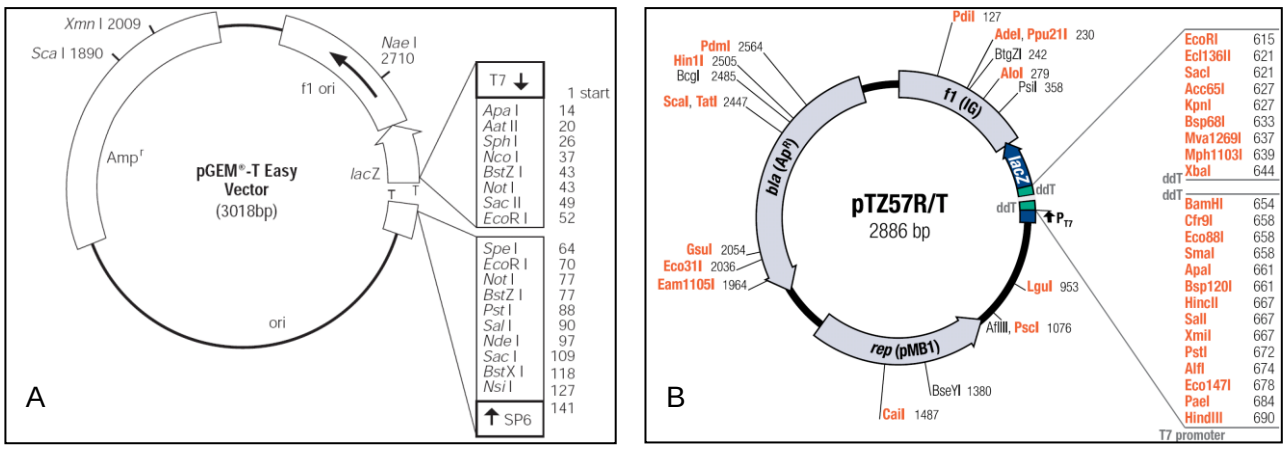


Figure 2.2 Structure of T-vectors used in the present study. A) pGEM-T Easy vector. **B)** pTZ57R/T vector. Both vectors carry f1 origins of replication, ampicillin resistance, lacZ markers, SP6 and T7 promoter sequences. The vectors are supplied in the linear form with a single thymidine (T) overhang at the 3' ends.

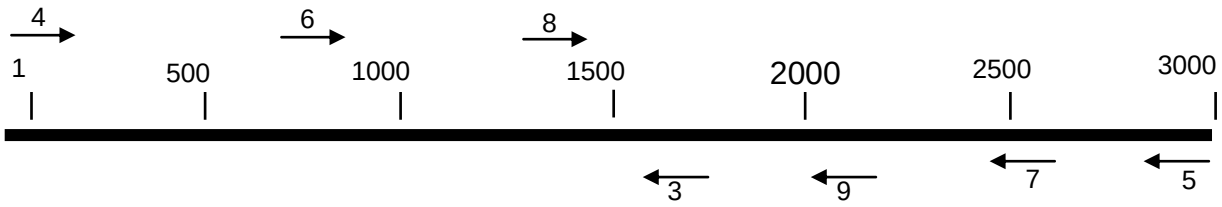


Figure 2.3 Location of the primers used for sequencing of the polyprotein coding region. Numbers correspond to the primer sequences in Table 2.4.

2.3.7 Sub-cloning into expression vectors

A single colony containing a recombinant T-vector was grown overnight in 2xYT broth [1.6% (w/v) tryptone, 1% (w/v) yeast extract, 86 mM NaCl] containing 100 µg/ml ampicillin (2xYT-Amp) and plasmid DNA was isolated, digested with EcoRI and NotI to release the insert DNA. All double digestion reactions with EcoRI and NotI were carried out in a 30 µl reaction containing 10x buffer O (3 µl), plasmid DNA (25 µl), EcoRI (1 µl or 10 U) and NotI (1 µl or 10 U). The digestion was carried out overnight at room temperature (RT, approximately 25°C) and analysed on a 1% (w/v) agarose gel (Section 2.3.3). For the full-length polyprotein coding region, the released insert DNA had the same size of 3 kb as that of the plasmid so the EcoRI/NotI digested plasmid was subjected to BgII digestion. The released inserts were purified from the agarose gel using the peqGOLD gel extraction kit.

For preparation of vectors (*E. coli* based pET-28b, pET-32b and pGEX-4T-1; *P. pastoris* based pPIC9), *E. coli* JM109 cells containing a plasmid in a glycerol stock were streaked on 2xYT-agar plates containing the appropriate antibiotic and incubated overnight at 37°C. The following day, a single colony was used to inoculate 2xYT medium containing the antibiotic and grown overnight with shaking at 37°C. Plasmid DNA was isolated and subjected to a double digest with EcoRI and NotI overnight at RT in a 30 µl reaction volume as before. About 2 µl of the reaction was analysed by agarose gel electrophoresis (Section 2.3.3) to check if the digestion was complete. The enzymes were deactivated by incubating at 65°C for 15 min prior to dephosphorylation. The dephosphorylation reaction, containing 10x SAP buffer (3 µl), digestion mix (26 µl) and SAP (1 µl), was incubated for 1 h at 37°C. The SAP was then inactivated by incubation at 65°C for 15 min. The dephosphorylated vectors were purified using DNA clean and concentration kit.

The purified inserts were ligated into dephosphorylated vectors using T4 DNA ligase and incubated overnight at 4°C. The ligation mixtures were used to transform *E. coli* JM109

cells using the TransformAid™ bacterial transformation kit and transformed cells were grown overnight on 2xYT plates containing appropriate antibiotics. Recombinant colonies were screened by colony PCR using vector primers i.e. pET primers, pGEX primers and AOX primers for pET vectors, pGEX-4T-1 and pPIC9, respectively (Appendix 2). Screening was further carried out by miniprep DNA isolation, restriction digestion and sequencing.

2.3.8 Sequence assembly and analysis

The nucleotide and translated amino acid sequences of the 52 IBDV strains (Table 2.1) were used for comparative analysis. The nucleotide sequence of the SA-KZN95 isolate was translated into an amino acid sequence and used for multiple sequence alignment using the BioEdit™ program. Phylogenetic analysis was conducted using the MEGA4 program (Tamura *et al.*, 2007). The tree was constructed from both aligned nucleotide and amino acid sequences using the neighbour-joining method. Branch topology was verified by generating 1000 bootstraps.

2.4 Results

2.4.1 IBDV identification by transmission electron microscopy

The presence of the IBDV particles in the infected bursal samples was confirmed by electron microscopy before RNA isolation. Electron microscopic analysis of the viral suspension revealed the presence of icosahedral particles with a diameter of approximately 60 nm (Fig. 2.4). This is in agreement with the described structure of IBDV particles (Bottcher *et al.*, 1997). The particles had a dark interior and a clearer margin which is typical for immature viral particles. It has been reported that some IBDV strains are unstable in CsCl gradients and may result in incomplete particles with abnormal protein compositions (Kibenge *et al.*, 1988).

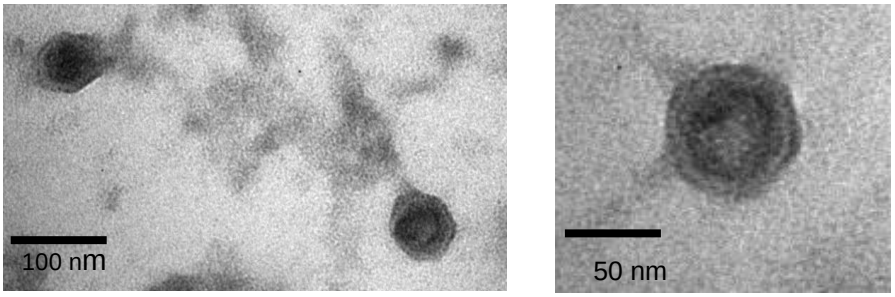


Figure 2.4 Electron micrograph of negatively stained infectious bursal disease virus from the infected bursal samples.

2.4.2 Isolation of viral dsRNA and RT-PCR

Double stranded viral RNA was purified from IBDV infected chicken bursae. The dsRNA segments were separated and visualised in a 1% (w/v) agarose gel, yielding two bands (A-segment and B-segment) of 3.3 and 2.9 kb, respectively (Fig. 2.5A). The dsRNA was used as template to synthesise segment A cDNA. The RNA component of the hybrid DNA-RNA was removed subsequently by the addition of RNase H and the cDNA was used as a template for the PCR reaction. The PCR step produced products of the expected sizes of approximately 3 kb (full-length ORF-A2), 2 kb (truncated ORF-A2), 1150 bp (VP4-AW), 900 bp (VP4-RA), 836 bp (VP4-RK) and 745 bp (VP4-AA) (Fig. 2.5B and C).

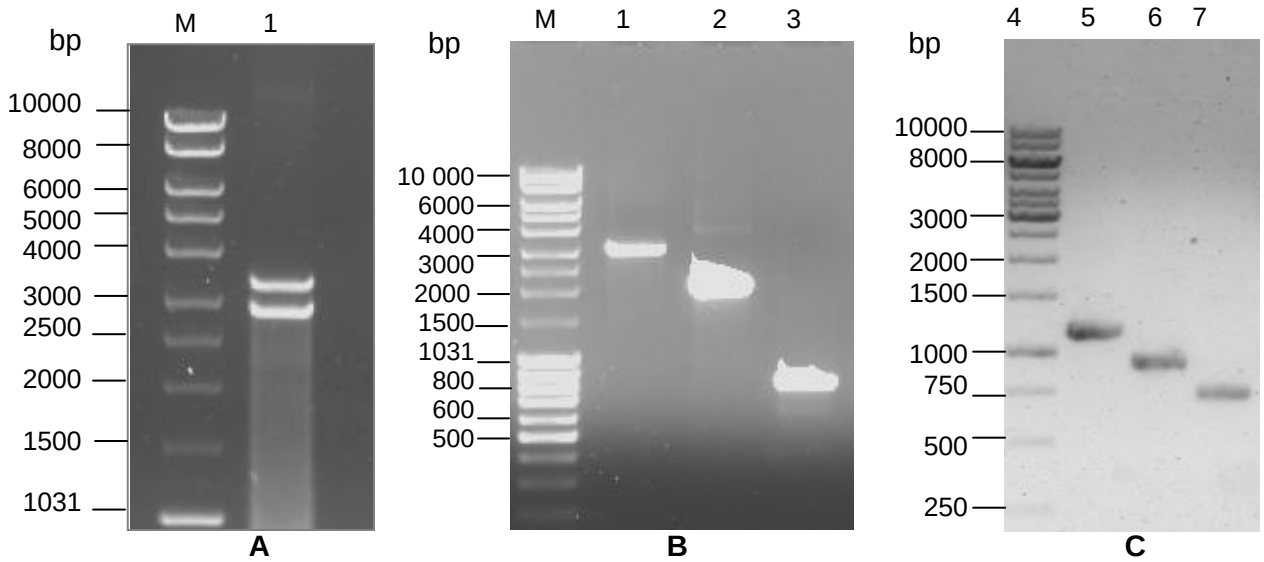


Figure 2.5 Agarose gel electrophoresis of viral RNA and RT-PCR products. A) IBDV dsRNA from infected bursal tissue. Lane M, MassRuler DNA ladder mix; lane 1, IBDV dsRNA segments (1 µg). **B)** and **C)** RT-PCR products. Lane M, MassRuler DNA ladder mix; lane 1, full-length ORF-A2; lane 2, truncated ORF-A2; lane 3, VP4-RK; lane 4, GeneRuler 1 kb DNA ladder mix; lane 5, VP4-AW; lane 6, VP4-RA and lane 7, VP4-AA.

2.4.3 Cloning into a T-vector

The PCR products were amplified with the addition of a 3' adenosine overhang, hence allowing for efficient TA cloning into a T-vector (pGEM-T Easy and pTZ57R/T). The white colonies selected by color screening on Xgal/IPTG plates were further screened for the presence of the inserts by miniprep DNA isolation followed by restriction digestion. The primers were designed with EcoRI and NotI restriction sites and a double digestion with these enzymes would therefore release the insert DNA from the recombinant plasmid. The full-length ORF-2A, truncated ORF-2A, VP4-RK and VP4-AA amplicons were

successfully cloned into the pGEM-T Easy vector (Fig. 2.6), whereas VP4-AW and VP4-RA were cloned into pTZ57R/T (Fig. 2.7).

The full-length polyprotein coding region had the same size of 3 kb as the vector, therefore the released insert and vector DNA could not be distinguished on the gel (Fig. 2.6, lanes 2 and 4). By using BgII which has a site in pGEM-T Easy, but not in the IBDV segment A sequence, it was possible to distinguish between the insert and the vector DNA. As can be seen in lanes 3 and 5, BgII digestion resulted in a 3 kb fragment corresponding to the insert DNA and two fragments at about 1.5 kb for the vector DNA. In the other cases, the vector was longer than the insert, therefore EcoRI and NotI digestion of the recombinant pGEM-T Easy released the 3 kb vector and the respective inserts: truncated ORF-A2 at 2 kb (lane 7), VP4-RK at 836 bp (lanes 9 and 10) and VP4-AA at 745 bp (lane 12).

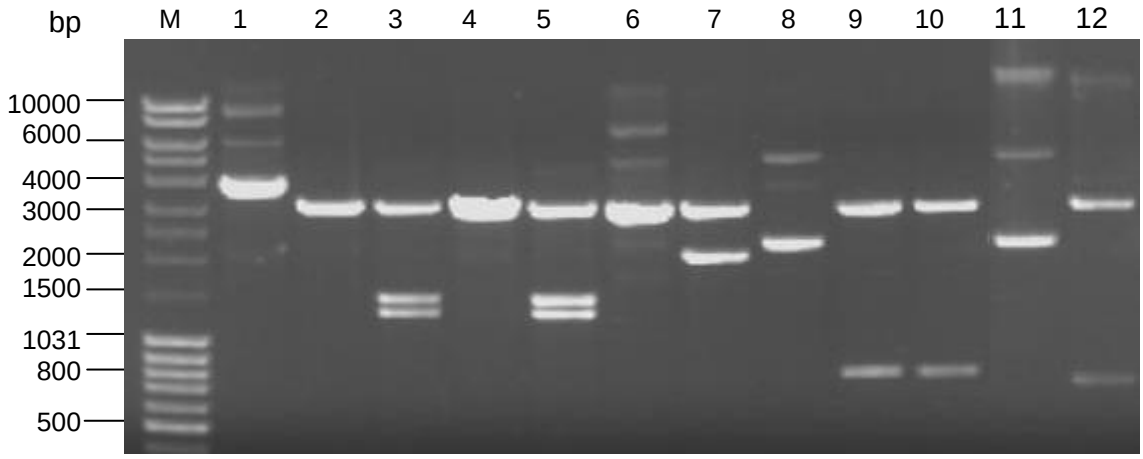


Figure 2.6 Agarose gel electrophoresis of the screening of recombinant pGEM-T Easy plasmids by restriction digestion with EcoRI and NotI. Lane M, MassRuler DNA ladder mix; lane 1, pGEM-T Easy-full-length ORF-A2 uncut; lanes 2 and 4, pGEM-T Easy-full-length ORF-A2 clone cut with EcoRI and NotI; lanes 3 and 5, pGEM-T Easy-full-length ORF-A2 clone cut with EcoRI/NotI and BgII; lane 6, pGEM-T Easy-truncated ORF-A2 uncut; lane 7, pGEM-T Easy-truncated ORF-A2 cut with EcoRI and NotI; lane 8, pGEM-T Easy-VP4-RK uncut; lanes 9 and 10, pGEM-T Easy-VP4-RK cut with EcoRI and NotI; lane 11, pGEM-T Easy-VP4-AA uncut; lane 12, pGEM-T Easy-VP4-AA cut with EcoRI and NotI.

Similarly, restriction digestion of recombinant pTZ57R/T plasmids with EcoRI and NotI released the VP4-AW and VP4-RA at the expected sizes of about 1150 (Fig 2.7A, lane 2) and 900 bp (Fig 2.7B, lane 1), respectively.

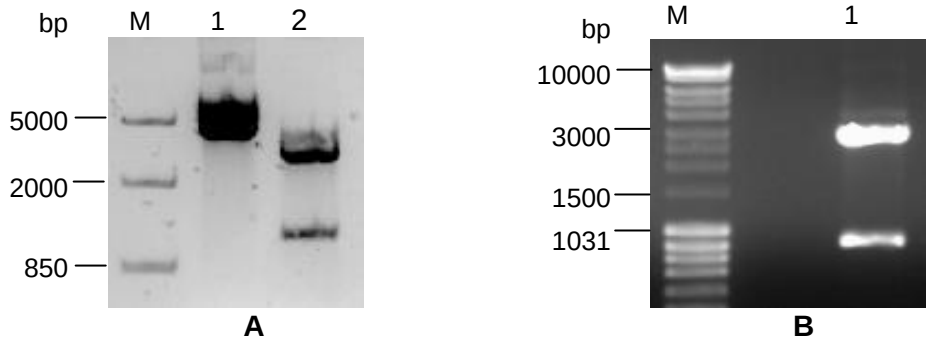


Figure 2.7 Agarose gel electrophoresis of the screening of recombinant pTZ57RT plasmids by restriction digestion with EcoRI and NotI. A) VP4-AW cloning. Lane M, FastRuler middle range DNA ladder; lane 1, pTZ57RT-VP4-AW uncut; lane 2, pTZ57R/T-VP4-AW cut. **B)** VP4-RA cloning. Lane M, MassRuler DNA ladder mix; lane 1, pTZ57RT-VP4-RA cut**.**

2.4.4 Sub-cloning into the expression vectors and screening of recombinants

The pET-28b, pET-32b, pGEX-4T-1 and pPIC9 expression vectors were subjected to EcoRI and NotI digestion, followed by dephosphorylation in preparation for cloning. The dephosphorylated plasmids were purified and analysed by agarose gel electrophoresis (Fig. 2.8). The observed and expected sizes were 5.5 kb (pET-28b), 6 kb (pET-32b), 5 kb (pGEX-4T-1) and 8 kb (pPIC9) (lanes 1-4).

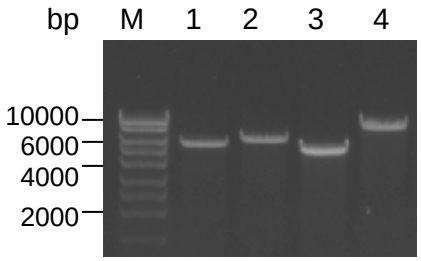


Figure 2.8 Agarose gel analysis of the gel purified vectors that were restricted by EcoRI and NotI followed by dephosphorylation. Lane M, MassRuler DNA ladder mix; lane 1, pET-28b; lane 2, pET-32b; lane 3, pGEX-4T-1; lane 4, pPIC9.

The released inserts were gel-purified and sub-cloned into expression vectors. All four fragments (full-length ORF-A2, truncated ORF-A2 and VP4-RK and VP4-AA), initially cloned into pGEM-T Easy, were each sub-cloned into all expression vectors. The colonies were firstly screened by direct colony PCR using the vector specific primers (Results not shown) prior to miniprep DNA isolation. The pET primers are expected to produce approximately 360 bp and 750 bp DNA fragments in the non-recombinant pET-28b and pET-32b vector, respectively. The pGEX primers produce a 150 bp fragment in a non-recombinant pGEX-4T-1 vector while the AOX primers produce a fragment of about 500 bp in a non-recombinant pPIC9 plasmid. Therefore the colony PCR product from a

recombinant plasmid would be the size of the insert plus the fragment produced from a non-recombinant plasmid.

The recombinant plasmids were further screened by restriction digestion with EcoRI which was expected to linearise the plasmids. All four fragments were successfully cloned into pET-32b (Fig. 2.9), pGEX-4T-1 (Fig. 2.10) and pPIC9 (Fig. 2.11). Sub-cloning into pET-28b was successful only for the full-length ORF-A2 and VP4-RK (Fig. 2.9). Upon linearisation, the non-recombinant pET-28b plasmid was observed at 5.4 kb (Fig. 2.9, lane 2), whereas recombinant plasmids containing the full-length ORF-A2 and VP4-RK were observed at the expected sizes of 8.4 kb (lane 4) and 6.2 kb (lane 6). However, the digestion of the recombinant pET-28b-VP4-RK was incomplete, as shown by the presence of the other conformations (lane 6). For pET-32b, the linear non-recombinant plasmid was observed at the expected size of 6 kb (Fig. 2.9, lane 8), although it was also incompletely digested. The recombinant plasmids containing the full-length ORF-A2, truncated ORF-A2, VP4-RK and VP4-AA fragments were also observed at the expected sizes of about 9 kb (lane 10), 8 kb (lane 12), 6.8 kb (lane 14) and 6.7 kb (lanes 16), respectively.

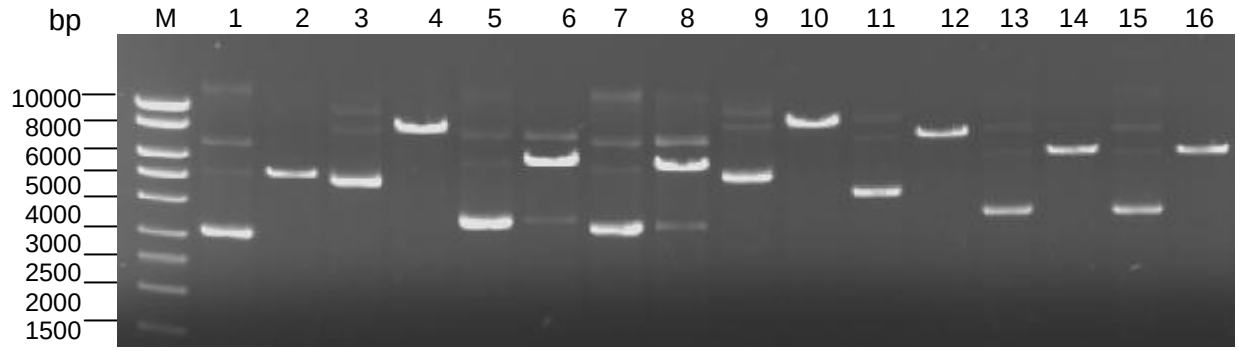


Figure 2.9 Agarose gel electrophoresis of the restriction digest of recombinant pET-28b and pET-32b plasmids with EcoRI only. Lane M, MassRuler high range DNA ladder; lane 1, non-recombinant pET-28b uncut; lane 2, non-recombinant pET-28b cut; lane 3, pET-28b-full-length ORF-A2 uncut; lane 4, pET-28b-full-length ORF-A2 cut; lane 5, pET-28b-VP4-RK uncut; lane 6, pET-28b-VP4-RK cut; lane 7, non-recombinant pET-32b uncut; lane 8, non-recombinant pET-32b cut; lane 9, pET-32b-full-length ORF-A2 uncut; lane 10, pET-32b-full-length ORF-A2 cut; lane 11, pET-32b-truncated ORF-A2 uncut; lane 12, pET-32b-truncated ORF-A2 cut; lane 13, pET-32b-VP4-RK uncut; lane 14, pET-32b-VP4-RK cut; lane 15, pET-32b-VP4-AA uncut; lane 16, pET-32b-VP4-AA cut.

Subcloning into pGEX-4T-1 was also successful for all four fragments (Fig. 2.10). Linearisation of the recombinant plasmids produced the expected sizes of 8 kb for full-length ORF-A2 (lanes 2 and 4), 7 kb for truncated ORF-A2 (lanes 6 and 8), 5.7 kb for VP4-RK (lanes 10 and 12) and 5.6 kb for VP4-AA (lanes 14 and 16). The linearised non-

recombinant pGEX-4T-1 plasmid was observed at the expected size of 5 kb (lane 18). Similarly, the linear non-recombinant pPIC9 plasmid was observed at the expected size of 8 kb (Fig. 2.11, lane 1). Linearisation of the pPIC9 recombinant plasmids produced the expected sizes of 11 kb for full-length ORF-A2 (lane 3), 10 kb for truncated ORF-A2 (lane 5), 8.8 kb for VP4-RK (lane 7) and 8.7 kb for VP4-AA (lane 9).

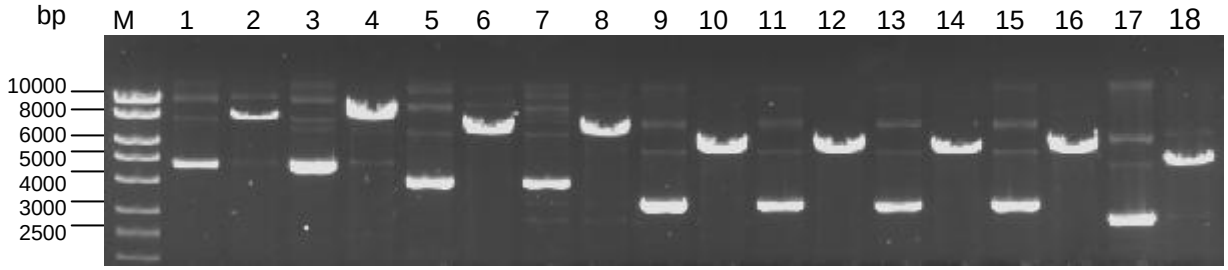


Figure 2.10 Agarose gel electrophoresis of the restriction digest of recombinant pGEX-4T-1 plasmids with EcoRI only. Lane M, MassRuler high range DNA ladder; lanes 1 and 3, pGEX-4T-1-full-length ORF-A2 uncut; lanes 2 and 4, pGEX-4T-1-full-length ORF-A2 cut; lanes 5 and 7, pGEX-4T-1-truncated ORF-A2 uncut; lanes 6 and 8, pGEX-4T-1-truncated ORF-A2 cut; lanes 9 and 11, pGEX-4T-1-VP4-RK uncut; lanes 10 and 12, pGEX-4T-1- VP4-RK clone cut; lanes 13 and 15, pGEX-4T-1-VP4-AA uncut; lanes 14 and 16, pGEX-4T-1-VP4-AA cut; lane 17, non-recombinant pGEX-T-1 uncut; lane 18, non-recombinant pGEX-4T-1 cut.

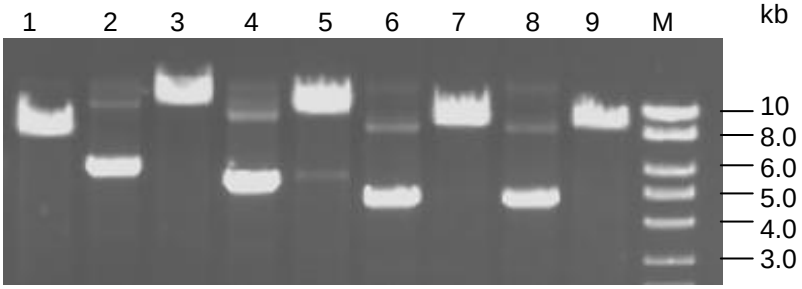


Figure 2.11 Agarose gel electrophoresis of the restriction digest of recombinant pPIC9 plasmids with EcoRI only. Lane M, MassRuler high range DNA ladder; lane 1, non-recombinant pPIC9 cut; lane 2, pPIC9-full-length ORF-A2 uncut; lane 3, pPIC9-full-length ORF-A2 cut; lane 4, pPIC9-truncated ORF-A2 uncut; lane 5, pPIC9-truncated ORF-A2 cut; lane 6, pPIC9- VP4-RK uncut; lane 7, pPIC9-VP4-RK cut; lane 8, pPIC9-VP4-AA uncut; lane 9, pPIC9-VP4-AA cut.

On the other hand, VP4-AW and VP4-RA, initially cloned into pTZ57R/T, were sub-cloned into pET-32b and pGEX-4T-1 only. The recombinants were successfully screened by colony PCR (results not shown) and subsequently by restriction digestion using EcoRI and NotI to release the inserts (Fig. 2.12). The double digestion of recombinant pET-32b and pGEX-4T-1 plasmids yielded VP4-AW at the expected size of 1150 bp (Fig. 2.12A and B, lane 2) and VP4-RA at 900 bp (Fig. 2.12A and B, lane 4).

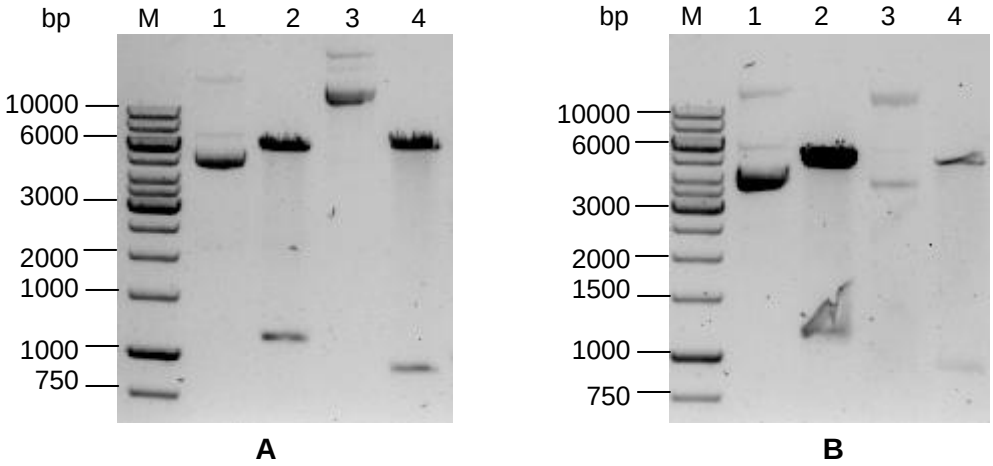


Figure 2.12 Agarose gel electrophoresis recombinant screening for VP4-AW and VP4-RA by restriction digestion with EcoRI and NotI. A) pET-32b recombinant plasmids. Lane M, GeneRuler 1 kb DNA ladder mix; lane 1, pET-32b-VP4-AW uncut; lane 2, pET-32b-VP4-AW cut; lane 3, pET-32b-VP4-RA uncut; lane 4, pET-32b-VP4-RA cut. **B)** pGEX-4T-1 recombinant plasmids. Lane M, GeneRuler 1 kb DNA ladder mix; lane 1, pGEX-4T-1-VP4-AW uncut; lane 2, pGEX-4T-1-VP4-AW cut; lane 3, pGEX-4T-1-VP4-RA uncut; lane 2, pGEX-4T-1-VP4-RA cut.

2.4.5 Sequence analysis

The recombinant plasmids were sequenced to verify that the inserts were cloned in frame and that no stop codons were introduced prior to expression (Chapter 3). The obtained sequence of the polyprotein coding region was also used for sequence analysis to determine the genetic relatedness of the isolate used in the present study. The sequence of the full-length polyprotein coding region was translated *in silico* into the corresponding amino acid sequence for comparison with polyprotein sequences of other IBDV strains.

2.4.5.1 Analysis of virus marker residues for the identification of isolate

Previous work by Xia *et al*. (2008) identified nineteen polyprotein residues that do not necessarily control virulence but are unique to the different pathotypes of IBDV and consequently may be used for virus classification. Two additional residues of this nature were identified in the present study. Thirteen of these residues are found in the VP2 sequence (Table 2.6), five in the VP4 sequence and three in the VP3 sequence (Table 2.7). The 21 residues were compared between the SA-KZN95 isolate and the 52 isolates shown in Table 2.1, and used to determine the genetic relatedness of the South African isolate (Tables 2.6 and 2.7).

The critical amino acid residues of the SA-KZN95 polyprotein were similar to those of the vvIBDV strains, particularly the Malaysian UPM94/273 strain. SA-KZN95 and UPM94/273 were different from other vvIBDV strains at position 254, where they have a Ser found in serotype II viruses and American variant strains. While SH/92, ks, BD 3/99 and SDH1

differed from SA-KZN95 only at position 254, BD 3/99 and SDH1 were different at position 685 where they had Ser instead of Asn. The VP2 residues of UPM97/61, UK661, Gx, GZ/96, 02015.1, P07, Korean, OKYM, HK46, GHUT-12, SH95, TO9 and Harbin-1 were identical. They differed from all other vvIBDV strains at position 279 where they had Asn found in variant and attenuated strains, instead of Asp found in all other vvIBDV strains. Some vvIBDVs showed unique characteristics, for example, while all other vvIBDV strains had Val^{990}, strain D6948 had Ala^{990} that is conserved in all classic, variant, attenuated and serotype II strains. TASIK was different in that it had Ser^{222} found in attenuated strains (Lukert, P3009 and K310) instead of Ala^{222} found in all vvIBDV strains.

Very virulent strains shared most of their residues except at positions 253, 284, 300 and 981 where they have the same residues as the classic and variant strains. The virulence determinants of 52/70 and IM were identical, while those of 002-73 were unique. The discriminating residues of HRP-2 were identical to those of STC, except at position 680 where HRP-2 has Cys and STC has Phe. OKYM was unique among the classic strains, and it was the only strain that contained Gly at position 279 instead of Asn or Asp found in all other IBDV strains. In addition, it contained Tyr at position 680 while other classic strains contained a Cys.

Pro^{222} appears to be unique among the classic strains. Variant strains were distinguished by Thr/Gln^{222}, Lys^{249}, Ser/Asn^{254} and Ile^{286}. Variant E was also distinguished by Gln at position 981 instead of Pro or Leu found in all other IBDV strains. Natural strains were distinguished by His^{253}, Thr^{284}, Arg^{330} and Leu^{981} when compared with the attenuated strains, thus showing that this residue may also be associated with tissue tropism. KK1 and KSH represent interesting recombinants strains in that the VP2 hypervariable region (residues 206 to 350) was identical to that of vvIBDV strains, while the residues in VP4 and VP3 were similar to those of attenuated strains. K310 appears to have these recombinant strains characteristics in that the VP2 region is similar to that in classic and antigenic variant strains while VP4 and VP3 are identical to the attenuated strains.

The two serotype II strains, OH and 23/82 were different from each other, however they share Glu at position 256. At this position the vvIBDV strains contain Ile while the rest contain mainly Val. Interestingly, they share His^{253} (OH strain only) and Thr^{284} with attenuated strains. Interesting differences were observed in the VP4 region at position 680-681. These positions are occupied by Tyr-Asn in vvIBDV strains or Cys-Lys in the other serotype I strains. At position 680, OH has a deletion while 23/82 contains Arg. At position 681, the apathogenic strains were distinguished by Ser, which is also found in two vvIBDV strains BD 3/99 and SDH1.

Table 2.6 IBDV markers in the VP2 region of polyprotein [(Adapted from Xia *et al.* (2008)]

Virulence	Strain	222	242	249	253	254	256	279	284	286	294	299	330	451
	SA-KZN95	**A**	**I**	**Q**	**Q**	**S**	**I**	**D**	**A**	**T**	**I**	**S**	**S**	**L**
vvIBDV	UPM94/273	•	•	•	•	•	•	•	•	•	•	•	•	•
	SH/92, ks	•	•	•	•	G	•	•	•	•	•	•	•	•
	BD 3/99, SDH1	•	•	•	•	G	•	•	•	•	•	•	•	•
	D6948	•	•	•	•	G	•	•	•	•	•	•	•	•
	TASIK	S	•	•	•	G	•	•	•	•	•	•	•	•
	UPM97/61, UK661, Gx, GZ/96, 02015.1, PO7, Korean, OKYM, HK46, GHUT-12, SH95, TO9, Harbin-1	•	•	•	•	G	•	N	•	•	•	•	•	•
Classic	52/70, IM	P	•	•	•	G	V	•	•	•	L	N	•	I
	HPR-2, STC	P	V	•	•	G	V	•	•	•	L	N	•	I
	002-73	P	V	•	•	G	V	G	•	•	L	•	•	I
Variant	GIs	T	V	K	•	•	V	N	•	•	L	N	•	I
	Variant E	T	V	K	•	•	V	N	•	I	L	N	•	I
	9109	T	V	K	•	N	V	N	•	I	L	N	•	I
	GZ29112	Q	V	K	H	•	V	N	T	I	L	N	•	I
Attenuated	Lukert, P3009	S	•	H	H	G	A	•	T	I	L	N	•	I
	Cu-1, D78, CEF94, KT1/99, NB, JD1, TL2004, HZ2, B87, ZJ2000, Harbin, GA-1	P	V	•	H	G	V	N	T	•	L	N	R	I
	CT, Soroa	P	V	R	H	G	V	N	T	•	L	N	R	I
	CS-2-35, H-30	P	V	•	•	G	V	N	T	•	L	N	R	I
Recombinants	KK1, KSH	•	•	•	•	G	•	•	•	•	•	•	•	•
	K310	S	V	•	•	•	V	N	•	•	L	N	•	I
Serotype II	OH	V	V	•	H	•	E	•	T	•	L	•	•	I
	23/82	P	V	•	•	•	E	•	T	•	L	N	•	I

(● represents residues identical to those of SA-KZN95)

Table 2.7 IBDV markers in VP4 and VP3 [(Adapted from Xia *et al.* (2008)]

Virulence	Strain	VP4		VP3					
		541	680	681	715	751	981	990	1005
	SA-KZN95	**I**	**Y**	**N**	**S**	**D**	**P**	**V**	**A**
vvIBDV	UPM94/273, SH/92, ks, TASIK, UK661, Gx, GZ/96, 02015.1, PO7, Korean, OKYM, HK46, GHUT-12, SH95, TO9, Harbin-1	•	•	•	•	•	•	•	•
	UPM97/61	V	•	•	•	•	•	•	•
	BD 3/99, SDH1	•	•	S	•	•	•	•	•
	D6948	•	•	•	•	•	•	A	•
Classic	52/70, IM, HPR-2, STC	V	C	K	P	H	•	A	T
	002-73	V	•	K	P	H	•	A	T
Variant	Gls, GZ29112	•	C	K	P	H	•	A	T
	Variant E	•	C	K	P	H	Q	A	T
	9109	•	•	K	P	H	•	A	•
Attenuated	Lukert, P3009	V	C	K	P	H	•	A	T
	Cu-1, D78, CEF94, KT1/99, NB, JD1, TL2004, HZ2, B87, ZJ2000, GA-1, Harbin, CT, Soroa, CS-2-35, H-30	V	C	K	P	H	L	A	T
Recombinants	KK1, KSH	•	C	K	P	H	L	A	T
	K310	V	C	K	P	H	L	A	T
Serotype II	OH	•	-	S	P	H	•	A	T
	23/82	•	R	S	P	H	•	A	T

(● represents residues identical to those of SA-KZN95; - represents a deletion)

2.4.5.2 Evolutionary relationships of SA-KZN95 with other IBDVs

Phylogenetic analysis was done at both nucleotide (Fig. 2.13) and amino acid sequence level (Fig. 2.14). A phylogenetic tree was constructed based on the nucleotide sequence of the polyprotein to determine the origin of the SA-KZN95 isolate and the phylogenetic relationship between the strains. The tree suggests that there are two main lineages, corresponding to the two serotypes. Serotype II consists of the American OH and European 23/82 strains, which are apathogenic. The serotype I lineage is divided into two main branches, one comprising the classic Australian 002-73 strain and the other contain the rest of the strains. The latter branch is further divided into two major groups, one consisting of vvIBDV strains and the other consisting of variants, classic and attenuated strains (Fig. 2.13).

Two clusters of attenuated strains were observed. The first cluster consists of European CT and Soroa, Chinese B87 and the American GA-1, H-30 and C2-2-35 strains. In the same cluster are the German CU-1 and the Indian KT1/99 strain which form a clade. The second cluster consists largely of the Chinese strains. These results indicate that CT, Soroa, GA-1, H-30, CS-2-35, KT1/99 and NB isolates are attenuated strains. The Korean K310 strain is distantly related to these two clusters. The classic strains, German IM and American STC are grouped together with HRP-2 in one cluster. The British classic 52/70 was clustered with two attenuated strains, P3009 and Lukert, which may suggest that the two attenuated strains are derived from an ancestor of 52/70. Between the classic and attenuated virus clusters are the Korean strains (KSH, KK1, K310) which seemed to be recombinant strains when analysed based on the critical residues (Table 2.6 and 2.7).

The American variant strains, Gls, 9109 and variant E are clustered together with the Chinese GZ29112. More than three noticeable clusters of vvIBDV strains were observed, the first one consisted of three Chinese strains (Gx, GZ/96 and Harbin-1) and the South African SA-KZN95 isolate. The second cluster consists of Asian strains and the Israeli ks strain. In the third cluster, the two Malaysian strains seem to be closely related to the SDH1 strain from Iran. In addition, the European strains D6948 and UK661, together with the Indonesian TASIK strain formed a distinct clade. The African vvIBDV strains TO9 and PO7 each seem to have originated separately. In the amino acid sequence based tree, the serotype I branch is divided into two major groups, one consisting of vvIBDV, hydrids and classic strains, and the other comprise attenuated and variant strains. vvIBDV strains are distinctly grouped into several clades. SA-KZN95 is among six strains which are not part of any clade i.e. OKYM, PO7, TASIK and SH/92 (Fig. 2.14).

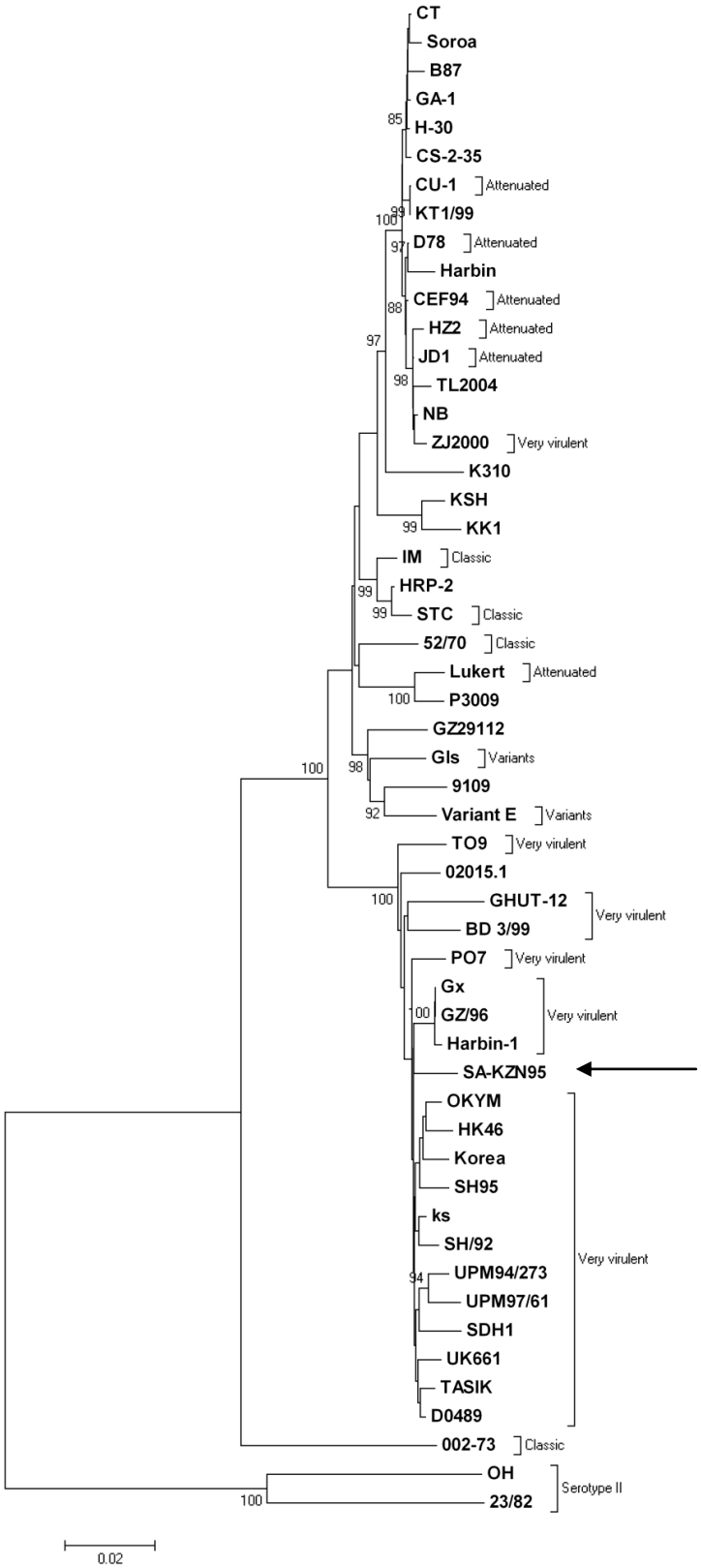


Figure 2.13 A phylogenetic tree based on the nucleotide sequence of the IBDV polyprotein. Multiple alignment and tree construction were performed by MEGA4 with the neighbour-joining method. Bootstrap probabilities (%) as determined for the 1000 replicates and only values greater than 80 are shown.

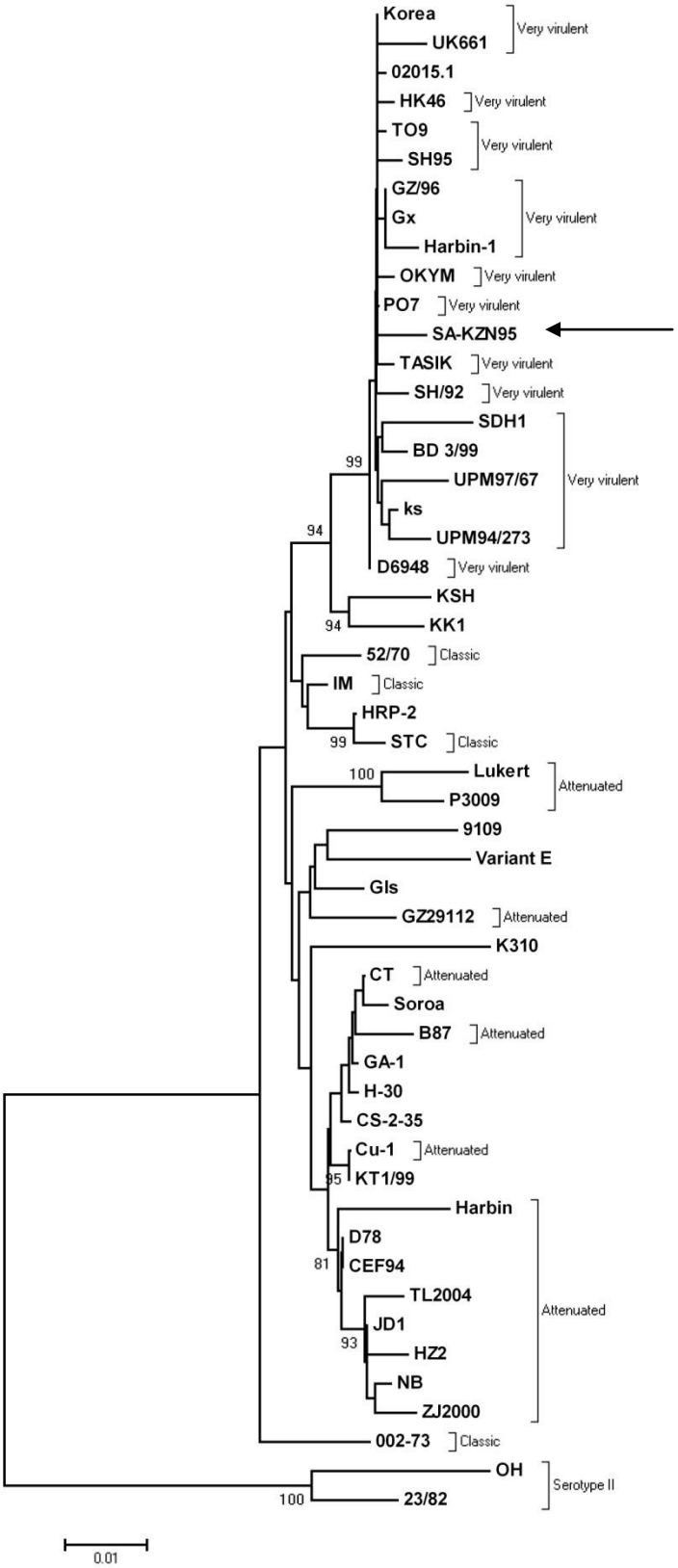


Figure 2.14 A phylogenetic tree based on the amino acid sequences of the IBDV polyprotein. Multiple alignment and tree construction were performed by MEGA4 with the neighbour-joining method. Bootstrap probabilities (%) as determined for the 1000 replicates and only values greater than 80 are shown.

2.5 Discussion

VP4 is initially an integral part of the polyprotein and it autocatalytically cleaves the ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} peptide bonds to produce the mature VP4 (Ala^{513}-Ala^{755}) (Sanchez and Rodriguez, 1999). The present study reports on the primer design, amplification and cloning of the IBDV polyprotein and VP4 constructs into T- and expression vectors. The IBDV infected bursal tissues used in the present study were from an IBDV outbreak in commercial broilers that took place in the KwaZulu-Natal province of South Africa in 1995, hence the isolate was named SA-KZN95. No genetic data was available on the isolate. Since field observation is not always sufficient to distinguish between virulent strains, genetic sequencing is required to distinguish between classic, very virulent and variant strains. In the present study, the cloned polyprotein coding region was also sequenced and the deduced amino acid sequence was compared to 52 other polyprotein sequences from different serotype I and II strains.

Two different types of primers were designed for RT-PCR. Nested primers without restriction site extensions were used in the reverse transcription step and specific primers containing restriction site extensions were used for PCR. It has been reported that more specific amplification of PCR products are obtained when different primers are used in the PCR step compared to the reverse transcription step (Boot *et al.*, 2000b). The cDNA synthesis by RT-PCR was one of the challenging steps in the study. The failure to initially accomplish this was attributed to the lack of denaturation of the template or the inability of the reverse transcriptase to synthesise full-length cDNA strands. Complete denaturation of dsRNA to convert it into ssRNA prior to cDNA synthesis is a fundamental requirement for RT-PCR. Heat denaturation along with the use of denaturants like DMSO (Qian and Kibenge, 1994) and methyl mercury hydroxide (Vakharia *et al.*, 1992) showed to be superior to heat denaturation alone. However, some IBDV studies have reported that heat denaturation alone was sufficient for amplification (Liu *et al.*, 1994; Boot *et al.*, 2000b; Mittal, 2005). In the present study, heat incubation of the viral dsRNA in the presence of DMSO was necessary for denaturation and success of cDNA synthesis.

Initially, reverse transcriptase systems such as AMV and Im PromII were attempted with no success under different denaturation conditions, from which it was speculated that these enzymes were unable to amplify long fragments. Most reverse transcriptases operate at 37°C or 42°C. For synthesis of a 3 kb long cDNA from a dsRNA, a reverse transcriptase that is stable at higher temperatures was required and Superscript III was the best choice. SuperScript III reverse transcriptase is a version of M-MLV reverse

transcriptase with reduced RNase H activity and provides increased thermal stability. The enzyme is used to synthesise cDNA at a temperature range of 42 to 55°C. Enzymes with lower RNase H activity are preferred for longer cDNAs (> 5 kb) (SuperScript™ III First-Strand Synthesis System Manual, Invitrogen Life technologies).

The use of a combination of nested primers without primer extension, Superscript III reverse transcriptase and denaturation of dsRNA in DMSO was necessary to synthesise segment A cDNA. In order to release the cDNA template from the RNA-DNA hybrid, RNase H was added to specifically digest RNA. The released cDNA was used as a template for the PCR amplification of the different fragments. By using specific primers, amplicons were obtained for the full-length ORF-A2, truncated ORF-A2, VP4-AW, VP4-RA, VP4-RK and VP4-AA, at their expected sizes. In order to limit the number of mutations that could be introduced during RT-PCR, a high fidelity enzyme mix consisting of Taq and Pwo polymerases was used (Long PCR Enzyme Mix Manual, Fermentas).

The RT-PCR products were successfully cloned into two different T-vectors. The full-length ORF-A2, truncated ORF-A2, VP4-RK and VP4-AA were cloned into pGEM-T Easy, while VP4-AW and VP4-RA were cloned into pTZ57R/T. For full-length ORF-A2 cloning, restriction digestion with EcoRI and NotI released both the vector and insert DNA at the same size. However, the use of BglI restriction enzyme made it possible to distinguish between the insert and the vector DNA by cutting the vector into two fragments. Cloning of the RT-PCR products into T-vectors made the sub-cloning into expression vectors possible. Full-length ORF-A2, truncated ORF-A2, VP4-RK and VP4-AA were successfully sub-cloned into pET-32b, pGEX-4T-1 and pPIC9. The pET-28b sub-cloning was successful for full-length ORF-A2 and VP4-RK. VP4-AW and VP4-RA were successfully sub-cloned into pET-32b and pGEX-4T-1. All six constructs were expressed and characterised for processing using anti-VP4 peptide antibodies (Chapter 3).

The second aim of the study was to identify the genetic relatedness of the SA-KZN95 isolate. The polyprotein consists of two structural proteins, VP2 and VP3, as well as the viral protease VP4. VP2 is the major protective antigen of IBDV. Neutralising epitopes have been located on the hypervariable region consisting of amino acid residues 206 to 350 (Bayliss *et al.*, 1990). There are two major hydrophilic regions (region A comprising residues 210 to 225 and region B comprising residues 312 to 324), which are very important in the conformation of neutralising epitopes (Heine *et al.*, 1991). Mutations in the hydrophilic regions are important for the ability to escape virus neutralising antibodies. A

mutation leading to the replacement of a single amino acid residue in the regions of peak hydrophilicity may establish a variant strain (Heine *et al.*, 1991), while the replacement of four amino acid residues leading to the loss of hydrophilicity, may generate a new serotype (Schnitzler *et al.*, 1993).

In a study by Xia *et al.* (2008)**,** aimed at identifying the virulence of a Chinese strain, Harbin-1, potential virulence determinants in both segments A and B were identified when the sequence was compared with those of 15 other IBDV strains. Nineteen of these residues were found to be in the polyprotein, eight in VP1 and four in VP5. The present study used a similar approach to determine the genetic relatedness of SA-KZN95, even though the sequence analysis was limited to the polyprotein only. The present study identified 21 amino acid residues in the polyprotein that are unique to the different pathotypes of IBDV and as a result they may be used for virus identification. Thirteen of these critical amino acid residues are found in the hypervariable region of VP2, five in VP4 and three in VP3. More variation in these residues was observed in the VP2 region than in either VP3 or VP4. The vvIBDV strains seem to be distinguished from other strains by residues Ala^{222}, Ile^{242}, Ile^{256}, Ile^{294} and Ser^{299}, in place of Pro/Thr^{222}, Val^{242}, Glu/Val^{256}, Leu^{294} and Asn^{299} in other strains. At position 222, the vvIBDV strains have Ala, the attenuated, classic and serotype II strains have Pro, while variants have Thr or Gln. In addition, variant strains were also distinguished by Lys^{249} in place of Gln^{249} in other strains. While changes at position 222 have been shown to induce antigenic differences, the Q249K mutation in the variants enables them to escape neutralisation (Lim *et al.*, 1999),

SA-KZN95 was similar to the vvIBDV strains, particularly the Malaysian UPM94/273 strain. SA-KZN95 and UPM94/273 were different from other vvIBDV strains at position 254, where they have a Ser found in this position in serotype II viruses and US variant strains. At this position, all other vvIBDV and classic strains have a Gly. The G254S change has been reported to distinguish variants from standard isolates (Mundt, 1999). Jackwood and Sommer-Wagner (2007) assessed the genetic variability among IBDV strains associated with high mortality, including some South African isolates. They distinguished two groups, one with vvIBDV characteristics (Ala^{222}, Ile^{242}, Ile^{256}, Ile^{294} and Ser^{299}) and the other with variant characteristics (Ala^{222}, Val/Ile^{242}, Val^{256}, Leu^{294} and Asn^{299}). They concluded that more data would be required to determine if these viruses are similar to US variants. Unfortunately, residue 254, which sets SA-KZN apart from other vvIBDV strains, was not analysed in their study; therefore no further comparisons between their South African isolates and SA-KZN95 could be made.

Very virulent strains are unable to propagate directly in tissue culture and attenuation by repeated passages of vvIBDV in CEFs (Yamaguchi *et al.*, 1996a; Yamaguchi *et al.*, 1996b) and Vero cells (Kwon *et al.*, 2000) was necessary for tissue culture adaptation. There is, however, one study which showed a vvIBDV propagated in LSCC-BK3 cells without adaptation (Setiyono *et al.*, 2001). Attenuated strains are distinguished by residues His^{253} and Ala^{284} in place of Gln^{253} and Thr^{284} found in the other strains. The H253Q mutation was shown to increase the virulence of an attenuated virus (Jackwood *et al.*, 2008). In agreement with this, the Q253H and A284T mutations of the variant E strain were necessary for its ability to infect chicken embryonic cells. By contrast, an A284T mutation in the GIs strain was sufficient for the alteration of tissue tropism (Mundt, 1999). Therefore, the Q253H and A284T mutations are associated with the ability of the bursa-derived strain to infect tissue culture cells (Mundt, 1999). Garriga *et al.* (2006) also showed that the exposed capsid residues, His^{253} and Thr^{284} are involved in particle-particle interactions. Attenuated strains also have Arg^{330} while all other strains have Ser^{330} residue in this position and may thus also play some role in the attenuated phenotype. The residues of the variable region of VP2 are associated with antigenicity and tissue tropism.

To date the molecular basis of the very virulent phenotype is not well understood. VP4 protease contains five of these critical residues: Tyr^{680}, Asn^{681}, Ser^{715} and Asp^{751} in vvIBDV strains and Cys^{680}, Lys^{681}, Pro^{715} and His^{751} in the other strains. A recent study showed that an anti-VP4 monoclonal antibody can be used to discriminate a pathogenic infection using a vvIBDV strain and a non-pathogenic infection using an attenuated strain (Wang *et al.*, 2009). It would be interesting to know if the failure of the monoclonal antibody to recognise VP4 in the non-pathogenic IBDV infected bursa is associated with the unique residues found in the VP4 region. Three of the critical residues were found in the VP3 region. Attenuated strains have Leu at position 981 that distinguish them from other strains that have a Pro in this position. The last two residues (Val^{990} and Ala^{1005}) distinguished the vvIBDV strains from other strains. These residues are outside the identified VP3 epitopes (Jagadish and Azad, 1991; Deng *et al.*, 2007). There is no data that shows the importance of these residues, however, it would be also interesting to know if these residues contribute to the very virulent phenotype. This can be achieved by constructing recombinant chimeric viruses with the different residues at these positions, followed by inoculation of chickens and analysing the bursa for pathological lesions (Brandt *et al.,* 2001).

In conclusion, the six polyprotein and VP4 fragments were successfully amplified from viral dsRNA, cloned into T-vectors and expression vectors, and sequenced. Sequence analysis of the polyprotein coding region based on the critical residues indicated that SA-KZN95 is undoubtedly a vvIBDV strain. Its critical residues were identical to those of the Malaysian UPM94/273. In a phylogenetic tree based on the nucleotide sequence, it formed a distinct clade with the Chinese vvIBDV strains Gx, GZ/96 and Harbin-1. However, at amino acid level, SA-KZN95 was among six vvIBDV strains that were not part of a clade therefore appearing to be a unique IBDV strain.

The designed constructs will be analysed for expression in *E. coli* and *P. pastoris*. Since the IBDV polyprotein was reported to be autocatalytically processed by the viral-encoded VP4 protease during its expression in *E. coli* (Hudson *et al.*, 1986; Jagadish *et al.*, 1988), VP4 peptides will also be designed for production of anti-VP4 peptide antibodies which will be used to detect the resulting cleavage products (Chapter 3).

CHAPTER 3
Characterisation of the expression of the embedded and mature VP4 protease by anti-peptide antibodies

3.1 Introduction

Polyprotein processing is a major strategy used by viruses to release their structural and non-structural proteins. Infectious bursal disease virus (IBDV) polyprotein (NH_2-VP2-VP4-VP3-COOH) is processed into two capsid proteins, VP2 and VP3, by the viral protease VP4 (Sanchez and Rodriguez, 1999). It has also been reported that the processing is an autocatalytic reaction and is co-translationally mediated (Hudson *et al.*, 1986; Jagadish *et al.*, 1988). IBDV VP4 is a serine protease that utilises a Ser/Lys catalytic dyad instead of the Ser/Asp/His triad (Lejal *et al.*, 2000). IBDV VP4 cleavage sites comprise a (Thr/Ala)-X-Ala↓Ala motif (Da Costa *et al.*, 2003), with the P_1 and P_1' residues conserved as alanyl residues, whereas the P_3 alaninyl residue can be substituted for Thr (Lejal *et al.*, 2000).

The polyprotein was shown to undergo processing at ^{512}Ala-Ala^{513} (pVP2-VP4 junction) and ^{755}Ala-Ala^{756} (VP4-VP3 junction), to yield pVP2, VP4 and VP3 (Lejal *et al.*, 2000; Sanchez and Rodriguez, 1999). The precursor VP2 (pVP2) undergoes secondary processing at ^{441}Ala-Phe^{442}, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502}. This secondary processing starts at the C-terminus and progresses towards the N-terminus to release VP2 (Met^{1}-Ala^{441}), and four structural peptides [pep46 (Phe^{442}-Ala^{487}), pep7a (Ala^{488}-Ala^{494}), pep7b (Ala^{495}-Ala^{501}) and pep11 (Ala^{502}-Ala^{512})] (Da Costa *et al.*, 2002). It was previously suggested that this secondary processing is carried out by the released VP4, based on the presence of dialaninyl peptides at the cleavage sites (Sanchez and Rodriguez, 1999). Irigoyen *et al.* (2009) demonstrated that pVP2 has autocatalytic endopeptidase activity that cleaves at the Ala^{441}-Phe^{442} site to release VP2 and pep70 (Phe^{442}-Ala^{512}).

Despite the critical role that polyprotein processing plays in capsid assembly and virus replication, some of the cleavage events are still not fully understood. The current model of the processing of IBDV polyprotein suggests that some cleavages occur in *cis*, implying an intramolecular reaction mechanism, and others are mediated in *trans*. For instance, cleavages at ^{441}Ala-Phe^{442}, ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} seem to occur in *cis* whereas other cleavages at the ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502} sites occur in *trans*. However, it is not known what determines the intra- and intermolecular activities. VP4 exists in two forms, the embedded and mature form. The embedded form exists as an

integral part of the polyprotein and it autocatalytically cleaves the ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} peptide bonds to produce the mature form (Ala^{513}-Ala^{755}). It has been shown that the mature form could intermolecularly cleave a VP4-VP3 junction (Lejal *et al*., 2000) indicating that it is active. Therefore, in IBDV, it seems that the processing generates a mature active protease from an autocatalytically active VP4 embedded in the polyprotein. What still remains unknown is the stretch of amino acids within the polyprotein that makes up the embedded protease**.** In addition, processing of the polyprotein at one site may be a prerequisite for processing at another site (Bartenschlager *et al.*, 1995). Such details regarding the processing of the IBDV polyprotein have not been described thus far.

In an attempt to study the two forms of VP4 and to shed more light on the polyprotein processing, different constructs of the IBDV polyprotein and the VP4 coding region were constructed (Fig. 3.1). The constructs vary with respect to the residues to the N- and C-termini of Ala^{513} and Ala^{755}, respectively, i.e. full-length polyprotein (Met^{1}-Glu^{1012}), truncated polyprotein (Ile^{227}-Trp^{891}), VP4-RA (Arg^{453}-Ala^{755}), VP4-RK (Arg^{453}-Lys^{722}), VP4-AW (Ala^{513}-Trp^{891}) and VP4-AA (Ala^{513}-Ala^{755}). The polyprotein constructs would represent the fully embedded VP4 protease. With cleavage occurring at ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756}, VP4-RA and VP4-RK would represent the N-terminally embedded protease while VP4-AW represents a C-terminally embedded form. In contrast, VP4-AA is the mature form (Figure 3.1).

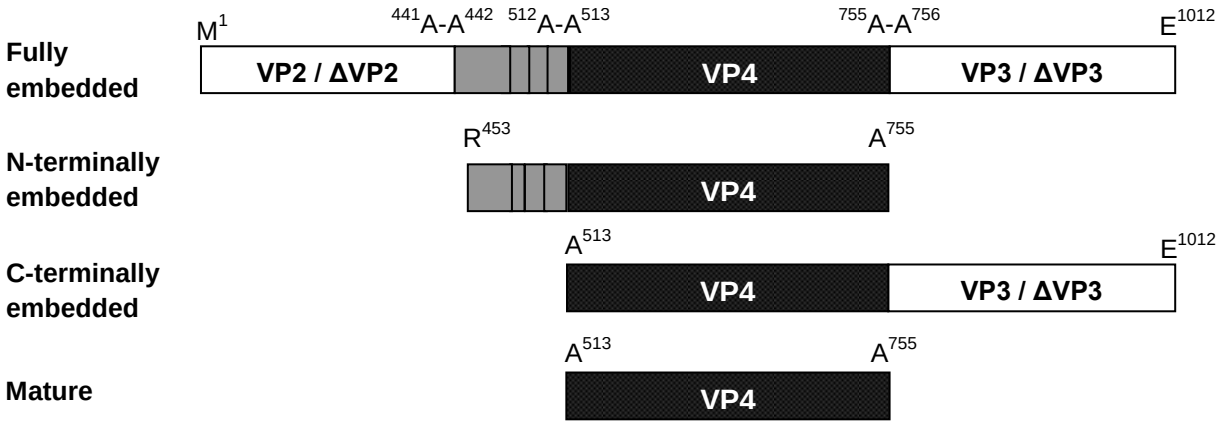


Figure 3.1 Schematic diagrams of the different forms of VP4 expressed in the present study. The embedded form differs from the mature form with respect to residues at the N- and C-terminus of Ala^{513} and Ala^{755}, respectively.

The inclusion of the Arg^{453}-Ala^{512} polypeptide in the N-terminally embedded constructs was motivated by the findings and suggestions from previous studies on VP4 protease. Prior to the identification of the vicinal dialanines, ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} as

cleavage sites of the pVP2-VP4 and VP4-VP3 junctions, respectively (Sanchez and Rodriguez, 1999), the VP4 amino acid sequence was thought to start at Arg^{453} and end at Lys^{722} in the IBDV polyprotein (Hudson *et al.*, 1986). We then speculated that VP4, as an integral part of the polyprotein, may not necessarily be the same as the one released after the processing. In other words, VP4-RK could be the protease responsible for cleavage at the ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} to release the VP4-AA (mature form).

Previous studies on the characterisation of the IBDV polyprotein processing were conducted by expression of the polyprotein in *E. coli* (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Lejal *et al.*, 2000) and mammalian systems such as African green monkey kidney epithelial BSC-1 cells (Sanchez and Rodriguez, 1999). In these studies, the processing was analysed by reducing SDS**-**PAGE and autoradiography or western blotting using anti-VP2 or -VP3 antibodies. In the present study, the designed constructs were also expressed in *E. coli*. Even though the IBDV structural proteins were shown not to be glycosylated (Kibenge *et al.,* 1997), the IBDV polyprotein sequence contains three potential N-glycosylation sites in the VP2 region and for that reason the expression was also attempted in the eukaryotic *P. pastoris.* In addition, anti-VP4 peptide antibodies were used to characterise the processing of the designed constructs. The use of anti-VP2 and anti-VP3 antibodies in previous studies could only identify VP2 and VP3, respectively. The idea was to produce anti-peptides antibodies that would distinguish the different cleavage products around the VP4 region. The antibodies were also expected to reveal the products that may not have been possible to recognise with the anti-VP2 or anti-VP3 antibodies.

Due to our speculation that VP4 may, as an integral part of the polyprotein not necessarily be the same as the mature form of VP4-AA, the VP4-RA (Arg^{453}-Ala^{755}) sequence was used for peptide selection. The preferred location of the selected peptides was within the VP4-AA sequence and the Arg^{453}-Ala^{512} polypeptide. It was also of interest to compare the expression profile of VP4-RK (Hudson *et al.*, 1986) and VP4-AA (Sanchez and Rodriguez, 1999), thus a further peptide within Lys^{722}-Ala^{755} was considered necessary. The antibodies against the peptide would be able to differentiate VP4-RK and VP4-AA during expression.

The constructs previously cloned into T-vectors and sub-cloned into various expression vectors (see Chapter 2) were expressed and the cleavage products resulting from the processing were detected by the anti-VP4 peptide antibodies. The peptide design and the production of anti-peptide antibodies in chickens are also described. The mature form of

VP4 was purified and characterised enzymatically. Attempts to purify the recombinant polyprotein are also described. Wang *et al.* (2009) used anti-VP4 monoclonal antibody to detect native VP4 in the cortex of lymphoid follicles of the bursa of Fabricius infected with a pathogenic IBDV strain. In the present study, the anti-VP4 peptide antibodies were also used to determine whether native VP4 could be detected in IBDV-infected bursal homogenates. The proteolytic activity of native VP4 in the infected bursal samples was also assessed on a gelatine-containing SDS-PAGE.

3.2 Materials

Recombinant protein expression: The antibiotics (chloramphenicol, kanamycin and ampicillin), guanidine-HCl, isopropyl thioglucopyranoside (IPTG) were obtained from Sigma-Aldrich-Fluka (Steinheim, Germany). The recombinant constructs in *E. coli* based expression vectors (pET-32b, pET-28b and pGEX-4T-1) and *P. pastoris* based pPIC9 were all cloned in the present study as described in Chapter 2. SacI was purchased from Fermentas Life Sciences (Vilinius, Lithuania).

Antibody preparation: The peptides chosen for antibody production were synthesised by Auspep (Parkville, Australia). Freund's complete and incomplete adjuvants, Sephadex® G-10 and Sephadex® G-25 resins, maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS), rabbit albumin and Ellman's reagent were purchased from Sigma-Aldrich-Fluka (Steinheim, Germany). Nunc Maxi Sorp™ 96-well enzyme-linked immunosorbent assay (ELISA) plates were from Nunc products (Roskilde, Denmark). Aminolink® coupling gel was purchased from Pierce Perbio Science (Erembodegem, Belgium).

Peptide substrates: The peptide substrates H-Val-Leu-Lys-7-amino-4-methylcoumarin (AMC), H-Ala-AMC, H-Ala-Phe-Lys-AMC, butyloxycarbonyl (Boc)-Leu-Arg-Arg-AMC and Boc-Leu-Gly-Arg-AMC were obtained from Bachem (King of Prussia, PA, USA). Benzyloxycarbonyl (Z)-Phe-Arg-AMC, Z-Gly-Pro-Arg-AMC, Z-Ala-Arg-Arg-AMC were obtained from Sigma (Munich, Germany). Z-Gln-Ala-Ala-AMC was purchased from GL Biochem (Shanghai) Ltd (China). Nunc Black 96-well plates were obtained from Nunc products (Roskilde, Denmark).

General Biochemistry: Bovine serum albumin (BSA) was purchased from Roche (Mannheim, Germany), polyethylene glycol (PEG) 6000 from Merck Biosciences (Damstadt, Germany) and 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) or ABTS was purchased from Boehringer (Mannheim, Germany), Anti-thrombin III and

glutathione-agarose beads from Sigma-Aldrich-Fluka (Steinheim, Germany) and PageRuler™ Plus Prestained Protein Ladder from Fermentas Life Sciences (Vilinius Lithuania). TALON® metal affinity resin was from Clontech (California, USA) whereas Ni-NTA His-Bind resin together with anti-His tag monoclonal antibody, enterokinase and thrombin were purchased from Novagen (Damstadt, Germany). BioTrace™ NT nitrocellulose and PVDF membranes were purchased from PALL Corporation (New York, USA). Amicon centricon centrifugal filters were from Millipore (Billerica, USA). Infected bursal tissues were obtained from Regional Veterinary Laboratory (Pietermaritzburg, South Africa) and fresh non-infected bursae harvested from broilers were obtained from the University of KwaZulu-Natal Ukulinga research and training farm. Inhibitors [*trans*-epoxysuccinyl-L-leucyl-amido(4-guanidino)butane (E-64), pepstatin A, soybean trypsin inhibitor (SBTI), N-tosyl-L-phenylalanine chloromethyl ketone (TPCK) and 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF)] were purchased from Sigma-Aldrich. BCA™ protein assay kit was obtained from Pierce (Belgium). Other general chemicals and reagents were of the highest purity available and obtained from Sigma-Aldrich-Fluka (Steinheim, Germany) and Merck Biosciences (Damstadt, Germany).

3.3 Methods

3.3.1 Anti-VP4 peptide antibody production

3.3.1.1 Selection and synthesis of peptides

Selection of the VP4 peptides was accomplished using the Predict7 program which analyses protein structural features using the following algorithms: hydrophilicity, surface probability, flexibility and antigenicity (Cármenes *et al.*, 1989). It was preferable that the selected peptides be located on the surface of VP4 to increase the accessibility of the corresponding antibodies. Due to the unavailability of the IBDV VP4 3-D structure, available structures of VP4 from other birnaviruses had to be used. The IBDV VP4-AA (Ala^{513}- Ala^{755}) equivalent has been crystallised in BSNV and IPNV (Feldman *et al.*, 2006; Lee *et al.*, 2007). A sequence alignment of BSNV VP4 with three other birnavirus proteases showed that BSNV VP4 is 23% identical in sequence to IBDV VP4, 18% identical to IPNV and DXV VP4 (Da Costa *et al.*, 2003). There are a number of amino acid deletions in the VP4 from IBDV, IPNV and DXV in comparison to the BSNV VP4 (Feldman *et al.*, 2006). Therefore the alignment of VP4 sequences, schematic presentation of the topology and 3-D structure of BSNV VP4 viewed in Cn3D (http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml) were used to predict whether

the selected peptides would be on the inside or outside of the IBDV VP4 structure (Fig. 3.2).

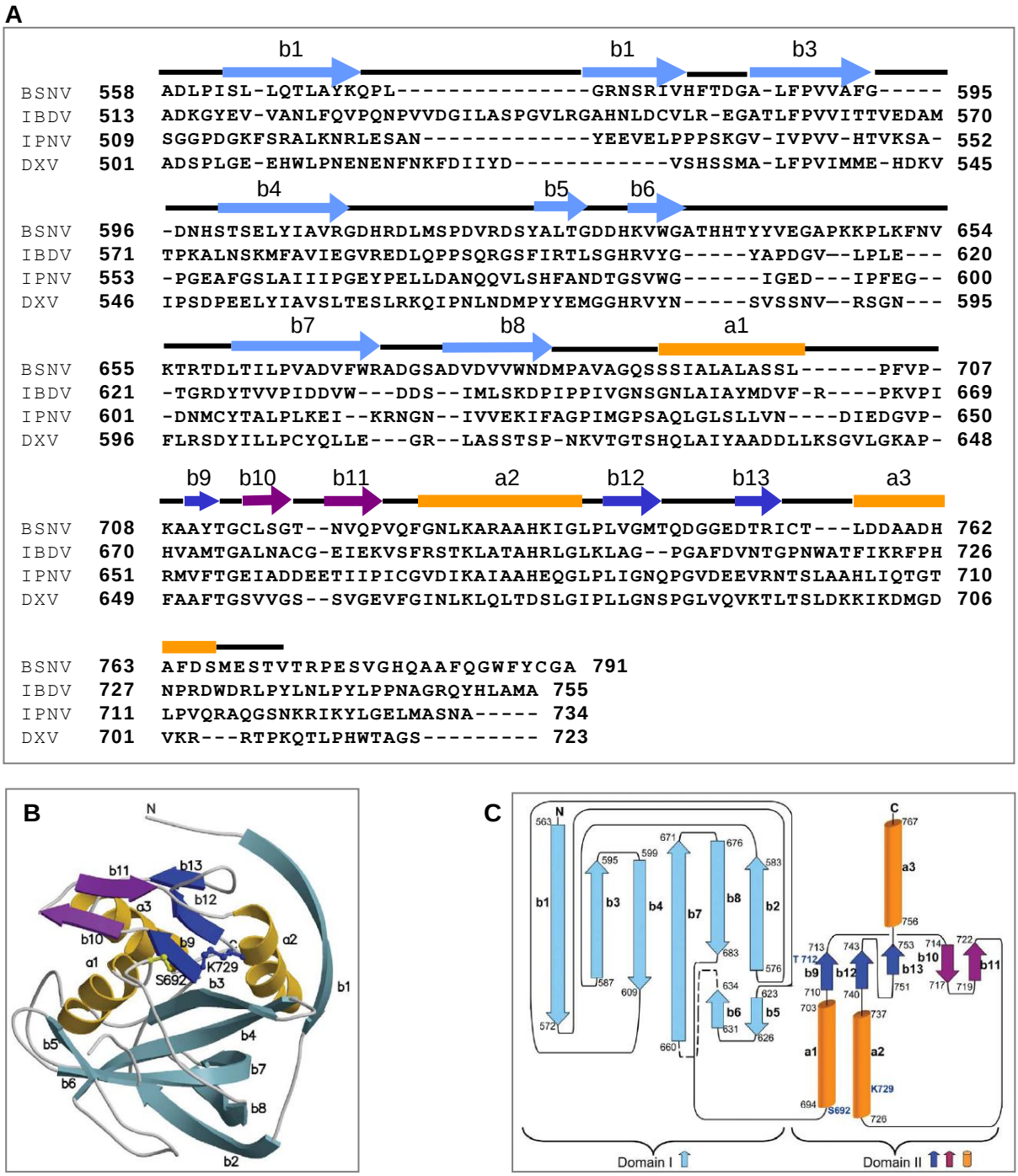


Figure 3.2 VP4 sequence alignment and structure. A) The alignment of the VP4 amino acid sequences from BSNV, IBDV, IPNV and DXV. **B)** Ribbon structure of BSNV VP4. **C)** Schematic diagram of the protein topology of VP4. The secondary structure and numbering of BSNV is shown above the alignment (Panel A). The N-terminal antiparallel β-sheet domain (domain I) is in light blue. The C-terminal α/β domain is shown in dark blue (three-stranded parallel β-sheet), purple (β hairpin), and gold (α-helices). The α-helices and β-sheets are numbered in order. The coloring is the same for the alignment, structure and topology (Feldman *et al.*, 2006).

Three 13-mer peptides were selected from the VP4-RA (Arg^{453}-Ala^{755}) sequence for raising chicken antibodies using the Predict7 program (Cármenes *et al.*, 1989). Selection of peptides was based on the high values for hydrophilicity, surface probability, flexibility and antigenicity. The selected peptides were modified before synthesis by the addition of an extra cysteine residue to the N- or C-terminus, to allow for coupling with MBS. Peptides were synthesised to greater than 70% purity as determined by HPLC (Auspep, Australia).

3.3.1.2 Coupling of peptides to rabbit albumin

The selected peptides were prepared for immunisation by coupling to rabbit albumin through the cysteine residue using MBS, which links the SH-group of a cysteine side chain to the ε-amino group of a lysine residue on the carrier protein (Peeters *et al.*, 1989). The peptides were coupled to rabbit albumin carrier at a peptide:carrier molar ratio of 40:1 (Briand *et al.*, 1985). Rabbit albumin was resuspended in 500 µl phosphate buffered saline (PBS, pH 7.4). MBS (1 mg) was resuspended in 200 µl DMF and made up to 500 µl with PBS. The rabbit albumin-MBS mixture was incubated with stirring at RT for 30 min before applying to a Sephadex® G-25 column. The MBS activated rabbit albumin was eluted with buffer A [100 mM sodium phosphate buffer pH 7, 0.02% (w/v) NaN_3] and fractions with $A_{280} > 0.5$ were pooled.

Previous studies in this laboratory determined that 4 mg of a synthetic peptide is sufficient to produce a high antibody titre in two chickens within 7-12 weeks after four immunisations (Mkhize, 2003). Peptide (4 mg) was dissolved in DMSO (50 µl) and the volume made up to 500 µl with peptide reducing buffer [100 mM Tris-HCl buffer pH 7, 1mM Na_2-EDTA, 0.02% (w/v) NaN_3]. DTT (10 mM) was added to make a final volume of 1 ml and the solution was incubated for 1 h at 37°C. The reduced peptide was separated from free reducing agent and uncoupled peptide on a Sephadex® G-10 column (15 x 110 mm, 10 ml/h). Fractions (500 µl) were collected and 10 µl of each fraction was analysed for the presence of free sulfhydryl groups using an equal volume of Ellman's reagent [0.004% (w/v) 5'5 Dithiobis (2-nitro-benzoic acid), 100 mM Tris-HCl buffer pH 8, 10 mM Na_2-EDTA, 0.1% (w/v) SDS]. The pooled MBS-activated rabbit albumin samples were mixed with the pooled reduced peptide solution and incubated at RT for 3 h with stirring. The solution was equally aliquoted into four microfuge tubes and stored at -20°C for immunisation.

3.3.1.3 Immunisations and IgY isolation

The study utilised Hy-Line Brown hens free of IBDV infections aged 25 to 40 weeks housed and cared for at the Ukulinga research and training farm (University of KwaZulu-Natal, Pietermaritzburg). Animal ethics clearance was obtained from the University of KwaZulu-Natal animal ethics committee (reference number 036/09/Animal). Each of the three peptide-rabbit albumin conjugates (200 µg) was mixed with Freund's complete adjuvant in a 1:1 (v/v) ratio, and triturated to give a stable water-in-oil emulsion. Two chickens were immunised per peptide-conjugate by intramuscular injections at two sites on either side of the breast bone. Booster injections were given at weeks 2, 4, and 6 using Freund's incomplete adjuvant. Eggs were collected prior to and throughout the immunisation period. IgY was isolated from chicken egg yolk according to Polson *et al.* (1985) with minor modifications (Goldring and Coetzer, 2003). Briefly, yolks were separated from egg white, and excess egg white removed by careful washing under running tap water. The yolk was dispersed in two volumes of 100 mM phosphate buffer, pH 7.6 and 3.5% (w/v) PEG 6000 added. The mixture was centrifuged (4420 x g, 20 min, RT) and filtered through a non-absorbing cotton wool. PEG 6000 [8.5% (w/v)] was added to the filtrate, dissolved and centrifuged (12000 x g, 10 min, RT). The pellet was dissolved in a volume of phosphate buffer equal to the original egg yolk volume and 12% (w/v) PEG 6000 added and mixed. The sample was centrifuged (12 000 x g, 10 min, RT) and the final pellet resuspended in of the original egg yolk volume of phosphate buffer. IgY concentration was determined by measuring A_{280} using the extinction coefficient = 1.25 (Goldring and Coetzer, 2003).

3.3.1.4 Affinity purification of anti-VP4 peptide antibodies

A peptide affinity matrix was prepared by coupling 7.5 mg of peptide to 1.5 ml of Pierce AminoLink® support according to the manufacturer's instructions. Antibodies isolated from chicken egg yolk were circulated overnight over the affinity matrix in a reverse direction at RT to ensure continuous flow. The column was washed with 20 column volumes of 100 mM sodium phosphate buffer, pH 6.5 containing 0.02% (w/v) NaN_3 and eluted with 100 mM glycine-HCl buffer, pH 2.8 into microfuge tubes containing 100 µl of neutralisation buffer (1 M sodium phosphate buffer, pH 8.5). Each 1 ml fraction was immediately mixed by inversion to neutralise the pH. The elution was monitored by measuring A_{280} of each fraction and pooled fractions were stored at 4°C with the addition of NaN_3 to a final concentration of 0.1% (w/v).

3.3.1.5 ELISA to monitor the production and purification of anti-peptide antibodies

The production of anti-peptide antibodies was monitored using ELISA over a 13 week period after the first immunisation. ELISA was also used to characterise the affinity purified anti-peptide antibodies. Microtitre plate wells were coated overnight at 4°C with 150 µl of 5 µg peptide or recombinant protein in PBS. Unoccupied sites in wells were blocked with 200 µl blocking buffer [0.5% (w/v) BSA in PBS] for 1 h at 37°C and washed three times with 0.1% (v/v) PBS-Tween 20. Chicken IgY isolated from egg yolk at the individual time points after immunisation (50 µg/ml) or affinity purified anti-peptide antibodies (1-50 µg/ml) diluted in blocking buffer were incubated for 2 h at 37°C and excess IgY washed out of the wells with 0.1% (v/v) PBS-Tween 20. Rabbit anti-chicken IgY-horseradish peroxidase conjugate in 0.5% (w/v) BSA-PBS (1:12000 dilution) was added and incubated for 1 h at 37°C. After washing as described, ABTS/H_2O_2 chromogen/ substrate solution [0.05% (w/v) ABTS in citrate-phosphate buffer pH 5, 0.0015% (v/v) H_2O_2 in PBS] was added and the plate incubated in the dark for 15 min prior to reading the absorbance values at 405 nm in a VersaMax™ ELISA plate reader using SOFTmax PRO software version 3.1, both purchased from Molecular Devices Corporation (California, USA).

3.3.2 Recombinant expression of the IBDV polyprotein and VP4 constructs

3.3.2.1 Expression using the pGEX system in E. coli

Recombinant pGEX-4T-1 plasmids were transformed using TransformAid™ bacterial transformation kit into *E. coli* BL21 cells for expression. A single colony of BL21 cells carrying recombinant pGEX-4T-1 was grown in 10 ml 2xYT-Amp broth and the culture was incubated overnight at 37°C. The overnight culture was diluted 1/100 with fresh 2xYT-Amp and incubated at 37°C until the OD_{600} was between 0.4 and 0.6. The culture was induced for expression by the addition of 1 mM IPTG and incubation continued at 37°C for another 4 h. The culture medium was continuously supplemented with 50 µg/ml ampicillin every hour during expression. The cells were harvested by centrifugation (5000 x g, 10 min, 4°C) and the cell pellet was resuspended into 1/20 culture volume of lysis buffer (20 mM Tris-HCl buffer pH 7.6, 100 mM NaCl, 5% (v/v) glycerol containing 1 mg/ml lysozyme) and frozen overnight at -20°C. The thawed samples were disrupted by sonication using the Virosonic™ cell disruptor (VirTis, New York, USA), analysed by reducing SDS-PAGE (Laemmli, 1970) and western blotting (Towbin *et al.*, 1979) (Section 3.3.3).

3.3.2.2 Expression using the pET system in E. coli

Both pET-32b and pET-28b recombinant plasmids were transformed into *E. coli* BL21 (DE3) and BL21 DE3 (pLysS) cells using TransformAid™ bacterial transformation kit. *E. coli* BL21 (DE3) cells were grown in medium containing 50 µg/ml ampicillin for pET-32b and 34 µg/ml kanamycin for pET-28b. For the BL21 (DE3) pLysS cells, the medium was also supplemented with chloramphenicol (30 µg/ml). Expression was carried out by both self-induction and IPTG induction. For self-induction, a single colony was grown overnight 37°C or at 16°C for 24 h in Terrific Broth [1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 0.017 M KH_2PO_4, 0.69 M K_2HPO_4] containing appropriate antibiotics in a shaking incubator. For expression by IPTG induction, a 2xYT-Amp overnight culture was diluted 1/100 in fresh medium and allowed to grow at 37°C. When the OD_{600} was between 0.4 and 0.6, the expression was induced by the addition of IPTG at a final concentration of 1 mM and the cultures (induced and uninduced) were further incubated for 4 h. The cells were harvested by centrifugation (5000 x g, 10 min, 4°C) and the cell pellet was resuspended into 1/20 culture volume of lysis buffer and frozen overnight at -20°C. The thawed samples were disrupted by sonication, analysed by reducing SDS-PAGE and western blotting (Section 3.3.3).

3.3.2.3 Expression in P. pastoris

Transformation into *P. pastoris* GS115 cells was carried out in two steps. Firstly, recombinant pPIC9 plasmids were linearised by *Sac*I to improve the efficiency of integration of the insert into the *P. pastoris* genome by homologous recombination. In the second step, recombinant pPIC9 plasmids were transformed into competent *P. pastoris* strain GS115 cells prepared according to Wu and Letchworth (2004), to select zeocin-resistant transformants. Briefly, *P. pastoris* cells were grown in 500 ml YPD [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose] until they reached an OD_{600} between 1 and 2. The cells were pelleted (2000 x g, 20 min, 4°C) and resuspended in 400 ml of 10 mM Tris-HCl buffer pH 7.5, containing 100 mM lithium acetate, 10 mM DTT, 0.6 M sorbitol, and incubated for 30 min at RT. The cells were subsequently washed with 1 M sorbitol (75 ml), and finally resuspended in 1.5 ml of 1 M sorbitol.

Linear plasmid DNA (200 ng) was used to transform competent GS115 yeast cells (200 µl) by electroporation (1.5 kV, 25 µF, 186 Ω) using a BioRad Gene Pulser™ in pre-chilled 2 mm Genepulser/micropulser electroporation cuvettes (BioRad, California, USA). The cells were immediately resuspended in chilled 1 M sorbitol (1 ml) and concentrated by

centrifugation (2000 x g, 10 s, RT) prior to plating on MD plates [1.34% (w/v) yeast nitrogen base without amino acids, 0.0004% (w/v) biotin, 2% (w/v) glucose, 15 g/l bacteriological agar] containing ampicillin (50 µg/ml). The plates were incubated at 30°C until the colonies were visible. Colony PCR using the AOX primers (Appendix 2) was used to screen for recombinants (Ayra-Pardo *et al.*, 1998).

A single colony was used to inoculate liquid YPD (50 ml) containing 50 µg/ml ampicillin and grown at 30°C in an orbital shaking incubator for 2-3 days. The culture (10 ml) was used to inoculate 100 ml of BMGY medium [1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate buffer, pH 6.5, and 1.34% (w/v) YNB] at 30°C for 16-20 h. For induction of protein expression, the cells were pelleted (6000 x g, 10 min, RT) and resuspended in 100 ml buffered minimal medium [BMM, 100 mM potassium phosphate buffer, pH 6.5, 1.34% (w/v) YNB, 0.0004% (w/v) biotin, and 5% (v/v) methanol]. The culture was grown in a baffled flask at 30°C in a shaking incubator and 0.5% (v/v) methanol was added daily during the 4-6 days expression period. Cells were pelleted (6000 x g, 10 min, RT) and the supernatant was analysed by SDS-PAGE with silver staining and western blotting (Section 3.3.3).

3.3.3 SDS-PAGE and western blotting

Reducing SDS-PAGE was performed as described by Laemmli (1970) using the BioRad Mini Protean III® vertical slab electrophoresis apparatus. For the preparation of non-reduced protein samples, half the total volume of the non-reducing treatment buffer [125 mM Tris-HCl buffer pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol] was added to the protein sample and vortexed to mix. For the preparation of reduced protein samples, equal volume of reducing treatment buffer [125 mM Tris-HCl buffer pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol] was added to the protein samples, the mixture was vortexed and boiled at 100°C for 5 min. Electrophoresis was performed at 18 mA per gel until the bromophenol blue tracker dye was about 0.5 cm from the bottom edge of the gel. The gel was stained with Coomassie blue R-250 staining solution [0.125% (w/v) Coomassie blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid], destained and the electrophoregram was imaged using the BioRad VersaDoc™ imaging system. The standard low molecular weight protein marker (MWM) consisting of phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (21 kDa) and α-lactalbumin (14.4 kDa) was prepared in-house.

Following separation of the bacterial lysate proteins or infected bursal homogenates (Section 3.3.6) by reducing SDS-PAGE, the proteins were electrotransferred onto a nitrocellulose membrane for western blotting. The nitrocellulose was transiently stained for 3 min with Ponceau S staining solution [0.1% (w/v) Ponceau S in 1% (v/v) glacial acetic acid] to visualise the bands for marking positions of molecular weight markers with pencil (when a non-rainbow protein marker was used) and destained with dH_2O. The non-specific sites on the nitrocellulose membrane were blocked with 5% (w/v) low fat milk in TBS (20 mM Tris-HCl buffer pH 7.4, 200 mM NaCl) for 1 h. The nitrocellulose was washed with TBS (3 x 5 min) and incubated for 2 h with primary antibody [1:2000 Novagen anti-His tag monoclonal antibody or 2 µg/ml mouse monoclonal anti-GST tag IgG (Anatech, South Africa) or 1 µg/ml chicken anti-VP4-peptide antibodies] made up in 0.5% (w/v) BSA-TBS.

Following washing with TBS (3 x 5 min), the nitrocellulose membrane was incubated for 1 h with a horse peroxidase-conjugated rabbit anti-IgY antibody or goat anti-mouse IgG (Jackson Immunochemicals, Pennsylvania, USA). The nitrocellulose membrane was washed with TBS (3 x 5 min) and incubated in chromogen/substrate solution [0.06% (w/v) 4-chloro-1-naphthol and 0.0015% (v/v) H_2O_2] in the dark until bands were clearly visible.

3.3.4 Solubilisation, refolding and purification of full-length polyprotein

The bacterial cell lysate was centrifuged (14000 x g, 30 min, 4°C), the supernatant discarded and the inclusion bodies solubilised and polyprotein refolded using the protocol described by Sijwali *et al*. (2001) with minor modifications. Firstly, the inclusion bodies were purified from the crude lysate by washing the pellet with 20 mM Tris-HCl buffer pH 8, 2.5% (v/v) Triton X-100, 2 M urea, followed by subsequent washes with 20 mM Tris-HCl buffer pH 8, 20% (w/v) sucrose. The inclusion bodies were pelleted (14000 x g, 30 min, 4°C) and solubilised in 20 mM Tris-HCl buffer pH 8, 0.5 M NaCl containing 8 M urea or 6 M guanidium-HCl. The crude insoluble material was separated by centrifugation (14000 x g, 30 min, 4°C) from the solubilised protein.

For purification under denaturing conditions, a Ni-NTA agarose affinity column (1.5 ml) was prepared and equilibrated with 10 column volumes of solubilisation buffer. The solubilised protein sample was incubated overnight with the resin at 4°C. The resin was washed with 10 column volumes each of the solubilisation buffer and wash buffer (20 mM Tris-HCl buffer pH 8, 500 M NaCl, 8 M urea). Bound protein was eluted from the affinity matrix with elution buffer (20 mM Tris-HCl buffer pH 8, 1 M imidazole, 8 M urea) in 1 ml

fractions and A_{280} readings of each fraction measured. Collected fractions were analysed by reducing SDS-PAGE (Section 3.3.3). The purification was also carried out using sodium phosphate-NaCl-urea solutions as follows: equilibration solution (20 mM sodium phosphate, 500 mM NaCl, 8 M Urea, pH 7.8); wash solution (20 mM sodium phosphate, 500 mM NaCl, 8 M Urea, pH 6) and elution solution (20 mM sodium phosphate, 500 mM NaCl, 8 M Urea, pH 4).

Refolding of the solubilised full-length polyprotein sample was attempted using both dilution and dialysis. Refolding by dilution was performed by drop-wise dilution of the protein sample with refolding buffer (100 mM Tris-HCl buffer pH 9.0, 1 mM EDTA, 30% (v/v) glycerol, 250 mM L-arginine, 1 mM GSH, 1 mM GSSG) at 4°C. The sample was diluted a 100-fold in refolding buffer and concentrated by ultrafiltration using an Amicon PM 50 filter membrane (50 kDa cut-off). Refolding by dialysis was performed against a series of buffers as illustrated in Table 3.1.

Table 3.1 Refolding by dialysis of the full-length polyprotein

Buffer	Time (h)	Volume (ml)	Temperature	Number of changes	Urea (M)
Dialysis buffer 1 [4 M urea, 20 mM Tris-HCl pH 8, 250 mM L-arginine, 5% (v/v) glycerol]	6	250	RT	2	4
Dialysis buffer 2 [2 M urea, 20 mM Tris-HCl pH 8, 250 mM L-arginine, 10% (v/v) glycerol]	18	1000	RT	1	2
Dialysis buffer 3 [1 M urea, 20 mM Tris-HCl pH 8, 250 mM L-arginine, 15% (v/v) glycerol]	6	500	RT	2	1
Dialysis buffer 4 [0.5 M urea, 20 mM Tris-HCl pH 8, 250 mM L-arginine, 20% (v/v) glycerol]	6	1000	4°C	2	0.5
Equilibration buffer [20 mM Tris-HCl pH 8, 250 mM L-arginine, 0.1% (w/v) NaN_3, 20% (v/v) glycerol]	6	1000	4°C	3	0

3.3.5 Purification and enzymatic characterisation of recombinant VP4-AA

Lyophilised glutathione-agarose beads (0.290 g) were allowed to swell in deionised water overnight at 4°C and washed thoroughly with 10 column volumes of equilibration buffer (PBS, pH 7.4) to remove the lactose. The resin was deaerated, the slurry carefully poured into a column (0.8 x 4 cm) and the column equilibrated with 10 column volumes of equilibration buffer. The cell lysate was clarified by centrifugation (10000 x g, 10 min, 4°C)

before application to the column. All the steps were performed at gravitational flow rate. The clarified supernatant was incubated overnight with the resin at 4°C. The resin was washed with 10 column volumes of 1% (v/v) PBS-Triton X-100 at 4°C. The GST fusion protein was eluted from the resin with elution buffer (5 mM reduced glutathione, 50 mM Tris-HCl buffer pH 8) and the eluted fractions were analysed by SDS-PAGE (Section 3.3.3).

For on-column cleavage of the fusion protein, following washing after incubating the clarified supernatant with the resin, the column was equilibrated with 20 ml of thrombin cleavage buffer [20 mM Tris-HCl buffer pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$]. Thrombin cleavage buffer (3 ml) was used to resuspend the resin followed by the addition of thrombin (5 U). The column was incubated overnight at RT with gentle rocking. Cleaved recombinant VP4-AA present in the buffer was directly collected from the column in 1 ml fractions. A wash fraction of 3 column volumes was also collected. Bound GST and any uncleaved fusion protein were eluted from the resin with elution buffer containing 10 mM reduced glutathione. The eluted fractions were analysed by SDS-PAGE (Section 3.3.3) and quantified by the BCA™ protein assay kit (Pierce, Belgium).

Substrate gel analysis and peptide hydrolysis assay were performed to determine the enzyme activity of recombinant VP4-AA. Purified VP4-AA (10 µg) and thrombin (5 U) were also separately pre-incubated overnight with inhibitors (0.5 U anti-thrombin III, 5 µg/ml SBTI or 1 mM AEBSF) at 37°C in a total volume of 20 µl. Gelatine SDS-PAGE gel (Heussen and Dowdle, 1980) was prepared by copolymerising the running gel with 0.1% (w/v) gelatine. Prior to electrophoresis, recombinant protein samples were mixed with non-reducing treatment buffer [125 mM Tris-HCl buffer pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol]. Following electrophoresis, the gel was washed in three changes of 2.5% (v/v) Triton X-100 for 1 h at RT to remove SDS. The washed gel was incubated overnight in assay buffer [20 mM Tris-HCl buffer pH 7.4, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% (v/v) glycerol) and stained with amido black [0.1% (w/v) amido black, 30% (v/v) methanol, 10% (v/v) acetic acid, 60% (v/v) water] and destained with several changes of destaining solution [30% (v/v) methanol, 10% (v/v) acetic acid, 60% (v/v) water] until clear bands indicating proteolytic activity were visible.

The *trans* activity of VP4-AA was evaluated using mutant VP4-RK (purified under denaturing conditions or refolded non-purified samples) as substrates. Mutant VP4-RK sample (25 µl; approximately 10 µg) was mixed with equal volumes of VP4-AA purified sample (10 µg), VP4-AA purified sample (10 µg) pre-incubated with anti-thrombin III (0.5

U) or thrombin (5 U). Assay buffer (50 µl) was added to the enzyme-substrate mix (50 µl) and the mixture incubated overnight at 37°C. The proteolytic activity assay was also evaluated without adding the assay buffer i.e in the storage buffer, to determine if activity required particular conditions. Mutant VP4-RK purified under denaturing conditions was stored in 20 mM Tris-HCl buffer pH 8, 1 M imidazole, 8 M urea, the refolded mutant VP4-RK was stored in 20 mM Tris-HCl pH 8, 250 mM L-arginine, 0.1% (w/v) NaN_3, 20% (v/v) glycerol) and VP4-AA was stored in 50 mM Tris-HCl buffer pH 8 containing 5 mM reduced glutathione. For the assay without assay buffer, mutant VP4-RK sample (50 µl; approximately 10 µg) was mixed with equal volumes of VP4-AA purified sample (10 µg). About 25 µl was used in an enterokinase activity assay containing 3 µl of 10 x cleavage buffer (200 mM Tris-HCl buffer pH 7.4, 500 mM NaCl, 20 mM $CaCl_2$), 1 µl enterokinase (0.5 U) and 1 µl dH_2O. Incubation was continued for 16 h at RT. The cleavage of mutant VP4-RK was analysed by SDS-PAGE and western blotting using anti-VP4 common peptide antibodies (Section 3.3.3).

Peptide hydrolysis assays for recombinant VP4-AA were carried out using the following substrates: H-Val-Leu-Lys-AMC, H-Ala-Phe-Lys-AMC, H-Ala-AMC, Boc-Leu-Arg-Arg-AMC, Boc-Leu-Gly-Arg-AMC, Z-Phe-Arg-AMC, Z-Gly-Pro-Arg-AMC, Z-Ala-Arg-Arg-AMC, and Z-Gln-Ala-Ala-AMC. Briefly, enzyme [20 ng diluted in 0.1% (v/v) Brij-35] was incubated in assay buffer (20 mM Tris-HCl buffer pH 7.4, 100 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol, 10 mM DTT] in a volume of 75 µl and incubated for 10 min at 37°C. Substrate stocks were diluted to 20 µM (final concentrations) and 25 µl added to 75 µl of enzyme-buffer mix. For the enzyme pre-incubated with anti-thrombin III, the mix was prepared as before and an equivalent of 20 ng VP4-AA and 0.15 U thrombin were used for the assay. Fluorescence was measured (excitation 360 nm; absorbance 460 nm) for 100 cycles using a FLUOstar Optima spectrophotometer (BMG Labtech, Offenburg, Germany).

3.3.6 Gelatine-gel analysis of infected and non-infected bursal samples

An equal volume of homogenisation buffer (20 mM Tris-HCl buffer, pH 7.8) was added to the bursal tissue and homogenised using a Potter S homogeniser (Braun Biotech International GmbH, Germany). Homogenate (100 µl) was resuspended in 400 µl of lysis buffer (20 mM Tris-HCl buffer pH 7.2, 1% (v/v) Triton X-100) and used for protein concentration determination using a BCA™ protein assay kit. The homogenate sample was added to half the volume of non-reducing treatment buffer and 35 µl (containing about 10 µg of total protein) was loaded into each well of a 10% gelatine containing SDS-

PAGE gel (Section 3.3.5). Following washing, the gels were incubated in the assay buffer in the presence and absence of protease inhibitors (AEBSF, SBTI, pepstatin A, TPCK and E-64), before staining and destaining as before.

3.3.7 Electroelution of VP4-RA and VP4-RK cleavage products for N-terminal sequencing

Electroelution of proteins was carried out according to Acil *et al.* (1997) with some modifications. Briefly, separation of the lysate proteins was performed on 10% reducing SDS-PAGE. After separation by electrophoresis, bands were visualised by staining with 0.5% (w/v) Coomassie blue R-250 in acetic acid-isopropyl alcohol-water (1:3:6) for 15-20 min, and destained with acetic acid:methanol:water (50:165:785). Bands were cut out with a scalpel and soaked in defixing buffer (25 mM Tris, 50 mM Tricine, 0.2% (w/v) SDS, 6 mM urea, 5 mM DTT, pH 8.5) for 30 min at RT. After incubation, the gel slices were equilibrated in electrophoresis buffer (25 mM Tris, 50 mM Tricine, 1.73 mM SDS, pH 8.5) for 5 min, cut into small pieces and placed into an electroelution glass tube. Elution was performed at 80 V for 24 h at RT on the Electro-Eluter Concentrator (CBS Scientific, California, USA). The electroeluted samples were separated on 12% SDS-PAGE gel, followed by transfer to a PVDF membrane, pre-treated in methanol (15 s), deionised dH_2O (5 min) and stored in N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) blotting buffer (10 mM CAPS buffer pH 11, 10% (v/v) methanol, 0.1 mM thioglycollic acid). After protein transfer overnight at 17 V on a semi-dry blotter, the PVDF membrane was stained with 0.1 % (w/v) amido black in methanol: acetic acid: water (30:10:60). The stained bands were subjected to N-terminal sequencing on the Applied Biosystems Procise sequencer at the Centre of Protein Engineering (CIP) (University of Liege, Belgium).

3.4 Results

3.4.1 Peptide selection

The VP4 sequence (Arg^{453}-Ala^{755}) was analysed with the Predict7 program for epitope prediction based on hydrophilicity, surface probability, side chain flexibility and antigenicity (Cármenes *et al.*, 1989) (Appendix 3). Three peptides were selected, one each specific to either VP4-AA or VP4-RK and one common to both VP4-AA and VP4-RK (Table 3.2). For VP4-RK, a 12 amino acid long sequence corresponding to amino acid residues 494-505 was chosen. The sequence seemed to be equally hydrophilic, flexible and surface located at its N- and C-terminus. The antigenicity was relatively lower at the C-terminus and a Cys-residue was added at the C-terminus for carrier conjugation (Fig. 3.3A).

Table 3.2 Amino acid sequences of the synthetic peptides and their residue positions within the VP4 protease

Peptide code	Sequence[a]	Specificity	Residue no.
VP4-RK	RAASGKARAASGR**C**	VP4-RK	494-505
VP4	EGVREDLQPPSQR**C**	VP4-RK and VP4-AA	584-596
VP4-AA	**C**FIKRFPHNPRDWD	VP4-AA	720-733

[a] Cys in bold was added for carrier conjugation

The VP4 common peptide EGVREDLQPPSQR is more hydrophilic at the N-terminus than at the C-terminus, but with a higher flexibility on the C-terminus than the N-terminus. A Cys-residue was added at the C-terminus for carrier conjugation. The sequence has a high surface probability but low antigenicity (Fig. 3.2B). The antigenicity plot provides an indication of the likelihood of a sequence to trigger an immune response. However, the algorithm for antigenicity prediction described by Welling *et al*. (1985) was based on data collected from 69 continuous epitopes of 20 different proteins. Therefore, the low antigenicity observed may not be a reliable prediction because of low number of proteins for which the epitopes have been identified (Van Regenmortel, 1999). The VP4-AA specific peptide corresponding to residues 720-733 has high surface probability but largely hydrophilic towards its C-terminus. The C-terminus also shows higher flexibility and antigenicity than the N-terminus which contained the added Cys- residue for carrier conjugation (Fig. 3.3C).

The IBDV VP4 common peptide GVREDLQPPSQR was matched with BSNV peptide GDHRDLMSPDVR which forms a loop between β strands 4 and 5 in the BSNV VP4 3-D structure (Fig. 3.1B). The VP4-AA specific peptide KRFPHNPRDWD was matched with BSNV VP4 peptide DAADHAFDSME which formed α-helix 3. Both these peptides were shown to be in an accessible position when the structure of BSNV VP4 protease was viewed using the cn3D protein viewer (Fig. 3.3B and C, right hand panels). The location of the VP4-RK specific peptide could not be confirmed since it is outside the VP4-AA region for which the structural information is available.

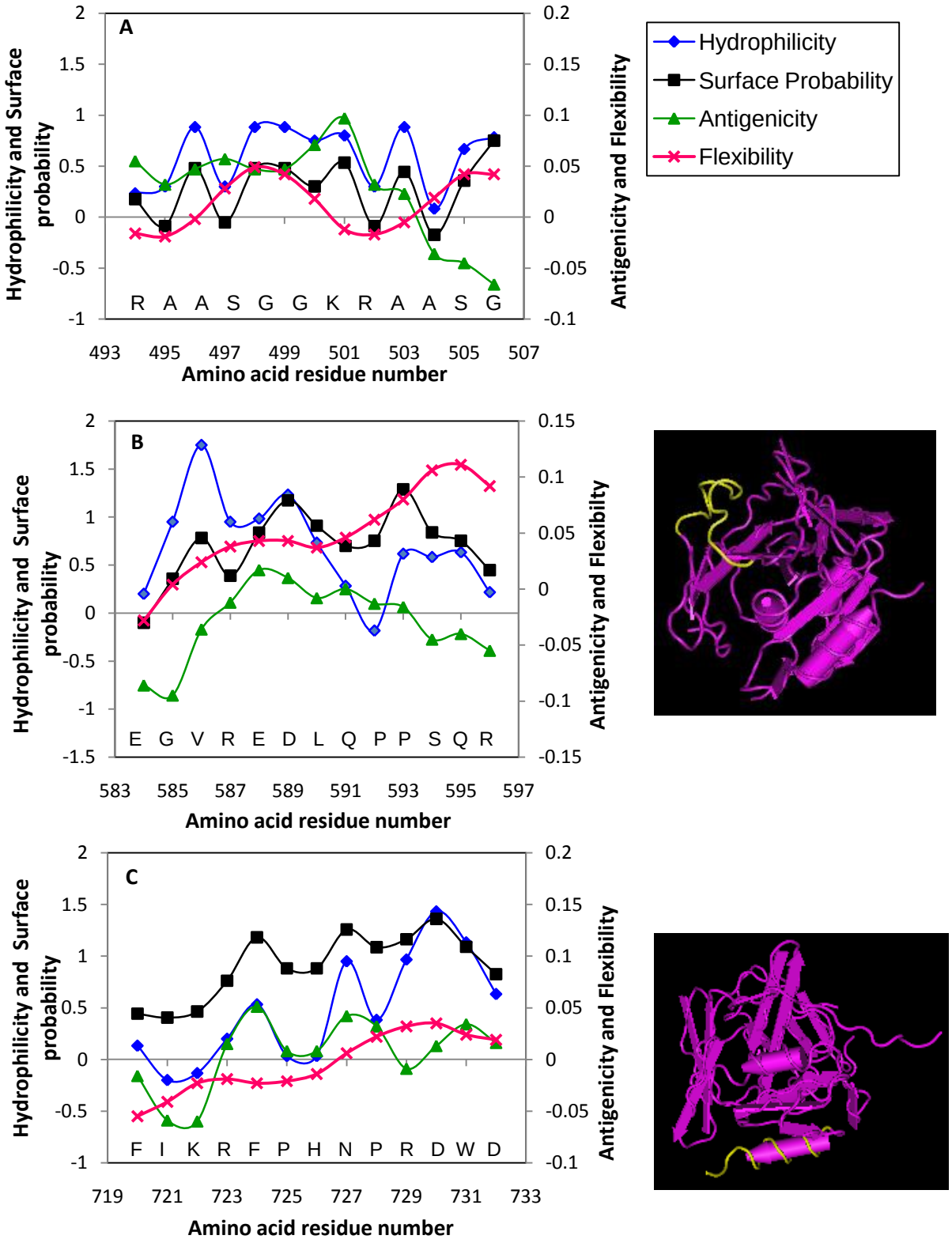


Figure 3.3 Epitope prediction plots for the selected VP4 peptides. A) VP4-RK specific sequence RAASGKARAASGR, **B)** VP4 common peptide EGVREDLQPPSQR and **C)** VP4-AA specific sequence FIKRFPHNPRDWD. Peptides were selected on the basis of four variables: hydrophilicity, surface probability, flexibility, and antigenicity (Cármenes *et al.*, 1989). The right hand panels represent the 3-D structure of BSNV mature VP4 viewed using the Cn3D program showing the positions of the corresponding peptides on the BSNV structure (in yellow). The VP4-RK peptide consists of residues 493 to 507 which are outside the mature VP4 sequence.

3.4.2 Anti-IBDV VP4 peptide antibody production

Two different chickens immunised with the IBDV VP4 peptide conjugates produced peptide specific antibodies as measured by ELISA (Fig. 3.4). In each case, the antibody titre peaked at week 4 following the first immunisation except for chicken 2 immunised with the VP4 peptide which peaked at week 5. There were no eggs collected during week 4 for chicken 2 immunised with VP4 common peptide, hence the absence of a week 4 data point. Antibody production showed a decline after week 12. The highest response was observed with the VP4-RK and VP4-AA specific peptides.

Antibodies isolated from egg yolks collected from weeks 4 to 8 and 9 to 13 after the first immunisation were pooled separately for each antigen and affinity purified on corresponding peptide affinity columns. The affinity purified antibodies showed increased binding to the corresponding peptides in an ELISA compared to non-affinity purified antibodies and the unbound fraction (Fig. 3.5).

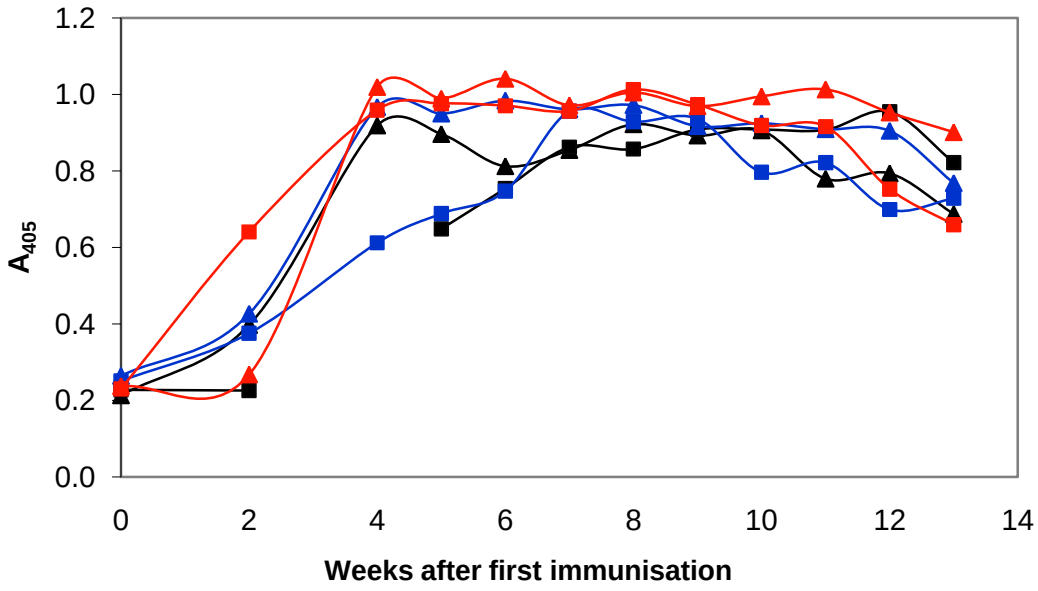


Figure 3.4 ELISA testing anti-VP4 peptide antibody production by immunised chickens. ELISA plates were coated with common VP4, VP4-RK specific and VP4-AA specific peptides and incubated with specific antibodies collected before and after the first immunisation at 50 µg/ml (in BSA-PBS, 100 µl per well, 2h, 37°C). Absorbance readings at 405 nm of IgY from two chickens immunised with VP4 (▲; ■), VP4-RK (▲; ■) and VP4-AA (▲; ■) peptides represent the average of duplicate experiments. IgY from chicken 1 is represented by a triangle, while chicken 2 is represented as a square.

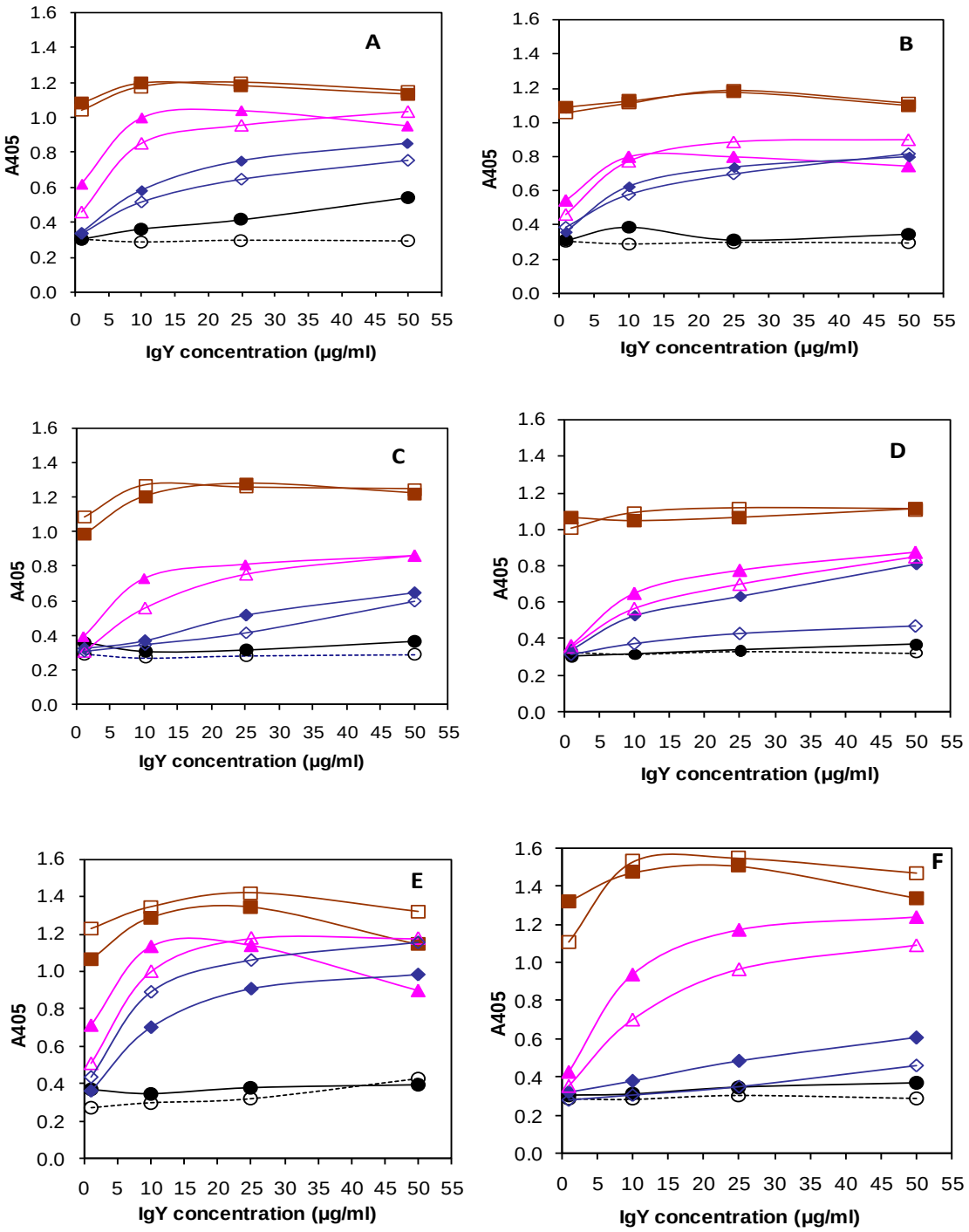


Figure 3.5 Recognition of corresponding peptides by affinity purified anti-peptide antibodies in an ELISA. **A and B)** Anti-VP4 common peptide antibodies from chickens 1 and 2; **C and D)** Anti-VP4-RK peptide antibodies from chickens 1 and 2; **E and F)** Anti-VP4-AA peptide antibodies from chickens 1 and 2. The ELISA plates were coated with the corresponding peptide and incubated with PBS (○), non-immune IgY control (●), non-affinity purified IgY weeks 4-8 (▲), non-affinity purified IgY weeks 9-13 (Δ), unbound IgY weeks 4-8 (♦), unbound IgY weeks 9-13 (◊), affinity-purified IgY weeks 4-8 (■); affinity-purified IgY week 9-13 (□).

3.4.3 Expression profile of the embedded and mature VP4

3.4.3.1 Expression in P. pastoris

Recombinant pPIC9 plasmids containing full-length ORFA2, truncated ORFA2, VP4-RK and VP4-AA were isolated from *E. coli* JM 109 cells by miniprep DNA isolation and were linearised by SacI digestion as shown by the presence of a single DNA band (Fig. 3.6A). The pPIC9- full-length polyprotein was not completely digested as shown by the presence of other plasmid conformations (Fig. 3.6A, lane 1).

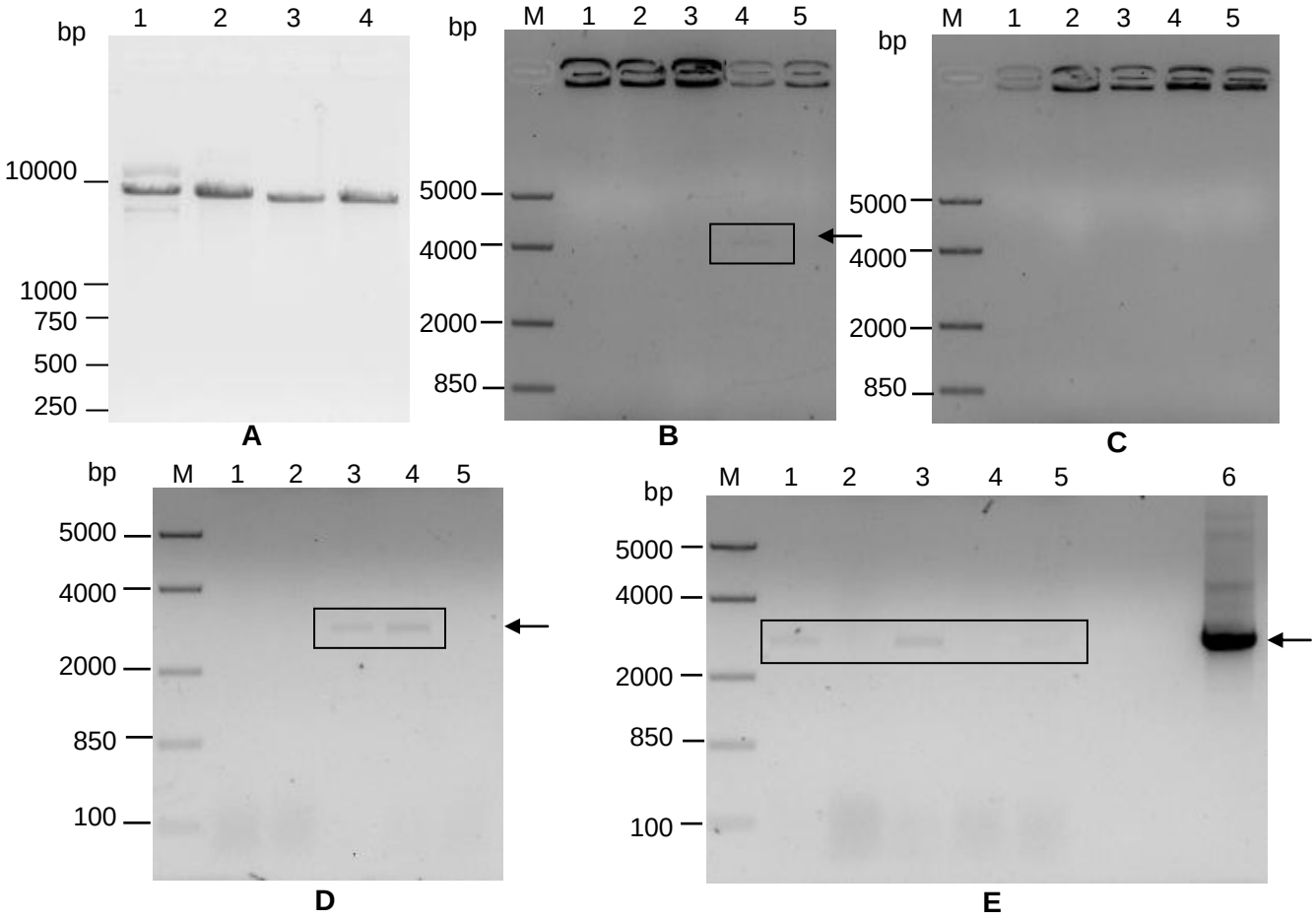


Figure 3.6 Integration of pPIC9 recombinant plasmids into the *P. pastoris* genome. A) Linearisation of recombinant pPIC9 plasmid DNA with SacI**.** Lane 1, pPIC9-full-length ORFA2; lane 2, pPIC9-truncated ORFA2; lane 3, pPIC9-VP4-RK and lane 4, pPIC9-VP4-AA. **B-D)** Colony PCR of P. *pastoris* GS115 cells transformed with SacI-linearised recombinant pPIC9 plasmids containing full-length ORFA2 (B), truncated ORFA2 (C), VP4-RK (D) and VP4-AA (E) was used to screen for successful integration into the genome. Lane M, Express middle range DNA ladder; lanes 1-5, screened colonies 1 to 5; lane 6, positive control using pPIC9-VP4-RK as template. The arrows indicate the colony PCR products confirming successful integration.

The SacI digested recombinant pPIC9 plasmids were transformed into *P. pastoris* GS115 cells and the transformants were screened by colony PCR for successful integration into

the *P. pastoris* genome (Fig. 3.6B-E). The results showed that the integration into the *Pichia* genome was only successful for the pPIC9 recombinant plasmids containing full-length polyprotein (Fig. 3.6B, lane 4), VP4-RK (Fig. 3.6D, lanes 3-4) and VP4-AA (Fig. 3.6D, lanes 1, 3, 5) coding regions. Recombinant *P. pastoris* transformants were used to inoculate minimal media and protein expression was induced by daily supplementation of the media with methanol. Analysis of protein expression by SDS-PAGE viewed by silver staining showed that there was no expression (results not shown).

3.4.3.2 Bacterial expression

All six fragments (full-length ORFA2, truncated ORFA2, VP4-RA, VP4-RK, VP4-AW and VP4-AA) were successfully cloned into pET-32b in both *E. coli* BL21 (DE3) and BL21 (DE3) pLysS as well as pGEX-4T-1 in *E. coli* BL21 cells. Only the full-length polyprotein coding region and VP4-RK were successfully cloned into pET-28b and the expression of both constructs was unsuccessful in this system. Expression in pGEX-4T-1 was only successful for the different VP4 constructs (results not shown). Only the pET-32b system was able to express all the constructs but only in the *E. coli* BL21 (DE3) cells, consequently this system was used for expression. The cell lysates were analysed by reducing SDS-PAGE followed by a western blot using anti-VP4 peptide antibodies and an anti-His tag monoclonal antibody (Fig. 3.7).

The full-length polyprotein construct was expressed as cleavage products which could not be easily distinguished from bacterial proteins by SDS-PAGE analysis (Fig. 3.7A). The full-length polyprotein construct encoded an N-terminal Trx-His tag, the pVP2-VP4-VP3 polyprotein and a C-terminal 6xHis tag (Fig. 3.7B), meaning that the resulting products would be recognised by all three anti-VP4 peptide antibodies and the anti-His tag antibodies. The expected size of the unprocessed polyprotein fused to the tag was 130 kDa, however, the presence of higher molecular weight proteins was observed on the SDS-PAGE gel (Fig. 3.7A, indicated by an arrow). These high molecular weight proteins were initially thought to be due to the relatively high alanine content of the polyprotein sequence as it has been shown that proteins with high alanine content migrate at a higher molecular weight than expected on SDS-PAGE (Nolan *et al.*, 2000; Thomson *et al.*, 2002; Baida *et al.*, 2006).

These high molecular weight proteins were also observed when pET-32b-VP2 was expressed in *E. coli* in another study in our laboratory (unpublished results). It was then thought that the high molecular weight proteins could be multimers of VP2 released after processing. However, the proteins were recognised by all the anti-VP4 peptide antibodies

as well as by the anti-His tag antibody suggesting that the antibodies could be recognising the assembled polyprotein.

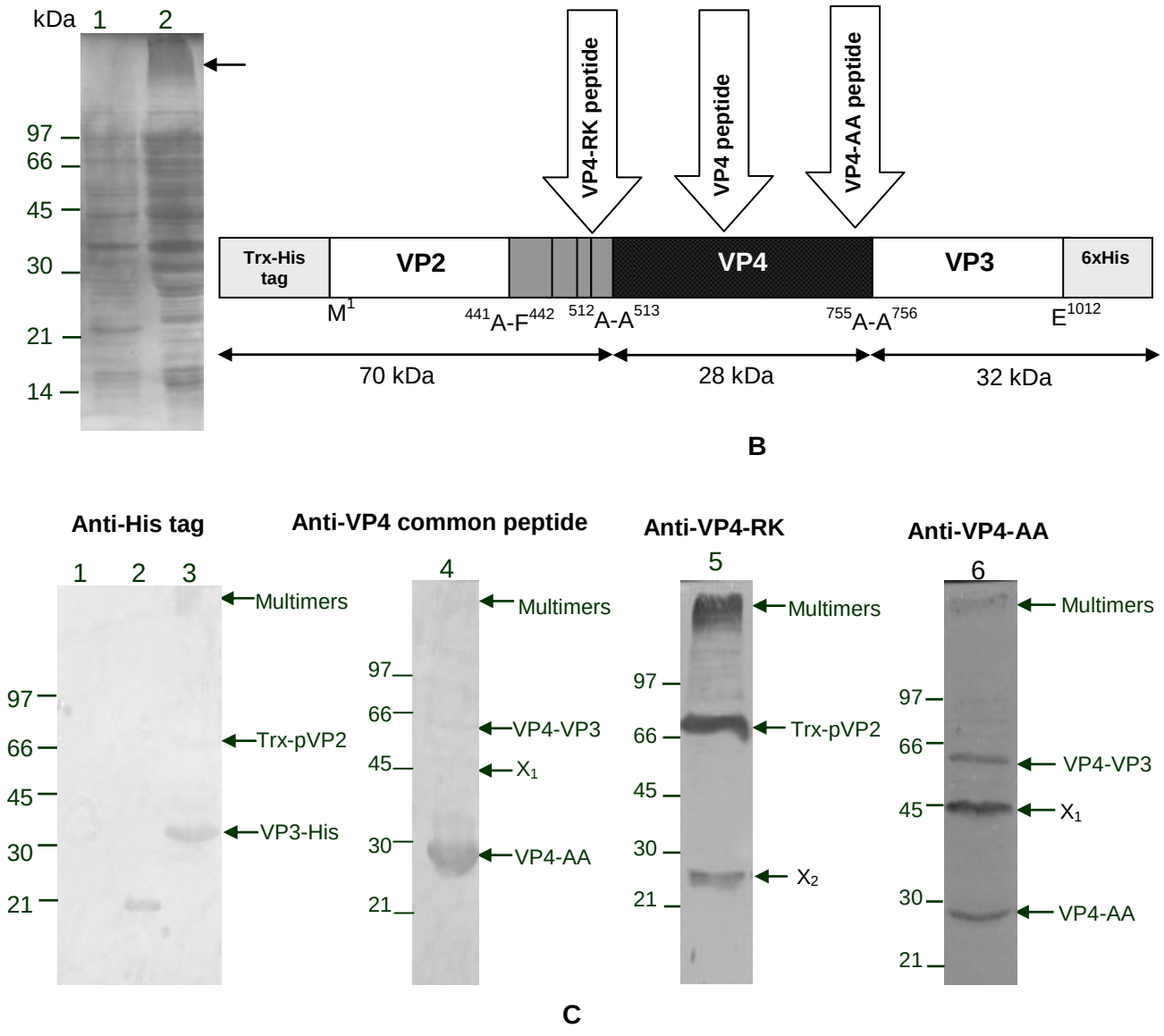


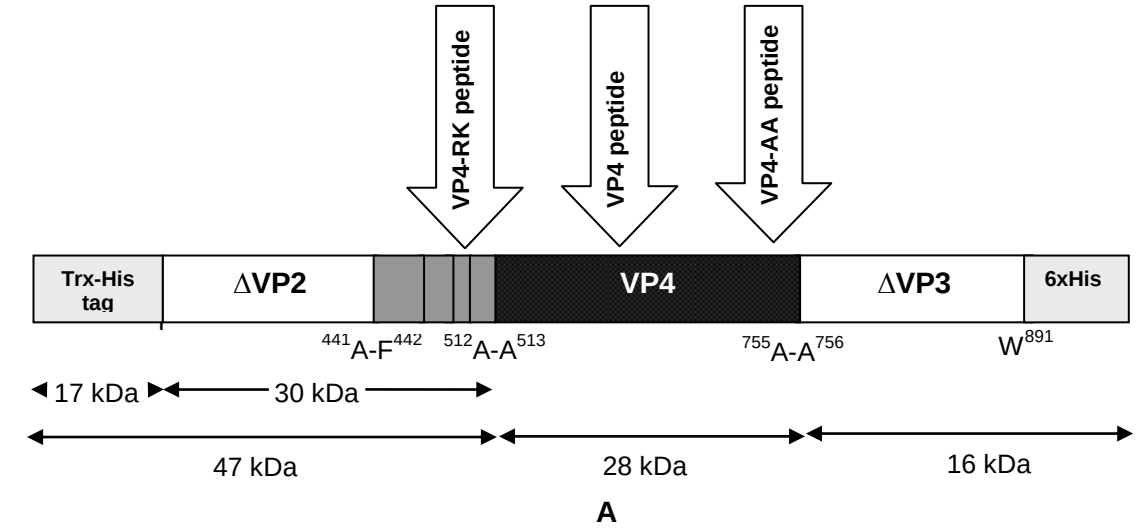
Figure 3.7 Expression profile of pET-32b-full-length polyprotein construct. A) Reducing SDS-PAGE analysis. Lane 1, *E. coli* BL21 (DE3*)* control cell lysate; lane 2, full-length polyprotein. **B)** Schematic representation of the construct showing anti-peptide antibody binding location. The expressed constructs would be expressed with a thioredoxin-His (Trx-His) tag and a C-terminal 6x His tag. **C)** Western blot analysis. Lane 1, *E. coli* BL21 (DE3); lane 2, non-recombinant pET-32b plasmid expression detected by anti-His tag monoclonal antibody; lane 3, full-length polyprotein detected by anti-His tag monoclonal antibody; lane 4, full-length polyprotein detected by anti-VP4 common peptide antibodies; lane 5, full-length polyprotein detected by anti-VP4-RK peptide antibodies; lane 6, full-length polyprotein detected by anti-VP4-AA peptide antibodies. All sizes are shown in kDa. The arrows indicate the expected products, the unexpected or unknown products are indicated with X_1 and X_2.

Western blot analysis using anti-His tag and anti-VP4 peptide antibodies showed different processing products. The anti-His tag monoclonal antibody recognised two products at 70 and 32 kDa representing the Trx-His-pVP2 fusion protein and the C-terminally 6xHis tagged VP3 (VP3-His), respectively (Fig. 3.7C, lane 3). This suggests that the Trx-pVP2 (approximately 70 kDa) released by the cleavage at ^{512}Ala-Ala^{513}, was not autocatalytically cleaved to yield Trx-His-VP2 fusion protein (approximately 64 kDa). Precursor VP2 (pVP2) was shown to be processed into VP2 in Hi-5 cells when the polyprotein was expressed using the Baculovirus system, but not in sf9 insects cells (Lee *et al.,* 2006b). Due to these authors' observation, there were some doubts regarding the processing of pVP2 in *E. coli,* especially since it has not yet been reported. The data presented here suggests that it was not processed.

The anti-VP4 common peptide antibodies recognised one prominent protein band at 28 kDa corresponding to VP4-AA and two less prominent bands at 60 and 45 kDa (Fig. 3.7C, lane 4). The 60 kDa may represent VP4-VP3 precursor polypeptide while the release of the 45 kDa product is unexpected with the currently understood processing strategy. Anti-VP4-RK peptide antibodies recognised two major cleavage products at 70 and 24 kDa. The 70 kDa product could represent the Trx-His-pVP2 fusion protein, while the product at 24 kDa did not match any predicted cleavage products (Fig. 3.7C, lane 5). The 70 kDa consisting of the N-terminal Trx-His tag (17 kDa) and pVP2 (54 kDa) would be the same product that was recognised by the anti-His tag monoclonal antibody. Anti-VP4-AA peptide antibodies, on the other hand, recognised three major cleavage products at 60, 45 and 28 kDa. The 60 kDa product matched the predicted VP4-VP3 precursor polypeptide (Fig. 3.7C, lane 6). The product at 28 kDa represents VP4-AA, while the product observed at 45 kDa did not match any of the predicted cleavage products. As expected, anti-VP4 common and anti-VP4-AA peptide antibodies recognised the same products. All the unknown or unpredicted products (X_1 and X_2) recognised by the anti-VP4 peptide antibodies suggest cleavages at different or additional sites.

Expression of the truncated polyprotein (ΔPP) produced an unexpected processing pattern in comparison to that of the full-length polyprotein. The truncated polyprotein construct encoded an N-terminal Trx-His tag (17 kDa), N-terminally truncated pVP2 (ΔpVP2, 30 kDa), VP4 (28 kDa) and a C-terminally 6xHis tagged truncated VP3 (ΔVP3-His, 16 kDa) (Fig. 3.8A). The unprocessed protein was observed at the expected size of approximately 90 kDa (Fig. 3.8B-E). A protein band below the 90 kDa protein, identified by all the anti-peptide antibodies was also observed. The protein was thought to be Trx-His-ΔpVP2-VP4, expected at about 75 kDa. However, this was questionable considering that

a corresponding product in the processing of the full-length polyprotein shown in Fig. 3.7 was not observed. High molecular weight products representing multimers and several distinct bands between 45 and 66 kDa (represented by X_3) were also observed. Anti-VP4-RK peptide antibodies recognised two other cleavage products at 45 and 24 kDa. The 45 kDa could represent the Trx-ΔpVP2 fusion protein, similar to the Trx-pVP2 fusion protein observed at 70 kDa for the full-length polyprotein. The product (X_2) at 24 kDa did not match any predicted cleavage products (Fig. 3.8C, lane 5).



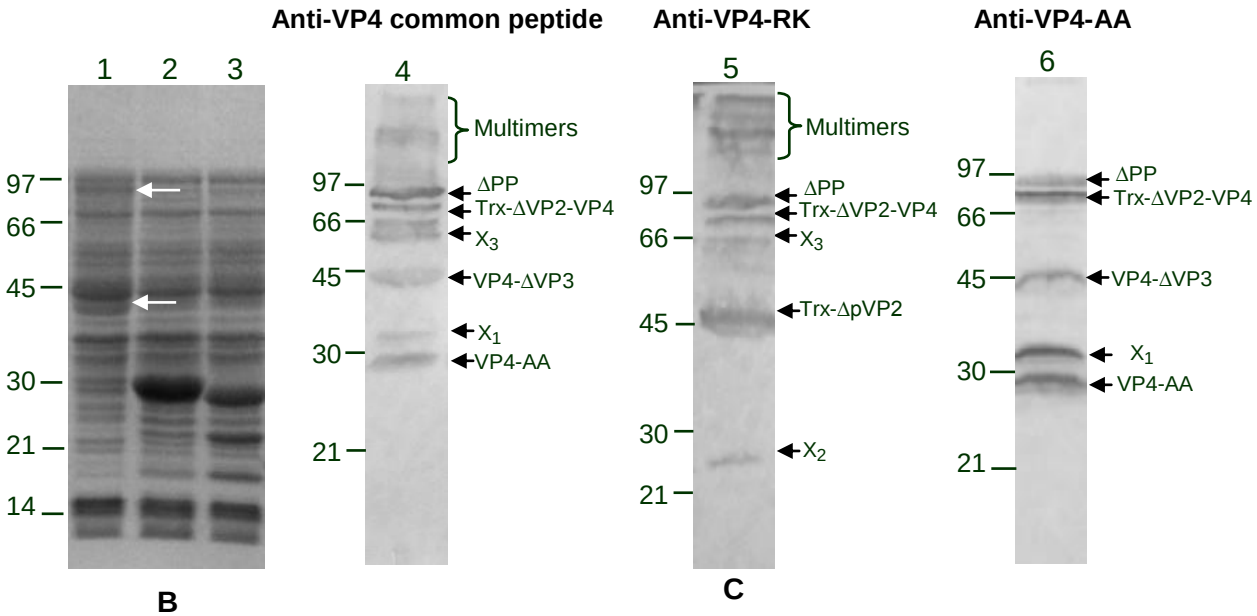


Figure 3.8 Expression profile of the wild-type pET-32b-truncated polyprotein construct. A) Schematic representation of the construct showing anti-peptide antibody binding location. **B)** Reducing SDS-PAGE analysis. Lane 1, truncated polyprotein; lane 2, VP4-RA and lane 3, VP4-RK. The white arrows show the expressed proteins observed in the SDS-PAGE gel. **C)** Western blot analysis of truncated polyprotein expression using different anti-VP4 peptide antibodies. Lane 4, anti-VP4 common peptide antibodies; lane 5, anti-VP4-RK peptide antibodies; lane 6, anti-VP4-AA peptide antibodies. All sizes are shown in kDa. The arrows indicate the expected products, the unexpected or unknown products are shown with $X_{1\text{-}3}$.

Anti-VP4 common peptide and anti-VP4-AA peptide antibodies recognised three other cleavage products at 45, 34, and 28 kDa. The 45 kDa product matched the predicted VP4-ΔVP3 precursor polypeptide, the product at 28 kDa represents VP4-AA, and the product observed at 34 kDa (X_1) did not match any of the predicted cleavage products (Fig. 3.8C, lanes 5 and 6). As was observed for the full-length polyprotein expression (Fig. 3.7), the observed products did not show the expected sizes. Even though the processing pattern was similar, there were more unknown products (X_1-X_3) observed for the truncated polyprotein than for the full-length polyprotein processing (X_1-X_2).

The expression of VP4-RA and VP4-RK was consistent with the use of a different cleavage site for the pVP2-VP4 junction. For VP4-RA, the cleavage at ^{512}Ala-Ala^{513} would be expected to yield an N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide of 24 kDa, and a C-terminally His-tagged VP4 (Ala^{513}-Ala^{755}) of 28 kDa (Fig. 3.9A). For VP4-RK, the cleavage at ^{512}Ala-Ala^{513} was expected to yield an N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide of 24 kDa, and a C-terminally His-tagged ΔVP4 (Ala^{513}-Lys^{722}) of 24 kDa (Fig. 3.9B). However, the products corresponding to the C-terminally His-tagged VP4 and ΔVP4 were observed at a higher size than expected. Two products of about 22 and 30 kDa were observed for VP4-RA. For VP4-RK, two products of molecular sizes 22 and 27 kDa were observed (Fig. 3.9C).

The anti-VP4 common peptide recognised the C-terminally His-tagged VP4 (30 kDa) and ΔVP4 (27 kDa) products as expected for VP4-RA and VP4-RK, respectively (Fig. 3.9D). The anti-VP4-AA peptide antibodies also recognised the C-terminally His-tagged VP4 (30 kDa) but not the ΔVP4 (27 kDa), as expected (Fig. 3.9F). The VP4-AA specific peptide is located within the Lys^{722}-Ala^{755} polypeptide which is not found in VP4-RK, explaining why the antibodies could not recognise the ΔVP4 product. The peptide for the anti-VP4-RK peptide antibodies is located within the Arg^{453}-Ala^{512} sequence of the polypeptide, therefore the antibodies should only recognise the N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide at 24 kDa, not the C-terminally 6xHis-tagged VP4 (30 kDa) or ΔVP4 (27 kDa). Instead, these antibodies seemed to recognise these 30 and 27 kDa VP4 products (Fig. 3.9E), which was very puzzling. The Arg^{453}-Ala^{512} polypeptide contains three cleavage sites: ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502}, and processing at these sites is not well understood. The VP4-RK peptide used for antibody production consists of residues 493-506, therefore the product containing this peptide should result from cleavage at ^{487}Ala-Ala^{488}. These results suggested that the site used for the processing of the pVP2-VP4 junction could be ^{487}Ala-Ala^{488} instead of ^{512}Ala-Ala^{513}.

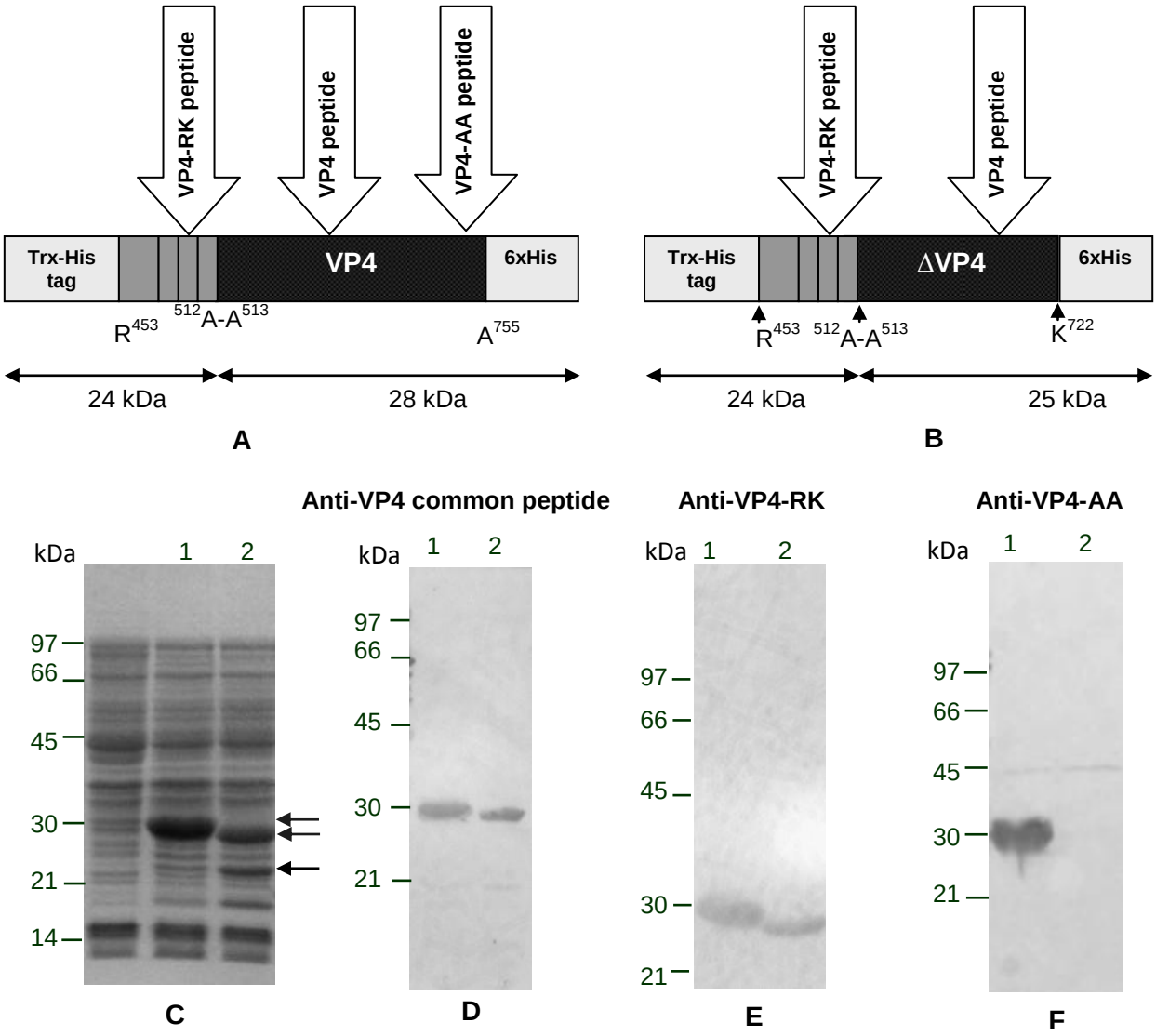


Figure 3.9 Expression profile of the wild-type VP4-RA and VP4-RK constructs. A) Schematic representation of the VP4-RA construct showing anti-peptide antibody binding locations. **B)** Schematic representation of the VP4-RK construct showing anti-peptide antibody binding locations. **C)** Reducing SDS-PAGE analysis. The arrows show the expressed proteins observed in the SDS-PAGE gel. **D-F)** Western blot analysis of the expression using anti-VP4 common peptide antibodies (D); anti-VP4-RK peptide antibodies (E) and anti-VP4-AA peptide antibodies (F). Lane1, VP4-RA and lane 2, VP4-RK.

The 30 and 27 kDa VP4 products recognised by the anti-peptide antibodies were purified by electroelution for N-terminal sequencing, to confirm the cleavage site used for the release of the products (Fig. 3.10). The N-terminal sequence of the two fragments was unambiguously identified as ADKGYE, meaning that cleavage occurs at the ^{512}Ala-Ala^{513} even in the presence of the other cleavage sites. It could not be explained how the anti-VP4-RK peptide antibodies were able to recognise these products when cleavage occurs at the ^{512}Ala-Ala^{513} site.

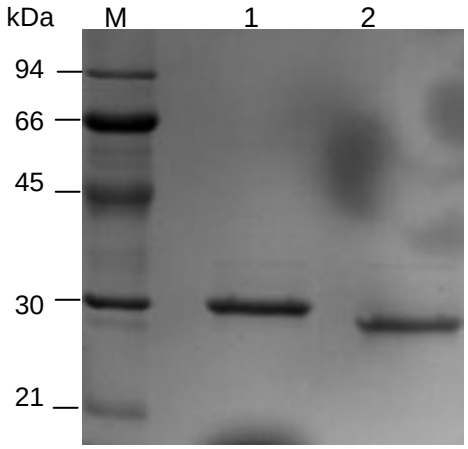


Figure 3.10 Purification of VP4-RA and VP4-RK cleavage products by electroelution. Lane M, Low MWM; lane 2, VP4-RA product; lane 3, VP4-RK product.

VP4-AW expression produced two bands corresponding to the unprocessed fusion protein at 60 kDa and a Trx-His tagged VP4-AA product at 46 kDa resulting from the processing at the VP4-VP3 junction (Fig. 3.11A and B, lane 2). Both the unprocessed fusion protein and the Trx-His tagged VP4-AA product were recognised by the anti-VP4 common peptide and anti-VP4-AA peptide antibodies as expected (Fig. 3.11C and E, lanes 2). The presence of the unprocessed protein may indicate that the processing reaction at this junction occurs in *trans*. VP4-AA was expressed as an unprocessed fusion protein at 46 kDa indicating that the mature form does not have the autocatalytic activity (Fig. 3.11A, lane 3). The Trx-His-VP4-AA fusion protein was recognised by the anti-VP4 common peptide and anti-VP4-AA peptide antibodies as expected (Fig. 3.11C and E, lanes 3). The VP4-RK peptide used for antibody production consists of residues 493-506 hence the antibodies did not recognise both VP4-AW and VP4-AA products (Fig. 3.11D, lanes 2 and 3).

The polyprotein and VP4 constructs containing the Arg^{453}-Ala^{512} polypeptide (VP4-RA and VP4-RK) were autocatalytic while the constructs without the polypeptide (VP4-AW and VP4-AA) did not display autocatalytic activity.

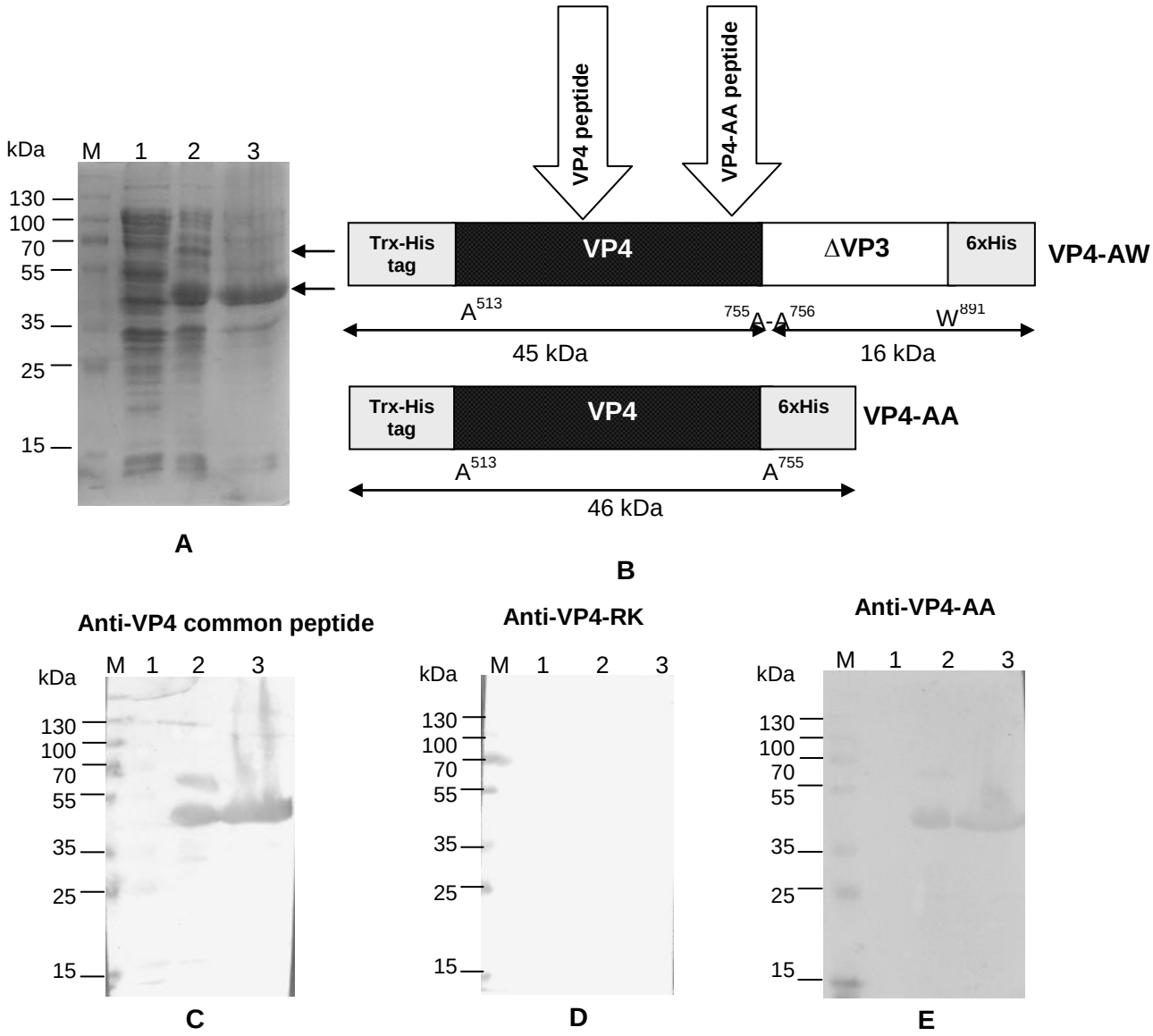


Figure 3.11 Expression profile of VP4-AW and VP4-AA constructs. A) Reducing SDS-PAGE analysis. Lane M, PageRuler™ plus prestained protein ladder; lane 1, *E. coli* BL21 (DE3*)* control cell lysate; lane 2, VP4-AW and lane 3, VP4-AA. The arrows show the 60 and 45 kDa expressed proteins observed in the SDS-PAGE gel. **B)** Schematic representation of the constructs showing anti-peptide antibody binding locations. **C-E)** Western blot analysis of the expression using anti-VP4 common peptide antibodies (C), anti-VP4-RK peptide antibodies (D) and anti-VP4-AA peptide antibodies (E). Lane M, PageRuler™ plus prestained protein ladder; lane 1, *E. coli* BL21 (DE3*)* control cell lysate; lane 2, VP4-AW and lane 3, VP4-AA.

3.4.4 Detection of VP4 in infected bursal homogenates by the anti-VP4 peptide antibodies

Since the anti-VP4 peptide antibodies successfully recognised recombinant VP4 on a western blot, the antibodies were also used to detect native VP4 in IBDV-infected bursal homogenates. Anti-VP4 common peptide antibodies detected a double band at about 28 kDa, which corresponds to the theoretical size of VP4-AA in IBDV infected bursal samples

(Fig. 3.12B). The anti-VP4-RK antibodies did not detect any protein in either infected or non-infected bursal samples (Fig. 3.12B). By contrast, the anti-VP4-AA peptide antibodies strongly detected two bands of 50 and 60 kDa in the infected sample and a 55 kDa band in the non-infected sample. The antibodies also cross-reacted to some extent with proteins of 100 and 120 kDa in both infected and non-infected samples (Fig. 3.12D).

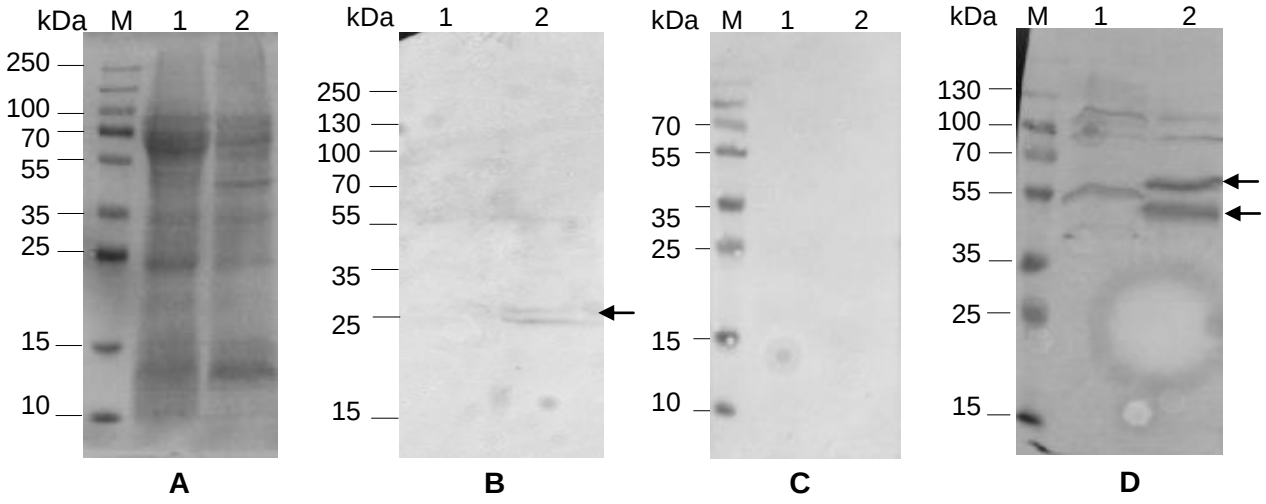


Figure 3.12 Detection of native VP4 in infected bursal samples by anti-peptide antibodies in a western blot. A) SDS-PAGE analysis of bursal samples. **B-D)** Western blot analysis using anti-VP4 common peptide antibodies (B), anti-VP4-RK peptide antibodies (C) and anti-VP4-AA peptide antibodies (D). Lane 1, non-infected bursal supernatant; lane 2, infected bursal supernatant. The arrows show the proteins detected in the infected bursal samples.

3.4.5 Purification of the full-length polyprotein

Full-length IBDV polyprotein was expressed as cleavage products and high molecular weight proteins in the form of insoluble inclusion bodies. The inclusion bodies were solubilised with 8 M urea and subjected to purification under denaturing conditions with the aim of purifying the high molecular weight proteins. As shown in Fig. 3.13, proteins bound poorly to the affinity matrix and some proteins were also washed off the column prior to elution, resulting in a very low protein yield. There was also co-purification of bacterial proteins. When the washing stringency was increased by including imidazole in the washing buffer, there was no yield of the proteins of interest which rendered this method of purification unsuccessful (results not shown).

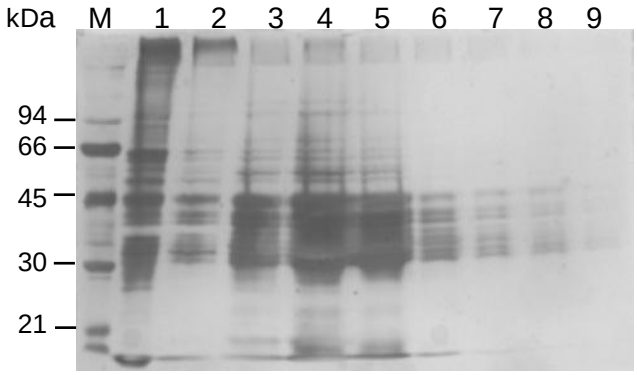


Figure 3.13 Analysis by reducing SDS-PAGE and silver staining of the purification of the polyprotein under denaturing conditions on a nickel affinity column. Lane M, low MWM; lane 2, solubilised sample; lane 3; unbound fraction; lane 4, wash fraction; lanes 5-9, imidazole elution fractions.

Since purification under denaturing conditions proved unsuccessful, the next method attempted was to refold the proteins before purification. Refolding by dilution resulted in a major loss of protein and did not seem to be successfully applied to this protein (results not shown). Refolding was also carried out by dialysis whereby the solubilised proteins were refolded by a series of dialysis steps against decreasing concentrations of urea in 20 mM Tris-HCl buffer pH 8. The refolding was analysed by western blotting using anti-VP4-RK antibodies (Fig. 3.14). Anti-VP4-RK peptide antibodies recognised the high molecular weight proteins and two cleavage products at 70 and 24 kDa. Refolded protein was expected to be in the final supernatant without urea. The results showed that some of the refolded protein formed precipitates; however, a substantial amount of the high molecular weight proteins and the two cleavage products were refolded (lane 14).

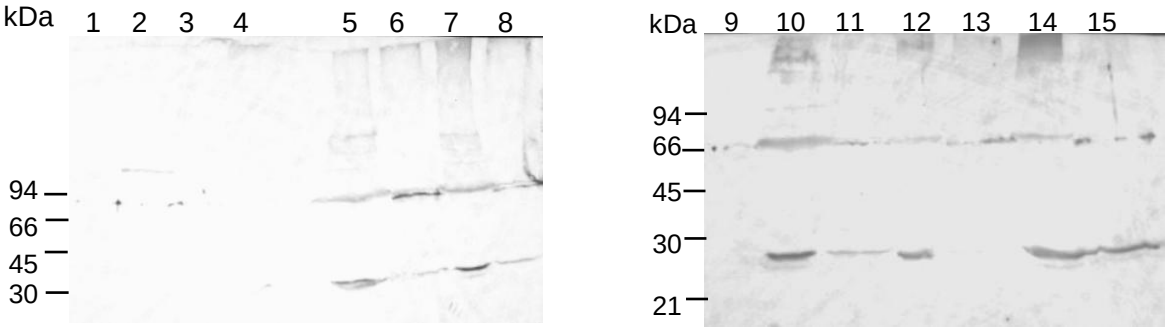


Figure 3.14 Western blotting analysis with anti-VP4-RK peptide antibodies of the refolding profile of recombinantly expressed full-length polyprotein. Lane 1, *E. coli* BL21 (DE3*)* control cell lysate; lane 2, soluble fraction; lane 3, inclusion body wash supernatant, lane 4, Inclusion bodies solubilised in 8 M urea; lane 5, supernatant at 4 M urea; lane 6, pellet at 4 M urea; lane 7, supernatant at 2 M urea; lane 8, pellet at 2 M urea; lane 9, *E. coli* BL21 (DE3*)* control cell lysate; lane 10, supernatant at 1 M urea; lane 11, pellet at 1 M urea; lane 12, supernatant at 0.5 M urea; lane 13, pellet at 0.5 M urea; lane 14, supernatant at 0 M urea; lane 15, pellet at 0 M urea.

Purification of the refolded full-length polyprotein sample was carried out using immobilised chelate affinity chromatography (nickel or cobalt). The low yield of proteins in the eluted fractions indicated that the polyprotein was not binding well to the nickel resin (Fig. 3.15). Even though the polyprotein bound to the cobalt resin, other bacterial proteins were co-purified. Further attempts to purify the full-length polyprotein using other methods of chromatography resulted in loss of sample (results not shown).

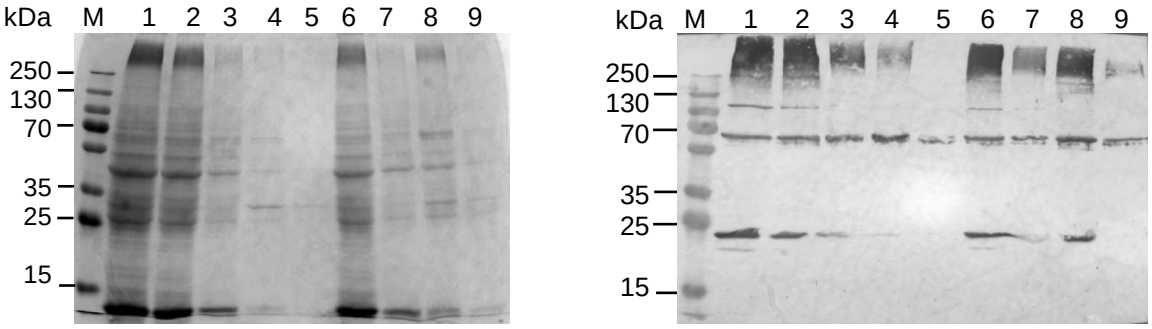


Figure 3.15 Analysis of the purification of recombinantly expressed full-length polyprotein by cobalt and nickel affinity chromatography by reducing SDS-PAGE (A) and western blotting using anti-VP4-RK peptide antibodies (B). Lane 1, refolded proteins; lane 2, unbound fraction from nickel column; lane 3, wash fraction from nickel column, lane 4-5, elution fractions from the nickel column; lane 6, unbound fraction from the cobalt column; lane 7, wash fraction from the cobalt column; lane 8-9, elution fractions from the cobalt column.

3.4.6 Purification of VP4-AA

VP4-AA was successfully expressed in both pET-32b and pGEX-4T-1. The latter expression system was more favourable since the expressed protein was only fused with an N-terminal GST, unlike pET-32b where the protein has an N-terminal Trx-His tag and a C-terminal 6xHis tag. In addition, VP4-AA expressed as a soluble protein in the pGEX-4T-1 system. The soluble fraction (Fig. 3.16A, lane 1) was applied to a glutathione-agarose column and the VP4-AA fusion protein bound to the glutathione as shown by the absence of the 54 kDa fusion protein in the unbound fraction (lane 2).

The purification of the 54 kDa fusion protein was successful as shown by a prominent band in the eluted fraction (lane 3). The cleavage of the fusion protein in solution and subsequent purification steps to separate the recombinant protein from GST was found to be difficult especially since GST and VP4-AA have similar sizes. The ‘on-column’ method of cleavage of the fusion protein eliminated these difficulties and significantly shortened the length of the purification procedure. Following on-column thrombin cleavage of the GST fusion protein, the recombinant VP4-AA eluted fractions had an apparent molecular weight of 28 kDa as expected (Fig. 3.16B, lane 3-8).

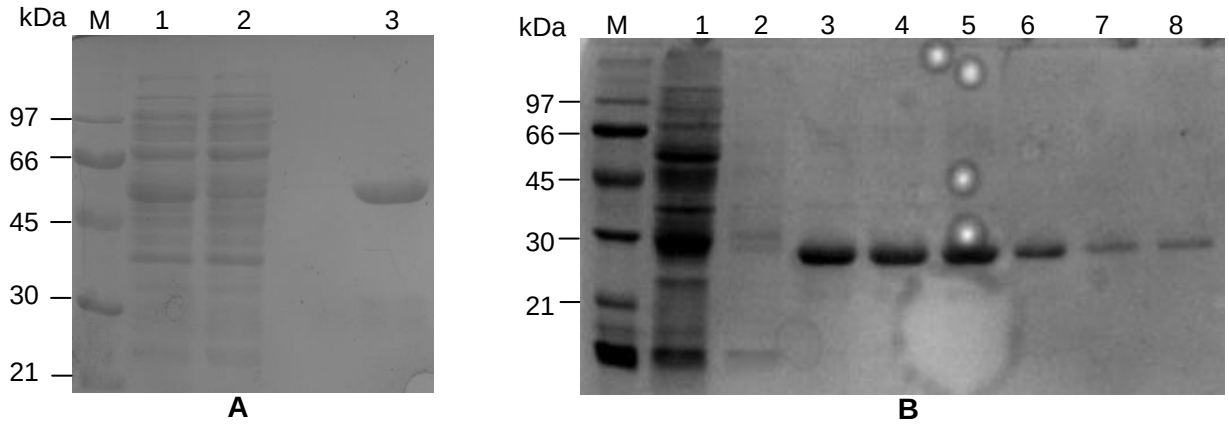


Figure 3.16 Purification of VP4-AA expressed by the pGEX system in *E. coli.* A) Affinity purification of GST-VP4-AA fusion protein on a glutathione-agarose resin. Lane M, Low range MWM; lane 1, soluble fraction; lane 2, unbound fraction; lane 3, eluted fraction. **B)** VP4-AA purification with on-column thrombin cleavage on glutathione-agarose resin. Lane M, low range MWM; lane 1, unbound fraction; lane 2, wash fraction; lanes 3-8, VP4-AA fractions eluted with reduced glutathione.

3.4.7 Reactivity of anti-peptide antibodies with recombinant VP4-AA in an ELISA

Once purified recombinantly expressed VP4-AA was available, the anti-peptide antibodies were also tested for reactivity against purified recombinant VP4-AA in an ELISA (Fig. 3.17) to determine whether the peptides are accessible to the antibodies in the IBDV recombinant VP4-AA protein. During peptide design, due to the unavailability of the IBDV VP4 structure, the BSNV VP4 3-D structure was used to predict the location of the selected peptides based on homology. The results from the ELISA would give some information regarding the location of the peptides. Recombinant VP4-AA was recognised by anti-VP4 common and anti-VP4-AA peptide antibodies as expected. This suggests that the folding of recombinant VP4-AA exposed these peptides. These results somewhat correlate with Fig. 3.3, which shows the position of the corresponding peptides on the BSNV VP4 3-D structure. As expected the anti-VP4-RK peptide antibodies did not recognise the VP4-AA since the VP4-RK peptide is not present in the VP4-AA region.

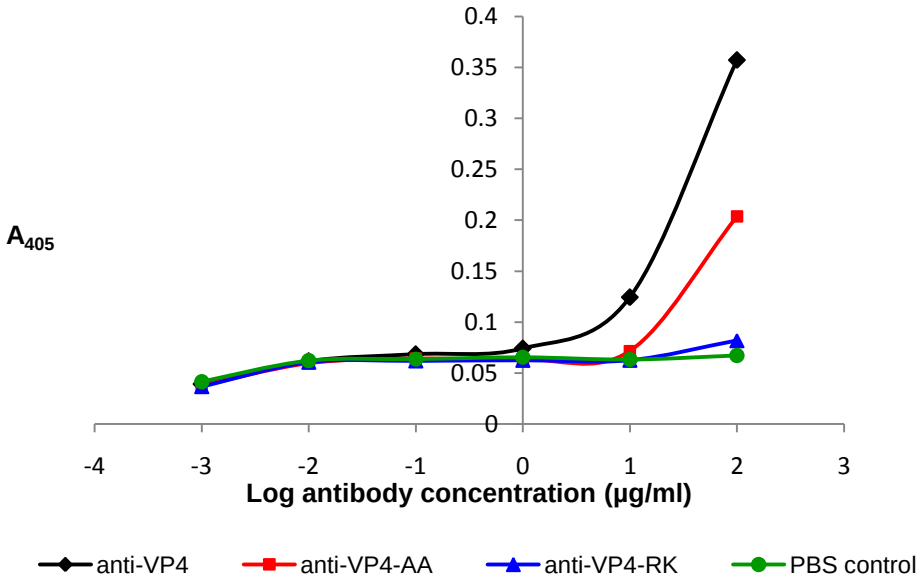


Figure 3.17 Recognition of recombinant VP4-AA by affinity purified anti-peptide antibodies in an ELISA. The ELISA plates were coated with the purified recombinant VP4-AA, and incubated with serial 10-fold dilutions of purified anti-peptide antibodies, followed by incubation with HRPO-linked rabbit anti-IgY secondary antibody and ABTS. Each point is the mean absorbance at 405 nm of triplicate samples.

3.4.8 Enzymatic characterisation of VP4-AA

3.4.8.1 Zymogram analysis of proteolytic activity

Gelatine SDS-PAGE analysis was performed to determine the activity of the purified VP4-AA against gelatine (Fig. 3.18). Activity as evidenced by a clear zone on the gelatine gel was observed at approximately 56 kDa (lane 1). Since VP4-AA was expressed as a GST fusion protein and the tag was removed through thrombin on-column cleavage, it could not be ruled out that the observed activity was due to thrombin. Preincubation of the VP4-AA sample with either anti-thrombin III, AEBSF or soybean trypsin inhibitor (serine protease inhibitors) did not inhibit the activity on gelatine. Surprisingly, samples pre-incubated with the inhibitors showed an additional prominent band and a further faint band in the case of anti-thrombin III and SBTI. Thrombin in the absence and presence of anti-thrombin III, AEBSF or SBTI were also run on the gelatine gel, and no clearing was observed (lanes 5-8). This was similar to the cysteine protease-inhibitor complexes observed with Cathepsin L and stefin B (Coetzer *et al.*, 1995). Protease inhibition by anti-thrombin III results from the formation of equimolar complexes between antithrombin and the protease which renders the active site inaccessible to the substrate (Olson and Bjork, 1994).

Repeated zymograms using VP4-AA purified by on-column thrombin cleavage also resulted in the clear zone at approximately 56 kDa, with two prominent bands and a faint band (lane 10). Therefore it was not clear whether the additional bands resulted from inhibitor-peptidase complexes or just VP4-AA. To further confirm that VP4-AA was responsible for the activity, the soluble fraction from recombinant expression used for VP4-AA purification was also analysed, and no activity was observed (lane 9). This indicated that the activity was not due to bacterial proteases. The purified GST-VP4-AA fusion protein was also analysed but showed no activity (lanes 11 and 13). However, a GST fraction eluted after on-column cleavage of the VP4-AA fusion protein showed clearing with low intensity (lane 12). This was expected since traces of VP4-AA could still be present. Interestingly, activity was only observed when VP4-AA was not fused to the fusion tag.

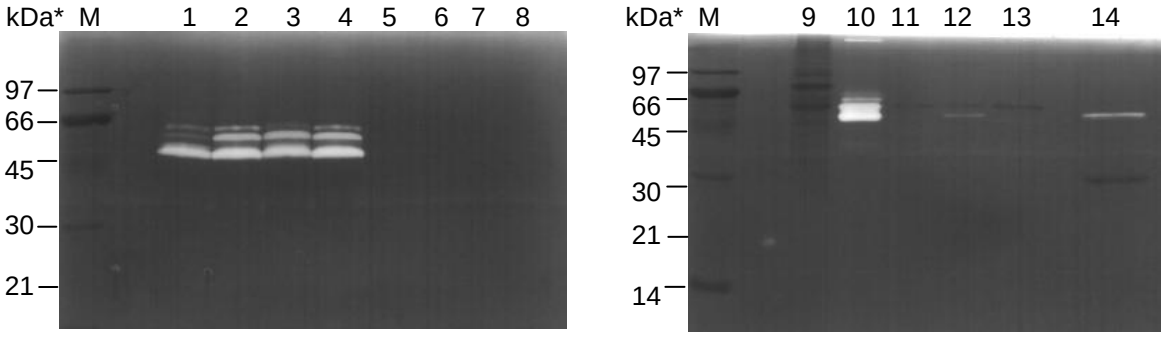


Figure 3.18 Analysis of proteolytic activity of purified VP4-AA on a gelatine-SDS-PAGE gel. Proteolysis is shown by a clear region on the gel. Lane M, Low range MWM; lane 1, VP4-AA purified sample; lane 2, VP4-AA purified sample pre-incubated with anti-thrombin III; lane 3, VP4-AA purified sample pre-incubated with AEBSF; lane 4, VP4-AA purified sample pre-incubated with trypsin inhibitor; lane 5, thrombin; lane 6, thrombin pre-incubated with anti-thrombin III; lane 7, thrombin pre-incubated with AEBSF; lane 8, thrombin pre-incubated with trypsin inhibitor; lane 9, VP4-AA soluble fraction; lane 10, thrombin cleaved purified VP4-AA sample, lanes 11 and 13, purified GST-VP4-AA fusion protein, lane 12, GST fraction; lane 14, electroeluted VP4-RA cleavage product. *Note that kDa sizes are approximate under non-reducing conditions.

The expression of the VP4-RA (Arg^{453}-Ala^{755}) construct in pET-32b resulted in a 30 kDa product corresponding to the C-terminally 6xHis tagged VP4-AA region (Fig. 3.9) and the product was electroeluted from a gel for N-terminal sequencing (Fig. 3.10). The electroeluted sample also digested the gelatine at 56 kDa, which confirmed conclusively that the VP4-AA region was active on a gelatine gel (Fig. 3.18, lane 14). However, only one band was observed as opposed to the three bands resulting from VP4-AA purified by on-column thrombin cleavage. It was interesting that some of the electroeluted protein was observed at the expected size of 30 kDa but showed no activity. However, activity was observed 56 kDa suggesting that VP4-AA was active as a dimer in a gelatine gel.

3.4.8.2 Trans-activity of the VP4-AA

The ability of VP4-AA to cleave in *trans* was also analysed using mutant VP4-RK (Fig. 3.19). The mutant construct carries the catalytic dyad (S625A and K692R) mutations, and was expressed as an N-terminally Trx-His-tagged protein at 49 kDa (Chapter 4). Four potential cleavage sites (P_1-P_1' position residues ^{488}Ala-Ala^{489}, ^{494}Ala-Ala^{495}, ^{501}Ala-Ala^{502} and ^{512}Ala-Ala^{513}) are present in the mutant VP4-RK. The mutant was expressed as inclusion bodies therefore a protein sample purified under denaturing conditions (Fig. 3.19A) and a non-purified refolded sample (Fig. 3.19B) were used in the assay. The results for both samples showed a similar pattern. The 49 kDa mutant VP4-RK protein remained intact when no enzyme was added (lane 1).

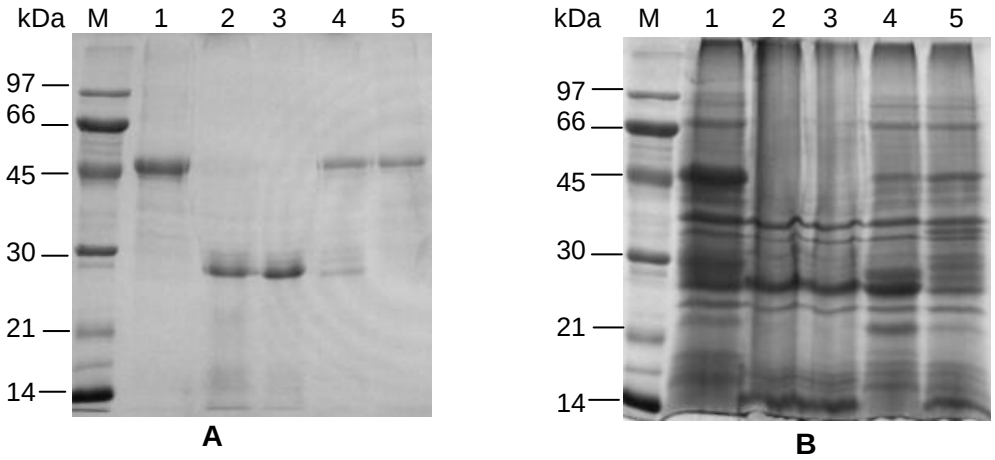


Figure 3.19 *Trans*-cleavage activity of IBDV VP4-AA using Ni-NTA purified mutant VP4-RK sample (A) and mutant VP4-RK refolded sample (B) as substrates. Lane M, Low range MWM; lane 1, no enzyme control; lane 2, VP4-AA sample in storage buffer (without assay buffer); lane 3, VP4-AA in the presence of the assay buffer; lane 4, VP4-AA sample pre-incubated with anti-thrombin III; lane 5, thrombin.

The addition of purified VP4-AA led to the hydrolysis of mutant VP4-RK as evidenced by the presence of a 26 kDa protein band, in the presence (lane 2) and absence (lane 3) of the assay buffer (Fig. 3.19A and B). Mutant VP4-RK purified under denaturing conditions was stored in a buffer containing 8 M urea, therefore the assay without the addition of assay buffer would be done in the presence of 4 M urea as compared to 2 M urea when the assay buffer was added. This suggests that VP4 was able to cleave in the presence of a denaturant.

The pre-incubation of VP4-AA with anti-thrombin III inhibited cleavage of mutant VP4-RK (lane 4). Mutant VP4-RK was expressed as a fusion protein with an N-terminal Trx-His tag, containing thrombin and enterokinase sites between the tag and the protein. Surprisingly, the thrombin control did not cleave both substrates (lane 5). The presence of

2 M urea in the sample of mutant VP4-RK purified under denaturing conditions should enhance the thrombin activity according to the manufacturer. The absence of activity could have resulted from the use of insufficient thrombin or a loss of the site during PCR.

To further verify that the observed cleavage was due to VP4, an experiment was set up where the mutant VP4-RK was first cleaved with each of the same enzyme samples as before (VP4-AA, VP4-AA pre-incubated with anti-thrombin III and thrombin) followed by incubation with enterokinase (Fig. 3.20). The rationale for this experimental design was that, should VP4-AA cleave mutant VP4-RK at the dialanine peptides, the resulting 24-26 kDa product would not be cleaved by enterokinase, whereas the 34 kDa thrombin cleavage product would still contain an enterokinase cleavage site (Fig. 3.20A). Cleavage of the 34 kDa product by enterokinase would result in a 30 kDa product. Since all these products contain the VP4 common peptide, they could be visualised on a western blot (Fig. 3.20D).

The uncleaved mutant VP4-RK was observed at about 49 kDa (Fig. 3.20C and D, lane 2) and after incubation with enterokinase a 30 kDa product was observed (lane 3) although the hydrolysis was incomplete as evidenced by the presence of a 49 kDa band. Mutant VP4-RK incubated with VP4-AA before and after incubation with enterokinase resulted in a 26 kDa protein (lanes 4 and 5), which is the size of the expected product. Had the cleavage been a result of thrombin activity, a 34 kDa would have been produced before addition of enterokinase. When VP4-AA was preincubated with anti-thrombin III, the cleavage was partially inhibited as shown by the presence of the 49 kDa uncleaved mutant VP4-RK and the 26 kDa VP4-AA protein band (lane 6). The addition of enterokinase simply cleaved the intact 49 kDa into the 34 kDa product (lane 7). Thrombin gave similar results to the no enzyme control, even though the amount of thrombin was increased (lanes 8 and 9). It was thus concluded that the thrombin cleavage site may have been inaccessible in the mutant VP4-RK. A parallel reaction using the enterokinase cleavage control protein (supplied with kit) was included to serve as a positive control. Enterokinase cleaves the 49 kDa control protein into two bands of 32 kDa and 16 kDa (lane 1).

The experimental design had some drawbacks, for example, the cleavage of mutant VP4-RK by VP4-AA would produce a C-terminally 6xHis-tagged truncated VP4 (24-26 kDa). In other words, both enzyme and product would have similar mobilities on SDS-PAGE. Secondly, enterokinase also migrates as a single band with a size of approximately 30

kDa. It was therefore necessary to analyse the cleavage products by western blot using anti-VP4 peptide antibodies.

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGT

Thrombin

APKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPR↓GSGM

Enterokinase

KETAAAKFERQHMDSPDLGTDDDDK↓AMAISDPN**R**IAVPVVSTLFPPAAPLAHAIGEGVDYLLGDEA

VP4-AA

QA↓ASGTARA↓ASGKARA↓ASGRIRQLTLA↓ADKGYEVVANLFQVPQNPVVDGILASPGVLRGAHNLD
CVLREGATLFPVVITTVEDAMTPKALNSKIFAVIEGVREDLQPPSQRGSFIRTLSGHRVYGYAPDGV
LPLETGRDYTVVPIDDVWDDSIMLSKDPIPPIVGNSGNLAIAYMDVFRPKVPIHVAMTGALNACGEI
EKVSFRSTKLATAHRLGLKLAGPGAFDVNTGPNWATFI**K**AALGHHHHHH

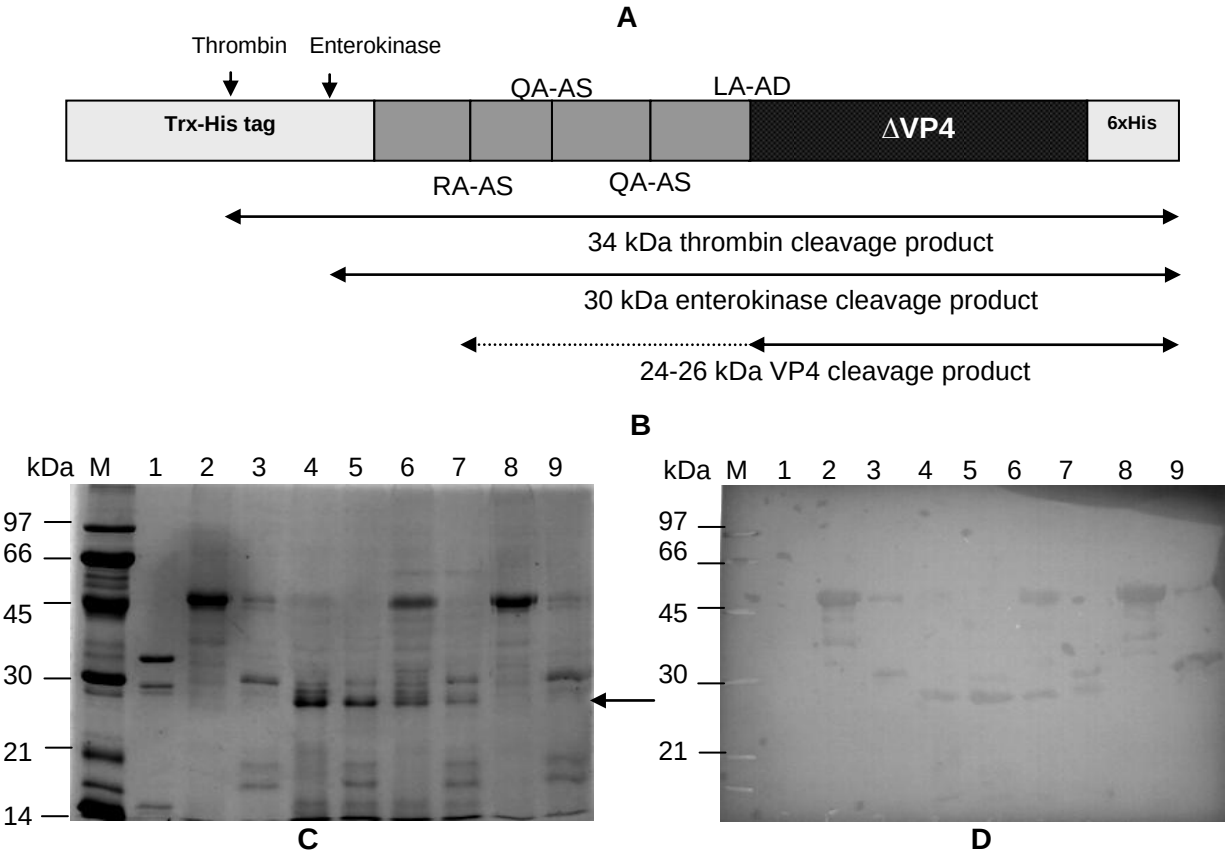


Figure 3.20 The analysis of the VP4-AA t*rans*-cleavage activity on the mutant VP4-RK fusion protein by enterokinase cleavage assay. A) The sequence of the mutant VP4-RK fusion protein with the cleavage sites for thrombin, enterokinase and VP4-AA underlined and indicated with an arrow. The N-terminal Arg and C-terminal Lys residues of VP4-RK are indicated in bold. **B)** The schematic diagram of the VP4-RK constructs showing the cleavage sites and the expected product sizes. For VP4, cleavage at any dialanyl peptides will produce a 24-26 kDa product. The mutant VP4-RK was incubated with purified VP4-AA or the necessary controls, followed by incubation with enterokinase. The hydrolysis products were analysed by reducing SDS-PAGE **(C)** and western blotting using anti-VP4 common peptide antibodies **(D)**. Lane M, Low range MWM; lanes 1, enterokinase cleavage positive control; Mutant VP4-RK was incubated with no enzyme (lanes 2-3); VP4-AA (lanes 4-5); VP4-AA pre-incubated with anti-thrombin III (lane 6-7); thrombin (lanes 8-9). Lanes 2, 4, 6 and 8 represent the samples before enterokinase cleavage while lanes 3, 5, 7 and 9 show the samples after enterokinase cleavage. The arrow indicates the size of the VP4-AA cleavage product at 26 kDa.

The uncleaved mutant VP4-RK protein at 49 kDa (lanes 2, 3, 6 and 8), VP4-AA cleavage product at 26 kDa (lanes 4 and 5) and enterokinase cleavage product at 34 kDa (lanes 3 and 9) were detected by the anti-VP4 common peptide antibodies as expected (Fig. 3.20D). The 28 kDa protein recognised by the anti-VP4 peptide common antibodies in lanes 6 and 7 represent the VP4-AA enzyme. This data further supports that VP4-AA cleaves the mutant VP4-RK in *trans* and may be inhibited by anti-thrombin III.

3.4.8.3 VP4-AA activity against synthetic substrates

Initial screening of VP4-AA hydrolysis of peptide substrates showed activity against H-Val-Leu-Lys-AMC, Boc-Leu-Gly-Arg-AMC and Z-Gly-Pro-Arg-AMC but no activity against H-Ala-AMC, H-Ala-Phe-Lys-AMC, Boc-Leu-Arg-Arg-AMC and Z-Phe-Arg-AMC (results not shown). All the hydrolysed substrates contained a basic residue such as arginine or lysine at P_1 and an aliphatic residue like glycine or leucine in P_2 and P_3. Substrates containing a bulky aromatic residues like phenylalanine or basic residues in P_2 were not hydrolysed. The assay was repeated with the substrates that were hydrolysed, in the presence or absence of anti-thrombin III or thrombin as a control (Fig. 3.21).

High background fluorescence was observed with the substrate H-Val-Leu-Lys-AMC. When VP4-AA was preincubated with antithrombin III, the activity was inhibited and with all the substrates the fluorescence was similar to that of no enzyme control. Thrombin hydrolysed H-Val-Leu-Lys-AMC and Boc-Leu-Gly-Arg-AMC much better than Z-Gly-Pro-Arg-AMC. Substrates containing Gly-Pro-Arg are generally known as Chromozym-TH, and their composition resemble the sequence in fibrinogen, which is susceptible to thrombin, although at a slower rate (Izquierdo and Burguillo, 1989). These three substrates have similar residues, a basic residue in P_1, hydrophobic residues in both P_2 and P_3, thus rendering them likely to be cleaved by thrombin (Morita *et al.*, 1977; Izquierdo and Burguillo, 1989). There were significant differences between the peptidolytic activies of VP4-AA and VP4-AA preincubated with antithrombin III for Boc-Leu-Gly-Arg-AMC and Z-Gly-Pro-Arg-AMC. Even though the substrates were cleaved by the thrombin control, considering the relative amounts of VP4-AA (20 ng) and thrombin (≤ 0.0006 ng), it could be concluded that the activity observed in the VP4-AA purified sample was due to VP4-AA. The results also suggest that VP4-AA is inhibited by anti-thrombin III.

VP4 protease shows preference for Ala at both P_1 and P_1' positions thus the activity of VP4-AA was also evaluated on a fluorogenic substrate containing Ala at P_1 position, Z-Gln-Ala-Ala-AMC. Surprisingly no activity was observed on this substrate. To confirm that

VP4-AA was not hydrolysing the substrate, trypanosome parasite lysate was used as a positive control (Fig. 3.22).

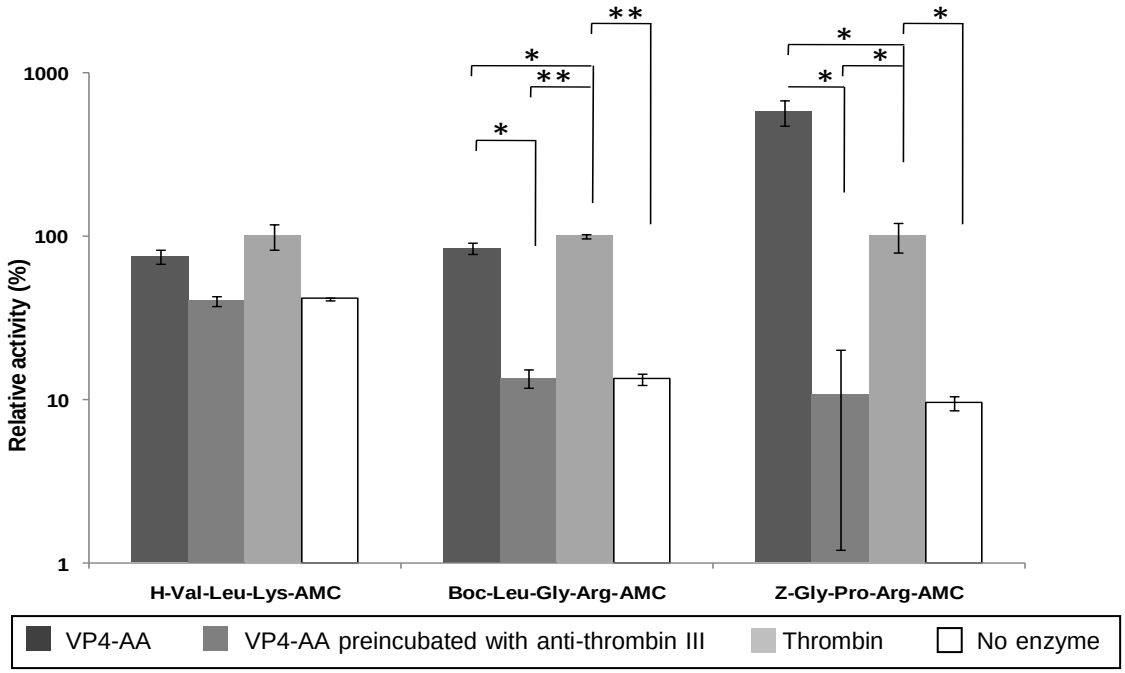


Figure 3.21 Peptidolytic activity of VP4-AA. The assay was carried out either using purified VP4-AA, thrombin, VP4-AA preincubated with anti-thrombin or no enzyme control incubated with each substrate. The activity presented was expressed as a mean of triplicate fluorescence values plotted relative to thrombin as a percentage for each substrate ±SD.*p<0.05 and **p<0.001.

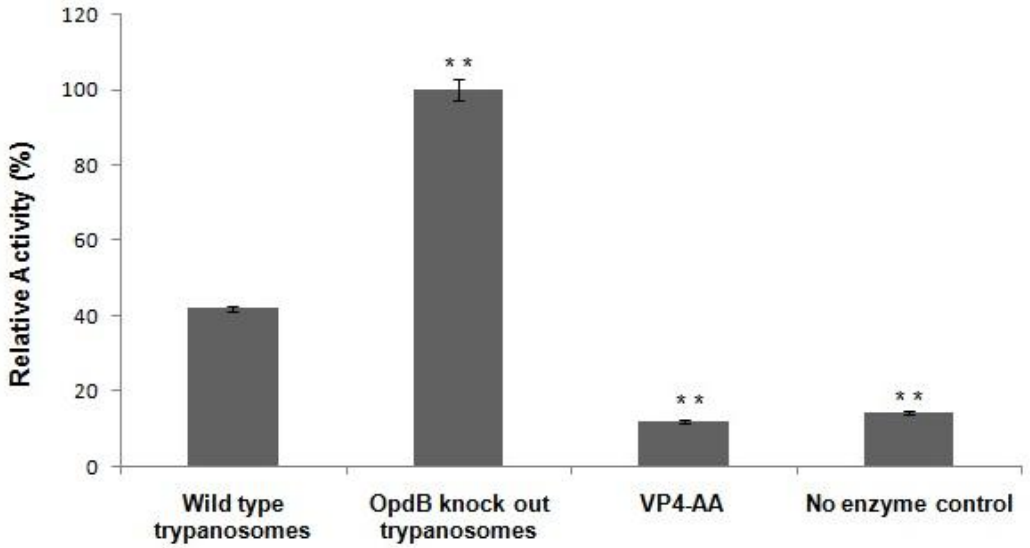


Figure 3.22 Comparison of Z-Gln-Ala-Ala-AMC hydrolysis by VP4-AA and trypanosome parasites. The data is presented as means ± SD (n = 3) of fluorescence values plotted relative to wild-type trypanosomes as a percentage. **p < 0.001 in comparison to wild-type trypanosomes.

Trypanosome parasites have serine peptidases such as oligopeptidase B (OpdB) and prolyl oligopeptidase (POP) that are released into host plasma to degrade host factors

such as hormones (Coetzer *et al.*, 2008; Bastos *et al.*, 2010). A study in our laboratory showed that OodB knock out trypanosome shows an up-regulation of POP activity (Kangethe *et al.*, 2012). POP hydrolyses substrates containing Pro or Ala in the P_1 position (Bastos *et al.*, 2010). POP activity in the knock-out parasites was higher than the activity in the wild type parasites, as expected. VP4-AA on the other hand displayed no activity thus confirming that it was not hydrolysing the substrate.

3.4.9 Protease activity in infected and non-infected bursa samples

Since recombinant VP4 was shown to be active on gelatine-containing SDS-PAGE, IBDV infected and non-infected bursal samples were subjected to the same analysis, to evaluate the proteolytic activity of native VP4 (Fig. 3.23). The proteolytic activity of recombinant VP4 (28 kDa) was evidenced by a clear zone on the gelatine gel at approximately 56 kDa (Fig. 3.18). There was no activity observed at either 56 or 28 kDa in the infected bursa samples that would correspond to the activity of native VP4 (Fig. 3.23).

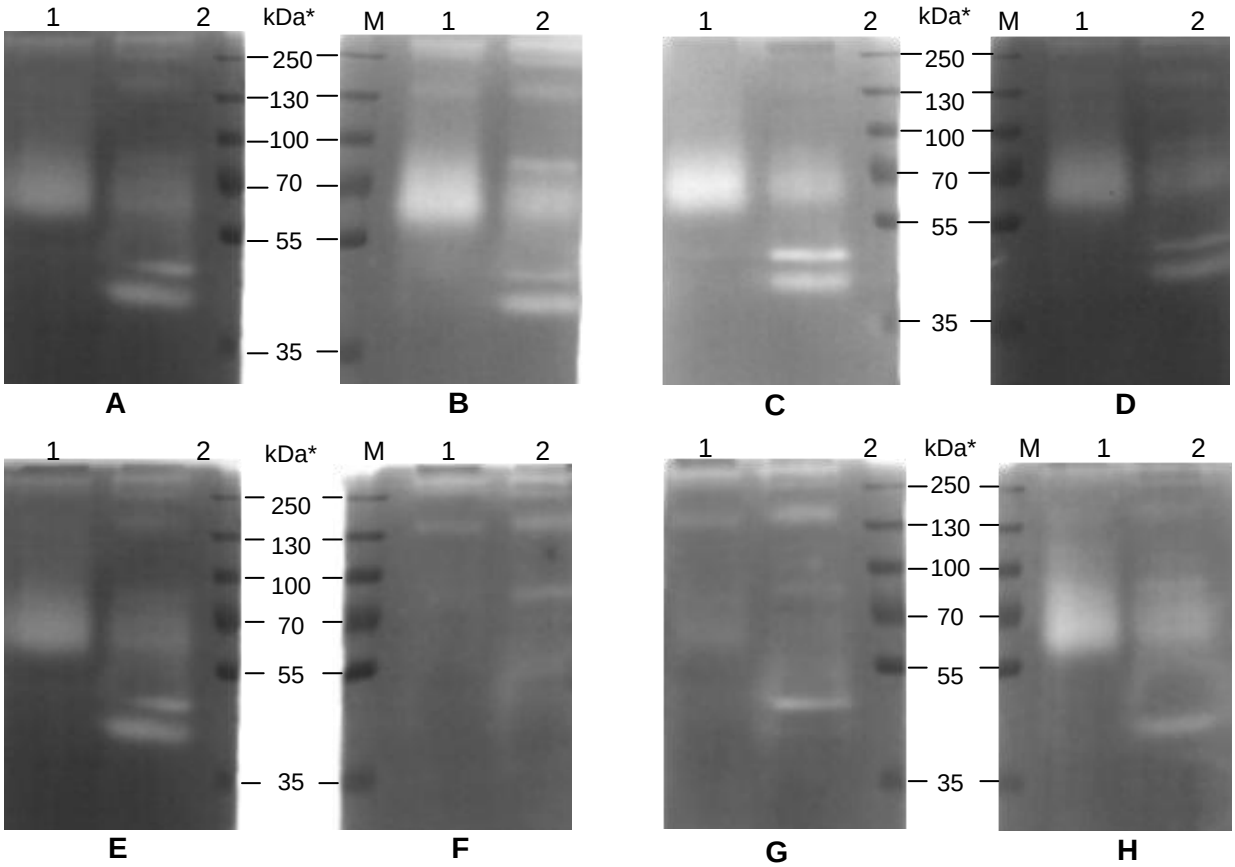


Figure 3.23 Comparison of protease activity on gelatine gels between IBDV infected and non-infected bursal samples. About 15 µg bursal homogenate total proteins was loaded per well on a gelatine gel and incubated overnight in assay buffer with **A and E)** No inhibitor; **B)** No EDTA; **C)** 1 µM E-64; **D)** 100 µM TPCK; **F)** 5 µg/ml SBTI; **G)** 1 mM AEBSF and **H)** 1 µg/ml Pepstatin A. Lane M, PageRuler™ plus prestained protein ladder; lane 1, infected bursal homogenate sample; lane 2, non-infected bursal homogenate sample. * Note that kDa sizes are approximate under non-reducing conditions.

However, there was an apparent difference in the clearing on the gelatine gel between the infected and non-infected bursal samples (Fig. 3.23A, E). The difference was in the two bands of hydrolysis at about 45 kDa seen in the non-infected samples, but not in the infected sample. In order to identify the class of proteases, the gelatine gels were incubated with class-specific protease inhibitors. The assay buffer contained EDTA and since metalloproteases are sensitive to metal-chelating agents, an EDTA free buffer was used as a control. The profile of gelatine clearing in the EDTA containing buffer was similar to that without EDTA (Fig. 3.23B). In the presence of E64, an irreversible inhibitor of cysteine proteases, digestion of gelatine between 130 and 250 kDa was inhibited (Fig. 3.23C). TPCK is an irreversible inhibitor of chymotrypsin and some cysteine proteases. The gelatine hydrolysis profile on the gels incubated with TPCK was not significantly different from gels incubated in the absence of inhibitors (Fig. 3.23D). SBTI and AEBSF inhibited the clearing of the bands at about 70 kDa in both the infected and non-infected samples. While SBTI also inhibited the two bands at 45 and 40 kDa in the non-infected samples, AEBSF inhibited the activity of the 40 kDa band (Fig. 3.23Fand G). Pepstatin A, an inhibitor of aspartic proteases seemed to inhibit the 45 kDa activity and the high molecular weight proteins (Fig. 3.23H).

3.5 Discussion

The present study revisited the processing of the IBDV polyprotein *in vitro.* Previous studies have reported that the precursor polyprotein is co-translationally processed by the viral-encoded protease, VP4 at pVP2-VP4 (^{512}Ala-Ala^{513}) and VP4-VP3 (^{755}Ala-Ala^{756}) junctions, to release pVP2, VP4 and VP3 (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000). Site-directed mutagenesis studies were used to demonstrate that VP4 is a serine protease that uses a Ser/Lys catalytic dyad (Lejal *et al.*, 2000), and that the cleavage sites consist of Ala-Ala in the P_1 and P_1' positions (Sanchez and Rodriguez, 1999).

Some events surrounding the cleavages are still not well understood. Firstly, it is not well understood whether the processing of pVP2 at ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502} is in *cis* or t*rans*. Secondly, it is clear that VP4 exists as an integral part of the polyprotein (embedded) and after processing the mature form is released. However, the exact stretch of amino acids within the polyprotein that make up the embedded protease has not been unequivocally established. Thirdly, processing of the polyprotein at one site may be a prerequisite for processing at another site (Bartenschlager *et al.*, 1995) but the order in

which processing of the IBDV polyprotein occurs has not yet been described. The present study was conducted in order to shed some light on some of the poorly understood aspects of the IBDV polyprotein processing. To this end different forms of the polyprotein were recombinantly expressed and the processing events investigated.

In the present study, constructs were designed for recombinant expression of the full-length polyprotein (Met^{1}-Glu^{1012}) and a truncated form of the polyprotein (Ile^{227}-Trp^{891}) representing the fully embedded VP4 form, VP4-RA (Arg^{453}-Ala^{755}) and VP4-RK (Arg^{453}-Lys^{722}) which contained the N-terminal cleavage sites namely ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495}, ^{501}Ala-Ala^{502} and ^{512}Ala-Ala^{513}, VP4-AW (Ala^{513}-Trp^{891}) with the C-terminal cleavage site ^{755}Ala-Ala^{756} and VP4-AA (Ala^{513}-Ala^{755}), the mature form without cleavage sites. The IBDV polyprotein was previously shown to undergo proteolytic processing when expressed in *E. coli* and mammalian systems, and anti-VP2 or anti-VP3 antibodies or SDS-PAGE autoradiography were used to identify the cleavage products (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000). It is believed that these methods would not be able to identify all the resulting products. For this reason, anti-VP4 peptide antibodies were used in this study for the detection of the cleavage products that could have been missed in previous studies.

The constructs were cloned, sequenced (Chapter 2) and expressed. Since it was previously observed that the IBDV VP4 protease is active in *E. coli* (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000), this expression system was chosen. Even though IBDV proteins were shown not to be glycosylated (Kibenge *et al*., 1997), the IBDV polyprotein sequence contains three potential N-glycosylation sites in the VP2 region. In addition, IBDV replicates in eukaryotic cells. For these reasons, the expression was also attempted in the eukaryotic system, *P. pastoris,* as this might better support the proper folding of the constructs than the *E. coli* expression system. Unfortunately the expression in *Pichia* was unsuccessful. For *Plasmodium* proteins, unfavorable codon bias of the AT rich regions cause premature termination of transcripts (Goh *et al.*, 2003). In the case of IBDV polyprotein constructs, the reason for the failure to express was unknown especially since Wu *et al.* (2005a) successfully expressed VP2 using this system.

As a result, expression was conducted in *E. coli*. Although the *E. coli* system is still an attractive source of heterologous protein expression, one of its drawbacks is the inability to produce soluble, correctly folded protein, thus leading to very low levels of expression or formation of inclusion bodies (Sorensen and Mortensen, 2005a). The present study

explored the possibilities of improving the solubility of the constructs via the use of different *E. coli* based expression systems (pET-32b, pET-28b and pGEX-4T-1) and manipulation of the expression conditions. Attempts to express using the pET-28b expression system were unsuccessful. The pGEX-4T-1 was able to express the VP4 constructs but failed to express the polyprotein constructs probably due to their sizes.

In the first instance, the pET-32b vector based on the T7 expression system was able to express all the constructs. The fact that expression was successful in pET-32b and failed in pET-28b was interesting and surprising since the only major difference between the two vectors lies in the antibiotic resistance. This observation has not been reported elsewhere. Efforts to produce soluble proteins in *E. coli* were unsuccessful for all constructs in the pET-32b system. This problem can sometimes be overcome by fusing the protein to highly soluble tags such as GST and Trx, which can enhance the solubility of foreign proteins (Sorensen and Mortensen, 2005b). Even though the proteins were expressed as fusion proteins with a tag, inclusion bodies were still formed regardless of the culture conditions. The proteins were expressed as inclusion bodies but processing was still observed in the expressed constructs. It is not known whether the processing of the IBDV polyprotein is determined by protein folding. If so, although uncommon, it is possible for inclusion bodies to be similar to the native form of the protein (Burgess, 2009).

The resulting cleavage products were detected by anti-VP4 peptide antibodies. The use of synthetic peptides to generate antibodies has increased in recent years and has proven successful. In light of the success in our laboratory (Coetzer *et al.*, 1991; Troeberg *et al.*, 1997; Hurdayal *et al.*, 2010) and other groups (Lee and Watson, 2002; Hirani *et al.*, 2008) to name a few, the anti-peptide antibody approach was used in the present study. Furthermore, to our knowledge, this is the first study using anti-VP4 peptide antibodies for the characterisation of polyprotein processing in IBDV. Several studies employed monoclonal or polyclonal antibodies against VP4 (Granzow *et al.*, 1997; Chevalier *et al.*, 2002; Wang *et al.*, 2009), but none has reported on anti-VP4 peptide antibodies.

The anti-peptide antibody approach was chosen due to its advantages. Firstly, this approach provides the means to produce antibodies before purified antigen is obtained through purification from the biological source or through recombinant expression and purification (Adams *et al.*, 1997). Secondly, the antibodies are directed to a dominant antigenic site of the target protein (Adams *et al.*, 1997). This predetermined site-specificity is one feature of anti-peptide antibodies that has led to its application in a variety of studies (Schulz-Utermoehl *et al.*, 2000; Lee and Watson, 2002; Saravanan *et al.*, 2004).

This was particularly important in the present study since the aim was to characterise the embedded and mature forms of VP4. The IBDV polyprotein has several cleavage sites, which makes the exact location of the processing site unclear, and thus the site-specific antibodies would shed more light on the processing events.

The most important step for the production of anti-peptide antibodies is the identification of suitable epitopes. Traditional immunological methods for epitope identification are time consuming (De Groot *et al.*, 2001; Morris, 2008) and have been replaced by computational analysis of protein sequences with epitope prediction tools, that provides an effective strategy for rapid identification of epitopes from different proteins (Sanchez-Burgos *et al.*, 2010). For identification of epitopes, the primary sequence of VP4 retrieved from the database was analysed for secondary structure, hydrophilicity, flexibility, surface probability and antigenicity (Cármenes *et al.*, 1989). Three peptides were selected, the first peptide located at the N-terminus of the pVP2-VP4 junction was specific to VP4-RK, the second located at the N-terminus of the VP4-VP3 junction was specific to VP4-AA, while the last peptide was within the VP4 region thus being a peptide common to both VP4-RK and VP4-AA.

The peptide-conjugates were successfully used to produce specific chicken anti-VP4 peptide antibodies that were able to recognise the cleavage products resulting from the processing of the polyprotein and the various VP4 constructs. The expression of the full-length and truncated polyproteins yielded high molecular weight proteins and cleavage products, which were identified by these antibodies. The high molecular weight proteins were thought to be multimers of the polyprotein. This was unexpected since the processing was described as co-translational (Hudson *et al.* 1986; Jagadish *et al.* 1988), implying that the processing is carried out as the polyprotein is being translated. The assembly of these multimers may thus oppose the idea of co-translational processing but suggest post-translational processing. The absence of a 130 kDa band may suggest that the polyprotein multimerisation may be a mechanism to prevent processing.

With IBDV polyprotein supposedly cleaved at ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} to release pVP2, VP4 and VP3, the anti-VP4 common peptide and anti-VP4-AA peptide antibodies should only recognise VP4 while the anti-VP4-RK peptide antibodies would recognise pVP2. If pVP2 cleaves itself at ^{441}Ala-Phe^{442} to release VP2 and the C-terminal pep70 (Irigoyen *et al.,* 2009), the anti-VP4-RK peptide antibodies would then not be able to recognise pVP2 (Fig. 3.7B). Therefore, these anti-VP4 peptide antibodies would recognise only two different cleavage products, pVP2 (if it was not processed) and VP4. The anti-

VP4 common peptide antibodies detected more products than expected, suggesting that other cleavage sites were used. Sanchez and Rodriguez (1999) showed that deletion of the ^{512}Ala-Ala^{513} site did not completely inhibit the processing at the pVP2-VP4 junction, indicating that VP4 may use other cleavage sites.

Both VP4-RA (Arg^{453}-Ala^{755}) and VP4-RK (Arg^{453}-Lys^{722}) contained the N-terminal cleavage sites, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495}, ^{501}Ala-Ala^{502} and ^{512}Ala-Ala^{513}. These fragments were expected to express as fusion proteins with an N-terminal Trx-His tag of 17 kDa, but instead, they were expressed as cleavage products. There was no uncleaved protein detected, suggesting that the cleavage was mediated in *cis* (intramolecular). For VP4-RA, cleavage occurring at ^{512}Ala-Ala^{513} would result in two products, an N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide and a C-terminally 6xHis tagged VP4-AA (Ala^{513}-Ala^{755}). The latter should be recognised by both anti-VP4 common peptide and anti-VP4-AA peptide antibodies, while the anti-VP4-RK antibodies would recognise the Trx-His tagged polypeptide. Similarly, for VP4-RK (Arg^{453}-Lys^{722}), cleavage at ^{512}Ala-Ala^{513} would produce the N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide and a C-terminally 6xHis tagged VP4-AK (Ala^{513}-Lys^{722}). The latter is recognised only by the anti-VP4-AA peptide antibodies, while the anti-VP4-RK antibodies would recognise the Trx-His tagged polypeptide.

For both VP4-RA and VP4-RK, the cleavage products corresponding to the VP4 region were recognised by anti-VP4-RK peptide antibodies. Once again, this suggested that the cleavage at the N-terminus was occurring at another site, namely ^{487}Ala-Ala^{488}, which consequently opposes the currently established polyprotein processing strategy. However, N-terminal sequencing of the cleavage products corresponding to the VP4 region showed that VP4 uses ^{512}Ala-Ala^{513} site even in the presence of the other three sites. Further investigation was therefore required to clarify these conflicting results.

Autocatalytic activity was observed with the expression of full-length and truncated polyprotein, VP4-RA and VP4-RK. The processing of VP4-AW (Ala^{513}-Trp^{891}) seemed to be intermolecular. VP4-AA, on the other hand, was expressed as a fusion protein. This implies that the embedded VP4 autocatalytically cleaves itself off to release the mature form (VP4-AA). In light of this, VP4-AA is simply a product of the autocatalytically active VP4, which is embedded within the polyprotein. The stretch of amino acids within the polyprotein that make up the embedded protease are not known, but these results suggest that the autocatalytically active VP4 sequence may include the Arg^{453}-Ala^{512} polypeptide. Since autocatalytic activity was only observed in the constructs containing

the Arg^{453}-Ala^{512} polypeptide, it was suggested that VP4 protease may be similar to *Flaviviridae* NS3pro which requires the presence of a cofactor to be fully functional (Yusof *et al.*, 2000; Niyomrattanakit *et al.*, 2004; Bessaud *et al.*, 2006). In flaviviruses like Dengue virus, the NS3 requires NS2B which is located N-terminal of NS3 for efficient proteolytic processing The NS2B functions as a cofactor and promotes the folding and functional activity of NS3pro. The NS3 protease domain lacking the NS2B part is catalytically inactive (Erbel *et al.*, 2006; Luo *et al.*, 2008).

The comparison of the processing products from the expression of VP4-RA and VP4-AW suggested that the cleavage may initially occur at the VP4-VP3 junction (^{755}Ala-Ala^{756}) followed by cleavage at the pVP2-VP4 junction (^{512}Ala-Ala^{513}) in order to support the idea of the polyprotein being autocatalytically processed. Cleavage of the polyprotein initially occurring at the VP4-VP3 junction would yield a fragment containing the Arg^{453}-Ala^{512} polypeptide, which render VP4 to still be autocatalytic. On the other hand, cleavage initially occurring at the pVP2-VP4 junction would result in a VP4 equivalent to VP4-AW, which loses the autocatalytic activity.

For further analysis, attempts were made to purify the polyprotein multimers (embedded form) and VP4-AA (mature form). The polyprotein in *E. coli* BL21 (DE3) was expressed as insoluble inclusion bodies using pET-32b. It was necessary to first solubilise the inclusion bodies. Purification was attempted with both solubilised protein, and in other cases, the solubilised protein was refolded before the purification. Purification before and after refolding by dilution and dialysis, proved unsuccessful. VP4-AA, on the other hand, was expressed as a soluble protein using the pGEX-4T-1 expression vector and was therefore successfully purified.

Even though VP4-AA does not display the autocatalytic activity observed in the embedded form of VP4, it was of interest to determine whether it can still act in *trans* or not. Although, the mature form of VP4 was shown to intermolecularly cleave a VP4-VP3 junction (Lejal *et al*., 2000), it has not been tested for activity on zymograms or against synthetic peptide substrates. It was thus analysed for intermolecular activity on a gelatine-containing SDS-PAGE gel and against fluorogenic peptide substrates. Activity of the mutant VP4-RK was included as a positive control, as this would mimic the study by Lejal *et al*. (2000). Thrombin used for on-column cleavage was not removed after purification and this seemed to compromise the obtained results. With sufficient controls included, it could be concluded that VP4-AA was active on a gelatine gel and also on the mutant VP4-RK which would mimic its natural substrate. The activity of VP4-AA also seemed to be

inhibited by anti-thrombin III. It was also interesting that *trans* proteolytic activity was observed acting on the mutant VP4-RK in the presence of urea. This has not yet been reported for any of the birnavirus VP4 proteases. White *et al.* (1995) also reported on a bacterial outer membrane protein T which was active in the presence of a denaturing agent.

On the gelatine gel, VP4-AA seemed active as a dimer. This was very interesting since a membrane bound serine protease that also uses a Ser-Lys catalytic dyad has a novel dimer structure (Yokoyama *et al.*, 2008). In addition, VP4-AA was not active as a fusion protein. Similar observations were made for the *Plasmodium falciparum* cysteine protease, falcipain-1, which was not active when fused to maltose binding protein (MBP), but showed activity when the MBP was cleaved off with factor Xa protease (Goh *et al.*, 2005). Regarding its activity against fluorogenic substrates, VP4 requires more than just dialanine peptides for activity. This can be supported by the fact that IBDV polyprotein sequence contains other dialanine peptides which are not cleaved by VP4. Just like other viral proteases, VP4 shows very stringent substrate specificities in both *cis* and *trans* activities.

The chicken anti-VP4 peptide antibodies were successfully used to identify cleavage products on a western blot and in so doing, revealed even those products that were not identified by the earlier methods of detection. The antibodies were also tested for the recognition of purified VP4-AA in an ELISA. This was done to establish whether the peptides used for antibody production were located in accessible sites of the protein. Although, anti-VP4 common and anti-VP4-AA peptides recognised the recombinant protein, the recognition was poor relative to the recognition of the peptides alone. The reason for the poor binding could be due to the expected differences in the 3D conformation of the peptides as a conjugate relative to its conformation in the protein itself. The antibodies' ability to recognise denatured proteins in western blot shows that the peptides conformation better resembled the region of the protein in its denatured state. It also has been reported that SDS used in electrophoresis increases the helical content of proteins which can also influence the binding of anti-peptide antibodies (Myers *et al.*, 1997).

Attempts to recognise native VP4 in infected bursal samples were not very successful, as the results were not reproducible. There were some protein bands that were recognised by anti-VP4 common peptide and anti-VP4-AA peptide antibodies. It is possible that the 28 kDa protein doublet detected by anti-VP4 common peptide antibodies could have been VP4. Wang *et al.* (2009) produced monoclonal antibodies to determine the expression

profile of VP4 during pathogenic and non-pathogenic IBDV infection. It was shown that VP4 can be detected mainly in the cortex of lymphoid follicles of the bursa of Fabricius infected with a pathogenic IBDV strain. This was in agreement with the previous report that VP4 is detected in infected cells but not as part of the virus particle (Granzow *et al.*, 1997). Anti-VP4-AA peptide antibodies recognised proteins that did not correspond to the expected sizes of IBDV proteins. *Ensemble* BLAST showed that the peptide (FIKRFPHNPRDW) is highly similar to peptides in chicken proteins (Table 3.3), which would explain why the antibodies would recognise chicken proteins.

Table 3.3 Alignment of the VP4-AA peptide to peptides from chicken proteins

Ensembl protein code	**VP4-AA peptide residues***												
	F	**I**	**K**	**R**	**F**	**P**	**H**	**N**	**P**	**R**	**D**	**W**	**D**
ENSGALP00000013949 Chloride channel	**F**	L	**K**	**R**	**F**	**P**							
ENSGALP00000036766 Xeroderma pigmentosum						**P**	**H**	**N**	**P**	**R**			
ENSGALP00000005610 Uncharacterised	**F**	**I**	**K**	**R**	**F**	H	**H**	T	R	**R**	**D**		
ENSGALP00000021094 Importin 8				**R**	**F**	**P**	**H**	**N**	**P**				

*The residues in bold match the peptide

Attempts to detect native VP4 in the infected samples was also carried out by proteolytic activity determination on gelatine-containing SDS-PAGE. The proteolytic activity of the 28 kDa recombinant VP4 was observed at about 56 kDa, but there were no clear zones of hydrolysis observed at either 28 or 56 kDa in infected bursal homogenates. However, there were clear differences between the proteolytic activity profile of the non-infected and infected bursal homogenates. The activity was due to three classes of proteases, i.e. serine, aspartic and cysteine proteases. There was no difference in the activity of cysteine proteases between infected and non-infected samples as was shown by E64 inhibition. Major differences were observed in the activity profiles of aspartic and serine proteases between infected and non-infected samples. One of the serine protease activities inhibited by SBTI seemed to be consistently more active in the infected than in the non-infected bursal homogenates. SBTI also inhibited a doublet at 40 and 45 kDa that was present and active only in the non-infected bursal samples. The 45 kDa band of the doublet was also inhibited by AEBSF, thus it could be confidently classified as a serine protease. The 40 kDa band, on the other hand was also inhibited by pepstatin A, an inhibitor of aspartic proteases. Its inhibition by both SBTI and pepstatin A made the classification uncertain.

Virus infection causes alteration in cellular processes. IBDV infection has been shown to change the potassium current properties in CEFs (Repp *et al.*, 1998), induce elevated levels of γ-interferon (Kim *et al.*, 2000) and induce apoptosis. Proteomics analysis of host cell responses to virus infected cells showed differential expression of proteins involved in the ubiquitin-proteasome pathway, cytoskeleton proteins, stress response, RNA processing and biosystems, protein translation, metabolism associated proteins (Zheng *et al.*, 2008). Neither cellular protease involvement on IBDV polyprotein processing nor the effect of IBDV infection on protease expression in the bursa of Fabricius has been reported. The observed results showed that there is a clear difference between proteolytic activity of infected and non-infected bursal homogenates. It is not known whether the observed difference in the protease activity between infected and non-infected bursal homogenates is a biological effect of IBDV infection or not. Host cellular proteases have been shown to play a role in polyprotein processing, assembly and entry of some picornaviruses (Morace *et al.*, 2008) and influenza virus (Chaipan *et al.*, 2009).

In summary, the expression of the polyprotein, VP4-RA, VP4-RK and VP4-AW resulted in cleavage products, while VP4-AA was expressed as an uncleaved fusion protein. VP4 is a serine protease that utilises the Ser/Lys catalytic dyad (Lejal *et al.,* 2000) and shows preference for Ala-Ala dipeptides (Sanchez and Rodriguez, 1999). The catalytic dyad and the cleavage sites are also characteristic of the Lon proteases, Lex A and leader proteases of *E. coli* (Paetzel and Dalbey, 1997), therefore it could not be ruled out that these bacterial proteases could have been cleaving the expressed proteins, especially since all the proteins that were cleaved contained the cleavage sites. It was therefore necessary to confirm that the observed processing was due to the autocatalytic activity of the embedded VP4 protease. In addition, further investigation was required to confirm that the pep70 residues are required as a cofactor for the autocatalytic activity of VP4. These results are reported in Chapter 4 and 5.

For the expression of VP4-RA and VP4-RK, western blot analysis using anti-VP4 peptide antibodies suggested that N-terminal cleavage occurred at the ^{487}Ala-Ala^{488} while N-terminal sequencing of the products showed it occurs at the ^{512}Ala-Ala^{513} site. Further investigation on the processing of these sites is reported in Chapter 5.

CHAPTER 4
Site-directed mutagenesis studies of the IBDV polyprotein constructs

4.1 Introduction

The IBDV polyprotein (NH_2-pVP2-VP4-VP3-COOH) is processed at the pVP2-VP4 and the VP4-VP3 junctions by VP4 protease (Sanchez and Rodriguez, 1999). Site-directed mutagenesis was previously used to show that the birnavirus VP4 protease uses a Ser/Lys catalytic dyad. The replacement of Ser^{652} (by Ala or Thr) and Lys^{692} (by Ala, Arg or His) completely inactivated polyprotein processing (Lejal *et al.*, 2000). The constructs prepared in the present study representing the embedded form of VP4 showed autocatalytic activity during expression in *E. coli* (Chapter 3). The present study was designed to rule out the possibility of cleavage by bacterial proteases, by mutating the Ser^{652}/Lys^{692} catalytic dyad in the constructs followed by expression and characterisation of the resulting mutant constructs.

A further consideration was that in IBDV, the pVP2 to VP2 conversion was initially observed in infected cells and was shown to involve one (or multiple) cleavage(s) of pVP2 at the C-terminus, and the processing was proposed to be associated with VP4 protease activity (Azad *et al.*, 1987; Kibenge *et al.*, 1997). The identification of the three vicinal dialanines at the C-terminus of pVP2 supported the idea of the association of VP4 with pVP2 processing (Sanchez and Rodriguez, 1999). The processing of pVP2 into VP2 has been shown to indeed occur at four cleavage sites, ^{441}Ala-Phe^{442}, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502} (Da Costa *et al.*, 2002). Precursor VP2 (pVP2) was shown to be efficiently processed into VP2 in Hi-5 cells in the absence of VP4, but not in Sf9 cells (Lee *et al.*, 2004). It was also suggested that VP4 and other cellular proteases may process pVP2 (Lee *et al.*, 2006b). Irigoyen *et al.* (2009) showed that pVP2 has an endopeptidase activity that cleaves the Ala^{441}-Phe^{442} site and the catalytic mechanism involves $Asp^{431.}$ Processing at the site was abrogated when a D431N mutation was introduced.

The expected sizes of pVP2 and VP2 are 54 kDa and 47 kDa, respectively. However, there have been a number of discrepancies regarding the sizes of pVP2 and VP2 when analysed on reducing SDS-PAGE. For example, several studies identified a 46 to 48 kDa protein to be pVP2 (Dobos *et al.*, 1979; Lombardo *et al.*, 1999; Caston *et al.*, 2001; Chevalier *et al.*, 2005). Studies by Da Costa *et al.* (2002) and Irigoyen *et al.* (2009)

reported pVP2 to be 43 and approximately 70 kDa, respectively. VP2, on the other hand, has been reported to have sizes ranging between 37 and 41 kDa (Dobos *et al.*, 1979; Lombardo *et al.*, 1999; Caston *et al.*, 2001; Da Costa *et al.*, 2002; Chevalier *et al.*, 2005). However, in the studies by Chen *et al.* (2005) and Irigoyen *et al.* (2009), VP2 was observed at 48 and 45 kDa, respectively on reducing SDS-PAGE.

In the present study (Chapter 3), characterisation of the wild-type full-length polyprotein processing with anti-VP4 peptide antibodies revealed certain cleavage patterns. The anti-His tag monoclonal and anti-VP4-RK peptide antibodies recognised a 70 kDa product, which was thought to be an N-terminally Trx-His tagged pVP2. This suggested that pVP2 was not processed, which is not consistent with the findings by Irigoyen *et al*. (2009), unless the processing depends on the expression system. For example, Kibenge *et al.* (1997) expressed the IBDV cDNAs in *vitro* with rabbit reticulocyte lysates in a coupled transcription-translation system and in the Sindbis virus expression system (with BHK-21 and Vero cell cultures) to study processing of the polyprotein, while Irigoyen *et al.* (2009) used the baculovirus-based system. Understanding of the pVP2 processing is further complicated by the discrepancies in the sizes reported in previous studies.

Since the D431N mutation was shown to abrogate pVP2 processing (Irigoyen *et al.*, 2009), the second aim of the present study was to introduce this mutation in the polyprotein construct, express it and determine if there is any change in the processing products. If pVP2 is not processed, then the 70 kDa product which was identified by the anti-His tag monoclonal and anti-VP4-RK peptide antibodies in the processing of the wild-type polyprotein should also be identified in the mutant polyprotein.

Site-directed mutagenesis has become a tool to study the role of individual residues in the structure and function of proteins. There are currently many mutagenesis techniques but PCR-based techniques are widely used due to their efficiency, simplicity and affordability. These include overlap-extension PCR (Ho *et al.*, 1989), the megaprimer method (Kammann *et al.*, 1989) and inverse PCR (Hemsley *et al.*, 1989). The overlap extension method comprises two PCR stages which use mutagenic and flanking primers to produce two fragments with an overlapping sequence. The two fragments are mixed, denatured and allowed to anneal in a PCR-ready buffer to produce a heteroduplex. The amplification by Taq DNA polymerase will only occur in a fragment carrying the 3’ extendable termini (Fig. 4.1A).

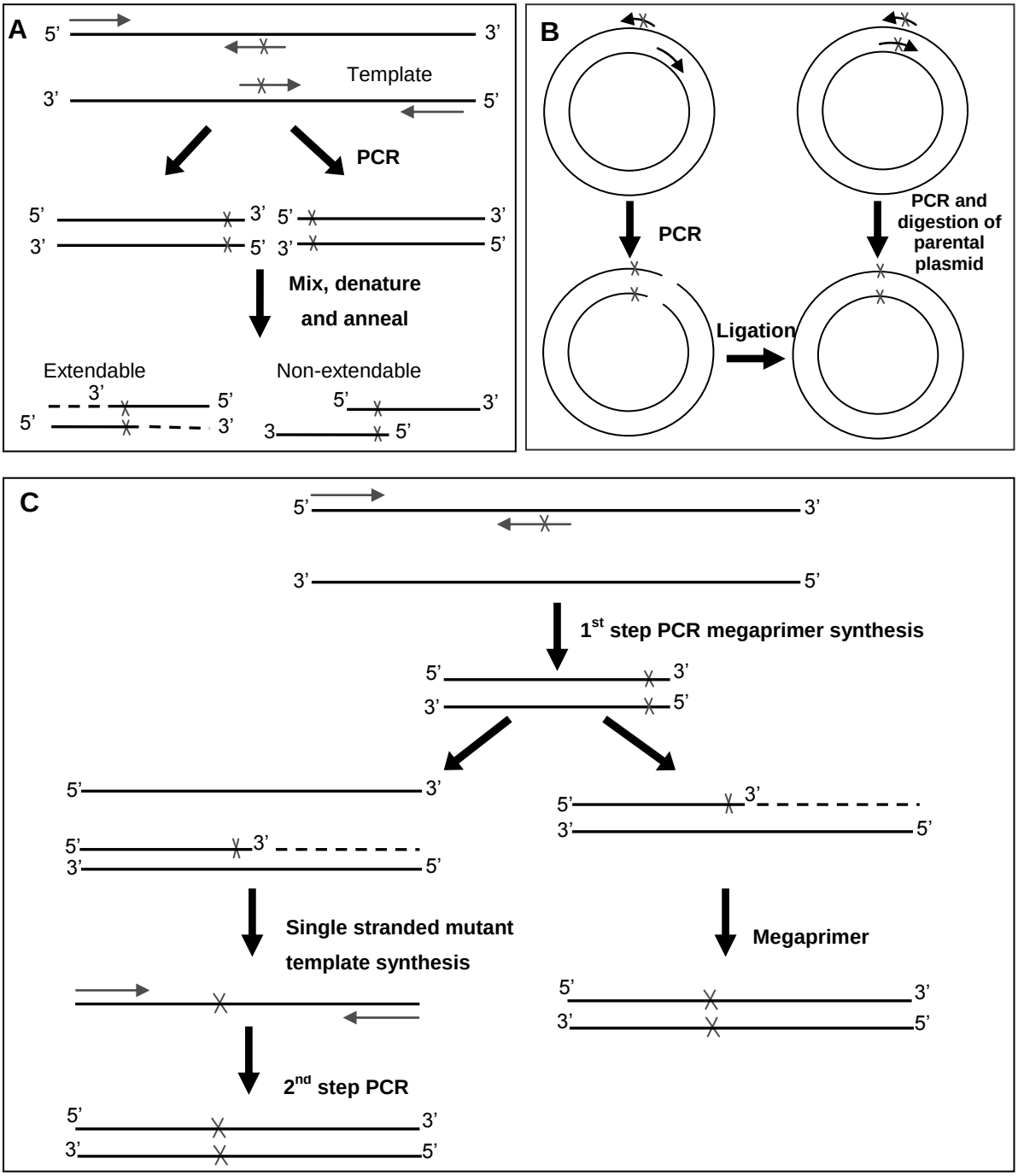


Figure 4.1 PCR based mutagenesis methods. A) Overlap extension PCR. Two PCR products carrying mutated overlapping sequences are mixed, denatured and annealed in a PCR-ready buffer to produce a final mutated product. **B)** Inverse PCR method. The entire plasmid is amplified either by using two complementary primers or a pair of tail-to-tail primers containing the desired mutation. **C)** Megaprimer method. A megaprimer is synthesised using a mutagenic primer and a flanking primer. The megaprimer can either be extended to form a single stranded mutant template that is then used to produce a final product or extended against a template to produce a final product [Adapted from Ling and Robinson (1997)].

The inverse PCR method uses a pair of complementary primers containing the desired mutation at the centre of the sequence (Ho *et al.*, 1989). From this technique, the QuickChange® Site-directed mutagenesis kit was developed (Stratagene, La Jolla, USA), which has been successfully used to introduce mutations in IBDV studies (Lejal *et al.*, 2000; Petit *et al.*, 2000). In other cases, a pair of tail-to-tail primers designed at the site of mutation is used to amplify the entire plasmid containing the gene of interest. The linear DNA is circularised by ligation of the two blunt-ends before transformation. In the case of using complementary primers, a nicked plasmid is produced and there is no need for ligation prior to transformation (Fig. 4.1B). In the megaprimer method, two rounds of PCR are used. The first-stage uses a flanking primer and internal mutagenic primer to synthesise a megaprimer. The mutagenic megaprimer is paired with the second flanking primer to produce the final mutant product (Fig. 4.1C). There have been some modifications to the method over the years. Wu *et al.* (2005b) used a high T_m mutagenic primer to synthesise a single stranded mutagenic DNA, which can be paired with a flanking primer to generate a full mutant PCR product that can be ligated to a vector.

Tseng *et al.* (2008) and Chapnik *et al.* (2008) adapted the megaprimer method where a megaprimer is generated in the first stage PCR from mutagenic and flanking primers. The megaprimer is extended in the second stage to produce a full-length mutagenic plasmid (Fig. 4.2). The method of Tseng *et al.* (2008) was used to introduce the D431N mutation into the pET-32b-full-length polyprotein construct, as it was a single mutation (Fig. 4.2, left hand panel). The same method was modified for the mutation of the catalytic dyad by synthesising the megaprimer using two mutagenic primers, each containing a different mutation (Fig 4.2, right hand panel).

In the present study, two different site-directed mutagenesis studies are described. Firstly, the mutagenesis of the Ser/Lys catalytic dyad in all the constructs that were shown to be processed (Chapter 3), i.e. full-length polyprotein (Met^{1}-Glu^{1012}), truncated polyprotein (Ile^{227}-Trp^{891}), VP4-RA (Arg^{453}-Ala^{755}), VP4-RK (Arg^{453}-Lys^{722}) and VP4-AW (Ala^{513}-Trp^{891}). The second is the introduction of a D431N mutation into the full-length construct. The study therefore describes the site-directed mutagenesis, expression and the results thereof.

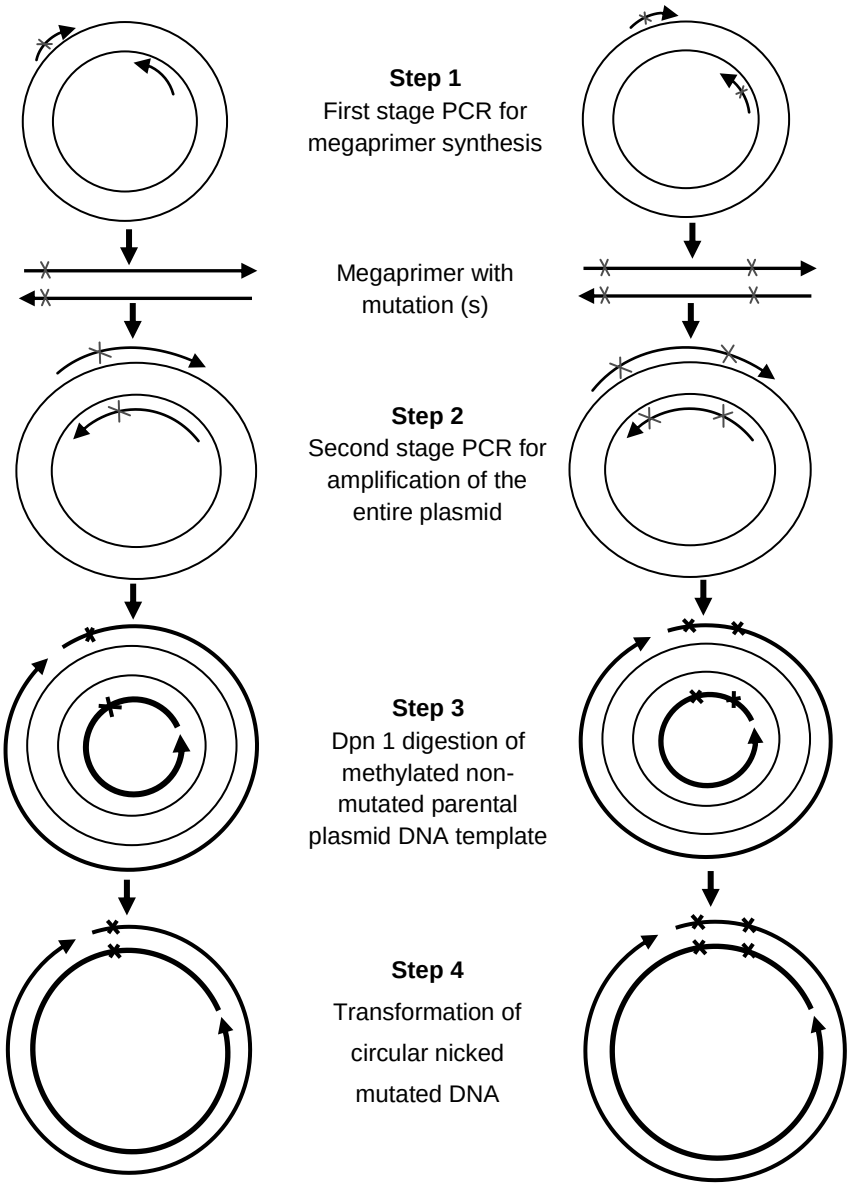


Figure 4.2 Schematic diagram of the megaprimed-inverse PCR used in the present study for introducing a single mutation (Tseng et al., 2008) (left panel) and for introducing double mutations (right panel). Both methods consist of two stages of PCR. The first stage PCR synthesises a megaprimer, which is then used to produce an entire mutant plasmid. The left panel shows the use of one mutagenic primer (—✕—►) and a flanking primer (——►), whereas the right panel shows the use of two mutagenic primers.

4.2 Materials

The following molecular biology reagents were obtained from Fermentas (Vilnius, Lithuania): Dpn1, 10 mM dNTP mix, FastRuler® middle range DNA ladder, GeneRuler 1kb DNA ladder mix, Pfu DNA polymerase, High fidelity enzyme mix, GeneJET™ plasmid miniprep kit and TransformAid™ bacterial transformation kit. The peqGOLD gel extraction kit was obtained from PEQLAB Biotechnologie (Erlangen, Germany) and DNA clean and

concentration kit was obtained from Zymo Research. All primers were synthesised at the MCB Synthetic DNA Unit (University of Cape Town, South Africa). PageRuler™ plus prestained protein Ladder was purchased from Fermentas Life Sciences (Vilinius Lithuania). Anti-His tag monoclonal antibody was purchased from Novagen (Darmstadt Germany). Anti-VP4 peptide antibodies were produced in the present study as described in Chapter 3. Anti-mouse IgG and rabbit anti-chicken IgG conjugated to horseradish peroxidase were obtained from Jackson Immuno-Research Laboratories (USA).

4.3 Methods

4.3.1 Plasmids

Recombinant pET-32b plasmids containing full-length polyprotein, truncated polyprotein, VP4-AW, VP4-RA and VP4-RK coding regions (Chapter 2) were used as templates. A glycerol stock of *E. coli* harbouring pET-32b recombinant plasmids was streaked on 2xYT-Amp plates and incubated overnight at 37°C. Liquid 2xYT-Amp (10 ml) was inoculated with a single colony and the culture was grown overnight at 37°C. All plasmid templates were prepared with the GeneJET^{TM} miniprep kit according to the manufacturer's instructions.

4.3.2 Primers

The VP4 protease catalytic residues (Ser^{652} and Lys^{692}) are 40 amino acid residues away from each other, which means the megaprimer containing the desired mutations would have an ideal size of about 120 bp. Two primers were designed and synthesised: a forward primer containing the S652A mutation (5' CCT ATT GTG GGA AAC **GC**C GGA AAC CTA GCC ATA GC 3') and the reverse primer containing the K692A mutation (5' GTG TGC AGT GGC GAG **CCT** GGT GCT TCT AAA GCT 3') (Fig 4.3). For the D431N mutation, two primers were designed for megaprimer synthesis; forward mutagenic primer (5' GAG GTG GCC **AAC** CTC AAC TCT 3') and the reverse flanking primer containing no mutation (5' GGC AGG TGG GAA CAA TGT 3'). The general characteristics of the primer sequences are summarised in Table 4.1. The S652A forward and K692R reverse primers were 35 and 33 bp long and they had 54 and 58% GC content, respectively. The primers for the D431N mutation were relatively shorter, with 21 and 18 base pairs for forward and reverse primer, respectively. The T_m of the mutagenic primer sequences ranged between 63 and 78^{o}C for the primers containing the mutations.

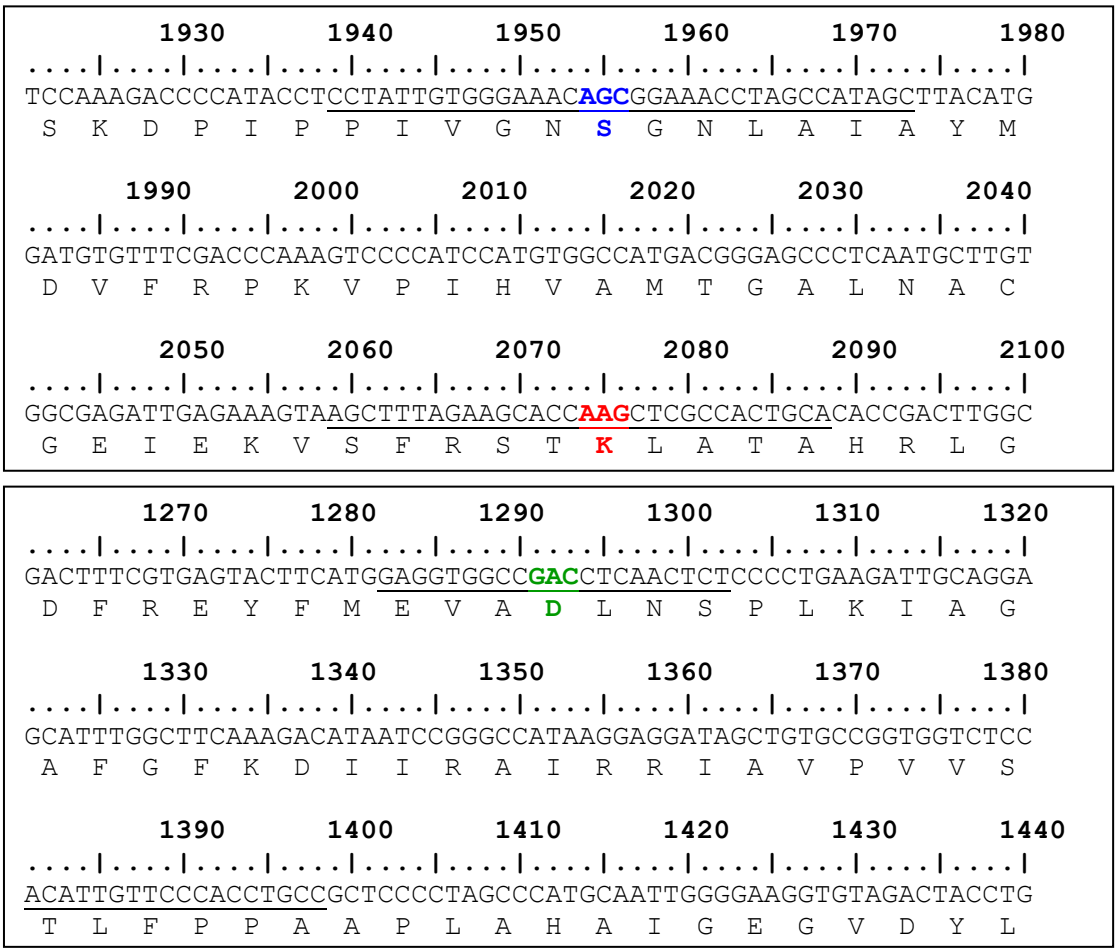


Figure 4.3 Partial nucleotide and amino acid sequence of the polyprotein showing the residues to be mutated. The top panel shows the catalytic residues Ser[652] (blue) and Lys[692] (red) and the bottom panel shows the Asp[431] residue (green). The numbering shown at the top is for the nucleotide sequence. The underlined sequences were used for designing primers for megaprimer synthesis.

Table 4.1 Properties of primers designed for megaprimer synthesis

Properties	VP4 catalytic residues mutation		pVP2 D431N mutation	
	Forward [a]	Reverse [a]	Forward [a]	Reverse
Primer sequence (5' to 3')	CCTATTGTGGGAAAC GCCGGAAACCTAGCC ATAGC	GTGTGCAGTGGC GAGCCTGGTGCTTC TAAAGCT	GAGGTGGCCAACC TCAACTCT	GGCAGGTGG GAACAATGT
Mutation	S652A (AGC →GCC)	K692R (AAG →AGG)	D431N (GAC →AAC)	None
DNA bp	35	33	21	18
GC content (%)	54	58	57	56
T_m (°C)	77.6	77.8	63.2	56.3

[a] The codon carrying the desired mutation is indicated in red

4.3.3 PCR mutagenesis

The mutagenic PCR was carried out in two stages. The first stage megaprimer synthesis was performed with the proofreading *Pfu* DNA polymerase in a 25 µl PCR reaction containing 0.25 µl IBDV segment A cDNA template (~200 ng), 1.25 µl of forward and reverse primer (0.5 µM each), 2 µl dNTP mixture (0.8 mM), 2.5 µl 10x *Pfu* PCR buffer with $MgSO_4$, 0.25 µl *Pfu* DNA polymerase (1.25 U) and 17.5 µl sterile dH_2O. The reagents were mixed by vortexing gently and centrifuging briefly to collect the sample. The megaprimer was synthesised in a thermal cycler with the following temperature profile: initial denaturing cycle of 3 min at 95°C followed by 30 cycles of 30 s at 94°C, 1 min at 55°C, 1 min at 72°C and a final extension step for 7 min at 72°C. The megaprimer was purified using the peqGOLD gel extraction kit and used for the amplification of the entire plasmid in the second stage PCR.

The second stage PCR was performed using the high fidelity PCR enzyme mix in a 50 µl volume containing 5 µl 10x PCR buffer, 4 µl $MgCl_2$ (2 mM), 1 µl megaprimer, 1 µl plasmid DNA, 2 µl dNTP mixture (0.4 mM), 0.5 µl High fidelity PCR enzyme mix (2.5 U) and 36.5 µl sterile dH2O. After mixing the reagents by vortexing gently and centrifuging briefly, the tubes were placed in a thermal cycler programmed with the following temperature profile: initial denaturation for 3 min at 94°C, 10 cycles of 30 s at 94°C, 1 min at 60°C, 12 min at 68°C; 15 cycles of 30 s at 94°C, 1 min at 60°C, 12 min + 10 s/cycle at 68°C and final extension for 10 min at 68°C.

The PCR products were purified using the DNA clean up and concentrator kit and resuspended in 15 µl. The purified DNA (8 µl) was subjected to Dpn1 digestion in a 10 µl reaction at 37°C for 1 h to degrade the original methylated non-mutated plasmid templates. The Dpn1 digested DNA was purified using the DNA clean up and concentrator kit and used to transform *E. coli* BL21 (DE3) competent cells using TransformAid™ bacterial transformation kit. The transformed cells were grown overnight on 2xYT-Amp agar plates and incubated overnight at 37°C. Colonies were selected and grown overnight in 10 ml 2xYT-Amp broth at 37°C. Plasmid DNA was isolated using the GeneJET™ plasmid miniprep kit and sequenced using pET primers and insert primers by Inqaba Biotech (Pretoria, South Africa) as shown before in Chapter 2 (Section 2.3.6, Fig. 2.3).

4.3.4 Expression of mutated constructs

A single colony from a 2xYT-Amp plate was grown overnight in Terrific Broth containing 100 µg/ml of ampicillin in a shaking incubator at 37°C. To verify expression, cells from 1 ml of the overnight culture were collected by centrifugation and disrupted directly in reducing treatment buffer and analysed by SDS-PAGE followed by western blotting (Section 3.3.3). For large scale harvesting, the cells were centrifuged (5000 x g, 10 min, 4°C) and to disrupt the cells, the cell pellet was resuspended in 1/20 culture volume of lysis buffer [20 mM Tris-HCl buffer pH 7.6, 100 mM NaCl, 5% (v/v) glycerol]. Lysozyme was added to a final concentration of 1 mg/ml and the cells were incubated on ice for 30 min. The cell lysate was frozen overnight at -20°C. After thawing, the cell lysate was disrupted by sonication followed by centrifugation (5000 x g, 4°C, 30 min). Insoluble protein pellet was washed 3 times with pellet wash buffer [20 mM Tris-HCl buffer pH 8, 2.5% (v/v) Triton X-100, 2 M urea]. Each washing step was followed by centrifugation (14 000 x g, 30 min, 4ºC). The inclusion bodies were solubilised as described in Section 3.3.4.

4.3.5 Molecular modeling

The attempted mutation of D431N resulted in an unexpected mutation of a Pro to Leu at position 377. To determine the structural effect of the mutation, a molecular model was generated using the crystal structure coordinates of VP2 obtained from the Protein Data Bank (PDB Accession Code 3BFM, http://www.rcsb.org/pdb/). To construct a model structure for the P377L variant of the protein, the mutation was introduced using the modeling program Swiss-PDB viewer version 3.7 (Guex and Peitsch, 1997). The conformational constraints were relaxed by energy minimisation. For comparison of structures, the VP2-Leu^{377} model was superimposed on the template homolog of VP2 using PYMOL version 1.3.

4.4 Results

4.4.1 Site-directed mutagenesis PCR

A combination of the megaprimer method and inverse PCR was used to introduce double mutations in the IBDV polyprotein (full-length and truncated forms) and VP4 encoding genes (VP4-AW, VP4-RA and VP4-RK) previously cloned into the pET-32b plasmid. The first stage PCR was performed by *Pfu* DNA polymerase using the forward and reverse mutagenic primers, carrying S652A and K692R mutations. Following the first stage PCR,

the products were analysed on a 1.5% agarose gel (Fig. 4.4). The PCR reaction produced a 120 bp DNA fragment for the Ser/Lys catalytic dyad mutation (Panel A, lane 1).

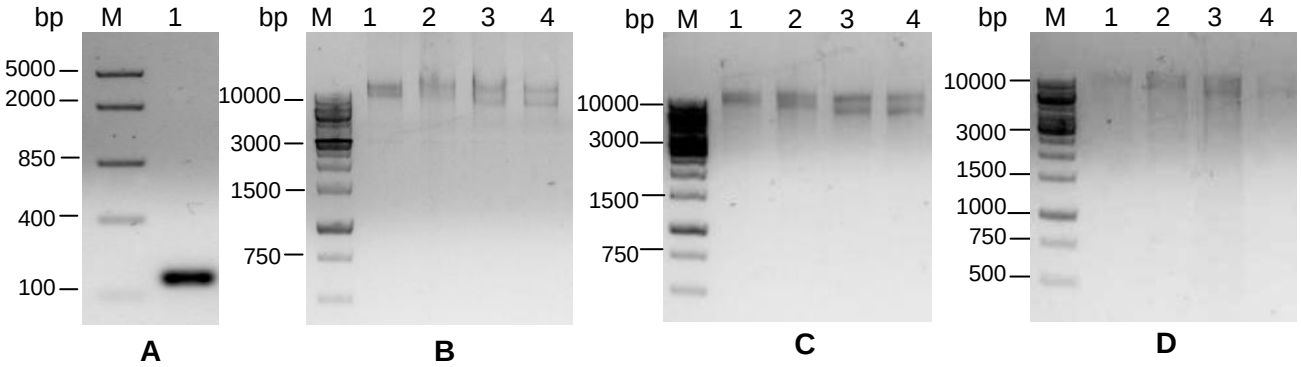


Figure 4.4 The first and second stage PCR for Ser/Lys catalytic dyad mutagenesis. **A)** First stage PCR for megaprimer synthesis. Lane M, FastRuler middle range ladder. Lane 1, 150 bp megaprimer. Megaprimer was used in the second stage PCR to produce the full-length plasmid. **B)** PCR products generated using megaprimers. **C)** Purified PCR products. **D)** Purified products after Dpn 1 digestion. Lane M, 1 kb GeneRuler DNA ladder; lane 1, pET-32b-Full-length ORFA2; lane 2, pET-32b truncated ORFA2; lane 3, pET-32b VP4-RA; lane 4, pET-32b VP4-RK.

The megaprimer was gel-purified and used as a primer for the second stage PCR which was designed to amplify the entire plasmid. After PCR amplification using mutagenic megaprimers, two prominent bands representing the circular nicked and linear forms of the amplified plasmid were observed at the corresponding sizes of the plasmids i.e. 9 kb pET-32b-full-length polyprotein coding ORFA2 (lane 1), 8 kb pET-32b-truncated ORF-A2 (lane 2), 7 kb pET-32b-VP4-RA (lane 3) and pET-32b-VP4-RK (lane 4), (Fig. 4.4B). The PCR products were purified prior to Dpn1 digestion (Fig. 4.4C). The agarose gel analysis of the final PCR products after Dpn1 digestion is shown in Fig. 4.4D. The mutagenic PCR for the amplification of pET-32b VP4-AW using mutagenic primers also yielded two bands at the expected size of approximately 7 kb (Fig. 4.5). The PCR products were also purified and subjected to Dpn1 digestion prior to transformation.

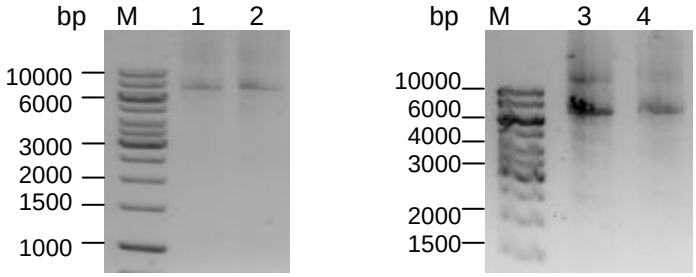


Figure 4.5 Agarose gel electrophoresis of the second stage PCR products for pET-32b VP4-AW at different stages prior to transformation. Lane M, 1 kb GeneRuler DNA ladder; lanes 1-2, PCR products generated using megaprimers; lane 3, purified PCR products; lane 4, purified products after Dpn1 digestion.

The megaprimer containing the D431N mutation was observed at the expected size of about 120 bp (Fig. 4.6A lane 1). The second stage PCR was also successful as shown by the two bands at the expected size for the pET-32b-full-length ORFA2 of 9 kb (Fig. 4.6B). The PCR products were successfully purified after the PCR (Fig. 4.6C) and after Dpn1 digestion (Fig. 4.6D).

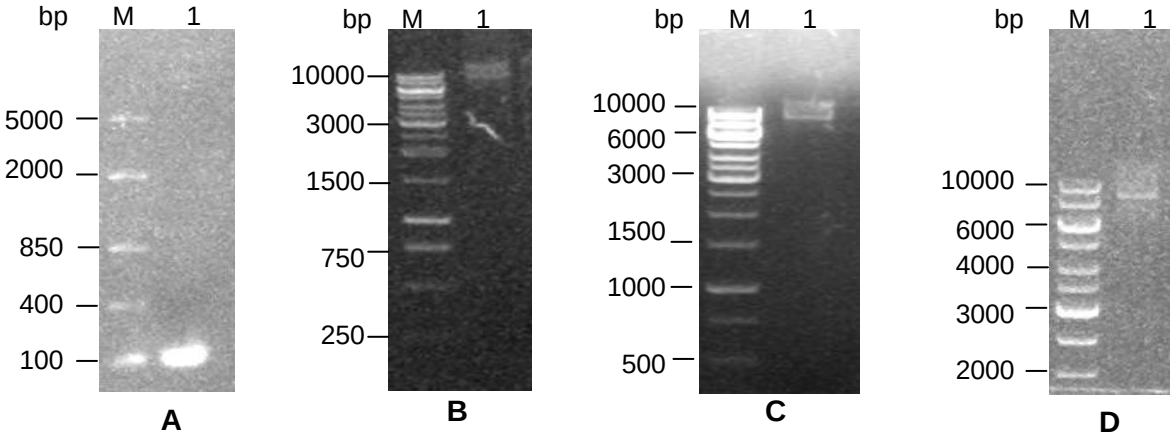


Figure 4.6 The first and second stage PCR for D431N site-directed mutagenesis. A) First stage PCR for megaprimer synthesis. Lane M, FastRuler middle range ladder. Lanes 1, megaprimer. The megaprimer was used in the second stage PCR to produce the full-length plasmid containing the mutation. **B)** PCR products generated using megaprimers. **C)** Purified PCR products. **D)** Purified products after Dpn 1 digestion. Lane M, 1 kb GeneRuler DNA ladder; lane 1, mutated pET-32b-full-length ORFA2 plasmid.

4.4.2 Expression of mutants

The recombinant pET-32b plasmids carrying the mutations of the VP4 catalytic residues were transformed into *E. coli* BL21 (DE3) cells. The effect of the mutations on the autocatalytic activity of the embedded VP4 was determined by expression and analysis by SDS-PAGE and western blotting using anti-VP4 peptide antibodies. As described in the preceding chapter, full-length polyprotein, truncated polyprotein and different forms of VP4 (VP4-AW, VP4-RA and VP4-RK) showed autocatalytic activity during expression in *E. coli*. In contrast to the wild-type constructs which expressed as cleavage products, all the mutant constructs expressed as fusion proteins (Fig. 4.7A). The expression of the mutant full-length polyprotein yielded an uncleaved protein with a molecular weight of approximately 130 kDa, corresponding to the size of the polyprotein of 114 kDa in fusion with the 17 kDa Trx-His tag (lane 2). High molecular weight proteins which were observed when the wild-type polyprotein was expressed were also observed in the expression of the mutant polyprotein. Similarly, other constructs were expressed at their expected molecular sizes i.e. truncated polyprotein at 90 kDa (lane 3), VP4-AW at 60 kDa (lane 4), VP4-RA at 52 kDa (lane 5) and VP4-RK at 49 kDa (lane 6).

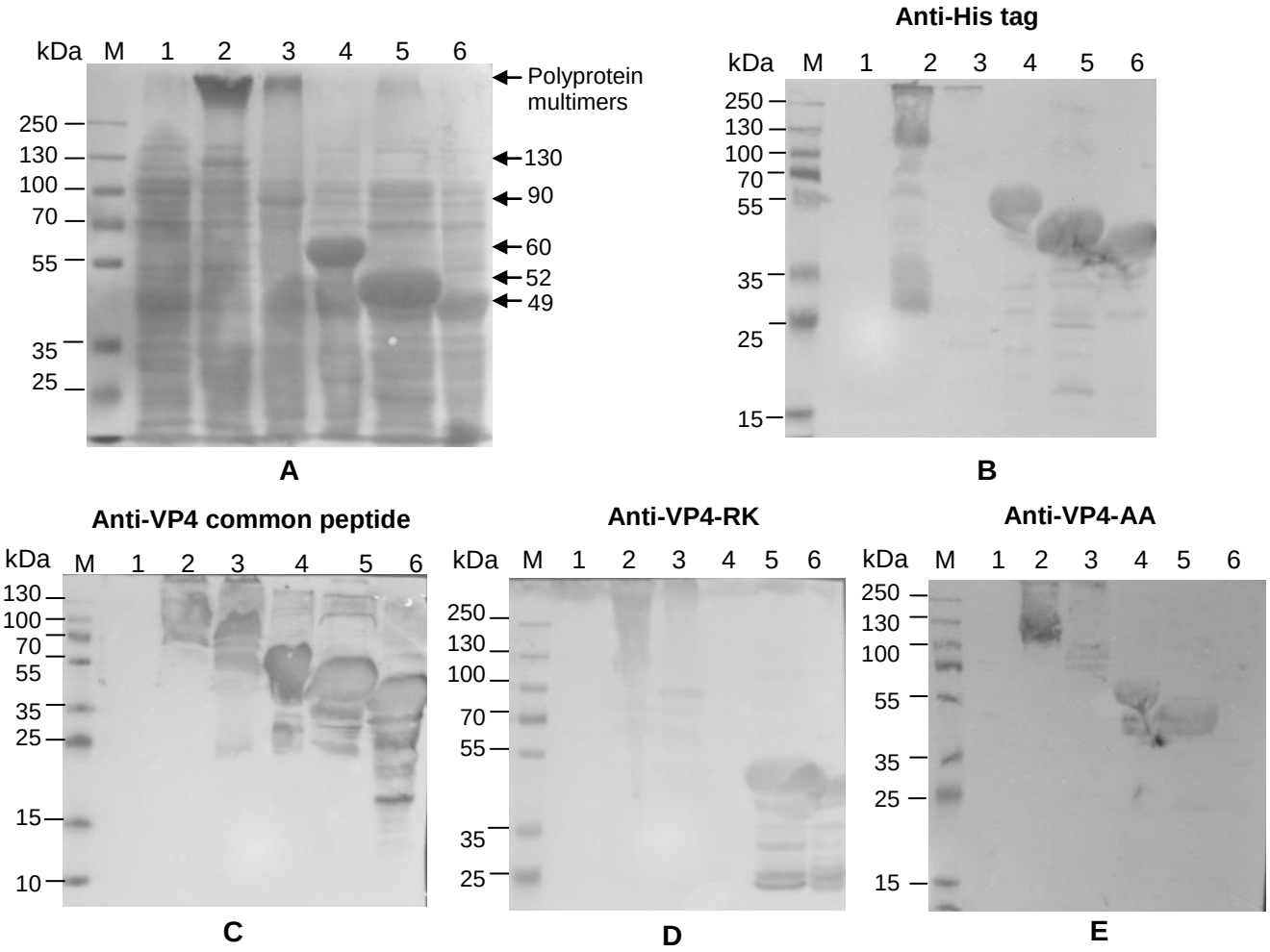


Figure 4.7 Analysis of the expression of polyprotein and VP4 constructs carrying the S652A and K692R mutations in *E. coli* by reducing SDS-PAGE (A) and western blot using anti-His tag monoclonal antibodies (B), anti-VP4 common peptide antibodies (C), anti-VP4-RK peptide antibodies (D) and anti-VP4-AA peptide antibodies (E). Lane M, PageRuler™ plus prestained protein ladder; lane 1, *E. coli* BL21 (DE3) cell control; lane 2, mutant full-length polyprotein; lane 3, mutant truncated polyprotein; lane 4, mutant VP4-AW; lane 5, mutant VP4-RA; lane 6, mutant VP4-RK. The arrows show the major mutant protein bands observed in the SDS-PAGE gel.

The expression of the mutants was confirmed by western blotting using anti-His tag monoclonal antibodies (Fig. 4.7B) and chicken anti-peptide antibodies (Fig. 4.7C to E). The mutant full-length and truncated polyprotein were recognised by anti-His tag monoclonal antibodies (Panel B) and chicken anti-VP4 peptide antibodies (Panel C-E). All the antibodies also recognised the high molecular weight proteins (Panel B-D, lanes 2 and 3). Mutant VP4-AW (lane 4) was only recognised by anti-VP4 common peptide (Panel C) and anti-VP4-AA peptide (Panel E) antibodies as expected and not by anti-VP4-RK peptide antibodies (Panel D) since the peptide is situated N-terminal to the pVP2-VP4 cleavage junction (Fig. 3.11B). Similarly, mutant VP4-RK (lane 6) was recognised by anti-VP4 common peptide (Panel C) and anti-VP4-RK peptide antibodies (Panel D), but not by anti-VP4-AA peptide antibodies (Panel E). All three anti-peptide antibodies were able to

recognise mutant VP4-RA (lane 5) since it contains all three peptides. Other bands of lower molecular weight were observed for the anti-VP4 common and VP4-RK peptide antibodies, however this was due to the high reactivity of these antibodies in comparison to the anti-VP4-AA peptide antibodies.

The solubility of the polyprotein mutants was also determined. The polyprotein mutants were expressed as insoluble inclusion bodies (Fig. 4.8A and B). The VP4 derived mutant constructs were insoluble, however, some of the protein was released in the 2 M urea washes. Results are not shown for VP4-AW and VP4-RA as the solubility profile was similar to that of VP4-RK shown in Fig. 4.8C.

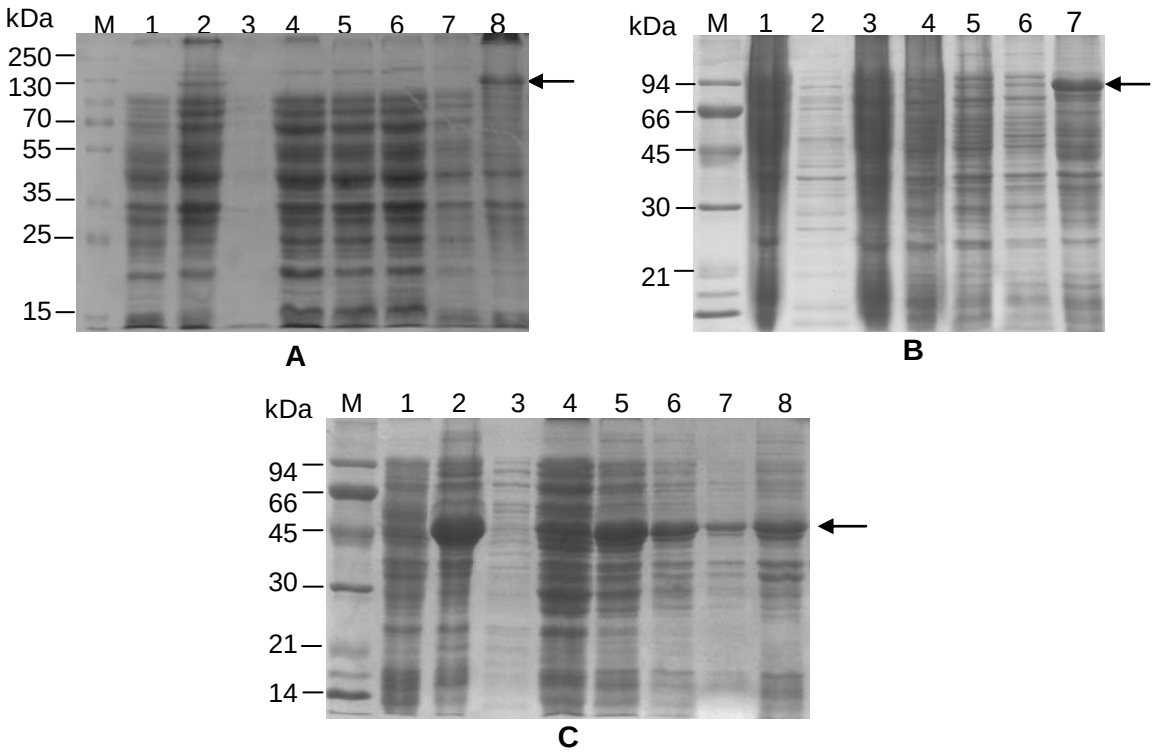


Figure 4.8 SDS-PAGE analysis of the solubilisation profile of the recombinantly expressed polyprotein and VP4 mutants. A) Mutant full-length polyprotein. Lane M, PageRuler™ plus prestained protein ladder; lane 1, *E. coli* BL21 (DE3) cell control; lane 2, total lysate; lane 3, soluble fraction; lanes 4-7, pellet washes; lane 8, solubilised inclusion bodies. **B)** Mutant truncated polyprotein. Lane M, Low molecular weight marker; lane 1, total lysate; lane 2, soluble fraction; lanes 3-6, pellet washes; lane 7, solubilised inclusion bodies. **C)** Mutant VP4-RK. Lane M, Low range MWM; lane 1, *E. coli* BL21 (DE3) cell control; lane 2, total lysate; lane 3, soluble fraction; lanes 4-7, pellet washes; lane 8, solubilised inclusion bodies.

The analysis of the transformants for the D431N site-directed mutagenesis induced mutation was carried out by sequencing. Sequencing revealed that in addition to the Asp^{431} not being mutated, neither the polyprotein cleavage sites nor the catalytic residues were mutated. Instead, a mutation of a Pro to Leu at position 377 in the VP2 region was observed (Fig. 4.9) and was the only mutation that distinguished the wild type and the

mutant transformants. The Pro^{377} was conserved among all IBDV strains of serotype I and II.

```
Mutant       GSVVTVAGVS NFELIPNLEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
SA-KZN95     GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
UPM94/273    GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
UK661        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
SH/92        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
SDH1         GSVVTVAGVS NFGLIPNPEL AKNLVTEYGR FDPGAMNYTK LILGERDRLG IKTVWPTREY  419
TASIK        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
D6948        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
52/70        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
STC          GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
002-73       GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
GLS          GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
Variant E    GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
GZ29112      GSVVTLAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSDGDRLG IKTVWPTREY  419
CEF94        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
D78          GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  419
23/82        GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  420
OH           GSVVTVAGVS NFELIPNPEL AKNLVTEYGR FDPGAMNYTK LILSERDRLG IKTVWPTREY  420
Consensus    *****:**** ** **** ** ********** ********** ***.: **** **********

Mutant       TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
SA-KZN95     TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
UPM94/273    TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
UK661        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
SH/92        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
SDH1         TDFREYFMEV VDLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
TASIK        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
D6948        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR ALRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
52/70        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
STC          TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
002-73       TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
GLS          TDFREYFMEV ADLSSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
Variant E    TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPV AHAIGEGVDY  479
GZ29112      TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
CEF94        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
D78          TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRRIAVPVV STLFPPAAPL AHAIGEGVDY  479
23/82        TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRKIAVPVV STLFPPAAPL AHAIGEGVDY  480
OH           TDFREYFMEV ADLNSPLKIA GAFGFKDIIR AIRKIAVPVV STLFPPAAPL AHANREGVDY  480
Consensus    ********** .**.****** ********** *:*:****** *********: ***  *****
```

Figure 4.9 Sequence alignment of the amino acid residues of the mutant and other IBDV serotype I and II strains. Only residues 360 to 480 are shown in the alignment. The Asp^{431} is shown in red whereas the P^{377} is shaded in grey.

To analyse the effect of the P377L site-directed mutagenesis induced mutation on the polyprotein processing during expression in *E. coli,* the cell lysate was analysed by SDS-PAGE and western blot using chicken anti-VP4 peptide antibodies (Fig. 4.10). A notable difference was observed in both SDS-PAGE and western blot between the cleavage products produced by the wild-type and the mutant polyprotein. In the SDS-PAGE analysis, a cleavage product at 28 kDa was observed in the wild-type lysate, in comparison to the 50 kDa band in the mutant lysate (Fig. 4.10A). The use of chicken anti-VP4 peptide antibodies revealed the other bands that were not visible on the SDS-PAGE gel.

The anti-VP4 common peptide antibodies recognised the high molecular weight proteins and the cleavage product at 28 kDa which corresponds to VP4-AA, in the cells transformed with the wild-type construct (Fig 4.10B, lanes 2-3). The other two products of 60 and 45 kDa which were recognised by anti-VP4 common peptide antibodies as shown in the preceding chapter (Fig. 3.7C, lane 4) were observed but not as prominent as the 28 kDa on the western blot. Proteins of molecular sizes 250, 100 and 50 kDa were recognised by the anti-VP4 common peptide antibodies in the mutant polyprotein expressing cell lysate (Fig. 4.10B, lane 4).

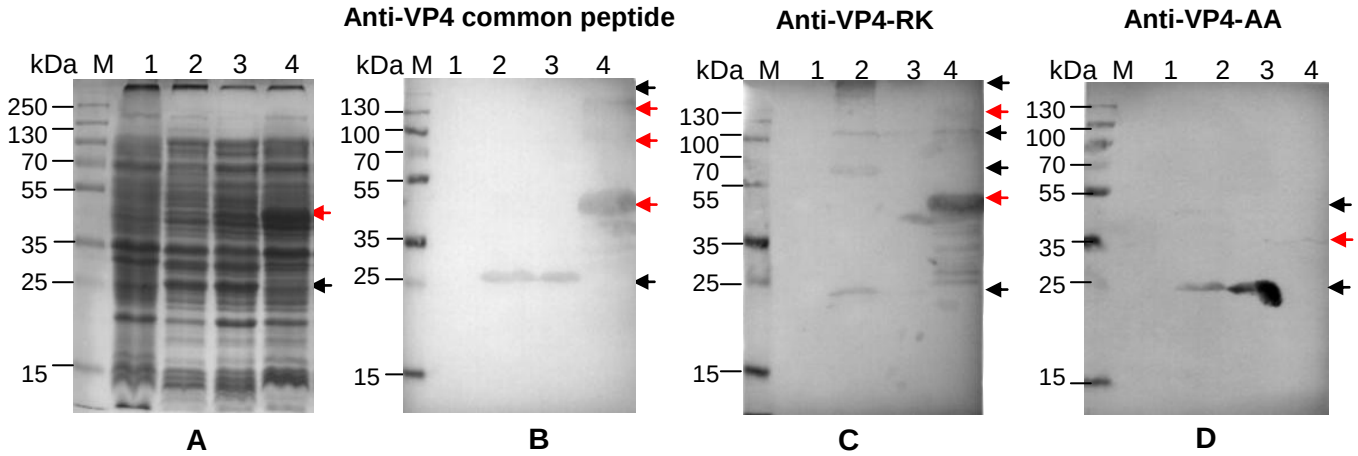


Figure 4.10 Analysis of the expression of pET-32b-full-length polyprotein carrying the P377L mutation in *E. coli* by reducing SDS-PAGE (A) and western blot using anti-VP4 common peptide antibodies (B), anti-VP4-RK peptide antibodies (C) and anti-VP4-AA peptide antibodies (D). Lane M, PageRuler™ Plus Prestained Protein Ladder; lane 1, *E. coli* BL21 (DE3) cell control; lanes 2 -3, wild-type construct; lane 4, mutant construct. (The black arrows indicate the cleavage products from the wild-type constructs whereas the red arrows the cleavage products from the mutant form).

The anti-VP4-RK peptide antibodies recognised the high molecular weight proteins, thought to be the assembled polyproteins and two cleavage products at 24 and 70 kDa in the wild type construct. Another band at about 100 kDa was observed, which seemed to result from non-specific antigen-antibody interaction (Fig. 4.10C, lane 2). The proteins were not as strongly recognised in the second wild type construct (lane 3). The same antibodies recognised proteins of molecular size 250 kDa and a major cleavage product at 50 kDa in the mutant construct. In addition to these proteins, several prominent bands were also recognised between 25 and 45 kDa in the mutant construct (lane 4). The anti-VP4-AA peptide antibodies strongly recognised the 28 kDa protein in the wild-type polyprotein construct, but very weakly bound the products from the mutant polyprotein, therefore the results were not taken into account for the conclusions made. It was clear, however that the mutation had an effect on the polyprotein processing in *E. coli*.

Given that Pro^{377} was conserved among all IBDV strains, and that it was the only mutation that distinguished the wild type and this mutant, it was concluded that the P377L was responsible for the observed differences in the expression profile. This Pro was within a Pro-Asn-Pro tripeptide, and there has been no special mention of these residues in any of the reported crystal analysis of VP2. In light of these findings, the next step was to determine the effect of the mutation on the protein structure. To determine whether the P377L mutation would have any effects on the structural fold, a model of the VP2 mutated version was built using the 3-D structure of VP2 as template (Irigoyen *et al.*, 2009) (Fig. 4.11).

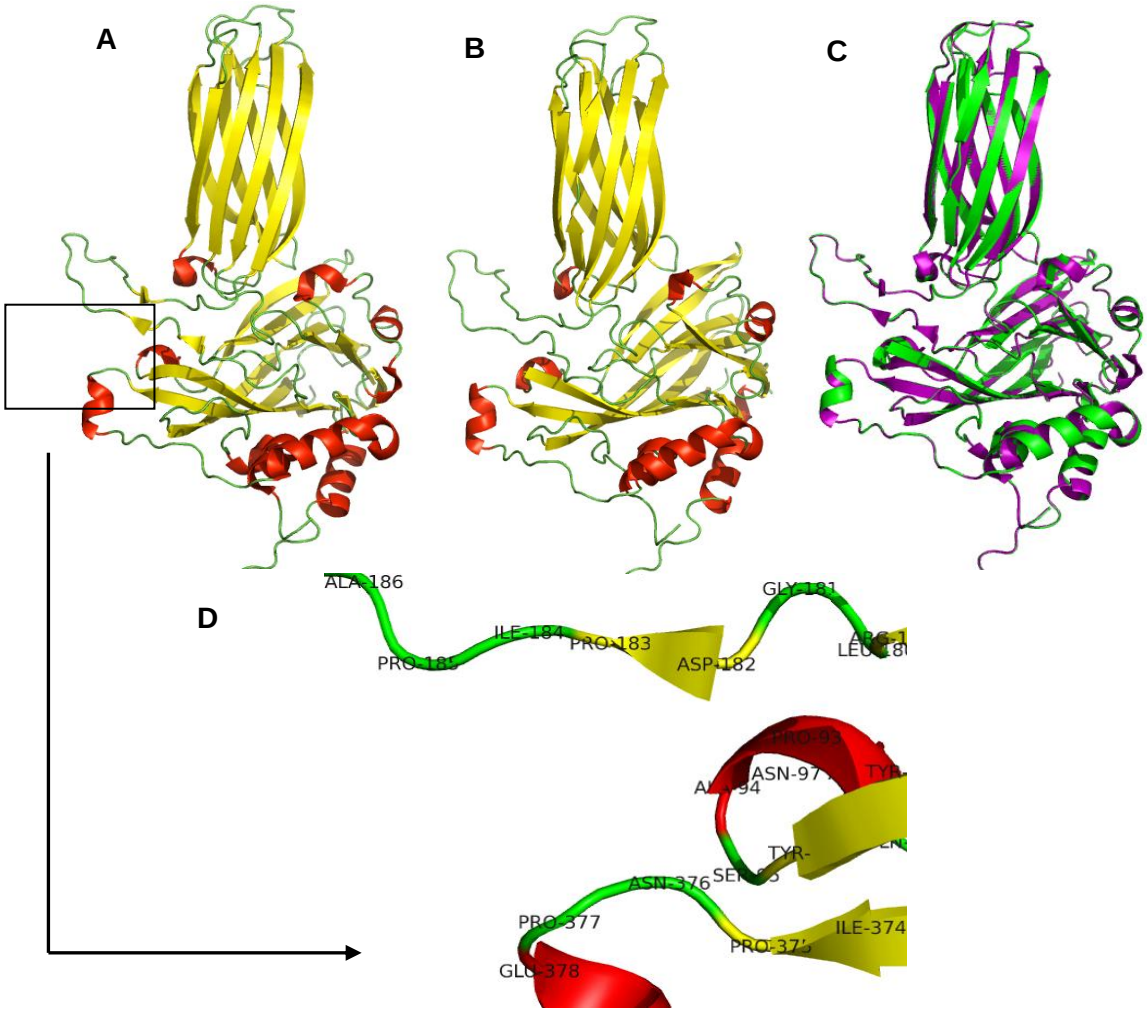


Figure 4.11 *In silico* modelling of VP2 3-D structures containing the P377L mutation. A) The crystal structure of VP2 obtained from the protein data bank (PDB code 3BFM). **B)** Mutant VP2 structure containing the P377L mutation. **C)** Superimposed structures of the wild-type and the mutant VP2. **D)** A closer view of the ribbon structure arrangement of Pro^{377} with other Pro residues. Loops are shown in green, while red and yellow represent α-helices and β-sheets, respectively. In the superimposed structure, the wild type P^{377} containing VP2 is shown in purple and the mutant is colored in green.

In the VP2 structure, Pro^{377} is present in a loop before an α-helix on a surface cavity in the B domain. Five Pro residues were found within this surface cavity, Pro^{185}, Pro^{183}, Pro^{93}, Pro^{375} and Pro^{377}. Even though the overall position of the domains remained relatively unchanged, some alterations were observed in the structure when the P377L mutation was introduced. While Pro^{377} was present in a loop, Leu^{377} was present in an α-helix. The model suggested that the mutation also affected the P domain, where one of the ten β-sheets was disrupted by an introduction of a loop. Major differences were also observed in the arrangement of loops, β-sheets and α-helices of the S domain where Pro^{377} is located. Taken together, these results suggest that the Pro^{377} residue plays an important role in the structure and function of the polyprotein.

4.5 Discussion

Previous studies on IBDV polyprotein processing have thus far focussed on identifying the catalytic residues (Lejal *et al.*, 2000) and cleavage sites (Sanchez and Rodriguez, 1999). From these studies, we speculated that VP4 exists in two forms, the embedded and mature forms. IBDV VP4 protease was shown to be active in *E. coli* where the expression of the full-length polyprotein yielded cleavage products pVP2, VP3 and VP4 (Hudson *et al.*, 1986; Lejal *et al.*, 2000). The fundamental question is whether the mature form retains its autocatalytic activity. Results presented in the previous chapter showed that the designed constructs corresponding to the embedded form were autocatalytic whereas the mature form was not. All the constructs which showed autocatalytic activity contained the Arg^{453}-Ala^{512} sequence. It was therefore deduced that the VP4 protease may use this polypeptide as a cofactor for its autocatalytic activity, as in the case of NS3 protease.

NS3 protease requires the presence of polypeptide cofactors for efficient proteolytic processing (Murthy *et al.*, 1999). The order of the non-structural proteins in the flavivirus polyprotein is P7-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 (Fig. 1.4). Flavivirus NS3 protease requires NS2B protein for activity, where NS2B functions as a cofactor and promotes the folding and functional activity of NS3 (Tsumoto *et al.*, 2002). The central hydrophilic region of NS2B (40 amino acids) is connected to NS3 via a flexible glycine linker (GGGGSGGGG) (Chappell *et al.*, 2007). A crystal structure of the NS2B40-G_4SG_4-NS3 shows that NS2B40 wraps around NS3pro in a belt-like manner (Erbel *et al.*, 2006). It has been suggested that NS2B may function as a prodomain needed only for the proper folding of NS3 and not necessarily for the proteolytic activity (Shiryaev *et al.*, 2006).

In the present study, all the constructs which were autocatalytically processed contained the Arg^{453}-Ala^{512} polypeptide situated at the N-terminus of VP4. The constructs lacking the

polypeptide i.e. VP4-AW (Ala^{513}-Trp^{891}) and VP4-AA (Ala^{513}-Ala^{755}) did not display the autocatalytic activity. This suggested that the polypeptide may serve a similar function to that of NS2B. Evidence was required to support that the processing observed in the present study was mediated by VP4, before making such deductions.

In addition, all the processed constructs had the vicinal Ala-Ala dipeptides which have been identified as the cleavage sites for VP4. It has been shown that birnavirus VP4 is similar to bacterial proteases (Paetzel and Dalbey, 1997), therefore it could not be ruled out that the expressed proteins could be cleaved by these bacterial proteases. It was previously shown that the replacement of Ser^{652} (by alanine or threonine) and Lys^{692} (by alanine, arginine or histidine) completely inactivated polyprotein processing (Lejal *et al.*, 2000). For these reasons, S652A and K692R mutations were introduced into the polyprotein and VP4 constructs, in order to determine if the observed processing was associated with VP4.

A PCR-based site-directed mutagenesis method was used to mutate the catalytic residues. Although the QuickChange site-directed mutagenesis method was successfully used to introduce mutations in previous IBDV mutagenesis studies (Lejal *et al.*, 2000; Petit *et al.*, 2000), it was initially attempted in the present study to introduce the mutations into the constructs but was unsuccessful. The failure of the method was attributed to several factors. Firstly, the method uses two complementary primers carrying the desired mutations to amplify an entire recombinant plasmid, which results in a high degree of primer dimer formation. Additionally, the sizes of plasmids to be amplified ranged from 7 kb to 9 kb, which are quite long. Such long amplicons can often be difficult to amplify with a single polymerase.

Secondly, the method can only introduce a single mutation per reaction, and in the present study the mutation of both of the catalytic residues was desired. This means that it would require a first stage PCR using a set of complementary primers containing the S652A mutation, transformation, plasmid isolation and sequencing to confirm the presence of the mutation. To obtain the second mutation, the mutated plasmid would then be used as a template in the second stage PCR using another set of primers containing the K692A mutation. The second stage PCR product would then be transformed into competent cells, followed by plasmid isolation and sequencing to ensure the introduction of the second mutation.

To solve these problems, a combination of the QuickChange site-directed mutagenesis and megaprimer methods was used. The forward and reverse primers were designed,

carrying the S652A and K692R mutation, respectively. These primers resulted in amplification of a 120 bp product carrying the desired mutations. The product was then used as a megaprimer to prime the amplification of the entire plasmid for each of the constructs used. For the amplification of the plasmids, the high fidelity PCR enzyme mix was used, a mix of *Taq* DNA polymerase and a thermostable DNA polymerase with proofreading activity which is able to produce longer PCR products with greater yield and fidelity than a single DNA polymerase (Fermentas Life Science).

The amplified products were subjected to Dpn1 digestion to degrade the original methylated non-mutated plasmid templates. The amplicons were purified prior to Dpn1 digestion since the digestion failed when the DNA was not purified from the PCR reaction mix. Dpn1 digested DNA was also purified before transformation, as it was observed that the transformation of the Dpn1 digestion reaction mix resulted in a very low transformation efficiency. The transformation of purified DNA after Dpn1 digestion yielded transformants and several colonies were randomly selected for sequencing to confirm the presence of mutations. This was a cost-effective and simple approach for introducing double mutations.

The constructs representing the fully-embedded protease (full-length and truncated polyprotein) and the N- and C-terminally embedded form of VP4 (VP4-RA and VP4-RK) (Fig. 3.1) contain the Arg^{453}-Ala^{512} polypeptide. The wild-type forms of these constructs were autocatalytically cleaved into products during expression in *E. coli*. The wild-type form of VP4-AW was also processed in *trans* during expression. After mutating the catalytic dyad, the mutant forms were expressed and analysed by SDS-PAGE and western blotting using the anti-VP4 peptide antibodies. All mutant constructs were expressed as fusion proteins without any cleavage products. These results indicate that the autocatalytic activity observed in the wild-type constructs was due to VP4. This also rules out the possibility of bacterial proteases being responsible for the observed cleavage in the wild-type constructs. From this study, it could be cautiously deduced that the autocatalytic activity of VP4 protease requires the Arg^{453}-Ala^{512} sequence.

The second aim was to introduce the D431N mutation to verify whether pVP2 is processed in *E. coli*. Surprisingly, the method failed to introduce the desired mutation but instead a P377L mutation was introduced into the construct. There is no explanation as to why the site-directed mutagenesis of the D431N mutation would fail. The P377L mutation could have been randomly introduced by the polymerase. Even though a high fidelity

polymerase was used for the mutagenesis, unintended mutations can be introduced especially when amplifying a 9 kb long plasmid (Volkov *et al.*, 1999).

The substitution of the Pro^{377} residue with Leu was shown to alter polyprotein processing. The conservation of this Pro residue in all the IBDV strains strongly suggests a functional role. Pro is a unique amino acid in that it lacks an amide proton. It is usually conserved in proteins and thus it plays an important role in protein structure and function in many organisms including viruses (Bajaj *et al.*, 2007). For example, two Pro residues at positions 7 and 13 were shown to alter Gag-Pol processing and abolished viral infectivity in HIV (Hill *et al.*, 2002; Hill *et al.*, 2007). A proline residue was also shown to influence replication and virus assembly in HCV (Hughes *et al.*, 2009). The functional role of Pro^{377} was shown in the processing of the IBDV polyprotein during expression in *E. coli.* A notable difference was observed between the cleavage products released from the wild-type and the mutant VP4 constructs. It was clear that the processing of the polyprotein was altered by the mutation, but it was not clear how it was altered. The IBDV polyprotein (Met^{1}–Glu^{1012}) is processed to pVP2 (Met^{1}-Ala^{512}, 54 kDa), VP4 (Ala^{513}-Ala^{755}, 28 kDa) and VP3 (Ala^{756}-Ala^{1012}, 32 kDa) (Sanchez and Rodriguez, 1999).

Theoretically, the processing of the wild type pET-32b-full-length polyprotein should produce an N-terminally Trx-His tagged pVP2 (approximately 70 kDa), VP4 (28 kDa) and a C-terminally 6xHis tagged VP3 (32 kDa) (Fig. 4.12). The expression of the P377L containing mutant pET-32b-full-length polyprotein plasmid produced a main cleavage product at 50 kDa which was detected by all the anti-peptide antibodies. This was not consistent with the known processing strategy. Processing at any of the known cleavage sites would not be able to release a 50 kDa protein that can be recognised by the three different anti-VP4 peptide antibodies as was observed here (Fig 4.12).

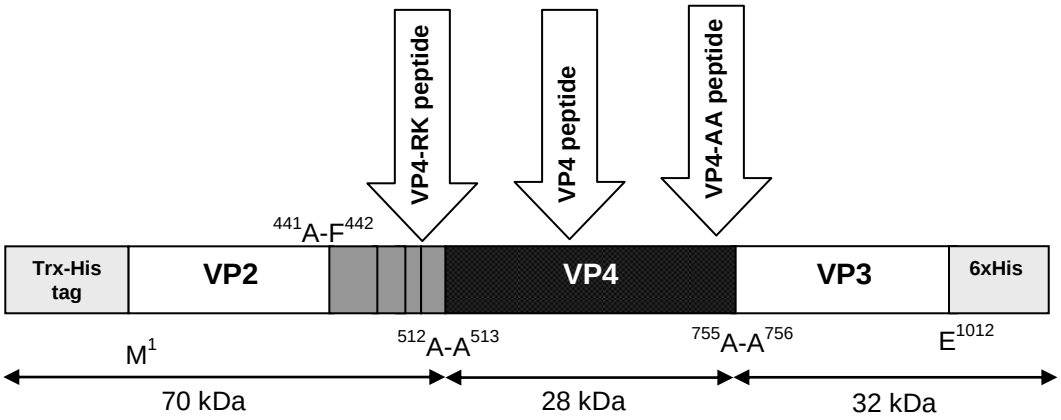


Figure 4.12 Schematic representation of the full-length polyprotein construct showing cleavage sites and anti-peptide antibody binding location. The sizes released according to the current polyprotein processing strategy are shown below.

The only scenario which would result in the observed cleavage pattern was if cleavage is occurring within the VP2 region and at the VP4-VP3 junction. The resulting polypeptide would contain the peptides required for recognition by the three anti-VP4 peptide antibodies. The IBDV VP4 cleavage sites are defined by the (Thr/Ala)-X-Ala↓Ala, motif (Da Costa *et al.*, 2003). Careful analysis of the polyprotein sequence revealed the presence of an Ala-Val-Ala-Ala sequence (residues 275-280), which matches the defined IBDV VP4 cleavage sites. This suggests that the observed 50 kDa product could have resulted from cleavage of the ^{279}Ala-Ala^{278} and ^{755}Ala-Ala^{756} dialanine peptides. The ^{279}Ala-Ala^{278} site has never been suggested to be a cleavage site, but it is the only site that would result in a 50 kDa product recognised by all three anti-VP4 peptide antibodies. If cleavage occurs at this site, it still remains unsolved why the cleavage at the ^{512}Ala-Ala^{513} cleavage site would be prevented.

A more logical and attractive explanation was that the Pro^{377} residue may influence the overall fold of the proteins within the polyprotein. This was motivated by the observations made by Rodriguez-Lecompte and Kibenge (2002) where a D590P mutation resulted in partial polyprotein processing. It was thus concluded that IBDV VP4 has the ability to act according to structural changes during translational and posttranslational processes. An interesting observation was that Pro^{377} was found in a conserved Pro-Asn-Pro-Glu motif. The Asx-Pro sequence and similar motifs adopt a turn conformation that is stabilised by hydrogen bonds (Wilson and Finlay, 1997) (Fig. 4.13).

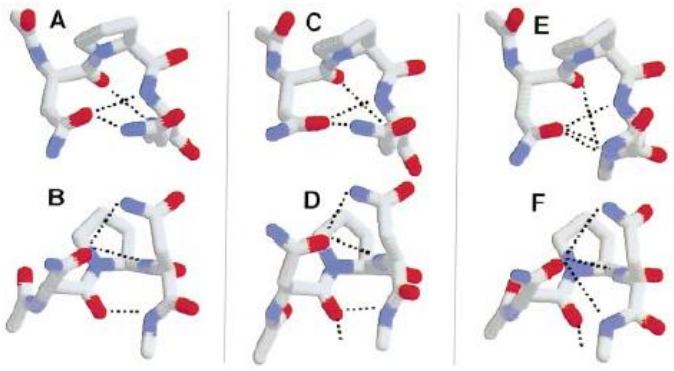


Figure 4.13 Molecular geometry of the hydrogen bonds that form Asx–Pro turns in three different Asn–Pro–Asn sequences. A and B) Residues 41-43 of glucose oxidase from *Aspergillus niger.* **C and D)** Residues 150-152 of nucleoside diphosphate kinase from *Dictyostelium discoideum.* **E and F)** Residues 268-270 of *E. coli* isocitrate dehydrogenase. In views A, C and E, The proline N-C^{α} and C^{α}-C bonds lie in plane of the paper. In views B, D and F, the models are rotated 90°around the axis (Wilson and Finlay, 1997).

These turns commonly function as helix caps which stabilise helices at the C-terminal end and terminate them at the N-terminal end (Wilson and Finlay, 1997). Given the presence of this motif and its characteristics for the formation of hydrogen bonds required for turns

and helix caps, it was now of interest to determine whether introducing the P377L mutation would change the 3-D structure of VP2. The generated model suggested that the mutation caused a rearrangement of some secondary structures within VP2 domains. The findings strongly suggest the importance of this residue in the polyprotein processing. However, it still remains unclear how the mutation may influence the arrangement of the catalytic residues or the cleavage sites, since there is no crystal structure available that encompass VP2 and VP4 or rather the whole polyprotein.

The autocatalytic processing observed in the constructs was certainly mediated by the VP4. It may also be cautiously suggested that the polypeptide Arg^{453}-Ala^{512} or the entire pep70 (Ala^{442}-Ala^{512}) may be required as a cofactor for the autocatalytic activity of VP4. Further investigation is required to verify this speculation. The processing pattern shown in the preceding chapter suggested the use of a different site for the cleavage of the pVP2-VP4 junction. Western blot analysis of the VP4-RA and VP4-RK processing products suggested that the N-terminal cleavage occurs at the ^{487}Ala-Ala^{488}. However, N-terminal sequencing of the products showed that cleavage occurs at the ^{512}Ala-Ala^{513} site. Further investigation was therefore required to clarify these conflicting results.

CHAPTER 5
The influence of the Arg-Ile dipeptides on the autocatalytic activity of VP4

5.1 Introduction

Infectious bursal disease virus (IBDV) polyprotein (NH_2-pVP2-VP4-VP3-COOH) has been shown to be co-translationally processed through the proteolytic activity of VP4 to produce pVP2 (Met^{1}-Ala^{512}), VP4 (Ala^{513}-Ala^{755}) and VP3 (Ala^{756}-Glu^{1012}). The idea of co-translational processing was not directly demonstrated, but was just strongly suggested (Jagadish *et al.*, 1988; Vakharia *et al.*, 1993). pVP2, the VP2 precursor, is further processed at its C-terminus to generate VP2 and structural peptides through the cleavage of the ^{441}Ala-Phe^{442} bond and three other Ala-Ala bonds between residues 487 and 488, 494 and 495, as well as 501 and 502 (Da Costa *et al.*, 2003).

Maturation through proteolytic processing of pVP2 is associated with virus assembly (Chevalier *et al.*, 2005). The resulting peptides i. e. pep46 (Phe^{442}-Ala^{487}), pep7a (Ala^{488}-Ala^{494}), pep 7b (Ala^{495}-Ala^{501}), and pep11 (Ala^{502}-Ala^{512}) are involved in the formation of the viral particle (Fig. 5.1). Pep46 was shown to induce large structural rearrangements in liposomes and to destabilise membranes in IBDV permissive and non-permissive cells, demonstrating its involvement in cell entry (Chevalier *et al.*, 2005; Galloux *et al.*, 2007). The deletion of pep46 and pep11 were shown to prevent virus recovery and infectivity, whereas substitutions and deletions of pep7a and pep7b did not have any detrimental effect on viral replication (Da Costa *et al.*, 2002). These two peptides (pep46 and pep11) were also shown to play a crucial role in IBDV assembly. Polyprotein constructs with a pep46 deletion could not assemble into IBDV-like particles (VLPs) but instead assembled into subviral particles (SVPs). Pep11 deletion resulted in the abrogation of the capsid morphogenesis (Chevalier *et al.*, 2005). These findings demonstrate the importance of these two peptides in viral assembly and replication.

The sequences of both pep46 and pep11 contain an Arg-Ile dipeptide. The Arg-Ile bond is common in the activation of serine proteases such as complement proteases (Arlaud and Gagnon, 1985; Spycher *et al.*, 1986; Dobo *et al.*, 1999; Arlaud *et al.*, 2002), mannan-binding lectin associated serine protease (Thiel *et al.*, 1997) and plasma proteases (Chung *et al.*, 1986). This mechanism of cleavage of an Arg-Ile bond is used for conversion of zymogens (or proenzyme) to a mature active form of the protease (Khan *et*

al., 1999). Viral proteins are initially expressed as a precursor polyprotein, which is subsequently autocatalytically cleaved by proteases which are encoded as integral parts of the polyprotein. Khan *et al.* (1999) suggested that the viral polyprotein resembles the organisation of proenzymes. The viral proteases usually cleave themselves from the polyprotein, and thereafter carry out the *trans* cleavages that are crucial for the virus life cycle.

Accurate and precise processing is required for the production of fully infectious viral particles. In picornaviruses, the polyprotein processing requires both host and viral proteases. In HAV, 3C protease cleaves the 2A-2B junction in the precursor polyprotein (VP0-VP3-VP1-2A-2B-2C-3A-3C-3D) to release the capsid protein precursor P1-2A. P1 comprise the N-terminal VP0-VP3-VP1 domains. The capsid protein precursor P1-2A is further cleaved into VP0, VP3 and VP1-2A by 3C protease. Domain 2A facilitates the folding of P1-2A required for cleavage by 3C protease (Cohen *et al.*, 2002; Rachow *et al.*, 2003). Following the cleavage of P1-2A precursor by 3C protease, five copies of each of VP0, VP3 and VP2-2A assemble into pentameric immature particles. The immature particles are then cleaved by host proteases at the VP1-2A junction to release 2A, which results in the maturation of these particles. In other words, the mature particles are devoid of 2A (Cohen *et al.*, 2002; Rachow *et al.*, 2003). In a study by Morace *et al.* (2008), the treatment of immature particles with factor Xa, trypsin and cathepsin L *in vitro* resulted in the conversion of VP1-2A to VP1, thus leading to the maturation of the particles. The VP1-2A junction comprises the Arg-Ile dipeptide. When Arg was mutated into Met or Glu, the capsid protein precursor P1-2A was inefficiently cleaved by 3C protease and no assembly was observed (Morace *et al.*, 2008).

Two of these Arg-Ile peptides (residues 453-454 and 505-506) are present in the IBDV polyprotein, within pep46 (Phe^{442}-Ala^{487}) and pep11 (Ala^{502}-Ala^{512}) which were shown to be important for VLP assembly and capsid morphogenesis, respectively (Chevalier *et al.*, 2005). Western blot analysis of the VP4-RA and VP4-RK processing products revealed that the N-terminal cleavage occurs at the ^{487}Ala-Ala^{488} (Chapter 3). However, N-terminal sequencing of the products showed that cleavage occurs at the ^{512}Ala-Ala^{513} site. These conflicting results necessitated further investigation for clarification.

The two dialanine peptides, ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} are the C-terminal cleavage sites for the release of pep 46 and pep11, respectively. The presence of Arg-Ile dipeptides before the ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} cleavage sites was interesting. It was therefore speculated that the ^{453}Arg-Ile^{454} dipeptide may be important for the cleavage at

^{487}Ala-Ala^{488} while ^{505}Arg-Ile^{506} is important for cleavage at the ^{512}Ala-Ala^{513} sequence. As described in the preceding chapters, the expression of VP4 constructs demonstrated that while VP4-RK (Arg^{453}-Lys^{723}) and VP4-RA (Arg^{453}-Ala^{755}) were autocatalytic, VP4-AA (Ala^{513}-Ala^{755}) was not. It was therefore suggested that the polypeptide Arg^{453}-Ala^{512} or the entire pep70 (Ala^{442}-Ala^{512}) may be required as a cofactor for the autocatalytic activity of VP4. However, this speculation also required further investigation.

To investigate the above speculations, different constructs were designed i.e. VP4-R^{1}W (Arg^{453}-Trp^{891}), VP4-R^{2}W (Arg^{505}-Trp^{891}), VP4-R^{1}A (Arg^{453}-Ala^{755}) and VP4-R^{2}A (Arg^{505}-Ala^{755}) (Fig. 5.1).

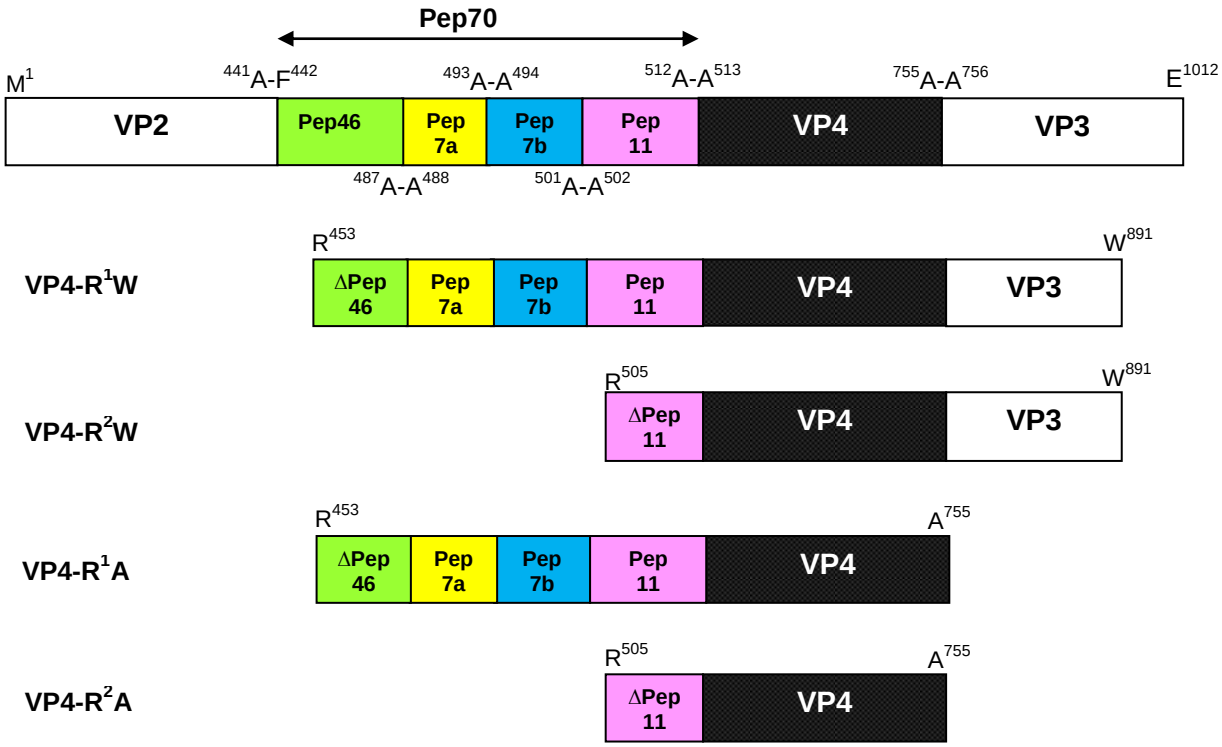


Figure 5.1 Schematic diagram of the IBDV polyprotein fragments constructed for analysing the influence of the Arg-Ile dipeptide on the autocatalytic activity of VP4. The constructs are shown in comparison to the full-length polyprotein organisation shown at the top.

It was shown that VP4-R^{1}A (Arg^{453}-Ala^{755}) is autocatalytically active (Chapter 3) therefore the rationale for designing these constructs was that, should VP4-R^{2}A be autocatalytically cleaved, the speculation that the Arg^{453}-Ala^{512} polypeptide or the entire pep70 (Ala^{442}-Ala^{512}) could be a cofactor for the autocatalytic activity of VP4, would be incorrect. Since VP4-AW (Ala^{513}-Trp^{891}) was not autocatalytically cleaved (Chapter 3), VP4-R^{1}W and VP4-R^{2}W were included in the present study to determine the effect of the polypeptide Arg^{453}-Ala^{512} and truncated pep11 (Δpep11) on the processing. The constructs were recombinantly expressed, analysed by SDS-PAGE and western blotting using anti-His tag and anti-VP4 peptide antibodies.

5.2 Materials

The following molecular biology reagents were obtained from Fermentas (Vilnius, Lithuania): EcoRI, NotI, SAP, T4 DNA ligase, 10 mM dNTP mix, X-gal, FastRuler® middle range DNA ladder, GeneRuler 1 kb DNA ladder mix, Taq DNA polymerase, Long PCR enzyme mix, GeneJET™ plasmid miniprep Kit, pTZ57R/T and TransformAid™ bacterial transformation kit. The peqGOLD gel extraction kit was obtained from PEQLAB Biotechnologie (Erlangen, Germany) and DNA clean and concentration kit™ was obtained from Zymo Research. The primers were synthesised at the MCB Synthetic DNA Unit (University of Cape Town, South Africa). Anti-His tag monoclonal antibody was purchased from Novagen (Darmstadt Germany). The general reagents were purchased from Sigma-Aldrich-Fluka (Steinheim Germany) and Merck Biosciences (Darmstadt, Germany).

5.3 Methods

5.3.1 Amplification of Arg-Ile dipeptide containing VP4 fragments

The designed constructs were amplified from IBDV segment A cDNA (Section 2.3.5) using the primers shown in Table 5.1.

Table 5.1 Forward and reverse primers for the Arg-Ile dipeptide containing constructs

Fragment	Forward primer	Reverse primer	Size (bp)
VP4-R^1W	VP4-RK-F (5' cc**gaattc**aggata gctgtgccggtggtctcc 3')	VP2I-3W-R (5' cgg**gcggccgc**cc agtactttagctggccggggcttgg 3')	1137
VP4-R^1A *	VP4-RK-F (5' cc**gaattc**aggata gctgtgccggtggtctcc 3')	VP4-AA-R (5' ac**gcggccgc**agcca tggcaaggtggtactggc 3')	909
VP4-R^2W	VP4-RI-F (5' gcc**gaattc**cgcata aggcagctaactctcgcc 3'	VP2I-3W-R (5' cgg**gcggccgc**cc agtactttagctggccggggcttgg 3')	1160
VP4-R^2A	VP4-RI-F (5' gcc**gaattc**cgcata aggcagctaactctcgcc 3'	VP4-AA-R (5' ac**gcggccgc**agcca tggcaaggtggtactggc 3')	753

* The construct was amplified and cloned as desribed in Chapter 2. The restriction extensions are shown in bold.

The PCR reaction was prepared in a total volume of 50 µl containing 10x *Taq* buffer with $(NH_4)_2SO_4$ (5 µl), 25 mM $MgCl_2$ (5 µl), 10 mM dNTP mix (5 µl), 10 µM forward and reverse primers (1 µl each), 10 U/µl Taq DNA polymerase (0.3 µl), 1 µl of the IBDV segment A cDNA and nuclease-free water (35.7 µl). Amplification was carried out using the following temperature profile: initial denaturation at 94°C for 5 min, 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 1 min, extension at 72°C for 1.5 min and a final

extension cycle at 72°C for 7 min. The PCR products were analysed on a 1% (w/v) agarose gel and stained with ethidium bromide (Section 2.3.3).

5.3.2 Cloning of Arg-Ile dipeptide containing VP4 fragments

The PCR products were purified from the agarose gel using peqGOLD gel extraction kit and ligated into pTZ57R/T. The ligation was carried out using T4 DNA ligase at an insert:vector ratio of 4:1, and the reaction was incubated overnight at 4°C. The ligation mix was transformed into competent *E. coli* JM109 cells prepared using a TransformAid™ bacterial transformation kit and transformed cells were grown overnight on 2xYT-Amp agar plates containing 0.1 mM IPTG and 0.005% (w/v) X-gal. Recombinant colonies were screened by colony PCR using Taq DNA polymerase and specific primers. A single colony containing a recombinant T-vector was grown overnight in 2xYT-Amp broth and plasmid DNA was isolated using the GeneJET™ plasmid miniprep kit, digested with EcoRI and NotI to release the insert DNA. All double digestion reactions with EcoRI and NotI were performed overnight in a 30 µl reaction containing 10x Buffer O (3 µl), plasmid (25 µl), EcoRI (1 µl) and NotI (1 µl) at RT and analysed on a 1% (w/v) agarose gel (Section 2.3.3). The insert DNA was purified from the agarose gel using peqGOLD gel extraction kit. The purified inserts were then ligated into dephosphorylated pET-32b prepared as before (Section 2.3.6) using T4 DNA ligase overnight at 4°C. The ligation mixtures were used to transform *E. coli* BL21 (DE3) cells using the TransformAid™ bacterial transformation kit and transformed cells were grown overnight on 2xYT-Amp agar plates. Transformants were screened by colony PCR using vector primers (Appendix 2) and recombinant colonies were grown overnight at 37°C in 2xYT-Amp broth for glycerol stock preparation.

5.3.3 Recombinant expression of Arg-Ile dipeptide containing VP4 constructs

Expression of pET-32b recombinant plasmids in *E. coli* BL21 (DE3) cells was carried out by self-induction. A single colony was grown overnight in Terrific Broth containing 100 µg/ml ampicillin in a shaking incubator at 37°C or for 24 h at 16°C. The cells were harvested by centrifugation (5000 x g, 10 min, 4°C) and the cell pellet was resuspended into 1/20 culture volume of lysis buffer and frozen overnight at -20°C. The thawed samples of *E. coli* BL21 (DE3) cell lysates were sonicated, analysed by reducing SDS-PAGE and western blotting according to Section 3.3.3.

5.4 Results

5.4.1 Amplification and cloning of Arg-Ile containing VP4 coding regions

The initial step was the amplification of the fragments from IBDV segment A cDNA. The expected products of approximately 1300 bp (VP4-R^1W), 750 bp (VP4-R^2A) and 1200 bp (VP4-R^2W) were produced (Fig. 5.2, lanes 1 to 3). The Taq polymerase amplified PCR products contained a 3' adenosine overhang hence allowing for efficient TA cloning into a T-vector. Successful cloning of the PCR products into a T-vector interrupts the β-galactosidase gene, which allows for the recombinant clones to be easily identified by colour screening on Xgal/IPTG plates. The selected white colonies were further screened for the presence of the inserts by miniprep DNA isolation followed by restriction enzyme digestion. The primers were designed with EcoRI and NotI restriction sites therefore a double digestion with these enzymes would release the insert DNA from the recombinant plasmid.

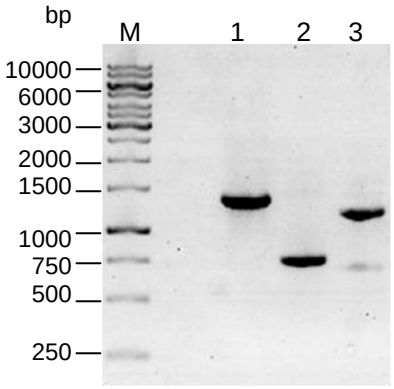


Figure 5.2 Agarose gel (1%) analysis of coding regions of Arg-Ile containing VP4 PCR products. Lane M, GeneRuler 1kb DNA ladder; lane 1, VP4-R^1W; lane 2, VP4-R^2A and lane 3, VP4-R^2W.

All three fragments were successfully cloned into the T-vector (Fig. 5.3). Double digestion of the recombinant pTZ57RT plasmids with EcoR1 and Not1 released inserts at the expected sizes for VP4-R^1W (1300 bp), VP4-R^2A (750 bp) and VP4-R^2W (1200 bp) (Fig. 5.3A). The released inserts were successfully purified (Fig. 5.3B) for cloning into a pET-32b expression vector that was also subjected to EcoRI and NotI digestion, followed by dephosphorylation. The dephosphorylated plasmid was purified and analysed by agarose gel electrophoresis (Fig 5.3C) and it was observed at its expected size of 6 kb.

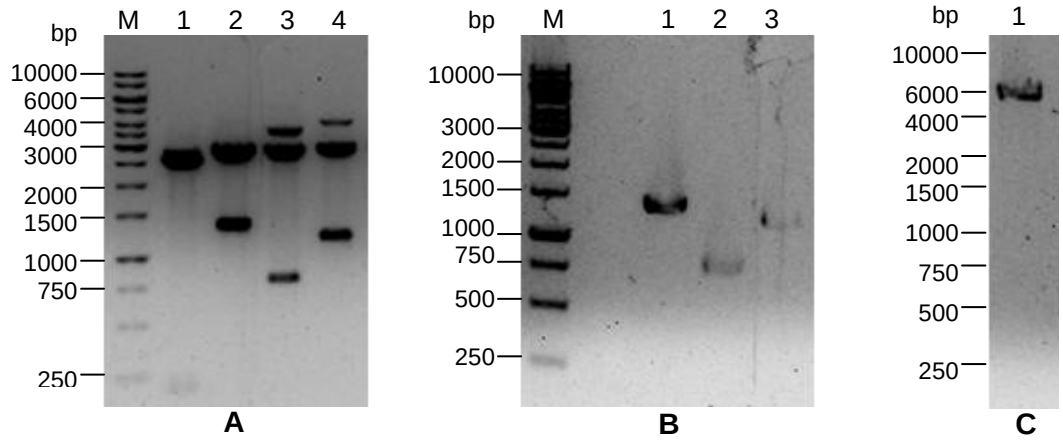


Figure 5.3 Agarose gel electrophoresis analysis pTZ5RT vector cloning of the VP4 fragments. A) Release of inserts by EcoR1/Not1 digestion of recombinant plasmids. Lane 1, positive control; lane 2, VP4-R^1W; lane 3, VP4-R^2A and lane 4, VP4-R^2W. **B)** Purification of insert DNA. Lane 1, VP4-R^1W; lane 2, VP4-R^2A and lane 3, VP4-R^2W. **C)** Purification of EcoR1/Not1 digested and dephosphorylated expression plasmids. Lane 1, pET-32b.

Subcloning into the pET-32b expression vector was screened by direct colony PCR using vector primers. For successful cloning, the PCR would produce amplicons with the size of the insert plus 750 bp. Judging by the obtained sizes of the amplicons (1950 bp for VP4-R^2W, 2050 bp for VP4-R^1W and 1500 bp for VP4-R^2A), all three fragments were successfully cloned into pET-32b (Fig. 5.4).

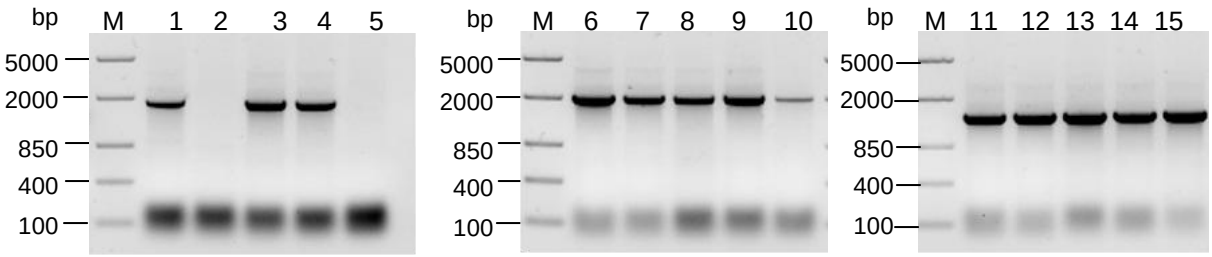


Figure 5.4 Colony PCR for screening recombinants from subcloning of VP4 fragments into pET-32b. Lane M, GeneRuler DNA ladder; lanes 1-5, colonies 1-5 for screening VP4-R^2W; lanes 6-10, colonies 1-5 for screening VP4-R^1W and lanes 11-15, colonies 1-5 for screening VP4-R^2A.

5.4.2 Analysis of expression

The cell lysates were analysed by reducing SDS-PAGE followed by western blotting using anti-VP4 peptide antibodies and an anti-His tag monoclonal antibody (Fig. 5.5). All the constructs were expressed as cleavage products, with the major cleavage product at 24 kDa protein as shown by SDS-PAGE analysis (Fig 5.5A). Western blotting was able to reveal the other cleavage products that could not be distinguished on a gel, with anti-His tag monoclonal antibody revealing more about processing than the anti-VP4 peptide antibodies.

The constructs starting at Arg^{505} contained the ^{512}Ala-Ala^{513} cleavage site only at the N-terminus. Therefore, with the N-terminal cleavage limited to this one site, for VP4-R^2A two products of 18 and 28 kDa corresponding to the Trx-His tagged Arg^{505}-Ala^{512} and Ala^{513}-Ala^{755} polypeptide would be expected (Fig 5.5B). Anti-His tag antibodies recognised the 18 kDa product (Fig 5.5C, lane 1) while the anti-VP4 common and anti-VP4-AA peptide antibodies recognised the 28 kDa product (Fig 5.5E and G, lane 1). The anti-VP4-RK peptide antibodies did not recognise any of the products as expected (Fig 5.5F, lane 1). For VP4-R^2W three products of 18, 28 and 16 kDa corresponding to the Trx-His tagged Arg^{505}-Ala^{512}, Ala^{513}-Ala^{755} polypeptide and a C-terminally 6xHis tagged truncated VP3 (∆VP3-His) would be expected (Fig 5.5B). The anti-His tag antibodies recognised the 16 and 18 kDa products as expected (Fig 5.5C, lane 3). The anti-VP4 common and anti-VP4-AA peptide antibodies recognised the 28 kDa product as expected (Fig 5.5E and G, lane 3) while the anti-VP4-RK peptide antibodies did not recognise any product as expected.

VP4-R^1A construct consisted of an N-terminal Trx-His tag and Arg^{453}-Ala^{755} polypeptide while VP4-R^1W construct comprised an N-terminal Trx-His tag, the Arg^{453}-Ala^{755} polypeptide and ∆VP3-His (Fig. 5.5D). The VP4-R^1A construct did not have the VP4-VP3 junction thus cleavage would only occur at the N-terminus. The anti-His tag monoclonal antibody recognised a doublet at about 24 kDa, thus confirming the cleavage at the ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} sites (Fig 5.5B and C, lane 2). The anti-VP4 common and anti-VP4-AA peptide antibodies recognised the 28 kDa product as expected (Fig 5.5E and G, lane 2) while the anti-VP4-RK peptide antibodies recognised the 24 kDa corresponding to Trx-His tagged Arg^{453}-Ala^{512} polypeptide (Fig 5.5F, lane 2).

For VP4-R^1W construct, the anti-His monoclonal antibody recognised a doublet at about 24 kDa and a 16 kDa protein. The 16 kDa product corresponds to ∆VP3-His (Fig 5.5 C, lane 4). The release of a doublet implies cleavage at two sites, the ^{487}Ala-Ala^{488} to release a 22 kDa product and at the ^{512}Ala-Ala^{513} site to release the expected 24 kDa. The anti-VP4 common and anti-VP4-AA peptide antibodies recognised the 28 kDa product as expected (Fig 5.5E and G, lane 4) while the anti-VP4-RK peptide antibodies recognised the 24 kDa product corresponding to Trx-His tagged Arg^{453}-Ala^{512} polypeptide (Fig 5.5F, lane 4).

The anti-VP4 common peptide antibodies also identified bands at about 30 kDa for VP4-R^1W and VP4-R^1A (Fig. 5.5E, lanes 2 and 4). For all the constructs, the anti-VP4-AA peptide antibodies bound weakly to the cleavage products (Fig. 5.5G).

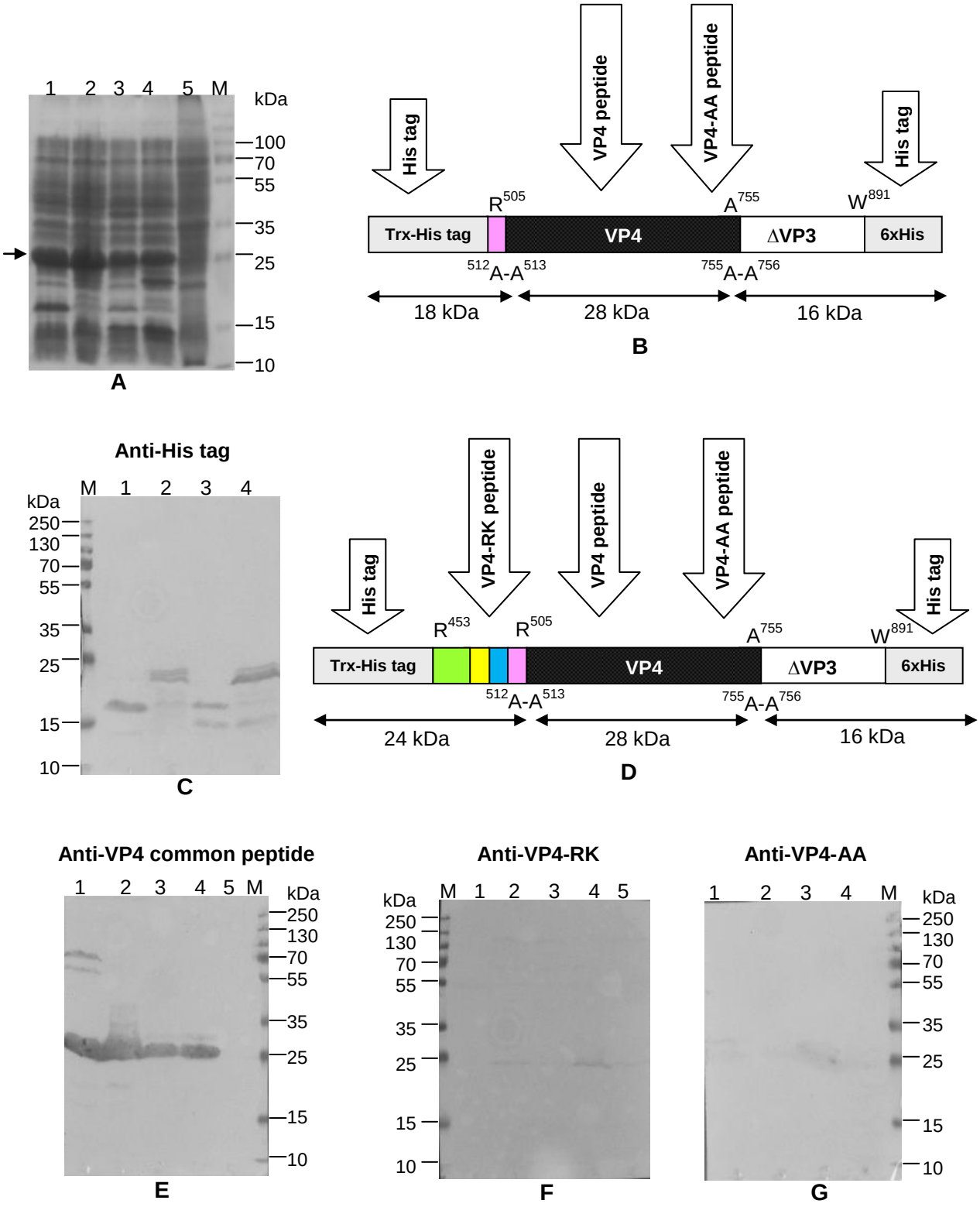


Figure 5.5 Expression profile of the VP4 constructs. A) Reducing SDS-PAGE analysis. The arrow shows the major expressed protein observed on the SDS-PAGE gel. **B)** Schematic representation of the constructs starting at R^{453} showing anti-peptide antibody binding locations. **D)** Schematic representation of the constructs starting at R^{505} showing anti-peptide antibody binding locations. Western blot analysis of the expression was carried out using anti-His tag monoclonal antibody **(C),** anti-VP4 common peptide antibodies **(E),** anti-VP4-RK peptide antibodies **(F)** and anti-VP4-AA peptide antibodies **(G).** Lane M, PageRuler™ plus prestained protein ladder; lane 1, VP4-R^{2}A; lane 2, VP4-R^{1}A; lane 3, VP4-R^{2}W; lane 4, VP4-R^{1}W and lane 5, *E. coli* BL21 (DE3*)* control cell lysate.

5.5 Discussion

In contrast to VP4-AA (Ala^{513}-Ala^{755}), VP4-RA (Arg^{453}-Ala^{755}) was shown to be autocatalytic. This observation led us to propose that the Arg^{453}-Ala^{513} polypeptide may be critical for the autocatalytic activity of VP4. The interesting feature of the polypeptide is the presence of two Arg-Ile dipeptides. The aim of the present study was to determine if the fragments starting from the second Arg-Ile dipeptide would also be autocatalytic. Western blot analysis of VP4-RA expression (Chapter 3) suggested that the cleavage occurred at the ^{487}Ala-Ala^{488} instead of the ^{512}Ala-Ala^{513} site. However, N-terminal sequencing of the product revealed that the cleavage occurred at the ^{512}Ala-Ala^{513} site. It was hypothesised that Arg-Ile dipeptides may influence cleavage site preference for the embedded VP4. This hypothesis was influenced by the fact that both the Arg-Ile dipeptides are present N-terminal to the cleavage sites.

The fragments starting from the first Arg-Ile dipeptide showed the same autocatalytic efficiency as the fragments starting from the second dipeptide. The first Arg-Ile is about 60 residues away from the ^{512}Ala-Ala^{513} cleavage site whereas the second is six residues away. This indicates that the whole polypeptide is not required for the autocatalytic activity of VP4. The six residues (Arg^{505}-Ala^{513}) had the same effect on the autocatalytic activity of VP4 as the Arg^{453}-Ala^{513} polypeptide. The observations made in the present study suggested that the cleavage was occurring at both ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} cleavage sites. It still remains unclear what determines the cleavage at a particular site. But it is clear that the Arg-Ile dipeptides do not influence the cleavage site preference. The findings further suggested that the cleavage starts at the C-terminus for the autocatalytic release of VP4. For example, for VP4-R^{1}W cleavage at the C-terminal site would release VP4-RA (Arg^{453}-Ala^{755}), which retains the autocatalytic activity to cleave the N-terminus, whereas initial cleavage at the N-terminus would release (Ala^{513}-Trp^{891}), which was shown not to cleave in *cis*. Even though these dipeptides may not play a crucial role in the autocatalytic activity of VP4, further studies are still required to determine their role in the viral infection cycle. Picornavirus HAV polyprotein has an Arg-Ile dipeptide at the VP1-2A junction cleaved by host proteases for efficient assembly (Morace *et al.*, 2008). Other cases of the involvement of host cellular proteases in the viral infection cycle have been described in SARS coronavirus (Belouzard *et al.*, 2009; Belouzard *et al.*, 2010) and influenza virus (Chaipan *et al.*, 2009; Choi *et al.*, 2009). The Arg-Ile peptides in the IBDV polyprotein are found within the pep46 stretch which plays a role in viral assembly and entry and thus would be interesting to know if the host cell proteases may be involved an any stage of the IBDV infection cycle.

CHAPTER 6
General Discussion

Infectious bursal disease virus (IBDV) is the causative agent of a highly contagious disease in chickens known as Gumboro disease, named after Gumboro, Delaware, USA where the first outbreak of the disease was reported (Cosgrove, 1962). The disease was later named infectious bursal disease (IBD) due to its tropism for the bursa of Fabricius, where it destroys the developing B-cells causing severe bursal damage, and consequently immunosuppression (Muller *et al.*, 2003). The disease is highly contagious and spreads quickly through flocks, thus resulting in major economic losses for the poultry industry. It is distributed through contaminated feed and water and is commonly diagnosed in 3-6 week old chickens (Muller *et al.*, 2003).

The disease became known in 1957 in the USA and now occurs in most parts of the world. Classical strains isolated from chickens in the USA in the early 1960's induced bursal lymphoid cell necrosis and immunosuppression with flock mortality up to 20%. In the late 1980s, variant strains that were highly cytolytic and caused rapid bursal atrophy were isolated from flocks, also in USA (Snyder *et al.*, 1988). In the late 1980s, very virulent IBDV (vvIBDV) strains, causing severe clinical symptoms and high mortality emerged in Europe, Asia and Africa (Chettle *et al.*, 1991; van den Berg and Meulemans, 1991). Attenuated strains were adapted to primary cell culture for use as vaccines (Brandt *et al.*, 2001).

Commercial vaccines are divided into three groups: mild, intermediate and intermediate-plus vaccines (van den Berg *et al*., 2000). Successful IBDV vaccination depends on several factors such as time of vaccination, the chosen vaccine strain, the maternal derived antibody levels and epidemiological field conditions (Rautenschlein *et al.*, 2005). The viruses in mild and intermediate vaccine are often destroyed by maternal derived antibodies, and are therefore used in chicks that have a low level of maternal derived antibodies. On the other hand the intermediate-plus vaccines are suitable for chickens with high levels of maternal derived antibodies (Tsukamoto *et al.*, 1995; Rautenschlein *et al.*, 2005). An IBDV recombinant vaccine in which the VP2 gene from variant E strain was inserted into the turkey herpesvirus genome (rHVT/IBD) was developed. The efficacy of the rHVT/IBD vaccine was evaluated in specific pathogen free chickens and commercial broilers. A single vaccination protected chickens against classic, variant and very virulent

strains with no interference of maternal derived antibodies (Tsukamoto *et al.*, 2002; Goutebroze *et al.*, 2003; Sato *et al.*, 2004).

Despite the success of existing vaccines, maternal derived antibodies and the emergence of very virulent strains still affect the efficacy of the vaccines (van den Berg *et al*., 2000). In view of these problems, new measures for disease control are required. A promising strategy is to develop drugs that target crucial processes in the virus life cycle such as polyprotein processing. Polyprotein processing is a mechanism used by viruses to release their structural and non-structural proteins. This processing is mediated by proteases which are encoded as integral parts of the viral polyprotein. These viral proteases are therefore attractive targets for anti-viral drug development (Feldman *et al.*, 2006). Viral protease inhibitor based anti-viral drugs have proved successful and have been approved for the treatment of HIV infections (Lee *et al.*, 2007).

IBDV belongs to the *Birnaviridae* family which consists of icosahedral non-enveloped viruses of about 60 nm in diameter containing a bisegmented dsRNA genome. Segment A contains two overlapping ORFs and the larger of these is translated into a polyprotein (NH_2-pVP2-VP4-VP3-COOH). The polyprotein is processed through the proteolytic activity of VP4 (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000). Polyprotein processing by VP4 results in the activation of VP1, the RNA-dependent RNA polymerase encoded by segment B (Birghan *et al*., 2000). In turn, VP1 binds to several VP3 molecules to promote the assembly of new virus particles. The assembly can therefore be correlated to polyprotein processing, demonstrating that VP4 protease is necessary for the appropriate assembly of the infectious viral particles (Chevalier *et al.*, 2004).

Several investigations have already been conducted on IBDV polyprotein processing. Firstly, an *in vitro* translation study suggested that the Australian strain 002-73 segment A codes for three structural proteins in a polycistronic fashion (Azad *et al.*, 1985). Hudson *et al.* (1986) then showed that these three proteins were encoded within a continuous ORF, which codes for a polyprotein that is processed into these three major proteins, VP2, VP4 and VP3. This processing was also observed when the polyprotein coding region was expressed as a fusion protein in *E. coli,* however, their results could not distinguish whether the processing was mediated by virus-encoded or cellular proteases. Although sequencing of the three polyprotein processing products was not successful due to the blockage of the N-termini of the VP3 and VP4 polypeptides, it was suggested that the dibasic residues, ^{452}Arg-Arg^{453} and ^{722}Lys-Arg^{723} could be the cleavage sites. The

subsequent study by Jagadish *et al.* (1988) showed that VP4 processes the polyprotein to release VP2 and VP3. They also showed that the mutation of ^{722}Lys-Arg^{723} into ^{722}Ile-Cys^{723} inactivated VP4 function.

Expression of IBDV segment A cDNAs by *in vitro* translation in rabbit reticulocyte lysates and in the Sindbis virus expression system also resulted in polyprotein processing. It was shown that the VP4 C-terminus is essential for processing at the pVP2-VP4 junction and and the VP4 N-terminus for processing at the VP4-VP3 junction. A translational time course was used to distinguish primary and secondary processing products, where it was shown that pVP2 is not further processed into VP2 in these expression systems (Kibenge *et al.*, 1997).

Sanchez and Rodriguez (1999) expressed mutant forms of the polyprotein containing either deletions or single amino acid mutations within the pVP2-VP4 and VP4-VP3 junctions in African green monkey kidney epithelial cells. The polyprotein was shown to undergo processing at ^{512}Ala-Ala^{513} (pVP2-VP4 junction) and ^{755}Ala-Ala^{756} (VP4-VP3 junction), to yield pVP2, VP4 and VP3. However, processing was still observed when the ^{512}Ala-Ala^{513} site was deleted suggesting the use of a secondary site within the pVP2-VP4 junction. In the study by Lejal *et al.* (2000) double amino acid mutations of the dialanine peptides, within the pVP2-VP4 and VP4-VP3 junctions, were introduced into the polyprotein. The different mutant constructs were expressed in *E. coli* and their expression profiles were compared to that of the wild-type polyprotein. Their study showed that the three sites at positions 487-488, 494-495 and 501-502 may be associated with the secondary processing of pVP2 into VP2. In addition, Ser^{652} and Lys^{692} were identified as the catalytic residues by site-directed mutagenesis. VP4 was also shown to cleave in *trans* at the ^{512}Ala-Ala^{513} cleavage site within the pVP2-VP4 junction.

Collectively, these studies showed that the IBDV polyprotein is processed by the VP4 serine protease at ^{512}Ala-Ala^{513} (pVP2-VP4 junction) and ^{755}Ala-Ala^{756} (VP4-VP3 junction) sites, to yield pVP2, VP4 and VP3. The precursor VP2 (pVP2) undergoes secondary processing at ^{441}Ala-Phe^{442}, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502}. A study by Da Costa *et al.* (2002) identified a 5 kDa peptide when the IBDV virus particles were analysed by SDS-PAGE. The peptide was identified to be pep46 by N-terminal sequencing. Mass spectrometry analysis on the virus particles was used to show the presence of all four peptides. The secondary processing of pVP2 was proposed to start at the C-terminus and progress towards the N-terminus to release VP2 (Met^{1}-Ala^{441}), and four structural peptides

[pep46 (Phe^{442}-Ala^{487}), pep7a (Ala^{488}-Ala^{494}), pep7b (Ala^{495}-Ala^{501}) and pep11 (Ala^{502}-Ala^{512})] (Da Costa *et al.*, 2002).

The processing of pep70 to release pep46, pep7a, pep7b and pep11 was previously associated with VP4 protease activity (Azad *et al.*, 1987; Kibenge *et al.*, 1997). However, cleavage at ^{441}Ala-Phe^{442} to release VP2 and pep70 (Phe^{442}-Ala^{512}) was shown to be mediated by an autocatalytic endopeptidase activity of the capsid protein VP2. It was shown that pVP2 is processed in the absence of VP4 protease (Lee *et al.,* 2004). In the study, pVP2 was processed into VP2 which resulted in the formation of SVPs, when it was expressed in Hi-5 insect cells. Irigoyen *et al*. (2009) proposed Asp^{391} and Asp^{431} as candidate catalytic residues for proteolytic activity at the ^{441}Ala-Phe^{442} based on the VP2 X-ray structure. By using site-directed mutagenesis of these two residues, they showed that the endopetidase activity of VP2 is associated with Asp^{431} which is situated in a flexible loop preceding the most C-terminal α-helix, thus suggesting that VP2 is a capsid protease (Irigoyen *et al.*, 2009).

Despite the findings in previous investigations on the IBDV polyprotein processing, some events surrounding the cleavages are still not clear. Firstly, it was previously believed that pVP2 is processed from the C-terminus by VP4 to release the four peptides, pep46, pep11, pep7a and pep7b. The fact that pVP2 processes itself at the ^{441}Ala-Phe^{442} site, makes the processing of pVP2 at ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495} and ^{501}Ala-Ala^{502} reported by Da Costa *et al.* (2002) unclear. Secondly, from observations made in previous studies, the present study hypothesised that IBDV VP4 exists in two forms, the embedded form and a mature form. The embedded form exists as an integral part of the polyprotein and it autocatalytically cleaves the ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} peptide bonds to produce the mature form (Ala^{513}-Ala^{755}). Therefore processing generates a mature protease from an autocatalytically active VP4 embedded in the polyprotein. It was of interest to determine whether the stretch of amino acids within the polyprotein that makes up the embedded protease is the same as the mature form.

Thirdly, processing of the polyprotein at one site may be a prerequisite for processing at another site (Bartenschlager *et al.*, 1995) but the order in which processing of the IBDV polyprotein occurs has not yet been described. Lastly, previous studies conducted on the IBDV polyprotein processing used anti-VP2, anti-VP3 or anti-IBDV antibodies (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Kibenge *et al.*, 1997; Sanchez and Rodriguez, 1999) and radiolabelling (Lejal *et al.*, 2000) to characterise the processing products. Until now, there has not been any report on the use of anti-VP4 antibodies to identify processing products.

It is believed that the previous methods of detection would not be able to identify all the resulting processing products. For this reason, anti-VP4 peptide antibodies were used for the detection of the cleavage products that could have been missed in previous studies. This is the first study to use anti-VP4 peptide antibodies for the characterisation of polyprotein processing.

Previous studies used the full-length polyprotein (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Kibenge *et al.*, 1997; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000) and selected fragments of the polyprotein such as Met^{1}-Leu^{686}, Met^{1}-Thr^{557}, Met^{145}-Glu^{1012} and Phe^{524}-Glu^{1012} (Kibenge *et al.*, 1997) to characterise polyprotein processing. The constructs used by Kibenge *et al.* (1997) were designed prior to the knowledge of the cleavage sites and VP4 catalytic residues. The aim of the present study was to revisit the processing of the IBDV polyprotein *in vitro.* Therefore, the following constructs were designed for characterisation of IBDV polyprotein processing in the present study: full-length polyprotein (Met^{1}-Glu^{1012}) and a truncated form of the polyprotein (Ile^{227}-Trp^{891}) representing the fully embedded VP4 form, VP4-RA (Arg^{453}-Ala^{755}) and VP4-RK (Arg^{453}-Lys^{722}) which contained the N-terminal cleavage sites (namely ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495}, ^{501}Ala-Ala^{502} and ^{512}Ala-Ala^{513}), VP4-AW (Ala^{513}-Trp^{891}) with the C-terminal cleavage site ^{755}Ala-Ala^{756} and VP4-AA (Ala^{513}-Ala^{755}) that represents the mature form without cleavage sites. All these fragments contained the Ser/Lys catalytic dyad.

Since the mature VP4 protease has the sequence Ala^{513}-Ala^{755}, the full-length and truncated polyprotein would represent a fully embedded protease. VP4-RA and VP4-RK would be N-terminally embedded whereas VP4-AW would be C-terminally embedded. VP4-RA and VP4-RK, containing Arg^{453} as their N-terminal residue, were included based on findings from previous studies (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999). Dibasic residues ^{452}Arg-Arg^{453} and ^{722}Lys-Arg^{723} were initially suggested by Hudson *et al.* (1986) as the polyprotein processing sites. The mutation of ^{722}Lys-Arg^{723} into ^{722}Ile-Cys^{723} mutation inactivated the proteolytic activity of VP4 (Jagadish *et al.*, 1988). However, another study ruled out the involvement of these dibasic residues by showing that their substitution with neutral residues did not affect the proteolytic processing of the polyprotein (Sanchez and Rodriguez, 1999). The aim was therefore to determine if the expression profile of VP4-RK will differ from VP4-AA.

All six fragments were amplified from the viral dsRNA by RT-PCR, cloned into a T-vector and sub-cloned into *E. coli* and *P. pastoris* expression vectors. Since it was previously observed that the polyprotein is processed during expression in *E. coli* (Hudson *et al.*,

1986; Jagadish *et al.*, 1988; Kibenge *et al.*, 1997; Lejal *et al.*, 2000), this expression system was chosen. Although the IBDV structural proteins were shown not to be glycosylated (Kibenge *et al.*, 1997), the VP2 sequence contains three potential N-glycosylation sites. Protein expression in *E. coli* is often problematic, due to either proteolytic degradation or formation of inclusion bodies (Sijwali *et al.*, 2001). The use of eukaryotic expression systems may overcome these insolubility and proteolytic degradation problems (Sorensen and Mortensen, 2005a). For these reasons, the fragments were sub-cloned into a eukaryotic *P. pastoris* expression vector, as this might better support the proper folding of the constructs than the *E. coli* expression system.

The viral RNA used for RT-PCR was isolated from IBDV infected bursae harvested from commercial broilers killed during an IBDV outbreak that took place in the KwaZulu-Natal province of South Africa in 1995, hence the isolate was named, SA-KZN95. The virulent form of IBDV reached South Africa in 1989 (Horner, 1994) and there has been no sequence information on the isolate. In the present study, the cloned polyprotein coding region was sequenced to establish the genetic relatedness of the local isolate. The deduced amino acid sequence was compared with 52 sequences of other serotype I and II strains.

An approach introduced by Xia *et al.* (2008) was used in the present study. They identified the virulence determinants and investigated the origin of a Chinese very virulent Harbin-1 strain, by comparing the nucleotide and the deduced amino acid sequences of its entire genome with 15 other IBDV strains. They identified 19 residues in the polyprotein which could be used as molecular markers of different IBDV pathotypes (Xia *et al.*, 2008). The present study only used the polyprotein coding region for the analysis and 21 of those markers were identified, thirteen in VP2, five in VP4 and three in VP3. The 21 marker residues of SA-KZN95 were identical to those of the Malaysian very virulent UPM94/273 strain. Phylogenetic analysis at nucleotide sequence level placed SA-KZN95 in a single clade with the Chinese very virulent strains, however, at amino acid level it appeared to be among unique very virulent strains that did not form part of a clade.

All the fragments were successfully expressed in the *E. coli* system but not in *P. pastoris.* The fully embedded (full-length and truncated polyprotein) and N-terminally embedded (VP4-RA and VP4-RK) forms of VP4 protease were autocatalytic, resulting in cleavage products. In contrast, the C-terminally embedded form (VP4-AW) did not show any autocatalytic activity but seemed to be cleaved in *trans*. VP4-AA, on the other hand expressed as an uncleaved fusion protein. This implies that the embedded VP4

autocatalytically cleaves itself off to release the mature form (VP4-AA). In light of this, VP4-AA is simply a product of the autocatalytically active VP4, which is embedded within the polyprotein.

Kibenge *et al.* (1997) stated that the VP4 C-terminus and N-terminus are essential for processing at the pVP2-VP4 and VP4-VP3 junctions, respectively. This statement was based on the fact that constructs Met^{1}-Leu^{686} (pVP2 and VP4 N-terminus) and Phe^{524}-Glu^{1012} (VP4 C-terminus and VP3) showed loss of proteolytic activity for VP4-VP3 and pVP2-VP4 junctions, respectively. The reason for their observed lack of activity was the absence of the N-terminal cleavage sites in the Phe^{524}-Glu^{1012} construct and of catalytic residues in the Met^{1}-Leu^{686} constructs. In the present study, VP4-RA (Arg^{453}-Ala^{755}) and VP4-RK (Arg^{453}-Lys^{722}) showed similar proteolytic activity, and that is contrary to the Kibenge *et al.* (1997) statement. Even though VP4-RK was truncated at the C-terminus, it was able to process the N-terminal sites.

The results presented here also suggested that the Arg^{453}-Ala^{512} polypeptide may be crucial for the autocatalytic activity of VP4. The Arg^{453}-Ala^{512} polypeptide is the C-terminal portion of pep70 that is found in the N-terminus of VP4. It was therefore suggested that VP4 protease may be similar to *Flaviviridae* NS3pro which requires the presence of a cofactor to be fully functional (Yusof *et al.*, 2000; Niyomrattanakit *et al.*, 2004; Bessaud *et al.*, 2006). In flaviviruses like Dengue virus, the NS3 requires the NS2B which is located N-terminal of NS3 for efficient proteolytic processing. The NS2B functions as a cofactor and promotes the folding and functional activity of NS3pro. The NS3 protease domain lacking the NS2B part is catalytically inactive (Erbel *et al.*, 2006; Luo *et al.*, 2008).

While previous studies used anti-VP2 or anti-VP3 antibodies or radiolabelling to identify the cleavage products released after processing (Hudson *et al.*, 1986; Jagadish *et al.*, 1988; Sanchez and Rodriguez, 1999; Lejal *et al.*, 2000), the present study used anti-VP4 peptide antibodies to characterise the processing of the polyprotein. The availability of published amino acid sequences of the IBDV polyprotein enabled the selection of immunogenic peptides using an epitope prediction algorithm for antibody production in chickens. Three peptides were selected, one each specific to either VP4-RK (residues 494-505) or VP4-AA (residues 720-733) and one common to both VP4-AA and VP4-RK (residues 584-596).

The anti-VP4 peptide antibodies allowed differentiation of the resulting cleavage products from the processing of the expressed polyprotein and VP4 constructs. The processing is

currently understood to initially occur at ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} to release pVP2, VP4 and VP3 (Sanchez and Rodriguez, 1999; Lejal et al., 2000). Precursor VP2 is autocatalytically processed at ^{441}Ala-Phe^{442} site to release VP2 and pep70. The processing of pep70 would then yield the four peptides, pep46, pep 11, pep7a and pep7b (Da Costa *et al.*, 2002; Irigoyen *et al.*, 2009). The observed cleavage products as identified by the anti-VP4 peptide antibodies, were not consistent with the expected products, should the cleavage occur at ^{512}Ala-Ala^{513} and ^{755}Ala-Ala^{756} for the processing of the pVP2-VP4 and VP4-VP3 junctions by VP4 protease, respectively. However it is worth mentioning that extra protein bands were also observed in some of the previous studies, but were not accounted for (Kibenge *et al.*, 1997).

The observed inconsistency suggested the use of other cleavage sites. Sanchez and Rodriguez (1999) showed that deletion of ^{512}Ala-Ala^{513} did not completely inhibit the processing at the pVP2-VP4 junction, thus indicating that VP4 may use other cleavage sites. There are three potential cleavage sites in the pep70 sequence, thus the unresolved question was which of these sites was used. Both VP4-RA (Arg^{453}-Ala^{755}) and VP4-RK (Arg^{453}-Lys^{722}) contained the three N-terminal cleavage sites, ^{487}Ala-Ala^{488}, ^{494}Ala-Ala^{495}, ^{501}Ala-Ala^{502} and ^{512}Ala-Ala^{513}. The anti-VP4 peptide antibodies revealed that cleavage occurred at the ^{487}Ala-Ala^{488} site. However, N-terminal sequencing of the major cleavage product corresponding to the VP4 region showed that the cleavage occurred at the ^{512}Ala-Ala^{513} site. Further investigation was therefore required to clarify these conflicting results.

Before further investigation, it was necessary to confirm that the observed processing was due to the autocatalytic activity of the embedded VP4 protease. The VP4 Ser/Lys catalytic dyad and the cleavage sites are also characteristic of the Lon proteases, Lex A and leader proteases of *E. coli* (Paetzel and Dalbey, 1997), therefore it could not be ruled out that these bacterial proteases could have cleaved the expressed proteins. Using site-directed mutagenesis, it was previously shown that the replacement of Ser^{652} (by alanine or threonine) and Lys^{692} (by alanine, arginine or histidine) completely inactivated polyprotein processing (Lejal et al., 2000). In order to confirm that the cleavage products observed were a result of processing by the VP4 protease alone, and not by the *E. coli* proteases, the Ser/Lys catalytic dyad was mutated in all the constructs that were shown to be processed during expression i.e. full-length polyprotein, truncated polyprotein, VP4-RA, VP4-RK and VP4-AW. The expression of all the mutant constructs resulted in a single band corresponding to uncleaved intact fusion proteins as opposed to the cleavage products observed in the expression of the wild-type constructs. This confirmed that the processing was indeed due to the proteolytic processing mediated by the VP4 protease.

In order to clarify the position of the cleavage site used in the processing of the pVP2-VP4 junction, further constructs were designed i.e. VP4-R^1W (Arg^{453}-Trp^{891}), VP4-R^2W (Arg^{505}-Trp^{891}) and VP4-R^2A (Arg^{505}-Ala^{755}). Arg-Ile dipeptides (residues 453-454 and 505-506) were present at the N-terminus of each of the two cleavage sites of interest, ^{487}Ala-Ala^{488}, and ^{512}Ala-Ala^{513}. Arg^{453} is one of the dibasic residues suggested by Hudson *et al.* (1986) and Jagadish *et al.* (1988) to be cleavage sites. The design of these constructs was motivated by the presence of these two Arg-Ile dipeptides: the first Arg-Ile (residues 453-454) is about 60 residues away from the ^{512}Ala-Ala^{513} cleavage site, whereas the second (residues 505-506) is six residues away.

The VP1-2A junction involved in the maturation of immature HAV particles comprises the Arg-Ile dipeptide. When Arg was mutated into Met or Glu, no assembly was observed (Morace *et al.*, 2008). In the IBDV polyprotein, these Arg-Ile peptides are present within pep46 (Phe^{442}-Ala^{487}) and pep11 (Ala^{502}-Ala^{512}) which were shown to be important for VLP assembly and capsid morphogenesis, respectively (Chevalier *et al.*, 2005). It was therefore speculated in the present study that the ^{453}Arg-Ile^{454} dipeptide may be important for the cleavage at ^{487}Ala-Ala^{488} while ^{505}Arg-Ile^{506} is important for cleavage at the ^{512}Ala-Ala^{513} sequence.

The constructs were analysed along with VP4-R^1A (also called VP4-RA, Arg^{453}-Ala^{755}) cloned and expressed in the present study. Although the fragments starting from the first Arg-Ile dipeptide (VP4-R^1W and VP4-R^1A) showed the same autocatalytic efficiency as the fragments starting from the second dipeptide (VP4-R^2W and VP4-R^2A), the use of anti-peptide antibodies to characterise the processing revealed more information than previously known. Firstly, the whole pep70 is not required for the autocatalytic activity of VP4. The six residues (Arg^{505}-Ala^{512}) immediately N-terminal to the cleavage site were as efficient as the 60 residues (Arg^{453}-Ala^{512}) in the self-cleavage of VP4 at both junctions. Secondly, the cleavage occurred at both ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} cleavage sites. Thirdly, the observations made suggested that the cleavage starts at the C-terminal cleavage site for complete processing to occur.

More processing products were detected in the designed constructs using the anti-VP4 peptide antibodies, especially for the polyprotein, VP4-R^1W and VP4-R^1A (Table 6.1). This suggested a different processing strategy from what is currently known. Taking all these observations together, a new processing strategy is hereby proposed (Fig. 6.1). The model of the processing shown in Fig. 6.1 is a tentative explanation of how the observed cleavage products are released from the polyprotein. The explanation of the proposed

mechanism is based on the observations made from the expression of the other constructs listed in Table 6.1.

Table 6.1 Summary of the observed cleavage products as compared to the expected ones

pET-32b construct and expected size	Approximate sizes of observed cleavage products as recognised by antibodies			
	Anti-His tag	**anti-VP4-RK**	**Anti-VP4 common**	**Anti-VP4-AA**
Polyprotein (Met^{1}-Glu^{1012}) (130 kDa)				
Expected	Trx-pVP2 (70 kDa) VP3-His (32 kDa)	Trx-pVP2 (70 kDa)	VP4-VP3 (60 kDa) VP4-AA (28 kDa)	VP4-VP3 (60 kDa) VP4-AA (28 kDa)
Observed	70 and 32 kDa	70 and 24 kDa	60, 45 and 28 kDa	60, 45 and 28 kDa
VP4-R^{1}W (Arg^{453}-Trp^{891}) (67 kDa)				
Expected	Trx-pep60 (24 kDa) [a] ΔVP3-His (16 kDa)	Trx-pep60 (24 kDa) [a]	VP4-AA (28 kDa)	VP4-AA (28 kDa)
Observed	24, 22 and 16 kDa	24 kDa	27 kDa	24 kDa
VP4-R^{1}A (Arg^{453}-Ala^{755}) (52 kDa)				
Expected	Trx-pep60 (24 kDa) [a]	Trx-pep60 (24 kDa) [a]	VP4-AA (28 kDa)	VP4-AA (28 kDa)
Observed	22 and 24 kDa	24 kDa	28 kDa	24 kDa
VP4-AW (Ala^{513}-Trp^{891}) (60 kDa)				
Expected	Trx-VP4-AW (60 kDa) ΔVP3-His (16 kDa) Trx-VP4-AA (45 kDa)	-	Trx-VP4-AW (60 kDa) Trx-VP4-AA (45 kDa)	Trx-VP4-AW (60 kDa) Trx-VP4-AA (45 kDa)
Observed	60 and 45 kDa	-	60 and 45 kDa	60 and 45 kDa
VP4-R^{2}W VP4-R^{1}W (Arg^{505}-Trp^{891}) (62 kDa)				
Expected	Trx-pep7b (18 kDa) ΔVP3-His (16 kDa)	-	VP4-AA (28 kDa)	VP4-AA (28 kDa)
Observed	18 and 16 kDa	-	28 kDa	28 kDa
VP4-R^{2}A (Arg^{505}-Ala^{755}) (47 kDa)				
Expected	Trx-pep7b (18 kDa)	-	VP4-AA (28 kDa)	VP4-AA (28 kDa)
Observed	18 kDa	-	28 kDa	28 kDa
VP4-AA (Ala^{513}-Ala^{755}) (46 kDa)				
Expected	Trx-VP4-AA (46 kDa)	-	Trx-VP4-AA (46 kDa)	Trx-VP4-AA (46 kDa)
Observed	46 kDa	-	46 kDa	46 kDa

[a] Pep60 includes amino acid residues 453-512

It seems that the expressed polyprotein is able to utilise alternative processing sites so that some of the polyprotein is cleaved at the ^{487}Ala-Ala^{488} site while some is processed at the ^{512}Ala-Ala^{513} site. The site preference for this initial cleavage seems to determine the sites of subsequent cleavages. This was based on the observations made from the cleavage products resulting from the expression of VP4-R^{1}W and VP4-R^{1}A. The VP4-R^{1}W construct comprised an N-terminal Trx-His tag and a C-terminally 6xHis tagged Arg^{453}-Trp^{891} sequence, while the VP4-R^{1}A construct consisted of an N-terminal Trx-His tag and a C-terminally 6xHis tagged Arg^{453}-Ala^{755} polypeptide.

With regards to the current processing strategy, cleavage occurring at the ^{512}Ala-Ala^{513} site within the pVP2-VP4 junction should release a 24 kDa N-terminally Trx-His tagged Arg^{453}-Ala^{512} polypeptide for both VP4-R^{1}W and VP4-R^{1}A. Cleavage occurring at the ^{487}Ala-Ala^{488} site within the pVP2-VP4 junction should release a 22 kDa N-terminally Trx-His tagged Arg^{453}-Ala^{487} polypeptide. However, the anti-His tag antibody recognised a doublet of 22 and 24 kDa, corresponding to two N-terminally Trx-His tagged Arg^{453}-Ala^{487} and Arg^{453}-Ala^{512} polypeptides, respectively. The 24 kDa protein was in turn identified by the anti-VP4-RK peptide antibodies, raised against residues 494-505 (Table 6.1). The doublet implied that the processing was occurring in one site or the other. This also supports the observation made by Sanchez and Rodriguez (1999) that deletion of ^{512}Ala-Ala^{513} did not completely inhibit the processing at the pVP2-VP4 junction.

The processing of some of the expressed polyprotein construct at the ^{512}Ala-Ala^{513} site would release an N-terminally Trx-His tagged pVP2 (recognised by both anti-His tag and anti-VP4-RK peptide antibodies at 70 kDa, Table 6.1) and a C-terminally 6xHis-tagged VP4-VP3 (recognised by both anti-VP4 common and anti-VP4-AA peptide antibodies at 60 kDa, Table 6.1; Fig. 6.1). An unexpected product with a size of 24 kDa, recognised by anti-VP4-RK peptide antibodies was observed from the processing of the polyprotein (Table 6.1). A similar product at about 24 kDa was observed and recognised by anti-IBDV serum in the study by Kibenge *et al.* (1997), but no further details were given about the protein. The 24 kDa product could only result from a cleavage occurring within the VP2 region of the N-terminally Trx-His tagged pVP2. Analysis of the pVP2 sequence showed the presence of a ^{275}Ala-Val-Ala↓Ala^{278} sequence, which has the IBDV VP4 specific cleavage site motif (Ala-X-Ala↓Ala). Cleavage of the Trx-pVP2 at this site would therefore release a 24 kDa product, which would be recognised by the anti-VP4-RK peptide antibodies (Fig. 6.1).

The 60 kDa C-terminally 6xHis-tagged VP4-VP3 resulting from polyprotein processing pathway 1, should behave similar to VP4-AW (Ala^{513}-Trp^{891}). The expression of VP4-AW produced an unprocessed fusion protein at 60 kDa and a Trx-His tagged VP4-AA product at 45 kDa resulting from the processing at the VP4-ΔVP3 junction (Table 6.1). The presence of the unprocessed protein indicated that the processing at this junction occurred in *trans*. This implies that the C-terminally 6xHis-tagged VP4-VP3 would undergo the same slow processing in *trans*. An unexpected 45 kDa product was also released from the processing of the polyprotein construct which was recognised by both anti-VP4 common and anti-VP4-AA peptide antibodies (Table 6.1). A product at about 45 kDa was also observed and recognised by anti-IBDV serum in the study by Kibenge *et al.* (1997),

once again no further details were given. A product of this size which is recognised by the anti-VP4 common and VP4-AA peptide antibodies is only possible with the cleavage of the C-terminally 6xHis-tagged VP4-VP3 within the VP3 region. The VP3 sequence also has a potential VP4 cleavage site, ^{922}Leu-Arg-Ala↓Ala^{925}. Therefore it can be proposed that the C-terminally 6xHis-tagged VP4-VP3 can be cleaved at the ^{924}Ala-Ala^{925} site to release the 45 kDa VP4-ΔVP3, which can be further cleaved in *trans* at the ^{755}Ala-Ala^{756} site to release VP4 (recognised by both anti-VP4 common and anti-VP4-AA peptide antibodies at 28 kDa, Table 6.1).

The processing of some of the expressed polyprotein construct at the alternative ^{487}Ala-Ala^{488} site (pathway 2, Fig. 6.1) would release an N-terminally Trx-His tagged VP2 containing pep46 and a C-terminally 6xHis tagged pep7a-pep7b-pep11-VP4-VP3 (Fig. 6.1). The 60 kDa N-terminally Trx-His tagged VP2 containing pep46 should theoretically be recognised by the anti-His tag antibody, but it was not. The C-terminally 6xHis tagged pep7a-pep7b-pep11-VP4-VP3 should behave similar to VP4-R^{2}W (Arg^{505}-Trp^{891}). VP4-R^{2}W was autocatalytic and was completely processed. For complete processing to occur, cleavage should start at the C-terminus. For example, cleavage at the VP4-VP3 junction of VP4-R^{2}W construct would release an N-terminally Trx-His tagged VP4-R^{2}A (Arg^{505}-Ala^{755}), which retains autocatalytic activity to cleave the ^{512}Ala-Ala^{513} site to further release VP4 (recognised by both anti-VP4 common and anti-VP4-AA peptide antibodies at 28 kDa, Table 6.1) and an N-terminally Trx-His tagged Arg^{505}-Ala^{512} peptide (recognised by the anti-His tag at 18 kDa, Table 6.1). In contrast, initial cleavage at the N-terminal site (pVP2-VP4 junction) would release VP4-AW (Ala^{513}-Trp^{891}), which was shown to lose the autocatalytic activity and has to be processed in *trans*.

Therefore, the C-terminally 6xHis tagged pep7a-pep7b-pep11-VP4-VP3 would be cleaved at the VP4-VP3 junction to release a C-terminally 6xHis tagged VP3 (recognised by the anti-His tag antibody at 32 kDa, Table 6.1) (Fig. 6.1) and pep7a-pep7b-pep11-VP4 which can be further processed in *cis* at ^{512}Ala-Ala^{513} to release VP4 (recognised by both anti-VP4 common and anti-VP4-AA peptide antibodies at 28 kDa, Table 6.1). Pep25 may then be processed by the 28 kDa mature VP4, to release pep7a, pep7b and pep11, as was shown by Lejal *et al.* (2000) that the mature VP4 can cleave in *trans*. Lastly, the N-terminally Trx-His tagged VP2 containing pep46 (recognised by anti-His tag at 68-70 kDa, Table 6.1) should process itself at the ^{441}Ala-Phe^{442} through its endopeptidase activity, as was shown by Irigoyen *et al.* (2009).

A

```
MTNLQDQTQQ IVPFIRSLLM PTTGPASIPD DTLEKHTLRS ETSTYNLTVG DTGSGLIVFF PGFPGSIVGA 70
HYTLQSNGNY KFDQMLLTAQ NLPASYNYCR LVSRSLTVRS STLPGGVYAL NGTINAVTFQ GSLSELTDVS 140
YNGLMSATAN INDKIGNVLV GEGVTVLSLP TSYDLGYVRL GDPIPAIGLD PKMVATCDSS DRPRVYTITA 210
ADDYQFSSQY QSGGVTITLF SANIDAITSL SIGGELVFQT SVQGLILGAT IYLIGFDGTA VITRAVAADN 280
GLTAGTDNLM PFNIVIPISE ITQPITSIKL EIVTSKSGGQ AGDQMSWSAS GSLAVTIHGG NYPGALRPVT 350
LVAYERVATG SVVTVAGVSN FELIPNPELA KNLVTEYGRF DPGAMNYTKL ILSERDRLGI KTVWPTREYT 420
DFREYFMEVA DLNSPLKIAG AFGFKDIIRA LRRIAVPVVS TLFPPAAPLA HAIGEGVDYL LGDEAQAASG 490
TARAASGKAR AASGRIRQLT LAADKGYEVV ANLFQVPQNP VVDGILASPG ILRGAHNLDC VLREGATLFP 560
VVITTVEDAM TPKALNSKMF AVIEGVREDL QPPSQRGSFI RTLSGHRVYG YAPDGVLPLE TGRDYTVVPI 630
DDVWDDSIML SKDPIPPIVG NSGNLAIAYM DVFRPKVPIH VAMTGALNAY GEIENVSFRS TKLATAHRLG 700
LKLAGPGAFD VNTGSNWATF IKRFPHNPRD WDRLPYLNLP YLPPNAGRQY DLAMAASEFK ETPELESAVR 770
AMEAAANVDP LFQSALSVFM WLEENGIVTD MANFALSDPN AHRMRNFLAN APQAGSKSQR AKYGTAGYGV 840
EARGPTPEEA QREKDTRISK KMETMGIYFA TPEWVALNGH RGPSPGQLKY WQNTREIPDP NEDYLDYVHA 910
EKSRLASEEQ ILRAATSIYG APGQAEPPQA FIDEVAKVYE INHGRGPNQE QMKDLLLTAM EMKHRNPRRA 980
PPKPKPKPNV PTQRPPGRLG RWIRAVSDED LE        1012
```

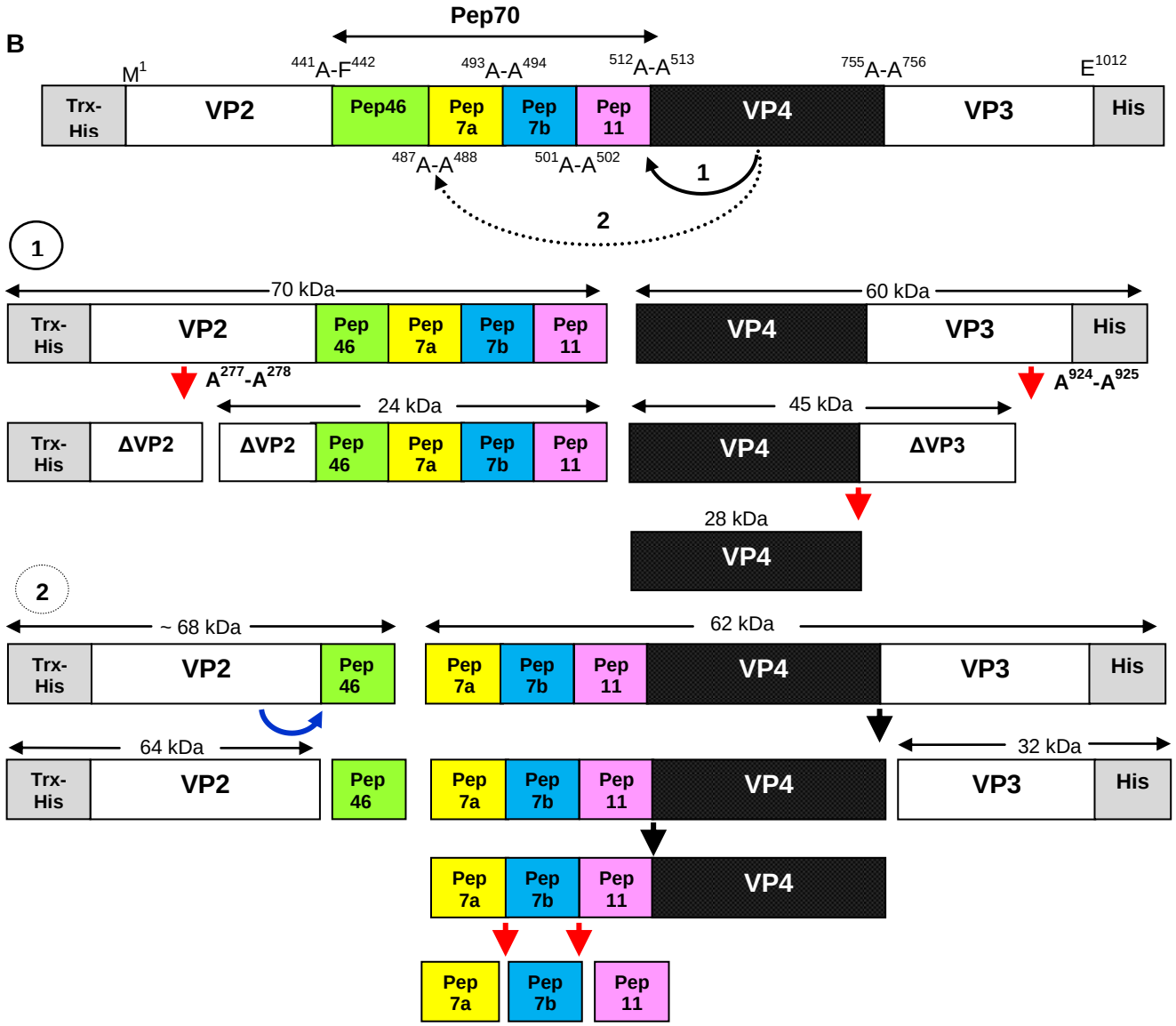


Figure 6.1 A proposed mechanism by which the IBDV polyprotein is processed in order to release the observed cleavage products. A) IBDV polyprotein sequence showing the previously identified cleavage sites (color junctions) and the additional two sites proposed in the present study (underlined and bold). **B)** The proposed polyprotein processing mechanism. Some of the expressed polyproteins are processed through pathway 1 where cleavage occurs at the ^{512}Ala-Ala^{513} site while some follows the second pathway where they are cleaved at ^{487}Ala-Ala^{488}. Processing at either site will determine how subsequent cleavages will occur. Sizes of the products listed in Table 6.1 are shown. The black arrows (▼) indicate autocatalytic proteolytic activities occurring in *cis*, the red arrows (▼) indicate cleavages mediated by VP4 in *trans* while the blue curved arrow (↷) indicate the processing carried out by the endopeptidase activity of VP2.

Therefore the use of specific anti-VP4 peptide antibodies was able to discriminate between the different products, which allowed the proposal of a modified strategy of polyprotein processing. This modified version is able to accommodate findings made by previous IBDV polyprotein investigators. This model may be verified through N-terminal sequencing of all the characterised products. The model of the polyprotein region without the tags is given in Appendix 4.

In the present study, the mutation of Pro^{377} was shown to alter the processing of the IBDV polyprotein. When the polyprotein construct containing the P377L mutation was expressed, a cleavage product of about 50 kDa recognised by all the anti-VP4 peptide antibodies was released. The release of this product is only possible if the cleavage occurs at ^{277}Ala-Ala^{278} and ^{755}Ala-Ala^{756} suggesting that the P377L mutation may prevent the processing at the ^{487}Ala-Ala^{488} and ^{512}Ala-Ala^{513} sites. However, this will also imply that the ^{277}Ala-Ala^{278} site is cleaved in *cis*. The functional significance of this residue is currently not known, but several studies showed the importance of proline residues in protein structure and function (Hu *et al.*, 1996; Ciani *et al.*, 2003; Bajaj *et al.*, 2007; Mazna *et al.*, 2008; Hughes *et al.*, 2009). Therefore, it is possible that Pro^{377} may influence the overall fold of the proteins within the polyprotein. Furthermore, the proline is part of a Pro-Asn-Pro motif that is not only conserved among all IBDV serotypes and strains, but also in all birnaviruses (IBDV, BSNV, IPNV and DXV) (Fig. 6.2). It was also very interesting that in the shown stretch of amino acids of the polyprotein sequence (Fig. 6.2), five other Pro residues are conserved among the four birnaviruses. The occurrence and conservation of this proline residue across the *Birnaviridae* polyprotein sequences seems to have a strong influence on the polyprotein folding and function.

```
IBDV        LVIPTNEITQ PITSIKLEIV TSKSGGQA-G DQMSWSARGS LAVTIHGGNY PGALRPVTLV  352
BSNV        AVFGETEITQ PVVGVQVLAK TN-------G DPIVVDS--Y VGVTVHGGNM PGTLRPVTII  328
DXV         SLFHENPPPE PVAAIKININ YGNN----TN GDSSFSVDSS FTINVIGGAT IGVNSPTVGV  342
IPNV        QSIPTEMITK PITRVKLAYQ LNQQTAIANA ATLGAKGPAS VSFSSGNGNV PGVLRPITLV  353
Clustal       :     .: *:. :::     .               .     . ..  .*    *.  * . :  117

IBDV        AYERVATGSV VTVAGVSNFE LIPNPELAKN LVTEYGRFDP GAMNYTKLIL SERDRLGIKT  412
BSNV        AYESVATGSV LTLSGISNYE LIPNPELAKN IQTSYGKLNP AEMTYTKVVL SHRDELGLRS  388
DXV         GYQGVAEGTA ITISGINNYE LVPNPDLQKN LPMTYGTCDP HDLTYIKYIL SNREQLGLRS  402
IPNV        AYEKMTPQSI LTVAGVSNYE LIPNPDLLKN MVTKYGKYDP EGLNYAKMIL SHREELDIRT  413
Clustal     .*: ::  :  :*::*:.*:* *:***:* ** :   **  :*   :.* * :* *.*:.*.:::  157

IBDV        VWPTREYTDF REYFMEVADL NSPLKIAGAF GFKDIIRAIR RIAVPVVSTL FPPAAPLAHA  472
BSNV        IWSIPQYRDM MSYFREVSDR SSPLKIAGAF GWGDLLSGIR KWVFPVVDTL LPAARPLTDL  448
DXV         VMTLADYNRM KMYMHVLTNY HVDEREASSF DFWQLLKQIK NVAVPLAATL APQFAPIIGA  462
IPNV        VWRTEEYKER TRAFKEITDF TSDLPTSKAW GWRDLVRGIR KVAAPVLSTL FPMAAPLIGA  473
Clustal     :    :*       :  :::        : :: .: :::  *: . . *:  **  *   *:     180
```

Figure 6.2 Sequence alignment of the polyprotein sequences from different birnaviruses showing the conserved proline residues (shown in red). The conserved PNP motif is boxed.

Even though no plausible explanation for this altered processing was demonstrated, a homology model built based on the 3-D structure of VP2 showed that the P377L mutation may cause a rearrangement of some secondary and tertiary structures within the VP2 domains. The findings strongly suggest the importance of this residue in polyprotein processing by possibly governing the overall conformation of the polyprotein, which in turn may determine the accessibility of the N-terminal cleavage sites of VP4. However, it still remains unclear how the mutation may influence the arrangement of the cleavage sites, since there is no published structure available that encompasses the whole polyprotein. To determine whether this mutation may change the polyprotein conformation, the mutation has to be introduced into the polyprotein which already contains the catalytic residue mutations. The resulting mutant polyprotein should therefore be expressed, purified and crystallised.

One of the objectives of the present study was also to characterise the embedded and mature forms of VP4. The embedded form was shown to have an autocatalytic activity that was completely abolished by mutating the catalytic dyad and was altered by the mutation of Pro^{377}. The mature form, on the other hand, has no autocatalytic activity. It was previously associated with the cleavage of the C-terminus of pVP2 (Kibenge et al., 1988). In a study by Lejal et al. (2000), VP4-AA with a Met residue added at the N-terminus was shown to cleave a mutant pET/VP2-VP4 fragment. In the present study, VP4-AA released from GST by thrombin was analysed for *trans* activity using a gelatine-substrate gel, fluorogenic substrates and mutant VP4-RK as a natural substrate.

This is the first study to analyse VP4-AA activity on a gelatine gel, and the observed results suggested VP4-AA to be active as a dimer. It only showed activity when released from GST but not as a fusion protein. GST exists as a dimer, and its fusion proteins also form dimers (Yamaguchi *et al.*, 2002), so GST may induce a different conformation. No activity was observed against fluorogenic substrates. The mutant VP4-RK, that mimicks a natural substrate, was cleaved by VP4-AA, thus supporting the observation that VP4-AA still retains activity after being released from the polyprotein following processing.

Based on the presence of vicinyl dialanines, the mature VP4 has been associated with the processing of the C-terminus of pVP2 (ASGTARA↓ASGKARA↓ASGRIRQLTLA) to release pep7a, pep7b and pep11 (Da Costa *et al.*, 2002). This sequence also contains three Ala-Gly-Ser tripeptides shown to be responsible for acetylthiocholinesterase activity (Jayanthi and Balasubramanian, 1992) and other additional enzyme activities such as aryl-acylamidase (AAA) and peptidase activities (Fujimoto, 1976; George and

Balasubramanian, 1980). It would be interesting to know whether these tripeptides play any role in the processing of this polypeptide.

The characterisation of the polyprotein using anti-VP4-peptide antibodies was conducted using recombinantly expressed polyprotein and VP4 protease. Additional studies on the native VP4 were also conducted. The first study used the chicken anti-VP4 peptide antibodies to detect VP4 in the infected samples by western blotting. Recently, Wang *et al.* (2009) detected VP4 in the cortex of lymphoid follicles of the bursa of Fabricius infected with a pathogenic IBDV strain using monoclonal antibodies, as was previously shown that VP4 is detected in infected cells (Granzow *et al.*, 1997). The anti-VP4 peptide antibodies detected some proteins in the IBDV-infected samples that were not present in the non-infected sample. Even though one of the proteins detected by anti-VP4 common peptide antibodies was 28 kDa in size, it could not be verified that it was VP4 in the absence of N-terminal sequencing data.

The second part of the study reported on the proteolytic activity of native VP4 in the IBDV infected samples by gelatine-gel SDS-PAGE analysis. The 28 kDa recombinant VP4 protease digested gelatine at about 56 kDa. No proteolytic activity was observed at 28 or 56 kDa for the infected bursal samples. Interestingly, there was a difference between the proteolytic profile of the non-infected and infected bursal homogenates. IBDV infection was shown to cause changes in cellular processes (Repp *et al.*, 1998; Zheng *et al.*, 2008), but the effect of IBDV on cellular proteases had not yet been studied. The present study also characterised protease activity in the infected and non-infected bursal samples. There was no difference in the activity of cysteine proteases whereas major differences were observed in the activity profiles of aspartic and serine proteases. It still remains to be confirmed whether the observed difference in the protease activity between infected and non-infected bursal homogenates is a biological effect of IBDV or not. Host cellular proteases have been shown to play a role in polyprotein processing, assembly and entry of some picornaviruses and influenza virus (Belouzard *et al.*, 2009; Chaipan *et al.*, 2009; Belouzard *et al.*, 2010). For IBDV, involvement of cellular proteases has not yet been shown for its replication, however the data from this study can be a platform for further studies.

The present study provided new insights into the processing of the IBDV polyprotein. A detailed model of the modified mechanism of IBDV polyprotein processing was deduced from the data obtained in the study. This model gives a better understanding of the

processing of the IBDV polyprotein, which may aid in the discovery of therapeutic agents. It was also shown that the embedded and mature forms of VP4 behave differently when expressed in *E. coli*. While the embedded form displayed a *cis* proteolytic activity, the mature form displayed a *trans* activity. It was also shown for the first time that mature VP4 is active on a gelatine gel. Even though VP4 from other birnaviruses (BSNV, IPNV and TV-1) have been crystallised, there is no structural information on the IBDV VP4 protease. Solving the structure of the embedded form of VP4 as a polyprotein and of the mature VP4 is required to provide insight into structural aspects of the processing. However, due to the autocatalytic activity observed, the polyprotein will have to be expressed and crystallised as an inactive mutant. Site-directed mutagenesis studies showed that the substitution of a proline that is conserved among birnaviruses may play a role in the processing. There has been no special mention of this proline before. There is still a need to understand the functional implications of the P377L mutation, and this may be accomplished by N-terminal sequencing of the products observed from the processing of the mutant polyprotein.

Furthermore, it was shown that the non-infected and infected bursal homogenates have different proteolytic activity profiles. This is of interest since IBDV infection was shown in previous studies to affect cellular proteases. Even though it was not established in the present study whether the differences were due to the virus infection, these findings lay the groundwork for an understanding of whether IBDV may have any effect on cellular protease activity. In order to verify this, IBDV could be replicated *in vitro* in cell culture. The infected cells and the control cells could then be analysed by gelatine-SDS-PAGE to determine whether there will be differences in the proteolytic activity.

The findings of the study have provided new information that will contribute to a better understanding of the polyprotein processing of IBDV. Polyprotein processing is crucial for the assembly of IBDV infectious particles. Hence, the findings may have applications in the development of anti-viral agents that may target the polyprotein processing. Inhibition of polyprotein processing could prevent IBDV capsid assembly, thus interfering with virus replication during infection.

References

Acil, Y., Brinckmann, J., Behrens, P., Muller, P. K. and Batge, B. (1997). Semipreparative isolation of collagen types I, II, III and V by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and electroelution. *J Chromatogr A* **758**, 313-318.

Adams, D. A., Edwards, R. J., Davies, D. S. and Boobis, A. R. (1997). Specific inhibition of human CYP1A2 using a targeted antibody. *Biochem Pharmacol* **54**, 189-197.

Adams, M. J., Antoniw, J. F. and Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Mol Plant Pathol* **6**, 471-487.

Akin, A., Wu, C. C. and Lin, T. L. (1998). A comparison of two RNA isolation methods for double-stranded RNA of infectious bursal disease virus. *J Virol Methods* **74**, 179-184.

Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J. & Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *Embo J* **21**, 3213-3224.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. and Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763-1767.

Anindya, R. and Savithri, H. S. (2004). Potyviral NIa proteinase, a proteinase with novel deoxyribonuclease activity. *J Biol Chem* **279**, 32159-32169.

Arlaud, G. J. and Gagnon, J. (1985). Identification of the peptide bond cleaved during activation of human C1r. *FEBS Lett* **180**, 234-238.

Arlaud, G. J., Gaboriaud, C., Thielens, N. M., Budayova-Spano, M., Rossi, V. and Fontecilla-Camps, J. C. (2002). Structural biology of the C1 complex of complement unveils the mechanisms of its activation and proteolytic activity. *Mol Immunol* **39**, 383-394.

Auld, D. S. (2004). Catalytic mechanisms of metallopeptidases. In Handbook of Proteolytic Enzymes. (ed. Barrett, A. J., Rawlings, N. D. and Woessner, J. F.). Elsevier, London. pp 268-289.

Ayra-Pardo, C. C., Martinez, G. and de la Riva, G. A. (1998). A single step screening procedure for Pichia pastoris clones by PCR. . *Biotechnologia Aplicada* **15**, 173-175.

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.

Babe, L. M. and Craik, C. S. (1997). Viral proteases: evolution of diverse structural motifs to optimize function. *Cell* **91**, 427-430.

Baida, R. C., Santos, M. R., Carmo, M. S., Yoshida, N., Ferreira, D., Ferreira, A. T., El Sayed, N. M., Andersson, B. and da Silveira, J. F. (2006). Molecular characterization of serine-, alanine-, and proline-rich proteins of Trypanosoma cruzi and their possible role in host cell infection. *Infect Immun* **74**, 1537-1546.

Bajaj, K., Madhusudhan, M. S., Adkar, B. V., Chakrabarti, P., Ramakrishnan, C., Sali, A. and Varadarajan, R. (2007). Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput Biol* **3**, 2465-2475.

Barbato, G., Cicero, D. O., Nardi, M. C., Steinkuhler, C., Cortese, R., De Francesco, R. and Bazzo, R. (1999). The solution structure of the N-terminal proteinase domain of the hepatitis C virus (HCV) NS3 protein provides new insights into its activation and catalytic mechanism. *J Mol Biol* **289**, 371-384.

Barrette-Ng, I. H., Ng, K. K., Mark, B. L., Van Aken, D., Cherney, M. M., Garen, C., Kolodenko, Y., Gorbalenya, A. E., Snijder, E. J. and James, M. N. (2002). Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. *J Biol Chem* **277**, 39960-39966.

Bartenschlager, R., Ahlborn-Laake, L., Yasargil, K., Mous, J. and Jacobsen, H. (1995). Substrate determinants for cleavage in cis and in trans by the hepatitis C virus NS3 proteinase. *J Virol* **69**, 198-205.

Bastos, I. M., Motta, F. N., Charneau, S., Santana, J. M., Dubost, L., Augustyns, K. and Grellier, P. (2010). Prolyl oligopeptidase of Trypanosoma brucei hydrolyzes native collagen, peptide hormones and is active in the plasma of infected mice. *Microbes Infect* **12**, 457-466.

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.

Belouzard, S., Chu, V. C. and Whittaker, G. R. (2009). Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A* **106**, 5871-5876.

Belouzard, S., Madu, I. and Whittaker, G. R. (2010). Elastase-mediated activation of the severe acute respiratory syndrome coronavirus spike protein at discrete sites within the S2 domain. *J Biol Chem* **285**, 22758-22763.

Bergmann, E. M., Mosimann, S. C., Chernaia, M. M., Malcolm, B. A. and James, M. N. (1997). The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. *J Virol* **71**, 2436-2448.

Bessaud, M., Pastorino, B. A., Peyrefitte, C. N., Rolland, D., Grandadam, M. and Tolou, H. J. (2006). Functional characterization of the NS2B/NS3 protease complex from seven viruses belonging to different groups inside the genus Flavivirus. *Virus Res* **120**, 79-90.

Birghan, C., Mundt, E. and Gorbalenya, A. E. (2000). A non-canonical lon proteinase lacking the ATPase domain employs the ser-Lys catalytic dyad to exercise broad control over the life cycle of a double-stranded RNA virus. *Embo J* **19**, 114-123.

Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J. and Curry, S. (2005). Crystal structure of foot-and-mouth disease virus 3C protease. New insights into catalytic mechanism and cleavage specificity. *J Biol Chem* **280**, 11520-11527.

Blevins, R. A. and Tulinsky, A. (1985). The refinement and the structure of the dimer of alpha-chymotrypsin at 1.67-A resolution. *J Biol Chem* **260**, 4264-4275.

Bolognesi, M., Spallarossa, A. and Ascenzi, P. (2001). Human rhinovirus 3C protease: a cysteine protease showing the trypsin(ogen)-like fold. *Biochemistry and Molecular Biology Education* **29**, 169-172.

Boot, H. J., ter Huurne, A. A., Hoekman, A. J., Peeters, B. P. and Gielkens, A. L. (2000a). Rescue of very virulent and mosaic infectious bursal disease virus from cloned cDNA: VP2 is not the sole determinant of the very virulent phenotype. *J Virol* **74**, 6701-6711.

Boot, H. J., ter Huurne, A. H. and Peeters, B. P. (2000b). Generation of full-length cDNA of the two genomic dsRNA segments of infectious bursal disease virus. *J Virol Methods* **84**, 49-58.

Bothner, B., Taylor, D., Jun, B., Lee, K. K., Siuzdak, G., Schultz, C. P. and Johnson, J. E. (2005). Maturation of a tetravirus capsid alters the dynamic properties and creates a metastable complex. *Virology* **334**, 17-27.

Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. and Gustchina, A. (2004). The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**, 8140-8148.

Bottcher, B., Kiselev, N. A., Stel'Mashchuk, V. Y., Perevozchikova, N. A., Borisov, A. V. and Crowther, R. A. (1997). Three-dimensional structure of infectious bursal disease virus determined by electron cryomicroscopy. *J Virol* **71**, 325-330.

Bouffard, P., Bartenschlager, R., Ahlborn-Laake, L., Mous, J., Roberts, N. and Jacobsen, H. (1995). An in vitro assay for hepatitis C virus NS3 serine proteinase. *Virology* **209**, 52-59.

Brandt, M., Yao, K., Liu, M., Heckert, R. A. and Vakharia, V. N. (2001). Molecular determinants of virulence, cell tropism, and pathogenic phenotype of infectious bursal disease virus. *J Virol* **75**, 11974-11982.

Bransom, K. L., Wallace, S. E. and Dreher, T. W. (1996). Identification of the cleavage site recognized by the turnip yellow mosaic virus protease. *Virology* **217**, 404-406.

Briand, J. P., Muller, S. and Van Regenmortel, M. H. (1985). Synthetic peptides as antigens: pitfalls of conjugation methods. *J Immunol Methods* **78**, 59-69.

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.

Buisson, M., Hernandez, J. F., Lascoux, D., Schoehn, G., Forest, E., Arlaud, G., Seigneurin, J. M., Ruigrok, R. W. and Burmeister, W. P. (2002). The crystal structure of the Epstein-Barr virus protease shows rearrangement of the processed C terminus. *J Mol Biol* **324**, 89-103.

Burgess, R. R. (2009). Refolding solubilized inclusion body proteins. *Methods Enzymol* **463**, 259-282.

Byrd, C. M., Bolken, T. a. and Hruby, D. E. (2004). Vaccinia Virus proteinase. In Handbook of Proteolytic Enzymes. (ed. Barrett, A. J., Rawlings, N. D. and Woessner, J. F.). Elsevier, London. pp 894-896.

Cármenes, R. S., Freije, J. P., Molina, M. M. and Martin, J. M. (1989). Predict7, a program for protein structure prediction. *Biochem Biophys Res Commun* **159**, 687-693.

Casanas, A., Navarro, A., Ferrer-Orta, C., Gonzalez, D., Rodriguez, J. F. and Verdaguer, N. (2008). Structural insights into the multifunctional protein VP3 of birnaviruses. *Structure* **16**, 29-37.

Caston, J. R., Martinez-Torrecuadrada, J. L., Maraver, A., Lombardo, E., Rodriguez, J. F., Casal, J. I. and Carrascosa, J. L. (2001). C terminus of infectious bursal disease virus major capsid protein VP2 is involved in definition of the T number for capsid assembly. *J Virol* **75**, 10815-10828.

Chaipan, C., Kobasa, D., Bertram, S., Glowacka, I., Steffen, I., Tsegaye, T. S., Takeda, M., Bugge, T. H., Kim, S., Park, Y., Marzi, A. and Pohlmann, S. (2009). Proteolytic activation of the 1918 influenza virus hemagglutinin. *J Virol* **83**, 3200-3211.

Chapnik, N., Sherman, H. and Froy, O. (2008). A one-tube site-directed mutagenesis method using PCR and primer extension. *Anal Biochem* **372**, 255-257.

Chappell, K. J., Stoermer, M. J., Fairlie, D. P. and Young, P. R. (2007). Generation and characterization of proteolytically active and highly stable truncated and full-length recombinant West Nile virus NS3. *Protein Expr Purif* **53**, 87-96.

Chen, C. S., Suen, S. Y., Lai, S. Y., Chang, G. R., Lu, T. C., Lee, M. S. and Wang, M. Y. (2005). Purification of capsid-like particles of infectious bursal disease virus (IBDV) VP2 expressed in E. coli with a metal-ion affinity membrane system. *J Virol Methods* **130**, 51-58.

Chettle, N., Stuart, J. C. and Wyeth, P. J. (1989). Outbreak of virulent infectious bursal disease in East Anglia. *Vet Rec* **125**, 271-272.

Chettle, N. J., Eddy, R. K., Saunders, J. and Wyeth, P. J. (1991). A comparison of serum neutralisation, immunofluorescence and immunoperoxidase tests for the detection of antibodies to chicken anaemia agent. *Vet Rec* **128**, 304-306.

Chevalier, C., Lepault, J., Erk, I., Da Costa, B. and Delmas, B. (2002). The maturation process of pVP2 requires assembly of infectious bursal disease virus capsids. *J Virol* **76**, 2384-2392.

Chevalier, C., Lepault, J., Da Costa, B. and Delmas, B. (2004). The last C-terminal residue of VP3, glutamic acid 257, controls capsid assembly of infectious bursal disease virus. *J Virol* **78**, 3296-3303.

Chevalier, C., Galloux, M., Pous, J., Henry, C., Denis, J., Da Costa, B., Navaza, J., Lepault, J. and Delmas, B. (2005). Structural peptides of a nonenveloped virus are involved in assembly and membrane translocation. *J Virol* **79**, 12253-12263.

Choi, G. H., Pawlyk, D. M. and Nuss, D. L. (1991). The autocatalytic protease p29 encoded by a hypovirulence-associated virus of the chestnut blight fungus resembles the potyvirus-encoded protease HC-Pro. *Virology* **183**, 747-752.

Choi, H. K., Lu, G., Lee, S., Wengler, G. and Rossmann, M. G. (1997). Structure of Semliki Forest virus core protein. *Proteins* **27**, 345-359.

Choi, S. Y., Bertram, S., Glowacka, I., Park, Y. W. and Pohlmann, S. (2009). Type II transmembrane serine proteases in cancer and viral infections. *Trends Mol Med* **15**, 303-312.

Chung, D. W., Fujikawa, K., McMullen, B. A. and Davie, E. W. (1986). Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. *Biochemistry (Mosc)* **25**, 2410-2417.

Chung, I. Y. and Paetzel, M. (2011a). Expression, purification and crystallization of VP4 protease from Tellina virus 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**, 157-160.

Chung, I. Y. and Paetzel, M. (2011b). Crystal structure of a viral protease intramolecular acyl-enzyme complex: insights into cis-cleavage at the VP4/VP3 junction of Tellina birnavirus. *J Biol Chem* **286**, 12475-12482.

Ciani, B., Layfield, R., Cavey, J. R., Sheppard, P. W. and Searle, M. S. (2003). Structure of the ubiquitin-associated domain of p62 (SQSTM1) and implications for mutations that cause Paget's disease of bone. *J Biol Chem* **278**, 37409-37412.

Coetzer, T., Elliot, E., Fortgens, P., Pike, R. and Dennison, C. (1991). Anti-peptide antibodies to cathepsins B, L and D and type IV collagenase: Specific recognition and inhibition of the enzymes. *J Immunol Methods* **136**, 199-210.

Coetzer, T. H., Dennehy, K. M., Pike, R. N. and Dennison, C. (1995). Baboon (Papio ursinus) cathepsin L: purification, characterization and comparison with human and sheep cathepsin L. *Comp Biochem Physiol B Biochem Mol Biol* **112**, 429-439.

Coetzer, T. H., Goldring, J. P. and Huson, L. E. (2008). Oligopeptidase B: a processing peptidase involved in pathogenesis. *Biochimie* **90**, 336-344.

Cohen, L., Benichou, D. and Martin, A. (2002). Analysis of deletion mutants indicates that the 2A polypeptide of hepatitis A virus participates in virion morphogenesis. *J Virol* **76**, 7495-7505.

Cosgrove, A. (1962). An apparently new disease of chickens: avian nephrosis. *Avian Dis* **6**, 385-389.

Coulibaly, F., Chevalier, C., Gutsche, I., Pous, J., Navaza, J., Bressanelli, S., Delmas, B. and Rey, F. A. (2005). The birnavirus crystal structure reveals structural relationships among icosahedral viruses. *Cell* **120**, 761-772.

Curry, S., Roque-Rosell, N., Zunszain, P. A. and Leatherbarrow, R. J. (2007). Foot-and-mouth disease virus 3C protease: recent structural and functional insights into an antiviral target. *Int J Biochem Cell Biol* **39**, 1-6.

Da Costa, B., Chevalier, C., Henry, C., Huet, J. C., Petit, S., Lepault, J., Boot, H. and Delmas, B. (2002). The capsid of infectious bursal disease virus contains several small peptides arising from the maturation process of pVP2. *J Virol* **76**, 2393-2402.

Da Costa, B., Soignier, S., Chevalier, C., Henry, C., Thory, C., Huet, J. C. and Delmas, B. (2003). Blotched snakehead virus is a new aquatic birnavirus that is slightly more related to avibirnavirus than to aquabirnavirus. *J Virol* **77**, 719-725.

Davison, A. J. (2007). Comparative analysis of the genomes. In Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. (ed. Arvin, A., Campadelli-Fiume, G., Mocarski, E., Moore, P. S., Roizman, B., Whitley, R. and Yamanishi, K.). Cambridge University Press, Cambridge. pp 10-26.

De Groot, A. S., Saint-Aubin, C., Bosma, A., Sbai, H., Rayner, J. and Martin, W. (2001). Rapid determination of HLA B*07 ligands from the West Nile virus NY99 genome. *Emerg Infect Dis* **7**, 706-713.

De, I., Fata-Hartley, C., Sawicki, S. G. and Sawicki, D. L. (2003). Functional analysis of nsP3 phosphoprotein mutants of Sindbis virus. *J Virol* **77**, 13106-13116.

de Vries, A., Horzinek, M., Rottier, P. and de Groot, R. (1997). The Genome Organization of the Nidovirales:Similarities and Differences between Arteri-, Toro-, and Coronaviruses. *Seminars in VIROLOGY* **8**, 33-47.

den Boon, J. A., Faaberg, K. S., Meulenberg, J. J., Wassenaar, A. L., Plagemann, P. G., Gorbalenya, A. E. and Snijder, E. J. (1995). Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. *J Virol* **69**, 4500-4505.

Deng, X., Gao, Y., Gao, H., Qi, X., Cheng, Y., Wang, X. and Wang, X. (2007). Antigenic structure analysis of VP3 of infectious bursal disease virus. *Virus Res* **129**, 35-42.

Dey, S., Upadhyay, C., Madhan Mohan, C., Kataria, J. M. and Vakharia, V. N. (2009). Formation of subviral particles of the capsid protein VP2 of infectious bursal disease virus and its application in serological diagnosis. *J Virol Methods* **157**, 84-89.

Diaz-Ruiz, J. R. and Kaper, J. M. (1978). Isolation of viral double-stranded RNAs using a LiCl fractionation procedure. *Prep Biochem* **8**, 1-17.

Ding, J., McGrath, W. J., Sweet, R. M. and Mangel, W. F. (1996). Crystal structure of the human adenovirus proteinase with its 11 amino acid cofactor. *Embo J* **15**, 1778-1783.

Dobo, J., Gal, P., Szilagyi, K., Cseh, S., Lorincz, Z., Schumaker, V. N. and Zavodszky, P. (1999). One active C1r subunit is sufficient for the activity of the complement C1 complex: stabilization of C1r in the zymogen form by point mutations. *J Immunol* **162**, 1108-1112.

Dobos, P., Hill, B. J., Hallett, R., Kells, D. T., 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.

Dobos, P. (2002). Birnaviruses. In Encyclopedia of Life Science 33. MacMillan Publishers, Nature publishing group, London. pp 289-283.

Domingo, E., Menendez-Arias, L., Quinones-Mateu, M. E., Holguin, A., Gutierrez-Rivas, M., Martinez, M. A., Quer, J., Novella, I. S. and Holland, J. J. (1997). Viral quasispecies and the problem of vaccine-escape and drug-resistant mutants. *Prog Drug Res* **48**, 99-128.

Duquerroy, S., Da Costa, B., Henry, C., Vigouroux, A., Libersou, S., Lepault, J., Navaza, J., Delmas, B. and Rey, F. A. (2009). The picobirnavirus crystal structure provides functional insights into virion assembly and cell entry. *Embo J* **28**, 1655-1665.

Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S. P., Yin, Z., Keller, T. H., Vasudevan, S. G. and Hommel, U. (2006). Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat Struct Mol Biol* **13**, 372-373.

Ertl, P., Russell, L. and Angier, J. (2000). Herpesvirus Protease Assays. *Methods Mol Med* **24**, 171-182.

Eterradossi, N., Arnauld, C., Toquin, D. and Rivallan, G. (1998). Critical amino acid changes in VP2 variable domain are associated with typical and atypical antigenicity in very virulent infectious bursal disease viruses. *Arch Virol* **143**, 1627-1636.

Feldman, A. R., Lee, J., Delmas, B. and Paetzel, M. (2006). Crystal structure of a novel viral protease with a serine/lysine catalytic dyad mechanism. *J Mol Biol* **358**, 1378-1389.

Fernandez-Arias, A., Martinez, S. and Rodriguez, J. F. (1997). The major antigenic protein of infectious bursal disease virus, VP2, is an apoptotic inducer. *J Virol* **71**, 8014-8018.

Fernandez-Arias, A., Risco, C., Martinez, S., Albar, J. P. and Rodriguez, J. F. (1998). Expression of ORF A1 of infectious bursal disease virus results in the formation of virus-like particles. *J Gen Virol* **79**, 1047-1054.

Fujimoto, D. (1976). Serotonin-sensitive aryl acylamidase activity of acetylcholinesterase. *FEBS Lett* **72**, 121-123.

Funk, P. E. and Thompson, C. B. (1998). Identification of a lectin that induces cell death in developing chicken B cells. *Cell Immunol* **186**, 75-81.

Galloux, M., Libersou, S., Morellet, N., Bouaziz, S., Da Costa, B., Ouldali, M., Lepault, J. and Delmas, B. (2007). Infectious bursal disease virus, a non-enveloped virus, possesses a capsid-associated peptide that deforms and perforates biological membranes. *J Biol Chem* **282**, 20774-20784.

Garriga, D., Querol-Audi, J., Abaitua, F., Saugar, I., Pous, J., Verdaguer, N., Caston, J. R. and Rodriguez, J. F. (2006). The 2.6-Angstrom structure of infectious bursal disease virus-derived T=1 particles reveals new stabilizing elements of the virus capsid. *J Virol* **80**, 6895-6905.

Gayathri, P., Satheshkumar, P. S., Prasad, K., Nair, S., Savithri, H. S. and Murthy, M. R. (2006). Crystal structure of the serine protease domain of Sesbania mosaic virus polyprotein and mutational analysis of residues forming the S1-binding pocket. *Virology* **346**, 440-451.

George, S. T. and Balasubramanian, A. S. (1980). The identity of the serotonin-sensitive aryl acylamidase with acetylcholinesterase from human erythrocytes, sheep basal ganglia and electric eel. *Eur J Biochem* **111**, 511-524.

Gibson, W. (2001). Action at the assemblin dimer interface. *Nat Struct Biol* **8**, 739-741.

Goh, L. L., Loke, P., Singh, M. and Sim, T. S. (2003). Soluble expression of a functionally active Plasmodium falciparum falcipain-2 fused to maltose-binding protein in Escherichia coli. *Protein Expr Purif* **32**, 194-201.

Goh, S. L., Goh, L. L. and Sim, T. S. (2005). Cysteine protease falcipain 1 in Plasmodium falciparum is biochemically distinct from its isozymes. *Parasitol Res* **97**, 295-301.

Goldring, J. and Coetzer, T. H. T. (2003). Isolation of chicken immunoglobulins (IgY) from egg yolk. *Biochem Mol Biol Edu* **31**, 185-187.

Golubtsov, A., Kaariainen, L. and Caldentey, J. (2006). Characterization of the cysteine protease domain of Semliki Forest virus replicase protein nsP2 by in vitro mutagenesis. *FEBS Lett* **580**, 1502-1508.

Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. and Snijder, E. J. (2006). Nidovirales: evolving the largest RNA virus genome. *Virus Res* **117**, 17-37.

Goutebroze, S., Curet, M., Jay, M. L., Roux, C. and Le Gros, F. X. (2003). Efficacy of a recombinant vaccine HVT-VP2 against Gumboro disease in the presence of maternal antibodies. *Br Poult Sci* **44**, 824-825.

Gradi, A., Foeger, N., Strong, R., Svitkin, Y. V., Sonenberg, N., Skern, T. and Belsham, G. J. (2004). Cleavage of eukaryotic translation initiation factor 4GII within foot-and-mouth disease virus-infected cells: identification of the L-protease cleavage site in vitro. *J Virol* **78**, 3271-3278.

Granzow, H., Birghan, C., Mettenleiter, T. C., Beyer, J., Kollner, B. and Mundt, E. (1997). A second form of infectious bursal disease virus-associated tubule contains VP4. *J Virol* **71**, 8879-8885.

Grubman, M. J., Zellner, M., Bablanian, G., Mason, P. W. and Piccone, M. E. (1995). Identification of the active-site residues of the 3C proteinase of foot-and-mouth disease virus. *Virology* **213**, 581-589.

Guarne, A., Tormo, J., Kirchweger, R., Pfistermueller, D., Fita, I. and Skern, T. (1998). Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition. *Embo J* **17**, 7469-7479.

Guarne, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I. and Skern, T. (2000). Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. *J Mol Biol* **302**, 1227-1240.

Guex, N. and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.

Guo, B., Lin, J. and Ye, K. (2011). Structure of the autocatalytic cysteine protease domain of potyvirus helper-component proteinase. *J Biol Chem* **286**, 21937-21943.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Sympo Ser* **41**, 95-98.

Hara, K., Shiota, M., Kido, H., Ohtsu, Y., Kashiwagi, T., Iwahashi, J., Hamada, N., Mizoue, K., Tsumura, N., Kato, H. and Toyoda, T. (2001a). Influenza virus RNA polymerase PA subunit is a novel serine protease with Ser624 at the active site. *Genes Cells* **6**, 87-97.

Hara, K., Shiota, M., Kido, H., Ohtsu, Y. and Toyoda, T. (2001b). Protease activity of influenza virus RNA polymerase PA subunit. *International Congress Series* **1219** 479–485.

Hedengren-Olcott, M., Byrd, C. M., Watson, J. and Hruby, D. E. (2004). The vaccinia virus G1L putative metalloproteinase is essential for viral replication in vivo. *J Virol* **78**, 9947-9953.

Hedstrom, L. (2002). Serine protease mechanism and specificity. *Chem Rev* **102**, 4501-4524.

Hehn, A., Fritsch, C., Richards, K. E., Guilley, H. and Jonard, G. (1997). Evidence for in vitro and in vivo autocatalytic processing of the primary translation product of beet necrotic yellow vein virus RNA 1 by a papain-like proteinase. *Arch Virol* **142**, 1051-1058.

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.

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. and Galas, D. J. (1989). A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res* **17**, 6545-6551.

Heussen, C. and Dowdle, E. B. (1980). Electrophoretic analysis of plasminogen activators in poly-acrylamide gels containing sodium dodecyl sulfate and copolymerised substrates. *Anal Biochem* **102**, 196-202.

Hill, M. K., Shehu-Xhilaga, M., Crowe, S. M. and Mak, J. (2002). Proline residues within spacer peptide p1 are important for human immunodeficiency virus type 1 infectivity, protein processing, and genomic RNA dimer stability. *J Virol* **76**, 11245-11253.

Hill, M. K., Bellamy-McIntyre, A., Vella, L. J., Campbell, S. M., Marshall, J. A., Tachedjian, G. and Mak, J. (2007). Alteration of the proline at position 7 of the HIV-1 spacer peptide p1 suppresses viral infectivity in a strain dependent manner. *Curr HIV Res* **5**, 69-78.

Hindiyeh, M., Li, Q. H., Basavappa, R., Hogle, J. M. and Chow, M. (1999). Poliovirus mutants at histidine 195 of VP2 do not cleave VP0 into VP2 and VP4. *J Virol* **73**, 9072-9079.

Hinton, T. M., Ross-Smith, N., Warner, S., Belsham, G. J. and Crabb, B. S. (2002). Conservation of L and 3C proteinase activities across distantly related aphthoviruses. *J Gen Virol* **83**, 3111-3121.

Hirani, V., Yarovoy, A., Kozeska, A., Magnusson, R. P. and Lasker, J. M. (2008). Expression of CYP4F2 in human liver and kidney: assessment using targeted peptide antibodies. *Arch Biochem Biophys* **478**, 59-68.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.

Horner, R., 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*. 312-5012.

Hu, M. C., Hsu, L. C., Hsu, N. C. and Chung, B. C. (1996). Function and membrane topology of wild-type and mutated cytochrome P-450c21. *Biochem J* **316** 325-329.

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.

Hughes, M., Gretton, S., Shelton, H., Brown, D. D., McCormick, C. J., Angus, A. G., Patel, A. H., Griffin, S. and Harris, M. (2009). A conserved proline between domains II and III of hepatitis C virus NS5A influences both RNA replication and virus assembly. *J Virol* **83**, 10788-10796.

Hurdayal, R., Achilonu, I., Choveaux, D., Coetzer, T. H. and Dean Goldring, J. P. (2010). Anti-peptide antibodies differentiate between plasmodial lactate dehydrogenases. *Peptides* **31**, 525-532.

Ingr, M., Uhlikova, T., Strisovsky, K., Majerova, E. and Konvalinka, J. (2003). Kinetics of the dimerization of retroviral proteases: the "fireman's grip" and dimerization. *Protein Sci* **12**, 2173-2182.

Irigoyen, N., Garriga, D., Navarro, A., Verdaguer, N., Rodriguez, J. F. and Caston, J. R. (2009). Autoproteolytic activity derived from the infectious bursal disease virus capsid protein. *J Biol Chem* **284**, 8064-8072.

Ismail, N. M. and Saif, Y. M. (1991). Immunogenicity of infectious bursal disease viruses in chickens. *Avian Dis* **35**, 460-469.

Izquierdo, C. and Burguillo, F. J. (1989). Synthetic substrates for thrombin. *Int J Biochem* **21**, 579-592.

Jackwood, D. J. and Sommer, S. E. (2002). Identification of infectious bursal disease virus quasispecies in commercial vaccines and field isolates of this double-stranded RNA virus. *Virology* **304**, 105-113.

Jackwood, D. J. and Sommer-Wagner, S. (2007). Genetic characteristics of infectious bursal disease viruses from four continents. *Virology* **365**, 369-375.

Jackwood, D. J., Sreedevi, B., LeFever, L. J. and Sommer-Wagner, S. E. (2008). Studies on naturally occurring infectious bursal disease viruses suggest that a single amino acid substitution at position 253 in VP2 increases pathogenicity. *Virology* **377**, 110-116.

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.

Jayanthi, L. D. and Balasubramanian, A. S. (1992). Isolation of a tripeptide (Ala-Gly-Ser) exhibiting weak acetylthiocholine hydrolyzing activity from a high-salt soluble form of monkey diaphragm acetylcholinesterase. *Neurochem Res* **17**, 351-359.

Jiang, B., Monroe, S. S., Koonin, E. V., Stine, S. E. and Glass, R. I. (1993). RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis. *Proc Natl Acad Sci U S A* **90**, 10539-10543.

Jungmann, A., Nieper, H. and Muller, H. (2001). Apoptosis is induced by infectious bursal disease virus replication in productively infected cells as well as in antigen-negative cells in their vicinity. *J Gen Virol* **82**, 1107-1115.

Kammann, M., Laufs, J., Schell, J. and Gronenborn, B. (1989). Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR). *Nucleic Acids Res* **17**, 5404.

Kangethe, R. T., Boulange, A. F., Coustou, V., Baltz, T. and Coetzer, T. H. (2012). Trypanosoma brucei brucei oligopeptidase B null mutants display increased prolyl oligopeptidase-like activity. *Mol Biochem Parasitol* **182**, 7-16.

Kaufer, I. and Weiss, E. (1980). Significance of bursa of Fabricius as target organ in infectious bursal disease of chickens. *Infect Immun* **27**, 364-367.

Khan, A. R., Khazanovich-Bernstein, N., Bergmann, E. M. and James, M. N. (1999). Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proc Natl Acad Sci U S A* **96**, 10968-10975.

Kibenge, F. S., Dhillon, A. S. and Russell, R. G. (1988). Biochemistry and immunology of infectious bursal disease virus. *J Gen Virol* **69** 1757-1775.

Kibenge, F. S., Qian, B., Cleghorn, J. R. and Martin, C. K. (1997). Infectious bursal disease virus polyprotein processing does not involve cellular proteases. *Arch Virol* **142**, 2401-2419.

Kim, I. J., You, S. K., Kim, H., Yeh, H. Y. and Sharma, J. M. (2000). Characteristics of bursal T lymphocytes induced by infectious bursal disease virus. *J Virol* **74**, 8884-8892.

Kwon, H. M., Kim, D. K., Hahn, T. W., Han, J. H. and Jackwood, D. J. (2000). Sequence of precursor polyprotein gene (segment A) of infectious bursal disease viruses isolated in Korea. *Avian Dis* **44**, 691-696.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lampisuo, M., Arstila, T. P., Liippo, J. and Lassila, O. (1998). Expression of chL12 surface antigen is associated with cell survival in the avian bursa of Fabricius. *Scand J Immunol* **47**, 223-228.

Lawrence, D. M., Rozanov, M. N. and Hillman, B. I. (1995). Autocatalytic processing of the 223-kDa protein of blueberry scorch carlavirus by a papain-like proteinase. *Virology* **207**, 127-135.

Lee, C. C., Ko, T. P., Chou, C. C., Yoshimura, M., Doong, S. R., Wang, M. Y. and Wang, A. H. (2006a). Crystal structure of infectious bursal disease virus VP2 subviral particle at 2.6A resolution: implications in virion assembly and immunogenicity. *J Struct Biol* **155**, 74-86.

Lee, J., Feldman, A. R., Delmas, B. and Paetzel, M. (2007). Crystal structure of the VP4 protease from infectious pancreatic necrosis virus reveals the acyl-enzyme complex for an intermolecular self-cleavage reaction. *J Biol Chem* **282**, 24928-24937.

Lee, K. J. and Watson, R. D. (2002). Antipeptide antibodies for detecting crab (Callinectes sapidus) molt-inhibiting hormone. *Peptides* **23**, 853-862.

Lee, M. S., Wang, M. Y., Tai, Y. J. and Lai, S. Y. (2004). Characterization of particles formed by the precursor protein VPX of infectious bursal disease virus in insect Hi-5 cells: implication on its proteolytic processing. *J Virol Methods* **121**, 191-199.

Lee, M. S., Doong, S. R., Lai, S. Y., Ho, J. Y. and Wang, M. Y. (2006b). Processing of infectious bursal disease virus (IBDV) polyprotein and self-assembly of IBDV-like particles in Hi-5 cells. *Biotechnol Prog* **22**, 763-769.

Lejal, N., Da Costa, B., Huet, J. C. and Delmas, B. (2000). Role of Ser-652 and Lys-692 in the protease activity of infectious bursal disease virus VP4 and identification of its substrate cleavage sites. *J Gen Virol* **81**, 983-992.

Lemke, C. T., Goudreau, N., Zhao, S., Hucke, O., Thibeault, D., Llinas-Brunet, M. and White, P. W. (2011). Combined X-ray, NMR, and kinetic analyses reveal uncommon binding characteristics of the hepatitis C virus NS3-NS4A protease inhibitor BI 201335. *J Biol Chem* **286**, 11434-11443.

Lemm, J. A., Rumenapf, T., Strauss, E. G., Strauss, J. H. and Rice, C. M. (1994). Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *Embo J* **13**, 2925-2934.

Lescar, J., Luo, D., Xu, T., Sampath, A., Lim, S. P., Canard, B. and Vasudevan, S. G. (2008). Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target. *Antiviral Res* **80**, 94-101.

Li, W., Ross-Smith, N., Proud, C. G. and Belsham, G. J. (2001). Cleavage of translation initiation factor 4AI (eIF4AI) but not eIF4AII by foot-and-mouth disease virus 3C protease: identification of the eIF4AI cleavage site. *FEBS Lett* **507**, 1-5.

Lim, B. L., Cao, Y., Yu, T. and Mo, C. W. (1999). Adaptation of very virulent infectious bursal disease virus to chicken embryonic fibroblasts by site-directed mutagenesis of residues 279 and 284 of viral coat protein VP2. *J Virol* **73**, 2854-2862.

Lim, K. P., Ng, L. F. and Liu, D. X. (2000). Identification of a novel cleavage activity of the first papain-like proteinase domain encoded by open reading frame 1a of the coronavirus Avian infectious bronchitis virus and characterization of the cleavage products. *J Virol* **74**, 1674-1685.

Ling, M. M. and Robinson, B. H. (1997). Approaches to DNA mutagenesis: an overview. *Anal Biochem* **254**, 157-178.

Liu, H. J., Giambrone, J. J. and Dormitorio, T. (1994). Detection of genetic variations in serotype I isolates of infectious bursal disease virus using polymerase chain reaction and restriction endonuclease analysis. *J Virol Methods* **48**, 281-291.

Liu, M. and Vakharia, V. (2006). Nonstructural Protein of Infectious Bursal Disease Virus Inhibits Apoptosis at the Early Stage of Virus Infection. *J Virol* **80**, 3369-3377.

Lombardo, E., Maraver, A., Caston, J. R., Rivera, J., Fernandez-Arias, A., Serrano, A., Carrascosa, J. L. and Rodriguez, J. F. (1999). VP1, the putative RNA-dependent RNA polymerase of infectious bursal disease virus, forms complexes with the capsid protein VP3, leading to efficient encapsidation into virus-like particles. *J Virol* **73**, 6973-6983.

Lombardo, E., Maraver, A., Espinosa, I., Fernandez-Arias, A. and Rodriguez, J. F. (2000). VP5, the nonstructural polypeptide of infectious bursal disease virus, accumulates within the host plasma membrane and induces cell lysis. *Virology* **277**, 345-357.

Luo, D., Xu, T., Hunke, C., Gruber, G., Vasudevan, S. G. and Lescar, J. (2008). Crystal structure of the NS3 protease-helicase from dengue virus. *J Virol* **82**, 173-183.

Luque, D., Saugar, I., Rodriguez, J. F., Verdaguer, N., Garriga, D., Martin, C. S., Velazquez-Muriel, J. A., Trus, B. L., Carrascosa, J. L. and Caston, J. R. (2007). Infectious bursal disease virus capsid assembly and maturation by structural rearrangements of a transient molecular switch. *J Virol* **81**, 6869-6878.

Luque, D., Saugar, I., Rejas, M. T., Carrascosa, J. L., Rodriguez, J. F. and Caston, J. R. (2009). Infectious Bursal disease virus: ribonucleoprotein complexes of a double-stranded RNA virus. *J Mol Biol* **386**, 891-901.

Maia, I. G., Haenni, A. and Bernardi, F. (1996). Potyviral HC-Pro: a multifunctional protein. *J Gen Virol* **77**, 1335-1341.

Manefeld, J. (2007). Metalloprotease. In Industrial Enzymes: Structure, functions and applications. (ed. Polaina, J. and MacCabe, AP). Springer, Dordrecht, Netherlands. pp 221-242.

Maraver, A., Clemente, R., Rodriguez, J. F. and Lombardo, E. (2003a). Identification and molecular characterization of the RNA polymerase-binding motif of infectious bursal disease virus inner capsid protein VP3. *J Virol* **77**, 2459-2468.

Maraver, A., Ona, A., Abaitua, F., Gonzalez, D., Clemente, R., Ruiz-Diaz, J. A., Caston, J. R., Pazos, F. and Rodriguez, J. F. (2003b). The oligomerization domain of VP3, the scaffolding protein of infectious bursal disease virus, plays a critical role in capsid assembly. *J Virol* **77**, 6438-6449.

Marcotte, L. L., Wass, A. B., Gohara, D. W., Pathak, H. B., Arnold, J. J., Filman, D. J., Cameron, C. E. and Hogle, J. M. (2007). Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. *J Virol* **81**, 3583-3596.

Martinez-Torrecuadrada, J. L., Caston, J. R., Castro, M., Carrascosa, J. L., Rodriguez, J. F. and Casal, J. I. (2000). Different architectures in the assembly of infectious bursal disease virus capsid proteins expressed in insect cells. *Virology* **278**, 322-331.

Matsuura, Y., Tohya, Y., Nakamura, K., Shimojima, M., Roerink, F., Mochizuki, M., Takase, K., Akashi, H. and Sugimura, T. (2002). Complete nucleotide sequence, genome organization and phylogenic analysis of the canine calicivirus. *Virus Genes* **25**, 67-73.

Mazna, P., Grycova, L., Balik, A., Zemkova, H., Friedlova, E., Obsilova, V., Obsil, T. and Teisinger, J. (2008). The role of proline residues in the structure and function of human MT2 melatonin receptor. *J Pineal Res* **45**, 361-372.

McGrath, W. J., Ding, J., Didwania, A., Sweet, R. M. and Mangel, W. F. (2003). Crystallographic structure at 1.6-A resolution of the human adenovirus proteinase in a covalent complex with its 11-amino-acid peptide cofactor: insights on a new fold. *Biochim Biophys Acta* **1648**, 1-11.

Mendez, E., Salas-Ocampo, M. P., Munguia, M. E. and Arias, C. F. (2003). Protein products of the open reading frames encoding nonstructural proteins of human astrovirus serotype 8. *J Virol* **77**, 11378-11384.

Merits, A., Vasiljeva, L., Ahola, T., Kaariainen, L. and Auvinen, P. (2001). Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *J Gen Virol* **82**, 765-773.

Merits, A., Rajamaki, M. L., Lindholm, P., Runeberg-Roos, P., Kekarainen, T., Puustinen, P., Makelainen, K., Valkonen, J. P. and Saarma, M. (2002). Proteolytic processing of potyviral proteins and polyprotein processing intermediates in insect and plant cells. *J Gen Virol* **83**, 1211-1221.

Mittal, D., Jindal, N., Gupta, S.L., Kataria, R.S. and Tiwari, A.K (2005). Detection of Infectious Bursal Disease Virus in Field Outbreaks in Broiler Chickens by Reverse Transcription-Polymerase Chain Reaction. *Int J of Poult Sci* **4**, 239-243.

Mkhize, P. (2003). Epitope mapping of a trypanosomal cysteine protease. *Department of Biochemistry*, University of KwaZulu-Natal, Pietermaritzburg. **MSc**

Morace, G., Kusov, Y., Dzagurov, G., Beneduce, F. and Gauss-Muller, V. (2008). The unique role of domain 2A of the hepatitis A virus precursor polypeptide P1-2A in viral morphogenesis. *BMB Rep* **41**, 678-683.

Morita, T., Kato, H., Iwanaga, S., Takada, K. and Kimura, T. (1977). New fluorogenic substrates for alpha-thrombin, factor Xa, kallikreins, and urokinase. *J Biochem* **82**, 1495-1498.

Morris, G. (2008). Epitope mapping. In Molecular Biomethods Handbook. (ed. Rapley R and Walker, J.). Humana Press Inc, Totowa, NJ. pp 619-630.

Muller, H., Islam, M. R. and Raue, R. (2003). Research on infectious bursal disease--the past, the present and the future. *Vet Microbiol* **97**, 153-165.

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. (1999). Tissue culture infectivity of different strains of infectious bursal disease virus is determined by distinct amino acids in VP2. *J Gen Virol* **80**, 2067-2076.

Murthy, H. M., Clum, S. and Padmanabhan, R. (1999). Dengue virus NS3 serine protease. Crystal structure and insights into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effects. *J Biol Chem* **274**, 5573-5580.

Myers, C. R., Porgilsson, B. and Myers, J. M. (1997). Antibodies to a synthetic peptide that react with flavin-containing monooxygenase (HLFMO3) in human hepatic microsomes. *J Pharmacol Toxicol Methods* **37**, 61-66.

Nair, S. and Savithri, H. S. (2010). Processing of SeMV polyproteins revisited. *Virology* **396**, 106-117.

Neiman, P. E., Thomas, S. J. and Loring, G. (1991). Induction of apoptosis during normal and neoplastic B-cell development in the bursa of Fabricius. *Proc Natl Acad Sci U S A* **88**, 5857-5861.

Nicola, A. V., Chen, W. and Helenius, A. (1999). Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nat Cell Biol* **1**, 341-345.

Nieminen, P., Liippo, J. and Lassila, O. (2002). Bursa of Fabricius. In Encyclopedia of Life Sciences. Macmillan Publishers Ltd, Nature Publishing Group. pp 1-6.

Niyomrattanakit, P., Winoyanuwattikun, P., Chanprapaph, S., Angsuthanasombat, C., Panyim, S. and Katzenmeier, G. (2004). Identification of residues in the dengue virus type 2 NS2B cofactor that are critical for NS3 protease activation. *J Virol* **78**, 13708-13716.

Nobiron, I., Galloux, M., Henry, C., Torhy, C., Boudinot, P., Lejal, N., Da Costa, B. and Delmas, B. (2008). Genome and polypeptides characterization of Tellina virus 1 reveals a fifth genetic cluster in the Birnaviridae family. *Virology* **371**, 350-361.

Nolan, D. P., Jackson, D. G., Biggs, M. J., Brabazon, E. D., Pays, A., Van Laethem, F., Paturiaux-Hanocq, F., Elliott, J. F., Voorheis, H. P. and Pays, E. (2000). Characterization of a novel alanine-rich protein located in surface microdomains in Trypanosoma brucei. *J Biol Chem* **275**, 4072-4080.

Oka, T., Katayama, K., Ogawa, S., Hansman, G. S., Kageyama, T., Ushijima, H., Miyamura, T. and Takeda, N. (2005). Proteolytic processing of sapovirus ORF1 polyprotein. *J Virol* **79**, 7283-7290.

Oka, T., Yamamoto, M., Yokoyama, M., Ogawa, S., Hansman, G. S., Katayama, K., Miyashita, K., Takagi, H., Tohya, Y., Sato, H. and Takeda, N. (2007). Highly conserved configuration of catalytic amino acid residues among calicivirus-encoded proteases. *J Virol* **81**, 6798-6806.

Olson, S. T. and Bjork, I. (1994). Regulation of thrombin activity by antithrombin and heparin. *Semin Thromb Hemost* **20**, 373-409.

Oppling, 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.

Paetzel, M. and Dalbey, R. E. (1997). Catalytic hydroxyl/amine dyads within serine proteases. *Trends Biochem Sci* **22**, 28-31.

Pan, J., Lin, L. and Tao, Y. J. (2009). Self-guanylylation of birnavirus VP1 does not require an intact polymerase activity site. *Virology* **395**, 87-96.

Pasternak, A. O., van den Born, E., Spaan, W. J. and Snijder, E. J. (2001). Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *Embo J* **20**, 7220-7228.

Peeters, J. M., Hazendonk, T. G., Beuvery, E. C. and Tesser, G. I. (1989). Comparison of four bifunctional reagents for coupling peptides to proteins and the effect of the three moieties on the immunogenicity of the conjugates. *J Immunol Methods* **120**, 133-143.

Peters, H., Kusov, Y. Y., Meyer, S., Benie, A. J., Bauml, E., Wolff, M., Rademacher, C., Peters, T. and Gauss-Muller, V. (2005). Hepatitis A virus proteinase 3C binding to viral RNA: correlation with substrate binding and enzyme dimerization. *Biochem J* **385**, 363-370.

Petersen, J. F., Cherney, M. M., Liebig, H. D., Skern, T., Kuechler, E. and James, M. N. (1999). The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. *Embo J* **18**, 5463-5475.

Petit, S., Lejal, N., Huet, J. C. and Delmas, B. (2000). Active residues and viral substrate cleavage sites of the protease of the birnavirus infectious pancreatic necrosis virus. *J Virol* **74**, 2057-2066.

Polgar, L. (2004). Catalytic mechanisms of serine and threonine peptidases. In Handbook of Proteolytic Enzymes. (ed. Barrett, A. J., Rawlings, N. D. and Woessner, J. F.). Elsevier, London. pp 1440-1448.

Polgar, L. (2005). The catalytic triad of serine peptidases. *Cell Mol Life Sci* **62**, 2161-2172.

Polson, A., Coetzer, T., Kruger, J., von Maltzahn, E. and van der Merwe, K. J. (1985). Improvements in the isolation of IgY from the yolks of eggs laid by immunized hens. *Immunol Invest* **14**, 323-327.

Pous, J., Chevalier, C., Ouldali, M., Navaza, J., Delmas, B. and Lepault, J. (2005). Structure of birnavirus-like particles determined by combined electron cryomicroscopy and X-ray crystallography. *J Gen Virol* **86**, 2339-2346.

Qian, B. and Kibenge, F. S. (1994). Observations on polymerase chain reaction amplification of infectious bursal disease virus dsRNA. *J Virol Methods* **47**, 237-242.

Qiu, X., Culp, J. S., DiLella, A. G., Hellmig, B., Hoog, S. S., Janson, C. A., Smith, W. W. and Abdel-Meguid, S. S. (1996). Unique fold and active site in cytomegalovirus protease. *Nature* **383**, 275-279.

Qiu, X., Janson, C. A., Culp, J. S., Richardson, S. B., Debouck, C., Smith, W. W. and Abdel-Meguid, S. S. (1997). Crystal structure of varicella-zoster virus protease. *Proc Natl Acad Sci U S A* **94**, 2874-2879.

Rachow, A., Gauss-Muller, V. and Probst, C. (2003). Homogeneous hepatitis A virus particles. Proteolytic release of the assembly signal 2A from procapsids by factor Xa. *J Biol Chem* **278**, 29744-29751.

Ratcliffe, M. J. H., Paramithiotis, E., Coumidis, A., Sayegh, C. E., Demaries, S. L., Martinez, O. and Jacobsen, K. A. (1996). The bursa of Fabricius and its role in avian B lymphocyte development. In Poultry immunology. (ed. Davison, T. R., Morris, T. and Payne, L.). Carfax Publishing Company, Abingdon, UK. pp 11-30.

Ratia, K., Saikatendu, K. S., Santarsiero, B. D., Barretto, N., Baker, S. C., Stevens, R. C. and Mesecar, A. D. (2006). Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc Natl Acad Sci U S A* **103**, 5717-5722.

Rautenschlein, S., Kraemer, C., Vanmarcke, J. and Montiel, E. (2005). Protective efficacy of intermediate and intermediate plus infectious bursal disease virus (IBDV) vaccines against very virulent IBDV in commercial broilers. *Avian Dis* **49**, 231-237.

Rawlings, N. D. and Barrett, A. J. (1993). Evolutionary families of peptidases. *Biochem J* **290**, 205-218.

Rawlings, N. D., Tolle, D. P. and Barrett, A. J. (2004). MEROPS: the peptidase database. *Nucleic Acids Res* **32**, D160-164.

Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2008). MEROPS: the peptidase database. *Nucleic Acids Res* **36**, D320-325.

Rawlings, N. D., Barrett, A. J. and Bateman, A. (2011). Asparagine peptide lyases: a seventh catalytic type of proteolytic enzymes. *J Biol Chem* **286**, 38321-38328.

Reddy, A., Schneemann , A. a. and Johnson, J. E. (2004). Nodavirus endopeptidase. In Handbook of Proteolytic Enzymes. (ed. Barrett, A. J., Rawlings, N. D. and Woessner, J. F.). Elsevier, London. pp 197-201.

Repp, H., Nieper, H., Draheim, H. J., Koschinski, A., Muller, H. and Dreyer, F. (1998). Infectious bursal disease virus changes the potassium current properties of chicken embryo fibroblasts. *Virology* **246**, 362-369.

Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S., Dryden, D. T., Dybvig, K., Firman, K., Gromova, E. S., Gumport, R. I., Halford, S. E., Hattman, S., Heitman, J., Hornby, D. P., Janulaitis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klaenhammer, T. R., Kobayashi, I., Kong, H., Kruger, D. H., Lacks, S., Marinus, M. G., Miyahara, M., Morgan, R. D., Murray, N. E., Nagaraja, V., Piekarowicz, A., Pingoud, A., Raleigh, E., Rao, D. N., Reich, N., Repin, V. E., Selker, E. U., Shaw, P. C., Stein, D. C., Stoddard, B. L., Szybalski, W., Trautner, T. A., Van Etten, J. L., Vitor, J. M., Wilson, G. G. and Xu, S. Y. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* **31**, 1805-1812.

Rodriguez-Lecompte, J. C. and Kibenge, F. S. (2002). Site-directed mutagenesis of Avibirnavirus VP4 gene. *Virology* **292**, 241-246.

Rosales-Leon, L., Ortega-Lule, G. and Ruiz-Ordaz, B. (2007). Analysis of the domain interactions between the protease and helicase of NS3 in dengue and hepatitis C virus. *J Mol Graph Model* **25**, 585-594.

Rudenskaya, G. N. and Pupov, D. V. (2008). Cysteine proteinases of microorganisms and viruses. *Biochemistry (Mosc)* **73**, 1-13.

Rumenapf, T., Stark, R., Heimann, M. and Thiel, H. J. (1998). N-terminal protease of pestiviruses: identification of putative catalytic residues by site-directed mutagenesis. *J Virol* **72**, 2544-2547.

Ryan, M. D., Monaghan, S. and Flint, M. (1998). Virus-encoded proteinases of the Flaviviridae. *J Gen Virol* **79**, 947-959.

Sanchez-Burgos, G., Ramos-Castaneda, J., Cedillo-Rivera, R. and Dumonteil, E. (2010). Immunogenicity of novel Dengue virus epitopes identified by bioinformatic analysis. *Virus Res* **153**, 113-120.

Sanchez, A. B. and Rodriguez, J. F. (1999). Proteolytic processing in infectious bursal disease virus: identification of the polyprotein cleavage sites by site-directed mutagenesis. *Virology* **262**, 190-199.

Saravanan, P., Satishkumar, Kataria, J. M. and Rasool, T. J. (2004). Detection of Infectious bursal disease virus by ELISA using an antipeptide antibody raised against VP3 region. *Acta Virol* **48**, 39-45.

Satheshkumar, P. S., Lokesh, G. L. and Savithri, H. S. (2004). Polyprotein processing: cis and trans proteolytic activities of Sesbania mosaic virus serine protease. *Virology* **318**, 429-438.

Saugar, I., Luque, D., Ona, A., Rodriguez, J. F., Carrascosa, J. L., Trus, B. L. and Caston, J. R. (2005). Structural polymorphism of the major capsid protein of a double-stranded RNA virus: an amphipathic alpha helix as a molecular switch. *Structure* **13**, 1007-1017.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A. and Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**, 234-258.

Sayegh, C. E., Demaries, S. L., Iacampo, S. and Ratcliffe, M. J. (1999). Development of B cells expressing surface immunoglobulin molecules that lack V(D)J-encoded determinants in the avian embryo bursa of fabricius. *Proc Natl Acad Sci U S A* **96**, 10806-10811.

Schechter, I. and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-162.

Schnitzler, D., Bernstein, F., Muller, H. and Becht, H. (1993). The genetic basis for the antigenicity of the VP2 protein of the infectious bursal disease virus. *J Gen Virol* **74**, 1563-1571.

Schregel, V., Jacobi, S., Penin, F. and Tautz, N. (2009). Hepatitis C virus NS2 is a protease stimulated by cofactor domains in NS3. *Proc Natl Acad Sci U S A* **106**, 5342-5347.

Schultheiss, T., Sommergruber, W., Kusov, Y. and Gauss-Muller, V. (1995). Cleavage specificity of purified recombinant hepatitis A virus 3C proteinase on natural substrates. *J Virol* **69**, 1727-1733.

Schulz-Utermoehl, T., Edwards, R. J. and Boobis, A. R. (2000). Affinity and potency of proinhibitory antipeptide antibodies against CYP2D6 is enhanced using cyclic peptides as immunogens. *Drug Metab Dispos* **28**, 544-551.

Schutze, H., Ulferts, R., Schelle, B., Bayer, S., Granzow, H., Hoffmann, B., Mettenleiter, T. C. and Ziebuhr, J. (2006). Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. *J Virol* **80**, 11598-11609.

Setiyono, A., Hayashi, T., Yamaguchi, T., Fukushi, H. and Hirai, K. (2001). Detection of cell membrane proteins that interact with virulent infectious bursal disease virus. *J Vet Med Sci* **63**, 219-221.

Shieh, C. C., Coghlan, M., Sullivan, J. P. and Gopalakrishnan, M. (2000). Potassium channels: molecular defects, diseases, and therapeutic opportunities. *Pharmacol Rev* **52**, 557-594.

Shiryaev, S. A., Ratnikov, B. I., Chekanov, A. V., Sikora, S., Rozanov, D. V., Godzik, A., Wang, J., Smith, J. W., Huang, Z., Lindberg, I., Samuel, M. A., Diamond, M. S. and Strongin, A. Y. (2006). Cleavage targets and the D-arginine-based inhibitors of the West Nile virus NS3 processing proteinase. *Biochem J* **393**, 503-511.

Sijwali, P. S., Brinen, L. S. and Rosenthal, P. J. (2001). Systematic optimization of expression and refolding of the Plasmodium falciparum cysteine protease falcipain-2. *Protein Expr Purif* **22**, 128-134.

Skern, T., Fita, I. and Guarne, A. (1998). A structural model of picornavirus leader proteinases based on papain and bleomycin hydrolase. *J Gen Virol* **79**, 301-307.

Skoging, U. and Liljestrom, P. (1998). Role of the C-terminal tryptophan residue for the structure-function of the alphavirus capsid protein. *J Mol Biol* **279**, 865-872.

Smooker, P. M., Jayaraj, R., Pike, R. N. and Spithill, T. W. (2010). Cathepsin B proteases of flukes: the key to facilitating parasite control? *Trends Parasitol* **26**, 506-514.

Snyder, D. B., Lana, D. P., Savage, P. K., Yancey, F. S., Mengel, S. A. and Marquardt, W. W. (1988). Differentiation of infectious bursal disease viruses directly from infected tissues with neutralizing monoclonal antibodies: evidence of a major antigenic shift in recent field isolates. *Avian Dis* **32**, 535-539.

Snyder, D. B., Vakharia, V. N. and Savage, P. K. (1992). Naturally occurring-neutralizing monoclonal antibody escape variants define the epidemiology of infectious bursal disease viruses in the United States. *Arch Virol* **127**, 89-101.

Sorensen, H. P. and Mortensen, K. K. (2005a). Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact* **4**, 1.

Sorensen, H. P. and Mortensen, K. K. (2005b). Advanced genetic strategies for recombinant protein expression in Escherichia coli. *J Biotechnol* **115**, 113-128.

Sosnovtseva, S. A., Sosnovtsev, S. V. and Green, K. Y. (1999). Mapping of the feline calicivirus proteinase responsible for autocatalytic processing of the nonstructural polyprotein and identification of a stable proteinase-polymerase precursor protein. *J Virol* **73**, 6626-6633.

Speroni, S., Rohayem, J., Nenci, S., Bonivento, D., Robel, I., Barthel, J., Luzhkov, V. B., Coutard, B., Canard, B. and Mattevi, A. (2009). Structural and biochemical analysis of human pathogenic astrovirus serine protease at 2.0 A resolution. *J Mol Biol* **387**, 1137-1152.

Spycher, S. E., Nick, H. and Rickli, E. E. (1986). Human complement component C1s. Partial sequence determination of the heavy chain and identification of the peptide bond cleaved during activation. *Eur J Biochem* **156**, 49-57.

Sreedevi, B., LeFever, L. J., Sommer-Wagner, S. E. and Jackwood, D. J. (2007). Characterization of infectious bursal disease viruses from four layer flocks in the United States. *Avian Dis* **51**, 845-850.

Strong, R. and Belsham, G. J. (2004). Sequential modification of translation initiation factor eIF4GI by two different foot-and-mouth disease virus proteases within infected baby hamster kidney cells: identification of the 3Cpro cleavage site. *J Gen Virol* **85**, 2953-2962.

Suzuki, T., Ishii, K., Aizaki, H. and Wakita, T. (2007). Hepatitis C viral life cycle. *Adv Drug Deliv Rev* **59**, 1200-1212.

Sweeney, T. R., Roque-Rosell, N., Birtley, J. R., Leatherbarrow, R. J. and Curry, S. (2007). Structural and mutagenic analysis of foot-and-mouth disease virus 3C protease reveals the role of the beta-ribbon in proteolysis. *J Virol* **81**, 115-124.

Tacken, M. G., Rottier, P. J., Gielkens, A. L. and Peeters, B. P. (2000). Interactions in vivo between the proteins of infectious bursal disease virus: capsid protein VP3 interacts with the RNA-dependent RNA polymerase, VP1. *J Gen Virol* **81**, 209-218.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-1599.

Tanimura, N. and Sharma, J. M. (1998). In-situ apoptosis in chickens infected with infectious bursal disease virus. *J Comp Pathol* **118**, 15-27.

ten Dam, E., Flint, M. and Ryan, M. D. (1999). Virus-encoded proteinases of the Togaviridae. *J Gen Virol* **80** 1879-1888.

Tham, K. M. and Moon, C. D. (1996). Apoptosis in cell cultures induced by infectious bursal disease virus following in vitro infection. *Avian Dis* **40**, 109-113.

Thiel, S., Vorup-Jensen, T., Stover, C. M., Schwaeble, W., Laursen, S. B., Poulsen, K., Willis, A. C., Eggleton, P., Hansen, S., Holmskov, U., Reid, K. B. and Jensenius, J. C. (1997). A second serine protease associated with mannan-binding lectin that activates complement. *Nature* **386**, 506-510.

Thomson, L. M., Lamont, D. J., Mehlert, A., Barry, J. D. and Ferguson, M. A. (2002). Partial structure of glutamic acid and alanine-rich protein, a major surface glycoprotein of the insect stages of Trypanosoma congolense. *J Biol Chem* **277**, 48899-48904.

Tian, X., Lu, G., Gao, F., Peng, H., Feng, Y., Ma, G., Bartlam, M., Tian, K., Yan, J., Hilgenfeld, R. and Gao, G. F. (2009). Structure and cleavage specificity of the chymotrypsin-like serine protease (3CLSP/nsp4) of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV). *J Mol Biol* **392**, 977-993.

Tizard, I. (2002). The Avian Antibody Response. *Seminars in Avian and Exotic Pet Medicine* **11**, 2-14.

Tong, L. (2002). Viral proteases. *Chem Rev* **102**, 4609-4626.

Torruella, M., Gordon, K. and Hohn, T. (1989). Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. *Embo J* **8**, 2819-2825.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354.

Troeberg, L., Pike, R. N., Lonsdale-Eccles, J. D. and Coetzer, T. H. (1997). Production of anti-peptide antibodies against trypanopain-Tb from Trypanosoma brucei brucei: effects of antibodies on enzyme activity against Z-Phe-Arg-AMC. *Immunopharmacology* **36**, 295-303.

Tsai, K. N. and Chen, G. W. (2011). Influenza genome diversity and evolution. *Microbes Infect* **13**, 479-488.

Tseng, W. C., Lin, J. W., Wei, T. Y. and Fang, T. Y. (2008). A novel megaprimed and ligase-free, PCR-based, site-directed mutagenesis method. *Anal Biochem* **375**, 376-378.

Tsukamoto, K., Tanimura, N., Kakita, S., Ota, K., Mase, M., Imai, K. and Hihara, H. (1995). Efficacy of three live vaccines against highly virulent infectious bursal disease virus in chickens with or without maternal antibodies. *Avian Dis* **39**, 218-229.

Tsukamoto, K., Saito, S., Saeki, S., Sato, T., Tanimura, N., Isobe, T., Mase, M., Imada, T., Yuasa, N. and Yamaguchi, S. (2002). Complete, long-lasting protection against lethal infectious bursal disease virus challenge by a single vaccination with an avian herpesvirus vector expressing VP2 antigens. *J Virol* **76**, 5637-5645.

Tsumoto, K., Misawa, S., Ohba, Y., Ueno, T., Hayashi, H., Kasai, N., Watanabe, H., Asano, R. and Kumagai, I. (2002). Inhibition of hepatitis C virus NS3 protease by peptides derived from complementarity-determining regions (CDRs) of the monoclonal antibody 8D4: tolerance of a CDR peptide to conformational changes of a target. *FEBS Lett* **525**, 77-82.

Tyler, S., Severini, A., Black, D., Walker, M. and Eberle, R. (2011). Structure and sequence of the saimiriine herpesvirus 1 genome. *Virology* **410**, 181-191.

Urcuqui-Inchima, S., Haenni, A. L. and Bernardi, F. (2001). Potyvirus proteins: a wealth of functions. *Virus Res* **74**, 157-175.

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., Snyder, D. B., He, J., Edwards, G. H., Savage, P. K. and Mengel-Whereat, S. A. (1993). Infectious bursal disease virus structural proteins expressed in a baculovirus recombinant confer protection in chickens. *J Gen Virol* **74** 1201-1206.

Vakharia, V. N., Snyder, D. B., Lutticken, D., Mengel-Whereat, S. A., Savage, P. K., Edwards, G. H. and Goodwin, M. A. (1994). Active and passive protection against variant and classic infectious bursal disease virus strains induced by baculovirus-expressed structural proteins. *Vaccine* **12**, 452-456.

van Aken, D., Snijder, E. J. and Gorbalenya, A. E. (2006a). Mutagenesis analysis of the nsp4 main proteinase reveals determinants of arterivirus replicase polyprotein autoprocessing. *J Virol* **80**, 3428-3437.

van Aken, D., Zevenhoven-Dobbe, J., Gorbalenya, A. E. and Snijder, E. J. (2006b). Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. *J Gen Virol* **87**, 3473-3482.

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. and Meulemans, G. (1991). Acute infectious bursal disease in poultry: protection afforded by maternally derived antibodies and interference with live vaccination. *Avian Pathol* **20**, 409-421.

van den Berg, T. P., Eterradossi, N., Toquin, D. and Meulemans, G. (2000). Infectious bursal disease (Gumboro disease). *Revenue Scientifique et Technique (International Office of Epizootics)* **19**, 509-543.

van Loon, A. A., de Haas, N., Zeyda, I. and Mundt, E. (2002). Alteration of amino acids in VP2 of very virulent infectious bursal disease virus results in tissue culture adaptation and attenuation in chickens. *J Gen Virol* **83**, 121-129.

Van Regenmortel, M. H. V. (1999). Molecular dissection of protein antigens and the prediction of epitopes. In Synthetic Peptides as Antigens (ed. Van Regenmortel, M. H. V. and Muller, S.). Elsevier, Amsterdam. pp 1-78.

Vasconcelos, A. C. and Lam, K. M. (1994). Apoptosis induced by infectious bursal disease virus. *J Gen Virol* **75**, 1803-1806.

Vasconcelos, A. C. and Lam, K. M. (1995). Apoptosis in chicken embryos induced by the infectious bursal disease virus. *J Comp Pathol* **112**, 327-338.

Verchot, J. and Carrington, J. C. (1995). Evidence that the potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. *J Virol* **69**, 3668-3674.

Volkov, A. A., Shao, Z. and Arnold, F. H. (1999). Recombination and chimeragenesis by in vitro heteroduplex formation and in vivo repair. *Nucleic Acids Res* **27**, e18.

Wang, Y., Wu, X., Li, H., Wu, Y., Shi, L., Zheng, X., Luo, M., Yan, Y. and Zhou, J. (2009). Antibody to VP4 protein is an indicator discriminating pathogenic and nonpathogenic IBDV infection. *Mol Immunol* **46**, 1964-1969.

Wang, Y. S., Fan, H. J., Li, Y., Shi, Z. L., Pan, Y. and Lu, C. P. (2007). Development of a multi-mimotope peptide as a vaccine immunogen for infectious bursal disease virus. *Vaccine* **25**, 4447-4455.

Welling, G. W., Weijer, W. J., van der Zee, R. and Welling-Wester, S. (1985). Prediction of sequential antigenic regions in proteins. *FEBS Lett* **188**, 215-218.

White, C. B., Chen, Q., Kenyon, G. L. and Babbitt, P. C. (1995). A novel activity of OmpT. Proteolysis under extreme denaturing conditions. *J Biol Chem* **270**, 12990-12994.

Wilson, D. R. and Finlay, B. B. (1997). The 'Asx-Pro turn' as a local structural motif stabilized by alternative patterns of hydrogen bonds and a consensus-derived model of the sequence Asn-Pro-Asn. *Protein Eng* **10**, 519-529.

Wirblich, C., Sibilia, M., Boniotti, M. B., Rossi, C., Thiel, H. J. and Meyers, G. (1995). 3C-like protease of rabbit hemorrhagic disease virus: identification of cleavage sites in the ORF1 polyprotein and analysis of cleavage specificity. *J Virol* **69**, 7159-7168.

Wlodawer, A. and Gustchina, A. (2000). Structural and biochemical studies of retroviral proteases. *Biochim Biophys Acta* **1477**, 16-34.

Wu, P. C., Su, H. Y., Lee, L. H., Lin, D. T., Yen, P. C. and Liu, H. J. (2005a). Secreted expression of the VP2 protein of very virulent infectious bursal disease virus in the methylotrophic yeast Pichia pastoris. *J Virol Methods* **123**, 221-225.

Wu, S. and Letchworth, G. J. (2004). High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol. *Biotechniques* **36**, 152-154.

Wu, W., Jia, Z., Liu, P., Xie, Z. and Wei, Q. (2005b). A novel PCR strategy for high-efficiency, automated site-directed mutagenesis. *Nucleic Acids Res* **33**, e110.

Xia, R. X., Wang, H. Y., Huang, G. M. and Zhang, M. F. (2008). Sequence and phylogenetic analysis of a Chinese very virulent infectious bursal disease virus. *Arch Virol* **153**, 1725-1729.

Yamaguchi, H., Miki, H. and Takenawa, T. (2002). Two verprolin homology domains increase the Arp2/3 complex-mediated actin polymerization activities of N-WASP and WAVE1 C-terminal regions. *Biochem Biophys Res Commun* **297**, 214-219.

Yamaguchi, T., Kondo, T., Inoshima, Y., Ogawa, M., Miyoshi, M., Yanai, T., Masegi, T., Fukushi, H. and Hirai, K. (1996a). In vitro attenuation of highly virulent infectious bursal disease virus: some characteristics of attenuated strains. *Avian Dis* **40**, 501-509.

Yamaguchi, T., Ogawa, M., Inoshima, Y., Miyoshi, M., Fukushi, H. and Hirai, K. (1996b). Identification of sequence changes responsible for the attenuation of highly virulent infectious bursal disease virus. *Virology* **223**, 219-223.

Yamaguchi, T., Ogawa, M., Miyoshi, M., Inoshima, Y., Fukushi, H. and Hirai, K. (1997). Sequence and phylogenetic analyses of highly virulent infectious bursal disease virus. *Arch Virol* **142**, 1441-1458.

Yamamoto, Y., Inagaki, N. and Satoh, K. (2001). Overexpression and characterization of carboxyl-terminal processing protease for precursor D1 protein: regulation of enzyme-substrate interaction by molecular environments. *J Biol Chem* **276**, 7518-7525.

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R. and Rao, Z. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* **100**, 13190-13195.

Yao, K., Goodwin, M. A. and Vakharia, V. N. (1998). Generation of a mutant infectious bursal disease virus that does not cause bursal lesions. *J Virol* **72**, 2647-2654.

Yao, N., Reichert, P., Taremi, S. S., Prosise, W. W. and Weber, P. C. (1999). Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* **7**, 1353-1363.

Yip, C. W., Yeung, Y. S., Ma, C. M., Lam, P. Y., Hon, C. C., Zeng, F. and Leung, F. C. (2007). Demonstration of receptor binding properties of VP2 of very virulent strain infectious bursal disease virus on Vero cells. *Virus Res* **123**, 50-56.

Yokoyama, H., Hamamatsu, S., Fujii, S. and Matsui, I. (2008). Novel dimer structure of a membrane-bound protease with a catalytic Ser-Lys dyad and its linkage to stomatin. *J Synchrotron Radiat* **15**, 254-257.

Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., Lv, Z., Ge, R., Li, X., Deng, T., Fodor, E., Rao, Z. and Liu, Y. (2009). Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* **458**, 909-913.

Yusof, R., Clum, S., Wetzel, M., Murthy, H. M. and Padmanabhan, R. (2000). Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. *J Biol Chem* **275**, 9963-9969.

Zeitler, C. E., Estes, M. K. and Venkataram Prasad, B. V. (2006). X-ray crystallographic structure of the Norwalk virus protease at 1.5-A resolution. *J Virol* **80**, 5050-5058.

Zheng, X., Hong, L., Shi, L., Guo, J., Sun, Z. and Zhou, J. (2008). Proteomics analysis of host cells infected with infectious bursal disease virus. *Mol Cell Proteomics* **7**, 612-625.

Zinovkin, R., Erokhina, T., Lesemann, D., Jelkmann, W. a. and Agranovsky, A. (2003). Processing and subcellular localization of the leader papain-like proteinase of Beet yellows closterovirus. *J Gen Virol* **84**, 2265-2270.

Zlotnick, A., Reddy, V. S., Dasgupta, R., Schneemann, A., Ray, W. J., Jr., Rueckert, R. R. and Johnson, J. E. (1994). Capsid assembly in a family of animal viruses primes an autoproteolytic maturation that depends on a single aspartic acid residue. *J Biol Chem* **269**, 13680-13684.

Zoll, J., van Kuppeveld, F. J., Galama, J. M. and Melchers, W. J. (1998). Genetic analysis of mengovirus protein 2A: its function in polyprotein processing and virus reproduction. *J Gen Virol* **79**, 17-25.

Appendix 1

Expression vectors used in the present study

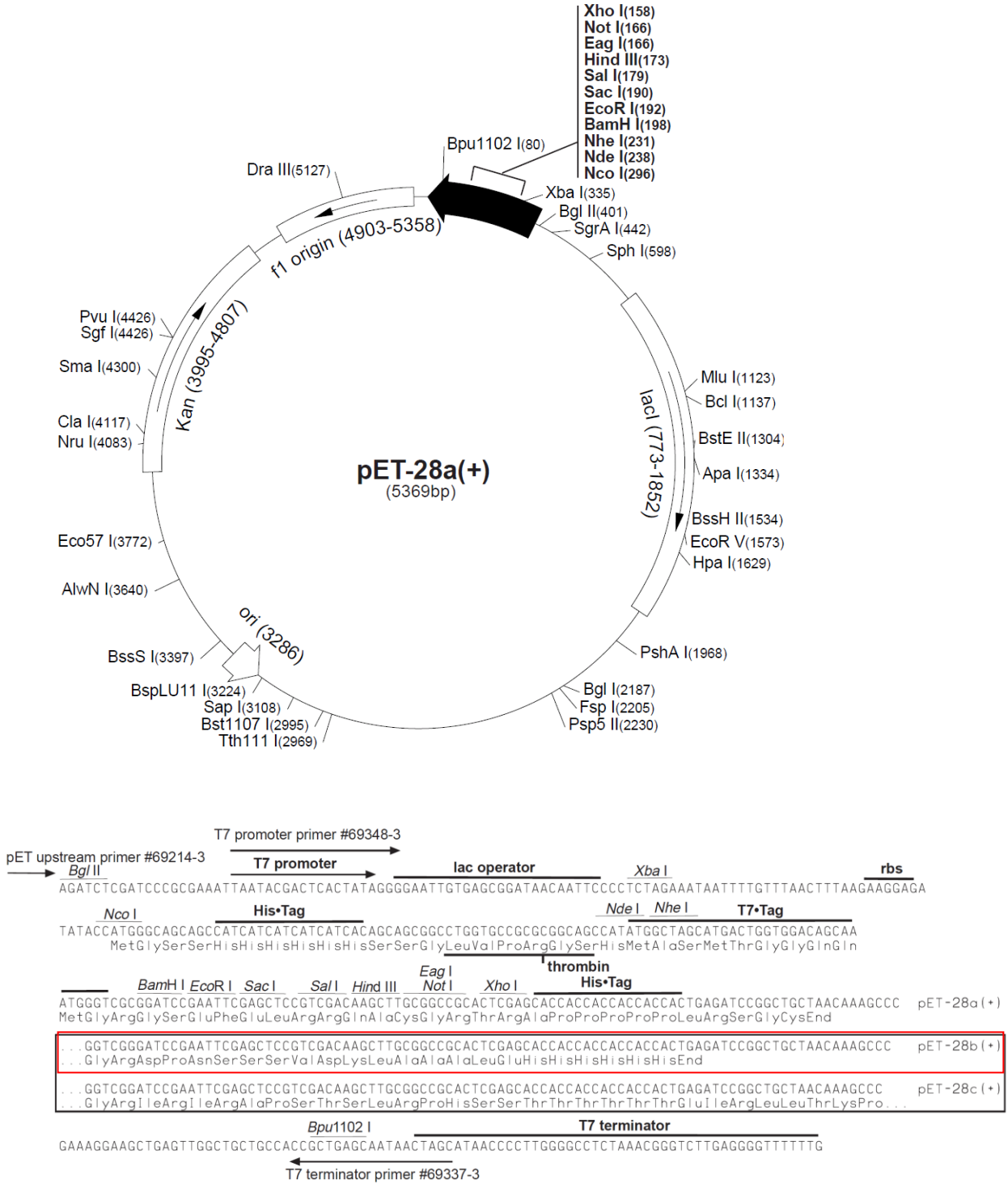


Figure 1. Structural features of the expression vector pET-28 plasmid. The top panel shows the circle map of pET 28 plasmids. The bottom panel contains the sequence of the cloning/expression region and unique restriction sites for cloning foreign DNA are shown. The kanamycin resistant pET-28b vector carries an N-terminal His•Tag/thrombin/T7•Tag configuration in addition to a C-terminal His•Tag sequence (Novagen).

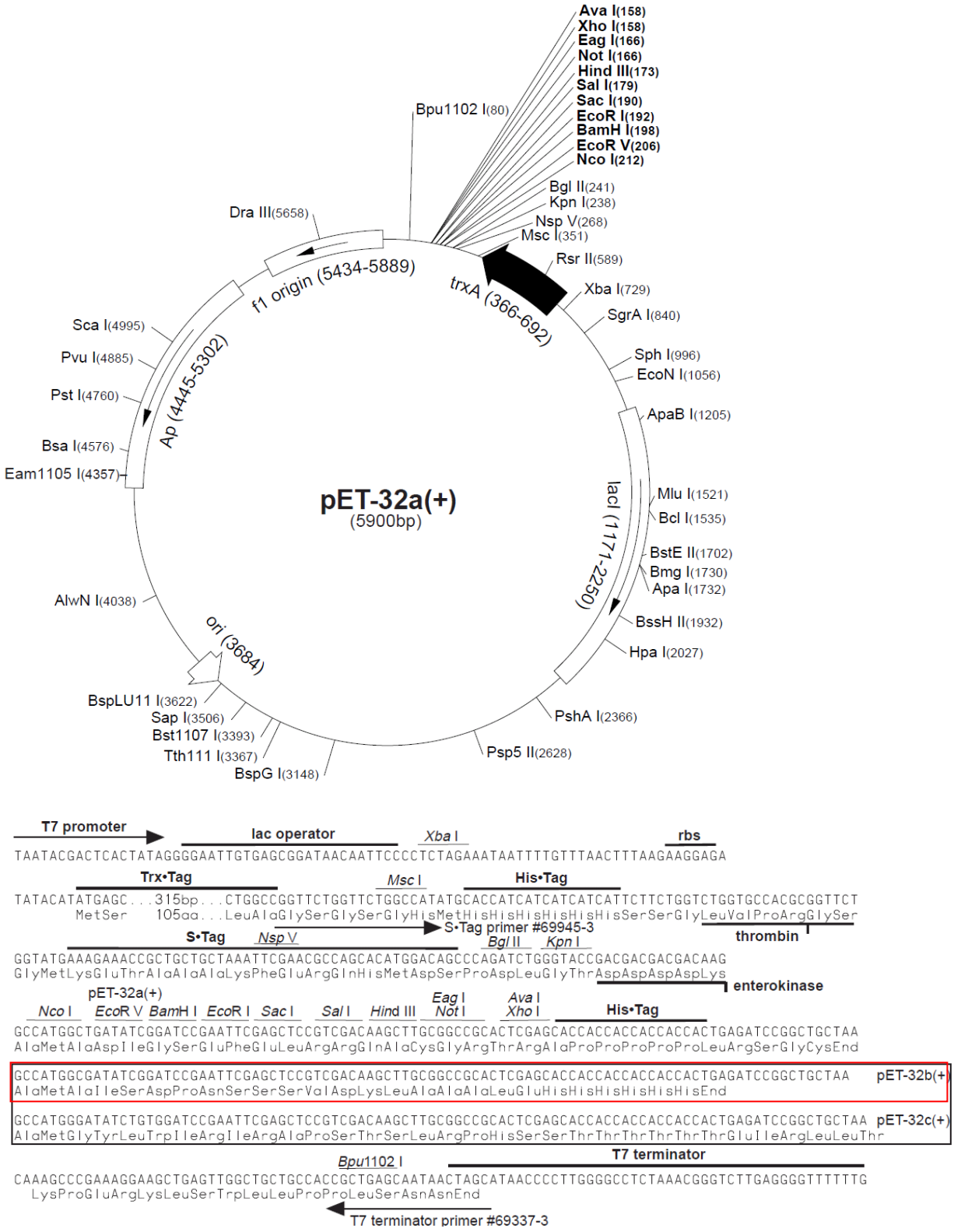


Figure 2. Structural features of the expression vector pET-32 plasmids. The top panel shows the circle map of pET-32 plasmids. The bottom panel contains the sequence of the cloning/expression region and unique restriction sites for cloning foreign DNA are shown. The ampicillin resistant pET-32b vector carries an N-terminal 109aa Trx•Tag thioredoxin protein/His•Tag/ thrombin/S•Tag configuration in addition to a C-terminal His•Tag sequence (Novagen).

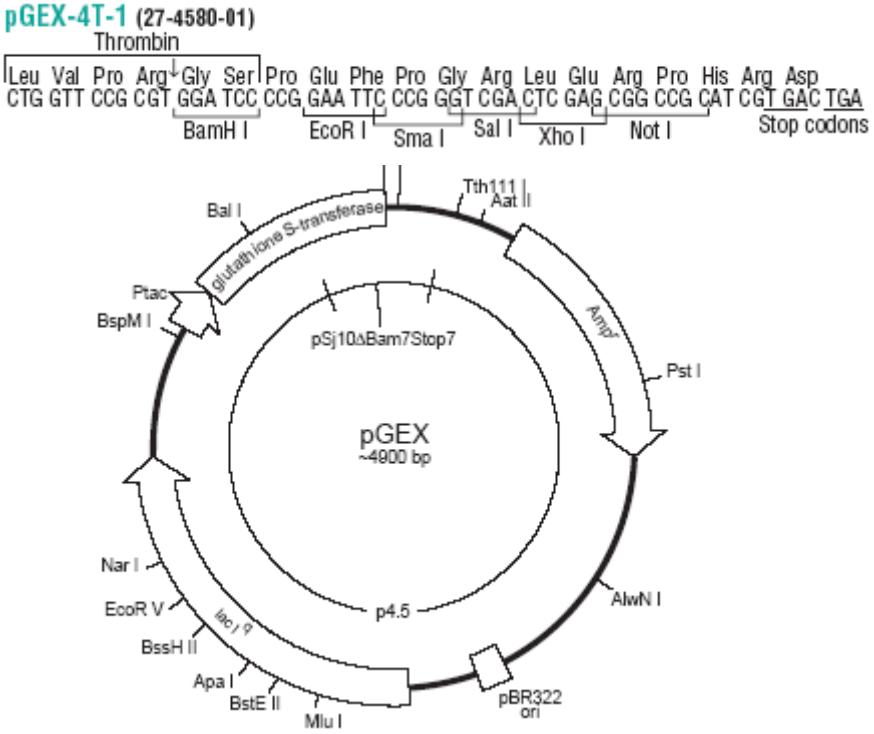


Figure 3. Structure of the protein expression vector pGEX-4T-1. The vector is approximately 4900 bp and carries the pMB1 origin of replication, amp resistance marker and tac promoter. The restriction sites that are used for cloning are shown above the circular plasmid map. The thin circle inside the main map shows the sources of the segments that were linked in the construction of the vector (Amersham Biosciences).

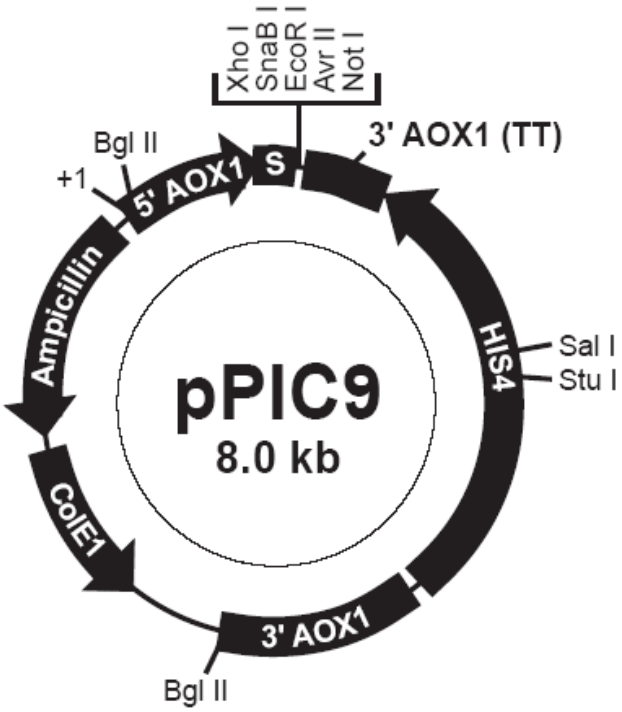


Figure 4. **Structure of the *Pichia* expression vector pPIC9.** The vector is 8028 bp and carries the f1 origins of replication, ampicillin res, ColE1 origin and AOX1 promoter sequence. The unique restriction sites that are used for cloning are shown (Invitrogen).

Appendix 2

Sequences of the universal primers used in the present study

Table 1 Sequences of the universal primers used for colony PCR or sequencing in the present study

Vector	Primer name	Sequence (5'→3')
pTZ57R/T	M13 forward	GTA AAA CGA CGG CCA GTG
	T7 reverse	TAT GCT AGT TAT TGC TCA G
PGEM T vector	T7 forward	TAA TAC GAC TCA CTA TAG GG
	SP6 reverse	AGC TAT TTA GGT GAC ACT ATA G
pET-32a; pET 28a	pET forward	TAA TAC GAC TCA CTA TAG GG
	pET reverse	GCT AGT TAT TGC TCA GCG G
pGEX-4T-1	pGEX forward	CCA GCA AGT ATA TAG CAT GG
	pGEX reverse	CCG GGA GCT GCA TGT GTC AGA GG
pPIC9	AOX forward	GAC TGG TTC CAA TTG ACA AG
	AOX reverse	GCA AAT GGC ATT CTG ACA TCC

Appendix 3

Predict 7 plot for VP4-RA sequence

```
        462         472         482         492         502         512         522         532         542         552         562         572
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
RIAVPVVSTL FPPAAPLAHA IGEGVDYLLG DEAQAASGTA RAASGKARAA SGRIRQLTLA ADKGYEVVAN LFQVPQNPVV DGILASPGIL RGAHNLDCVL REGATLFPVV ITTVEDAMTP
        582         592         602         612         622         632         642         652         662         672         682         692
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
KALNSKMFAV IEGVREDLQP PSQRGSFIRT LSGHRVYGYA PDGVLPLETG RDYTVVPIDD VWDDSIMLSK DPIPPIVGNS GNLAIAYMDV FRPKVPIHVA MTGALNAYGE IENVSFRSTK
        702         712         722         732         742         752
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ...
LATAHRLGLK LAGPGAFDVN TGSNWATFIK RFPHNPRDWD RLPYLNLPYL PPNAGRQYDL AMA
```

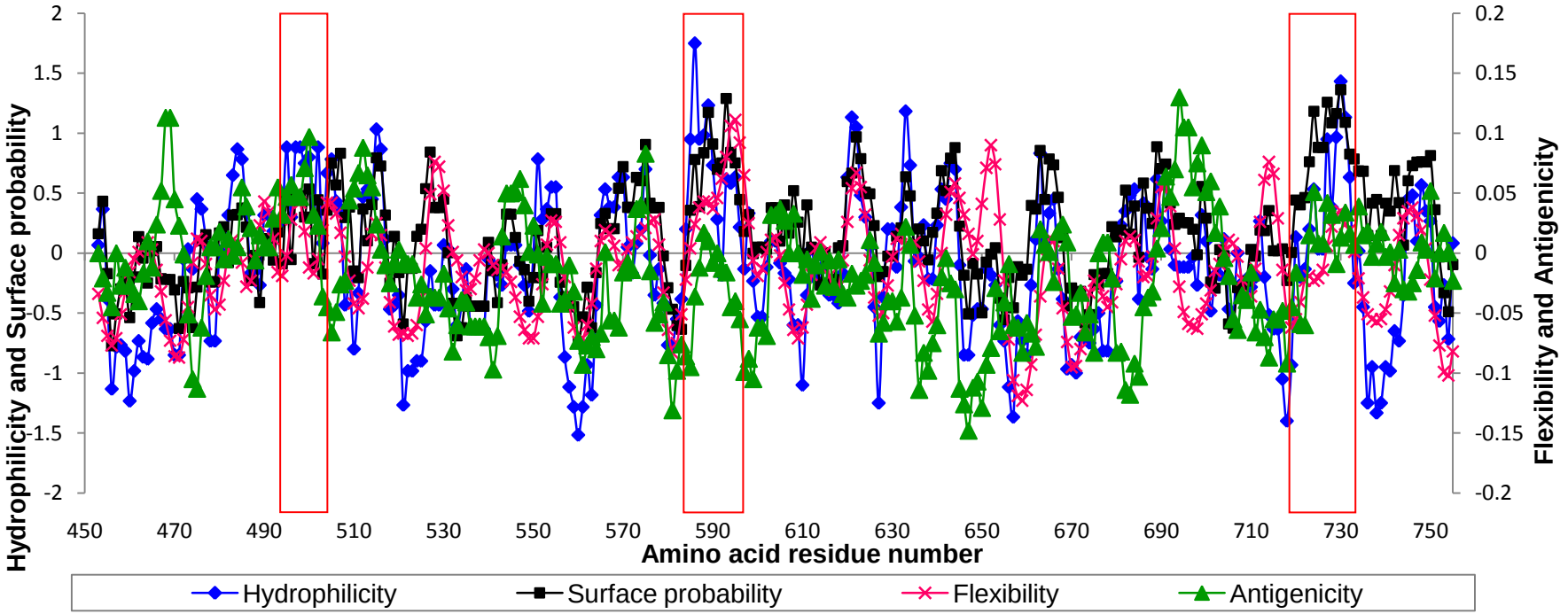


Figure 5 Predict 7 plot for VP4-RA sequence (Arg^{453}-Ala^{755}). Four variables were plotted: hydrophilicity, surface probability, flexibility and antigenicity (Cármenes *et al.*, 1989). The VP4-RA sequence is shown above. The selected peptides are boxed in red.

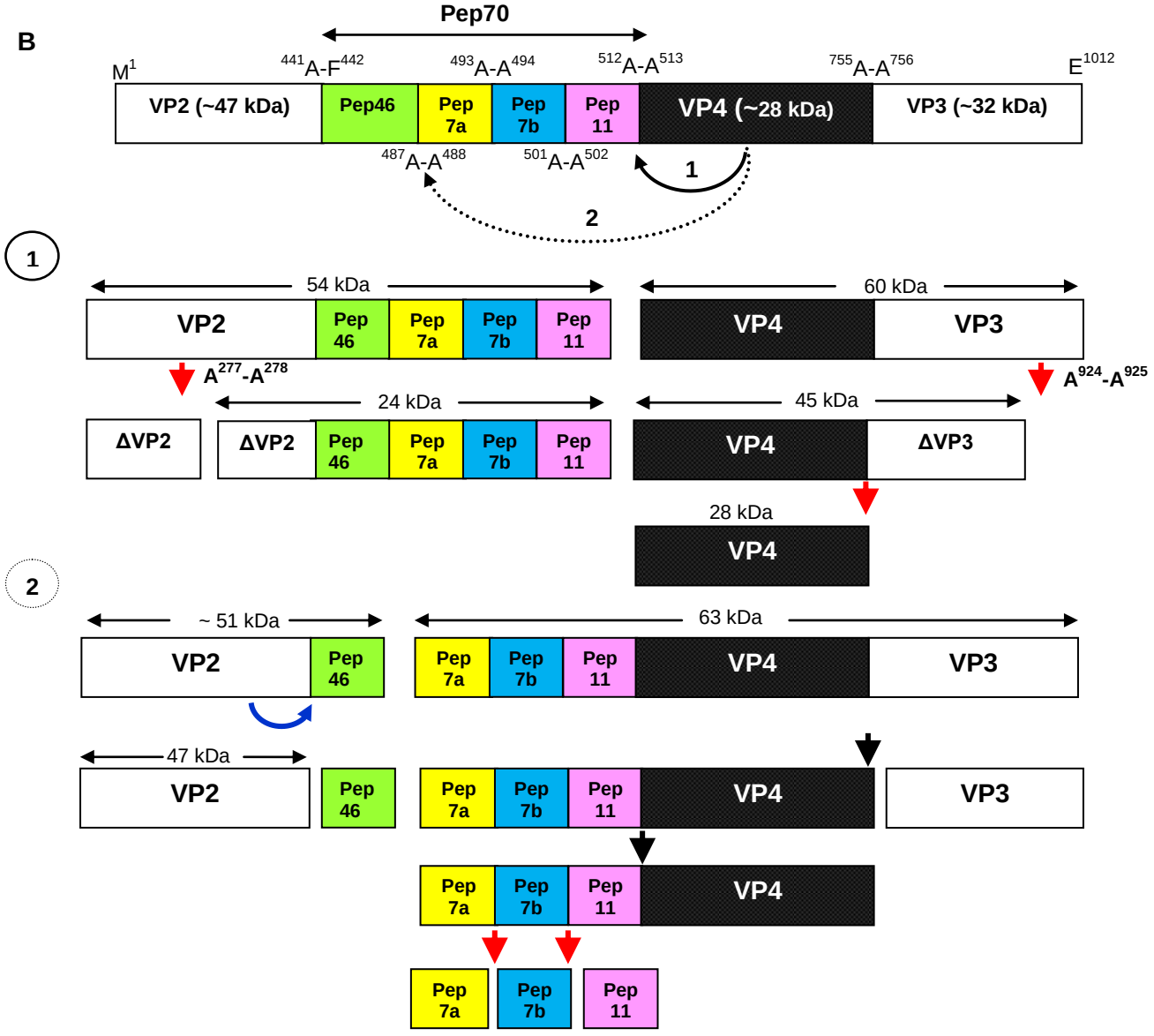
Appendix 4

The proposed model of polyprotein processing

A

```
MTNLQDQTQQ IVPFIRSLLM PTTGPASIPD DTLEKHTLRS ETSTYNLTVG DTGSGLIVFF PGFPGSIVGA 70
HYTLQSNGNY KFDQMLLTAQ NLPASYNYCR LVSRSLTVRS STLPGGVYAL NGTINAVTFQ GSLSELTDVS 140
YNGLMSATAN INDKIGNVLV GEGVTVLSLP TSYDLGYVRL GDPIPAIGLD PKMVATCDSS DRPRVYTITA 210
ADDYQFSSQY QSGGVTITLF SANIDAITSL SIGGELVFQT SVQGLILGAT IYLIGFDGTA VITRAVAADN 280
GLTAGTDNLM PFNIVIPISE ITQPITSIKL EIVTSKSGGQ AGDQMSWSAS GSLAVTIHGG NYPGALRPVT 350
LVAYERVATG SVVTVAGVSN FELIPNPELA KNLVTEYGRF DPGAMNYTKL ILSERDRLGI KTVWPTREYT 420
DFREYFMEVA DLNSPLKIAG AFGFKDIIRA LRRIAVPVVS TLFPPAAPLA HAIGEGVDYL LGDEAQAASG 490
TARAASGKAR AASGRIRQLT LAADKGYEVV ANLFQVPQNP VVDGILASPG ILRGAHNLDC VLREGATLFP 560
VVITTVEDAM TPKALNSKMF AVIEGVREDL QPPSQRGSFI RTLSGHRVYG YAPDGVLPLE TGRDYTVVPI 630
DDVWDDSIML SKDPIPPIVG NSGNLAIAYM DVFRPKVPIH VAMTGALNAY GEIENVSFRS TKLATAHRLG 700
LKLAGPGAFD VNTGSNWATF IKRFPHNPRD WDRLPYLNLP YLPPNAGRQY DLAMAASEFK ETPELESAVR 770
AMEAAANVDP LFQSALSVFM WLEENGIVTD MANFALSDPN AHRMRNFLAN APQAGSKSQR AKYGTAGYGV 840
EARGPTPEEA QREKDTRISK KMETMGIYFA TPEWVALNGH RGPSPGQLKY WQNTREIPDP NEDYLDYVHA 910
EKSRLASEEQ ILRAATSIYG APGQAEPPQA FIDEVAKVYE INHGRGPNQE QMKDLLLTAM EMKHRNPRRA 980
PPKPKPKPNV PTQRPPGRLG RWIRAVSDED LE        1012
```

B



See Fig. 6.1 for information