CHAPTER 1: LITERATURE REVIEW

1.1 INTRODUCTION

In 1898 researchers identified infectious agents that were smaller than the smallest known bacteria. Viruses are obligate intracellular parasites that use the machinery of the host cell to replicate. They consist of a DNA or RNA genome surrounded by a protein capsid, which may be surrounded by an envelope in some viruses (Fields, 1998, Gale *et al*., 2000). Virus classification rests largely on the characteristics of the viral genome and virus structure.

Numerous viruses perform virion associated enzymatic activities, which vary according to the strategy for replication of their nucleic acid. Genome transcription is an important stage in the life cycle of a virus as the product of this process must be recognized by the machinery of the host cell for the production of viral proteins necessary for genome replication and the assembly of progeny virions. In the case of viruses with dsRNA genomes, host cells do not have the endogenous enzymes for transcribing from a dsRNA template. Also, dsRNA is never released into the cell because it would evoke a cellular defence response (Bamford, 2000). As a result, dsRNA viruses produce their own enzymes to synthesize mRNA transcripts within the core particle (Lawton *et al*., 2000).

It is thought that the mechanisms involved in the production of mRNA are similar in most viruses having dsRNA genomes. Functional similarities include a RNA-dependent RNA polymerase that produces single-stranded mRNA transcripts from the dsRNA genome as well as replicating the viral genome from single-stranded RNA templates. Viruses infecting eukaryotes have a capping enzyme for synthesizing 5' capped mRNA for translation by the eukaryotic machinery (Lawton *et al*., 2000). Double-stranded RNA viruses require unwinding of the genome segments by a helicase prior to transcription (Kadaré and Haenni, 1997).

The International Committee for the Taxonomy of Viruses (ICTV) recognizes eight families of dsRNA viruses (Mertens, 2004). This includes the family Reoviridae, members of which infect mammals, other vertebrates, insects and plants. Within the Reoviridae family is the genus *Orbivirus* (Levy *et al*., 1994). Virus species within the *Orbivirus* genus that are known to infect equids include African horsesickness virus (AHSV), equine encephalosis virus (EEV) (Mertens *et al.*, 2000) and Peruvian horse sickness virus (PHSV) which was recently assigned to the genus.

African horsesickness virus causes an infectious but non-contagious disease of equines (Coetzer and Erasmus, 1994). Mortality rates in naive populations of horses can exceed 98% (Mertens *et al*., 2000). It is accordingly considered to be one of the most lethal diseases of horses and has been given Office International des Epizooties (OIE) list A

status. Although mortality of horses as a result of AHS occurs annually in South Africa, major epidemics of the disease also occur (Coetzer and Erasmus, 1994).

In BTV, the prototype virus of the *Orbivirus* genus, inner core protein VP6 has been identified as the dsRNA helicase of the virus (Stäuber *et al*., 1997). The focus of this study is an investigation of the primary structure and nucleic acid binding characteristics of the proposed helicase of African horsesickness virus (VP6). In order to provide a background to the study, the literature review will be focussed on aspects of African horsesickness virus and viral helicase activities.

1.2 FAMILY REOVIRIDAE

The viruses classified within the family Reoviridae have dsRNA genomes. The name Reovirus comes from respiratory enteric orphan virus, which is the name proposed for a group of viruses originally classified as echo virus 10 (Sabin, 1959). Gomatos and Tamm (1963) reported that these viruses have a dsRNA genome. There are 11 genera within this family (*Orthoreovirus, Orbivirus, Cypovirus, Aquareovirus, Rotavirus, Coltivirus, Fijivirus, Phytoreovirus, Seadornavirus, Mycoreovirus* and *Oryzavirus*). A new genus has been proposed for insect reoviruses as well as some unassigned viruses (Mertens, 2004). They share specific characteristics despite an enormous diversity in hosts, transmission mechanisms, geographical distribution and pathological outcomes. Members of the family *Reoviridae* share numerous structural properties. These properties include: isometric virions with a 60 – 80 nm diameter; an inner protein coat with one or two icosahedral capsids; an outer capsid which may appear incomplete or missing altogether leaving a naked core. The core consists of the inner protein coat and the genome and has transcriptase activity. The virion has a characteristic molecular weight and buoyant density in CsCl (Urbano and Urbano, 1994).

The dsRNA genome is segmented, consisting of 10 – 12 segments. Genetic recombination by means of reassortment of genome segments occurs between closely related viruses (Gorman and Taylor, 1985; Urbano and Urbano, 1994). Virus classification and taxonomy has recently been reviewed. In this review, parameters have been outlined for the identification of virus species within the genera of the family *Reoviridae* (Mertens *et al*., 2000). The capacity for reassortment of genome segments is the primary determinant of virus species (Calisher and Mertens, 1998; Mertens *et al*., 2000). In the absence of evidence of segment reassortment, serological comparisons are usually used to investigate the relatedness of virus isolates. In some cases, such as bluetongue virus and African horsesickness virus, the serogroup specific antigen, VP7, is used for these serological comparisons. There is a list of species parameters for the genus that is used to identify the members of the same or distinct virus species or genera (Mertens *et al*., 2000).

Viral replication, which occurs in the cytoplasm, involves conservative transcription of each genomic segment into full-length plus strands which are capped and methylated

by endogenous enzymes. These function as mRNAs for viral protein synthesis or templates for progeny viral RNA (Urbano and Urbano, 1994). Members of the family *Reoviridae* are clustered into genera on the basis of minor structural differences and major extrinsic properties such as host range. These genera consist of viruses isolated from terrestrial and nonterrestrial vertebrates and invertebrates as well as plants (Calisher and Mertens, 1998). A large number of varied viruses are grouped under the genus *Orbivirus* (Urbano and Urbano, 1994).

1.3 ORBIVIRUSES

The outer shell of orbiviruses has no discernible capsomeric structure, while the inner shell consists of 32 ring shaped capsomeres arranged in icosahedral symmetry (Oellermann *et al*., 1970; Gorman and Taylor, 1985; Gould and Hyatt, 1994). The name originates from the Latin word *orbis* which means a ring or circle (Gorman and Taylor, 1985; Gould and Hyatt, 1994). Although orbiviruses are not a homogeneous group, recognized serological groups contain genomes consisting of 10 dsRNA segments (Verwoerd *et al.*, 1970; Gorman and Taylor, 1985).

Orbiviruses are able to multiply in insects and in vertebrates (Gorman and Taylor, 1985). There are 20 serogroups or species of orbiviruses as well as 12 tentative species (Mertens, 2004). The classification of viruses within this genus was originally determined by serologic tests (Urbano and Urbano, 1994) and later supported by molecular characteristics (Calisher and Mertens, 1998). Orbiviruses include pathogens of man (Changuinola virus, Corriparta virus, Great Island virus and Lebombo virus); domestic and wild animals (bluetongue virus, African horsesickness virus, Equine encephalosis virus, Peruvian horse sickness virus, Epizootic haemorrhagic disease of deer virus, Orungo virus, Wongorr virus, Wad Medani virus and Palyam virus), birds (Chenuda virus, Ieri virus and Umatilla virus), bats (Chobar gorge virus) and marsupials (Wallal virus and Warrego virus) as well as Eubenangee virus and St Croix River virus with unknown vertebrate hosts (Gorman, 1979; Calisher and Mertens, 1998; Mertens *et al*., 2000). The type species of the genus is bluetongue virus (BTV). Orbiviruses are transmitted by *Culicoides* midges, ticks, phlebotomine flies and anopheline and culicine mosquitoes (Calisher and Mertens, 1998).

1.4 AHSV AND BTV EPIDEMIOLOGY, TRANSMISSION AND GEOGRAPHICAL DISTRIBUTION

Pitchford and Theiler proposed in 1903 that AHS may be transmitted by biting insects as they found that horses housed in mosquito-proof enclosures could be protected against infection (Coetzer and Erasmus, 1994). In 1944, it was demonstrated that certain species of the biting midge *Culicoides* transmit bluetongue virus. It was further proposed that horsesickness may be transmitted in the same fashion (Du Toit, 1944). AHSV and BTV cause annual outbreaks of disease in the southern hemisphere. This is

characteristic of many *Culicoides*-transmitted orbiviruses including EHDV (Mellor and Boorman, 1995).

C. imicola has been considered the most important stock associated species in summer rainfall areas of southern Africa and the only proven vector of BTV and AHSV in this subregion (Burrage and Laegreid, 1994; Venter and Meiswinkel, 1994; Meiswinkel, 1997). Various regions of South Africa have been found to be free of *C. imicola*, these include Port Elizabeth (Meiswinkel, 1997) and the eastern Orange Free State (Venter and Meiswinkel, 1994). In an outbreak of AHS in the cooler, mountainous region of South Africa between February and May of 1998, of twenty seven species of Culicoides captured in the area, AHSV was only isolated from *C. bolitinos* (Meiswinkel and Paweska, 2003). In a survey of *Culicoides* species in Portugal, although *C. imicola* was found to be the most abundant summer vector, *C. obsoletus* and *C. pulicaris* also have a wide distribution in the area. Capela *et al.* (2003) suggest a role for these species in the transmission of BTV and AHSV. BTV-9 caused outbreaks in Bosnia Herzegovina in 2002 which is further north than the range of *C. imicola* suggesting a role for *C. pulicaris* and/or *C. obsoletus* which are abundant across most of northern Europe (Takamatsu *et al.*, 2003). Temperature has been shown to play a role in infection rate and virogenesis in *Culicoides* species with a decline in replication at lower temperatures. Infectious viruses persist, however, and latent virus replication occurs in flies transferred to higher temperatures (Bradley *et al*., 1995; Wellby *et al*., 1996). The global distributions of BTV and AHSV are limited to geographical conditions where competent vectors are present and to seasons climatically favourable for vector activity (Mellor and Boorman, 1995). The disease disappears after the first frosts (Mellor 1993; Barnard, 1998). Some evidence has been presented suggesting that emergence of AHS from its enzootic zones may be due to long range flights of infected arthropod hosts (Mellor 1993). The reappearance of AHS in successive years is often caused by different serotypes of AHSV. This suggests that outbreaks are caused by new introductions of the virus and not reactivation of over wintering virus (Barnard, 1998). Changes in the global distribution of arthropod vectors of BTV and AHSV may result in an alteration in the distribution of the viruses (Mellor and Boorman, 1995). Viraemia in the mammalian hosts lasts for a maximum of one hundred days. If winter conditions last in excess of one hundred days, BTV and AHSV was not expected to survive from one year to the next. BTV and AHSV have been known to survive the winter in the absence of adult insect vectors in western Turkey, Corsica/Sardinia, Calabria (Italy), Serbia and Kosovo (approximately two years after previous reports of the disease). In these locations, *Culicoides* are absent for several months each year when severe winter conditions prevail. Takamatsu *et al*. (2003), propose that sheep $\gamma\delta$ T-cells that are continuously infected with BTV can be converted to lytic infection following interaction with skin fibroblasts or an antibody to $\gamma\delta$ T-cell-restricted surface protein WC-1. When *Culicoides* midges feed, the feeding area becomes inflamed. Large numbers of $\gamma\delta$ T-cells move to the site of inflammation. The interaction of $\gamma\delta$ T-cells that are continuously infected with BTV and skin fibroblasts would result in an increase in the number of viruses at the

feeding areas. This may increase transmission of BTV to the insect vector. Takamatsu *et al.* (2003), suggest this as a new mechanism that may account for BTV over wintering in the absence of vector insects. White *et al*. (2005), have proposed a mechanism for overwintering of BTV in their *Culicoides* hosts. They detected BTV RNA of segment 7 and to a lesser extent segment 2 RNA in larvae and pupae of *C. sonorensis.* They suggest that BTV may overwinter in the invertebrate host with possible down regulation of segment 2 expression during continuous virus infection.

BTV is distributed world wide between the latitudes 35°S and 40°N. Although bluetongue has occurred in Europe it is not endemic to the continent (Mellor and Boorman, 1995). A serological survey for BTV in Kazakhstan showed seropositive animals at the most northern limit of the BTV distribution range (Lundervold *et al.*, 2003). Vertebrate hosts for BTV include domestic and wild ruminants such as cattle, sheep and goats as well as camels, elephants and predatory carnivores (Calisher and Mertens, 1998; Mertens *et al*., 2000).

AHSV is distributed across sub-Saharan Africa with an enzootic band from Senegal and Gambia to Ethiopia and Somalia. AHS occurs in South Africa and has occurred in Egypt (Mellor, 1993; Mellor and Boorman, 1995). Although thought to be confined to Africa, outbreaks of AHSV have occurred in the Middle East, S.W. Asia, India, Cyprus and Spain. Such outbreaks are thought to be the result of increased movement of horses (Powell, 1985, Mellor and Hamblin, 2004). Low levels of antibodies against AHSV have been detected in free-living elephants. This is, however, thought to be the result of frequent exposure to infected biting midges. Significant levels of neutralizing antibodies against AHSV have not been detected in elephants and they are accordingly not regarded as a source of AHSV (Barnard *et al*., 1995). Dogs contract a highly fatal form of AHS after natural infection or ingestion of infected horse meat (Coetzer and Erasmus, 1994). They are also thought to be an unlikely reservoir for African horse sickness virus (Braverman and Chizov-Ginzburg, 1996). The most susceptible hosts for AHSV in southern Africa are horses, donkeys and zebra. The most likely candidate for a game vertebrate reservoir is zebra which are present in areas endemic for AHSV (Barnard, 1998).

1.5 ORBIVIRUS INFECTION

Bluetongue virus particles adsorb to and penetrate the cells to be infected by means of a pinocytotic vesicle (Lecatsas, 1968). From 8 h.p.i., virus specific structures such as virus inclusion bodies (VIB), virus specific tubules and virus particles may be observed in infected cells (Lecatsas, 1968; Hyatt *et al*., 1989). Virus is released from infected cells by means of budding and extrusion through the cell membrane. Released particles enter the cells by endocytosis (Hyatt *et al*., 1989). BTV is thought to multiply in the regional lymph nodes following the bite of an infected midge. From there it spreads to the rest of the body. The development of characteristic viral structures has shown that replication occurs primarily in the endothelial cells, pericytes of capillaries and small

blood vessels. Cytopathic changes in cells eventually result in hypoxia, oedema and haemorrhaging (Verwoerd and Erasmus, 1994).

In the case of AHSV, regional lymph nodes appear to be the site of primary replication. AHSV spreads throughout the animal by means of viraemia (Burrage and Laegreid, 1994). The target organs are the lungs and lymphoid tissue throughout the body (Coetzer and Erasmus, 1994). Theiler initially recognized the association of AHSV with erythrocytes (Henning, 1956). As a result of severe injury to endothelial cells, it seems that this is an important site of secondary replication. Virulence appears to be associated with the ability to cause endothelial damage (Burrage and Laegreid, 1994).

1.6 AFRICAN HORSESICKNESS PATHOGENESIS

The incubation period and severity as well as overall outcome of the disease are determined mainly by virus virulence and susceptibility of the animal. AHSV may cause severe pathological symptoms in the vertebrate host (Mertens *et al*., 2000) that can progress rapidly from first symptoms to death (Koekemoer and van Dijk, 2004). Serotype 9 is considered slightly less virulent than serotypes 1 – 8 (Coetzer and Erasmus, 1994). There are four distinct clinicopathological syndromes of AHS described by Theiler in 1921. In order of severity: fever form, subacute cardiac form (“dikkop” or “thick head” where subcutaneous swelling of the head is present), acute or mixed form and pulmonary form (“dunkop” or “thin head” where subcutaneous swelling is absent). The different forms vary in mortality between 0 and 100% (Henning, 1956; Mellor, 1993; Burrage and Laegreid, 1994; Coetzer and Erasmus, 1994). Oedematous changes and effusions into body cavities, especially the lungs, and visceral and serosal haemorrhages indicate endothelial damage and are often found in fatal cases of AHS (Coetzer and Erasmus, 1994).

1.7 DETECTION OF AHSV IN INFECTED AND VACCINATED HORSES

Methods of detection of AHSV include the use of viral antigens or nucleic acids. In the case of antigens, the use of anti-AHSV antiserum is serogroup specific. It lacks sensitivity with only 50% success in a trial (Laegreid, 1994). In another study utilizing immunological screening, the earliest serological markers corresponded mainly to VP5, VP6 and NS2 and may potentially be used as group-specific diagnostic reagents (Martinez-Torrecuadrada *et al*., 1997). RT-PCR detection based on the nucleic acid sequence of the gene encoding NS2 (segment 8) was found to be 100% successful in a trial using DNA extracted from blood or tissue. RT-PCR of the NS2 gene was also found to be serogroup specific (Laegreid, 1994; Stone-Marschat *et al.*, 1994). Single tube RT-PCR of segment 7 of AHSV has been proposed as a mechanism for diagnosis of AHSV. Using AHSV specific primers the authors were able to detect all 9 serotypes (Zientara *et al*., 1994; Zientara *et al*., 1995). Koekemoer and van Dijk (2004), developed

a rapid serotyping procedure based on a single RT-PCR of AHSV segment 2. A universal primer set allowed amplification of a short fragment at the 5' terminal of segment 2 of all 9 AHSV serotypes. DsRNA isolated from tissue or organ samples from infected horses was used to produce a serotype-specific probe for reverse line blot hybridization to identify the AHSV serotype responsible for the infection. This procedure would be extremely effective in the case of an AHS virus outbreak where a rapid response is required to control the outbreak.

Purified VP2 of AHSV may be used as a diagnostic tool to detect AHSV antibodies in field samples, serotyping and evaluation of protective antibody levels in horses (Martínez-Torrecuadrada *et al*., 1994).

Some 80 – 90% of horses vaccinated with inactivated purified AHSV did not develop antibodies against NS3 of AHSV expressed as a fusion protein in *Escherichia coli* cells. Laviada *et al*. (1995) have suggested the usefulness of NS3 in discriminating between infected and vaccinated horses.

1.8 ORBIVIRUS MORPHOLOGY

Most of the initial knowledge of the molecular biology of orbiviruses arose from research conducted on BTV. Verwoerd *et al.* (1972), found that as in the case of reovirus, seven polypeptides were present in the capsid of the bluetongue virus. Of these, 3 are minor proteins (VP1, VP4 and VP6) which are located in the inner core. The inner core is made up of two major proteins VP7 and VP3. The outer capsid consists of major proteins VP2 and VP5. There are at least 3 non-structural proteins, NS1, NS2 and NS3 (Verwoerd *et al*., 1972; Gould and Hyatt, 1994).

There are three distinct BTV particles: virions, cores and subcores. The virion is an icosahedral particle with a blurred appearance distinct from reovirus and rotavirus (Huismans and van Dijk, 1990; Roy, 1996). Removal of the outer capsid protein layer gives rise to core particles. The core particles are characterized by 32 capsomeres arranged in icosahedral symmetry with a triangulation number of T-3 (Els and Verwoerd, 1969). Immediately after infection, most of the virions are uncoated to core particles, following which a large number are converted to subcore particles during further infection (Huismans *et al*., 1987b). The subcore particle is thought to be the scaffold for the assembly of the capsomeres (VP7) (Burroughs *et al.*, 1995). According to Grimes *et al*. (1998), the core capsid is separated into two layers: a thin inner shell made up of 120 triangular plates of VP3 which is made rigid by the addition of 260 trimers of VP7 (Prasad *et al*., 1992; Stuart *et al*., 1998). The core particle is associated with an RNA-dependent RNA polymerase activity (Verwoerd and Huismans, 1972; Roy, 1992). The subcore particles consist of one major protein, three minor proteins and dsRNA (Huismans and van Dijk, 1990).

Electron microscope evidence showed the similarity between BTV and AHSV (Oellermann *et al*., 1970). Both the BTV and AHSV genomes consist of 10 dsRNA

segments (Verwoerd, 1969; Verwoerd *et al*., 1970). The segments are designated 1 to 10 in order of decreasing size on a polyacrylamide gel (Huismans and van Dijk, 1990) or L (large) 1 – 3, M (medium) 4 – 6 and S (small) 7 – 10 (Roy, 1992). Each genome segment codes for at least one polypeptide (Mertens *et al*., 1984). The coding assignments are summarized in table 1.1.

Table 1.1: Coding assignments for BTV and AHSV.

Segment	Size in base pairs		Protein [a, b]	Size in amino acids		Location [a, b]	Estimated number / particle (BTV)	Function [a]
	BTV-10 [b]	AHSV		BTV-10 [a,b]	AHSV			
1 (L1)	3954	[c]3965$_{(9)}$	VP1	1302	1253[c]	Inner core	~12[q]	RNA polymerase
2 (L2)	2926	[f]3205$_{(9)}$ [n]3221$_{(3)}$	VP2	956	1020[f] 1057[n]	Outer capsid	180[b,q]	Outer capsid protein, serotype specific antigen, cell entry
3 (L3)	2772	[d]2792$_{(6)}$	VP3	901	842[d]	Core	120[q]	Scaffold for VP7
4 (M4)	[o]2011	[h]1978$_{(3)}$	VP4	654	599[h]	Inner core	20[q]	Capping enzyme and guanylyl transferase
5 (M5)	1769[q]	[g]1748$_{(6)}$	NS1	552[m,q]	516[g]	Infected cell	NA	Nonstructural protein tubules [k,g]
6 (M6)	1638[q]	[d]1564$_{(6)}$	VP5	526[q]	518[d]	Outer capsid	360[q]	Structural protein – outer capsid, interacts with core
7 (S7)	1156	[d]1050$_{(6)}$	VP7	1156	280[d]	Core	780[b,q]	Group specific antigen
8 (S8)	1124[p]	[e]1166$_{(9)}$	NS2	357[p]	356[e]	Infected cell	NA	Binds ssRNA, found in VIBs, involved in virus replication
9 (S9)	[p]1046	[i]1169$_{(3/6)}$	VP6	328	369[i]	Inner core	60[q]	Binds ssRNA and dsRNA, NTPase and helicase
10 (S10)	822	[j]758$_{(3)}$	NS3	229	217[j]	Infected cell	NA	Virus release
			NS3A	216	207[j]		NA	

a) Huismans and van Dijk, 1990; b) Roy, 1992; c) Vreede and Huismans, 1998, d) Williams *et al*., 1998, e) van Staden *et al*., 1991, f) Venter *et al*., 2000, g) Maree and Huismans, 1997, h) Van den Bout and Huismans, unpublished data, I) Turnbull *et al*., 1996, j) van Staden and Huismans, 1991 k) Roy, 1996 l) Hall *et al*., 1989, m) Lee and Roy, 1987, n) Vreede and Huismans, 1994 o) Yu *et al*., 1987 p) Fukusho *et al*., 1989 q) Mertens, 2004.

1.9 STRUCTURE AND FUNCTION RELATIONSHIPS OF ORBIVIRUS GENES AND GENE PRODUCTS

In the last 15 years a wealth of information about the structure and function of the *Orbivirus* gene products has accumulated. The genomes of BTV and AHSV and other orbiviruses have been cloned and sequenced and the proteins expressed in baculovirus or other expression systems. High resolution X-ray crystallography has elucidated structure-function relationships of *Orbivirus* genes and gene products. The following section will highlight the most important aspects of these relationships.

1.9.1 THE OUTER CAPSID: VP2 AND VP5

Iwata *et al*. (1992) investigated the evolutionary relationships among the four major capsid proteins of African horsesickness virus, bluetongue virus and Epizootic haemorrhagic disease virus. Of the four capsid proteins, VP2 is the most variable with only 19-24% identical amino acids. As it is found on the outer capsid, it may be under immune pressure so that it evolves to accommodate host responses. Amino acid sequences of VP2 proteins in BTV vary between 40 and 70% conservation between different serotypes (Fukusho, 1987). VP2 of AHSV serotypes 3 and 4 were found to have 50.5% identity with two regions of high variability (Vreede and Huismans, 1994). Multiple alignment of the VP2 amino acid sequences for all nine serotypes of AHSV showed that homology between serotypes varied from 71.4% between AHSV-1 and AHSV-2 and 47.6% between AHSV-2 and AHSV-9 (Potgieter *et al*., 2003). VP2 contains serotype specific determinants for the virus (Huismans and Erasmus, 1981; Mecham *et al*., 1986; Roy, 1992).

In AHSV it was found that the major antigenic domain of VP2 is located in the central domain, with neither the N- half nor C-terminal regions demonstrating immunogenicity (Martínez-Torrecuadrada and Casal, 1995). Using truncated peptides of VP2 of AHSV, Venter *et al*. (2000) demonstrated the presence of a strong linear epitope in a large hydrophilic domain between amino acids 369 and 403. A filamentous phage library was used to determine sequences that may form part of a discontinuous neutralization epitope in VP2 of AHSV-3. Antigenic regions were compared with corresponding regions on three other serotypes revealing regions which may potentially be used for serological discrimination between AHSV serotypes (Bentley *et al*., 2000).

Huismans *et al*. (1987c) showed that BTV VP2 induced neutralizing antibodies in sheep which provided full protection against virulent challenge. Stone-Marschat *et al*. (1996), inoculated horses with a vaccinia construct of segment 2 cDNA. They found that immunized horses developed serum neutralizing antibodies and were clinically normal after virulent challenge. These results demonstrate that VP2 may be used alone to elicit a protective response and suggest that a subunit vaccine based on AHSV VP2 alone may be effective.

The main function of the outer capsid is cell attachment and penetration (Grimes *et al*., 1998). Hassan and Roy (1999), determined the oligomeric nature of VP2 of BTV. Their results show that most purified VP2 protein forms dimers and to a lesser extent trimers. Purified VP2 was found to have virus haemagglutinin activity. VP2 of BTV is responsible for entry into mammalian cells and may be responsible for BTV transmission by the *Culicoides* vector to vertebrate hosts (Hassan and Roy, 1999).

VP5 which is largely unexposed on the capsid surface shows 43 – 45% identical amino acids between the different orbiviruses (Iwata *et al*., 1992). Although VP5 is located on the outer capsid it does not appear to have any neutralizing activity when applied on its own. However, when used in conjunction with VP2, vaccinated sheep demonstrate higher levels of neutralizing antibodies to BTV (Roy, 1992; DeMaula *et al*., 2000). This indicates an involvement of VP5 in protection and virus neutralization. Roy (1992) proposed that VP5 may enhance the immune response as a result of interaction with VP2. O'Hara *et al.* (1998), have suggested a role for VP5 in expression of the intermediate phenotype and thus control of virulence. They also propose that VP2 which has a role in cell entry may influence virulence by affecting tissue or cell tropism. VP5 of BTV was found to bind to mammalian cells but was not involved in cell entry. VP5 induces cytotoxicity by permeabilizing mammalian and *Culicoides* insect cells (Hassan *et al*., 2001). Cytotoxicity has been localized to two amphipathic helices at the amino terminal. Another feature of VP5 of BTV that is consistent with membrane penetration activity is the formation of trimers in solution (Hassan *et al*., 2001). It seems that the function of VP5 lies in the mediation of virus-cell penetration.

Synthetic BTV VLPs consisting of VP2 and VP5 of different serotypes have been assembled using recombinant baculoviruses. These VLPs were used to produce serum with high neutralizing antibody titres against live BTV suggesting a potential use as an anti-BTV vaccine (Loudon *et al*., 1991). The same proteins expressed in *Saccharomyces cerevisiae* failed to produce a protective response to a BTV challenge as a result of conformational differences to the native proteins (Martyn *et al*., 1994). VP2 expressed in a baculovirus system has been demonstrated as a potential vaccine for AHS in horses (Roy *et al.*, 1996). However, du Plessis *et al.* (1998) have demonstrated that baculovirus expressed VP2 of AHSV is present in the form of insoluble aggregates with only 10% soluble protein. Their results further established that only soluble VP2 was capable of inducing neutralizing antibodies. The insoluble nature of VP2 is a limitation to its usefulness as a vaccine. Co-expression of VP2, VP5 and VP7 of AHSV serotype 4 has been shown to induce low levels of neutralizing antibodies and give protection against virulent virus challenge in horses (Martínez-Torrecuadrada *et al*., 1996).

Grimes *et al*. (1998) report that core particles of BTV from which the outer capsid has been removed retain infectivity for insect vectors even though their infectivity for mammalian cells has been greatly reduced.

1.9.2 THE CORE

BTV enters the cell by removal of the outer capsid and release of the core into the cytoplasm of the cell. The genome remains within the protein core which functions as a molecular engine (Grimes *et al*., 1998).

There is a large amount of conservation within the capsid genes of BTV, AHSV and EHDV (Williams *et al*., 1998). The most conserved capsid protein VP3 shows 57 – 58% amino acid identity between serogroups. The main function of VP3 appears to be structural. It provides protection for the dsRNA genome (Hwang *et al.*, 1994). Although VP7 is also very conserved it is not as highly conserved as VP3 (44 – 46% identity) (Iwata *et al*., 1992; Williams *et al*., 1998).

AHSV VP7 expressed in Sf cells self assembles to form flat, mostly hexagonal crystals such as are seen in BHK cells (Oellermann *et al.*, 1970; Chuma *et al*., 1992; Burroughs *et al*., 1994). VP7 is exceptionally hydrophobic (Roy, 1992). X-ray crystallography studies have shown that the core capsid has icosahedral symmetry. Two hundred and sixty VP7 trimers form a lattice on top of a thin shell of 120 VP3 monomers (Grimes *et al*., 1997). Thirteen icosahedrally independent copies of VP7 form the outer layer of the core capsid. They are arranged so that four trimers are in general positions and one has its molecular three-fold axis aligned with the icosahedral three-fold axis. The core has small pores at the icosahedral three-fold axes and larger pores at the five-fold axes (Grimes *et al*., 1998).

VP3 subunits display icosahedral symmetry within the subcore layer. There are two sets of VP3 subunits with a triangular wedge type appearance. The amino acids of the protein are arranged in a plate-like three-domain structure. The three domains are: the apical domain (near the 5-fold axes of the icosahedron); the carapace (forms a rigid plate) and the dimerization domain (forms a quasi-two-fold interaction). Interaction between the VP7 and VP3 layers of the core capsid occurs at flattish, mostly hydrophobic surfaces. Grimes *et al*. (1998), propose that the core may be thought of as being constructed by the crystallization of VP7 trimers onto the subcore of VP3 subunits. The driving force of this reaction may be the interaction of each trimer with the underlying VP3 layer.

Six epitopes have been mapped to the VP7 molecule of BTV. One such epitope located at the N-terminus is accessible on the surface of both intact virions and core particles (Wang *et al*., 1996a). Some evidence has been presented for the use of VP7 as a subunit vaccine for BTV (Wade-Evans *et al*., 1996) and AHSV (Wade-Evans *et al*., 1997). A series of substitutions based on the X-ray structure of VP7 were designed for BTV. These site-directed mutations targeted amino acids believed to be involved in interaction within the VP7 subunit and between subunits. The purpose being to determine the effect of certain substitutions on VP7 trimer formation and core assembly (Limn *et al*., 2000). Such information has important biological implications especially for

development of VP7 as a epitope delivery system. Similar research has been conducted on AHSV VP7 in our laboratory.

Co-expression of VP3 and VP7 of BTV and AHSV in insect cells results in the spontaneous formation of core-like particles (Hewat *et. al*., 1992; Maree *et al.*, 1998). Recombinant baculoviruses expressing VP2, VP5, VP7 and VP3 of BTV allow synthesis of non-infectious, double-shelled, virus-like particles when co-infected in insect cells. These VLPs are the same size and have the same appearance as authentic BTV virions (French *et al*., 1990).

VP7 appears to be involved in the binding of BTV vector insect membranes. Xu *et al*. (1997) suggest that VP7 may initiate specific binding to membrane receptors and facilitate VP2 and VP5 in virus cell entry. This evidence suggests a role in mediation of cell attachment and penetration for the core proteins (Grimes *et al*., 1998). Enzymatic functions of the virus are performed within the inner core.

1.9.3 THE INNER CORE

The inner core contains the minor structural proteins VP1, VP4 and VP6 and the dsRNA genome. The core particle functions as a “molecular engine” by producing full-length capped mRNAs from each of the genome segments. These mRNAs function as templates for viral protein synthesis as well as for negative strand synthesis for the production of progeny dsRNA genome segments (Grimes *et al*., 1998).

For efficient transcription, the dsRNA should be laid down within a pre-existing protein shell. Grimes *et al*. (1998) propose that the minor inner core proteins may attach to VP3 decamers with the result that the complete subcore (VP3, dsRNA genome and minor core proteins) has a transcription complex at each of the 12 five-fold axes. Gouet *et al.* (1999), propose a model for the packing of dsRNA within the inner core of the BTV particle. In their model they suggest that the dsRNA coils around the transcription complex at the five-fold axis (figure 1.1). The transcription complexes are made up of the three minor structural proteins namely VP1 (polymerase), VP4 (capping enzyme) and VP6 (the helicase) (Diprose *et al*., 2001, Mertens and Diprose, 2004).

During transcription, the polymerase utilizes the negative RNA strand as template for the synthesis of a positive daughter strand. The reaction requires NTP and the by-product is pyrophosphate (PPi). The VP4 protein synthesizes a 5’ cap on the daughter strand (Diprose *et al*., 2001). All of the 10 dsRNA segments are transcribed separately and repeatedly (Gouet *et al*., 1999). Each dsRNA segment is thought to be attached to a transcription complex. The dsRNA segments have to move through the polymerase active site during transcription. VP6 (helicase) may be involved in both the unwinding of the parental strands parental-daughter strand complex (fully conservative replication) and export of daughter strands from the core. VP6 uses ATP hydrolysis as energy source producing ADP and phosphate (Pi) as by-products.

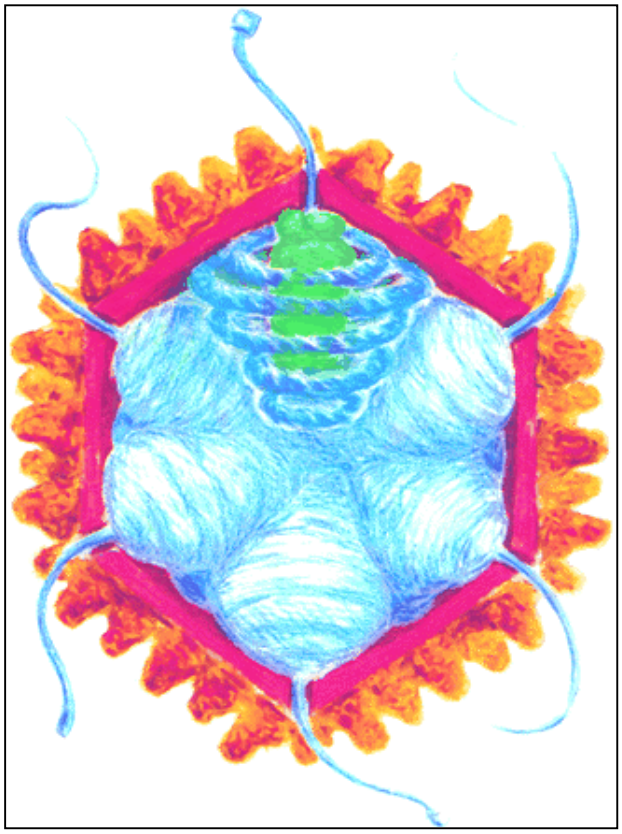


Figure 1.1 Model for the packaging of dsRNA in the BTV core particle.

The RNA is depicted as a blue coil around the transcription complex (in green) showing the position of the RNA with regards to the five-fold vertices (Gouet *et al*., 1999).

mRNAs produced by transcription must be exported from the core for translation in the host cells cytoplasm. Therefore, there must be entry points in the core for NTP and other substrates. Exit points for reaction by-products and mRNA molecules into infected cells are also required. DsRNA must remain inside the core to prevent the host immune response on contact of dsRNA with the cytoplasm of host cell. Diprose *et al*. (2001), have used X-ray crystallography to investigate how the core manages these activities. They have proposed that mRNA is extruded from the core at the pores at the icosahedral five-fold axes. These twelve pores are the largest openings through the VP3 layer. There are positively charged amino acid groups in the five-fold related pores that may guide the RNA by means of electrostatic charge (Grimes *et al*., 1998). The pore between the A and B molecules of VP3 (site N) is used for entry of substrates and release of by-products. This pore carries little overall charge. Further binding sites in the VP7 layer occur between monomers of VP7 between adjacent trimers. Large numbers of nucleotides on the particle surface may act as a substrate sink. This may serve to

increase the local concentration and supply of nucleotides around the core particle thus facilitating polymerase activity. Ion binding sites are located on the core. Two cation sites are found at the base of the VP7 trimers and are occupied by Mg^{2+} ions. A general anionic binding site has been observed between the VP3 and VP7 layers and is occupied by phosphate or sulphate ions. The entry and exit ports are summarized in figure 1.2.

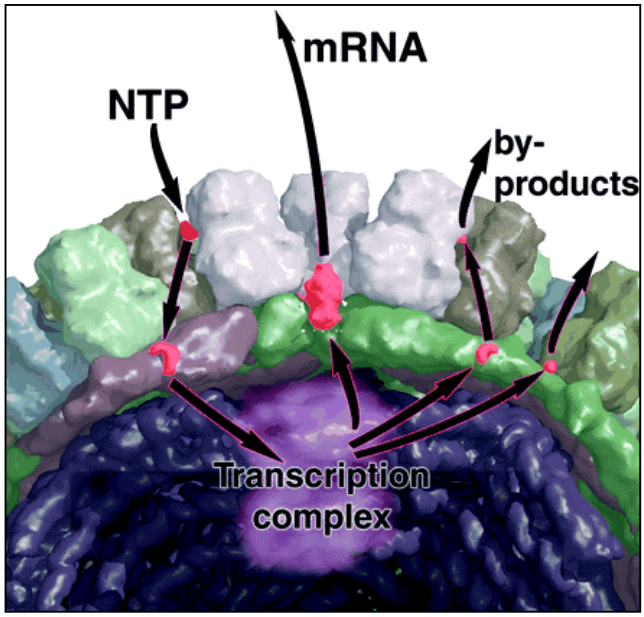


Figure 1.2 Cartoon illustrating the movement of the NTP raw material into the core particle and the site of exit of the mRNA transcript and by-products.

NTPs are sequestered at the inter-trimer site and NTPs are thought to cross the protein barrier between molecules A and B of VP3. The mRNA product leaves via the pore at the five-fold axis of the particle. By-products leave via pores in the VP7 layer at the icosahedral five-fold axes and between molecules A and B of VP3 (Diprose *et al*., 2001).

Most functions in the cell are thought to be carried out by macromolecular complexes or "molecular machines" as opposed to single protein enzymes (Nogales and Grigorieff, 2001). This would seem to be the case in BTV based on the X-ray crystallography derived structural modules. The function of each of the components of the transcription complex of BTV namely, VP1, VP4 and VP6 have been assayed individually.

The putative RNA polymerase VP1 is encoded by segment 1. There is strong amino acid sequence similarity between various RNA and DNA polymerases (Roy, 1992; Huang *et al*., 1995).

VP1 of AHSV has been proposed to include a motif characteristic of RNA-dependent RNA polymerases of positive-strand RNA viruses and some dsRNA viruses (Vreede

and Huismans, 1998). Urakawa *et al*. (1989), demonstrated RNA polymerase activity by means of a poly (U) template -oligo (A) primer polymerase assay for VP1 of BTV. It has further been demonstrated that BTV VP1 has replicase activity, producing dsRNA from plus strand viral template RNA in what appears to be a template independent manner (Boyce *et al*., 2004).

The capping and methylation of the virus mRNA is another enzymatic function performed by VP4 (Roy, 1992). Le Blois *et al*. (1992) demonstrated covalent binding of GTP by VP4 of BTV and proposed VP4 as the candidate guanylyl transferase of the virus. Sequence analysis revealed a potential leucine zipper motif near the C-terminus of VP4 which is conserved among five US BTV viruses (Huang *et al*., 1993). Mutation analysis has revealed that this leucine zipper motif is essential for dimerization of the VP4 molecule which in its native form exists as a dimer in solution. It also appears that dimerization is necessary for assembly of VP4 into BTV cores (Ramadevi *et al*., 1998b). VP4 of BTV has two methyltransferase activities: it caps existing RNA molecules by transferring a methyl group to the guanosine capping residue at the 5' terminal and catalyzes methylation of the ribose of the nucleotide preceding the 5' terminal nucleotide. It is also able to condense GTP to form a capped dinucleotide which confirms its guanylyl transferase activity (Ramadevi *et al*., 1998a). VP4 of BTV has nucleoside triphosphatase (NTPase) activity with a preference for GTP as substrate (Ramadevi and Roy, 1998).

The second smallest RNA segment codes for VP6 which together with VP1 and VP4 make up the mRNA polymerase complex. BTV VP6 is a hydrophilic protein, which is very basic, containing a large number of charged amino acids (Fukusho *et al*., 1989; Roy, 1992). Wade-Evans *et al*. (1992), observed a doublet on expression of VP6 of BTV-1. This doublet has been attributed to two in-phase start codons which supposedly give rise to these two forms of VP6 as opposed to post-translational modification.

VP6 of BTV is a nucleic acid binding protein which binds dsRNA, ssRNA and dsDNA in a mechanism independent of its tertiary structure. In accordance with its nucleic acid binding ability, VP6 is found within the VIBs where virus morphogenesis is believed to occur (Roy *et al*., 1990). Hayama and Li (1994), further characterized BTV VP6 nucleic acid binding activity. By deletion mutation analysis, they localized binding to two domains which correspond to three of six identified antigenic epitopes.

VP6 has long been proposed as the putative viral helicase (Roy, 1992). In BTV, this activity has been demonstrated. It has also been demonstrated to have ATP binding activity and RNA-dependent ATPase activity. Previously unidentified in viral helicases, RNA duplexes with 3' and 5' overhangs as well as blunt-ended dsRNA were unwound by VP6 (Stäuber *et al*., 1997).

Due to its amino acid composition, deletion and substitution mutants of BTV VP6 encompassing the amino and carboxy terminals were generated for investigation of intracellular location and structure-function relationships. This investigation has shown

that residues in the amino terminal are necessary for retention of VP6 in the cell cytoplasm, while residues in the carboxy terminal may allow for nuclear localization. However, the intact protein is located in the cytoplasm (Yi *et al*., 1996).

In a recent study, Kar and Roy (2003), targeted putative helicase domains of BTV VP6. These included sites thought to be involved in ATP binding activity, helicase activity and RNA binding. By site directed mutagenesis, they identified motifs important for ATP binding and hydrolysis as well as RNA binding. According to their results, Kar and Roy (2003) propose that purified BTV VP6 forms multimers which can be stabilized into ring-like particulate structures (hexamers) in the presence of RNA.

1.9.4 THE NONSTRUCTURAL PROTEINS: NS1, NS2, NS3 AND NS3A

There are four virus specific nonstructural proteins found in cells infected with orbiviruses. A characteristic feature of *Orbivirus* infected cells is the polymeric morphological structures assembled from NS1 which is expressed abundantly in infected cells (Huismans and Els, 1979; Roy, 1996). The BTV dsRNA segment 5 codes for the NS1 protein which has a predicted M_r of 64kDa (Lee and Roy, 1987). In BTV, NS1 tubules are found throughout the infection cycle and can be demonstrated at 2 – 4 hr p.i. (Huismans and Els, 1979). These tubules are attached to the intermediate filaments of the cell cytoskeleton (Eaton and Hyatt, 1989) and are associated with virus inclusion bodies (VIB) and virus particles (Eaton *et al.,* 1987). They are thought to be involved in virus transportation in infected cells (Roy, 1996). There is high sequence identity between NS1 within the BTV serogroup (Hwang *et al*., 1993), between serogroups the amino acid identity ranges between 20 and 50% (Moss and Nuttall, 1995). These authors found that five cysteine residues are conserved in AHSV, BRDV, BTV and EHDV. In the case of BTV-10, substitution of two of the conserved cysteines prevents tubule formation (Monastyrskaya *et al.,* 1994). Using a panel of monoclonal antibodies, Monastyrskaya *et al*. (1995) established an NS1 antigenic site near the carboxy terminus which appears to be exposed on the surface of the tubules. In addition, they found that adding 16 amino acids at the carboxy terminus of BTV-10 NS1 does not affect tubule formation. Mikhailov *et al.* (1996), produced various chimeric BTV-10 NS1 genes by inserting foreign sequences at the C-terminus. All of the constructs formed tubules expressing the foreign antigenic sequences on the surface of the tubules and were highly immunogenic, recognizing the native foreign proteins in ELISA. Due to high expression levels and easy purification, NS1 tubules may be a useful system for the presentation of foreign epitopes (Mikhailov *et al.,* 1996). NS1 tubules of AHSV have been characterized (Maree and Huismans, 1997) and are currently under investigation as particulate immunogen delivery systems.

Other virus-specific structures found in cells infected with orbiviruses such as BTV are the virus inclusion bodies (VIBs) (Lecatsas, 1968). A major component of VIBs is nonstructural protein NS2 which is expressed in large amounts in infected cells (Roy, 1996). The assembly of virus particles appears to occur in VIBs. These VIBs contain

structural and nonstructural proteins as well as virus particles at various stages of morphogenesis (Brookes *et al*., 1993). Thomas *et al*. (1990), has reported that baculovirus expressed NS2 of BTV-10 forms inclusion bodies similar to viral inclusion bodies found in BTV infected cells. A comparison of the nucleotide sequences of segment 8 of AHSV, EHDV and BTV showed a similarity of 53 – 60% (van Staden *et al*., 1991). BTV NS2 has been demonstrated to bind ssRNA by means of a nitrocellulose-binding assay where NS2 bound both cellular ssRNA and BTV mRNA (Huismans *et al*., 1987a). Uitenweerde *et al*. (1995) demonstrated that NS2 is a multimer of six or more NS2 molecules. NS2 is a phosphorylated protein which is expressed throughout the virus replication cycle (Huismans *et al*., 1987a). It has been found that NS2 is the only BTV protein that is phosphorylated which is different from equivalent proteins with an affinity for ssRNA in other *Reoviridae* viruses such as reovirus (Huismans and Joklik, 1976) and rotavirus (Boyle and Holmes, 1986). Phosphorylation was found to reduce the efficiency of ssRNA binding of NS2 (Theron *et al*., 1994). Van Staden *et al.* (1991), identified a motif in NS2. This motif plays an important role in the conformation of the protein necessary for ssRNA binding and the formation of VIBs in infected cells (Theron *et al.,* 1996a). It has been proposed that NS2 may have a role in selection and condensation of the virus mRNA (Huismans *et al*., 1987a).

In a study using UV cross-linking assays, NS2 of BTV was found to have the highest affinity for the 3’ regions on viral mRNA. Mutant transcripts lacking 5’ and 3’ highly conserved terminal hexanucleotides bound as effectively as wild-type transcripts. Therefore, these sequences are not necessary for binding (Theron *et al*., 1996b; Theron and Nel, 1997). Three regions with non-specific ssRNA binding activity *in vitro* have been identified in BTV NS2 protein (Fillmore *et al*., 2002). Lymperopoulos *et al*. (2003) have found that BTV NS2 recognizes and binds to a sequence in the coding region of BTV-10 segment 10. They suggest that the secondary structure (hairpin-loop structure with two minor stem-loops in the main-loop area) is important for binding activity. The main stem and part of the main loop are proposed to bind NS2. Horscroft and Roy (2000), have reported NTP binding and phosphohydrolase activity in NS2 of BTV. They have suggested that the NTPase activity may be important in providing energy for selection, transport or packaging of ssRNA.

NS3 and NS3A are the gene products of segment 10 (Mertens *et al.,* 1984; van Dijk and Huismans, 1988; French *et al*., 1989). Peptide mapping of AHSV, Palyam and BTV indicated that they arise from translation of two in-phase start codons (van Staden and Huismans, 1991). The hydrophobic domains, putative transmembrane sequences, possible glycosylation sites and a cluster of proline residues are conserved between BTV and EHDV S10 gene products (Jensen *et al*. 1994). NS3 and NS3A are associated with the plasma membrane of infected cells. They are localized to regions of plasma membrane disruption (Hyatt *et al*., 1991). It was further demonstrated that NS3 / NS3A mediate budding and release of virus-like particles from infected cells (Hyatt *et al*.,

1993). AHSV NS3 has been demonstrated to have a cytotoxic effect on infected cells. It was found to disrupt the cell membrane and immunofluorescence suggested possible membrane association. This membrane association may lead to increased permeability of the cell membrane resulting in the loss of osmotic regulation and subsequent cell death (Van Staden *et al*., 1995).

De Sá *et al*. (1994), investigated the phylogenetic relationships between segment 10 of AHSV serotypes 1, 3, 4, 8 and 9. They found a close relationship between serotypes 4 and 9; 1 and 8 and serotype 3 was found to be closer to the 4 and 9 group. Higher variation across and within serotypes of NS3 of AHSV was found than in EEV and BTV NS3. These findings indicate that NS3 is the second most variable AHSV protein after VP2 (Van Niekerk *et al*., 2001b; Van Niekerk *et al*., 2003). Van Staden and Huismans (1991) reported two highly conserved regions in the amino acid sequence of NS3 of serotypes 3 and 9. These were found to be conserved in all five serotypes (De Sá *et al*., 1994). Comparisons of the hydrophobic profiles of AHSV and BTV NS3 revealed two conserved hydrophobic domains (De Sá *et al*., 1994). Van Staden *et al*. (1995) identified these two hydrophobic domains as regions capable of forming transmembraneous helices in BTV, AHSV, Palyam, BRD and EHDV. Using this data, they proposed a model for the membrane associated topology of NS3 with the transmembrane regions spanning the membrane and the N and C termini on the cytoplasmic side of the membrane. Some support for this model was provided by an investigation of membrane organization of bluetongue virus NS3 (Bansal *et al*., 1998). Their data indicated that both the hydrophobic domains span the cell membrane with the likelihood that the N and C termini are cytoplasmic.

Mutations in either of the hydrophobic domains destroys membrane anchoring of NS3 of AHSV (van Niekerk *et al*., 2001a). As a result of this, NS3 is no longer localized to the cell surface and the cytotoxic effect is eliminated. Van Niekerk *et al*. (2001a) have demonstrated that cytotoxicity of NS3 is dependent on membrane association determined by both the hydrophobic domains.

BTV NS3 / NS3A exist as N-linked glycoproteins (Wu *et al.,* 1992). Bansal *et al*. (1998) have demonstrated that only one of the two glycosylation sites in bluetongue virus NS3 is utilized. Using mutants in which the active glycosylation site is deleted, no NS3 could be detected in association with the cell membrane. They suggest that this may be as a result of failure of transport due to improper folding or proteolysis resulting from the lack of protection by the carbohydrate.

In a study using an avirulent and a virulent strain of African horse sickness virus, three phenotypes were observed in a mouse model namely, fully virulent, fully avirulent and a new intermediate virulence phenotype. The intermediate phenotype arose from reassortment of segment 10 between the avirulent and the virulent strains (O'Hara *et al.*, 1998).

1.10 VIRAL ENZYMATIC FUNCTIONS

Virus replication is performed using the host cell machinery. As a result of this, viruses contain a few enzymes at most. The genomes of double-stranded nucleic acid viruses function as templates for replication and transcription (Fields, 1998).

The synthesis of mRNA in BTV is performed by the RNA polymerase associated with the core particle (Huismans and van Dijk, 1990). Replication and transcription are initiated after unwinding of the double-stranded nucleic acid occurs. After virus entry into permissive cells, the outer capsid layer is removed and transcriptase activity occurs inside the viral core (Huismans and van Dijk, 1990; Stäuber *et al*., 1997). Capped and methylated mRNA extrudes from the viral core to be translated into viral proteins in the cytoplasm of infected cells (Huismans and van Dijk, 1990; Patton *et al*., 1990; Roy, 1992; Stäuber *et al*., 1997; Diprose *et al*., 2001).

1.10.1 VIRAL HELICASES

DNA helicases, first reported in 1976 (Abdel-Monem *et al*., 1976) and RNA helicases, first reported by Lain *et al.* in 1991, are found in both prokaryotes and eukaryotes. Helicases enzymatically unwind duplex nucleic acid structures of cellular and viral genomes by translocation along one strand (Geider and Hoffmann-Berling, 1981; Gibson and Thompson, 1994). They generate single-stranded intermediates which are required for replication, recombination and repair (Matson and Kaiser-Rogers, 1990; Runyon and Lohman, 1993; Lohman and Bjornson, 1996). Helicase activity is linked to NTP hydrolysis which presumably provides energy for the reaction. The reaction also requires the presence of a divalent cation, usually Mg^{2+} (Gomez de Cedron *et al*., 1999; Bideshi and Federici, 2000). All helicases described to date have NTPase activity (Matson and Kaiser-Rogers, 1990; Lohman and Bjornson, 1996; Kadaré and Haenni, 1997).

1.10.1.1 Structural features of helicases

Mechanisms for helicase action propose two features: ATP-driven conformational rearrangement that generates energy and a minimum of two nucleic acid binding sites between which energy can be employed to execute unwinding (Gibson and Thompson, 1994; Kadaré and Haenni, 1997). The two sites would allow intermediates to bind either two single-stranded nucleic acid regions or one single-stranded and one duplex region at an unwinding junction. Evidence for such intermediate complexes exists for the Rep protein of *E. coli* (Lohman, 1992; Lohman, 1993).

The most direct mechanism for a helicase to obtain multiple nucleic acid binding sites is oligomerization (Lohman and Bjornson, 1996). Enzymes that are monomeric in solution may assume active oligomeric forms upon binding of a nucleic acid or a nucleotide cofactor. An example is the Rep helicase of *E. coli* which dimerizes only on DNA binding (Lohman, 1992; Lohman, 1993; Kadaré and Haenni, 1997).

DNA helicases can be divided roughly into two groups namely: hexameric helicases and non-hexameric helicases. Recent structural and biochemical data has shown that non-hexameric helicase bacteriophage T4 helicase Dda is active in a monomeric form (Marians, 2000).

Comparisons of amino acid sequences of proteins involved in nucleic acid metabolism have yielded several different motifs. These motifs include ATPase activity, helicase activity, NTP and nucleic acid binding sites (Matson and Kaiser-Rogers, 1990; Kadaré and Haenni, 1997).

1.10.1.2 Helicase families

Based on sequence comparisons, helicase proteins may be divided into five major groups. These superfamilies (SF) include cellular and viral DNA and RNA helicases (Lohman and Bjornson, 1996; Kadaré and Haenni, 1997, Soultanas and Wigley, 2001).

The superfamilies are classified as follows: in SF1 are the alphavirus-like nsP2 proteins, this includes nsP2 of Semliki forest virus (Gomez de Cedron *et al*., 1999). The SF1 proteins are associated with a zinc finger structure (Seybert *et al*., 2000; van Dinten *et al*., 2000). The helicase of vaccinia virus and proteins encoded by the potyvirus-flavivirus-pestivirus NS3-like proteins such as the flavivirus, West Nile virus NS3 (Borowski *et al*., 2001) are found in SF2. SF2 includes the DexH/DEAD protein family (Koonin, 1991, Gibson and Thompson, 1994). SF1 and SF2 are characterized by seven conserved motifs (Caruthers and McKay, 2002). SF3 includes picornavirus-like 2C-like proteins. It also includes the parvovirus NS1. SF3 proteins have three conserved motifs including the “A” and “B” sites and a third site designated “C” (Kadaré and Haenni, 1997). Helicases that are related to the DnaB protein of *E. coli* fall into the fourth group. These helicases usually form hexameric ring structures and have five conserved motifs. The fifth family was recognized as having sequence similarity to the β subunit of proton-translocating ATPases, represented by the transcription termination factor Rho (Caruthers and McKay, 2002, Singleton and Wigley, 2002). Proteins are usually classified into one of the above families on the basis of characteristic helicase motifs. Some proteins have been biochemically investigated. It is becoming evident that only a small percentage of these proteins actually possess helicase activity. This may be because the helicases are only functional when part of a protein complex. Another explanation is that these helicase motifs are characteristic of NTP dependent nucleic acid translocases (Soultanas and Wigley, 2001; Singleton an Wigley, 2002). The translocase molecular motor may not be able to unwind duplex nucleic acid. Helicase activity may be conferred by other extra protein domains, excluding the helicase motifs (Caruthers and McKay, 2002).

Certain helicases, for example RuvB, belong to the AAA^+ family rather than the previously described helicase super families (Singleton and Wigley, 2002). The AAA^+ family is thought to represent a broad class of mechanoenzymes which in many cases form hexameric rings that change their conformation during the ATPase cycle. Amongst

a large array of functions, some AAA^+ proteins are molecular motors with ATP binding sites (Vale, 2000).

In 1990, Dorer and others reported the discovery of a new member of the RNA helicase family. This helicase of *Drosophila* differs from the SF helicases in that it has glycine repeats. Glycine rich proteins (GRPs) have been described in numerous organisms including RNA binding proteins (Sachetto-Martins *et al*., 2000). Most helicases that have glycine repeats or are glycine rich have other characteristic motifs such as DEAD/DexH, RNP-1 and RNP-2 or dsRBDs, (Roussell and Bennet, 1993; Sato, 1994; Zhang and Grosse, 1997). Glycine rich sequences have been reported to play a role as anchors for ATP (Bossemeyer, 1994).

1.10.1.3 Function of the conserved motifs

The conserved motifs in helicases may divided according to primary function. Some are primarily involved in binding MgATP/MgADP, others bind oligonucleotides and there are those that play a role in coupling the ATPase cycle to the intramolecular conformational changes involved in the unwinding and strand displacement activities (Caruthers and McKay, 2002).

The most obvious common factor of the superfamilies is the NTP binding motif. This motif consists of the "A" and "B" sites of the "Walker box". "A" may be [$GxxxxGK^S/_T$, where x is any amino acid] or the abridged [$GxGK^S/_T$] (Gorbalenya *et al*., 1989). It forms the P-loop within the NTP binding site. "B" consists of an aspartate (D) which interacts with the NTP by means of Mg^{2+} (Lohman and Bjornson, 1996; Kadaré and Haenni, 1997). The NTP binding motif alone is not sufficient to propose a helicase signature as there are a variety of proteins that utilize NTP but are not necessarily helicases. The energy generated by NTP hydrolysis may be used for processes such as RNA encapsidation, intercellular movement of the virus or vesicular trafficking of replication complexes (Kadaré and Haenni, 1997). Another important aspect of helicase activity is the ability to bind nucleic acids.

Numerous nucleic acid binding motifs have been identified for both DNA and RNA binding proteins. Focusing on RNA binding motifs, possibly the most common and best characterized RNA binding motif is the RNP (or otherwise known as the RNA recognition motif - RRM) which consists of the RNP-1 ((K/R)G(F/Y)(G/A)FVX(F/Y)) and RNP-2 ((L/I)(F/Y)(V/I)(G/K)(G/N)L)) sites (Burd and Dreyfuss, 1994; Sachetto-Martins *et al*., 2000). Another common motif is the arginine-rich motif (ARM). It is characterized by 10 – 20 amino acids, rich in arginine (Mattaj, 1993). The RGG box has 20 – 25 amino acids with RGG repeats which are closely spaced with other (often aromatic) amino acids. Usually the RGG box is found in combination with other types of RNA binding domains. Binding is generally sequence non-specific (Burd and Dreyfuss, 1994). The KH motif is a complex motif identified in the hnRNP K protein. The KH motif proteins of known function are associated with RNA and bind RNA in vitro (Mattaj, 1993; Burd and Dreyfuss, 1994). The double-stranded RNA binding motif (dsRBM) consists of a 70

amino acid region that binds double-stranded RNA characterized by an α-β-β-β-α fold. There are nine families of proteins with the dsRBM. The proteins are classified according to their number and distribution of dsRBMs. Other additional domains characterize these protein families (Fierro-Monti and Mathews, 2000). These proteins are involved in diverse cellular functions. The Zinc finger-knuckle ($CX_{2-5}CX_{4-12}C/HX_{2-4}C/H$) is found in a small number of RNA-binding proteins which include RNA polymerases and retroviral nucleocapsid proteins (Burd and Dreyfuss, 1994). The cold shock domain (CSD) has been found in RNA binding proteins such as CspA and CspB. The RNA binding domain has conserved aromatic and basic residues and binds dsDNA, ssDNA and RNA (Graumann and Marahiel, 1996). Helicase motif VI (RXGRXXR), characteristic of SF2 helicases, has an important role in RNA binding (Kadaré and Haenni, 1997; Kim *et al*., 1997; Kar and Roy, 2003). This motif is characterised by a high number of positively charged amino acid residues. Finally, GRPs (Glycine-rich proteins) are a family of proteins found in many different organisms, they are proposed to be involved in protein-protein interactions. The glycine-rich regions are thought to result in a highly flexible secondary structure. They are often found in association with RNP-1 and RNP-2. The consensus motif may be GGGX; GGXXXGG or GXGX (Sachetto-Martins et al., 2000). Bossemeyer (1994) reported two consensus sequences found in glycine rich nucleotide binding proteins. $GxGxGL^{S}/_{T}$ is usually found in proteins with a preference for binding mononucleotides. GxGxxG on the other hand may be found in proteins that bind dinucleotides preferentially.

1.10.1.4 Helicase activity

NTPase activity of helicases is generally stimulated by the presence of single-stranded nucleic acids, preferentially by the type of nucleic acid to which the helicase is functionally related (i.e. DNA or RNA) (Lohman, 1993; Kadaré and Haenni, 1997). Characterized RNA helicases hydrolyze ATP in the presence of different RNAs including synthetic homopolymers and also single-stranded DNA. They all bind single-stranded RNA and DNA. It was thought that most RNA helicases cannot unwind DNA, however, notable exceptions include vaccinia virus NPHII (Bayliss and Smith, 1996) and SV40 T antigen (Stahl *et al*., 1986; Scheffner *et al*., 1989) both of which may act as DNA or RNA helicases depending on the context, eIF-4A (Matson and Kaiser-Rogers, 1990; Kadaré and Haenni, 1997) Hepatitis C NS3 (Gwack *et al*., 1997) and λ1 protein of Reovirus (Bisaillon *et al*., 1997). Some DNA helicases such as the Rho protein of *E. coli* are capable of unwinding RNA/DNA hybrids in an ATP dependent manner. This type of reaction could be significant in pre-mRNA splicing, translation and other cellular processes (Matson and Kaiser-Rogers, 1990). There is evidence indicating that some helicases may be sensitive to the structure of the duplex substrate. For example, in unwinding studies using hepatitis C virus (HCV) NS3 it was found that heteroduplexes of peptide nucleic acid (PNA) and DNA or RNA adopt a structure that is unfavourable for unwinding by NS3 of HCV. While morpholino-DNA and phosphorothioate-DNA and DNA-DNA substrates were utilized efficiently (Tackett *et al*., 2001).

During helicase activity, the helicase enzyme remains bound to one strand using the energy generated by ATP hydrolysis to disrupt hydrogen bonding of a duplex structure. It is thought that a protein conformation leading to optimized binding and hydrolysis of the NTP may be induced by nucleic acid binding. High salt concentrations do not stimulate NTPase activity suggesting that in this condition the enzyme may not remain bound to the nucleic acid (Kadaré and Haenni, 1997) or, as in the case of the T-antigen, helicase of SV40, the enzyme assumes a conformation not suitable for helicase activity (Scheffner *et al.*, 1989; Kadaré and Haenni, 1997). Most DNA helicases function either 5' to 3' or 3' to 5' (Matson and Kaiser-Rogers, 1990). The polarity of an unwinding reaction is defined by convention with respect to the strand of nucleic acid to which the helicase is bound (Matson, 1991). *E. coli* RecBCD which is a recombination helicase, unwinds blunt ended DNA duplexes and is inhibited by flanking ssDNA regions that exceed 25 nucleotides. The RecB subunit interacts with the 3' end of the ssDNA strand and the RecC subunit interacts with the 5' end of the ssDNA strand. This provides direct evidence for helicase interaction with both strands. There is no direct evidence supporting strict unidirectional translocation of a helicase even though it may preferentially unwind DNA with a 3' or 5' ssDNA flanking region. A functional initiation complex may only be formed when an appropriate single-stranded flanking region is available (Lohman, 1993). RNA helicases described up to 1997 such as the helicase of tamarillo mosaic potyvirus and vaccinia virus helicase unwind unidirectionally 3' to 5' (Shuman, 1992; Eagles *et al*., 1994; Kadaré and Haenni, 1997). However, as mentioned above, RNA helicases such as VP6 of BTV possess bi-directional helicase activity as well as the ability to unwind blunt ended substrates (Stäuber *et al*., 1997). Binding of a helicase to a region of unpaired nucleic acid was thought to be a prerequisite for the initiation of the unwinding activity (Geider and Hoffman-Berling, 1981), this, however, is clearly not the case with BTV VP6 (Stäuber *et al*., 1997).

1.10.1.5 Models of helicase activity

Several mechanisms for DNA unwinding and translocation have been proposed. These may be active or passive depending on whether the helicase plays a role in the unwinding or whether it functions only to stabilize the ssDNA. According to the passive mechanism, the helicase sequesters single-stranded nucleic acid that arises as a result of thermal fluctuations at single-stranded/double-stranded nucleic acid junctions. Based on kinetic measurements of duplex DNA unwinding, NS3 of hepatitis C virus (HCV) has been proposed to function using a passive mechanism (Porter, 1998). However, other researchers have concluded that NS3 of HCV interacts weakly with the displaced strand. Therefore, conclusive evidence for a passive mechanism is not available (Paolini *et al*., 2000; Tackett, *et al*., 2001).

Two types of active helicase mechanism have been proposed. The first type is the "torsinol" model. In this model, the helicase binds to both ssDNA strands at the ss/dsDNA junction with no interaction with the duplex. Unwinding occurs as a result of the distortion of the adjacent duplex region through NTP induced conformational

changes. The other type of active unwinding mechanism includes both the “inchworm” and “rolling” models. In the inchworm model, there are two different binding sites: one site binds with polarity to ssDNA while the other site (the leading site) binds both single-stranded and dsDNA. Translocation occurs by means of conformational changes coupled to the binding and hydrolysis of one or more NTP molecules. The enzyme travels along the dsDNA and separates the duplex in the fashion of a snowplough. This model has been proposed for PcrA, a monomeric helicase, based on crystal structures of PcrA bound to a DNA substrate (Bird *et al*., 1998; Velankar *et al*., 1999). The “inchworm” model is illustrated in figure 1.3.

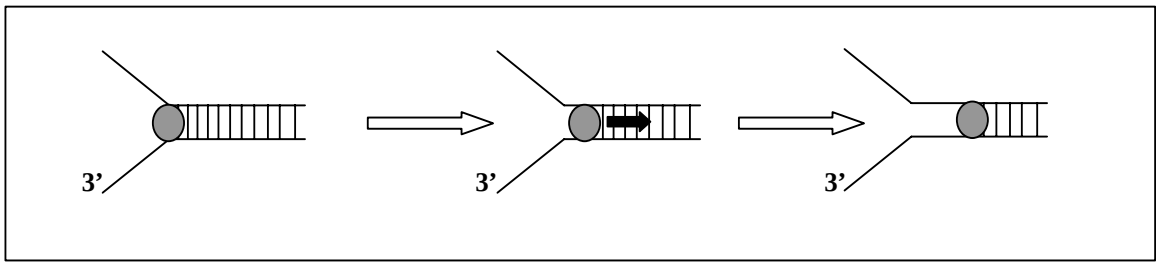


Figure 1.3 The inchworm model of helicase activity.

The helicase molecule (represented by the grey circle) moves along the DNA in discrete steps “melting” the DNA strands as it goes. The helicase in this model does not have to be oligomeric (redrawn from Bird et al., 1998).

The “rolling” model is based on the Rep molecule of *E. coli*. This model is based on observed allosteric effects of nucleotide cofactors on the single-stranded and dsDNA binding properties of the Rep dimer. It could be applied to any homo-oligomeric helicase. One subunit of the Rep dimer is always bound to 3’ ssDNA at the fork while the other may either bind to the same single strand or to the adjacent duplex region ahead of the fork. Each subunit alternates between binding of single-stranded and dsDNA controlled by the allosteric effects of ATP and ADP binding. In this model the Rep dimer translocates by rolling along the DNA with translocation coupled to ATP binding and DNA unwinding coupled to ATP hydrolysis (Lohman, 1993; Lohman and Bjornson, 1996). The “rolling” model is illustrated in figure 1.4.

Of the various models proposed, work on the SF1 and SF2 enzymes (usually monomeric) have suggested that the inchworm model may be the most likely helicase mechanism (Soultanas and Wigley, 2000, Singleton and Wigley, 2002). There is some debate on whether the mechanism is active or passive with regards to ATP hydrolysis. Is the energy derived from hydrolysis utilized for translocation and duplex separation or translocation alone? There is considerable thermal fraying of duplex DNA at room temperature. However, the RecBCD helicase translocates at a rate excessive of 1kb/s which is thought to be faster than passive duplex fraying (Singleton and Wigley, 2002).

Recently, equilibrium binding studies have demonstrated that the HCV helicase single-stranded binding activity is increased by the presence of ADP and the absence of ATP (Levin *et al*., 2003). When the HCV helicase bound ATP, a hundred fold reduction in its affinity for ssDNA was observed. This has not been observed for duplex DNA. These results would appear to suggest that the nucleic acid binding site of the HCV helicase is allosterically modified by the ATPase reaction. The binding energy that is used to change the conformation (brings the HCV helicase out of a tightly bound state) may be used to facilitate translocation. The ATP hydrolysis reaction and product release are thought to promote tight rebinding of the nucleic acid and the helicase. Levin *et al.* (2003), propose a Brownian motor model for the unidirectional translocation of HCV helicase along the length of the nucleic acid strand.

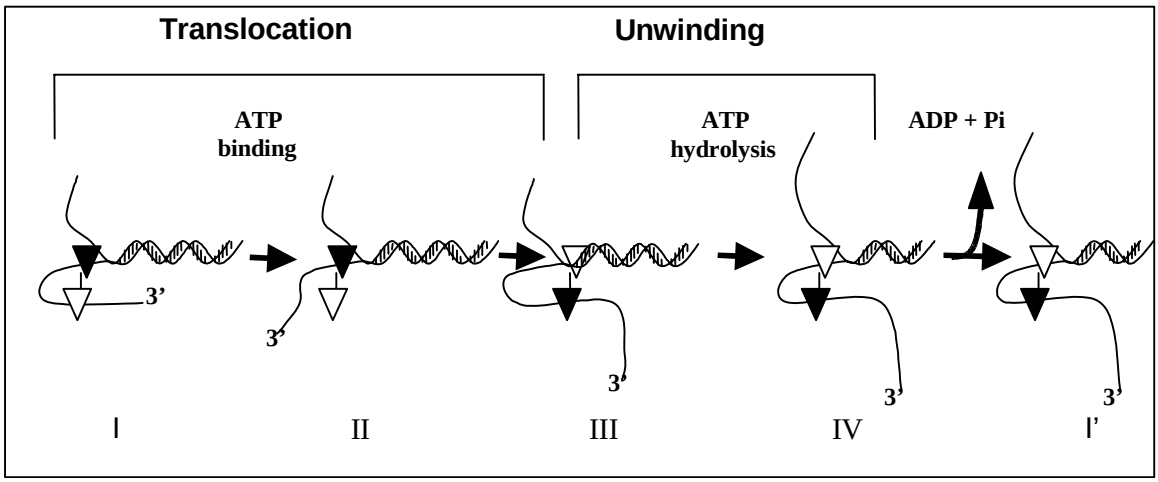


Figure 1.4 "Rolling" mechanism for Rep-catalyzed DNA unwinding.

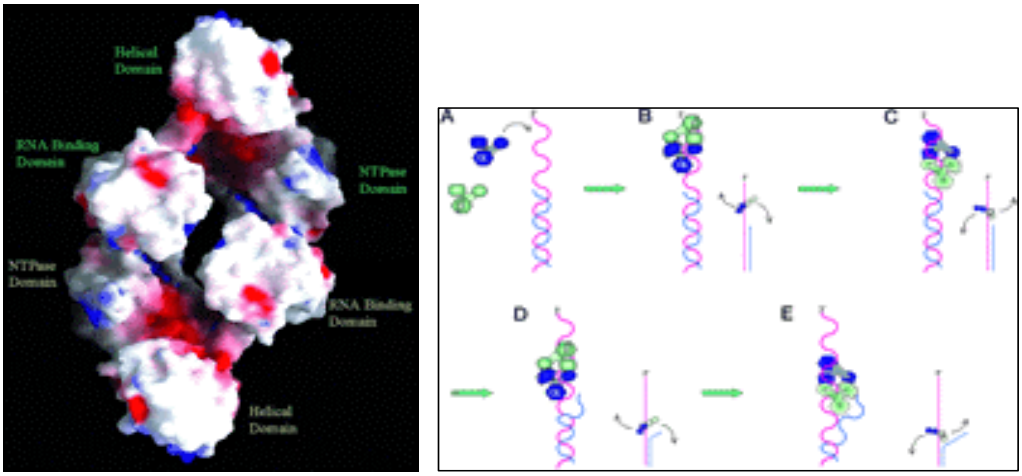
A dimeric helicase in which the subunits possess C2 symmetry, unwinds by direct interaction with both ssDNA and dsDNA. I is an intermediate where both subunits are bound to the 3' single strand. When ATP is bound, the affinity of one subunit for ssDNA decreases and the intermediate II is formed. As its affinity for ssDNA decreases, its affinity for dsDNA increases resulting in the formation of intermediate III. In steps III – IV, ATP hydrolysis brings about a protein conformation change that results in destabilization of basepairing within a region of duplex DNA bound to the Rep protomer. As a result, the 5' single strand is displaced while the Rep protomer remains bound to the 3' strand. The cycle is completed with the release of ADP and inorganic phosphate by the formation of the I' complex. This is functionally equivalent to intermediate I because Rep is a homodimer (Lohman, 1992; Lohman, 1993; Bird *et al*., 1998).

Progress towards a more complete picture of the structure and mechanism of helicases has recently come from high resolution crystallography. The crystal structures for a number of helicases have been determined. The structural similarities between

hexameric and non-hexameric helicases suggest the possibility of a single mechanism (Bird *et al.*, 1998; Soultanas and Wigley, 2000).

Cho *et al.* (1998), determined the crystal structure of the RNA helicase domain of HCV type 1b at 2.3 Å resolution. The oligomeric state of a functional HCV helicase is unknown. These authors investigated the packing interfaces and propose a “symmetry-related monomer-monomer” interaction possibly reflecting a functional form of the RNA helicase. Interfaces arise from contacts between the NTPase and RNA binding domains of the two molecules. The RNA binding motif of each molecule forms a helical shaped channel in the middle of the dimer (figure 1.5a).

Cho *et al*. (1998), have proposed a ‘descending molecular see-saw’ model for HCV NS3. This model is consistent with properties previously observed for the protein. According to their model, duplex RNA is unwound coupled with NTP hydrolysis. As ATP hydrolysis results in the dissociation of RNA from the enzyme, it is hypothesized that NTP hydrolysis causes a hinge bending motion. This transforms the enzyme from its activated conformation to its resting conformation coupled with the release of the bound ssRNA from the interdomain cleft of the resting molecule of the proposed dimer. The dimer thus rotates relative to the bound ssRNA. The rotation axis must be towards the 5’ end of the bound RNA in order for the dimer to be able to translocate on the ssRNA. In this model, the RNA binding motifs are thought to play a very important role in the helicase mechanism by serving as the pivot region for the rotation. The dimer rotates so that the resting molecule goes into the activated conformation. Repeated cycles of rotation and translocation are described as a “descending molecular see-saw” (figure 1.5b).



(a) (b)

Figure 1.5 Crystal packing interaction of the crystals of HCV RNA helicase

a) ssRNA is thought to pass through the channel created by the two molecules. There is a 2-fold crystallographic symmetry axis approximately perpendicular to the figure. The interdomain cleft shown in red is thought to be mostly negative in charge. The dimer shown above is proposed to be similar in conformation to the putative active dimer. (b) Schematic diagram of the 'descending molecular see-saw' model for the translocation of HCV RNA helicase along the ssRNA. In (A) the α helicase molecule binds ssRNA. In (B), another helicase molecule, β, binds and forms a functional dimer. A small conformational change (induced by the ssRNA bound to the interdomain cleft of α) is thought to increase the NTP hydrolysis activity. In (C), as a result of NTP hydrolysis the ssRNA detaches and the dimer rotates at the RNA binding motif of α. Thus the dimer translocates in a 5' direction on the ssRNA. β binds the ssRNA at its interdomain cleft. In (D), the point where the ssRNA becomes dsRNA is reached by repeated cycles of translocation. In (E), interactions between the interdomain cleft and the ssRNA provide energy for the disruption of basepairs. The strand moving through the channel, generated by the dimer interface, is separated from the strand that hangs out of the dimer (Cho *et al*., 1998).

The first high resolution X-ray data on a hexameric helicase (2.4 Å resolution) was provided for RepA of plasmid RSF1010 (Niedenzu *et al*., 2001). The RepA helicase is a DNA helicase that forms a hexamer in its functional form. It has an annular shape with a central hole and 6-fold rotational symmetry (C6). The hexamer has a positive surface charge with negatively charged areas restricted to the bottom side of the central hole. The size of the central hole only allows ssDNA to thread through the hexamer (figure 1.6).

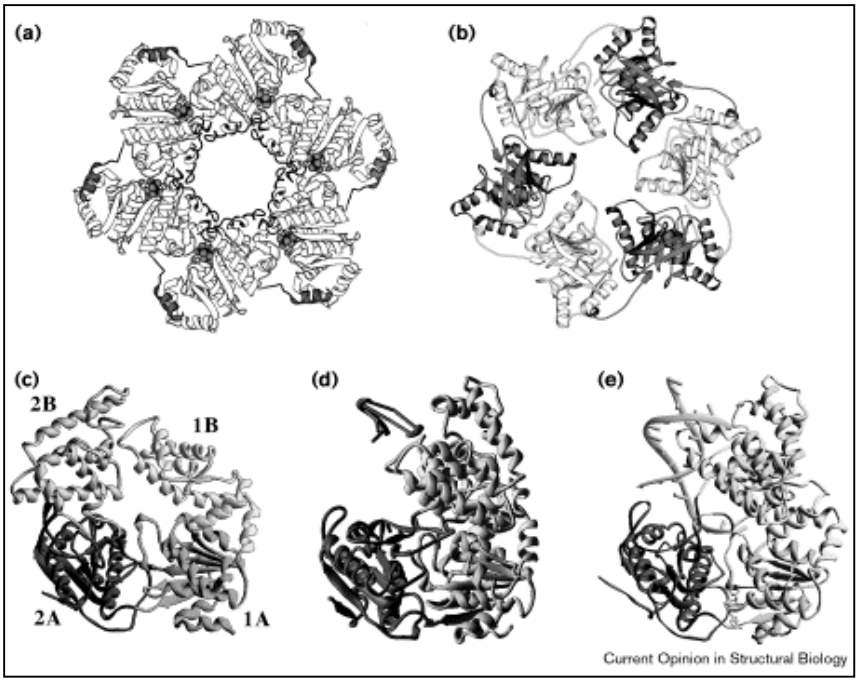


Figure 1.6 Crystal structures of two hexameric helicases

a) T7 helicase domain from the helicase-primase protein and b) the RepA hexameric ring. Both are viewed parallel to the 61 crystallographic axis clearly illustrating the six molecules forming the crystallographic unit. c) – e) Represent the three forms of the monomeric helicase PcrA. The four different subdomains identified by Velankar et al., (1999) are indicated on the apo form (c). d) Product form and e) is the substrate complex. (Soultanas and Wigley, 2000).

Two different structures of PcrA DNA helicase complexed with a DNA duplex with a ssDNA tail have been used to propose a model for a general helicase mechanism (figure 1.7) (Velankar *et al*., 1999). This model is loosely based on the 'inchworm' model and may be divided into two processes namely, translocation and duplex destabilization. These two processes together are responsible for helicase activity. When the helicase binds to ssDNA, translocation begins. Domain swivelling is initiated which allows binding of a DNA duplex. At this stage, ssDNA is bound to domains 1A and 2A (figure 1.6c). The cleft between the two domains closes when ATP is bound. For this to occur, one of the domains (2A) releases its hold on the ssDNA. The bound ATP

is hydrolized and the cleft opens. The helicase moves along the DNA across domain 2A while domain 1A maintains its hold on the DNA. This continues until the helicase reaches a duplex. The protein has a greater affinity for dsDNA and accordingly binds to the duplex region next to the fork. The duplex DNA is dragged onto the surface (domains 1B and 2B figure 1.6c) and there is pressure on the basepairing at the fork. The structure of the dsDNA becomes distorted and several basepairs near the fork are destabilized. Therefore, binding energy generates more ssDNA along which the helicase can translocate. This model (otherwize referred to as the “Mexican wave model” figure 1.7b) has several implications for the understanding of helicase activity. According to this model, helicase activity is active rather than passive. Components required for DNA translocation are found in domains 1A and 2A which are widely conserved in RNA and DNA helicases. Helix destabilization and/or substrate specificity is controlled by domains 1B and 2B which are considerably more variable. In this model, domain swivelling creates a site for binding dsDNA once the protein is bound to ssDNA. Thus the helicase cannot interact with dsDNA facilitating strand separation within a sealed duplex because it must be activated by the presence of ssDNA. This model may be a general model for helicase activity (Velankar *et al*., 1999; Soultanas and Wigley, 2000).

1.10.1.6 Prevalence and role of helicases in viruses

There appears to be a correlation between the genome size of a virus and the presence of a putative helicase. For example, birnaviruses which have small genomes have no known helicase motifs (Koonin, 1991). All dsDNA viruses, for which complete sequence data is available, contain helicase motifs. RNA viruses with large genomes such as reoviruses (Harrison *et al*., 1999) have helicase motifs. Genome sequences of single-stranded viruses have yielded putative helicases. The genomes of single-stranded viruses may be partially duplex in form. During replication of positive stranded RNA viruses, templates must be available for the next round of replication. It has, therefore, been proposed that single-stranded viruses may require helicases to disrupt intramolecular base pairing in template RNA and to prevent extensive base pairing between template and complementary strands (Kadaré and Haenni, 1997).

Three general hypothesis regarding the functions of helicase enzymes in viral life cycles have been proposed. Firstly, helicases may have a proof reading energy dependent mechanism responsible for the fidelity of replication. This may explain why viruses with large genomes (larger than 5kb) encode putative helicases, as they are needed for efficient replication (Koonin, 1991).

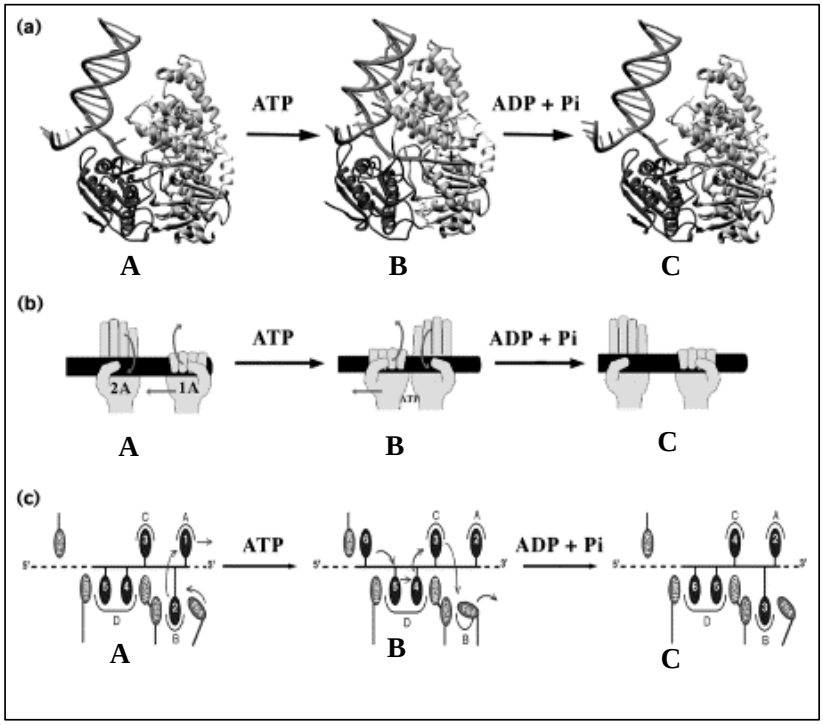


Figure 1.7 Model for helicase activity based on changes in the conformation of both the protein and nucleic acid substrate

(a) The intermediates are based on the three forms of PcrA (Velankar et al., 1999)(figure 1.6 c – e). In step A, the helicase binds to the ssDNA. In step B, following ATP binding, a conformational change in the protein allows it to bind dsDNA at domains 1B and 2B. This is coupled with the disruption of several basepairs at the fork. In step C, after ATP hydrolysis, the protein conformation reverts to that in (A) and the helicase moves along the ssDNA by one base and lets go of the dsDNA. (b) Cartoon showing the alternating ssDNA affinity of domains 1A and 2A. A loose grasp of the ssDNA is illustrated by an open hand while a closed hand represents a tight grip. (A – C reflect the same events in (a); (b) and (c)). (c) Cartoon illustrating the ssDNA binding region at each stage showing the conformational changes as the bases move between binding pockets when the helicase translocates along the ssDNA (Velankar *et al*., 1999; Soultanas and Wigley, 2000).

According to the second hypothesis, helicase activity may be required for strand separation of double-stranded structures during transcription and prevention of R-loop formation behind the elongating RNA polymerase. This activity may be necessary for the initiation of (early) transcription and has been proposed for vaccinia virus helicase (Gross and Shuman, 1996). The third hypothetical role for virus encoded helicases is during initiation of translation. This hypothesis is based on eIF-4A. Together with eIF-4B, eIF-4A is thought to disrupt secondary mRNA structures upstream of the initiation codon thus facilitating 40S ribosome attachment (Matson and Kaiser-Rogers, 1990; Rozen *et al*., 1990). It is sometimes difficult to determine unambiguously which process

a specific helicase is involved in as many organisms encode multiple helicases such as *E. coli* which has at least 10 helicases (Matson, 1991; Lohman, 1993).

1.10.1.7 Past and future aspects of viral helicases

The signatures of the helicase super families are found in both prokaryotic and eukaryotic helicases. This suggests that their common ancestor existed very early in the evolution process and the main branching point may possibly have occurred prior to the eubacterial/eukaryotic divergence. For this reason, it has been proposed that the helicases of the SFs form a fairly compact domain within the large number of proteins with NTP-binding motifs (Kadaré and Haenni, 1997; Anantharaman *et al*., 2002).

Comparison of the three dimensional structures of ABC (ATP-binding cassette) transporters and helicases have shown that there are similarities in the structure of their nucleotide binding sites, their ATP binding domains and the positioning of important residues. It is possible that these two protein families share a common mechanism for harnessing the energy produced by ATP hydrolysis. This may indicate a common ancestor and the two families may be considered to be closely related (Geourjon *et al*., 2001).

Regions of DNA involved in helicase binding can theoretically be identified by Dnase I and chemical protection methods (i.e. footprinting). However, unless there is high specificity for helicase binding, it may be difficult to distinguish between productive and non-productive complexes. As a result, determining functional contacts is difficult. A case in point is SV40 large T antigen. Two different groups applied DNA footprinting techniques and obtained different results (Lohman and Bjornson, 1996). In the case of viral RNA helicases, no specific RNA sequence requirements have been reported. However, it is presumed that such specificity does exist *in vivo* and may arise from a combination of sequence, secondary and tertiary structure of the RNA substrate (Mattaj, 1993; Kadaré and Haenni, 1997). Helicases and their substrates may offer targets for drug development specifically with a view to interfering with viral infection cycles as these proteins appear to fulfil important roles for the virus (Kadaré and Haenni, 1997). In particular, NS3 of hepatitis C virus which has both protease and helicase activities is considered an important target for anti-HCV drug therapy (Tackett *et al*., 2001). Borowski *et al*. (2002), have suggested possible mechanisms for using the NS3 of the hepatitis C virus for antiviral therapy. NTPase activity may be inhibited by the introduction of a competitive NTPase inhibitor, for example one which blocks the ATP binding site and inhibits the ATPase activity of the enzyme. Alternatively, allosteric interference may be used to inhibit unwinding by sterically blocking translocation. NTPase and helicase activities may be uncoupled by disrupting the energy transfer within the helicase. Competitive inhibition of RNA binding, for example, RNA aptamers with many secondary structures interact with the HCV helicase NS3 protein and stop helicase activity. Intercalating agents stabilize both DNA and RNA and, therefore, the energy required for unwinding is higher.

Helicase proteins differ in size and may have additional N- and C-terminal sequences that determine the specific action of each protein. These sequences may include membrane targeting signals, nuclear localization signals or regions for interaction with modulating accessory proteins (Kadaré and Haenni, 1997). This may explain the diversity of functions that are achieved by variations on the common structural theme. In some viruses, different enzymatic functions are assigned to different proteins as is the case with BTV, AHSV and other orbiviruses. In other viruses, helicases may function in complexes with other proteins (Lohman, 1993). For example, in brome mosaic virus (Ahola *et al*., 2000), one protein is responsible for RNA capping and helicase activity, while a different protein is related to polymerases. A similar situation exists for λ1 protein of Reovirus (Bisaillon and Lemay, 1997). NS3 of hepatitis C virus has both protease and helicase functions and is thought to be involved in viral replication (Jin and Peterson, 1995; Banerjee and Dasgupta, 2001). The function of viral enzymes is to ensure that viral proteins are expressed and new virions are produced. This involves the processes of transcription and replication. Together with helicases, important viral enzymes include polymerases and mRNA capping enzymes.

1.10.2 VIRAL TRANSCRIPTASE ACTIVITIES

In the case of Reovirus, the λ1 protein has NTPase and helicase activity as well as 5' RNA triphosphatase activity. Besides its role in transcription of the viral genome, λ1 protein is thought to participate in the capping of the 5' ends of new viral mRNAs. Capping involves the removal of the 5' end phosphate by means of the polynucleotide phosphohydrolase function of the RNA triphosphatase. GMP derived from GTP is donated by guanylyltransferase. The result is a 5'-5'-triphosphate linkage which is characteristic of the mRNA cap structure (Bisaillon and Lemay, 1997). The hepatitis E virus capping enzyme P110, shows similarities to the methyltransferase and guanylyltransferase of alphavirus nsP1, tobacco mosaic virus P126, brome mosaic virus replicase protein 1a and bamboo mosaic virus nonstructural protein. These similarities in properties suggest a common evolutionary origin for the above plant and animal virus families (Magden *et al*., 2001). These authors further propose methyltransferase and guanylyltransferase reactions as potential antiviral drug development targets as both reactions are strictly virus specific.

RNA viruses are replicated in infected cells by viral RNA-dependent RNA polymerases, referred to as replicases (Ishihama and Barbier, 1994; Kanjanahaluethai and Baker, 2000). As well as polymerase activity, the replicase enzyme of influenza virus has a proofreading function (Ishihama *et al*., 1986).

A number of viruses have either a multifunctional enzyme (NS3 of HCV) or several proteins co-localize at the site of viral replication (mouse hepatitis virus - MHV). NS3 of hepatitis C virus (HCV), is a multifunctional enzyme with three known catalytic activities separated into two independent domains. Serine protease activity is located at the N-terminal end of the protein while NTPase and helicase activities are found in the C-

terminal region. It has been further suggested that a protein expressed immediately downstream of NS3 (NS4A) in the viral polyprotein, may play an integral structural role in the NS3 protease domain. There is some evidence suggesting that the domains of NS3 may influence the activity of one another and that interactions between the genomic elements of HCV may regulate the enzymatic activities of this HCV replicase component (Morgenstern *et al.*, 1997).

Proteolytic processing of the replicase gene product of MHV has been shown to be necessary for viral replication (Teng *et al*., 1999). The replicase gene (gene 1) is translated into two polyproteins that are processed to generate in excess of 15 mature proteins. Several of these proteins have been demonstrated to localize at the site of viral RNA synthesis in tightly associated membrane populations (Sims *et al*., 2000). It has further been demonstrated that complexes which are known to be involved in RNA synthesis contain several gene 1 products closely associated with structural proteins located at sites of virion assembly (Bost *et al*., 2000).

In the case of Dengue virus, a mosquito-borne flavivirus, replication occurs at the membrane associated replicase complex (RC). The RC complex includes a number of proteins which are derived from the viral polyprotein. Amongst these are NS3, protease/ NTPase / helicase and NS5 the methyltransferase / replicase. It would seem that the RC may be viewed as a “protein machine” with important protein-protein and RNA-protein interactions determining its replication function. Amongst these interactions, it has recently been demonstrated that the C-terminal region of NS3 interacts with the N-terminal region of NS5 (Johansson *et al*., 2001).

1.10.3 REPLICATION AND TRANSCRIPTION IN BTV

In the case of BTV, virus replication and transcription occur in the virus inclusion bodies (VIBs). The proteins responsible for replication and transcription are the three minor core proteins. VP1 has been identified as the replicase (Urakawa *et al*., 1989; Boyce *et al*., 2004). The role of capping of the mRNA has been ascribed to VP4 (Ramadevi *et al*., 1998a). VP6 has been identified as the helicase with associated NTPase activity (Stäuber *et al*., 1997). An important characteristic relating to helicase function is the ability to bind nucleic acids. BTV VP6 has been identified as a nucleic acid binding protein (Roy *et al*., 1990; Hayama and Li, 1994). As yet, the activity of each of these proteins has been examined independently and their association has not been demonstrated. Very little is known about the corresponding proteins in AHSV. Sequence analysis has suggested similarities in function.

1.11 AIMS

When this study was initiated, very little information was available with regards to the VP6 gene of AHSV or its product. This study focused on the characterization of the structure and function of AHSV VP6. These long term, broad-based objectives were met by addressing the following short term aims:

a) To determine the primary structure of the genome segment encoding VP6 and the VP6 protein in order to identify structural elements or sequence homologies that will assist us in identifying the function of VP6.

b) To characterize the VP6 protein of AHSV with a particular focus on optimizing the expression of the protein and investigating protein features such as solubility, post-translational modification and in particular nucleic acid binding activity.

Both the above aims will enable us to predict a possible helicase function for VP6 of AHSV.