CHAPTER 1

LITERATURE REVIEW

1.1 INTRODUCTION

African horsesickness (AHS), which is endemic to sub-Saharan Africa, is a non-contagious but infectious disease of equines. The severity of the disease is well illustrated by a recent outbreak in the Western Cape, South Africa, in which only 3 of the 34 affected horses survived (P.G. Howell, personal communication). The aetiological agent of the disease, African horsesickness virus (AHSV), was identified as early as 1901 by Sir Arnold Theiler (Theiler, 1921).

1.2 AFRICAN HORSESICKNESS VIRUS

1.2.1 Classification

AHSV is a member of the *Orbivirus* genus within the *Reoviridae* family (Verwoerd *et al.*, 1979). The *Reoviridae* family is characterised by (i) a segmented double-stranded RNA genome of 10 to 12 segments, and (ii) virus particles (60 – 80nm) with an inner protein coat and one or two icosahedral capsids. On the basis of differences in morphology and physiological properties this family can be sub-divided into nine genera, the *Orbiviruses* being the largest subgroup (reviewed in Urbano & Urbano, 1994).

Orbiviruses are distinguished from other *Reoviridae* by their ability to multiply in insects and vertebrates, and their sensitivity to low pH conditions, lipid solvents and detergents. Bluetongue virus (BTV) is the type species of this genus (Borden *et al.*, 1971) and is comparable both in morphology and molecular constitution to AHSV (Roy *et al.*, 1994). On the basis of cross-neutralisation studies of the major outer capsid protein, VP2, 25 serotypes have been defined for BTV and 9 for AHSV (McIntosh, 1958; Howell, 1962; Gorman & Taylor, 1985).

1.2.2 Epidemiology

AHSV infects equids, the usual hosts being horses, mules, zebras and donkeys. The disease however is more severe in horses than donkeys or mules (Coetzer & Erasmus, 1994). Occasional hosts include elephants, onager, camels and dogs (after eating infected blood or horsemeat). AHS is not directly contagious, but is transmitted by biting midges (*Culicoides imicola*)(Du Toit, 1944; Moore & Lee, 1972). Venter and co-workers (2000) recently identified *Culicoides botilinos* as another potential vector of AHSV. Moist mild conditions and high temperatures favour the presence of these insect vectors.

AHS is endemic in the central and eastern tropical regions of Africa, from where it spreads regularly to southern Africa and occasionally to northern Africa (Coetzer & Erasmus, 1994). Isolated outbreaks have occurred in Spain (1966, 1987-90), in the Near and Middle East (1959-63) and in Portugal (1989) (Mellor, 1993). The disease not only affects the export of horses from southern Africa, but also has a large impact on the export of zebras to international conservation reserves (Lubroth, 1988).

1.2.3 Pathogenesis

African horsesickness has been allocated Office International des Epizooties (OIE) List A status. The disease caused by AHSV ranges from sub-clinical to acute with mortality rates of 0-100% in horses (reviewed in Burrage & Laegreid, 1994). In 1921 Sir Theiler described four clinical forms of AHS on the basis of time of onset, clinical presentation, mortality and the organs affected (Laegreid *et al.*, 1992a; Laegried *et al.*, 1992b). These are the pulmonary, cardiac, mixed and horsesickness fever forms. The form of AHS seen in infected horses may be as a result of several factors, including the route of infection, tropism of sub-populations of viral particles (Erasmus, 1972) and host immune status, genetic susceptibility and permissivity (Laegried *et al.*, 1993). Laegreid and co-workers, (1993) showed that the viral virulence phenotype is the primary determinant of the form of disease expressed by experimentally infected naïve horses.

Pulmonary AHS or 'dunkop' is an acute disease characterised by high fever, severe pulmonary oedema, pleural effusion and mortality approaching 100% of affected horses. This form has an incubation period of only 4 to 5 days and affects the lungs, spleen, hydrothorax, bronchial and mediastinal nodes. Cardiac AHS or 'dikkop' is subacute and results in pronounced oedema of the subcutaneous and intermuscular tissues of the head and neck. Large amounts of fluid are also present in the pericardial sac, while multifocal haemorrhages of the epi-, endo-, and myocardium are observed. Although some animals do recover, the mortality rate is nevertheless high (50-80%). The incubation period is variable, between 7 to 14 days. The mixed form of the disease is the most common and displays features of both the pulmonary and cardiac forms. This form has a greater than 80% mortality rate and an intermediate incubation period of 5 to 7 days. The mildest form of AHS is the fever form, here affected horses develop a transient subacute fever, some scleral infection and mild depression. The incubation period is 5 to 14 days, and all affected horses recover (Coetzer & Erasmus, 1994; Laegreid *et al.,* 1993; Burrage & Laegreid, 1994).

The principal pathological features of African horsesickness are, therefore, oedema, effusion and haemorrhage. Skowneck *et al.,* (1995) suggest that the loss of endothelial cell (EC) barrier function may be the cause of these prominent features. AHSV has been experimentally shown to infect pulmonary microvascular endothelial cells. Severely affected cells display pronounced cell swelling with some discontinuity of the plasma membrane and loss of structural detail (Laegreid *et al.,* 1992a). Multiple mechanisms may be involved in the development of this pulmonary microvascular leakage. Damage to the EC and loss of integrity of intercellular junctions could result in the loss of EC barrier function and the subsequent development of oedema (Laegreid *et al.,* 1992a; Laegried *et al.,* 1992b). Thus the lethal event in acute AHS, pulmonary oedema, may be related to the infection and damage of pulmonary endothelial cells (Laegried *et al.,* 1992a; Laegreid *et al.,* 1992b). In an experiment conducted by Laegreid *et al.* (1992b) the ability of two virulence variants of AHSV serotype 4 (AHSV-4) to infect EC correlated to the pathogenicity of these variants for horses. The difference in ability of virulence variants of AHSV to cause acute disease may therefore be as a result of differences in their ability to infect and damage EC.

1.2.4 Prevention and control

There is currently no efficient treatment for AHS. In an attempt to control outbreaks, horses in southern Africa are vaccinated annually with a polyvalent attenuated vaccine. Immunisation with this vaccine does however have side-effects, in severe cases fatal enchaphilitis has been reported (reviewed in Coetzer & Erasmus, 1994). The horses may, furthermore, not be exerted for at least three weeks following immunisation. Once an outbreak of AHS has occurred a monovalent vaccine may be administered once the serotype has been identified. The first reverse transcription polymerase chain reaction (RT-PCR) amplification method for serotyping of AHSV was recently developed, and will allow for rapid and reliable identification and discrimination between the nine AHSV serotypes using L2. This is vital at the start of an outbreak to enable the early selection of a vaccine to control the spread of the disease (Sailleau *et al.*, 2000). The only other methods of control include quarantine and slaughter of infected animals, destruction of cadavers, and control of insect vectors (insecticides, repellants and screens). Alternative subunit vaccines for AHSV are currently being investigated.

1.2.5 Molecular biology

AHSV possesses a double stranded RNA (dsRNA) genome composed of 10 segments packaged within the core of a double layered capsid (Bremer, 1976; Huismans, 1979; Mecham & Dean, 1988). The genome segments are of different sizes, three large, designated L1-L3, three medium, M4-M6, and four small, S7-S10. The 5' and 3' non-coding regions of the genome segments range in length from 12 to 35 base pairs (bp) and 29 to 100 bp respectively (Roy *et al.*, 1994). Like BTV, the 5' and 3' end sequences of each segment show partial inverted complementarity and may form secondary structures when in the single stranded RNA (ssRNA) form (Roy *et al.*, 1994). Each genome segment encodes at least one protein. Seven structural (VP1-VP7) and four non-structural (NS1-NS3, NS3A) proteins have been identified (Oellerman *et al.*, 1970; Bremer, 1976; Grubman & Lewis, 1992; Laviada *et al.*, 1993).

1.2.5.1 Proteins of the outer capsid:

The outer capsid of the virus is composed of VP2 and VP5. Lewis and Grubman (1991) confirmed VP2 as the major exposed protein on the virus particle. This protein is encoded by the L2 genome segment and is the most variable AHSV protein with up to 64% variation on the amino acid level across serotypes (Vreede & Huismans, 1994). VP2 contains the major serotype specific and neutralising epitopes of AHSV (Burrage *et al.*, 1993).

Bentley and co-workers (2000) have identified a number of antigenic determinants located in the amino terminal half of VP2. The antigenic structure of AHSV-4 VP2 was recently determined by Martínez-Torrecuadrada and co-workers (2001). Fifteen antigenic sites were identified, of which three induced neutralising antibodies for AHSV-4. These characteristics make VP2 an ideal candidate for a subunit vaccine. Still, the relatively low neutralisation titres found by Martinez-Torrecuadrada *et al.* (2001) make the possibility of producing a synthetic vaccine for AHSV unlikely.

VP5 is encoded by segment M6 and is less exposed on the virion surface than VP2. This protein is present in small amounts in infected cells, and is insoluble and possibly cytotoxic when expressed in insect cells (Roy *et al.*, 1994). Other than its structural role little is known about the function of AHSV VP5. Hassan and co-workers (2001) propose a role for BTV VP5 in virus-cell penetration, based on their finding that purified VP5 was able to permeabilise mammalian and *Culicoides* insect cells, inducing cytotoxicity. Both VP2 and VP5 of AHSV may influence the virulence phenotype expressed by a given virus isolate (O'Hara *et al.*, 1998).

1.2.5.2 Proteins of the inner capsid:

The inner core is composed of two major proteins, VP3 and VP7, and three minor proteins VP1, VP4 and VP6 (Bremer, 1976).

VP3, encoded by L3, is the innermost component of the core and forms a scaffold on which clusters of VP7 trimers are arranged (Hewat *et al.*, 1992; Prasad *et al.*, 1992). Both these proteins are highly conserved amongst serotypes and contain the group specific antigenic determinants. When co-expressed these proteins spontaneously aggregate into empty core-like particles (French & Roy, 1990; Maree *et al.*, 1998). In

contrast to the cognate protein in BTV, AHSV VP7, encoded by S7, is highly insoluble and hydrophobic (Roy *et al.*, 1991). When expressed in insect cells VP7 spontaneously assembles into immunogenic disc-shaped crystals (Chuma *et al.*, 1992). The possibility of using these crystals as antigen delivery systems is currently under investigation (Maree, 2000). Tan and co-workers (2001) recently identified an RGD motif between residues 168 and 170 of BTV VP7 as being responsible for the attachment of the core to *Culicoides* cells.

Little is known of the functions of the minor core proteins VP1, VP4 and VP6 of AHSV, presumably they have similar enzymatic functions to their cognate proteins in BTV. VP1, encoded by L1, is postulated to be the virus replicase-transcriptase, capable of elongating ssRNA (Roy *et al.*, 1994). VP4 expressed from the M4 segment may have guanyl transferase activity for the methylation of the mRNAs during transcription (Roy *et al.*, 1994). VP6, encoded by S9, has strong ds- and ssRNA binding ability and may have helicase activity. This is supported by sequence similarities between VP4 and the helicase enzyme of *E. coli* (Turnbull *et al.*, 1996).

1.2.5.3 Non-structural proteins:

The nonstructural proteins NS1, NS2, NS3/A of AHSV are encoded by the M5, S8 and S10 segments respectively. While NS1 and NS2 are highly conserved, the NS3 proteins are the second most variable AHSV gene products (Van Niekerk *et al.*, 2001b).

AHSV NS1 dimers assemble as helically coiled tubules in the cytoplasm of infected cells and are believed to play a role in virus transport (Huismans & Els, 1979; Urakawa & Roy, 1988; Roy, 1996). The protein contains 16 cysteine residues of which 9 are conserved at the same relative positions in BTV NS1. The residues at position 337 and 340 in BTV NS1 have been shown to be necessary for tubule formation and may be as essential in AHSV NS1. The AHSV protein however has a lower sedimentation rate than BTV NS1 and does not display the same ladder-like structure (Maree & Huismans, 1997, Van Staden *et al.*, 1998).

AHSV NS2 is a major component of the virus inclusion bodies routinely observed in AHSV infected cells. This protein is thought to be involved in recruiting mRNA during

virion assembly at these sites, substantiated by its ability to bind ssRNA (Uitenweerde *et al.,* 1995). NS2 is phosphorylated at specific serine residues and is, in fact, the only viral phosphorylated protein present in AHSV infected cells (Devaney *et al.,* 1988). Theron and co-workers (1994) found that the phosphorylated version of EHDV NS2 bound ssRNA less efficiently than the unphosphorylated form. Another important feature of NS2 is the presence of conserved hydrophobic domains and potential β-turns at the N-terminus of the protein. Zhao *et al.,* (1994) found that the amino terminus of BTV NS2 was essential for ssRNA binding.

The smallest genome segment of AHSV encodes two related proteins NS3 and NS3A, where NS3A is a truncated version of NS3 lacking the 10 or 11 N-terminal amino acids (Van Staden & Huismans, 1991; Mertens *et al.,* 1984; Van Dijk & Huismans, 1988). Initiation of synthesis of these two proteins has been shown to occur at two in-phase AUG initiation codons in the same open reading frame (Van Staden & Huismans, 1991). The presence of the second in-frame start codon is conserved in all 9 serotypes of AHSV and in the S10 genes of other *Orbiviruses* including BTV, epizootic hemorrhagic disease virus (EHDV), Palyam virus, Broadhaven virus (BRD) and Chuzan virus (Van Staden & Huismans, 1991; Moss *et al.,* 1992; De Sá *et al.,* 1994; Jensen *et al.,* 1994; Yamakawa *et al.,* 1999). The sequence and length of the N-terminal extension present in NS3 but not NS3A is, however, variable.

Other regions or features of AHSV NS3 and NS3A conserved in BTV NS3/A and the NS3/A proteins of other orbiviruses include:

- A proline rich region between amino acids 22 and 34
- A highly conserved domain between residues 43 and 92, and
- Two hydrophobic domains (HD) at amino acids 116 to 137 and 154 to 176 (Van Staden *et al.,* 1998).

NS3 has been shown to be present in the membrane components of infected cells associated with virus release (Stoltz *et al.,* 1996). The protein is possibly retained on the cell membrane by the conserved hydrophobic domains and may mediate the egress of the virion from infected cells in the final stages of virus morphogenesis as proposed for BTV NS3 (Hyatt *et al.,* 1993). In the computational conformational model proposed by Van Staden and co-workers (1995) both the hydrophobic regions form transmembrane

domains that span the membrane with the terminal regions located in the cytoplasm. Alternatively only one of the putative transmembrane regions could span the membrane with the amino-terminus orientated in the endoplasmic reticulum (ER) lumen as seen in the analogous protein in rotavirus, NSP4 (Chan *et al.,* 1988). This model was also proposed for the NS3 protein of EHDV, a related orbivirus (Jensen & Wilson, 1995).

Only small amounts of NS3 are detected when expressed as a recombinant in the baculovirus system or in Vero cells infected with AHSV, unlike BTV NS3 which is expressed at higher levels. No NS3A was detected in *Spodoptera frugiperda* (*Sf9*) cells infected with recombinant baculoviruses expressing the full length S10 gene (Van Staden *et al.,* 1995). Expression of NS3A was achieved by the truncation of the full length NS3 gene (Van Staden *et al.,* 1998).

AHSV NS3 was found to be cytotoxic when expressed in the baculovirus expression system in insect cells causing disruption of the cellular membrane and increased membrane permeability (Van Staden *et al.,* 1995; Smit, 1999). Substitution mutations in the hydrophobic domains of NS3 abrogated this cytotoxic effect, while mutations in the proline rich and highly conserved regions had little effect (Van Staden *et al.,* 1998; Van Niekerk *et al.,* 2001a). The cytotoxicity of NS3 appears, therefore to be dependant on its membrane association (Van Niekerk *et al.,* 2001a). No significant difference between the cytotoxicity of NS3 and NS3A in insect cells was observed (Van Niekerk *et al.,* 2001a).

Many aspects of the functioning of, role of and reasons for the conservation of NS3 and NS3A in the viral life cycle remain unclear. As these proteins are the subject of this investigation, a comparison to the cognate proteins in other orbiviruses is, therefore, necessary. In addition, a greater understanding of viral membrane proteins, cytotoxic proteins and proteins expressed from overlapping reading frames is essential. Each of these aspects will, therefore, be discussed in the following sections.

1.3 NS3 RELATED PROTEINS OF VIRUSES IN THE REOVIRIDAE FAMILY

1.3.1 BTV NS3 and NS3A:

The S10 genome segment of BTV encodes NS3 and NS3A from the same open reading frame (ORF)(Van Dijk & Huismans, 1988). Expression of the BTV NS3 gene in the baculovirus system results in the production of both proteins. NS3A expression however is more variable than NS3 and usually occurs at a lower level (French *et al.,* 1989; Bansal *et al.,* 1998).

These proteins, unlike the AHSV NS3 proteins, display a high degree of sequence conservation and homology amongst the BTV serotypes (Hwang *et al.,* 1992). Conserved features of BTV NS3 and NS3A include two hydrophobic domains, two N-linked glycosylation sites (63 and 150), a cluster of six proline residues near the N-terminus, two conserved cysteine residues (137 and 181) and a tyrosine residue at position 159 (Hwang *et al.,* 1992). Sera from sheep infected with homologous and heterologous BTV react with NS3 and NS3A expressed in the baculovirus system, suggesting that these proteins are highly conserved group specific antigens (French *et al.,* 1989).

NS3 and NS3A have been detected in regions of the plasma membrane associated with membrane perturbation (Hyatt *et al.,* 1991). Hyatt and co-workers (1993) found that NS3 and NS3A mediate the release of virus-like particles (VLPs), but not core-like particles (CLPs), from infected cells. These proteins are therefore believed to be involved in the final stages of BTV morphogenesis.

Both the hydrophobic domains (HDs) of NS3 and NS3A are proposed to span the membrane and only the site at amino acid 150 between the transmembrane regions is glycosylated (Bansal *et al.,* 1998). Deletion of part of the first HD (H1) abolishes glycosylation. H1 appears therefore to be responsible for the introduction of the protein into the ER for the addition of the carbohydrate group (Bansal *et al.,* 1998). This glycosylation serves to protect the protein from degradation either before or after cell membrane insertion (Bansal *et al.,* 1998).

Neither BTV proteins have been reported to be cytotoxic. The BTV and AHSV NS3 proteins may have some functional similarity to the NSP4 protein of rotavirus, a member of the *Reoviridae* family.

1.3.2 Rotavirus NSP4:

The non-structural glycoprotein NSP4 is an integral membrane protein localised in the endoplasmic reticulum (ER) of rotavirus infected cells. NSP4 is a protein of 175 amino acids with three internal hydrophobic domains. The second hydrophobic domain acts as the only transmembrane domain while the C-terminal domain lies on the cytoplasmic side and the N-terminal domain in the lumen of the ER (Taylor *et al.,* 1992). NSP4 assembles as homotetramers mediated by α-helical coiled coil structures adopted by the cytoplasmic region between residues 95 and 137 (Maass & Atkinson, 1990; Taylor *et al.,* 1996). These NSP4 oligomers act as receptors for the binding of and translocation across the ER membrane of single-shelled rotavirus particles (SSP) (Meyer *et al.,* 1989). The C-terminal amino acids 161 to 175 of NSP4 facilitate binding to SSP (Au *et al.,* 1993). O'Brein *et al.,* (2000) used solid-phase binding assays to show that the binding of NSP4 to SSP requires at least 17 of the C-terminal 20 amino acids (156 to 175), including the final methionine residue. Binding to the SSP involves VP6, the major structural protein on the surface of rotavirus SSP (Au *et al.,* 1989). NSP4 has also been shown to bind VP4, an outer capsid protein (Mattion *et al.,* 1994). After translocation of the rotavirus SSP, removal of the transient envelope requires the glycosylation of NSP4 (Petrie *et al.,* 1983).

Several of the cytopathic effects seen in rotavirus infection are also evident when NSP4 is expressed alone in mammalian or bacterial cells or when added extracelluarly to cells such as gastrointestinal epithelial cells. Expression of NSP4 from a recombinant vaccinia virus in cultured mammalian cells causes a loss in plasma membrane integrity, profound morphological changes and eventually cell death (Newton *et al.,* 1997). In *Sf9* cells the expression of NSP4 results in a potentially toxic increase in the concentration of Ca^{2+} in the cytoplasm (Tian *et al.,* 1994). Purified NSP4 releases calcein incorporated into liposomes, disrupts microsomes and appears, therefore, to have direct membrane-

destabilising activity (Tian *et al.,* 1996b). NSP4 furthermore induces paracellular leakage in polarised MDCK-1 epithelial cells (Tafazoli *et al.*, 2001).

A membrane-proximal domain of NSP4 (residues 54 to 74) was identified as being important in mediating the cytopathic effects of NSP4 in MA104 cells (Newton *et al.,* 1997). This region is rich in basic amino acids and has the potential to form a cationic amphipathic helix structure, a feature common to many membrane-destabilising proteins (Newton *et al.,* 1997). Browne *et al.,* (2000) confirmed these findings by identifying the domain of NSP4 between residues 48 and 91 as the mediator of internal membrane destabilisation when NSP4 is expressed in *E. coli* cells. Localised disruption of the ER membrane may affect the permeability of the ER and cause an increase in the rate of Ca^{2+} leakage. An increase in intracellular calcium concentration causes cell lysis and the release of viral particles during the final stages of rotavirus morphogenesis (Tian *et al.* 1996a; Ruiz *et al.,* 2000). The oligomerisation of several rotavirus proteins as well as the integrity of rotavirus particles, furthermore, also requires the presence of Ca^{2+} (Ruiz *et al.,* 2000).

Free NSP4 has also been identified as a viral enterotoxin due to its ability to promote diarrhoea, similar to that in virus-induced disease, when administered to infant mice (Ball *et al.,* 1996). A peptide (residues 114 to 135) derived from the cytoplasmic domain of NSP4 has been shown to possess enterotoxic activity (Ball *et al.,* 1996). This enterotoxic activity is possibly mediated by a putative plasma membrane receptor that triggers a phospholipase C-mediated increase in intracellular Ca^{2+}, following binding of NSP4 to human intestinal cells (Dong *et al.,* 1997). An increase in intracellular Ca^{2+} results in enhanced Cl^- secretion in intestinal epithelial cells and diarrhoea. Purified NSP4 and the NSP4 114-135 peptide, however, also released calcein from calcein-loaded liposomes and disrupted microsomes and may, therefore, have direct membrane destabilising activity (Tian *et al.,* 1996b). Zhang *et al.,* (2000) have identified a secreted, cleavage product of NSP4 (aa 112-175) in the early media from rotavirus-infected cells and demonstrated that this peptide was a functional enterotoxin. Browne *et al.,* (2000) suggest that the membrane-destabilising and enterotoxic activities of NSP4 are mediated by different regions of the protein. The mechanism by which endogenous NSP4 causes

an increase in intracellular calcium, furthermore, appears to be unrelated to phospholipase C activity (Tian *et al.*, 1995).

One of the key mechanisms in the development of diarrhea during rotavirus infection in infants, increased intestinal secretion, is therefore stimulated by the intracellular or extracellular action of the rotavirus non-structural protein, NSP4 (Estes *et al.,* 2001). NSP4, in conjunction with other rotavirus proteins, VP3, VP4 and VP7, therefore, plays an important role in determining viral pathogenicity (Kirkwood *et al.,* 1996).

1.4 VIRAL MEMBRANE PROTEINS

The majority of viral membrane proteins (VMP) are complex high molecular weight transmembrane oligomers (Doms *et al.,* 1993). Topologically, membrane proteins can be classified into three groups. Type 1 proteins span the lipid bilayer once with their N-termini orientated towards the ER lumen and their C-termini exposed on the cytoplasmic side. These proteins generally possess a cleavable signal peptide. Type 2 proteins span the membrane in the opposite orientation with their N-termini in the cytoplasm, their C-termini in the ER lumen and typically possess a permanent insertion signal. Type 3 membrane proteins span the membrane more than once (Garoff, 1985).

Viral membrane proteins are typically translated on membrane-bound ribosomes and co-translationallly inserted into the ER in an unfolded form. This co-translational translocation begins in the cytosol with the synthesis of the first hydrophobic segment of a nascent polypeptide, either a signal or transmembrane (TM) sequence. Translocation, folding and conformational maturation of these proteins occurs by the same mechanisms used by cellular membrane proteins. These processes are not spontaneous events but require the participation of numerous folding enzymes and molecular chaperones within the ER (Doms *et al.,* 1993). An increasingly complex macromolecule, the translocon (Walter & Lingappa 1986) is responsible for the transport and biogenesis of proteins at the ER membrane. The translocon recognises potential membrane-spanning domains in membrane proteins, orientates these domains with respect to each other and the ER

membrane and facilitates their integration into the lipid bilayer (Hedge & Lingappa, 1997).

Once inserted into the ER membrane, folding of the various topological domains of a membrane protein occurs under different conditions in different environments, namely the ER lumen, the ER membrane or the cytosol. The ectodomain of the protein folds in the ER lumen and typically comprises the bulk of the protein's mass and all of the carbohydrate moieties and disulphide bonds. The transmembrane domain or domains adopt an α-helical configuration within the hydrophobic interior of the ER membrane, each domain containing at least 20 to 22 hydrophobic amino acids flanked by charged residues. The cytoplasmic domain of the protein folds within the cytosol, presumably according to the same rules that apply to cytosolic proteins and often contains sites necessary for interaction with viral components (Doms *et al.,* 1993).

After folding in the ER, membrane proteins are often assembled into oligomers immediately or during transit in the Golgi apparatus. Oligomerisation may be as a result of disulphide bond formation or, more commonly, non-covalent interactions. For example, coiled-coil domains often stabilise the protein and allow for protein dimerisation where the α-helical coils wind around each other (Bassel-Duby *et al.,* 1985).

Post-translational targeting occurs via the default pathway used by cellular membrane proteins, unless the protein contains additional sorting signals. This pathway is from the ER to the Golgi apparatus, across the Golgi from the cis to the medial and finally the trans components and on to the cell surface. In the Golgi apparatus further modifications such as glycosylation may occur (Hedge & Lingappa, 1997). Once correctly folded, modified, assembled and targeted to the cell membrane the VMP can perform its specific function.

The functions of viral membrane proteins are diverse and include binding viral particles to receptors on the host cell plasma membrane, mediating membrane fusion, directing virus morphogenesis at the budding site and acting as receptor-destroying enzymes for virus release (Doms *et al.,* 1993). A further, novel function for VMP was discovered by Pinto and co-workers in 1992, when it was found that the small integral membrane protein, M2, of Influenza A virus had ion channel activity that allowed the protein to modify the permeability of the membrane into which it was inserted.

Subsequent to this, many small viral membrane proteins have been discovered that modify membrane permeability. This membrane permeabilisation may play a critical role not only in the viral life cycle but also in viral pathogenesis.

1.5 MODIFICATION OF MEMBRANE PERMEABILITY BY ANIMAL VIRUSES:

A number of animal viruses induce membrane permeability changes during infection. These modifications are observed either early during the infection of host cells or later in the replicative cycle following viral gene expression (Carrasco, 1994). Early and late membrane permeabilisation occur via different mechanisms. Early membrane permeabilisation may be as a result of co-entry of toxins with virus particles during infection. Late membrane permeabilisation requires the expression of viral genes and, therefore, one or more viral protein is responsible for the permeability changes observed. A greater understanding of the changes that accompany the disruption of the cell membrane by a viral protein is necessary.

Alterations in the permeability of host cell membranes most commonly enhance the permeability of the cell to monovalent cations with a subsequent decrease in membrane potential (Carrasco *et al.,* 1989). Sodium ions accumulate in the cell while potassium ions leak out. Viral mRNAs translated late during infection are, therefore, adapted with special structures for optimum translation under these altered conditions (Carrasco, 1994). Late virus enduced membrane permeabilisation may also affect the concentration of calcium ions, the pH of infected cells and enhance the permeability of the cell to other compounds for example hydrophilic antibiotics such as Hygromycin B (Carrasco, 1994). Viral gene products involved in the modification of membrane permeability may, furthermore, activate phospholipase C and, possibly, phospholipase A2 leading to the release of choline and the formation of high concentrations of inositol-3-phosphate (IP3) and arachidonic acid in the medium. This, in turn, would lead to the formation of prostaglandins or diacylglycerol and lysophosphatidylcholine which destabilise and permeabilise the membrane (Carrasco, 1994).

1.5.1 Viroporins:

Viroporin is the name given to a family of virus proteins that alter membrane permeability (Carrasco, 1989). In many cytolytic viruses a single gene product or viroporin is the primary cause of the cytopathic effect (CPE) observed. General characteristic structures or structural motifs of viroporins have been identified (Carrasco, 1994):

- Viroporins are generally short proteins of between 50 and 120 amino acids
- Viroporins frequently contain a higher than normal content of leucine and isoleucine residues and a lower overall glycine content
- Viroporins are integral membrane proteins with at least one membrane spanning domain, which may form an amphipathic helix
- Viroporins often occur as oligomers, usually tetramers, that form hydrophilic pores allowing non-specific diffusion of low molecular weight compounds and ions
- Viroporins may contain regions of basic amino acids that could destabilise the lipid bilayer.

Viroporins appear, therefore, to possess activities like some ionophores or membrane-active toxins. The primary function of viroporins is to permit the release of mature virions from infected cells, accompanied by modifications in the morphology and metabolism of the cell. The rotavirus protein NSP4 and the influenza virus protein M2 discussed earlier are examples of viroporins. Other examples include: HIV Env and Vpr proteins (Piller *et al.,* 1996), poliovirus 2BC (Barco & Carrasco, 1998), hepatitis virus A 2C (Jecht *et al.,* 1998), vaccinia virus A38L (Sanderson *et al.,* 1996), togavirus 6K protein (Sanz *et al.,* 1994) and Japanese encephalitis virus NS2B-NS3 (Chang *et al.,* 1999). Some of these proteins, for example NSP4 (Ball *et al.,* 1996) and Vpr (Piller *et al.,* 1998), have been shown not only to alter membrane permeability when expressed in cells as recombinants but also when added extracellularly to cells in purified form. All these proteins play an integral role in the viral replication cycle and possibly in the pathogenesis of the related disease.

1.6 DUAL IN-FRAME TRANSLATION INITIATION

1.6.1 Mechanisms for translation initiation at dual initiation site:

Three mechanism for generating more than one protein from a single mRNA in eukaryotes have been described. These are context-dependant leaky scanning, reinitiation and possibly direct internal initiation (Kozak, 1999). All three mechanisms represent an escape from the constraints imposed by the first-AUG rule in eukaryotes.

Reinitiation involves the production of two proteins from separate ORFs on the same mRNA (Kozak, 1987) and will not be discussed here. Direct internal initiation was first proposed for the translation of the picornavirus polyprotein. In this model the ribosome and other protein factors recognise a nucleotide sequence with significant secondary structure, termed the internal ribosome entry site (IRES), on the substrate mRNA. This internal ribosome entry is believed to facilitate translation initiation at internal start codons on the mRNA independent of the m^7cap structure (Gale *et al.*, 2000). Although interesting, further experimental evidence is necessary to substantiate this hypothesis (Kozak, 1999). Little consensus exists between the IRES sites identified.

Leaky scanning is the most frequently observed mechanism and occurs when the 40S ribosomal subunit bypasses the first AUG and instead initiates at a downstream AUG codon. The absence of a good context around the first AUG is the most probable cause of leaky scanning (Kozak, 1995, Kozak, 1999). Initiation sites in eukaryotic mRNAs usually conform to all or part of the sequence GCCRCCaugG (Kozak, 1987). The most highly conserved positions within this consensus sequence are the purine, usually A, at –3 and the G at +4 (where the AUG codon is numbered +1 to +3). Mutations affecting these two positions strongly impair initiation (Kozak, 1999) while the rest of the sequence contributes only marginally. An AUG codon can, therefore, be classified broadly as weak or strong based on positions –3 and +4.

1.6.2 Examples of dual initiation in viruses:

The use of in-frame AUG codons for the production of long and short protein isoforms appears to be more common in viruses than in cellular mRNAs (Kozak, 1987, refer to Table 1.1). This allows the virus to maximise its genome coding capacity and potentially encode functionally distinct proteins from a common mRNA. Some examples are discussed here to illustrate this.

a. Foot-and Mouth Disease Virus L proteins:

Initiation of translation can occur at more than one start codon in some picornaviruses, probably the best example being the Leader protein in foot-and-mouth disease virus (FMDV). This protein is produced in two distinct forms, Lab and Lb. Lab has an additional 28 amino acids at its amino-terminus compared to Lb and both are encoded from dual initiation sites *in vivo* and *in vitro* (Belsham, 1992). The in-frame AUG is conserved in all 7 serotypes of FMDV. Studies indicate that leaky scanning is the mechanism for translation initiation (López de Quinto & Martinez-Salas, 1999). Expression of the two proteins separately was achieved through mutation of the start codons, however both Lab and Lb appear to function with the same activity and specificity (Medina *et al.,* 1993). Although the additional sequences in Lab are quite variable between different strains of FMDV (Sangar *et al.,* 1987), they therefore appear not to adversely affect the activity of the protein. It is unclear why FMDV has conserved this feature of dual initiation sites. Interestingly, the presence of the AUG corresponding to the second initiation codon, but not the first, is essential for viral replication (Cao *et al.,* 1995, Piccone *et al.,* 1995).

b. Hepatitis B Virus Surface Glycoproteins:

Hepatitis B virus (HBV) is an enveloped DNA virus. The nucleocapsid of this virus is surrounded by a host-derived lipid envelope containing three viral surface proteins, termed S, M and L. These three HBV surface antigens (HBsAgs) are translated from a single ORF from three in-phase start codons. The S protein is the smallest of these polypeptides and the most abundant. The next largest protein, M, contains the S domain

Table 1.1 Examples of genes that produce two proteins from the same open reading frame.

GENE	PROTEIN PRODUCT
Simian virus 40 late 19S mRNA	VP2 and VP3 (Sedman & Mertz, 1988)
Rotavirus SA11 segment 9	37K and 35K (VP7) (Kozak, 1991)
West Nile flavivirus	V2 core proteins (Kozak, 1991)
Dengue (type 3) flavivirus	C and C' (Kozak, 1991)
Foot-and-mouth disease virus	Lab and Lb (Belsham, 1992)
Hepatitis B virus	Human, S,M and L; Duck, S and L (Heermann *et al.*, 1984)
Feline leukemia virus	$gPr80^{gag}$, $gPr65^{gag}$ (Koza, 1991)
Rift Valley fever (bunya)virus	M proteins (Kozak, 1991)
Cucumber necrosis virus	p20 and p21 (Johnston & Rochon, 1996)
Cowpea mosaic virus RNA-M	95K and 105K (Kozak, 1991)
Barley stripe mosaic virus	βb and βb' (Petty & Jackson, 1990)
African horsesickness virus segment 10	NS3 and NS3A (Van Staden & Huismans, 1991)
Bluetongue virus segment 10	NS3 and NS3A (Mertens *et al.*, 1984; Van Dijk & Huismans, 1988)
Herpes simplex virus, thymidine kinase gene	43K, 39K and 38K (Haarr *et al.*, 1985)
Epstein-barr virus, BBXLF1	Deoxythymidine kinase (Holton & Gentry, 1996)
Rabies virus, P gene	P1 – P5 (Chenik *et al.*, 1995)
Rice Dwarf phytoreovirus, segment S11	P11a and P11b (Suzuki *et al.*, 1991)
Chicken brain cells	Creatine kinase (Kozak, 1991)
Human tumor cell lines	N-myc (Kozak, 1991)
Yeast	CCAse: Small (Kozak, 1991)
Yeast	MOD5p: Small (Kozak, 1991)
Plants	AlaRS: Small (Kozak, 1991)
Epithelial cells, annexin XIII	Annexin XIIIa and XIIIb (Lecat *et al.*, 2000)
Rat liver, fumarase	Cytosolic and mitochondrial fumarases (Suzuki *et al.*, 1992)
Aradopsis thaliana, valyl- and threonyl-tRNA synthetases	Mitochondrial and cytosolic forms (Soiciet *et al.*, 1999)
p55PIK, regulatory subunit of phosphoinositide 3-kinase	54kDa and 50kDa (Xia & Serrero, 1999)

and a hydrophilic extension of 55 amino acids arising from the translation of the PreS2 domain. The L protein is the largest of the HBsAgs and contains a domain of 119 residues encoded from the PreS1 region in addition to the 55 PreS2 amino acids and the N-terminal S domain (Heermann *et al.,* 1984). The S protein is the major constituent of infectious virions. The L protein is necessary for virion maturation, while S also plays a role in this but alone is not sufficient (Ueda *et al.,* 1991). The M protein appears to be dispensable and is not found in all HBV subtypes (Santantonio *et al.,* 1992; Summers *et al.,* 1991). Virions contain differential amounts of the three surface proteins. Initiation codon context and leaky scanning are the prominent factors involved in the differential synthesis of the HBsAgs and represent a type of translational control (Sheu & Lo, 1992; Gallina *et al.,* 1992). Poisson and co-workers (1997) demonstrated that both PreS1 and S domains interact with core particles and propose that this double interaction may be necessary for virion morphogenesis. PreS1 peptide binds specifically to plasma membranes derived from human liver cells and may be responsible for receptor binding during infection (Lin *et al.,* 1991). The PreS2 domain found in the L and M proteins contains a cell permeable motif that may play a crucial role in the internalisation of the viral particle (Oess & Hildt, 2000). The S and L proteins are, therefore, both essential and functionally distinct in HBV.

c. Rabies Virus Phosphoprotein:

The P gene of Rabies virus encodes five related proteins (P1-P5) from alternative in-frame AUG codons *in vivo* (Chenik *et al.,* 1995). Mutational analysis identified leaky scanning as the mechanism for translation initiation. Immunofluorescence studies indicated that these P products were found in the cytoplasm of transfected cells, however, the proteins initiated from the third, fourth and fifth AUG codons were found mostly in the nucleus. The N-terminal deletions could remove a cytoplasmic retention signal or expose a nuclear localisation signal. However, no such domains were found by sequence searches of the P gene. The migration of these small proteins may simply be due to passive diffusion, although deletion of the 120 C-terminal amino acids of the P protein did not result in migration to the nucleus. This nuclear localisation may have biologic effects on virus multiplication (Chenik *et al.,* 1995).

d. Simian Virus 40 VP2 and VP3:

Structural proteins VP2 and VP3 of Simian Virus 40 (SV40) are encoded within the same open reading frame by dual in-frame initiation so that the amino acid sequence of VP3 corresponds to the C-terminal two-thirds of VP2 (Sedman & Mertz, 1988). VP2 and VP3 are found in infected cells at a steady-state ratio (Dabrowski & Alwine, 1988). Leaky scanning and not proteolytic processing was found to be the mechanism for VP2 and VP3 synthesis. Good and co-workers (1988) propose that the ratio of synthesis of VP2 relative to VP3 is largely determined by the efficiency with which ribosomes initiate translation at the two start codons, and so may represent a means of translational regulation. SV40 appears, therefore, to co-ordinate the expression of pairs of genes so that each structural protein is present at the correct molar ratio for virion production (Sedman & Mertz, 1988).

1.6.3 Examples of dual initiation in eukaryotes:

The use of alternative translation initiation at two in-phase AUG codons is a less common occurrence in cellular mRNAs, probably due to less stringent genome size restrictions. Some examples where this has been found to occur are: rat liver fumarases (Suzuki *et al.,* 1992), valyl- and threonyl-tRNA synthetases (Souciet *et al.,* 1999), and annexin XIII (Lecat, 2000) (Table 1.1). An interesting similarity between some of the examples mentioned here is that the two protein isoforms are differentially localised, due to the presence of either retention, secretion or signal sequences at the N-terminus of the longer isoform. This dual targeting seems to be one of the mechanisms used by eukaryotic cells to localise proteins to different cellular compartments where the same function is required (Small *et al.,* 1998).

1.6.4 Functional differences in protein isoforms:

From the examples discussed above, it is clear that the use of alternate in-frame start codons represents a general mechanism used in viruses for producing related polypeptides with similar but distinctly different biological properties. This would increase

the potential complexity of protein function with minimal redundancy of genetic information (Haarr *et al.,* 1985). Increased complexity and functional differentiation could, for example, be achieved in the following ways:

- Through oligomerisation where the presence of different polypeptide monomers could extend the range of potential complexes formed.
- Through a functional domain in the N-terminal region that could for example control basal protein turnover, the localisation of the full-length protein or even change the activity or specificity of the protein.
- Through alterations in the ratio of full-length to shorter proteins, thus offering an additional level of regulation.
- Through the usage of the second AUG in a cell-type dependant manner.

The simplicity of this mechanism allows viruses to sample new versions of a polypeptide as they evolve.

1.7 AIMS OF THIS STUDY

The wide conservation of the second start codon in the NS3 gene of orbiviruses particularly in AHSV, indicates a possible distinct role for NS3 and NS3A in the viral life cycle. When expressed in insect cells as baculovirus recombinants, both proteins have been shown to be cytotoxic, causing membrane permeabilisation and cell death. This appears to be unique to the NS3 proteins of AHSV when compared to the cognate proteins of other orbiviruses. NS3 and NS3A of AHSV may be more closely related, on a functional level, to the NSP4 protein of rotavirus. In fact, NS3 and NS3A appear to share many of the characteristics of the lytic viroporin proteins discussed and may play an important role in the virulence and pathogenicity of AHSV. The effect of the AHSV NS3 proteins on mammalian cells has, as yet, not been studied.

The aims of this study were to characterise and compare the NS3 and NS3A proteins of AHSV, in terms of their cytotoxicity and localisation in infected cells. This implies elucidating an approach for distinguishing between these almost identical proteins in AHSV infected cells. In order to achieve these aims the following aspects were investigated:

1. The effect, if any, of extracellular NS3 on the membrane permeability of mammalian cells (Chapter 2).

2. A comparison of the cytotoxic effect of individual and co-expression of NS3 and NS3A in insect cells (Chapter 2).

3. The individual expression of NS3 and NS3A and large scale purification of these proteins (Chapter 2).

4. The production of polyclonal antibodies directed against the N-terminal amino acids of NS3, for use in the investigation of the localisation of NS3 in AHSV infected cells (Chapter 3).