CHAPTER 1

LITERATURE SURVEY

The study of viruses and their interactions with host cells has benefited greatly from the ability to engineer specific mutations into viral genomes, a technique known as reverse genetics. This ability to generate infectious viruses from cloned sequences has contributed greatly to our biological understanding of pathogens and their replication and hence to disease control. It also enables the exploitation of the viral replication machinery for the expression of heterologous proteins.

1. REVERSE GENETICS

1.1 DNA viruses

DNA-containing viruses were the first to become amenable to reverse genetic techniques. This breakthrough was achieved when DNA of SV40 (~5000 bp in length) was found to be infectious, giving rise to new viral particles when a cloned copy was transfected into cells. This allowed the first rescue of defined viral mutants from mutated DNA molecules (Goff & Berg 1976).

Subsequently, the molecular engineering of large DNA-containing viruses such as herpes and pox viruses was enabled by methodology that involved homologous recombination of the viral genome with plasmids bearing foreign sequences flanked by viral sequences, under appropriate selection conditions (Post & Roizman 1981; Mackett *et al.* 1982; Panicali & Paoletti 1982). Similarly, techniques have been developed to specifically alter the genomes of adenoviruses and many other DNA viruses (Jones & Schenk 1978; Samulski *et al.* 1989). Reverse genetic strategies to recover infectious and mutant viruses by recombination between cotransfected cosmids containing overlapping portions of

large viral genomes have also been developed and successfully used (Cunningham & Davison 1993; Kemble *et al.* 1996; Cohen *et al.* 1989b). Thus, extraordinary progress has been made in harnessing the genomes of DNA viruses to facilitate study and understanding of structure-function relationships of the viral components and as viral vectors expressing foreign proteins.

1.2 RNA viruses

1.2.1 Retroviruses

Unique among the RNA viruses are the retroviruses whose replication involves a dsDNA phase, making these viruses an easy target for genetic manipulation. Transfection of full-length cDNA molecules results in the formation of replicating virus particles and integration of the viral genetic information into the host genome, as first demonstrated by Wei *et al.* (1981). Engineering of retroviruses has been widely applied in the study of viral gene expression and of protein-structure analysis and has enabled them to be used as vectors for gene transfer and gene therapy (Mulligan 1993).

1.2.2 Single-strand RNA viruses

The study of the molecular biology of non-retroviral RNA viruses has long been hampered by the fact that these viruses do not encompass a DNA intermediate step in their replication cycle.

One of the original distinctions between positive- and negative-strand RNA viruses was based on the ability or not of their purified RNAs to initiate an infectious cycle after transfection of appropriate host cells (Baltimore *et al.* 1970).

Positive-sense ssRNA viruses

In the case of positive-strand RNA viruses, the full-length genomic RNA functions as mRNA, directing the production of some or all viral proteins necessary for the initiation of virus propagation, and as template for viral RNA replication, making these viruses highly amenable to genetic engineering. Advances in molecular techniques have enabled direct genetic manipulation of positive-strand RNA viruses, through the use of cDNA intermediates to produce biologically active RNA molecules. Thus, infectious positive-strand RNA viruses

can be generated from cloned cDNAs, by transfecting plasmids (or RNA transcribed from plasmids) containing the viral genome directly into cells, as was first demonstrated with *Poliovirus* (PV; Racaniello & Baltimore 1981).

Due to their generally smaller genome sizes compared to DNA viruses, whole RNA virus genomes can be cloned as cDNA and manipulated at will. This approach has been successfully achieved for multiple small and medium sized positive-strand RNA viruses (see Boyer & Haenni 1994), greatly enhancing the potential of investigations. Indeed, they can facilitate studies of viruses that are present only in low titres in infected cells or whose isolation is problematic. For instance, the development of a reverse genetics system for caliciviruses (Sosnovtsev & Green 1995) was anticipated to assist in the identification of the molecular basis for the strong host- and tissue-specific restrictions of many members of the *Caliciviridae* and to lead to the development of recombinant DNA-based systems for the non-cultivatable caliciviruses.

Clearly, the synthesis and cloning of full-length cDNA of larger positive ssRNA viral genomes, with correct termini, and the instability of these clones in bacteria, can be troublesome (reviewed in Boyer & Haenni 1994). However, remarkable success has been achieved in studies using alphaviruses such as *Sindbis virus* (Rice *et al.* 1987) and *Semliki forest virus* (Liljestrom *et al.* 1991). cDNA-derived RNAs of these positive-strand RNA viruses were used to efficiently rescue infectious viruses, thus allowing extensive analyses of promoter elements of the viral RNAs as well as structure-function studies of the viral proteins. Furthermore, these viruses have shown excellent potential for expressing large quantities of heterologous proteins via recombinant constructs.

Obvious advantages to the use of positive-strand RNA viruses as vectors for the expression of heterologous sequences include the easy and rapid engineering of the DNA constructs and, in contrast to DNA viruses like *Vaccinia virus* (VACV), the possibility of avoiding any wild type virus background by *de novo* generation of viruses entirely from cloned sequences. In addition, RNA virus vectors have the advantage as vaccine vectors in that they don't appear to (down) modulate the immune system, as do many large DNA viruses, including pox and herpes viruses (Ploegh 1998). Moreover, because these RNA viruses lack a DNA phase, there is no concern about unwanted integration of foreign sequences into chromosomal DNA.

An alternative reverse genetic approach, particularly with larger (30 kb) RNA genomes such as that possessed by the coronaviruses, utilises defective interfering particles (DIs). It is based on the generation of DIs by transcription of smaller RNAs that contain deletions and are not able to replicate autonomously. DIs are able to replicate with the help of normal viruses that provide the functions not encoded by the DI itself. Since DI genomes have to contain certain *cis*-acting sequences to be replicated and packaged into infectious particles, construction of synthetic DI genomes can be used for the identification of replication and packaging signals (Makino *et al.* 1990; Van der Most *et al.* 1991). In addition, recombinant DI-RNAs can serve as replicons designed for the expression of foreign genes (Finke & Conzelmann 1999).

Recent advances

In order to achieve successful infection, viral transcripts must interact with virus-encoded proteins, most particularly with the viral replicase, and with host cell components such as the translation machinery; therefore the structure of viral transcripts has to mimic that of the virion RNA as closely as possible. Investigations using the reverse genetics strategies described above were markedly promoted by *in vivo* expression of infectious viral RNAs. This was initially achieved by T7 promoter-driven transcription from transfected cDNA-containing vectors in cells infected a recombinant VACV (vTF7-3) encoding T7 RNA polymerase (Fuerst *et al.* 1986). Alternatively, a host range-restricted VACV recombinant (MVA-T7) that expresses T7 polymerase, but does not replicate, in many mammalian cells (Schneider *et al.* 1997) and a recombinant baculovirus expressing T7 RNA polymerase (Yap *et al.* 1997) have also been developed. Cell lines constitutively expressing T7 polymerase alone or together with helper proteins have also been used (Radecke *et al.* 1995). Recently, cellular RNA polymerase II promoters have also been used for intracellular synthesis of foreign transcripts (Johnson & Ball 1997).

In order to retain intact 3' terminal sequences for functional genome transcripts, the development of plasmid vectors that allowed intracellular generation of RNAs with discrete termini was a further major technical breakthrough in optimising the system. This was achieved by the discovery (Cech 1986) and subsequent exploitation of the autolytic activity of ribozyme sequences, as first successfully used for the intracellular generation of functional

nodavirus RNA (Ball 1992; Ball 1994) and *Vesicular stomatitis virus* (VSV) RNAs (Pattnaik *et al.* 1992). This system allowed much more efficient production of appropriate RNA inside a cell, as compared to transfection with *in vitro* transcribed RNA or with linearised DNA constructs to obtain intracellular runoff transcripts. The latter in particular is not very effective in the presence of VACV, most likely owing to ligation and modification of DNA by VACV enzymes. The ribozyme sequence of *Hepatitis delta virus* (HDV) has generally been used, which has the advantage of requiring only sequences downstream of the cleavage site for autolytic activity and is indiscriminate with regard to sequences 5′ of its cleavage site. Thus RNAs ending with a specific 3′ terminal nucleotide can be generated by autolytic cleavage of primary transcripts containing the HDV ribozyme sequence immediately downstream of the viral sequences. According to the ribozyme cleavage mechanism, the 3′ terminal ribose of the (upstream) genome analogue possesses a cyclic 2′-3′ phosphate instead of a hydroxyl group (Long & Uhlenbeck 1993). This modification might be speculated to contribute to the success of the approach, in preventing polyadenylation of the RNA by VACV enzymes, for example, or in delaying degradation of the RNA 3′ terminus.

Negative-sense ssRNA viruses

The genomes of negative-strand RNA viruses have been less amenable to artificial manipulation as neither naked genomic viral RNA nor anti-genomic complementary RNA can serve as a direct template for protein synthesis and are therefore not infectious. Both genomic and positive-sense anti-genomic RNAs exist as viral ribonucleoprotein (RNP) complexes, and the viral RNA polymerase is essential for transcribing both mRNA and anti-genome template RNA from these RNP complexes. Thus the minimal biologically active replication unit is formed by the genomic or anti-genomic RNA complexed with nucleoprotein and the viral RNA polymerase (Emerson & Wagner 1972). In addition, precise 5′ and 3′ ends are required for replication and packaging of the genomic RNA (Zheng *et al.* 1996). The absence or extremely low frequency of homologous recombination during the replication cycle of negative-sense RNA viruses eliminates the possibility of inserting novel genes into the genome of these viruses by targeted recombination with synthetic nucleic acids.

Unfortunately, this lack of systems for genetic manipulation of RNP viruses long limited experimental approaches to studying the genetics and

biology of negative-strand RNA viruses. This also prevented the enormous potential of RNP viruses from being exploited as tools for basic and applied biomedical research. This potential derives from the high integrity of RNP genomes within the cell, the mode of gene expression from simply organised genomes, the cytoplasmic replication cycle of most RNP viruses and also from the relatively simple structure of virions.

The genomes of segmented negative-sense RNA viruses allowed some genetic manipulation through the isolation of reassortant viruses, but manipulation of the complete genome progressed slowly, hampered by the very fact that the genome is segmented, requiring a separate viral RNA for each segment.

Site-specific manipulation of a negative-strand RNA virus was first made possible in 1990 for the segmented *Influenza A virus* (FLUAV; Enami *et al.* 1990). Biologically active RNP complex (comprised of synthetic cDNA-derived RNA of one viral genomic segment complexed with purified nucleoprotein and polymerase proteins) was reconstituted *in vitro* and then transfected into helper virus-infected cells. The helper virus provides the viral proteins required for the rescue and isolation (under selection) of infectious genetically engineered (reassortant) virus. More recently, intracellular reconstitution of RNP complexes by expression of a viral RNA-like transcript from plasmid-based expression vectors containing a truncated human polymerase I promoter and a 3′ ribozyme sequence, demonstrated efficient transcription and replication of a reporter in FLUAV-infected cells (Pleschka *et al.* 1996).

However, it took some time to produce active RNP complexes of non-segmented negative-strand RNA viruses *in vitro*, most likely owing to the tighter RNP structure of full-length genomes of 11 kb or more (Baudin *et al.* 1994). Park *et al.* (1991) demonstrated that a synthetic RNA generated *in vitro* by T7 RNA polymerase from a linearised plasmid to create precise *Sendai virus* (SeV) genome-specific untranslated ends could be amplified and expressed in SeV-infected cells. Similar helper virus-driven rescue of transfected RNAs was also successful in other paramyxovirus systems, such as *Respiratory syncytial virus* (Collins *et al.* 1991) and *Measles virus* (Sidhu *et al.* 1995). The first recovery of infectious non-segmented negative-strand RNA virus from cDNA came when Schnell *et al.* (1994) succeeded in constructing a plasmid that expresses a full-length *Rabies virus* (RABV) RNA transcript (in plus-sense) from the T7 RNA

polymerase promoter. Co-transfection of the plasmid DNA containing this viral insert with plasmids expressing the viral polymerase complex proteins led to the formation of recombinant RABV. This system has been elegantly exploited to study the promoter elements of RABV RNA and to elucidate the interaction of this virus with cells (Mebatsion & Conzelmann 1996; Mebatsion *et al.* 1996), demonstrating that genetic engineering can redirect the host range and cell tropism of rabies viruses.

The key to reproducibly recovering recombinant virus was the use of a plasmid directing transcription of anti-genome (positive-sense) RNA rather than genome RNA. This strategy avoided a potentially deleterious anti-sense problem, in which the vast amounts of positive-sense mRNA transcripts encoding polymerase complex proteins would hybridise to the complementary negative-sense vRNA genome sequences and interfere with critical assembly of the genome into RNP. In starting with anti-genome RNA, only one successful round of replication driven by the plasmid-encoded support proteins is required to yield an infectious genome RNP. The same strategy was also used for the successful recovery and manipulation of paramyxoviruses (Collins *et al.* 1995) and of the prototype rhabdovirus VSV (Lawson *et al.* 1995; Whelan *et al.* 1995). In the latter case, this reverse genetics system was utilised in a fascinating study to systematically alter the phenotype of the virus by manipulation of the genome (Wertz *et al.* 1998; Ball *et al.* 1999). Relative levels of gene expression in VSV, as in other members of the order *Mononegavirales*, is controlled by the highly conserved order of the genes relative to the single transcriptional promoter at the 3′ end of the viral genome through progressive transcriptional attenuation at the intergenic junctions. By rearranging the gene order in an infectious cDNA clone, the authors were able to alter their expression levels and thereby the viral phenotype. Viable viruses were recovered from all ten nucleocapsid (N) gene (Wertz *et al.* 1998) and phosphoprotein (P), matrix protein (M) and glycoprotein (G) gene (Ball *et al.* 1999) rearrangements constructed to date. Levels of gene expression were found to vary in concordance with their distance from the 3′ promoter, yielding rearranged variant viruses with altered replication potential and virulence. Such gene rearrangements yielding stable variant viruses were envisaged to facilitate the study of many aspects of virus biology and the rational development of attenuated live vaccines.

The application of similar reverse genetics strategies for segmented negative-strand RNA viruses posed a formidable challenge, as multiple separate viral RNAs with their RNP complex proteins need to be produced through transfection and co-expression. In one study, Bridgen and Elliott (1996) used reverse genetics to generate *Bunyamwera virus*, a bunyavirus with a tri-segmented negative-sense RNA genome, entirely from cloned cDNAs. This helper-free system involved the co-transfection of plasmids expressing the three anti-genomic viral RNA segments from a T7 promoter, and terminating in the self-cleaving HDV ribozyme sequence at the 3' end, with T7 plasmids expressing the viral mRNAs encoding all the viral proteins, into cells infected with vTF7-3. This eliminated the need for selection to retrieve a small number of transfectants from a vast number of helper viruses. This allows the virus from the initial transfection to be characterised immediately, thus limiting the chance of viruses containing reversions or spontaneous mutations from becoming significant contaminants.

Recently, the generation of influenza A viruses, which possess eight negative-sense RNA genome segments, entirely from cloned cDNAs was reported (Neumann *et al.* 1999). The viral RNP complexes were generated *in vivo* through intracellular plasmid-derived synthesis of the eight vRNAs under control of the cellular RNA polymerase type I promoter and transcription terminator, and simultaneous expression of the viral polymerase complex proteins and nucleoprotein under control of RNA polymerase type II promoters in cotransfected protein expression plasmids. This achievement required minimum co-transfection of 12 different plasmids. However, the authors reported that the addition of plasmids expressing all of the remaining viral structural proteins, requiring co-transfection of up to 17 plasmids, led to a substantial increase in virus recovery. The primary RNA transcripts produced by RNA polymerase I are ribosomal RNAs that possess neither a 5' cap nor a 3' poly(A) tail. The major difference in this system from the reverse genetics of non-segmented negative-sense RNA viruses described above lies in the use of plasmids expressing anti-genome- or genome-sense RNA transcripts. It was speculated that the use of anti-genomic plasmids might further increase the already high efficiency of influenza virus recovery. Shortly following this report, Fodor *et al.* (1999) provided independent evidence for similar plasmid-based rescue of FLUAV, albeit with lower efficiency. The latter also expressed negative-sense genomic viral

RNA, but utilised the HDV ribozyme sequence downstream of the vRNA-coding genes in order to obtain the correct 3' end of vRNA.

This technology thus permits the generation of transfectants with defined mutations in any gene segment, enabling investigators to address issues such as the nature of regulatory sequences in non-translated regions of the viral genome, structure-function relationships of viral proteins, and the molecular basis of host range restriction and viral pathogenicity. Furthermore, this may translate into rationally designed vaccines and the enhanced use of influenza viruses as vaccine vectors and gene delivery vehicles.

1.2.3 **Double-strand RNA viruses**

Double-strand RNA viruses have been found infecting vertebrate and invertebrate hosts, plants, fungi, bacteria and protozoans, and are taxonomically divided into six families (Murphy *et al.* 1995). As the host cells do not possess enzymes for transcribing dsRNA, these viruses share a common requirement to introduce their own virion-associated RNA polymerase into the host cell together with the viral genome. This requirement is reflected in their specialised structures and infection mechanisms.

In theory, infectious dsRNA viral particles should be formed in cells in which a full complement of viral mRNA is introduced. Indeed, Chen *et al.* (1994a) demonstrated that the electroporation of fungal spheroplasts with synthetic plus-sense RNA transcripts corresponding to the non-segmented dsRNA *Hypovirus*, an uncapsidated fungal virus, yields mycelia that contain cytoplasmically replicating dsRNA. Mundt and Vakharia (1996) subsequently described the development of a system for the generation of infectious IBDV (*Infectious bursal disease virus*; *Birnaviridae*) by transfection of synthetic plus-sense transcripts derived from full-length cDNA clones of the entire coding and non-coding regions of the 2 viral genomic segments.

In the case of bacteriophage ϕ6, a *Cystovirus* with a dsRNA genome consisting of three linear segments, acquisition of the plus-strands of the small (S), medium (M) and large (L) genomic segments into procapsids is serially dependent, involving the exposure and concealment of binding sites on the outer surface of the procapsid. This is effected by the amount of RNA causing the empty procapsid to expand. The plus-strand of segment S can be packaged alone, while packaging of the plus-strand of segment M depends on prior

packaging of S. Packaging of M is a prerequisite for the packaging of the plus-strand of L (reviewed in Mindich 1999). Proteins P1, P2, P4 and P7, encoded by the L genomic segment, have been found to assemble into empty procapsids when expressed in *E. coli* (Gottlieb *et al.* 1988). Such cDNA-derived procapsids were found capable of packaging and replicating viral plus-strand RNA to double-stranded genomic segments, as well as producing transcripts using the dsRNA as a template (Gottlieb *et al.* 1990). This approach was recently used by Poranen and Bamford (1999) to demonstrate that the 5' end of the L genome segment in single-stranded form is required to switch from packaging to minus-strand synthesis and the same sequence in double-strand form switches on plus-strand synthesis. By combining previous efforts, Olkkonen *et al.* (1990) demonstrated that such cDNA-derived ϕ6 polymerase complexes that have replicated the viral RNA *in vitro* can be rendered infectious by assembly with *E. coli*-expressed coat protein P8, enabling the generation of recombinant virus.

However, it has generally proved extremely difficult to introduce heterologous genetic information into more complex dsRNA viral genomes, such as those of the family *Reoviridae*, whether by transfection of ssRNA or dsRNA, or by intracellular generation of plus-sense ssRNA (Moody & Joklik 1989). Roner *et al.* (1990) reported the development of a complex system for the recovery of infectious reovirus from *in vitro* synthesised components. Cells were transfected with a combination of ssRNA, dsRNA and *in vitro* translated viral proteins, and complemented with a helper virus of a different serotype. Resulting viruses were distinguished from helper virus by plaque assay. The study determined that dsRNA is 20 times as infectious as ssRNA but that dsRNA and ssRNA together yield 10 times as much infectious virus as dsRNA alone. The addition of *in vitro* translated protein was not found to be absolutely essential, but increased virus yields up to 100 fold, depending on the time for which translation was allowed to proceed. However, destruction of the RNA template following translation abolished the activity, suggesting that the active molecular species was RNA-protein complexes and not protein alone.

2. MOLECULAR BIOLOGY OF THE FAMILY *REOVIRIDAE*

The family *Reoviridae* includes nine genera, infecting a variety of vertebrates, invertebrates and plants. The most highly characterised genera of the *Reoviridae* are *Orbivirus*, *Rotavirus* and *Reovirus*. These viruses share a similar yet unique double layered capsid structure. The viral genome, comprising 10 to 12 dsRNA segments, is encapsidated within an inner core with icosahedral symmetry (Fields 1996). The complete nucleotide sequences of the genome segments of certain species in all three genera are available (Fukusho *et al.* 1989; Wiener & Joklik 1989; Mitchell *et al.* 1990). De-proteinised viral dsRNAs are not infective, reflecting the fact that the virus particles contain their own RNA-dependent-RNA-polymerase for transcription of dsRNA into active mRNAs. The dsRNA segments are base-paired end to end, and the plus-sense strand possesses a 5' cap structure. The 5' and 3' terminal sequences of the genomic segments within a species are highly conserved suggesting that they contain signals important for RNA transcription, replication or assembly during viral morphogenesis. mRNAs of members of the *Reoviridae* are largely mono-cistronic, possessing 5' guanylate residues with a cap structure, but lacking 3' poly(A) tails. The open reading frame is ensconced between non-coding regions of varying lengths, possessing initiation codons in a strong context for initiation (Kozak 1981).

2.1 Infection cycle

The life cycle of the *Reoviridae* is unique and can be summarised as follows. The first step in infection is attachment of the virion or infectious sub-viral particle (ISVP) to various receptor molecules (often unknown) on the cell surface via specific interactions with the viral haemaglutinin (HA) and cell-attachment proteins, in a strain-specific manner. Attached particles are internalised by endocytosis, although alternatives such as phagocytosis or direct penetration of the cell membrane by some viruses and sub-viral particles have also been shown. Following internalisation, viral particles are contained in vacuoles (endosomes or lysosomes) located in the cytoplasm. Within these vacuoles, morphological changes involving uncoating of the viral particles occur, to yield structures very similar to ISVPs and cores. These processes appear to be

essential for further viral infection, specifically penetration of the vacuolar membrane and entry into the cytoplasm.

The uncoated viral core remains intact during the early stages of infection, carrying viral enzymes into the host cell. These include a transcriptase and helicase, as well as guanylyl transferase and transmethylase activities, required to synthesise, cap and methylate mRNA copies of the viral genome segments.

Coincident with uncoating, the sub-viral particle-associated transcriptase is activated. Transcription in the *Reoviridae* is asymmetric and conservative *i.e.* the negative-sense strand serves as template for mRNA synthesis, but remains in the dsRNA form. The dsRNA is also retained within the core. Distinct ssRNA transcripts representing full-length copies of the genomic plus-strands are synthesised and extruded from the core particle into the cytoplasm. The ssRNA can function as message for the translation of viral proteins by the cellular machinery or as templates upon which progeny dsRNA genomes are made by the synthesis of the complementary negative-sense strand (replication). The latter occurs in particles that are assembled from newly synthesised viral protein and positive-sense template RNAs. Following replication, the minus-strand RNA replica remains associated with the plus-strand template, reconstituting the genomic dsRNA. The core particles bearing the dsRNA can either support additional rounds of transcription or alternatively undergo further maturation to form infectious progeny particles. The steps involved in assortment, assembly and packaging of single copies of each different segment that makes up the viral genome into progeny virions are still not well defined.

The significant complexities of these mechanisms may explain why no truly effective reverse genetics system is as yet available for the members of the *Reoviridae* and why the few reports of initial successes (discussed previously) use methods that are poorly understood.

2.2 Development of reverse genetic systems for the *Reoviridae*

The applicability of reverse genetics to segmented dsRNA viruses, by definition, depends on (i) the availability of a system that permits the introduction of a full complement of genome segments into cells and their assembly into functional genomes and (ii) the ability of this system to insert

foreign or altered genetic information. Various individual aspects of the life cycle of the *Reoviridae*, including transcription, replication and assembly, have been investigated with the aim of understanding the requirements for viral infection, replication and assortment, which are all more or less crucial to the development of effective reverse genetics systems. This approach has involved recombinant VACV and baculovirus expression of virion components, the analysis of reassortants and temperature sensitive mutants.

2.2.1 **Transcription**

In contrast to the members of the *Birnaviridae*, in which the intact virion is an active polymerising complex (Mertens *et al.* 1982; Spies *et al.* 1987), intact *Reoviridae* virions are not transcriptionally active; activation of mRNA synthesis requires a structural alteration involving uncoating. However, although virions can't make full-length transcripts, they have been found in the case of reoviruses to readily synthesise short oligonucleotides when given appropriate substrates, suggesting that the transcriptase is constitutively active in virions, but only capable of limited elongation (Yamakawa *et al.* 1981). Therefore, activation of transcription during uncoating may be a misnomer in the sense that this process does not actually modify the enzyme complex, but rather releases the complex from structural constraints.

Investigation of the viral transcriptase activity in the *Reoviridae* has largely been achieved by *in vitro* uncoating of the virion (removal of specific outer capsid polypeptides) by various chemical or physical treatments, resulting in activation of the transcriptase. These approaches have yielded important information regarding the basic essentials for transcriptase activity, including temperature, salt and pH requirements. In addition, data concerning the relative levels of mRNAs in infected cells have been collected. Thus, it was determined that the *in vitro* transcription reaction by reovirus and rotavirus cores is absolutely dependent on magnesium ions and has an unusually high temperature optimum of 45°C to 50°C (Kapuler 1970; Cohen 1977; Yamakawa *et al.* 1982; Yin *et al.* 1996). The *in vitro* transcriptase reaction of orbiviruses is also dependent on magnesium ions but it has a lower temperature optimum of 28°C to 37°C (Verwoerd & Huismans 1972; Van Dijk & Huismans 1980; Van Dijk & Huismans 1982). Transcription in the orbivirus *Bluetongue virus* (BTV) also differs from that of reoviruses in such a way that the different mRNA species are

not transcribed at a frequency proportional to their molecular weight (Van Dijk & Huismans 1988). Similar results have been reported for other orbiviruses (Huismans *et al.* 1979; Namiki *et al.* 1983). Furthermore, the relative frequency of transcription of the respective BTV genome segments remains the same throughout the infection cycle, another distinctive feature of BTV transcription (Huismans & Verwoerd 1973).

Both moving transcriptase and moving template models have been proposed (Yamakawa *et al.* 1982; Joklik 1983; Shatkin & Kozak 1983); the latter is generally accepted. This model states that complexes of the transcriptase enzymes are bound at specific sites in the inner capsid, necessitating movement of both product and template RNAs during transcription. The latter has been corroborated by numerous structural studies investigating the localisation of the RNA-dependent RNA polymerase within viral particles (Coombs 1998; Dryden *et al.* 1998; Loudon & Roy 1991; Gouet *et al.* 1999), suggesting a fixed binding position for the transcriptase complex within the core on the icosahedral five-fold axes near the base of channels spanning the outer capsids. The function of these channels is not known, but it is believed possible that they are involved in exporting nascent RNA transcripts from the core (Bartlett *et al.* 1974; Prasad *et al.* 1988; Dryden *et al.* 1993). It is proposed that the entire length of each dsRNA gene segment moves past the fixed transcriptase and that the nascent mRNA is directed past the capping enzyme and out through the spikes on the core surface.

2.2.2 **Assortment**

Roner *et al.* (1990) concluded that their results from the infectious reovirus system focussed attention on the assortment process *i.e.* the formation of complexes that contain one of each of the 10 progeny genome segments and then mature into infectious virus particles, as the most critical during the reovirus replication cycle. Indeed, in contrast to other viruses such as influenza virus, where virus particles probably contain random 11-segment collections of the eight actual influenza genome segment species, so that roughly 1 in 25 particles contains at least one of each genome segment and is therefore infectious (Lamb & Choppin 1983; Enami *et al.* 1991), the ratio of virus particles to infectious units is essentially 1 (Spendlove *et al.* 1970; Larson *et al.* 1994).

As such, the assembly of genome segments in the reoviruses is an extraordinarily efficient and precise process.

Assembly of a reovirus particle requires specific signals for encapsidation and selective sorting of individual genome segments. Characterisation of functional deletion fragments of specific genome segments (DI RNAs) has proved useful for studies on viral genome encapsidation and replication (Levis *et al.* 1986; DePolo *et al.* 1987). The analysis of DI RNAs associated with *Wound tumour virus*, a plant virus member of the *Reoviridae*, provided an emerging view of the mechanism underlying packaging (Anzola *et al.* 1987). The minimal sequence information required for replication and packaging of a genome segment was found to be located within the terminal domains of a genome fragment. In addition, packaging of one pair of terminal structures excluded the subsequent packaging of a structure with identical termini. This exclusion mechanism implies the presence of two operational sorting signals in each segment: one signal specifies it as a viral and not a cellular RNA molecule, and the second that it is a particular RNA segment (Anzola *et al.* 1987). All the *Reoviridae* RNA segments have strictly conserved terminal sequences about 4 to 8 bp long, perhaps representing the first sorting signal. Additionally, the genome segments also appear to have a 6- to 9-nucleotide segment-specific inverted repeat immediately adjacent to the conserved terminal sequences (Anzola *et al.* 1987). This could represent the putative second sorting signal needed for segment recognition during encapsidation. The significance of both the 5′ and 3′ terminal regions has been demonstrated by Zou and Brown (1992). The smallest packaging-competent genomic deletion fragment was only 344 bp long (132-135 5′ nucleotides and 183-185 3′ nucleotides) and still contained all the necessary signals for encapsidation. They also predicted a panhandle hairpin structure in the non-coding region at the terminal consensus sequences, suggesting a role for RNA secondary structures in defining the packaging signals.

Recent studies looking to discover and more precisely characterise the nature of the molecular interactions involved in the assembly of reovirus genomes have utilised both the infectious reovirus RNA system described earlier and genome segment reassortment (Joklik 1998). Genome segment reassortment occurs when cells are infected with two species of reovirus particles *e.g.* particles of different serotypes, but not with different genera *e.g.* rotaviruses or orbiviruses. In cells infected with viruses belonging to any two of

the three reovirus serotypes, roughly 15%, but not more, of the progeny are reassortants, the genomes of which contain all possible combinations of parental genome segments in roughly equal proportions (Brown *et al.* 1983). However, no such reassortants were found using the infectious RNA system (Joklik & Roner 1995; Roner *et al.* 1995; Joklik & Roner 1996). It was then shown that two mutations in the S4 genome segment (G74 to A and G624 to A) function as acceptance signals (Roner *et al.* 1995). The presence of these signals appears essential for the acceptance of heterotypic genome segments into the genome, provided that the incoming genome segments possess appropriate recognition signals. The effect of the mutations was traced to the function of protein σ3, most likely through interaction with RNA.

Specific non-structural proteins with intrinsic ssRNA binding activity have been identified that may act as condensing agents to bring together the ssRNA templates for dsRNA synthesis in reoviruses (σNS and μNS; Antczak & Joklik 1992), rotaviruses (NSP2; Gombold *et al.* 1985; Kattoura *et al.* 1992) and orbiviruses (NS2; Thomas *et al.* 1990). Detection of protein-RNA complexes containing both ssRNA and dsRNA suggest that genome segment assortment into progeny genomes is linked to minus-strand synthesis (Antczak & Joklik 1992).

2.2.3 Replication

In the case of rotaviruses, the eleven dsRNAs have been shown to be synthesised in equimolar amounts within replicase particles, leading to virions containing equimolar concentrations of the genome segments (Patton 1990). Analysis of RNA products detected in replicase particles suggests that RNA replication is regulated such that synthesis of full-length dsRNAs proceeds from the smallest to the largest genome segment (Patton & Gallegos 1990). Replicase particles appear to undergo a continuous change in size during RNA replication due apparently to plus-strand RNA templates moving into the replicase particle during synthesis of dsRNA (Patton & Gallegos 1990). Rotavirus single-shelled particles are assembled by the sequential addition of VP2 and VP6 to pre-core replication intermediates consisting of VP1, VP3, VP9, NSP2 and NSP3 (Gallegos & Patton 1989).

With the report of the development of a groundbreaking *in vitro* template-dependent replicase assay for rotaviruses, considerable research has been

carried out to determine the *cis*-acting signals regulating replication of the viral genome (Chen *et al.* 1994b). This assay describes replicase activity associated with sub-viral particles derived from native virions or baculovirus co-expression of rotavirus genes, on native rotavirus mRNA templates or *in vitro* transcripts with bona fide 5′ and 3′ termini. Essentially, the combined data indicate that the core proteins VP1 (the RNA-dependent RNA polymerase) and VP2 (core scaffold) of rotaviruses constitute the minimal replicase particle in the *in vitro* replication system (Zeng *et al.* 1996; Patton *et al.* 1997). Rotavirus VP1 specifically binds to the 3′ end of viral mRNA, but this interaction alone, although required (Chen & Patton 1998), is not sufficient to initiate minus-strand synthesis, requiring the presence of VP2 for replicase activity (Patton 1996). In addition, it was shown that the single-stranded nature of the 3′ end of rotavirus mRNA is essential for efficient dsRNA synthesis (Chen & Patton 1998) and that the 3′ terminal consensus 7 nucleotides of rotavirus mRNA is the minimal promoter of negative-strand RNA synthesis (Wentz *et al.* 1996). Recently, the open reading frame in rotavirus mRNA has also been shown to specifically promote synthesis of dsRNA (Patton *et al.* 1999).

2.2.4 **Assembly**

Another process in the life cycle of the *Reoviridae* playing a role in the development of a reverse genetics system is clearly assembly of the virion. Significant advances in orbivirus research have been made in recent years through the use of the baculovirus expression system. Both BTV virus-like particles (VLPs) and core-like particles (CLPs) (French & Roy 1990; French *et al.* 1990) and AHSV CLPs (Maree *et al.* 1998) have thus been synthesised through self-assembly of the individual components. In addition, protein-protein interaction studies on the components of the virion particle have yielded considerable insight into the intricate organisation and topography of the individual viral components (Le Blois *et al.* 1991; Loudon & Roy 1991; Loudon & Roy 1992). The structure of the core particle of BTV has recently been determined by X-ray crystallography at a resolution approaching 3.5Å (Grimes *et al.* 1998). This transcriptionally active compartment, 700Å in diameter, represents the largest molecular structure determined in such detail.

3. CONCLUSIONS AND AIMS

Thus, many avenues have been explored in the development of reverse genetic systems for DNA and RNA viruses, both segmented and non-segmented. Despite numerous successes in all categories, as enlarged upon above, many viruses still resist breakthroughs, whether due to complexity or simply to genomic size. In the *Reoviridae* considerable progress has been made to the elucidation of the mechanisms of viral transcription and replication. However, the development of a truly efficient reverse genetic system for these complex viruses still eludes researchers. The benefits that such a system would embody are clear, as has been touched upon above. Not least, these include elucidation of structure-function relationships of viral components and the possibility of constructing highly efficient and safe vaccine strains for clinically and economically very important pathogens.

In order to enable the development of reverse genetics systems for the segmented dsRNA viruses, clones of the entire viral genome are required. In the case of *African horse sickness virus* (AHSV), no such library of all the genome segments of a single serotype exists. Specifically, no full-length cDNA clone of the largest genome segment, encoding VP1, of any AHSV serotype has been reported to date. In addition, very little is currently known about the molecular details of the transcription and replication mechanisms of the orbiviruses. These processes are central to the life cycle of the virus and determination of the nature of VP1, the putative viral RNA-dependent RNA polymerase, is thus crucial to an understanding of the molecular biology of the virus and the development of a reverse genetics system. Accordingly, the following aims were envisaged for this study:

- Development of an efficient technique to enable the cloning of complete genomes of the orbiviruses, and cloning of the AHSV VP1 gene.
- Characterisation of the AHSV VP1 gene by sequence determination and analysis.
- Expression and analysis of VP1 as the putative RNA-dependent RNA polymerase of AHSV.

CHAPTER 2

AHSV cDNA SYNTHESIS AND CLONING

1. INTRODUCTION

The ability to clone isolated genes represents an extremely powerful, yet simple technology to allow thorough molecular investigation of the encoded proteins. In the case of a dsRNA genome, such as that of AHSV, it is necessary that cDNA clones of the genes encoding the proteins in question be constructed. A wide array of methods has been developed for this purpose, but the synthesis and cloning of full-length cDNA of the larger genome segments of AHSV, such as the 4 kb genome segment 1, which encodes VP1, has proved difficult.

Historically, the polyadenylation strategy of Cashdollar *et al.* (1982) has served as the foundation for much of the cloning of *Reoviridae* genes carried out to date, including those of reoviruses (Cashdollar *et al.* 1984), rotaviruses (Both *et al.* 1982) and orbiviruses (Purdy *et al.* 1984; Fukusho *et al.* 1989; Yamakawa *et al.* 1999a). This approach involves polyadenylation of genomic dsRNA and cDNA synthesis on denatured dsRNA with oligo(dT) primers, followed by either blunt-ended cloning into a suitable vector or dC-tailing and cloning into dG-tailed *Pst* I-cut pBR322. However, this strategy has important limitations. Firstly, it biases the cloning towards smaller genome segments or truncated cDNAs (unpublished observations). Cashdollar *et al.* (1984) surmounted this problem by fractionation of the cDNA by alkaline agarose gel electrophoresis to optimize the cloning of complete gene copies. Secondly, blunt-ended ligation is notoriously inefficient, whereas the addition of homopolymeric G/C tails has been shown to inhibit the expression of cloned genes (Galili *et al.* 1986). Nel and Huismans (1991) introduced an additional PCR amplification step, with primers specific for the sequenced termini of the cloned gene, in order to remove the homopolymer tails. A recently published modification of the polyadenylation method utilises an adaptor oligo(dT) primer for cDNA synthesis (Shapouri *et al.* 1995). Restriction enzyme sequences were incorporated at the 5' end of the oligo(dT) primer to

simplify cloning of the synthesised cDNA. However, the largest clone obtained using this approach represented a truncated gene of 1505 bp.

PCR amplification of cDNA increases the efficacy of these cloning methods, but requires knowledge of the terminal flanking sequences of the gene of interest. Kowalik *et al.* (1990) and Cooke *et al.* (1991) utilised segment termini-specific primers (based on the sequence conservation of termini within the serogroups of the *Reoviridae*) to selectively synthesise and amplify specific full-length cDNA of the dsRNA genes. However, researchers amplifying cDNAs of genome segments larger than 3 kb resorted to an overlapping RT-PCR approach, utilising primers specific for sequences internal in the gene (Hwang & Li 1993; Hwang *et al.* 1994; Huang *et al.* 1995).

In order to overcome a lack of terminal sequence information, Lambden *et al.* (1992) devised a novel strategy for the cloning of non-cultivatable rotavirus through single primer amplification: a universal oligonucleotide ligated to genomic dsRNA serves as template for cDNA synthesis and amplification with a single complementary primer. However, this approach has been reported to only yield full-length clones of smaller dsRNA segments (Lambden *et al.* 1992; Bigot *et al.* 1995). Bigot *et al.* (1995) used internal segment primers for genes larger than 1.7 kb, thereby obtaining overlapping clones of the 5' and 3' ends of the gene.

In this chapter, the cloning of the 4 kb AHSV genome segment 1 using a novel strategy derived from the above methods is described. The major advantage of this approach is that clones with convenient flanking restriction enzyme sites can be obtained, without any prior sequence information.

2. MATERIALS AND METHODS

2.1 Cells and viruses: A South African isolate of AHSV serotype 1 (AHSV-1), was propagated by limited passaging in suckling mice and thereafter propagated firstly in Vero cells and then in CER (chicken embryo rabbit) cells using modified Eagles' medium supplemented with 5% bovine serum.

2.2 Isolation and purification of viral dsRNA: Genomic AHSV dsRNA was isolated from infected cells and purified by the SDS-phenol extraction method essentially as described by Sakamoto *et al.* (1994). Monolayers of CER cells infected with AHSV-1 at a

multiplicity of infection (MOI) of 10 plaque forming units (pfu)/cell were harvested at 48 h post infection by low speed centrifugation. Cells from 24 Roux flasks were resuspended in 80ml 2mM tris(hydroxymethyl)aminomethane (Tris) pH9.0. Sodium acetate (NaAc) pH5.0 and ethylenediaminetetra-acetic acid (EDTA) were added to final concentrations of 10mM each and then sodium dodecyl sulfate (SDS) was added to a final concentration of 1% (m/v). The pH of the solution was adjusted to 5.0 with glacial acetic acid before extracting the solution with an equal volume of phenol at 60°C (15 min 60°C, 15 min on ice, 15 min centrifugation at 10000g). Phenol residues were removed with two chloroform extractions and the RNA precipitated by the addition of 0.1M NaCl and two volumes ethanol. The precipitate was dissolved in 0.01M STE (10mM NaCl, 10mM Tris-HCl pH7.6, 1mM EDTA) and the ssRNA was removed by salt precipitation in 2M LiCl. The supernatant was diluted with 0.01M STE to 0.2M LiCl and the dsRNA precipitated with two volumes ethanol.

2.3 Oligonucleotide ligation: Three oligonucleotides, with sequences as follows, were synthesised. The molecular weights given are as supplied by the supplier (Syngene or Boehringer Mannheim).

Oligo-1: 5'-GGATCCCGGGAATTCGG-3' (molecular weight = 5413g/mol)

Oligo-2: 5'-CCGAATTCCCGGGATCC-3' (molecular weight = 5115g/mol)

Oligo-3: 5'-GGATCCCGGGAATTCGG$(A)_{17}$-3' (molecular weight = 10933 g/mol)

Oligo-1 and –3 were 5'-phosphorylated to allow ligation to the 3' ends of dsRNA genome segments using T4 RNA ligase as described by Lambden *et al.* (1992), and 3'-terminally linked to an amino group to prevent concatenation of oligonucleotides during ligation. 2µg freeze-dried oligonucleotide was incubated with approximately 10 - 50µg AHSV-1 dsRNA in 60mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) pH 8.0, 18mM $MgCl_2$, 1mM dithiothreitol (DTT), 1mM ATP, 0.6µg bovine serum albumin (BSA), 1/10 volume concentrated dimethyl sulfoxide (DMSO) and 10 units T4 RNA ligase for 16 h at 4°C. Unligated oligonucleotides were removed by spin column chromatography with Promega Sephacryl^R S400 matrix.

2.4 Determination of efficiency of oligonucleotide ligation: 250ng samples of AHSV-1 dsRNA ligated to ^{32}P-labelled oligo-1, and an equivalent amount of unligated dsRNA as a ligation-negative control, were separated by 1% agarose gel electrophoresis. The RNA was transferred to Nylon membrane by electroblotting with a Biorad mini trans-blot cell at 50V for 2 h in 1X TAE (0.04M Tris-acetate, 1mM EDTA pH8.5). The membrane was cut into strips corresponding to the lanes on the agarose gel. Two ten-fold dilution series of the unlabelled oligo-1, from 1µg to 1fg and from 900ng to 9fg, were also slot-blotted onto Nylon membrane strips as standard controls. The membrane strips were hybridised overnight to complementary ^{32}P-labelled oligo-2 in hybridisation

buffer (5X Denhardt's reagent, 5X SSC (750mM NaCl, 75mM trisodium citrate, pH7.0), 0.1% SDS, 50% formamide). One membrane strip blotted with ligated dsRNA was mock hybridised to serve as the hybridisation-negative control. All the membranes were washed in 1X SSC, 0.1% SDS for 15 min at varying temperatures. Membranes blotted with ligated dsRNA were washed at 37°C, 42°C, 50°C, 60°C or 72°C, whereas the control membranes (negative and standard controls) were all washed at 37°C. Following autoradiography, the profiles were analyzed by scanning densitometry using the Roche LumiImager™ and LumiAnalyst™ version 2.1 software.

2.5 dsRNA size fractionation and purification: Oligonucleotide-ligated dsRNA was enriched for the larger genome segments by centrifugation on a 5ml linear density gradient of 5-40% sucrose in 1xTE buffer (10mM Tris pH 7.4, 1mM EDTA pH 8.0). Centrifugation was carried out for 16 h at 48 000 rpm in a Beckman SW50.1 rotor at 4°C. Gradients were fractionated using a gradient tube fractionator (Hoefer Scientific Instruments) and collecting 8-10 drops per fraction. Following agarose gel electrophoretic analysis, fractions containing predominantly (>80%) genome segments 1-3 were pooled, diluted in an equal volume of water and ethanol precipitated.

2.6 cDNA synthesis: Oligonucleotide-ligated dsRNA was denatured in 20mM methylmercuric hydroxide (MMOH) for 10 min at room temperature. cDNA synthesis was carried out at 42°C for 1 h using 1µg oligo(dT)$_{15}$ (Boehringer Mannheim) as primer and 18 units *Avian myeloblastosis virus* (AMV) reverse transcriptase (Promega) in the presence of 50mM Tris-HCl pH8.3, 10mM $MgCl_2$, 70mM KCl, 3mM β-mercaptoethanol, 100 units human placental ribonuclease inhibitor (Amersham), 0.5mM each dNTP and 20µCi α^{32}P-dCTP (>400Ci/mmol, Amersham) in a final volume of 60µl. Thereafter the reaction was diluted to a final volume of 100µl with 1mM Tris-HCl pH8.0 and the cDNA was separated from unincorporated nucleotides by Sephadex G-100 column chromatography in 1mM Tris pH8.0. α^{32}P-dCTP incorporation was monitored by Cerenkov counting. The cDNA fractions in the leading peak were pooled and lyophilised.

2.7 Size separation and purification of cDNA: Lyophilised cDNA was resuspended in a suitable volume of water, with the addition of an equal volume of 10X alkaline buffer (0.3M NaOH, 20mM EDTA) to hydrolyse the RNA. The cDNA samples were then diluted, and bromophenol blue in 40% sucrose was added, to a final concentration of 2.5X alkaline buffer. Separation was effected by vertical 1.5% agarose gel electrophoresis in 1X alkaline buffer at 100mA for 4 h (with one buffer change), followed by autoradiography of the wet agarose gel wrapped in Gladwrap. Gel slices containing cDNA of individual (or multiple moderately separated) genome segments were excised from the gel. An equal volume of

30mM HCl, 10mM Tris pH8.0 was added to the gel slices prior to recovery of the cDNA by Geneclean™ II kit methodology. Alternatively, the cDNA was fractionated by centrifugation on a linear density gradient of 5-40% sucrose in 1X alkaline buffer, as described before. Fractions were analysed by Cerenkov counting and three pools of fractions in the leading peak were collected. Pool samples were analysed by 1.5% vertical alkaline agarose gel electrophoresis and autoradiography to determine the segment representation of each pool. cDNA was recovered by NENsorb (NEN) column chromatography.

2.8 Annealing of cDNA: Recovered cDNA was allowed to anneal in 50mM Tris-HCl pH 7.5, 100mM NaCl, 10mM $MgCl_2$, 1mM DTT buffer by heating to 80°C for 5 min, incubating at 65°C for 16 h and finally cooling to 30°C over 3 h. Partial duplexes were filled in using Klenow as described by Sambrook *et al.* (1989).

2.9 G/C-tailed cloning of cDNA: Double-stranded cDNA samples were cleaned by NENsorb (NEN) column chromatography and lyophilisation prior to dC-tailing with 15 units terminal deoxynucleotide transferase (Gibco BRL) in 100mM potassium cacodylate pH7.2, 2mM $CoCl_2$ and 0.2mM DTT buffer in the presence of 100µM dCTP for 20 min at 37°C. Following repurification by NENsorb (NEN) column chromatography and lyophilisation, cDNA was annealed to 200ng dG-tailed *Pst* I-cut pBR322 (Gibco BRL, stock supplied by Prof. H Huismans, University of Pretoria) in 10mM Tris-HCl pH8, 150mM NaCl, 2mM EDTA for 5 min at 80°C followed by 1 h each at 65°C, 56°C, 42°C and room temperature. Annealed DNA was transformed into competent HB101 cells.

2.10 PCR amplification of cDNA: Oligo-2 was used as single primer for PCR amplification of the cDNA. 1/100 of double-stranded cDNA samples were incubated in a reaction mixture containing 10mM Tris pH8.8, 50mM KCl, 1.5mM $MgCl_2$ 0.1% Triton X-100, 0.2mM each dNTP, 500ng primer-2 and 2.5 units DyNAZyme™ II (Finnzymes Oy). Optimization of PCR conditions allowed amplification of distinct cDNA species representing full-length 3 to 4 kb AHSV genome segments. This entailed 30 cycles of denaturation at 95°C for 30 seconds (120 seconds on the first cycle), annealing at 67°C for 30 seconds and extension at 72°C for 270 seconds (extended to 420 seconds on the final cycle). Amplified material was either first purified from agarose gels by Geneclean™ II methodology or cloned directly.

2.11 Cloning of cDNA: PCR-amplified cDNA was TA cloned into the pMOS*Blue* T-vector (Amersham Life Science) according to the manufacturer's instructions. This system exploits the template-independent preferential addition of a single 3' A residue to

dsDNA by many thermostable polymerases, enabling ligation of the PCR product to compatible single 3' T overhangs at the insertion site in pMOS*Blue*.

2.12 Northern blotting of dsRNA: AHSV-1 dsRNA genome segments were separated by polyacrylamide gel electrophoresis (PAGE) using the buffer system described by Loening (1967). Preparative 6% acrylamide, 0.16% bisacrylamide gels were prepared by polymerization in Loening buffer (40mM Tris-HCl pH7.8, 20mM NaAc, 2mM EDTA) containing 0.08% ammonium peroxodisulfate (m/v) and 0.0008% N,N,N',N'-tetramethylethylenediamine (TEMED) (v/v) and electrophoresis was carried out at 80V for 22 h. After staining in ethidium bromide, the genome segments were visualised by UV fluorescence and their positions blueprinted. The gel was soaked in 0.1N NaOH for 30 min to denature the dsRNA and then washed in 0.5X TAE. The RNA was transferred to Hybond N (Amersham) nylon membrane by electroblotting with a Biorad trans-blot cell in 0.5X TAE for 3 h at 0.8A and fixed to the membrane by UV exposure. The blueprint was used to pinpoint the positions of the genome segments on the membrane. Strips cut from the membrane were then used for hybridisation.

2.13 Sequencing of plasmid DNA: DNA sequencing was carried out by the Sanger *et al.* (1977) dideoxynucleotide chain termination method, using the Sequenase™ Version 2.0 kit (USB) according to the manufacturer's instructions.

2.14 Labelling of probes: Radioactive labelling of plasmid DNA was carried out using the Promega nick translation system according to the manufacturer's instructions. Nicks introduced in DNA by DNase I are translated by a combination of the exonuclease and polymerase functions of DNA polymerase I, incorporating radioactively labelled nucleotides. Oligonucleotides were radiolabelled by phosphorylation with T4 polynucleotide kinase, as described by Sambrook *et al.* (1989).

2.15 Geneclean™ purification of DNA: DNA fragments were isolated and purified from agarose gels by binding to glassmilk using Geneclean™ II kit (Bio101) methodology.

2.16 *In vitro* translation: *In vitro* synthesised RNA was translated with the Promega rabbit reticulocyte lysate or wheat germ extract systems, wherein the lysate or extract contains the cellular components necessary for protein synthesis, according to the manufacturer's instructions.

2.17 Molecular biological manipulation of DNA: All further standard molecular biological manipulations of DNA were carried out as described by Sambrook *et al.* (1989).

3. RESULTS

Prior to and at the commencement of this study, the protocols described by Lambden *et al.* (1992) for oligonucleotide ligation and cDNA synthesis were imitated in our laboratory in pursuit of clones of the large genome segments of AHSV, specifically genome segments 1 and 2. However, this approach only yielded clones of the smaller genome segments, including full-length or partial clones of genome segments 5, 6, 7, 8 and 10 of AHSV-5 (Viljoen & Cloete, personal communication). Accordingly, a more thorough investigation of this approach was initiated to confirm its relevance to our application.

3.1 Efficiency of oligonucleotide ligation to AHSV dsRNA

An experiment to investigate the efficiency of oligonucleotide ligation was performed by hybridisation of an oligo-2 probe to oligo-1 ligated dsRNA and oligo-1 standard dilutions, as described in Materials and Methods. The resultant autoradiograph and the densitometrically-scanned profiles of the pertinent membranes are shown in Figure 2.1. The ligation-negative control (lane 4) represents hybridised but unligated dsRNA (= background), whereas the hybridisation-negative control (lane 10) represents label incorporated by ligation of radiolabelled oligo-1. Representative background profiles were subtracted from each lane to minimize background noise. Thus, the background-corrected intensity of the bands in the ligated dsRNA samples following hybridisation, subtracted by the intensity of the corresponding band in the hybridisation-negative control, reflects the amount of oligo-1 ligated to dsRNA. Lane width was adjusted to the maximum width of the slots and kept constant for all lanes during densitometric analysis, to include all blotted target and minimize effects of uneven distribution. Band width was determined by peaks in the profiles. An equilibration curve was prepared by densitometric analysis of bands 5, 6 and 7 in lane 2, corresponding to radiolabelled oligo-2 hybridising to 900pg, 90pg and 9pg oligo-1 respectively. Higher concentrations of the oligo-1 dilution series were excluded from the equilibration curve due to excessive overspill. The background-corrected intensities of each band on the autoradiograph, determined

as a function of the equilibration curve and given as pg oligo-1, are shown in Table 2.1. Calculations were done as follows:

Considering that 250ng AHSV dsRNA, with a genome size of 19528 bp, was blotted per lane, the number of moles of dsRNA per lane = 250×10^{-9} g / (19528 x 649 g/mol) = 1.97×10^{-14} moles

Therefore, there are $1.97 \times 10^{-14} \times 2$ moles of 3' ends for each AHSV genome segment represented in every lane.

The number of moles of oligo-1 represented in each band (as determined by densitometric profiling) was divided by the calculated number of moles of AHSV genomic 3' ends in that specific band.

Using the sum of band intensities in every lane to make allowances for variation in dsRNA transfer efficiency and varying levels of background noise, the measure of ligation efficiency in lane 3 was calculated as follows:

[(1759.8 – 490.36) x 10^{-12} g / 5413 g/mol] / (1.97 x 10^{-14} mol dsRNA x 20 3' ends) = 0.6

Excluding lane 3 band 5, which is clearly an outlyer, this value varies to a minimum of approximately 0.45 (for lane 3 band 1) with the calculations based on individual bands.

Figure 2.1 Autoradiograph (A) and densitometric profiles (B) following hybridisation of an oligo-2 probe to Southern dot blotted oligo-1 standard dilutions (lanes 1 and 2) and electrophoretically separated and Northern blotted oligo-1 ligated dsRNA (lanes 3 to 10). Densitometrically detected bands are indicated on the left for lanes 1 and 2 and on the right for lanes 3 to 10. Standard 10 times dilutions of oligo-1 range from 1µg (band 1) to 10pg (band 6) in lane 1 and from 900ng (band 2) to 9pg (band 7) in lane 2. The bands in lanes 3 to 10 represent AHSV-1 genome segments 1 (band 1), 2 and 3 (band 2), 4 to 6 (band 3), 7 to 9 (band 4) and 10 (band 5). Lane 4 represents the ligation-negative control and lane 10 the hybridization-negative control. The respective membranes were washed at 37°C (lanes 1 to 4 and 10), 45°C (lane 5), 52°C (lane 6), 60°C (lanes 7 and 8) and 72°C (lane 9).

Table 2.1 Background-corrected densitometric intensities, given as pg oligo-1 relative to lane 2 band 5, of pertinent bands detected on the autoradiograph shown in Figure 2.1.

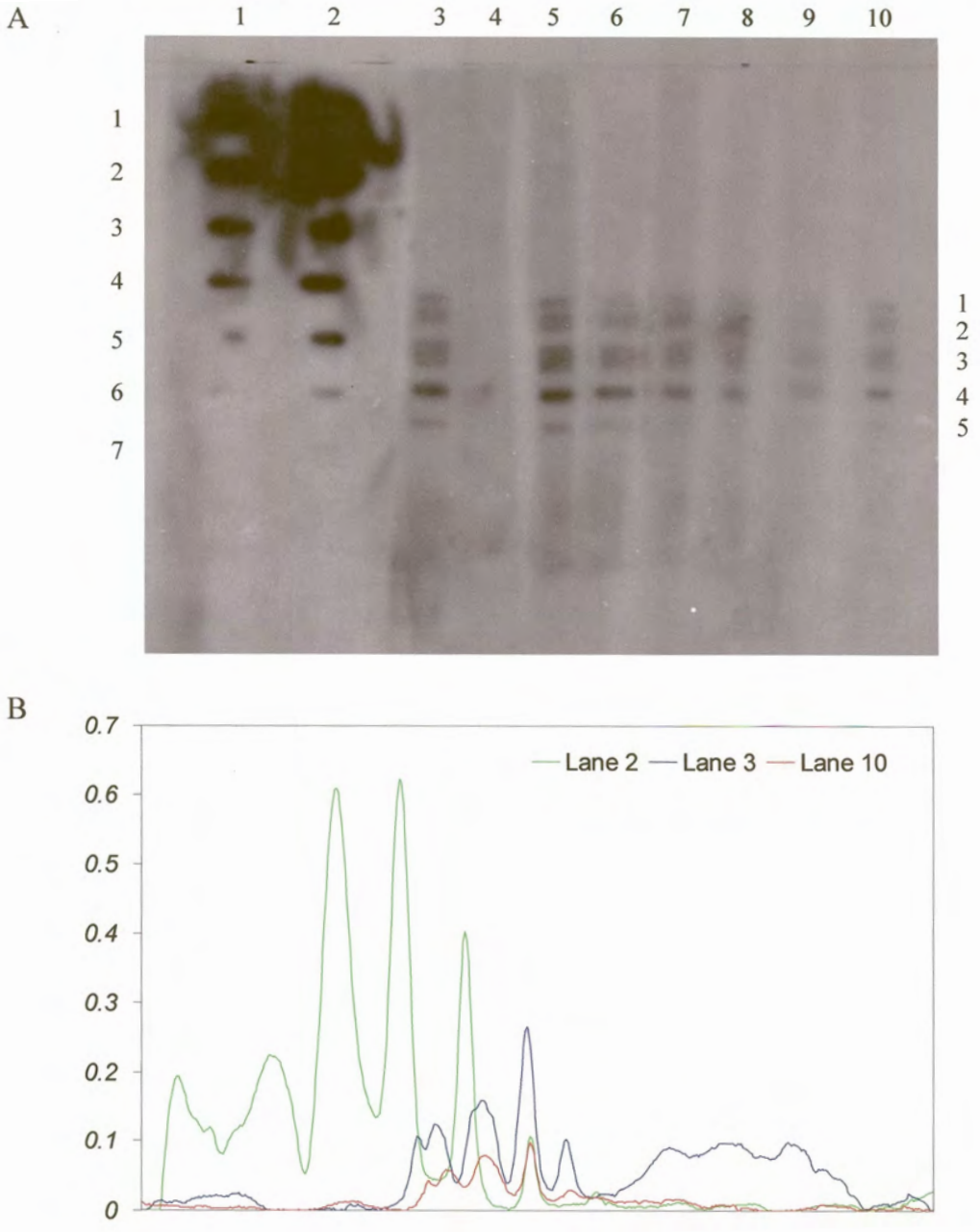


Figure 2.1 (Legend footnoted in text)

Table 2.1 (Legend footnoted in text)

Band	Lane 1	Lane 2	Lane 3	Lane 5	Lane 10
1	956.77		118.14	126.46	21.808
2	284.95	1635.9	285.39	345.71	81.451
3	1361.6	2807.2	623.92	650	249.88
4	1113.2	1932.9	595.26	468.34	110.66
5	180.24	900	137.11	43.373	26.558
6	10.037	92.156			
7		9.0525			
Total			1759.8	1633.9	490.36

Based on this experiment, it appeared that at least 45% of the dsRNA was ligated with oligo-1 after overnight incubation. Thereafter, autoradiographical analysis of ^{32}P-labelled cDNA synthesised from ligated dsRNA by the method described by Lambden *et al.* (1992) and separated by alkaline agarose gel electrophoresis revealed ill-defined smears (results not shown). In contrast, previous work (by the author (Vreede 1994) and others) in the laboratory of Prof. H Huismans at the University of Pretoria using the oligo(dT)-primed strategy for orbivirus cDNA synthesis described by Huismans and Cloete (1987), yielded defined bands corresponding in profile to the AHSV genomic segments.

3.2 Poly(dA)-oligonucleotide ligation strategy for cloning of AHSV genome segments

A hybrid approach utilising facets of the published protocols of Lambden *et al.* (1992) and Huismans and Cloete (1987) was designed and investigated. A schematic diagram of this approach is shown in Figure 2.2

3.2.1 Poly(dA)-oligonucleotide ligation

An oligonucleotide (oligo-3) comprising convenient restriction enzyme sequences and modified by the inclusion of a poly(dA) tail to facilitate oligo(dT) priming for cDNA synthesis was synthesised and ligated to purified AHSV-1 dsRNA, as described. The ^{32}P-labelled oligo-3 ligated dsRNA was purified and enriched for larger genome segments by sucrose gradient centrifugation. Gradient fraction samples analyzed by agarose gel electrophoresis are shown in Figure 2.3. Fractions 4 to 7 were pooled, yielding oligonucleotide-ligated dsRNA enriched for genome segments 1, 2 and 3, as shown.

3.2.2 cDNA synthesis

Following precipitation, this purified and enriched oligo-3 ligated dsRNA served as template for the synthesis of cDNA with an oligo(dT) primer, using the protocols described by Huismans and Cloete (1987). cDNA was size fractionated by either vertical alkaline agarose gel electrophoresis or by alkaline sucrose gradient centrifugation. An autoradiograph of a cDNA sample separated by

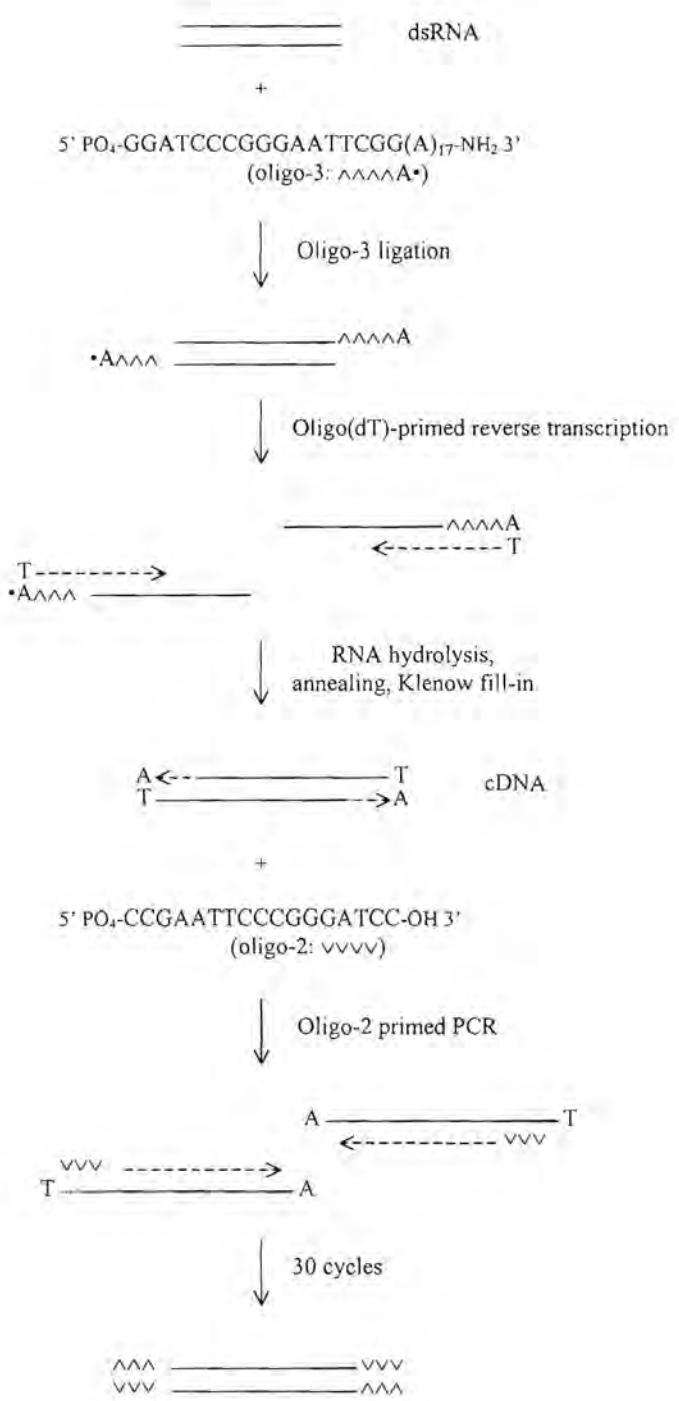


Figure 2.2 Schematic representation of the strategy for synthesis and amplification of full-length cDNA of AHSV dsRNA.

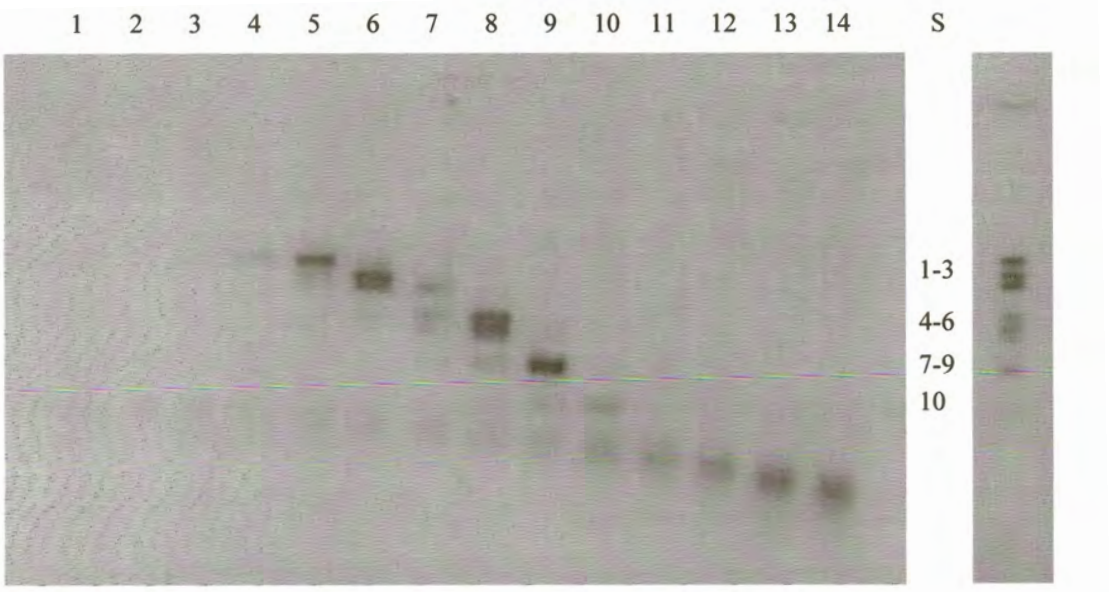


Figure 2.3 Autoradiographs of agarose gel electrophoretic analyses of oligonucleotide-ligated dsRNA fractions, separated by sucrose gradient centrifugation (left), and enriched for large genome segments (right) by combining the fractions shown in lanes 4 to 7. The AHSV genome segments are labelled (S).

alkaline agarose gel electrophoresis is shown in Figure 2.4. cDNA of genome segments 1, 2 and 3 was isolated and purified.

3.2.3 **cDNA amplification and cloning**

Purified cDNA was allowed to anneal and then filled in with Klenow enzyme. In separate experiments with similarly prepared cDNA fractions of smaller genome segments (segments 7 – 9) of AHSV-1 and AHSV-5, attempts at direct cloning by C-tailing and ligating into dG-tailed *Pst* I-cut pBR322 yielded various cloned fragments of approximately 0.4 to 1.6 kb, with single or no terminal oligo-3 specific restriction enzyme sites (results not shown). Subsequently, samples of the larger genome segment cDNA fractions were therefore subjected to PCR amplification with the single complementary primer, oligo-2. In the initial exploratory experiments, purified cDNA fractions of AHSV-1 genome segment 2 and segments 4 to 6 were used as templates for amplification, yielding cDNA of genome segments 1 (very faintly), 2 and 3, as shown in Figure 2.5, or segments 4, 5 and 6 (results not shown) respectively. The individual genome segment 1 cDNA amplicon was isolated and purified by Geneclean™ methodology prior to cloning into pMOS*Blue* (Amersham).

3.2.4 **Clone analysis**

Putative recombinants were analyzed by restriction enzyme digestions and agarose gel electrophoresis and their identity and status as full-length AHSV-1 genome segment 1 clones were confirmed by Northern blot hybridisation (results not shown) and sequencing of the termini (Figure 2.6 and Figure 2.7). The conserved terminal hexanucleotides of AHSV genome segments (Mizukoshi *et al.* 1993) were identified abutting directly onto oligo-2 specific restriction enzyme sequences. One clone containing cDNA representing only the 5' 970 bp of the AHSV-1 genome segment 1 was identified. Despite the non-directional cloning approach, it was observed that all clones of the AHSV-1 VP1 gene in pMOS*Blue* were found to be in the same orientation, and not suitable for *in vitro* transcription from the T7 RNA polymerase promoter. As no other promoter was available for *in vitro* transcription, the genes from four clones were subcloned into pBS (Stratagene), a vector possessing T7 and T3 RNA polymerase promoters flanking the multiple cloning site, using *Xma* I. Once again, despite the non-directionality of the cloning strategy, all subclones were found to be in

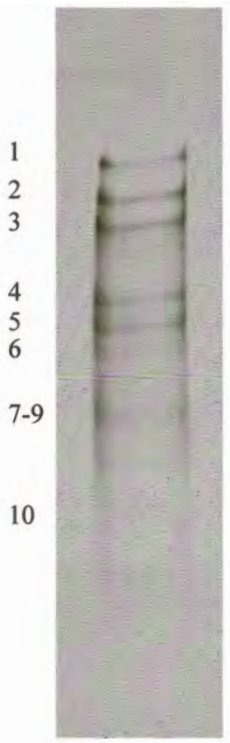


Figure 2.4 An autoradiograph of an alkaline agarose gel electrophoretic analysis of oligo(dT)-primed AHSV-1 specific cDNA synthesized on poly(dA)-oligonucleotide ligated dsRNA enriched for genome segments 1 to 3. The AHSV genome segments are labelled on the left.

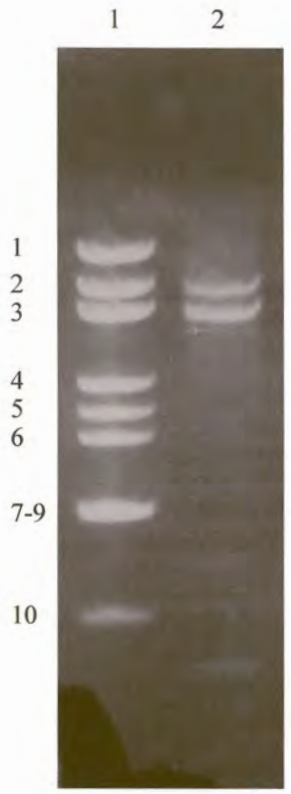


Figure 2.5 Electrophoretic analysis of the PCR amplification products of AHSV segments 2 and 3 cDNA (lane 2). The AHSV genomic dsRNA segments (lane 1) are labelled on the left.

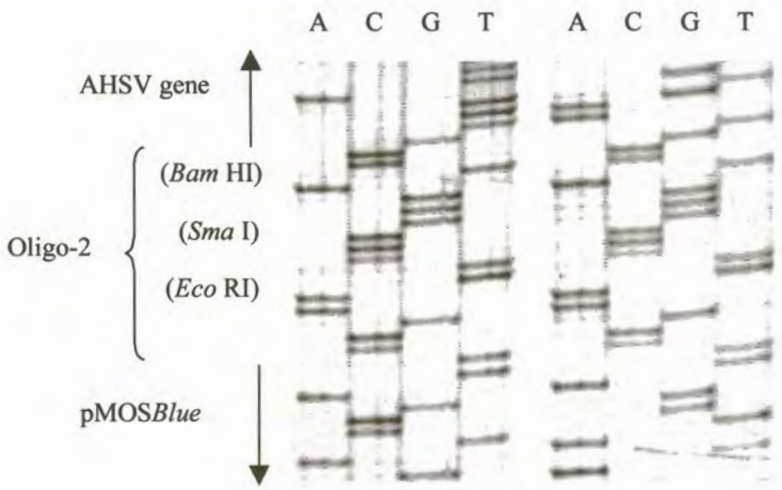


Figure 2.6 Sequences of the pMOS*Blue* vector and AHSV gene clone junction site confirming the full-length status of the cloned gene and demonstrating the incorporation of terminal flanking restriction enzyme sequences. The sequence on the left (GTTTAT...) represents the 5′ consensus sequence of AHSV gene segments and that on the right (GTAAGT...) the 3′ consensus sequence.

```
1     GTTTATTTGAGCGATGGTCATCACCGTGCAAGGTGCAGATCTAGTCAGGAGGGCTTTAAA
      TCGATTATTTAAATATGGGAGGATAGATGGAACTAAAATGTATTATGAGTATTATAGATA
      TTCAAGTAAAATGAGGGAGACTAGGAGGAAGAAAGG.....                    156

756       .....GTTTGCCGCATCCTAAGAAGATAAACAACACGTTGCGCTCACCGTATTCGTGGTT
      TATCAAGAACTGGGGTATCGGATGTCGAAGAGTGAAGGTTTTAACATCGATTGGAGGTGA
      GGATCGGAATTCAAAGGAAGTTTTTTATACCGGTTACCACGAAACAGAGAACCTATACTC
      AGAGATTGTCCAGAAATCAAAATTTTATAGAGAAA.....                     965

3761            .....ACGGGGTTAACGCCATCACGATATGATATTAATGTATCGGGAGACGAAAG
      GGTACGATTTAAGCAGCGCGTTGCCCGATTTAATACACATTTACCCAAGATGCGGATGGT
      CAAAAGATTGATCGAAACGGAGAGGTTGTCCGCGAGGTTGGTTCAGAACCAGTTTGTCTG
      ATTAGAACTAGCACCCCACAGCTCAAAACACTTAC                          3965
```

Figure 2.7 Partial nucleotide sequence of AHSV-1 genome segment 1. Numbering of the nucleotides is based on homology with the sequence of the AHSV-9 genome segment 1, discussed in chapter 3.

the same orientation. T7 RNA polymerase-driven transcripts of *Sph* I-linearised recombinant pBS plasmids were translated with the Promega rabbit reticulocyte lysate and wheat germ extract systems. SDS-PAGE analysis of the translation products revealed no translation products from the AHSV-1 VP1 genes using the rabbit reticulocyte lysate expression system. However, SDS-PAGE analysis of translation products obtained with wheat germ extract established the presence of a protein of the expected size (150kDa) from one clone, confirming the presence of an intact open reading frame, whereas transcripts of at least two further clones yielded products of approximately 28kDa (Figure 2.8).

4. DISCUSSION

As discussed earlier, several sequence-independent methods for the cloning of dsRNA genes have been investigated, but none have proved truly efficient at cloning of the 3 to 4 kb genome segments of the *Reoviridae*. Accordingly, a novel approach to clone the 4 kb AHSV VP1 gene was sought. Hence, an amalgamation of the original polyadenylation method of Cashdollar *et al.* (1982), as described by Huismans and Cloete (1987), for the cloning of orbivirus genes, with the single primer amplification approach of Lambden *et al.* (1992), was investigated. This was anticipated to combine the optimised effectiveness and functionality of the former with the advantage of the sequence independence and convenience of the latter.

The resultant poly(dA)-oligonucleotide ligation method described here is a significant technological advance in routinely obtaining full-length clones of large dsRNA genes. The novel modification presented in this report involved the terminal ligation of an oligonucleotide with a 3' poly(dA) tail to the dsRNA genome segments with T4 RNA ligase as template for oligo(dT)-primed cDNA synthesis. This sequence-independent procedure is rapid and yields full-length cDNA with convenient terminal restriction enzyme sites.

During the course of the development of this technique, an investigation to determine the efficiency of ligation of oligodeoxyribonucleotides to dsRNA with T4 RNA ligase was carried out. This enzyme was first identified for its ability to catalyze the circularisation of ssRNA chains in the presence of ATP (Silber *et al.* 1972), through the formation of a phosphodiester bond between a

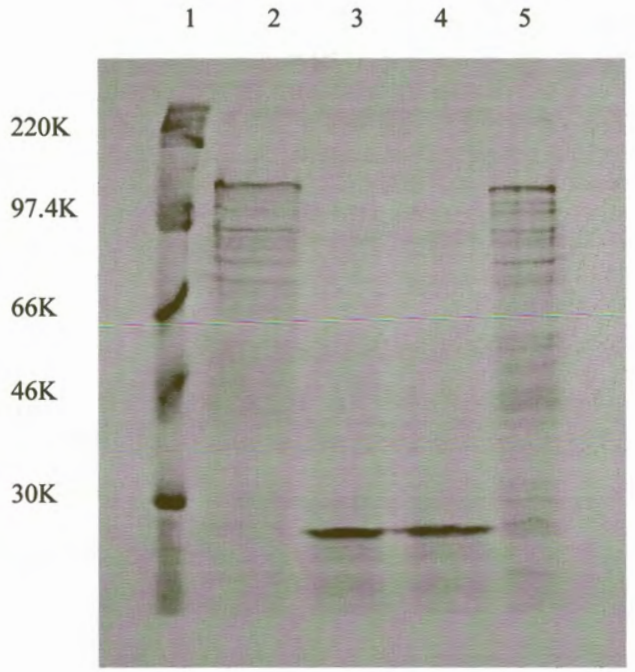


Figure 2.8 Autoradiograph of SDS-PAGE separated *in vitro* translations of AHSV VP1 gene clones. Lane 1 represents ^{14}C-labelled molecular weight marker (labelled on left), lanes 2, 3 and 4 *in vitro* translations of AHSV-1 VP1 gene clones obtained by the poly(dA)-oligonucleotide ligation strategy, and lane 5 *in vitro* translation of an AHSV-9 VP1 gene clone as control (see chapters 3 and 4).

5'-phosphoryl terminated donor and a 3'-hydroxyl terminated acceptor. The enzyme was subsequently also described to ligate oligodeoxyribonucleotides to RNA or DNA acceptors (Snopek *et al.* 1976). More recently, Tessier *et al.* (1986), reinvestigating the ligation of oligodeoxyribonucleotides, pointed out that the yield of ligation decreased from 67% to 40% when the length of the acceptor oligonucleotide increased from 25 to 40 nucleotides. In our hands, analysis of ligation of oligodeoxyribonucleotide to dsRNA indicated that at least 45% of the dsRNA 3' termini were ligated, closely matching the reported figures. However, it should be noted that the analytical approach utilised, namely hybridisation and densitometry of autoradiographed blots, was not necessarily accurately quantitative. This is reflected by a lack of linearity in the signals associated with hybridisation to the dilution series of oligo-1, particularly at higher concentrations. On the other hand, the proportional intensities of bands representing one (segment 1 and segment 10), two (segments 2/3) or three (segments 4/5/6 and segments 7/8/9) genome segments lends validity to the approach.

Utilising the described conditions, excellent yields of full-length cDNA of AHSV genome segments 1, 2 and 3 were achieved through reverse transcription of oligo-3 ligated dsRNA. The quality of cDNA obtained, as adjudged by electrophoresis on alkaline agarose gels, was at least as good as that described by Cashdollar *et al.* (1982) with publication of the original polyadenylation method and that previously obtained by this author using the latter method for synthesis of AHSV genome segment 2 cDNA (Vreede 1994). The former represents the only group thus far to have directly cloned full-length 4 kb *Reoviridae* cDNA successfully. In the course of this study, attempts at direct cloning by G/C-tailing of poly(dA)-oligonucleotide ligation-generated cDNA fractions of genome segments 7 to 9, using severely dated stocks of dG-tailed *Pst* I-cut pBR322, yielded only incomplete segments lacking terminal oligo-3 specific sequences. These clones are probably derived from the slight smearing observed on autoradiographs of electrophoretically separated cDNA. Conservatively, fractions of larger segments were not subjected to direct cloning attempts. It may be postulated that this approach may have proved successful considering the limited presence of incomplete genomic segment cDNA fragments in this region of the alkaline agarose gel.

Nonetheless, PCR amplification of the cDNA using a single complementary primer yielded full-length AHSV genome segment size-specific amplicons, including the largest genome segments, confirming the efficiency of the approach. An intact full-length clone of the 4 kb AHSV-1 VP1 gene was obtained following PCR amplification, as adjudged by the presence of conserved 5′ and 3′ terminal sequences and *in vitro* translation of a 150kDa protein. This represented the first full-length clone of AHSV VP1 obtained to date. Effectively, approximately 200ng total dsRNA was utilised as source material for this cloning. This is proportionally similar to the quantity of material reportedly used by Lambden *et al.* (1992) for cloning the 728 bp rotavirus gene 10 by single primer amplification.

At least two other full-length clones of the AHSV-1 VP1 gene which were obtained in this study following amplification of segment 2-specific cDNA yielded partial proteins on translation, implying the presence of chain terminating mutations in the cloned genes. Taking the renowned lack of fidelity of DNA polymerases used in the PCR into account (Eckert & Kunkel 1991), combined with the 4 kb length of the VP1 gene, it can be postulated that these mutations were most likely introduced during PCR amplification. The more appropriate use of a high fidelity DNA polymerase with 3′ – 5′ exonuclease-dependent proofreading activity (Cline *et al.* 1996) for PCR amplification may have reduced the incidence of mutant clones.

Clones of a number of other genome segments of AHSV, both full-length and partial, were also obtained using this method during this study. Most notably, in terms of its large size, this included an intact full-length clone of the AHSV-5 VP2 gene.

The protocol described here requires careful attention to detail and maintenance of fresh stocks of reagents, but it has nonetheless proved repeatable in different researchers' hands. Indeed, subsequent to the present study, this protocol (with some minor modifications for rationalization) has been used routinely for the cloning of complete genomes of a number of segmented dsRNA viruses, including AHSV, *Equine encephalosis virus* (EEV) and rotavirus, utilising minimal quantities of dsRNA as starting material (Potgieter, personal communication). pGEM-T (Promega), as opposed to pMOS*Blue* (Amersham), was found to be the vector of choice for cloning of the PCR-amplified cDNA, apparently improving yields of full-length clones. Furthermore, by manipulating

the sequence of the oligonucleotides used for ligation, cDNA produced can be cloned into any suitable vector with the appropriate restriction enzyme.

Considering the above discussion, it is clear that the full-length AHSV-1 VP1 gene clone obtained may also contain undetected base substitutions and missense mutations. With this in mind, and with the concomitant availability of an AHSV VP1 gene clone obtained without PCR amplification (described in the following chapter), it was considered prudent to continue the analysis of the AHSV VP1 gene and gene product with this cDNA clone of AHSV-9 genome segment 1.

CHAPTER 3

SEQUENCING OF GENOME SEGMENT 1 OF AHSV-9

1. INTRODUCTION

Virally-encoded RNA-dependent RNA polymerases which mediate the replication of the viral genome are a common feature to all RNA viruses, with the exception of retroviruses. Despite the rapid mutational change that is typical of RNA viruses, enzymes mediating the replication and expression of virus genomes contain functional sequence motifs that appear as most conserved. Kamer and Argos (1984) identified several similar motifs between the known PV RNA-dependent RNA polymerase and the putative RNA-dependent RNA polymerases of several other positive-strand RNA viruses of plant, animal and bacterial origin. Subsequent analyses have extended the range of viruses and identified further conserved motifs (Habili & Symons 1989; Poch *et al.* 1989; Bruenn 1991; Koonin 1991; Koonin & Dolja 1993). Three such conserved sequence motifs show unequivocal conservation, containing a few absolutely conserved amino acids between the positive-strand, negative-strand and dsRNA viruses. These motifs have been suggested to form the catalytic centre of this class of polymerases (Poch *et al.* 1989; Bruenn 1991; Koonin 1991), defining a "signature" by which putative RNA-dependent RNA polymerases can be identified.

In the case of the *Reoviridae*, two processes requiring polymerase activity can be distinguished, namely transcription of dsRNA genome segments into plus-strands and replication of plus-strands to form progeny dsRNA genome segments. As already discussed in chapter 1, considerable advances have been made in unraveling the molecular details of the transcription and replication mechanisms of both rotaviruses and reoviruses, but progress with orbiviruses has been somewhat slower. These processes are central to the viral life cycle and determination of the nature of the RNA polymerase is thus crucial to an understanding of the molecular biology of the virus.

RNA-dependent RNA polymerase sequence motifs have been identified in the deduced amino acid sequences of λ3 of reovirus (Morozov 1989) and VP1 of rotavirus (Cohen *et al.* 1989a). Similarly, the VP1 amino acid sequence predicted from the genome segment 1 gene sequence of BTV has been found to include the proposed characteristic signature of RNA-dependent RNA polymerases (Koonin *et al.* 1989).

VP1, VP4 and VP6 of orbiviruses are minor proteins closely associated with the viral genome and encapsidated by the major proteins VP3 and VP7 to form the core particle. The core particles of AHSV, BTV and *Epizootic haemorrhagic disease virus* (EHDV) have been shown to possess transcriptase activity (Verwoerd & Huismans 1972; Van Dijk & Huismans 1982). VP3 is believed to form the scaffold of the core particle (Roy 1996). VP4 within BTV-derived cores has been demonstrated to covalently bind GTP, advocating VP4 as the candidate guanylyl transferase of the virus (Le Blois *et al.* 1992), whereas sequence analyses of VP6 of AHSV (Turnbull *et al.* 1996) and BTV (Roy 1992) have revealed motifs common to several helicases.

In the case of VP1 of AHSV, genome segment 1 represents the only genome segment that has not been cloned or characterized by sequencing (Roy *et al.* 1994). This chapter describes the cloning, without PCR amplification, and sequencing of the AHSV-9 VP1 gene, and reports the similarities of the predicted VP1 amino acid sequence to other viral RNA-dependent RNA polymerases.

2. MATERIALS AND METHODS

2.1 Cloning and construction of the AHSV-9 VP1 gene: Partial clones of the VP1 gene of AHSV-9, prepared by the author (Vreede 1994) and GB Napier, were available from the laboratory of Prof. H Huismans at the Department of Genetics of the University of Pretoria. Briefly, libraries of AHSV-9 specific clones were generated by standard shotgun cloning techniques using homopolymer tails as previously described (Bremer *et al.* 1990). The method entails the addition of poly(A) tails to the genomic RNA (with or without prior denaturation), cDNA synthesis using oligo(dT) primers and subsequent dC-tailing of the cDNA to permit cloning into dG-tailed *Pst* I-cut pBR322. Genome segment 1-specific clones were identified by Northern blotting and characterized by terminal sequencing and restriction enzyme analysis. Although no full-length segment 1 clones

were identified, as determined by the presence of the terminal consensus sequences, two overlapping clones of 3.2 kb and 3.4 kb, together representing the entire gene, were identified. A unique *Xho* I site within the overlap was exploited to construct a full-length clone (91.pBR). The full-length gene was subsequently also cloned into the pLITMUS29 (NEB; 91.pLITMUS), pGEM3 (Promega; 91.pGEM3) and pGEM3z f(+) (Promega; 91.pGEM3z) vectors.

2.2 Sequencing of the AHSV-9 VP1 gene: An extensive restriction map of the gene was drawn up using a panel of available restriction enzymes. A strategy was devised whereby the complete genome segment sequence could be determined by terminal (<300 nucleotides) sequencing of 17 restriction enzyme fragments (Figure 3.1). The devised strategy involved isolation of the 17 fragments from 8 restriction enzyme digestion reactions of the full-length clones (using different enzyme combinations), and cloning into 6 different linearised vectors (Table 3.1). The fragments as shown in Table 3.1 and Figure 3.1 are hereinafter referred to as # 1 to 17. The fragments were subcloned by standard techniques and terminally sequenced by the dideoxy chain termination method using the Sequenase 2.0 kit (USB) with M13/pUC forward and reverse primers.

2.3 Analysis of the AHSV-9 VP1 gene and gene product sequences: All computer analyses of the RNA-dependent RNA polymerase gene nucleotide sequences and RNA-dependent RNA polymerase amino acid sequences were carried out using the GCG version 8.1 software package (Genetics Computer Group 1994). The fragment assembly programs, namely 'Gelstart', 'Gelenter', 'Gelmerge', 'Gelassemble', 'Gelview' and 'Geldisassemble' were utilised to assimilate the raw sequence data into a complete gene sequence. Nucleotide and amino acid sequence comparisons were carried out with the 'Bestfit' program, or the 'Pileup' program in the case of multiple sequences. Hydrophobicity plots (Kyte & Doolittle 1982) were determined using the ANTHEPROT version 4.0 protein analysis software (written by G. Deleage), available at http://www.ibcp.fr/ANTHEPROT. ClustalX (Thompson *et al.* 1997), available from http://www-igbmc.ustrasbg.fr/BioInfo/ClustalX/Top.html, and MEGA (Kumar *et al.* 1994), available at http://evolgen.biol.metro-u.ac.jp/MEGA/manual/default.html, were used to construct phylogenetic trees, The PHD method of Rost and Sander (1994), as found on the PredictProtein server at http://www.embl-heidelberg.de/predictprotein/, was used to predict the secondary structure of the polymerase sequences.

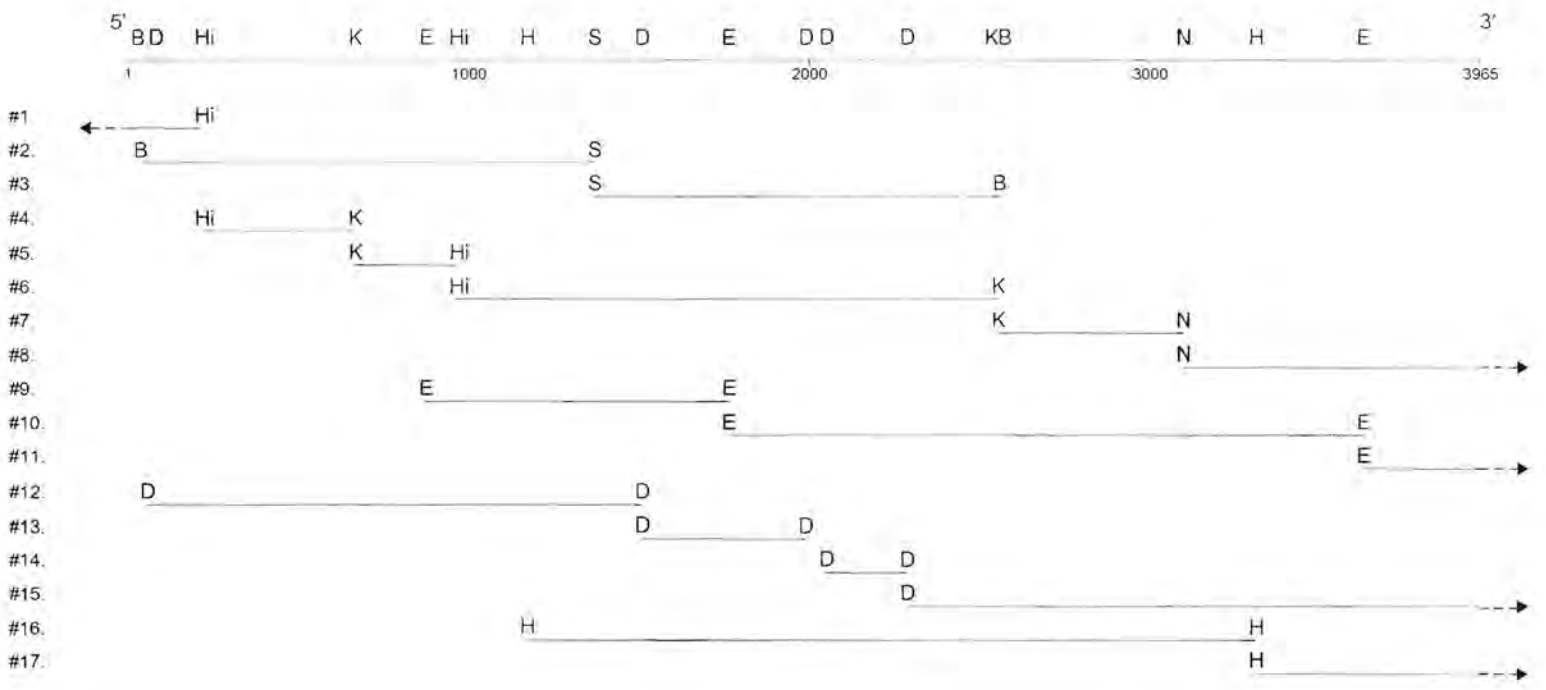


Figure 3.1 Schematic representation of the subcloning strategy for sequencing of the AHSV-9 VP1 gene. A partial restriction enzyme map of the full-length gene is represented. The subclones are numbered according to Table 3.1.

B = *Bgl* II D = *Dra* I E = *Eco* RI H = *Hpa* I Hi = *Hind* III K = *Kpn* I N = *Nsi* I S = *Spe* I

Table 3.1 Summary of the AHSV-9 segment 1 (AHSV-9.1) restriction enzyme fragment subclones utilised for the complete sequencing of the VP1 gene.

#	Restriction Fragment	Origin	AHSV-9.1 nucleotides	Cloning Vector
1.	*Pst* I $_{pBR322}$ - *Hind* III $_{215}$	91.pBR $_{Pst\ I\ -\ Hind\ III}$	1 - 215	pUC19 $_{Pst\ I\ -\ Hind\ III}$
2.	*Bgl* II $_{37}$ - *Spe* I $_{1370}$	91.pBR $_{Bgl\ II\ -\ Spe\ I}$	37 - 1370	pGEM3z f(+) $_{Bam\ HI\ -\ Xba\ I}$
3.	*Spe* I $_{1370}$ - *Bgl* II $_{2561}$	"	1370 - 2561	"
4.	*Hind* III $_{215}$ - *Kpn* I $_{667}$	91.pBR $_{Hind\ III\ -\ Kpn\ I}$	215 - 667	pGEM3z f(+) $_{Hind\ III\ -\ Kpn\ I}$
5.	*Kpn* I $_{667}$ - *Hind* III $_{964}$	"	667 - 964	"
6.	*Hind* III $_{964}$ - *Kpn* I $_{2557}$	"	964 - 2557	"
7.	*Kpn* I $_{2557}$ - *Nsi* I $_{3099}$	91.pGEM3 $_{Kpn\ I\ -\ Nsi\ I}$	2557 - 3099	pGEM3z f(+) $_{Kpn\ I\ -\ Pst\ I}$
8.	*Nsi* I $_{3099}$ - *Kpn* I $_{pGEM3}$	"	3099 - 3965	"
9.	*Eco* RI $_{877}$ - *Eco* RI $_{1762}$	91.pBR $_{Eco\ RI}$	1762 - 877	pGEM3z f(+) $_{Eco\ RI}$
10.	*Eco* RI $_{1762}$ - *Eco* RI $_{3626}$	"	1762 - 3626	"
11.	*Eco* RI $_{3626}$ - *Eco* RI $_{pBR322}$	"	3626 - 3965	"
12.	*Dra* I $_{57}$ - *Dra* I $_{1509}$	91.pBR $_{Dra\ I}$	57 - 1509	pGEM3z f(+) $_{Sma\ I}$
13.	*Dra* I $_{1509}$ - *Dra* I $_{1990}$	"	1509 - 1990	"
14.	*Dra* I $_{2050}$ - *Dra* I $_{2290}$	"	2050 - 2290	"
15.	*Dra* I $_{2290}$ - *Eco* RV $_{pLITMUS29}$	91.pLITMUS $_{Dra\ I\ -\ Eco\ RV}$	2290 - 3965	"
16.	*Hpa* I $_{1174}$ - *Hpa* I $_{3310}$	91.pGEM3 $_{Hpa\ I}$	1174 - 3310	"
17.	*Hpa* I $_{3310}$ - *Sma* I $_{pGEM3}$	"	3310 - 3965	"

3. RESULTS

3.1 Cloning and sequence determination of AHSV-9 genome segment 1

Two partial clones of AHSV-9 genome segment 1, previously obtained by shotgun cDNA synthesis and cloning by the polyadenylation method of Cashdollar *et al.* (1982), were identified as overlapping and representative of the full-length gene by restriction enzyme mapping and terminal sequencing. The two clones were spliced using an internal restriction enzyme site to prepare a full-length cDNA clone. The complete sequence was determined by terminal (less than 300 nucleotides) sequencing of 11 restriction enzyme fragment subclones (# 1 to 11), as shown in Table 3.1 and Figure 3.1. Ambiguous sequences or areas with strong secondary structures were subsequently elucidated by subcloning and terminal sequencing of 6 additional restriction enzyme fragments (# 12 to 17). The AHSV-9 genome segment 1 nucleotide sequence is shown in Figure 3.2, together with the predicted amino acid sequence of the VP1 protein.

3.2 Sequence analysis of AHSV-9 genome segment 1 and its encoded protein

Genome segment 1 of AHSV-9 is 3965 nucleotides in length with a base composition of 31.6% A, 15.8% C, 24.8% G and 27.8% U residues. The 5′ and 3′ terminal hexanucleotides are 5′ GUUUAU and ACUUAC 3′ respectively, supporting the consensus AHSV terminal sequences proposed by Mizukoshi *et al.* (1993), namely 5′ GUU$^{A}_{U}$A$^{A}_{U}$ and $^{A}_{C}$C$^{A}_{U}$UAC 3′. The longest open reading frame was defined by an AUG at position 14-16 and a UGA at position 3929-3931, delineating terminal non-coding regions of 13 and 34 nucleotides respectively. The AUG flanking sequences, specifically a purine at position -3 and a guanosine at position +4, place this codon in a favourable context for initiation of translation according to the consensus sequences identified by Kozak (1984). Translation of this open reading frame yields a protein comprising

---

Figure 3.2 Nucleotide sequence of AHSV-9 genome segment 1 and the translated amino acid sequence (numbered) of AHSV-9 VP1. Non-coding nucleotides are shown in lower case. Motifs IV, V and VI (underlined) represent conserved amino acid sequence motifs in RNA-dependent RNA polymerases.

```
                                                gtttatttgagcgATGGTCATCACCGTGCAAGGTGCAGATCTAGTCAGGAGGGCTTTAAATCGATTATTTAAGTATGGGAGGATAGATGGAACTAAGATGT
1                                                M  V  I  T  V  Q  G  A  D  L  V  R  R  A  L  N  R  L  F  K  Y  G  R  I  D  G  T  K  M  Y                  30

ATTATGAATATTACAGATATTCAAGTAAAATGAGGGAGACTAGGAGGAAAAAAGGGACGAAGTATAAAACAGATGATGAGTTTTTAGAGCGCGAGAGGGATGCGGGGAGGTTGAAGCTTT
31   Y  E  Y  Y  R  Y  S  S  K  M  R  E  T  R  R  K  K  G  T  K  Y  K  T  D  D  E  F  L  E  R  E  R  D  A  G  R  L  K  L  Y   70

ATGATCTACAAGTGATACGGGAAGCGTCTTGGGAAGATTTGTTGTATGAAAATGTTCACACTGCTGAGTTAGATATTTATGTTAGATCAATCTTGAAATTAGAGGACTTAGAGCCAGAAG
71   D  L  Q  V  I  R  E  A  S  W  E  D  L  L  Y  E  N  V  H  T  A  E  L  D  I  Y  V  R  S  I  L  K  L  E  D  L  E  P  E  E  110

AAGAATTTTTGCGAAATTACGCAGTTTATGATGGTGTGCACCCGCTGAAGGATTTCGTTGAAATGAGAGCGAAAAATGAAATGCAGATTTTTGGTGATATGCCCATTAAGGCGTGGATCT
111  E  F  L  R  N  Y  A  V  Y  D  G  V  H  P  L  K  D  F  V  E  M  R  A  K  N  E  M  Q  I  F  G  D  M  P  I  K  A  W  I  S  150

CAGTATTGATGCAAATCTCGCGCGAAACAAAGCATAAACCTCTTGGGTTAATGGTAGCTTCGGATTTTGTCGGAAGATTTGGTTCGCCTTTTGAACAGAATTTTAGAGATTTGTCACAGA
151  V  L  M  E  I  S  R  E  T  K  H  K  P  L  G  L  M  V  A  S  D  P  V  G  R  F  G  S  P  F  E  Q  N  F  R  D  L  S  Q  I  190

TTAACGAATATGGTTATTGTTACTCAAGTCCATTATTATTTGAGATGTGTGTAACTGAGTCAATTTTGGAATTTAATATGTGGTACCGGATGCGTGAAGAAAGAATACAGTCGCTGAAAT
191  N  E  Y  G  Y  C  Y  S  S  P  L  L  F  E  M  C  V  T  E  S  I  L  E  F  N  M  W  Y  R  M  R  E  E  R  I  Q  S  L  K  F  230

TCGGATTGGAAGTGATTGATCCTTTCAAACTGATACGAGAATTTTTTGAAATATGTTTGCCGCATCCTAAGAAGATAAACAACACGTTGCGCTCACCGTATTCGTGGTTTGTCAAGAATT
231  G  L  E  V  I  D  P  F  K  L  I  R  E  F  F  E  I  C  L  P  H  P  K  K  I  N  N  T  L  R  S  P  Y  S  W  F  V  K  N  W  270

GGGGTATTGGATGTTCGAGAGTGAAGGTTTTAACATCGATTGGAGGTGAGGATCGGAATTCAAAGGAAGTTTTTTATACTGGTTACCATGAAACAGAGAACCTATACTCAGAGATTGTCC
271  G  I  G  C  S  R  V  K  V  L  T  S  I  G  G  E  D  R  N  S  K  E  V  F  Y  T  G  Y  H  E  T  E  N  L  Y  S  E  I  V  L  310

TGAAATCGAAATTTTATAGAGAAAGCTTGAAACAAAATATGACGAAAACGGAAGAGGCGATTACGTATTCGCAAAAACTTGGCAATCATGGGAGAACGATGCCCATTTTCCTGAAAATGT
311  K  S  K  F  Y  R  E  S  L  K  Q  N  M  T  K  T  E  E  A  I  T  Y  S  Q  K  L  G  N  H  G  R  T  M  P  I  F  L  K  M  L  350

TAAAAGCGGTATATACTACAGAGTTTGATCCAACAAAAATAAGTCACGTTATATTGGCGTCATTATGCTTAAGTATACAAACAATAACAGGATATGGGAGGGCATGGGTAGTTAACAAAT
351  K  A  V  Y  T  T  E  F  D  P  T  K  I  S  H  V  I  L  A  S  L  C  L  S  I  Q  T  I  T  G  Y  G  R  A  W  V  V  N  K  S  390

CTAGTGATTTGGAAGCGCAGATGAAACCAAGTAGTGACAACTATGTACAGCGCGTATGCGATTATACAAAAAATAACTTTATAAAAGCCTATGAAGAGGCGAGGCGAGGGGGTGAAGAGA
391  S  D  L  E  A  Q  M  K  P  S  S  D  N  Y  V  Q  R  V  C  D  Y  T  K  N  N  F  I  K  A  Y  E  E  A  R  R  G  G  E  E  I  430

TTGTAATGCCTGAAGATATGTATACATCGATTTTACGTCTCGCTAAAAATACAAGCTCGGGTTTCTCAACTAGTATTGATGTTTTTAAGCGGTATGGCCCCAACGCGAAGGGtGGACGTG
431  V  M  P  E  D  M  Y  T  S  I  L  R  L  A  K  N  T  S  S  G  F  S  T  S  I  D  V  F  K  R  Y  G  P  N  A  K  G  G  R  G  470

GAGAAAAGATTCAAATTACGTCGCGGATTAAAGCGTTAGTCATTTTTACAAAAGGTCATGAAATATTTACGCCAAAAAACCTAGCTTTAAAGTATAACACTACTGAATTTTTCCAAACGA
471  E  K  I  Q  I  T  S  R  I  K  A  L  V  I  F  T  K  G  H  E  I  F  T  P  K  N  L  A  L  K  Y  N  T  T  E  F  F  Q  T  K  510

AGGGTTCGAGAGATGTTCCGATTAAGTCGACCAGAATAGTGTATTCCATCAACCTCTCAATTTTAGTTCCGCAGCTCATAGTGACATTACCTTTGAATGAGTATTTTGCAAGGGCGGGAG
511  G  S  R  D  V  P  I  K  S  T  R  I  V  Y  S  I  N  L  S  I  L  V  P  Q  L  I  V  T  L  P  L  N  E  Y  F  A  R  A  G  G  550

GAAGTACTTTACCTGAAACTCAACGTATGGGTGGAAAAATTATCGTTGGAGATTTGGAGGCTACAGGATCGCGTGTGATGGACGCAGCTGATACATTTAGGAATTCTTCAGATCCACTTA
551  S  T  L  P  E  T  Q  R  M  G  G  K  I  I  V  G  D  L  E  A  T  G  S  R  V  M  D  A  A  D  T  F  R  N  S  S  D  P  L  N  590

ACTTAACTATCGCGATTGACTACAGCGAATTTGATACGCATTTAACACCTTATAATTTCAGGAACGGGATGCTAGATGGTATACGAGAAGCGATGCGCAGATATCAACATCTGAGATATG
591  L  T  I  A  I  D  Y  S  E  F  D  T  H  L  T  P  Y  N  F  R  N  G  M  L  D  G  I  R  E  A  M  R  R  Y  Q  H  L  R  Y  E  630
                                                motif IV

AGGGTTATACATTAGACGAATTAATAGAGTTTGGATATGGAGAGGGTCGCGTAATGAATACATTATGGAATGGAAAGAGGCGAGTGTTTAAAGTGGCGTTTGAGGATTACGTTATGTTGA
631  G  Y  T  L  D  E  L  I  E  F  G  Y  G  E  G  R  V  M  N  T  L  W  N  G  K  R  R  V  F  K  V  A  F  E  D  Y  V  M  L  S  670

GTGATGAAGATAAAGTTCAGGGAACCTTTAAACCACCCATTGGGGTGAAGCCGGTAAAGAATATTAAAATTTGTGAAGAGCTGGAGAAAAAGGCGGATGGCCGTGATTTGATTTTGGTTT
671  D  E  D  K  V  Q  G  T  F  K  P  P  I  G  V  K  P  V  K  N  I  K  I  C  E  E  L  E  K  K  A  D  G  R  D  L  I  L  V  S  710

CGCCTACGGATGGAAGCGACCTAGCGCTGATTAACACTCATTTATCCGGGGAGAACTCCACGTTGATTGCGAACAGTTTACATAATCTAGCGATTGGAACAGTTATTCGTGAGGAGGTTA
711  P  T  D  G  S  D  L  A  L  I  N  T  H  L  S  G  E  N  S  T  L  I  A  N  S  L  H  N  L  A  I  G  T  V  I  R  E  E  V  K  750
                                                motif V

AACGCATTTTTGGCGATGATATTTCATTTAAATCAGAACAGTATGTTGGTGACGACACCTTATTTTACACTGAATTGAGAACGCGGTCAGTTGAGCGGTTCGATTCAATCGTGGATACGA
751  R  I  F  G  D  D  I  S  F  K  S  E  Q  Y  V  G  D  D  T  L  F  Y  T  E  L  R  T  R  S  V  E  R  F  D  S  I  V  D  T  I  790
                                                motif VI

TTTTTGAAGTCATTAAGAAGAGTGGCCACGAAGCCTCAATGTCAAAAACTCTTATAGCTCCCTTTTCGGTCGAGAAAACACAGACGCACGCTAAGCAGGGAATATATATTCCACAGGATC
791  F  E  V  I  K  K  S  G  H  E  A  S  M  S  K  T  L  I  A  P  F  S  V  E  K  T  Q  T  H  A  K  Q  G  I  Y  I  P  Q  D  R  830

GAATGATGCTAGTTTCGTCCGAACGAAGAAAAGATATAGAGGATGTTGCTGGGTACCTGAGATCTCAAGTTCAAACGTTGACAACAAAAATTAGTCGTGGTTTCTCGCATGAGCTGGCGC
831  M  M  L  V  S  S  E  R  R  K  D  I  E  D  V  A  G  Y  L  R  S  Q  V  Q  T  L  T  T  K  I  S  R  G  F  S  H  E  L  A  Q  870

AAATCATTTTTATGATGAAATCGTCGATTATCGGGCATAGGAAACTAAAGAGGACAATTAAGGATGGAGGATATAGAGATAGAAAATACGATGATGACAAAGAGGACGGCTTCACTCTGA
871  I  I  F  M  M  K  S  S  I  I  G  H  R  K  L  K  R  T  I  K  D  G  G  Y  R  D  R  K  Y  D  D  D  K  E  D  G  F  T  L  I  910

TAATGTTGCGAGATCCGCTTATTGCTTTTTATCCCGTTGAATGGAACGGCTTTGGAGCGCACCCGGCAGCGATGAATATAATAATGACTGAAGATATGTTCGTAGACTCAGTGATGAGAG
911  M  L  R  D  P  L  I  A  F  Y  P  V  E  W  N  G  F  G  A  H  P  A  A  M  N  I  I  M  T  E  D  M  F  V  D  S  V  M  R  G  950

GGGAATGTAGAGCGTGGATGGAACCCTTAGTTAAGTTGATTGATCAATCTCCTCCCTTGTGGAACGAAACGAGCGCGGATAAAAGAATGATTGGTACGGATAGTACGATGTCTTTTTTCT
951  E  C  R  A  W  M  E  P  L  V  K  L  I  D  Q  S  P  P  L  W  N  E  T  S  A  D  K  R  M  I  G  T  D  S  T  M  S  F  F  S  990

CGAGAATGGCAAGGCCAGCGGTACGAACGGTATTAACCAACTCTGAGGTGGGAGATGCCGTAAGAAGTTTGCCTCTGGGAGATTTCTCTCCTTTCAATATCTCTAAAACTATGATGCATT
991  R  M  A  R  P  A  V  R  T  V  L  T  N  S  E  V  G  D  A  V  R  S  L  P  L  G  D  P  S  P  F  N  I  S  K  T  M  M  H  S  1030

CGGCGTTATTGAAAGAAAAGAATGCCCGTTCATTACTTACCCCAGCGTATGAGATGGAGTATCAAAAAGAGTTGCAAGGTTGGAGACCAAGGCAGAAAAAGTTCTTAGTTACGTCTAACG
1031  A  L  L  K  E  K  N  A  R  S  L  L  T  P  A  Y  E  M  E  Y  Q  K  E  L  Q  G  W  R  P  R  Q  K  K  F  L  V  T  S  N  E  1070

AGATGGAAATTACTACAAATTATATGAAGATGTTTAACGTTGGTAAAATTCCACTGCACGGTTTGGCGTTGAAGTTTTTCCCAGATGTTAACTTGTCGAAAGAATTTTTCTTACAGAAGA
1071  M  E  I  T  T  N  Y  M  K  M  F  N  V  G  K  I  P  L  H  G  L  A  L  K  F  F  P  D  V  N  L  S  K  E  F  F  L  Q  K  S  1110

GTGTTTTAGGGAATAGAGAGAGTCCGAGAGCACGAATGTCATACGTCGATAGAATTGACTCAATACTACGTGGCGATGTCGTAATGCGAGGGTTTATCACTGCGAATACTATTATCAACA
1111  V  L  G  N  R  E  S  P  R  A  R  M  S  Y  V  D  R  I  D  S  I  L  R  G  D  V  V  M  R  G  F  I  T  A  N  T  I  I  N  I  1150

TATTGGAGAAATTAGGGCACACACATTCGGCGAGTGACTTAACCACATTATTTGAGATAATGAATTTGTCATCAAGCGTCGCTCAGAGGCTATCAGAGTACATCACGACCGAACGCGTTC
1151  L  E  K  L  G  H  T  H  S  A  S  D  L  T  T  L  F  E  I  M  N  L  S  S  S  V  A  Q  R  L  S  E  Y  I  T  T  E  R  V  R  1190

GCTTCGATGCCATGAAACTCTCGAAGAGGGGTATATGTGGTGATGAATTCTCAATGTCATTGGATGTTTGCACACAAACGATGGTAGATAGATATATACGTGCGCCGACGCAGTTCACTA
1191  F  D  A  M  K  L  S  K  R  G  I  C  G  D  E  F  S  M  S  L  D  V  C  T  Q  T  M  V  D  R  Y  I  R  A  P  T  Q  F  T  K  1230

AGACGGAGCTTGATGCCGTAAACCTATATGTTGCCCAACATATTATGTTAGATGCAGCGACTGGATTGACGCCATCACGATATGATATTAATGTGTCTGGAGACGAAAGGGTGCGATTTA
1231  T  E  L  D  A  V  N  L  Y  V  A  Q  H  I  M  L  D  A  A  T  G  L  T  P  S  R  Y  D  I  N  V  S  G  D  E  R  V  R  F  X  1270

AGCAGCGTGTTGCCCGATTTAACACACACTTGCCTAAGATGCGGATGGTTAAAAGATTGATTGAAACAGAAAGGTTGTCCGCGAGATTGGTTCAGAACCAGTTTGTCTGAtcagaactag
1271  Q  R  V  A  R  F  N  T  H  L  P  K  M  R  M  V  K  R  L  I  E  T  E  R  L  S  A  R  L  V  Q  N  Q  F  V  *                1305

caccccctcagctcaaaacacttac
```

1305 amino acids (Figure 3.2) with a predicted molecular weight of 150.292kDa and an overall charge of +25.5 at neutral pH.

The AHSV-9 VP1 amino acid sequence was found to include the proposed characteristic signature

$DX_3[FYWLCA]X_{0-1}DX_n[STM]GX_3TX_3[NE]X_n[GS]DD$

of RNA-dependent RNA polymerases of positive-strand and some dsRNA viruses (X indicates an unspecified amino acid residue; alternative amino acids at particular sites are shown in square brackets; Koonin & Dolja 1993). The three conserved motifs defining this signature are indicated in Figure 3.2 (labelled IV, V and VI as specified by Koonin (1991), located between residues 591-605, 720-741 and 762-771 respectively.

3.3 Observation with reference to colony morphology and insert orientation

An incidental observation made during non-directional subcloning of complete and partial AHSV genome segment 1 cDNA is of interest. In certain recombinants one orientation of the segment 1-specific insert in the plasmid appeared to affect the morphology, or specifically to decrease the size, of the transformed bacterial colony on agar plates (Figure 3.3) and its growth in liquid culture. This effect will loosely be referred to hereinafter as 'toxicity'. The most clear-cut and repeatable distinctions appeared with non-directional subcloning of the full-length AHSV-9 VP1 gene (with terminal GC tails) and the 5'-proximal *Dra* I fragment (nucleotides 57 through 1509; # 3) into pGEM3z f(+), with repeatable, though not absolutely unequivocal, prediction of the orientation of the insert based on the colony morphology possible. Deletion of the 3' *Sal* I fragment (nucleotides 1567 through 3965) from clones of 91.pGEM3z in the 'toxic' orientation did not alter or restore 'normal' morphological characteristics of the colony. In all these cases, small colony sizes translated into noticeably reduced DNA yields in standard mini-preparations from liquid culture relative to DNA prepared from 'normal' colony-size cultures (Figure 3.4). Although similar orientation-dependent colony sizes were distinguishable in the subcloning of segment 1 5' distal fragments (# 10, 15 and 16), the distinctions were less apparent, and no significant differences in DNA yield from mini-preparations could be observed. A summary of the observed effects is provided in Table 3.2 and Figure 3.5.

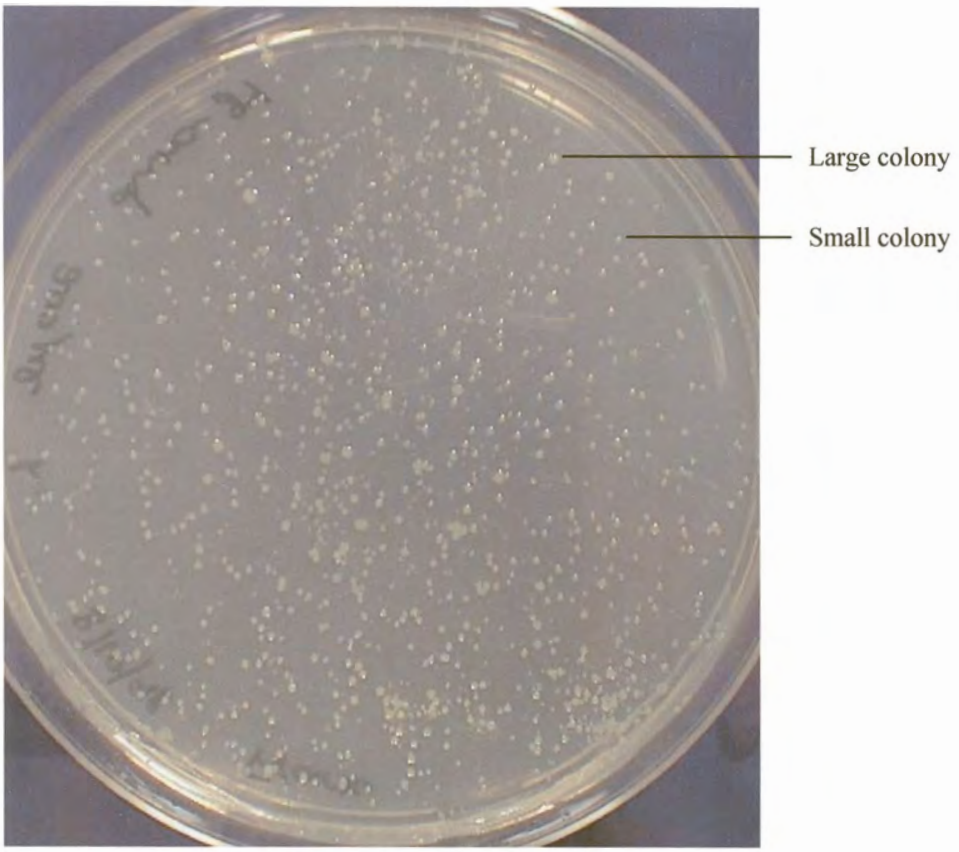


Figure 3.3 Agar plate showing colonies of *E. coli* (DH5α) transformed with recombinant pGEM3z f(+) plasmid containing full length AHSV-9 genome segment 1 cDNA. Examples of a distinctly larger and a distinctly smaller colony are labelled, although colonies of less distinguishable size are also present.

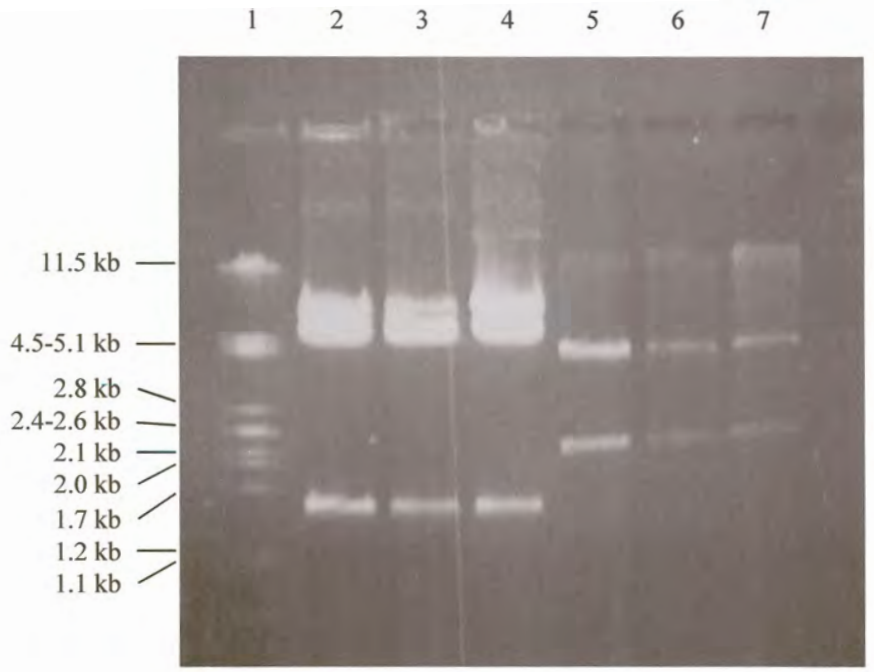


Figure 3.4 Agarose gel electrophoretic analysis of 91.pGEM3z plasmid DNA extracted from *E. coli* colonies and digested with *Sal* I. Lane 1 represents λ-*Pst* I marker. Plasmids originate from either larger (lanes 2 to 4) or smaller (lanes 5 to 7) colonies and possess the AHSV-9 VP1 gene in opposite (lanes 2 to 4) or like (lanes 5 to 7) orientation to the *lacZ* gene on the pGEM3z vector.

Table 3.2 Summary of the phenotypic effects of cloning the AHSV-9 VP1 gene-specific fragments shown in Figure 3.5 in opposite orientations in plasmid vectors. The clones are graphically depicted in Figure 3.5 below. Clone numbers are cross-referenced (in parentheses) to Table 3.1 where applicable. "*lacZ* +" indicates the VP1 reading frame of the insert being in the same orientation as the reading frame of the β galactosidase gene (*lacZ*) in the vector, and "*lacZ* –" in the opposite orientation.

Clone #	'Toxic' (*lacZ* +) *cf* 'non-toxic' (*lacZ* -) orientation		Comment
	Colony size	Plasmid yield	
a	smaller	lower	Terminal dG tails
b	smaller	lower	5' dG tail
c (#12)	smaller	lower	
d (#9)	unknown	unknown	Only *lacZ* – orientation obtained
e (#16)	smaller	equal	VP1 ORF out of *lacZ* reading frame
f (#10)	smaller	equal	
g (#15)	indistinctly smaller	equal	VP1 ORF out of *lacZ* reading frame

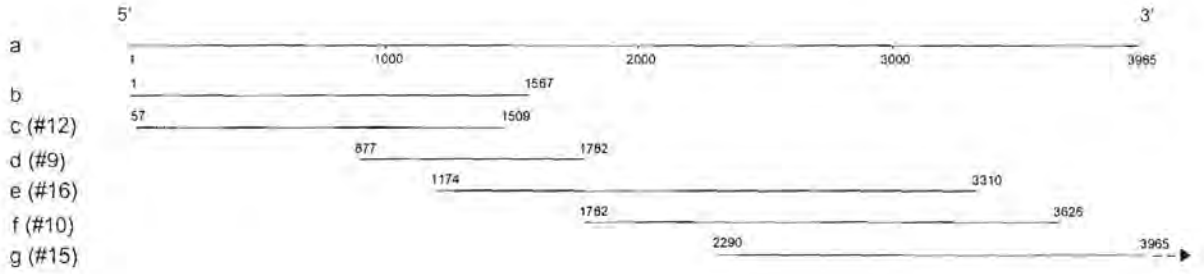


Figure 3.5 Summary of AHSV-9 VP1 gene-specific fragments (a to g) displaying phenotypic discrimination when cloned in different orientations into plasmid vectors. The clone numbers are cross-referenced (in parentheses) to Table 3.1 where applicable. Phenotypic effects are summarized in Table 3.2 above.

4. DISCUSSION

The complete genomes of a number of different members of the family *Reoviridae* have been cloned, including reovirus serotype 3 (*Orthoreovirus*; Cashdollar *et al.* 1984), rotavirus SA11 (*Rotavirus*; Mitchell *et al.* 1990), BTV-10 (*Orbivirus*; Fukusho *et al.* 1989), Chuzan virus (*Orbivirus* (*Palyam virus* serogroup); Yamakawa *et al.* 1999b), *Rice dwarf virus* (RDV, *Phytoreovirus*; Uyeda *et al.* 1995) and the *Nilaparvata lugens* reovirus (NLRV, *Fijivirus*; Nakashima *et al.* 1996). Besides the expression of the encoded proteins, cloning of the genome greatly facilitates investigation into the molecular biology of the virus. Characterization of the cloned genes through nucleotide sequencing allows derivation of the amino acid sequence of the encoded gene product, enabling detailed comparisons of cognate genes and proteins and provides insights into the possible function of the protein and hence its role in the morphogenesis of the virus.

In this chapter, the cloning and characterization by sequencing of the AHSV-9 genome segment 1 gene has been described. This completes the cloning and sequencing of all the AHSV genome segments, albeit of different serotypes. While taking note of the inter-serotype variation, combining the sequences of AHSV-9 segments 1, 5, 6, 8 and 10, AHSV-3 segments 2 and 9 and AHSV-4 segments 3 and 4 yields the first composite AHSV genome (summarized in Table 3.3) comprising 19528 base pairs (*cf.* the BTV-10 genome comprises 19218 bp (Fukusho *et al.* 1989) and Chuzan virus 18915 bp (Yamakawa *et al.* 1999b).

Limited AHSV inter-serotype comparisons of the VP1 gene were possible using sequences generated from AHSV-1 genome segment 1 clones obtained earlier in this study, as described in chapter 2. Conservation of nucleotides varied from 92% to 96%, whereas the conservation of predicted amino acid sequences varied from 95% to 100%, in different short sequenced regions (Table 3.4).

Inter-serogroup comparisons of sequences of the VP1 genes (and the deduced proteins) of AHSV-9, BTV-10 (Accession # X12819; Roy *et al.* 1988) and Chuzan virus (Accession # AB018086; Yamakawa *et al.* 1999b) were performed by best-fit alignments using the default parameters of GCG. Results of pairwise comparisons are shown in Table 3.5. The amino acid sequence of

Table 3.3 AHSV genome segments and the encoded proteins.

Genome segment[a]	serotype	segment length	terminal sequences of coding strand		5' UTR[b]	3' UTR[b]	viral protein	protein length	predicted M_r	location[c]
		(bp)	5'	3'	(bp)	(bp)		(aa)		
1[1]	9	3965	GUUUAU	ACUUAC	13	34	VP1	1305	150292	IC
2[2]	3	3221	GUUUAA	ACUUAC	12	35	VP2	1057	123063	C
3[3]	4	2792	GUUAAU	ACUUAC	26	48	VP3	905	103269	OC
4[4]	4	1978	GUUUAU	CCUUAC	11	38	VP4	642	75826	IC
5[5]	9	1566	GUUUAU	ACAUAC	19	29	VP5	505	56771	C
6[6]	9	1748	GUUAAA	ACUUAC	35	66	NS1	548	63377	NS
7[7]	9	1167	GUUUAA	ACUUAC	17	103	VP7	349	37916	OC
8[8]	9	1166	GUUUAA	ACAUAC	22	46	NS2	365	41193	NS
9[9]	3	1169	GUUAAA	ACUUAC	17	42	VP6	369	38464	IC
10[10]	9	756	GUUUAA	ACUUAC	18	84	NS3	217	23659	NS
							NS3A	206	22481	NS

[a] The relevant Genbank accession numbers and source references are footnoted:

[1] U94887 (present study); [2] U01832 [24]; [3] M94681 [6]; [4] D14402 [12]; [5] U74489 [2]; [6] U01069 (Nel, unpublished); [7] U90337 (Maree *et al.*, submitted); [8] M69090 [21]; [9] U19881 [18]; [10] D12480 [20].

[b] untranslated region.

[c] IC = inner core; OC = outer core; C = outer capsid; NS = non-structural.

Table 3.4 Inter-serotype conservation of AHSV-1 and AHSV-9 VP1 genes and proteins.

Location	Nucleotides			Amino acids			
	Residue #	Length	% identity	Residue #	Length	% identity	% similarity
5' terminal	1-156	156	96.0	1-47	47	100	100
internal	756-965	210	95.0	249-317	69	95.0	96.0
3' terminal	3761-3965	205	91.7	1250-1305	56	100	100

Table 3.5 Inter-serogroup conservation of orbivirus VP1 genes and proteins.

	Nucleotides		Amino acids		
	Gaps (# residues)	% identity	Gaps (# residues)	% identity	% similarity
AHSV : BTV	30 (115)	60.4	9 (10)	55.9	66.5
AHSV : Chuzan	17 (58)	62.9	7 (10)	64.0	72.5
BTV : Chuzan	26 (99)	61.2	8 (10)	55.5	67.2

Chuzan virus VP1 used in this study was as published, utilising the second ATG to initiate the open reading frame. However, based on these comparisons, it is suspected that there may be a minor error in the published sequence between the first and second ATGs.

Further alignments of the amino acid sequences of the other AHSV proteins (as listed in Table 3.3) with the cognate BTV-10 (Fukusho *et al.* 1989) and Chuzan virus (Yamakawa *et al.* 1999a; Yamakawa *et al.* 1999b) proteins revealed that the inter-serogroup conservation of the putative viral RNA polymerase is second only to the inner core protein VP3, which forms the basic scaffold of the core particle (results not shown).

Single examples of each of the genera of the family *Reoviridae* for which RNA-dependent RNA polymerase amino acid sequences have been determined were selected, as listed in Table 3.6, for intra-family comparisons. Automated alignments of these RNA-dependent RNA polymerase sequences reveal very limited homology, even in low stringency pairwise alignments requiring gaps of 40 to 50 residues: identities of approximately 18% and similarities of approximately 50% were obtained in best-fit alignments with a gap weight of 3 and a gap length weight of 0.1. This lack of similarity among the dsRNA viruses of higher eucaryotes has been previously noted (Bruenn 1993). A plot of the similarity along all the RNA-dependent RNA polymerase amino acid sequences listed in Table 3.6, aligned by 'Pileup' with a gap weight of 1.0 and a gap length weight of 0.1, is shown in Figure 3.6. Despite a generally very low level of similarity, there are clearly 4 peaks of higher similarity in the central portion of the RNA-dependent RNA polymerase. These peaks represent polymerase motifs I/II, IV, V and VI as defined by Koonin (1991). Through visual inspection and motif-constrained sequence similarity alignments based on this plot, the alignments were manipulated to assist the identification and definition of further conserved motifs. Specifically, more stringent alignment of the central portion of these RNA-dependent RNA polymerases by 'Pileup', using a gap weight of 1.6 and a gap length weight of 0.1, exposed a number of conserved residues defining putative motifs, which are shown and correlated with previously identified motifs in Table 3.7. The alignments and the identity of some motifs as shown here are sometimes partially at variance with the versions suggested by other workers (Bruenn 1991; Suzuki *et al.* 1992; Buck 1996; Nakashima *et al.* 1996; Upadhyaya *et al.* 1998). The identities of motif D of Poch *et al.* (1989)

Table 3.6 Summary of the viral RNA-dependent RNA polymerases of various genera within the *Reoviridae* utilised for multiple sequence alignments.

Genus	Virus	RdRp # amino acids	Molecular weight (kDa)	Accession #
Orbivirus	AHSV-9	1305	150.3	U94887
Rotavirus	*Bovine rotavirus*	1088	125.0	J04346
Orthoreovirus	*Reovirus-3*	1267	142.2	M31058
Fijivirus (previously *Phytoreovirus*)	RDV	1444	164.4	U73201
Oryzavirus	*Rice ragged stunt virus* (RRSV)	1255	141.4	U66714
Unclassified (*Fijivirus*?)	NLRV	1442	165.9	D49693

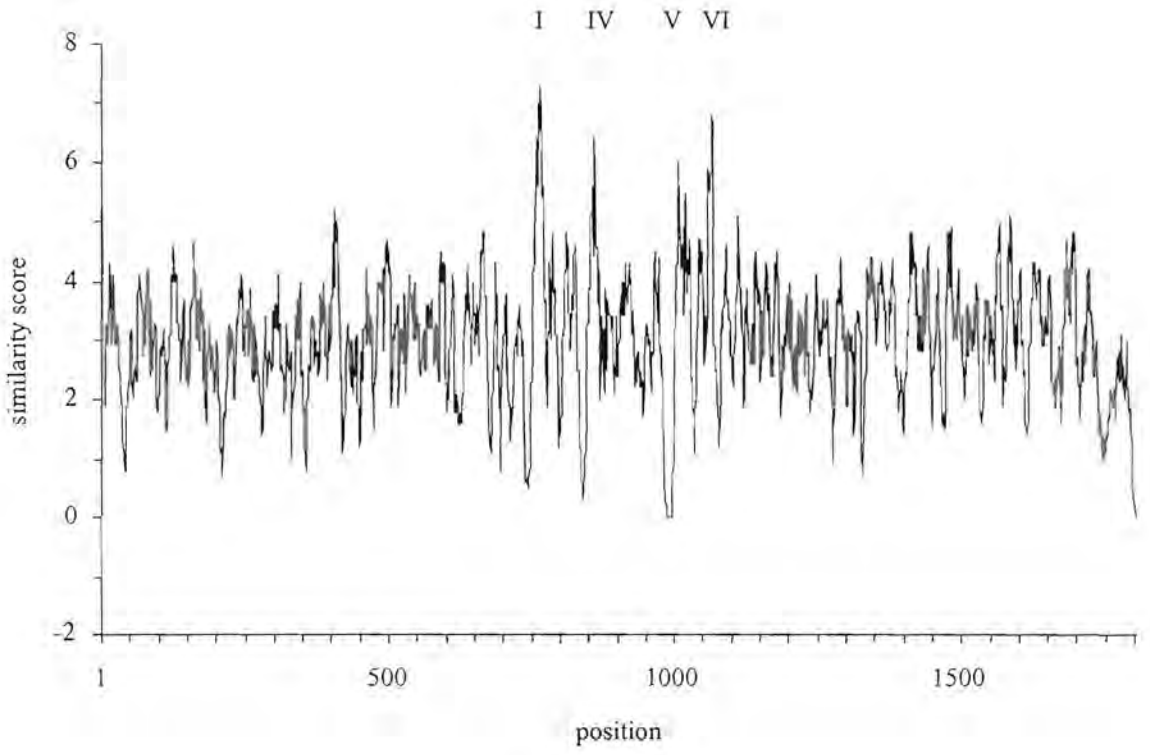


Figure 3.6 A similarity plot of the RNA-dependent RNA polymerase amino acid sequences of the dsRNA viruses listed in Table 3.5. The positions of the conserved RNA-dependent RNA polymerase motifs I, IV, V and VI in the aligned amino acid sequences are indicated.

Table 3.7 Alignment of conserved sequence motifs of viral RNA-dependent RNA polymerases within the family *Reoviridae*. The motifs are designated according to a) Koonin (1991; I to VIII), b) Poch *et al.* (1989; A to C) and c) Bruenn (1991; 1 to 6). The most highly conserved residues defining the polymerase domain are highlighted in bold. The number of amino acid residues between the motifs and the first and last amino acid residues in the alignment are indicated. Gaps introduced to optimize the alignments are represented by dots. Conserved amino acid residues are indicated by *, while : denotes conserved substitutions and . semi-conserved substitutions.

a		I		II		IV		V		VI		VII		VIII	
b						A		B		C					
c						1		2		3		5		6	
AHSV	511	GSRDVPI.KSTRIVYSIN	66	AIDYSEFDT	121	LSGENSTLIANSLHN	24	QYVGDDT	43	VEKTQTHAKQGI	7	MLVSSERR	839
rotavirus	450	GRRDVPG.RRTRIIFILP	51	YTDVSQWDS	63	ASGEKQTKAANSIAN	22	RVDGDDN	41	IEIAKRYIAGGK	7	NLLNNEKR	701
RDV	637	GSRSTTAWRPVRPIY.IN	68	LADCSSWDQ	76	WSGRLDTFFMNSVQN	23	QVAGDDA	44	GEYAKIYYYAGM	7	QLHESEKD	922
reovirus	516	GLRNQVQ.RRPRSIMPLN	50	NIDISACDA	89	PSGSTATSTEHTANN	34	VCQGDDG	42	AEYLKLYFIFGC	8	PIVGKERA	806
RRSV	495	VGRTQIG.RRQRAIAGVN	51	SADVDAMDA	88	PSGQPFTTVHHTFTL	24	TVQGDDT	37	SEYLQQRVSCGT	8	SLFAAERP	770
NLRV	642	V.RHQID.RRGRIIVIVP	54	SSDMSGMDA	90	MSGLFATSGQHTMFL	22	YVMGDDI	40	AQFLQQVALNGV	8	GVFCDEKS	922
		* : * : :		* . *		** * ..		***		: : *		: *:	

and motif 4 of Bruenn (1991) were especially unpronounced in these alignments and could not be unequivocally identified.

Based on the alignments, a signature for the RNA-dependent RNA polymerases of the *Reoviridae* can be defined as follows:
$Rx_{4-5}[RK]x_2Rx[VI]x_nDx[SD]x_2Dx_nSGx_3Tx_3[NH][ST]x_nGDDx_n[QE]x_2[KQ]x_5Gx_{8-10}[LIV]x_3E[RK]$
(where x indicates an unspecified amino acid residue and alternative amino acids at particular sites are shown in square brackets).

The RNA-dependent RNA polymerase amino acid sequences of all the members of the *Reoviridae* for which these have been determined thus-far fit this bill. A second GDD sequence motif found between residues 287 and 289 of VP1 of five U.S. BTV's (Huang *et al.* 1995) was also present in Chuzan virus VP1, but aligned to a GED motif in AHSV-9 VP1. The significance, if any, of these motifs is presently not known. A second GDD sequence motif identified between residues 754-756 of AHSV-9 VP1 was not present in either BTV or Chuzan virus VP1.

The vast majority of RNA-dependent RNA polymerases have been identified solely on the basis of sequence similarity. Availability of the amino acid sequence enables the prediction of secondary structure that is presumed to be relatively inflexible due to the functionality of the protein. Although these predictions are of limited reliability, their concordance over numerous sequences may lend more credence. Hydropathicity plots of the RNA-dependent RNA polymerases of the viruses given in Table 3.6 are shown in Figure 3.7. The central polymerase "modules" of the individual proteins are indicated, without any clear indication of structural conservation.

Following determination of the crystal structure of the RNA-dependent RNA polymerase of PV, Hansen *et al.* (1997) found that, whereas the structures of the 'fingers' and 'thumb' sub-domains of PV polymerase differ from those of other polymerases, the 'palm' sub-domain contains a core structure very similar to that of other polymerases. This core structure is composed of the conserved amino acid sequence motifs A to D (IV to VII) described for RNA polymerases. O'Reilly and Kao (1998) have recently published an analysis of RNA-dependent RNA polymerase structure and function guided by known polymerase structures and computer predictions of secondary structure. In this case, the PHD method of Rost and Sander (1994) was used to predict the secondary structure of the RNA-dependent RNA polymerases from six different positive ssRNA viral families

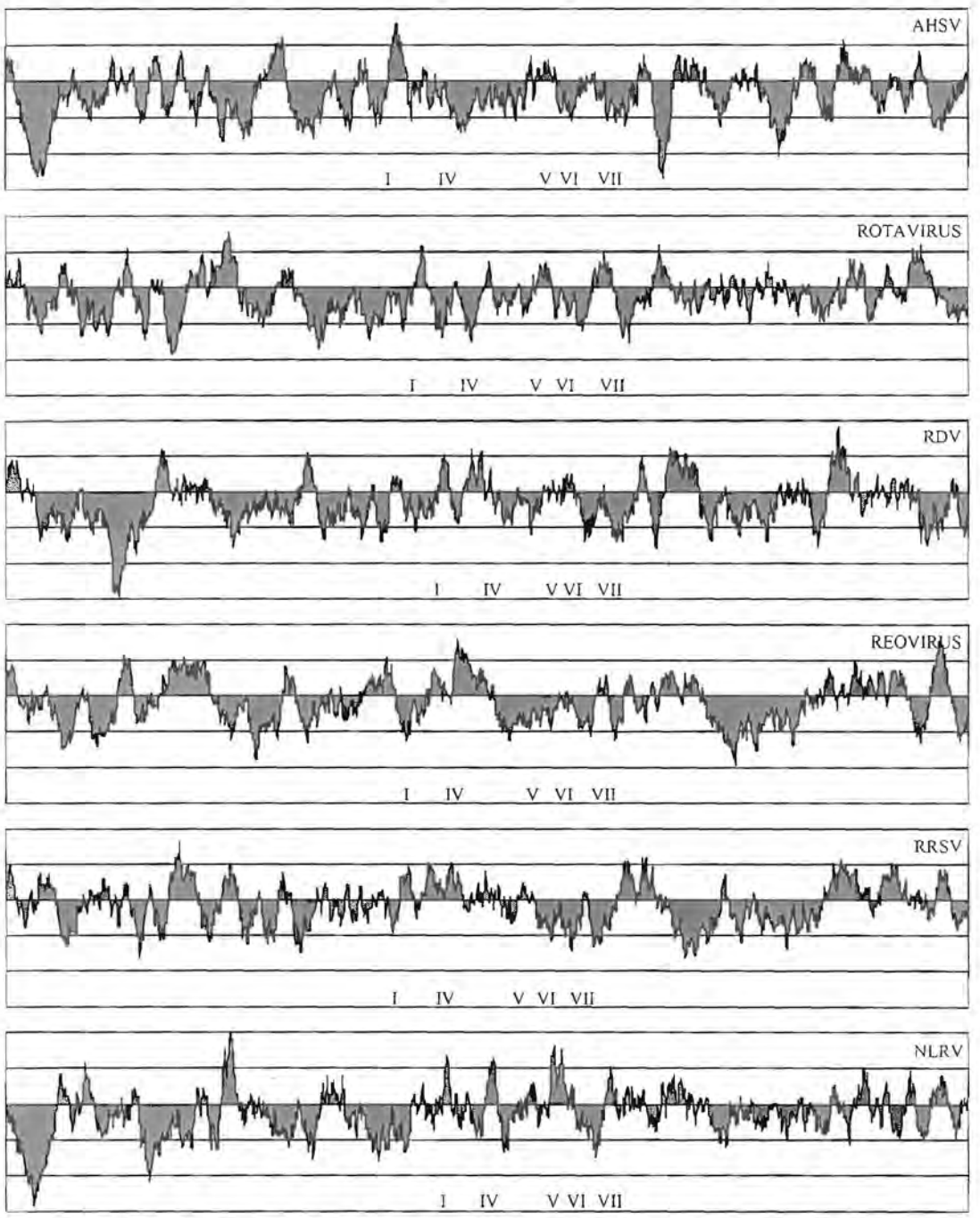


Figure 3.7 Kyte-Doolittle hydropathicity plots of the RNA-dependent RNA polymerase amino acid sequences of the dsRNA viruses listed in Table 3.5, showing regions with a net hydrophobicity (positive values) and a net hydrophilicity (negative values) using a window size of 21. I, IV, V, VI and VII indicate the positions of the conserved RNA-dependent RNA polymerase motifs in the respective amino acid sequences.

(bromoviruses, tobamoviruses, tombusviruses, leviviruses, hepatitis C-like viruses and picornaviruses) and compared with the known crystal structure of PV. The PHD method (for Predict at Heidelberg) uses a neural network that has been trained on more than 130 crystallized protein chains of known crystal structures. To generate more accurate predictions, the program uses sequence alignments rather than single sequences. However, the best alignments are those without redundant information that have sequence identities ranging from 90% to 30% identity. Within the *Reoviridae*, only the AHSV, BTV and Chuzan virus RNA-dependent RNA polymerase sequence alignment fulfils this criterion and was submitted for PHD protein prediction to the 'PredictProtein' server at http://www.embl-heidelberg.de/predictprotein. The predicted secondary structures in the vicinity of the most highly conserved polymerase motifs (A through C) showed significant similarity with the results of O'Reilly and Kao (1998). The identity of motifs D and E, the latter of RNA-dependent polymerases, in AHSV VP1 could be determined by comparison of the predicted secondary structure with these and other (Hansen *et al.* 1997) structure-based sequence alignments. The motifs, along with their predicted secondary structures and those of the PV RNA-dependent RNA polymerase as identified by Hansen *et al.* (1997), are shown in Table 3.8.

Due largely to their functional conservation, viral RNA-dependent RNA polymerases have long been regarded as a primary phylogenetic marker, although this premise has recently been questioned (Zanotto *et al.* 1996). The latter authors argue that RNA-dependent RNA polymerase sequences cannot be used to construct a single phylogenetic tree including all RNA viruses because of a lack of both basic sequence similarity and reliable phylogenetic signal. It was felt more appropriate to present the evolutionary relationships between RNA viruses as a set of distinct sub-trees, the links between which are unclear. Phylogenetic analysis of selected members of the *Reoviridae* was performed, based on a multiple alignment of the amino acid sequences of their RNA-dependent RNA polymerases. The multiple alignment was carried out using 'Pileup' with a gap weight of 1 and a gap length weight of 0.1, thereby allowing maximum similarity and alignment of the conserved polymerase motifs. A homology matrix of the alignment is shown in Table 3.9. Phylogenetic analyses were subsequently carried out with the full data set of complete amino acid sequences (comprising a best possible alignment of 1771 residues) and a data

Table 3.8 Comparative predicted secondary structure (indicating α-helical and β-strand regions) of conserved sequence motifs of the RNA-dependent RNA polymerases of AHSV and PV. The motifs are designated A to E according to Poch *et al.* (1989). Only the extents of structural agreements of PV RNA-dependent RNA polymerase with other polymerases as identified by Hansen *et al.* (1997) are shown, with exceptions that include highly conserved residues printed in italics. The most highly conserved residues are highlighted in bold and underlined. The number of amino acid residues between the motifs, and the first and last amino acid residues in the alignment are indicated.

		A		B		C		D		E	
AHSV	589	LNLTIAI**D**YSEF**D**T**T**L	110	SDLALINTHLS**G**ENS**T**LIANSLHNLAIGTVI	12	SFKSEQYV**GDD**TLFYTEL	9	SIVDTIFEVIKKSGHEASMS**K**TLI	14	GIYIPQDRMM	832
		ββββββββ		α αααααααααααααααα		βββ ββ βββ		αααααααααααα β βββ		βββ βββ	
PV	226	EEKLFAF**D**YTGY**D**ASL	36	*KTYCVKGGMPS**G**CS*G**T**SIFNSMINNLIIRTL	10	DHLKMIAY**GDD**VIASYPH	2	DASLLAQSGKDYGLTM*TPAD**K**S*AT	6	ENVTFLKRFF	378
		βββββββββ ααααα		αααααααααααααααα		βββββββ βββββββ		αααααααααααα βββββ		βββTTβββ	

Table 3.9 Homology matrix of the RNA-dependent RNA polymerases of some representative members of different genera within the family *Reoviridae*, showing percentage identity of amino acid sequences in a multiple alignment.

	AHSV	Chuzan	BTV	RDV	rotavirus	RRSV	NLRV	reovirus
AHSV	100%							
Chuzan	64.5%	100%						
BTV	56.6%	56.3%	100%					
RDV	19.5%	20.7%	19.4%	100%				
rotavirus	21.2%	21.2%	20.3%	31.0%	100%			
RRSV	14.1%	15.3%	13.4%	13.4%	14.7%	100%		
NLRV	15.6%	15.8%	15.5%	11.3%	13.0%	29.4%	100%	
reovirus	14.1%	13.5%	13.4%	13.0%	10.4%	25.4%	23.3%	100%

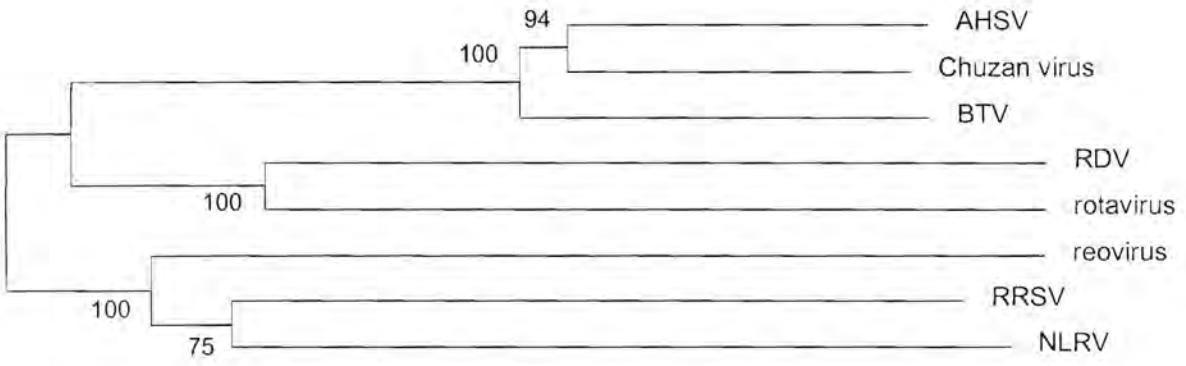
set including only the region of conserved sequence motifs shown in Table 3.7, with the 10 flanking amino acids (comprising a best possible alignment of 371 amino acids). Neighbour joining trees were constructed using Clustalx (Thompson *et al.* 1997) and MEGA (Kumar *et al.* 1994). In the latter case, various genetic distance calculation methods (p-distances, Poisson correction and gamma distances), with or without complete deletion of the gaps introduced for alignment purposes, were employed. Identical topologies, shown in Figure 3.8A and B, were obtained throughout. Furthermore, the resulting phylogenetic trees show high levels of bootstrap support (based on 1000 replications) for intra-genus clustering of the orbivirus species and inter-genus distinction within the family. However, phylogenetic analyses wherein a more distantly related virus (IBDV, family *Birnaviridae*), was included as an outgroup, failed to distinguish between the families (Figure 3.8C). Thus, the data suggest that the sequence similarities and phylogenetic signals from RNA polymerase sequences are insufficient to support evolutionary groupings or classifications higher than the genus level of RNA viruses, at least on the basis of simple amino acid sequence alignments.

The observations of orientation-dependent 'toxicity' induced by insertion of fragments of the AHSV VP1 gene into plasmids appears to relate to the source proximity of the insert fragment to the 5' terminus of the AHSV genome segment 1. Cloning of 5′ proximal fragments causes orientation-dependent inhibition of host bacterial cell growth and plasmid yields whereas the effect is diminished with more 3′ proximal fragments. The AHSV VP1-specific open

---

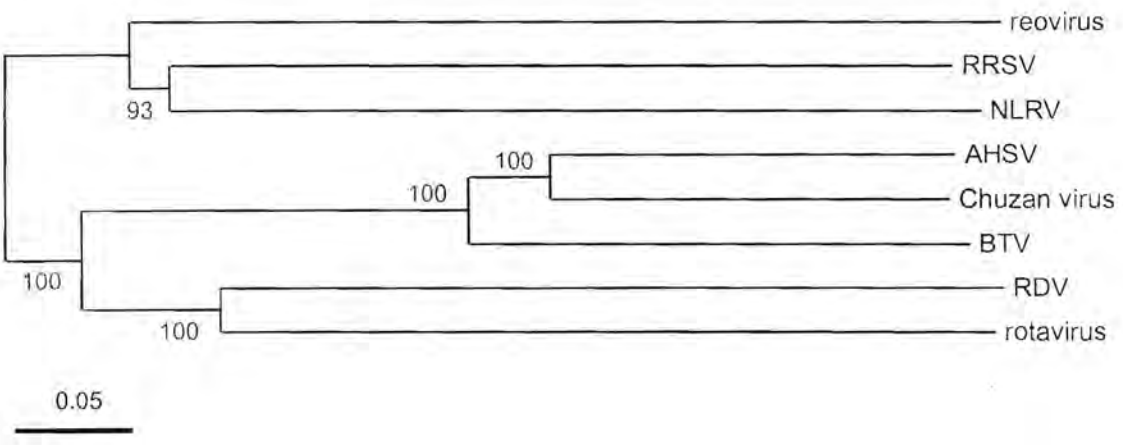
Figure 3. 8 Phylogenetic trees of some representative members of different genera in the family *Reoviridae*, based on alignments of the amino acid sequences of the respective RNA-dependent RNA polymerases. Tree A (which is based on alignments of the polymerase motif) was drawn by neighbour-joining using MEGA (Kumar *et al.* 1994); p-distances, pairwise deletion of gaps and 1000 bootstrap replications were applied. Trees B and C are based on alignments of the full-length polymerase sequences and were prepared by neighbour-joining with Clustalx (Thompson *et al.* 1997) using 1000 bootstrap replications. The dsRNA virus IBDV (family *Birnaviridae*) was included as an outgroup in tree C. Bootstrap values are shown as percentages.

A

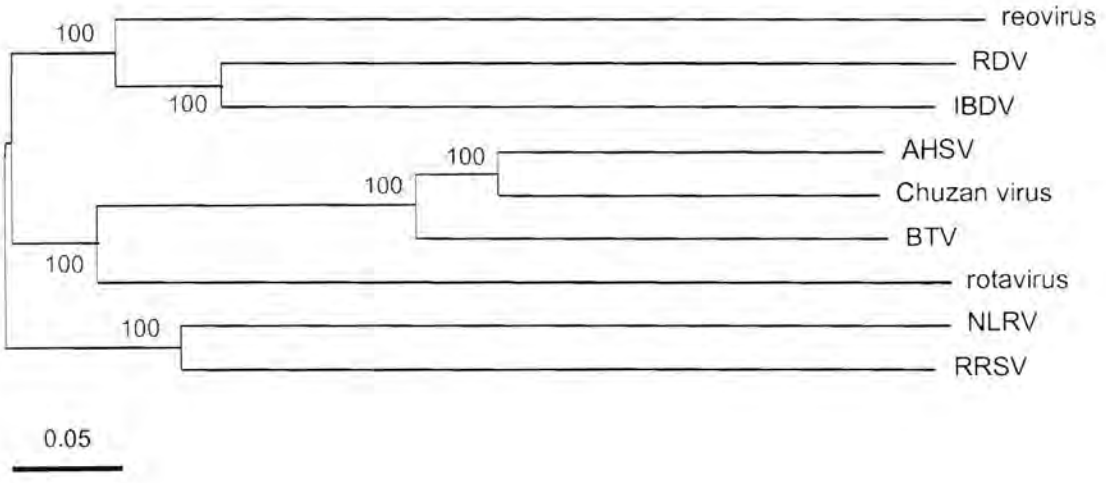


Scale: each ——— is approximately equal to the distance of 6.9%

B



C



reading frame of the inserts of all 'toxic' recombinants was found to be in the same orientation as the β-galactosidase gene (*lacZ*) on the plasmid ('*lacZ*-positive'), leading to speculation that the 'toxicity' might be related to the orientation of the insert relative to *lacZ*. However, the possibility that this may reflect in-frame translation of the gene from the *lacZ* promoter was negated by the presence of the 5' terminal G-tails in full-length clones and the identification of 'toxic' subclones (# 15 and 16) in which the open reading frame of the insert is out of frame with the *lacZ* promoter. Also of possible relevance in this regard is the recovery of only single orientations of AHSV genome segment 1-specific recombinants in non-directional subcloning experiments. This includes recovery of only one orientation of full-length AHSV-9 segment 1 cDNA cloned into pCi (Promega), a mammalian expression vector which does not contain *lacZ*, and only the '*lacZ*-negative' orientation of the 5'-proximal *Eco* RI fragment (nucleotides 877 through 1762) being obtained from subcloning into pGEM3z f(+). In addition, all clones of the AHSV-1 VP1 gene in pMOS*Blue*, described in chapter 2, and subclones of these full-length genes in pBS (Stratagene), were found to be only in the '*lacZ*-negative' orientation.

A cloning plasmid described by Thomson and Parrott (1998), pMECA, features a comparable attribute, whereby recombinants are selected on the basis of colony size. In this case, insertion of a DNA fragment into the multiple cloning site within *lacZ* on the plasmid negates the 'toxicity' of the wild type vector, restoring the normal morphology in the recombinant. In this case, it was found that reduction in bacterial colony size did not affect plasmid yields and it was speculated that a large poly-linker hairpin or a synthetic protein product of the poly-linker affected the bacterial colony.

Analysis of the VP1 amino acid sequences of AHSV and other viruses within the *Reoviridae* can at best merely provide an indication of the function of the protein and its domains. In order for an in depth analysis of viral RNA polymerase activity to be carried out, it was necessary to express AHSV VP1 in a eucaryotic system in conjunction with other putative components of the polymerase complex.