Development of experimental systems for studying the biology of *Nudaurelia capensis* β virus

Cheryl Tracy Walter

Dissertation submitted in fulfillment of the requirements of Masters of Science in Microbiology

Rhodes University, Grahamstown
South Africa
February 2005

Acknowledgements

To my supervisor over the past 3 years, Prof. Rosie Dorrington, thank you for all of you help, advice and guidance. You have taught me much more than just the art of science.

My co-workers of lab 417 who have got to be the most entertaining bunch of people I have ever known. Mez and Gwynn, thank you for all you advice and for all the conversations we had. Michele thanks for all the funny stories and your help with some of the virus work and finally a big thank you to Petra for not only helping me so much over the past three years but also for being such a wonderful and generous friend.

Val Hodgson and Prof. Don Hendry for all your valuable advice on our favourite virus and a big thank you to Val for maintaining tissue culture for me.

Mr. Robin Cross and Mrs. Shirley Pinchuck. Thank you both for spending hours searching for viruses under the TEM, taking amazing photographs and for all the valuable advice. Both of you are gifted professionals whom I greatly admire.

My parents up in Johannesburg and my friends, old and new, thank you for keeping my feet on the ground, distracting me when I needed to be distracted and for always being there for me through the good times and the bad. Your support has been unwavering and I will never forget it.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

After 20 years, *Nudaurelia β virus* (NβV) was re-isolated from a population of *Nudaurelia capensis* larvae exhibiting similar symptoms to those described in 1974 for a tetravirus infection. NβV is a member of the *Tetraviridae*, a family of positive sense insect RNA viruses that exclusively infect Lepidopteran insects. In addition to NβV, there was evidence that the insects were infected with another small RNA virus. SDS-PAGE and Western analysis revealed two proteins (p56 and p58), that cross-reacted with anti-NβV antibodies. Transmission Electron Microscopy (TEM) analysis showed the presence of particles exhibiting a morphology described for NβV and majority of particles of a diameter of 37 nm. In addition there was a second, minor population of particles with a diameter of 34 nm, which also exhibited the characteristic pitted surface of NβV, raising the possibility of two species of NβV in the *N. capensis* population.

To further investigate this, cDNA corresponding to the 3` end of the replicase gene as well as the entire capsid gene of NβV was synthesized and sequenced. Alignments of the cDNA sequence showed a 99.46 % identity to the published sequence of NβV. Two amino acid substitutions were observed in the capsid coding sequence, one of which was a conservative substitution. Both of these substitutions were found in the β-sandwich domain of the capsid protein. Inspection of the capsid coding sequence showed a second methionine (Met50) at the VCAP amino terminus raising the possibility that p56 might arise from a translation product starting at this site. To investigate this, a full length and truncated capsid coding sequence starting at Met50, were expressed in a baculovirus expression system. VLPs were examined by TEM and Western analysis showed the presence of virus like particles with NβV morphology, but smaller in diameter than the wild-type with an average of 33.33 nm, similar to the smaller particles observed in the virus preparations of NβV. This result supported the hypothesis that NβV translates a smaller coat protein from the second in-frame methionine residue.

Chapter 1

LITERATURE REVIEW

It is surprising that very little is known about viruses that infect insects, when one considers the fact that the number of insect species is probably greater than most other classes of eukaryotes, (Rossmann and Tao, 1999). Insects comprise over 80 % of existing species on the planet and viruses infect most orders of insects including both beneficial and pest species. The first recorded case of insect viral disease can be traced back as far as the 16th and 17th centuries where scientific records exist describing the “wilting” disease of silkworms, indicating that insect viruses have been of interest over at least four centuries (Fields *et al*., 1996).

The study of small RNA viruses has numerous applications in fields ranging from protein assembly to organismal biology making important contributions to the understanding of the biology of living cells. Research on small RNA viruses has resulted in the first demonstration of an RNA-dependent RNA polymerase, the identification of a helical base-paired double-stranded replicative form of viral RNA, and the discovery that poliovirus synthesizes its gene products by proteolytically cleaving a large polyprotein (van Regenmortel *et al*., 2000). A number of families of viruses are known to infect insects. These families include both DNA and RNA viruses as well as retroviruses (Table 1.1).

Only a small number of the families listed have a host range limited to only insects. These include the *Ascoviridae*, *Baculoviridae*, *Tetraviridae* *Polydnaviridae* and the *Dicistroviridae* (van Regenmortel *et al*., 2000).

Table 1.1. Summary of virus families known to infect insects (van Regenmortel *et al*., 2000).

Genome type	Family	Examples of infected species
dsDNA	*Poxviridae, Iridoviridae, Baculoviridae, Polydnaviridae (Ichnoviridae* and *Bracoviridae)* and the *Ascoviridae*	Red Hairy caterpillar, *Spodoptera* spp., and the Codling moth.
dsRNA	*Reoviridae* and *Birnaviridae*	Silk worms, *Spodoptera* spp., *Drosophila* spp.
ssDNA	*Circoviridae* and *Parvoviridae*	Silk worms, Yellow fever mosquito
-ssRNA	Rhabdoviridae and Bunyaviridae	*Drosophila* spp. and mosquitoes
+ssRNA	*Togoviridae, Flaviviridae, Picornaviridae, Nodaviridae, Tetraviridae* and *Dicistroviridae*	Pine Emperor moth, Cotton bollworm, various cricket species, Gypsy moth
RNA	*Metaviridae*	*Drosophila* spp.

ss = single stranded; **ds**= double stranded; **-**=negative sense; **+**=positive sense.

As mentioned earlier, very little is known about insect viruses, in particular small unenveloped icosahedral (a geometrical shape with 20 identical faces) RNA viruses. Much work has been done on viruses with medical or agricultural applications and a classic example of such a family is the *Picornaviridae*. The *Tetraviridae* are a family of icosahedral viruses that have a host range limited to a single order of insects, the *Lepidoptera.* They are structurally closely related to the *Nodaviridae*, a family of small RNA viruses known to infect both animals and insects, and more distantly to the *Picornaviridae* (Johnson *et al*., 1994). There are many similarities between positive sense RNA viruses which allows for a detailed understanding of the evolutionary relationships which exists among these viruses (Rossmann and Johnson, 1989). The subject of this literature review is the biology of the *Tetraviridae*, which will be discussed in detail below.

1.1. The *Tetraviridae.*

The first members of the *Tetraviridae* were observed to infect the larvae of the Pine Emperor moth, *Nudaurelia cytherea capensis*. As their name suggests, the larvae of the Pine Emperor moth feed off the needles on pine trees. The larvae *of N. capensis* were originally found to feed off indigenous vegetation but with the establishment of commercial pine plantations, were found in great numbers feeding from pine trees (Hendry *et al*., 1968). The indigenous plants are preferred but, if found intergrowing in *Pinus radiata* (exotic) plantations, ovipositing adults prefer the exotic plants. Food shortages force the larvae to attempt feeding on unsuitable plants (Geertsema and Giliomee, 1972).

Damage to windbreaks and smaller plantings of *P. radiata* as well as to some exotic and indigenous trees and shrubs used for the reclamation of coastal sand dunes is caused annually by these larvae. Varying degrees of defoliation occur and this leads to corresponding damage to trees, i.e. scorching of the bark, deformation of the crown and young trees, and probably loss of wood increment. In the case of continuous defoliation the trees may suffer so severely as to cause them to lose their vigour and die. Unhealthy and struggling trees are also more attractive to secondary insects. Young trees of 1-2 years of age are rarely defoliated although eggs may be laid on them. Trees from 3 to 15 years old are most prone to attack and suffer the most injury. In the older trees, defoliation still takes place but foliage loss is not so obvious due to the pine needles being too hard for the freshly emerged larvae to feed on (Geertsema and Giliomee, 1972).

In 1974 five small viruses were purified from the larvae *N. capensis* (Juckes, 1974). The viruses were termed *Nudaurelia* α, β, γ, δ, and ε *viruses* in order of increasing prevalence. It was later discovered during structural studies on *Nudaurelia* β *virus* (NβV) that this particular virus was unique and made the first member of a new family of viruses, the *Tetraviridae* (Olson *et al*., 1990). More than 10 years later another dominant virus was isolated from *Nudaurelia* larvae, the *Nudaurelia* ω *virus* (NωV), which had biophysical properties similar to those of other tetraviruses but differed by the presence of a second genomic RNA strand (Hendry *et al*., 1985). This virus became the type virus

for a second genus within the *Tetraviridae*, the ω-like genus and was joined by a second similar virus from Australia, the *Helicoverpa armigera stunt virus* (HaSV) which was first isolated from a laboratory-bred colony of *Helicoverpa armigera*. The viruses containing either a monopartite and bipartite genome were thus separated out into two genera, the beta-like now known as the betateraviruses and omega-like now known as the omegatetraviruses (Hanzlik and Gordon, 1997). Viruses in this family are positive sense single stranded RNA viruses with a distinct icosahedral symmetry and average diameter of 40 nm (Hendry and Agrawal, 1994).

1.2. Quasiequivalence structure and crystal structure of the capsids of tetraviruses.

The quasiequivalence theory provided by Caspar and Klug (1962) provides a basis for the classification of the protein shells of isometric viruses in terms of the triangulation number *T*, which specifies an icosahedral surface lattice accommodating 60T equal protein subunits. According to this theory, *T*, can take the values 1, 3, 4, 7, etc., of which *T*=1 denotes the shell built from 60 structurally equivalent subunits; while for *T*> 1 the subunits are structurally quasi-equivalent. The quasi-equivalence theory, proposed by Caspar and Klug (1962), describes the formation of virus capsid protomers. The theory of icosahedral virus construction was based on the recognition that structural proteins are adaptable molecules that can self-assemble by bonding together in different ways within a highly ordered structure (Caspar and Klug, 1962).

Three-dimensional image reconstructions from electron micrographs showed that NβV is built from 240 structure units clustered in Y-shaped trimers and arranged with the symmetry of the *T*=4 (Olson *et al*., 1990). This unique symmetry was used to name the family, the *Tetraviridae*. The micrographs showed that the icosahedral surface of the virus was constructed from 240 protein structural subunits clustered in trimers with four Y-shaped trimers per icosahedral face (20 triangular faces per virus particle) (Olson *et al*., 1990).

NβV was shown to be isometric with a diameter of 35-41 nm and the molecular weight calculated by Struthers and Hendry (1974) corresponded with each particle comprising of 240 subunits (Finch *et al*., 1974).

Olson and co-workers (1990) constructed three-dimensional structures of frozen-hydrated NβV and NωV using electron microscopy and image analysis. The two virus particles at 3.2 nm resolution showed that each face has distinct structures at points of symmetry protruding approximately 4 nm above the contiguous surface. However, the NβV face is distinct from NωV by being more planar with deeper grooves between the faces and having three prominent pits (Olson *et al*., 1990) (Figure 1.1).

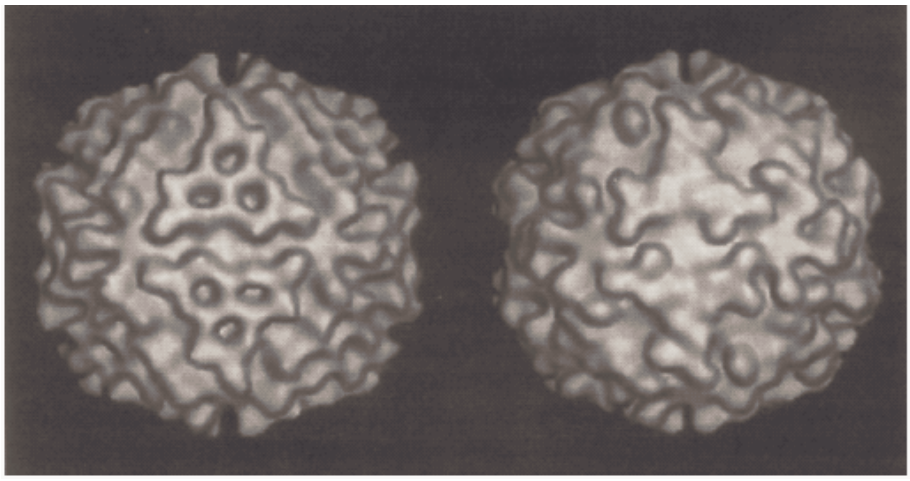


Figure 1.1. Image reconstruction of frozen-hydrated *Nudaurelia* viruses analysis by cryo-electron microscopy. The virus particle on the left is NβV and on the right NωV. (Olson and Baker, 1990)

The capsid consists of multiple domains created by protein subunits. Crystallographic studies of RNA viruses that infect eukaryotic cells have shown that all capsid proteins contain an eight-stranded antiparallel β-barrel motif that forms the contiguous shell of the particle. Although the image reconstruction gives no detail of the subunit organization in the contiguous shell portion of the NβV capsid, it was been proposed that it may be formed by β-barrels arranged according to architectural principals similar to those found in the shells of *T*=3 viruses (Olson *et al*., 1990).

1.3.1. Distribution and Host Range.

Table 1.2. Species in the Betatetravirus and Omegatetravirus genera (Adapted from van Regenmortel *et al*., 2000).

Virus name	Abbreviation	Host genera	Diameter (nm)	Reference	Genome sequence accession number
Genus: Betatetravirus					
Antheraea eucalypti virus	AeV	*Saturniidae*	n.d	Grace and Mercer, 1965	n.d
Nudaurelia capensis β virus	NβV	*Saturniidae*	39.7	Gordon *et al*. 1999	AF102884
Darna trima virus	DtV	*Limacodidae*	n.d	Reinganum *et al*. 1978	n.d
*Dasychira pudibunda virus**	DpV	*Lymantriidae*	n.d	Greenwood and Moore, 1984	n.d
Philosamia cynthia x ricini virus	PxV	*Saturniidae*	n.d	Reinganum *et al*. 1978	n.d
Pseudoplusia includens virus	PiV	*Noctuidae*	n.d	Chao *et al*. 1983	n.d
*Thosea asigna virus***	TaV/SaV	*Limacodidae*	35	Pringle *et al*. 1999	AF062037
Trichoplusia ni virus	TnV	*Noctuidae*	35-38	Morris et al. 1979	n.d
Euprostena elaesa virus	EeV	*Limacodidae*	n.d	Zeddam *et al*. up.	AF461742
Providence virus	PrV	*Noctuidae*	40	Pringle *et al*. 2002	AF548354
Genus: Omegatetravirus					
Nudaurelia capensis ω virus	NωV	*Saturniidae*	41	Agrawal and Johnson, 1992	RNA 2, S43937
Helicoverpa armigera stunt virus	HaSV	*Noctuidae*	38	Hanzlik *et al*., 1995	RNA1, U18246 RNA 2, L37299

*Previously named *Calliteara pudibunda virus*. n.d: not determined
**Previously named *Sethosea asigna virus*. up: unpublished.

The known host range of tetraviruses is confined to a single insect order, the *Lepidoptera* (moths and butterflies). Transmission of these viruses appears to be both horizontal via ingestion and vertical from parent to progeny, although evidence strongly suggests the former (Juckes, 1974). How the latter could occur is not well understood however emerging larvae may become infected as they bite their way out of eggs superficially contaminated with virus (transovum transmission) (Juckes, 1974). Virus could also be present in insect frass, but whether this could play a role in transmission is unclear.

Another possibility is transovarial transmission from infected adults (Hendry and Agrawal, 1994).

1.3.2. Pathology of tetraviruses.

Tetravirus disease ranges from unapparent infections to acutely lethal infections (Hendry and Agrawal, 1994). The disease response for any particular insect is dependent on host/virus combinations and is affected by host stress, host life stage, virus dose and the presence of other viruses (Hanzlik and Gordon, 1997). In lethal infections, larvae become moribund, discoloured and flaccid at about 7-9 days post infection, and hang by their prolegs (Hendry and Agrawal, 1994). An associated internal liquefaction of the organs leaves the integument intact, differing from the pathology caused by baculoviruses in the final stages of virogenesis (Hanzlik and Gordon, 1997). On occasion, larvae infected with NβV pupate but either do not emerge or emerge with stunted wings (Hendry and Agrawal, 1994). Replication occurs only in the cytoplasm of midgut cells of the larvae (Agrawal and Hendry, 1994) and only at the very late stages of infection however viral antigens have been detected in fat body and gonadal tissue using enzyme-linked immunosorbent assay (ELISA). This is thought to be a result of viruses leaking into these tissues rather than by infection (Greenwood and Moore, 1984).

It has been found that *Helicoverpa armigera* larvae exhibit developmental resistance towards a tetravirus infection with increasing instars explaining why earlier larval instars react more severely to infection compared to larvae in later instars. This is thought to occur by developmental resistance as well as the midgut cells becoming more refractory to infection and, combined with cell sloughing, allows the midguts of older larvae to recover more readily from *Helicoverpa armigera stunt virus* (HaSV) infection (Brooks *et al*., 2002).

1.3.3. Agricultural and economical importance.

As for most other RNA viruses not pathogenic to humans, little attention has been paid to tetraviruses as biological control agents despite their known effects on some pest hosts. The insect hosts of the tetraviruses are among the most damaging pests worldwide for which new control strategies are urgently required (Christian *et al*., 2000). Agricultural interest in the possible utility of the tetraviruses stems from the ability of these viruses to infect the midgut cells, rapidly retarding the growth of the host and eventually liquefying the body contents (Hendry and Agrawal, 1994). Some laboratory studies suggest that this group of viruses may have good potential as viral insecticides (Christian *et al*., 2000).

Small RNA viruses pathogenic to insects are not generally thought of as viable biological control agents for agricultural pests as their use in the field is hindered by their potential instability and concern over their relatedness to vertebrate viruses. There are, however, at least three instances in the literature where small RNA viruses have been effective against pest insects. The *Darna trima virus* (DtV), also a tetravirus, was used successfully against a moth pest of oil palms in Malaysia (Moore, 1991*a*), and a picorna-like virus, the *Gonometa virus*, controlled a moth pest of pines in Uganda (Moore, 1991*b*). The applications of tetraviruses as a biological insecticide are immense since the hosts are widely spread throughout the world. HaSV, which infects the cotton bollworm has come under scrutiny as a possible heliothine biological control agent because of the agricultural importance of heliothine pests (Christian *et al*., 2000). EeV infects *Euprosterna elaeasa* (Oil-palm leaf-eater) larvae, which are pests found on oil-palms in certain Latin American countries (Gorbalenya *et al*., 2002)

1.4. Molecular Biology of the *Tetraviridae*.

The *Tetraviridae* are grouped into genera according to genome organization. In this case, all betatetraviruses are monopartite, with the replicase and capsid coding sequences present on a single species of RNA (Hanzlik and Gordon, 1997). The capsid coding sequence is subgenomically produced. Examples of betatetraviruses include TaV, EeV, NβV and PrV (Table 1.2).

1.4.1. The Omegatetraviruses.

Omegatetraviruses are also grouped according to genome organization and characteristically contain two species of RNA. RNA1 encodes the RNA dependent RNA polymerase (RdRp) and RNA2, the capsid gene (Figure 1.2) (Hanzlik and Gordon, 1997). The only known species to exist within this genus are HaSV and NωV (van Regenmortel *et al*., 2000).

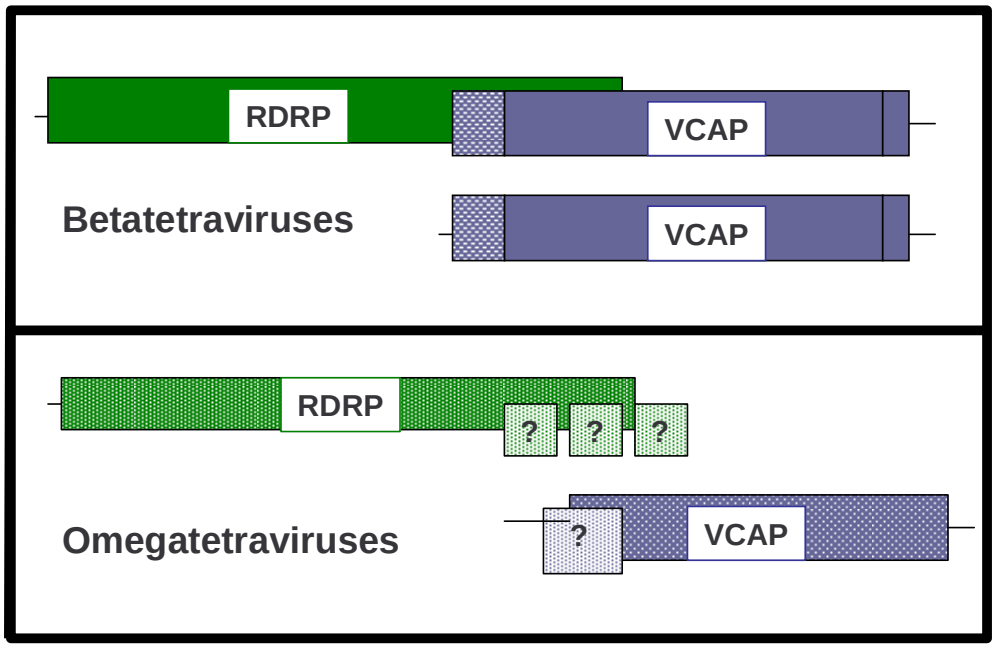


Figure 1.2. Schematic representation of the genome organization of omegatetraviruses and betatetraviruses.

1.4.2. Genome organization.

The entire sequence of both species of RNA of HaSV, the best studied member of the Omegatetravirus genera, has been elucidated (Gordon *et al*., 1995). The genes and the sizes of the proteins translated from these genes have been calculated making this virus a model for other omegatetravirus genome organization. RNA1 of this virus encodes a 187 kDa protein identified as the RdRp. The 3` end of RNA1 suggests the presence of three minor ORFs encoding peptides of 11, 15 and 8 kDa respectively (Gordon *et al*., 1995). The second RNA species of this virus encodes the caspsid protein precursor which is translated into a 71 kDa (alpha) protein. This protein is cleaved to produce a 64 kDa (beta) major coat protein and 7 (gamma) kDa minor coat protein. Translation from the 5` end also results in the production of a 17 kDa protein. The function of this protein is appears to be non-structural (Hanzlik *et al*., 1995). The genome organization of HaSV and the protein sizes of translated ORFs are summarized below (Figure 1.3).

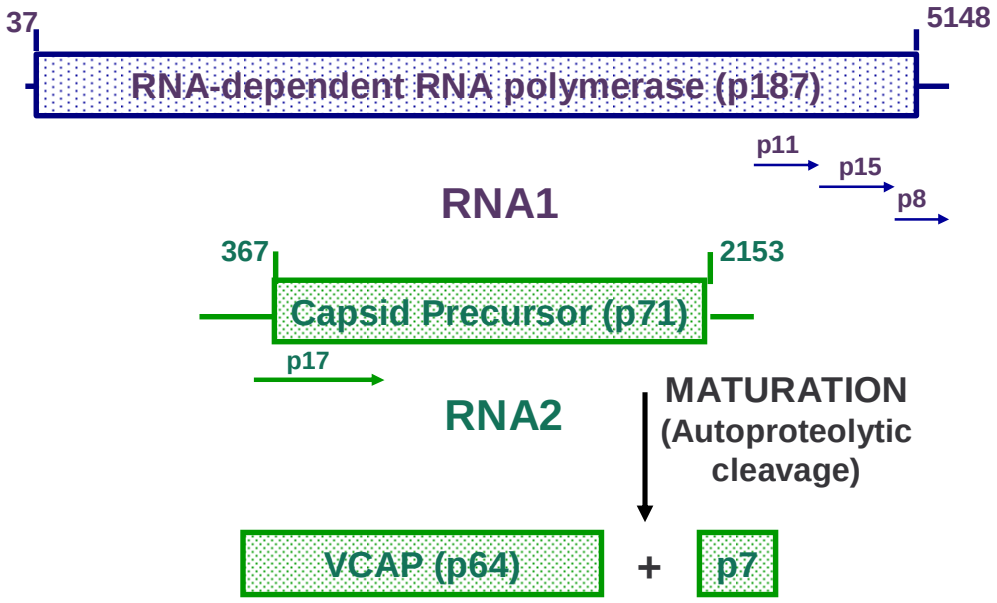


Figure 1.3. Genome organization of HaSV RNA 1 and RNA 2.

1.4.3. Capsid structure

Members of the *Tetraviridae* produce a capsid protein precursor that is cleaved to produce the actual capsid protein. NωV produces a capsid precursor protein of 70 kDa (alpha peptide) that is cleaved to produce a major coat protein of 62 kDa (beta peptide) and a minor protein of 8 kDa (gamma peptide) (Hendry and Agrawal, 1994). This leads on to the question as to how the capsid precursor protein is cleaved and also, the structure of the capsid monomer. NωV is the only tetravirus to date whose crystal structure has been elucidated (Munshi *et al*., 1996). The capsid particles of NωV assemble from 240 copies of the 70 kDa coat protein precursor which adopt four slightly different conformations in the capsid (designated A, B, C and D) (Munshi *et al*., 1996) according to the quasi-equivalence theory devised by Casper and Klug (1962). Only after assembly, are all of the subunits autocatalytically cleaved resulting in the major and minor capsid proteins (Agrawal and Johnson, 1992).

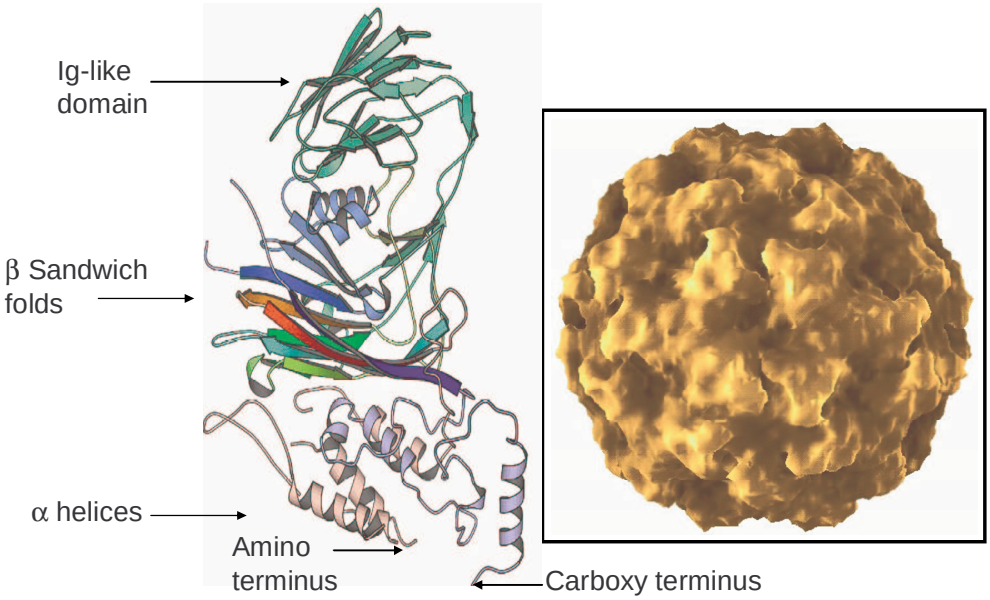


Figure 1.4. Ribbon structure of a capsid monomer of NωV and cryo-EM image reconstruction of NωV particle. Dark colours indicate the viral jelly roll (β sandwich) structure and the rest of the molecule in pale colours. Domains are indicated by the arrows (Adapted from Helgstrand *et al*, 2004). Image of the NωV capsid particle (Reddy *et al*., 2001)

The elucidation of the crystal structure of NωV yielded much information on the structure and the processing of each subunit. Each subunit is composed of a helical inner domain (where the cleavage occurs) containing residues preceding and following a canonical, viral, eight stranded β-sandwich that forms the contiguous shell (Munshi *et al*., 1996). Inserted between two strands of the shell domain is an immunoglobulin c-type fold (Figure 1.4) (Munshi *et al*., 1996). The cleavage event of the capsid precursor protein occurs between the amino acid residues Asn570 and Phe571 (Agrawal and Johnson, 1992).

This cleavage event occurs with a decrease in pH from 7.0 to 5.0 (Canady *et al*., 2000). After maturation and at pH 5, the NωV VLPs were termed a capsid and the particles resembled wild type particles with a diameter of 41 nm and distinctive trimeric bundles visible on the surface of the capsid (Canady *et al*., 2000). An Asn570 cleavage defective mutant formed a capsid that could be reverted to a procapsid when the surrounding pH was increased (Taylor *et al*., 2002). It has since been proven that a control mechanism exists for the pH-transition of NωV virus particles and that the conformational changes that occur upon lowering the pH are responsible for the activation of these changes (Helgstrand *et al*., 2004). The proposed control mechanisms involved in this conformational change are a result of stabilised protein subunit interactions and the protonation of exposed amino acid residues allowing for closer proximity between these exposed residues and tightly packaged capsid particles (Helgstrand *et al*., 2004).

1.5. Betatetraviruses.

The Betatetraviruses were the first genus of the *Tetraviridae* to be discovered (Matthews, 1982)). At present the genus is comprised of 10 members (van Regenmortel *et al*., 2000). PrV, TaV, EeV and NβV represent members that have been studied in detail. Although little is known about EeV, it has been classified as a betatetravirus, with a monocistronic genome including an overlap between replicase and capsid ORFs (Gorbalenya *et al*., 2002). The

size of the genome is approximately 5.7 kb and contains two open reading frames. The capsid precursor protein is translated from a subgenomic copy of the gene and is cleaved to form a major and minor capsid (Gorbalenya *et al*., 2002)

The physical characteristics of TaV particles support the inclusion of this virus in the Betatetravirus genus of the *Tetraviridae,* as do the genome size and organization (Reinganum *et al*., 1978, Pringle *et al*., 1999). The presence of three pits on each face of the TaV capsid and the organization of the genome, with the capsid gene present twice, once on the genomic fragment and again on the subgenomic RNA molecule, are characteristics that have been described for NβV (Pringle *et al*., 1999). This makes TaV a more suitable model compared to NβV even though it has been shown that TaV produces the p17 product which has not been discovered in NβV (Pringle *et al*., 2001).

In 2002, Pringle and co-workers discovered another tetravirus –PrV. PrV was found to persistently infect a midgut cell line (MG8) derived from the corn earworm (*Helicoverpa zea*) (Pringle *et al*., 2003). PrV thus proved to be the first tetravirus shown to consistently replicate in tissue culture. It was also demonstrated that the purified virus could also infect two other insect cell lines, one from *H.zea* fat body and one from the larvae of the beet armyworm, *Spodoptera exigua* (Pringle *et al*., 2003). PrV virions are isometric particles made up of two structural proteins (60 and 7.4 Ka) which encapsidate both genomic (6.4 kb) and subgenomic RNA (2.5 kb). PrV has a monopartite genome but has demonstrated a homology to tetraviruses in both genera (Pringle *et al*., 2003). The work achieved on PrV and TaV has updated the knowledge of previously unclassified tetraviruses and shed light on the phylogenetic relationships which exist between the two genera and even species of *Tetraviridae*.

Due to similarities of confusing evolutionary paths of these viruses, the *Tetraviridiae* have been compared on the basis of a number of characteristics such as replicase protein domains (Gorbalenya *et al*., 2002), capsid protein precursor similarities (Pringle *et al*., 2003) as well as general genomic

comparisons (Pringle *et al*., 2001). Analysis compiled by comparing the replicase genes of EeV to TaV, NβV and HaSV has found that TaV and EeV do not have characteristic methyltransferase and helicase domains but rather a putative RdRp with a unique C-A-B motif arrangement in the palm subdomain that is also found in dsRNA birnaviruses (Gorbalenya *et al*., 2002).

The study of the replicase domains alone shows distinct differences between the types and numbers of recognizable domains with the replicase genes of TaV and EeV versus NβV (Gorbalenya *et al*., 2002).

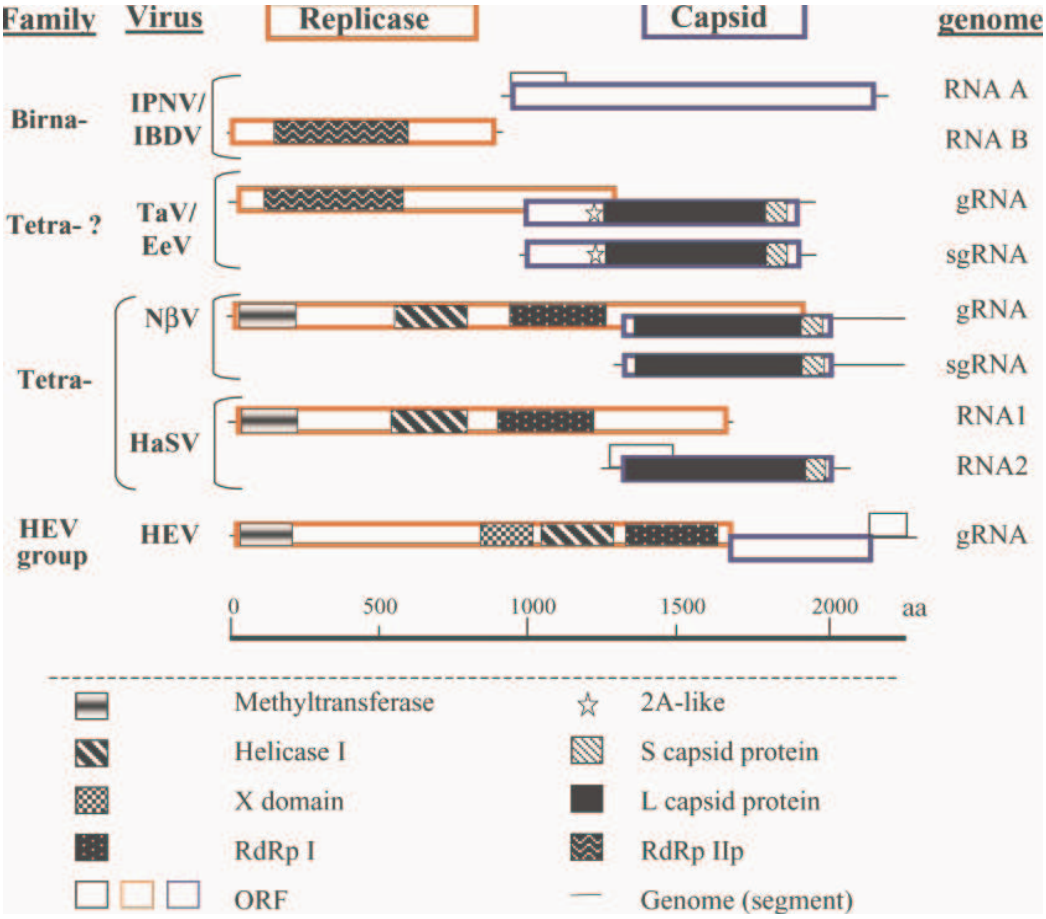


Figure 1.5. Identification and comparison of domains of both the replicase and capsid proteins of NβV, HaSV, TaV, EeV, Hepatitis E Virus (HEV), Infectious Bursal Disease Virus (IBDV) and Infectious Pancreatic Necrosis Virus (IPNV). Note the genome organization of each virus on the right of the figure. (Gorbalenya *et al*., 2002)

Further analysis of the capsid precursor protein of five viruses of the *Tetraviridae* revealed, in particular instances, relatively low percentages of identity between viruses of the same genus (Table 1.3.).

Table 1.3. Comparison of the capsid protein precursor of five tetraviruses (Pringle *et al*., 2003)

Identity \ Similarity	Virus	PrV	TaV	HaSV	NβV	NωV
	PrV		35	47	35	46
	TaV	28		33	43	35
	HaSV	38	27		36	75
	NβV	27	35	29		35
	NωV	37	27	66	29	

When comparing the percentage identity between the capsid precursor protein of PrV and TaV, it was observed that more homology exists between this virus and members of the omegatetraviruses. The same protein of NβV was also found to be marginally more homologous to PrV than to HaSV, confirming that genome organization alone is not a sufficient characteristic to group these viruses into the two genera. This work has provided evidence that, when including the capsid precursor gene of *Providence virus* (PrV), the exact phylogenetics of the *Tetravirdae* become confusing. PrV also has a monopartite genome with its capsid ORF expressed subgenomically. Even though, on a genomic and structural level, the virus is very similar to the betateraviruses, its capsid gene shares more homology with the omegatetraviruses (Pringle *et al*., 2003). This new discovery has set the wheels in motion for what seems to be another rearrangement of the *Tetraviridae* family and perhaps the inclusion of a third genus.

1.6. Capsid processing strategies of PrV, TaV and NβV

Much has yet to be elucidated about the capsid processing strategies of omegatetraviruses and betatetraviruses. Since the single RNA species of viruses with a monopartite genome is infectious, the genome must contain the information for the synthesis of non-structural as well as capsid proteins. Consequently, such viruses have resorted to several strategies to produce their different proteins. These include:

1. Premature termination at the level of a stop codon different from the termination codon leads to the formation of two proteins that differ in size but whose amino acid sequences are identical over the total length of the shorter protein
2. Read-through or suppression of a termination codon also leads to the presence of two proteins. This strategy is thought to be mediated by suppressor tRNA
3. Post-translational cleavage of a polyprotein that yields mature functional proteins.
4. Multiple reading frames. This often involves subgenomic RNA.
5. Subgenomic RNA thought to be derived from RNA fragmentation.
6. Internal initiation (Rowlands *et al*., 1987).

In 2001, Pringle *et al*. elucidated the capsid expression strategy of TaV. TaV was first isolated in the larvae of *Setothosea asigna* by Reinganum *et al*. (1978). Upon re-classification of TaV, it was discovered that TaV particles consist of two capsid proteins of 56 and 6 kDa, and the capsid gene expressed subgenomically, typical for the betatetraviruses (Pringle *et al*., 2001). Interestingly, it is proposed that the capsid protein precursor undergoes two cleavage events to produce the major and minor capsid proteins as well as a 17 kDa N-terminal cleavage product. Comparisons of TaV with other beta- and omegatetraviruses revealed that TaV was most closely related to NβV (Pringle *et al*., 2001).

Although TaV has been proven to have various structural and physical differences compared to other members of the Betatetravirus genus it does have important similarities to NβV such as capsid structure and monopartite genome. Interestingly, Pringle *et al.* (1999) proposed that the capsid precursor underwent two cleavage events before becoming a stable virus capsid (Figure 1.6).

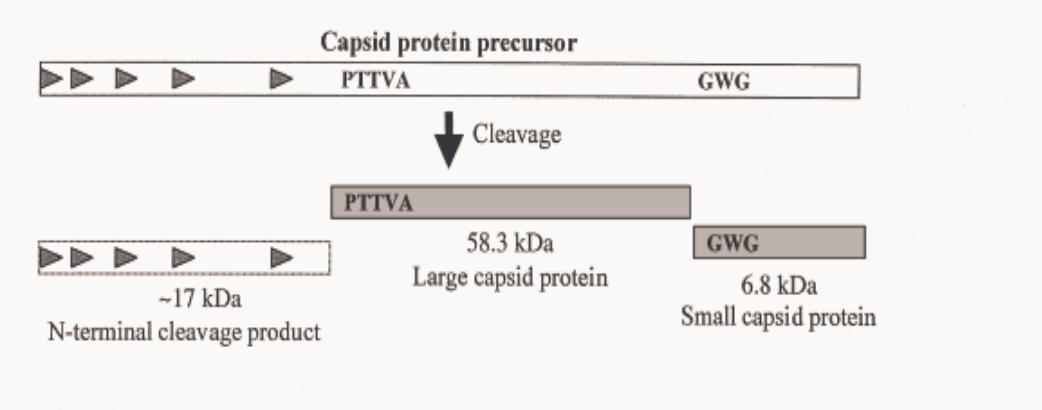


Figure 1.6. The capsid processing strategy of TaV (not to scale). The capsid precursor protein of TaV is produced from one of the five ATG codons (shaded arrows) at the start of the capsid contig. The proteins that are produced by cleavage of the precursor are shown. The 58±3 and 6±8 kDa proteins (shaded) are the large and small capsid proteins. A putative protein of up to 17 kDa is also produced by cleavage of the precursor. This protein has no known function and has not been detected in TaV virions (Pringle *et al*., 1999).

Pringle *et al*. (2001) investigated the processing strategy of TaV and in particular, the role of the p17 protein. It was discovered that the protein is not essential for virus assembly and therefore has a non-structural function as opposed to the p17 protein produced by HaSV which was speculated to produce tubules when expressed in *E. coli* cells (Hanzlik *et al*., 1995). Since NβV, which is closely related to TaV, does not produce this protein it can be assumed that this protein does not play a structural role for this virus either (Pringle *et al*., 2001). This was substantiated by the expression of the TaV capsid coding sequence in the absence of p17 resulting in the production of VLPs regardless of the presence of this protein (Pringle *et al*., 2001). It was also discovered that a cleavage event between the major and minor capsid proteins is required for assembly of VLPs and that these two proteins both need to be present for assembly of particles. While an equivalent protein

(p17) has yet to be identified in the N-terminal region of the capsid coding sequence of NβV (Gordon *et al*., 1999), a similar protein might exist explaining why two methionine residues are found on the N-terminal of the capsid gene ORF of both NβV and TaV. The data produced by Pringle and co-workers has been important in piecing together the number of functional genes, evolutionary relationships and possible methods of capsid processing in the *Tetraviridae,* in particular, Betatetraviruses.

1.7. *Nudaurelia β virus*

Nudaurelia β virus (NβV) is the type member of the Betatetraviruses, a genus of insect viruses with small (35-41 nm) isometric *T*=4 particles and monopartite ss+RNA genomes (Hendry and Agrawal, 1994; Hanzlik and Gordon, 1997; Reinganum, 1991). Like all the other viruses in this family, NβV has a host range restricted to the *Lepidopteron* insect order, *Nudaurelia capensis* (Hendry *et al*., 1968).

The +ss RNA genome consists of 6625 nucleotides containing two open reading frames (ORFs) and contributes 11 % of the weight to each virion (van Regenmortel *et al*., 2000). The 5' proximal ORF of 5778 nucleotides encodes a protein of 215 kDa containing three functional domains characteristic of a viral replicase: the methyltransferase, helicase and RNA-dependent RNA polymerase sequence motifs (Gordon *et al*., 1999). The replicase gene covers 87% of the NβV genomic RNA, terminating at a stop codon over nucleotides 5868–5870 (Gordon *et al*., 1999). A surprising feature about the genome of NβV is the degree of overlap, with no intervening non-coding sequence, between the replicase gene and that for the capsid precursor gene, which is located from nucleotides 4039-5874. The reading frames for these two genes overlap by 1827 nucleotides which represents virtually the entire capsid gene (Gordon *et al*., 1999).

The 3' proximal ORF, of 1836 nucleotides, encodes the predicted 70 kDa capsid precursor protein and overlaps the replicase gene by more than 99% and is in the +1 reading frame relative to the replicase reading frame (Hanzlik and Gordon, 1997). The capsid precursor (66 kDa) gene is located from nucleotide 4039-5874 and is cleaved to produce a 61 kDa capsid coat protein (Gordon *et al*., 1999). The size of this major coat protein has since been reported as 58.4 kDa with a suspected minor protein of 8 kDa (van Regenmortel *et al*., 2000) although no smaller minor capsid protein of the predicted molecular weight has been detected. The method by which the capsid precursor is cleaved to produce two smaller proteins is uncertain (D. Hendry pers. comm., 2004). The amino acid sequence determined for the N-terminus of the minor capsid protein by sequence analysis of the complete virus particle produced results which suggested that the cleavage event to produce these two proteins occurs between amino acids 536 and 537 which would produce two proteins of 58448 and 7975 Da (Gordon *et al*., 1999). No sequence corresponding to the N-terminus of the major capsid protein was obtained which suggested that this terminus was blocked. Similar results were obtained during the N-terminal sequence of the major capsid coat protein of NωV (Agrawal and Johnson, 1992) and HaSV (Hanzlik *et al*., 1993)

The method of gene expression differs between the mono- and bipartite genomes. The 3’ proximal ORF encodes capsid precursor proteins that are translated from a sub-genomic fragment of RNA (Figure 1.6) (Gordon *et al.*, 1999). The 6625 nucleotide NβV genomic RNA encodes both the replicase open reading frame (ORF) at the 5’end and the capsid protein at the 3’ end. The initiation codon of the 5’ ORF is located at nucleotides 93-95, consistent with translation of the replicase directly from the genomic RNA as suggested by du Plessis *et al*., 1991. The average molecular weight of the single RNA strand is about 2.0×10^6 (Gordon *et al*., 1999).

The sequence of NβV suggests a model for replication of the genome in infected host cells involving the generation of the genomic RNA species and of the positive subgenomic capsid mRNA from the same replication complex carrying a full-length negative sense RNA (Gordon *et al*., 1999). The size of

the coding sequence of the subgenomic RNA is 1.9 kb (Gordon *et al*., 1999). There are no other significant ORFs on the NβV sequence, including the 636 nucleotide stretch between the end of the capsid gene and the beginning of the tRNA-like structure at nucleotide 6514 (Gordon *et al.*, 1999).

The 3' terminal structure of 120 nucleotides has a valine anticodon, similar to the tRNA-like secondary structures which have also been identified on the HaSV RNAs. Unlike known plant viruses, this tRNA-like structure lacks pseudoknot in the stem but instead, a cloverleaf folding of canonical tRNA. This feature is present in both the Omegatetraviruses and Betatetraviruses (Gordon *et al*., 1999). These structures are most likely involved in RNA replication (Gordon *et al*., 1999). Therefore the lack of a poly (A) tail on the 3' terminus suggest this might be a prominent structural feature, with one purpose being to protect the RNA from exonucleolytic degradation (Hanzlik and Gordon, 1997).

1.8. Problem statement

NβV has not been isolated since 1985 which has created a lack of an experimental system for the type member of the betatetraviruses. The re-isolation of this virus is imperative so as to accumulate viable virus for further studies both fundamental and comparative. Also, at present an efficient heterologous expression system for the betatetraviruses does not exist. The expression of the capsid gene of TaV using the baculovirus expression system yielded intact VLPs (Pringle *et al*., 2001) albeit the concentrations of which are unknown. Heterologously expressed VLPs need to be compared to viable wild type particles of the same species. Much of the biology of this virus has yet to be elucidated and it is proposed that this virus has a unique capsid processing strategy involving a second in-frame start codon within the capsid gene ORF that can be monitored by the heterologous expression of the capsid gene and truncated form (by 50 amino acids) of NβV using a baculovirus expression vector. The expression of the TaV capsid gene with and without the N-terminal (deletion of the p17 protein) yielded mature VLPs albeit in very small quantities.

1.9. Objectives of the project.

- Isolation of naturally occurring tetraviruses from *N. capensis* larvae exhibiting symptoms of infection.
- Molecular characterization of newly isolated tetraviruses.
- Expression of NβV capsid gene and truncated form of capsid coding sequence.

Chapter 2

METHODS AND MATERIALS

2.1. Virus purification from infected *N. capensis* larvae.

The virus purification protocol used in this study was modified from Morris *et al*. (1979) and is summarized in Figure 2.1. Approximately 250 g of frozen insects were blended in a Warring homogenizer with Tris-EDTA [ethylenediaminetetra-acetic acid di-sodium salt] buffer (0.2 % β-mercaptoethanol, 0.5 M Tris-Cl pH 7.5, 0.01 M Na_2EDTA, and 250 ml of butan-1-ol). This was homogenized at 4°C and the supernatant was clarified by centrifugation at 8 000 rpm for 15 minutes at 4°C using a JA 10 rotor in a J2-21 series Beckman centrifuge. The pellet was re-extracted for a second time, the two supernatants pooled and the volume measured. To the clarified supernatant, 8 % PEG (polyethyleneglycol) 6000 and 0.1 M NaCl was added and stirred at room temperature for 1 hour. The amount of PEG and NaCl added was calculated based on the volume of the supernatant. The proteins in the supernatant were allowed to precipitate at 4°C for 3 hours and the supernatant containing the precipitated proteins centrifuged at 8 000 rpm for 15 minutes at 4°C. The supernatant was discarded and the pellet resuspended in one-tenth the volume measured of homogenized supernatant of 0.1 M Tris-EDTA (pH 7.5) followed by centrifugation at 8 000 rpm for 10 minutes at 4°C to pellet any insoluble proteins. The pellet was discarded and the virus particles pelleted through a 1.5 ml 30 % sucrose cushion (w/v in EB [50 mM Tris pH7.5, 250 mM NaCl] buffer) in a Beckmann SW 41T at 40 000 rpm at 11°C for 1.25 hours. The pellet was resuspended in 500 µl of EB (Elution buffer) and left at 4°C overnight. A Beckman L-70 ultracentrifuge was used for all high-speed centrifugations.

The resuspended pellet was then centrifuged through either a caesium chloride or sucrose gradient. Caesium chloride gradients were prepared by adding approximately 5 ml of 30 % caesium chloride (w/v in EB buffer) to a Beckman SW 41 tube. An equal volume of 60 % caesium chloride (w/v in EB buffer) was used to underlay the 30 % caesium chloride. The tube was filled to the top with the resuspended sucrose pellet. The gradients were spun at 30 000 rpm for 18 hours at 18°C in a Beckman L-70 ultracentrifuge.

Sucrose gradients were prepared in a Biocomp Gradient maker 107 ip, by adding 6 ml of 10 % sucrose in extraction buffer (EB) (pH 5 and 7) to a Beckman SW 41 tube. The 10 % sucrose was underlayed with 6 ml of 40 % sucrose in EB after which the gradient was generated according to the manufacturer's instructions. A volume equivalent to the resuspended sucrose virus was removed from the top of the gradient and the virus added slowly to the pre-formed gradient. The tubes were then centrifuged at 40 000 rpm for 1.25 hours at 11°C.

Both sucrose and caesium chloride gradients were fractionated using an ISCO Density Gradient fractionator Model 640. Alternatively if visible bands were sufficiently separated, the bands were extracted by piercing the side of the tube with a 2 cm^3 syringe and 22Ga needle. The fractions were analyzed by measuring the optical densities at 280nm and 260nm and the refractive indices measured using an ABBE-type refractormeter (Bellingham and Stanley, Ltd., England). Selected fractions were either stored at -20°C or used directly for RNA extraction.

Purification of viruses infecting Pine Emperor Moth larvae

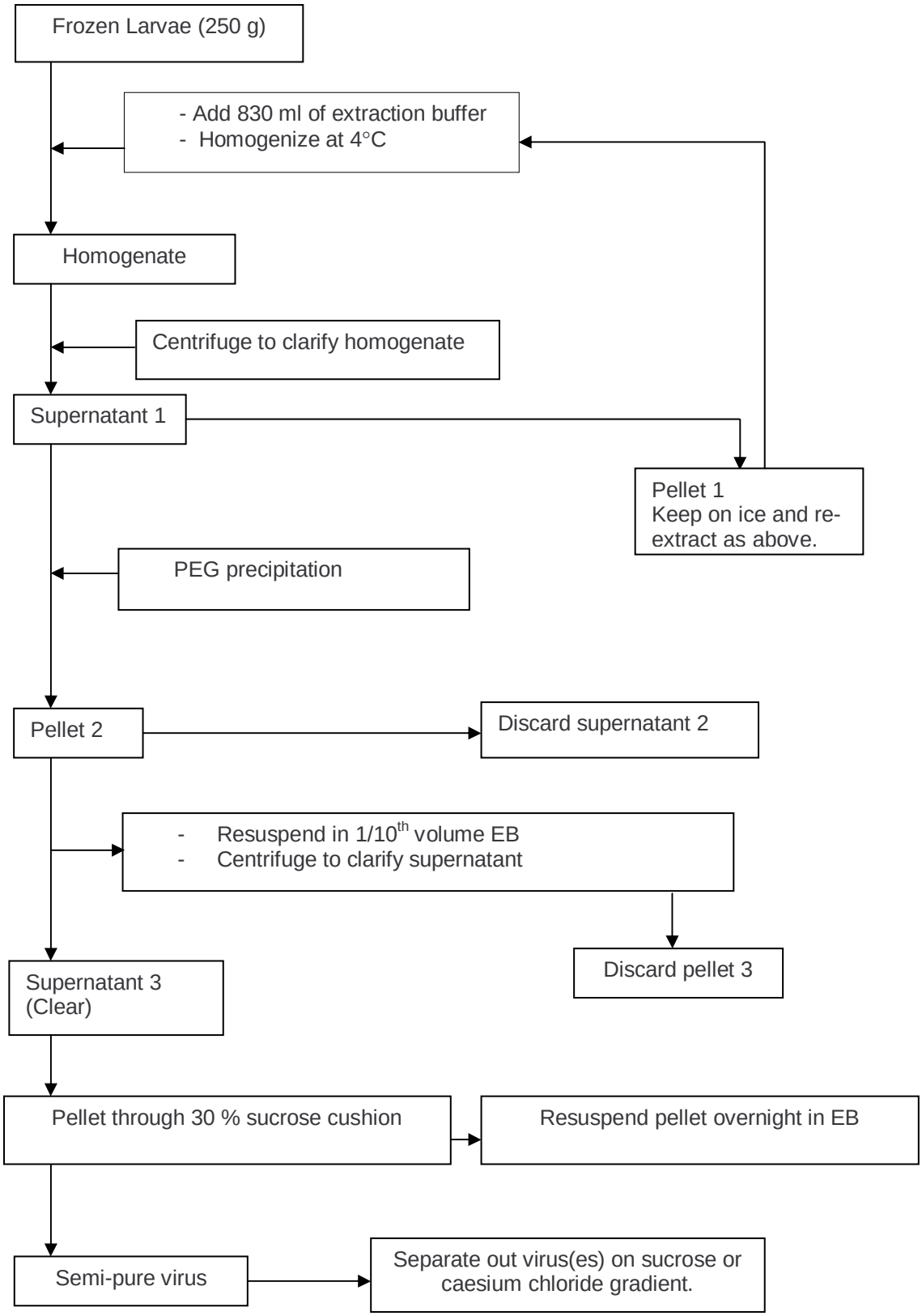


Figure 2.1. Flow diagram for the purification of tetraviruses from diseased *N. capensis* larvae

2.2. Analysis of viral RNA

2.2.1. RNA extraction from virus particles.

The protocol below was modified from Agrawal and Johnson (1992) and Hendry *et al*. (1985). All reagents were prepared in DEPC [Diethyl Pyrocarbonate]-treated triple distilled water and all containers, autopipettes, and disposables treated with RNAse Away (Molecular Bioproducts). 300 µl of a CsCl fraction containing virus was placed in a sterile Eppendorf tube and 15 µl of 10 % SDS (Sodium dodecylsulphate) added. This was thoroughly mixed, heated to 80°C for exactly 50 seconds, cooled rapidly on ice for 5 minutes and 15 µl of 1 X TAE buffer (Appendix C) added. 300 µl of phenol (equilibrated to pH 8 with TE buffer – Appendix C) was added, the tube vortexed for 1 minute followed by the addition of 300 µl of chloroform and again vortexed for 1 minute. This was microfuged at 13 000 rpm for 10 minutes in a Heraeus Biofuge microfuge. The aqueous phase was placed in a fresh Eppendorf tube and the phenol/chloroform extraction repeated twice. Following this, 300 µl of chloroform was added to the aqueous layer, vortexed for 1 minute and centrifuged at 13 000 rpm for 10 minutes. The aqueous layer was placed a fresh Eppendorf tube and the chloroform extraction process repeated twice as above. To precipitate out nucleic acids, 15 µl of 2 M sodium acetate (pH 5.2), 30 µl of 2 M acetic acid and 600 µl of ice cold 95 % ethanol was added. The solution was left to precipitate at -20°C overnight and microfuged at 13 000 rpm for 30 minutes at 4°C in a refrigerated microfuge (Eppendorf). The supernatant was removed and 600 µl of ice cold 70 % ethanol was added and the tube microfuged for a further 5 minutes at 4°C. The pellet was dried and resuspended in 40 µl of DEPC-treated water. This was left to resuspend for 1 hour at 4°C, aliquoted into volumes of 5 µl and stored at -80°C.

2.2.2. RNA denaturing gels

In order to determine the quality, concentration and size of the RNA extracted from caesium chloride fractions of purified virus, 5 µl of extracted RNA was analysed by formaldehyde RNA denaturing electrophoresis. 10 X MOPS [4-Morpholine-Propane Sulfonic Acid] (pH 7 with 1 M sodium hydroxide) buffer was prepared by combining 200 mM MOPS, 50 mM sodium acetate, and 10 mM EDTA and filter sterilized through Whatman 1 Qualitative paper before use. RNA loading buffer was prepared by mixing 6.5 ml of formamide, 2 ml of formaldehyde, 1.25 ml of 10 X MOPS buffer, 0.25 ml of 10 mg/ml ethidium bromide and 0.5 ml of water into an appropriate container and stored at 4°C. Running buffer (1 X) was prepared by mixing 100 ml of 10 X MOPS buffer, 20 ml of formaldehyde, and 880 ml of water. RNA were treated by adding 5 µl of RNA to 20 µl of RNA loading buffer and incubated at 60°C for 15 minutes. An RNA molecular weight marker (Sigma, 0.28-6.6 kb) was treated in exactly the same manner. The gel was prepared by adding 1.65 g agarose to 15 ml of 10 X MOPS buffer and 125 ml of water. This was boiled until the agarose was dissolved, cooled to 50°C and 25 ml of formaldehyde added. All 25 µl of sample was loaded and the gel electrophoresed at 70 V for 15 minutes after which the voltage was increased to 100 V for 1.5 hours. The RNA was visualized on a U.V. transilluminator and a photograph obtained using the Kodak *ds* ID Digital science ™ v.2.0.3 program.

2.2.3. Northern blotting protocol.

Amersham Hybond-N+ membranes were used and solutions made according to the protocol of Sambrook *et al*. (1989). The agarose gel containing the electrophoresed RNA was rinsed and the RNA transferred to membrane as described by Sambrook *et al*. (1989). Transfer occurred via capillary action overnight after which the apparatus was dismantled and a photograph of the agarose gel taken to confirm that transfer had indeed taken place. The membrane was first orientated according to the gel and checked for the

successful transfer of RNA by using a shortwave hand held U.V light. The membrane was then fixed by placing it between two sheets of filter paper at 80°C for 2 hours after which it was sealed in a plastic sleeve and stored at -80°C.

2.2.4. *In vitro* transcription of positive controls.

Three constructs named pCW4, pGEM6625 and pCW17R were used to generate specific transcripts. The clone pCW4 contained the full capsid coding sequence of NβV, pGEM6625 (Luke, 2001), the entire genome of NβV and pCW17R, the coding sequence of the capsid gene of NωV. pCW4 and pCW5 (truncated version of NβV capsid gene) were generated by cloning polymerase chain reaction (PCR) fragments obtained from pGEM6625 into the vector pGEM-T Easy (Promega). The presence of a T7 promoter on the 5` side of the pGEM-T Easy cloning cassette and the Sp6 promoter sequence on the 3`end of the cloning cassette enabled production of both positive and negative sense transcripts for later use.

A T7 Transcription kit (MBI Fermentas) was used to create transcribed RNA in large quantities to be used as positive and negative sense probes for Northern analysis of viral RNA. Plasmid DNA of the three above-mentioned constructs were purified (High Pure plasmid Isolation kit - Roche Applied Science) and 1 μg of DNA used to generate RNA transcripts according to the manufacturer's instructions. In order to prevent accessory transcripts complementary to the synthesized ones, the DNA to be transcribed was digested with a restriction endonuclease creating 5`overhangs before or after the promoter and gene of interest. *Nco* I was used to linearize pCW17R, *Nde* I for pGEM6625 and *Bam* HI for pCW4. The DNA was quantified and a transcription reaction prepared according to the manufacturer's instructions. T7 RNA polymerase was used to create positive sense RNA transcripts from the construct pGEM6625 and Sp6 RNA polymerase, to create negative sense RNA transcripts from the constructs pCW4 and pCW17R. The RNA was

quantified, sized relative to the other transcribed constructs, diluted accordingly, and loaded on a denaturing agarose gel along with extracted viral RNA as described in Section 2.2.1.

2.2.5. Production of radiolabelled RNA probes.

The plasmid pCW4, containing the capsid coding sequence of NβV, was used to create a negative sense RNA probe for the detection of this gene in the Northern analysis of extracted viral RNA . Similarly, pGEM6625 was also used to create a positive sense probe for the detection of NβV genome. A negative sense probe in the form of transcribed RNA from the construct pCW17 containing the capsid coding sequence of NωV, was also prepared as described in Section 2.2.4. The labelling reaction was completed using Sp6 RNA Polymerase and T7 RNA polymerase Transcription kit (MBI Fermentas) as per the manufacturer's instructions using 50 μCi of α^{-32}P CTP [10mCi/ml] (Amersham Biosciences). Similarly, a cold reaction was prepared as above. The transcribed RNA from both sets of reactions was precipitated as described in Section 2.3.1 and the pellet resuspended in 100 μl of DEPC-treated water. The hot reaction was denatured at 92°C for five minutes before use and the cold reaction electrophoresed on an agarose gel to confirm transcription of DNA into RNA.

2.2.6. Hybridization conditions.

The nylon membranes were placed in a hybridization bottle containing 12 ml of pre-warmed (65°C) hybridization solution (Appendix D). This was incubated with rotation at 60°C for 4 hours in a hybridization oven (Techne Hybridiser HD-1B) after which the hybridization solution was replaced with approximately 2 x 10^6 cpm of each probe in 12 ml pre-warmed hybridization solution, and left to hybridize overnight at 60°C with rotation. The blot was then washed twice with 20 ml of wash solution (Appendix D) at room

temperature and a further 20 ml of wash solution (at room temperature) was added to the membrane. This was incubated at 60°C for 20 minutes with rotation. During this time, 20 ml of wash solution was heated to 65°C and added after the first wash. This second wash step was left for 4 hours with rotation at 60°C. The blot was removed from the hybridization tube, excess liquid blotted off with Whatman 3MM paper and a Geiger counter passed over it to check for hybridization and background noise. The blot was then sealed in a clean plastic sleeve and exposed to X-Ray film (Agfa CP-BU Medical X-ray film, 180 X 240 mm) for various time intervals in a Kodak BioMax intensifying screen and cassette at -80°C.

2.2.7. Reverse transcription of viral RNA and sequence analysis of cDNA.

The Expand Reverse Transcriptase kit (Roche Applied Sciences) was used to generate all of the cDNA. The viral RNA extracted above was treated with DNAse (MBI Fermentas) to eliminate any contaminating DNA. A reverse transcription polymerase chain reaction (RT-PCR) was set up as follows: 5 µl of viral RNA was used in each reaction set up according to the manufacturer's instructions. A negative control was also prepared by substituting the RNA with water. A PCR reaction using the cDNA generated from the previous step was performed using the Expand High Fidelity PCR system [Roche Applied Sciences] as per the manufacturer's instructions. The primers RDNBVCAPF1 and RDNBVCAPR (Appendix A) corresponding to nucleotides 4030-4061 and 5860-5885 of the NβV genome were used with an annealing temperature of 52°C and a 3.5 minute extension time for 25 cycles.

A similar reaction was performed on a further 5 µl of viral RNA using the primers NβV REP2 and NβV CAP1 (Appendix A) corresponding to nucleotides 3248-3265 and 4300-4317 of the NβV genome. The reaction was set up as above with the exception of the annealing temperature which was decreased to 48°C. The amplified products were concentrated and purified

using a Zymo column (Zymo Research) according to the manufacturer's instructions and the DNA eluted in a final volume of 16 µl using $dddH_2O$. The eluted RT-PCR products were ligated into the pGEM-T Easy vector (Promega) as per the manufacture's instructions. The ligated DNA was transformed into competent *E. coli* DH5α cells using the heat-shock method described by Sambrook *et al*. (1989) and colonies containing recombinant constructs selected by growth on MacConkey agar plates containing 100 µg/ml Ampicillin (Roche). Colonies containing recombinant plasmids were inoculated into 5 ml Luria broth containing ampicillin and incubated at 37°C overnight. Plasmid DNA was extracted using the protocol of Berghammer and Auer (1993) (Appendix C). The extracted plasmid DNA was analysed by restriction endonuclease digestion. *Eco* RI was used to determine the presence of inserts since there are two restriction sites flanking the cloning site in the pGEM-T Easy vector.

Plasmid DNA was prepared using the High Pure Plasmid isolation kit (Roche) and further screened by restriction endonuclease digestion. Recombinant clones were selected on the basis of similarity of restriction fragments with the NβV genome sequences as described by Gordon *et al*. (2001). The primers pucF and pucR (Appendix A), which are complementary in sequence to nucleotides on either side of the pGEM-T Easy vector cloning cassette, and NβV CAP1 and NβV REP2 were used to determine the nucleotide sequence of the recombinant constructs. The Sanger dideoxy sequencing method and the DNA sequencing kit Big DYE Terminator V 3.0 were used as per the manufacturer's instructions. A GeneAmp® PCR cycler version (9700) was used to extend the DNA. The extension products were then concentrated and purified using a Zymo column, dried under vacuum using a SpeedVac (Savant) and the nucleotide sequence determined using an ABI 3100 instrument (Rhodes University DNA Sequencing Unit). The sequences were analysed using the Chromas® V. 2 program and the BLAST 2.2.6 algorithm hosted by the NCBI database. Plasmid maps of the constructs were created in Vector NTI and appropriately named and described (Appendix G). Finally a composite alignment of the all the generated

sequences of the cDNA clones was produced and compared to the original sequence posted in the NCBI Database by Gordon *et al*. (1999).

2.3. Viral protein analysis

Gradient fractions containing viral protein were analysed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) on 7.5 % SDS-PAGE gels prepared according to Laemmli (1970). Exactly 10 µl of each sample was mixed with 10 µl of 2 X Sample buffer (Appendix E), heated to 100°C for 5 minutes and loaded into individual wells. A protein molecular weight marker (MBI Fermentas) was treated in the same manner as the samples. The gel was electrophoresed at 150 V for 1.5 hours, placed in Coomassie stain overnight, and then destained as described in Appendix E.

Samples containing viral proteins were electrophoresed on a SDS-PAGE gel as described above and transferred to Hybond C membrane (Amersham Pharmacia Biotech) as per the manufacturer's instruction using a Biorad protein transfer apparatus. The proteins were transferred under constant resistance of 250 mA for 1 hour, the apparatus disassembled and the membrane stained with 1 X Ponceau in order to detect the marker bands. The membrane was then blocked overnight in 5 % BLOTTO in TBS-TWEEN. The membrane was rinsed in TBS-TWEEN for 30 minutes followed by another four 5 minute washes. A 1 in 10 000 dilution of primary antibody (anti-NβV or anti-NβV+NεV polyclonal antibodies courtesy of D. Hendry) diluted in TBS-TWEEN containing 1 % BSA was added directly to the membrane. The membrane was left to rock at room temperature for 2 hours, rinsed and 1 µl of secondary antibody (anti-mouse/rabbit IgG-POD antibody – Roche Applied Science) added directly to the membrane diluted in 12.5 ml of 1 % BSA in TBS-TWEEN. Following the manufacturer's instructions, the blot was incubated, rinsed and the chemiluminescence substrate added. The blot was then exposed to X-Ray film (Agfa CP-BU Medical X-ray film, 180 X 240

mm) for various time intervals in a Kodak BioMax cassette and film developed at varying time intervals.

Protein samples were concentrated using the trichloroacetic acid (TCA) precipitation method. This was done by adding 2 % (w/v) of deoxycholate to a sample a final concentration of 0.02 %. This was incubated at room temperature for 15 minutes and 24 % freshly prepared TCA added to a final concentration of 6 % followed by incubation on ice for 1 hour. The sample was then microfuged at 4°C for 10 minutes at 13 000 rpm, the supernatant carefully removed and 200 µl of ice-cold acetone added to the pellet. The mixture was left to incubate on ice or 15 minutes and microfuged as above. The pellet was left to dry and then resuspended in EB buffer (pH 7.5).

The removal of CsCl from fractions containing viral protein was achieved by the addition of one or more fractions to the top of a centrifuge tube filled with EB buffer followed by centrifugation for 5.5 hours at 4°C at 26 200 rpm using a SW 28 Beckman rotor. The pelleted virus was left to resuspend overnight in 100 µl EB buffer at 4°C, transferred to an Eppendorf tube and stored at -20°C.

2.3.1. Analysis of virus particles using transmission electron microscopy (TEM).

Virus preparations and virus-like particles (VLPs) were examined by transmission electron microscopy using a Phillips JEOL 1210 Transmission Electron Microscope. Carbon/formvar copper grids were prepared with 2 % uranyl acetetate in order to visualise negatively stained particles. A second method, described by Dong *et al*. (1998), was used to study purified virions and virus-like particles. This method involved the adsorbtion of sample onto a copper grid (repeated if necessary), followed by the removal of excess liquid with filter paper. The grid was then floated on bubbles of 50 mM HEPES (pH 7) three times, excess liquid blotted with filter paper and finally, the grid stained three times with 1 % uranyl acetate.

2.4 Construction of recombinant baculoviruses.

2.4.1. Transformation and isolation of recombinant bacmid DNA.

The constructs pCW4 and pCW5 were used to create pCW6 and pCW7, respectively. These constructs were the expression vectors used to create recombinant baculoviruses containing the capsid gene and a truncated version to be used to transfect insect cells. This was achieved by excising the *Bam* HI/*Sac* I fragment from both of these constructs and ligating this into pFASTBAC Dual that had been cut with *Bam* HI/*Sac* I. pCW4 and pCW5 were constructed by the PCR of the full and truncated capsid coding sequences of the NβV capsid ORF from the plasmid pGEM6625, using the forward primers RDNBVCAPF1 and RDNBVCAPf2, respectively (Appendix A), creating a *Bam* HI site at the start of the coding sequence. The cycle parameters for this PCR are summarized in Table 2.1 and the reaction prepared as described in Section 2.3.7 with the exception of substituting the cDNA with 50 ng of pGEM6625.

RDNBVCAPF2 binds 150 nt downstream from the +1 nt and creates a *Bam* HI site upstream on the second in-frame AUG codon which will act as a second translational start upon the expression of the capsid coding sequence. The PCR product from this reaction was purified and concentrated using a Zymo column and ligated into the pGEM-T Easy vector. The ligation reactions were transformed into competent *E. coli* DH5α cells and treated as described in Section 2.2.7. Restriction analysis was used to screen for recombinant constructs using the enzymes *Bam* HI and *Sac* I. Colonies exhibiting the correct sized insert were sequenced with the oligonucleotides RDNBVCAPFI for pCW4, RDNBVCAPF2 for pCW5 and RDNBVCAPR as a reverse primer as described in Section 2.2.7. A summary of the sub-cloning process for pCW4 to create pCW6 (Figure 2.2).

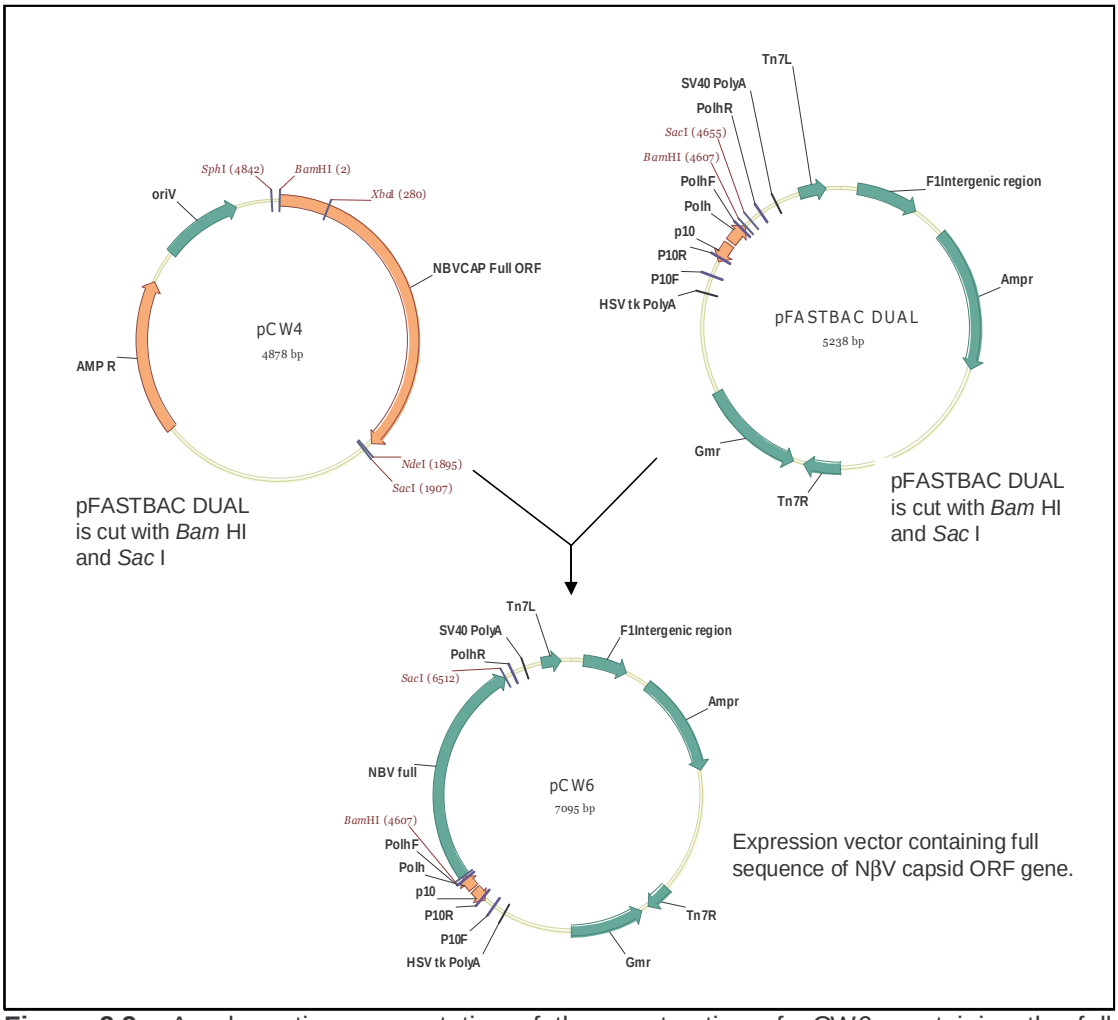


Figure 2.2. A schematic representation of the construction of pCW6, containing the full length capsid coding sequence of NβV.

The transformation of pCW6 and pCW7 into *E.coli* DH10BAC™ cells and subsequent isolation of recombinant bacmid DNA was carried out as described in the instruction manual for the Bac-to-Bac® Baculovirus Expression Systems from GibcoBRl (Invitrogen™ Life Technologies). Verification of the insertion of the NβV full length and truncated capsid ORF into the bacmid DNA was done by PCR analysis of the isolated bacmid DNA using the primers pucR and pucF (Appendix A) which bind to regions which flank the point of insertion in pFASTBAC Dual. PCR products from this reaction should be 4101 and 3951 bp in size with pCW6 and pCW7 as template respectively, would denote the presence of an insert, while a PCR product of 273 bp indicates an unmodified vector lacking an inserted gene or a gentamycin resistance gene and flanking sequences (Figure 2.3). The

amplification reactions were prepared using 1 µl of template bacmid DNA and *Taq* DNA polymerase (Bioline) following the manufacturer's instructions. The reaction was cycled in a Hybaid PCR sprint machine using an annealing temperature of 55°C and an extension time of 5.5 minutes with 30 cycles. Figure 2.3 demonstrates the principle of the integrated gene of interest, surrounding genes and the expected sizes of an amplified product, using the primers pucF and pucR.

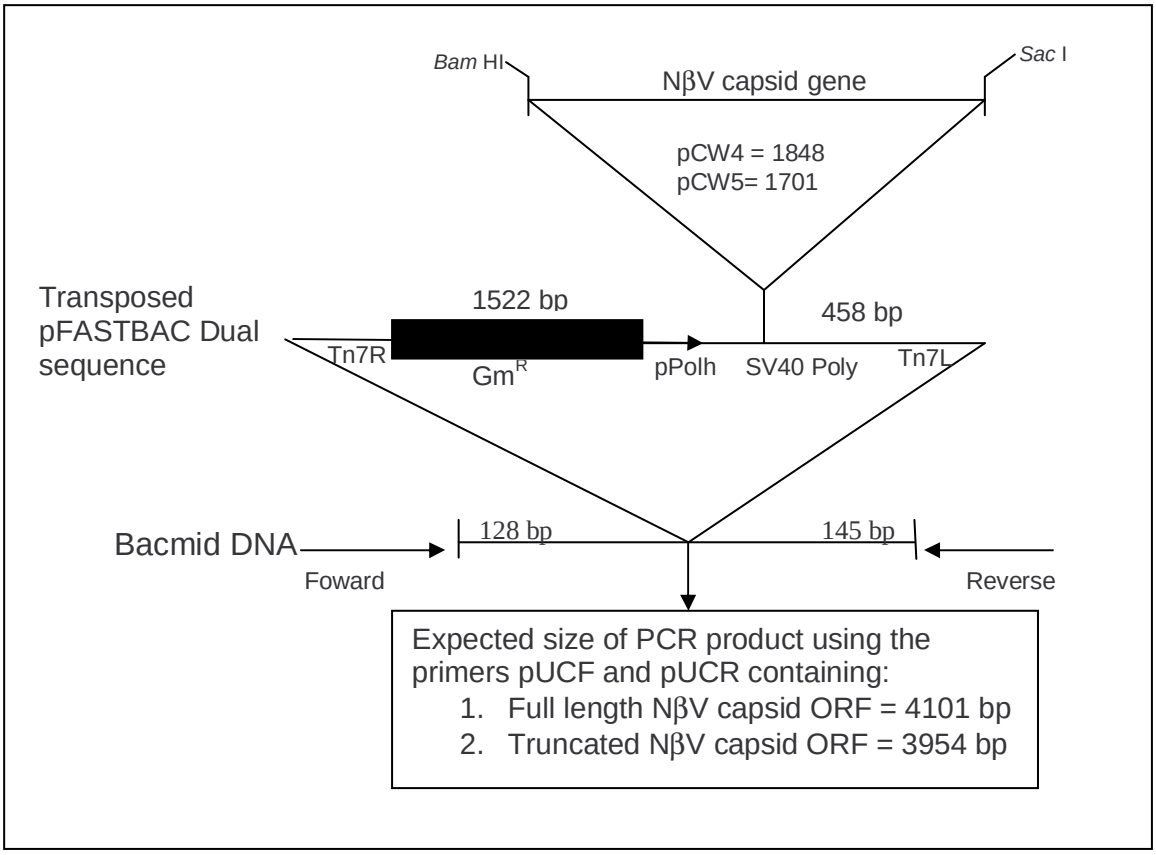


Figure 2.3. Schematic representation of the approach used to confirm the presence of inserts in recombinant bacmid DNA.

2.5. Baculoviral expression of NβV capsid constructs.

2.5.1. Transfection of insect cells with recombinant baculovirus DNA.

The protocols described below were adapted from the Bac-to-Bac® Baculovirus Expression Systems instruction manual from Invitirogen™ Life Technologies. All *Spodoptera frugiperda* 9 (*Sf*9) cell cultures were grown as monolayers in Grace's complete medium (545 ml Grace's medium, 10% foetal calf serum, and 10 µl of 100 x penicillin (10 000 U/µl), and streptomycin (10 mg/ml) (Highveld Biological Pty Ltd.) per litre of medium at 27°C in an incubator (Labcon). The cells were obtained from P. de Philepe, (Department of Biomedical Sciences, St. Andrews University, Scotland). A single 35 mm 6 well tissue culture dish was seeded with $1x10^6$ cells/well and incubated at room temperature for 1 hour to allow the cells to attach. The growth medium was removed and replaced with 2 ml of fresh Grace's insect medium. Recombinant bacmid DNA and vector DNA was then transfected into the cells as follows: Solution A was prepared by adding 5 µl Fugene transfection reagent (Roche) drop wise to 95 µl of Grace's medium and incubated at room temperature for 5 minutes. Solution B was prepared by diluting 6 µl of recombinant bacmid DNA to a final volume of 25 µl using Grace's medium. Solution A was added drop wise to solution B to a final volume of 100 µl and incubated at room temperature for 30 minutes. The mixed solutions were added slowly to the seeded wells. The 35 mm dish was placed in a plastic container with wet paper towels to prevent drying out of wells and incubated at 28°C for 4 days. After 4 days the supernatants were transferred to sterile tubes and microfuged for 2 minutes at 2000 rpm to pellet the cells and debris while the recombinant baculovirus remained in the supernatant.

2.6.1. Virus amplification.

Passage 1 virus stocks: 35 mm 6 well plates were seeded with *Sf*9 cells at a density of $1x10^6$ cells/well and incubated at room temperature for 1 hour to allow the cells to attach. The Grace's medium was then removed and the cells infected with 0.5 ml of purified supernatant from transfected *Sf*9 cells. The cells were gently agitated at room temperature for 1 hour and 1.5 ml of Grace's medium added to the infected cells. The cells were incubated at 28°C for four days and the supernatant transferred to a sterile tube. The supernatant was microfuged at 2000 rpm for 5 minutes. The resultant supernatant was stored at 4°C to be used to generate virus passage 2 stocks.

Passage 2 virus stocks: T75 (75 cm^2) tissue culture flasks containing confluent cells were used to generate passage 2 stocks. The 15 ml of medium used to grow the cells was removed and replaced with 8 ml of fresh Grace's medium. 250 µl of passage 1 virus stock was added directly to the medium and the flask incubated at 28°C with gentle agitation for 2 hours. A further 7 ml of Grace's medium was added to each flask and incubated at 28°C for 4 days. The 15 ml supernatant was transferred to a sterile Beckman JA20 centrifuge tube and centrifuged at 2000 rpm for 10 minutes in a Beckman JA20 rotor to pellet cellular debris. 1 ml aliquots of the supernatant were made, stored at 4°C and used to generate virus passage 3 stocks.

Passage 3 virus stocks: T75 flasks were prepared in the same manner as described for the passage 2 stocks. The cells were infected with 250 µl of the passage 2 virus stock and incubated at 28°C with gentle agitation for 2 hours. 9 ml of Grace's medium was added to each flask and incubated at 28°C for 4 days. The 15 ml of supernatant was transferred to a sterile Beckman JA20 centrifuge tube and centrifuged at 2000 rpm for 10 minutes at 4°C in a Beckman JA20 rotor to pellet cellular debris. The passage 3 stocks supernatants were stored at 4°C and used to infect insect cells and generate a multiplicity of infection by using the $TCID_{50}$ assay.

2.6.2. 50 % Tissue culture infectious dose ($TCID_{50}$) Assay.

The virus titre of the supernatant of viral passage 3 stocks of baculovirus-mediated expression of bCW6 and bCW7 was determined using the $TCID_{50}$ assay described by O'Reilly *et al*. (1992). *Sf*9 cells were used to seed 96 well plates at $2x10^4$ cells/well and left to incubate for 1 hour to allow the cells to attach. A ten-fold dilution series of the virus in Grace's medium was made up to 10^{-12}. Once the cells had attached, the 10^{-8} to 10^{-12} dilution was used to infect the wells. A microtitre plate was used for the assay and each dilution was repeated 8 times so as to fill the plate with the above-mentioned dilution of recombinant bacmid DNA containing the genes from bCW6 and bCW7. Two columns of media and cells only were prepared between columns of the two recombinant constructs. All of the remaining wells were infected with 50 µl baculovirus per well and incubated at 28°C for 10 days. After 5 days of incubation, each well was fed with 10 µl of Grace's medium and after 10 days each well was scored for virus by evaluating the cytopathic effect.

2.6.3. Extraction of baculovirus DNA

A total volume of 6 ml of high titre recombinant virus was centrifuged at 12 000 g for 15 minutes at 4°C and the pellet resuspended in 100 µl of virus disruption buffer (10 mM Tris pH7.6, 10 mM EDTA and 0.25 % SDS). Proteinase K was added to a final concentration of 500 µg/ml and the sample incubated at 37°C until the resuspended pellet appeared translucent. The resuspended pellet was extracted with equal volumes of phenol, phenol:chloroform:isoamyl [P.C.I (25:24:1)], and chloroform:isoamyl (24:1). The first organic phase from the phenol extraction was back-extracted in equal volumes of TE buffer (pH 8). The baculovirus DNA was precipitated in 205 µl of 3M sodium acetate (pH 5.2) and 410 µl of ice cold 96 % ethanol. The DNA was left to precipitate overnight at -20°C, centrifuged at 12 000 g for 16 minutes at 4°C, the pellet washed twice with 200 µl of ice cold 70% ethanol, air dried and stored at 4°C.

The baculovirus DNA was analysed by PCR reactions to confirm the presence and size of inserted capsid genes using the primers RDNBVCAPF1/ RDNBVCAPF2 and CWR1 which will generate amplified products of 1302 and 1152 bp for bCW6 and bCW7, respectively. The verification on an intact polyhedrin promoter and accompanying gene was achieved by PCR using the primers CWR1 and CWF3 (Appendix A) which bind to the centre of the NβV capsid gene ORF and the start of the polyhedrin promoter, respectively. The expected sizes of the PCR products were 1279 and 1398 bp for the constructs bCW7 and bCW6, respectively. The PCR reactions were set up according to the manufacturer's instructions and O'Reilly *et al*. (1992) using 3 µl of baculovirus DNA and Expand High Fidelity Polymerase (Roche). An annealing temperature of 47°C and extension time of 1 minute for 30 cycles was employed as well as an extension of each elongation cycle by 5 seconds.

2.6.4. Expression of recombinant NβV capsid constructs in *Sf*9 cell lines.

Four 100 % confluent T75 flasks containing *Sf*9 cells per recombinant bacmid construct were infected with an MOI of 5 using the supernatant of virus passage. One flask was left uninfected and used as a negative control. These 9 flasks were incubated at 28°C for 10 days and the VLPs extracted in acidic (pH 5.0) buffers as described below.

2.6.5. VLP purification.

The protocol used for the purification of VLPs from cell culture was modified from the protocol used by Canady *et al*. (2000). All buffers were adjusted to a pH of 5 instead of the reported 7.5. The flasks containing infected cells described above were harvested by lysing the infected cells of each flask with the addition of 0.5 % (v/v) of NP40 (Roche) to each flask, followed by incubation of 15 minutes on ice. The cell lysate was centrifuged at 5 000 rpm in a Beckman JA20 rotor for 10 minutes at 4°C. The supernatant was transferred to a 50 ml Falcon tube, RNAse A (10 mg/ml) [Appendix C] added

to the supernatant to a final concentration of 10 μg/ml and incubated at 28°C for 30 minutes with gentle agitation. The cell lysate was then centrifuged through a 30 % sucrose cushion made in EB buffer [50 mM sodium acetate, 250 mM NaCl] (pH 5.0), in a Beckman SW 28 for 5.5 hours at 26 200 rpm at 4°C. The sucrose pellets were resuspended overnight in 2 ml of EB buffer (pH 5.0).

The resuspended pellets were centrifuged through a 10-40% continuous sucrose gradient made in EB buffer as described in Section 2.1. The gradients were centrifuged in a SW41 rotor at 40 000 rpm for 1.25 hours at 4°C. 400 μl fractions were collected using an ISCO density gradient fractionator and the absorbance readings of each fraction at both 260nm and 280nm measured. Fractions containing protein were analysed further using SDS-PAGE, Western analysis and transmission electron microscopy. The 9 flasks containing infected *Sf*9 cells were harvested as above using EB buffer at a pH of 5.0 in order to mature VLPs during the purification process as a result of a decrease in pH to 5.0 as described by Taylor *et al*. (2002). 50 mM sodium acetate was adjusted to pH 5 and EB buffer at pH 5, was used to make up 10, 30 and 40 % sucrose in order to induce maturation of virus-like particles during the extraction process. These acid-matured particles were analysed by SDS-PAGE, Western and TEM analysis. Anti-NβV+NεV antibodies were used in this set of experiments due to diminished stocks of anti-NβV antibodies. Although these antibodies were raised against both *Nudaurelia* β Virus and *Nudaurelia* ε Virus, they were used in the Western analysis of expressed VLPs since little cross-reactivity with other cellular proteins and the absence of other viruses was expected.

Chapter 3

RESULTS AND DISCUSSION

3. Introduction

In 2002 a large population of *N. capensis* was observed in the MTO pine plantations on the slopes of the Simonsberg near Stellenbosch in the Western Cape Province of South Africa. By September 2002, foresters reported signs of severe mortality with large numbers of dead or dying larvae on the forest floor (Figure 3.1). The symptoms resembled those reported before by Hendry *et al.* (1985) and Juckes (1974) raising the possibility of a potential tetravirus infection. The larvae were flaccid, discoloured and those that were still alive, had ceased to feed, were vomiting and often found to be hanging from their prolegs from the trunks of pine trees. Several kilograms of dead or dying larvae were collected in 2002, 2003 and 2004 and frozen at -20^{0}C. A general tetravirus extraction protocol modified from Morris *et al*. (1979) was used for virus extraction.

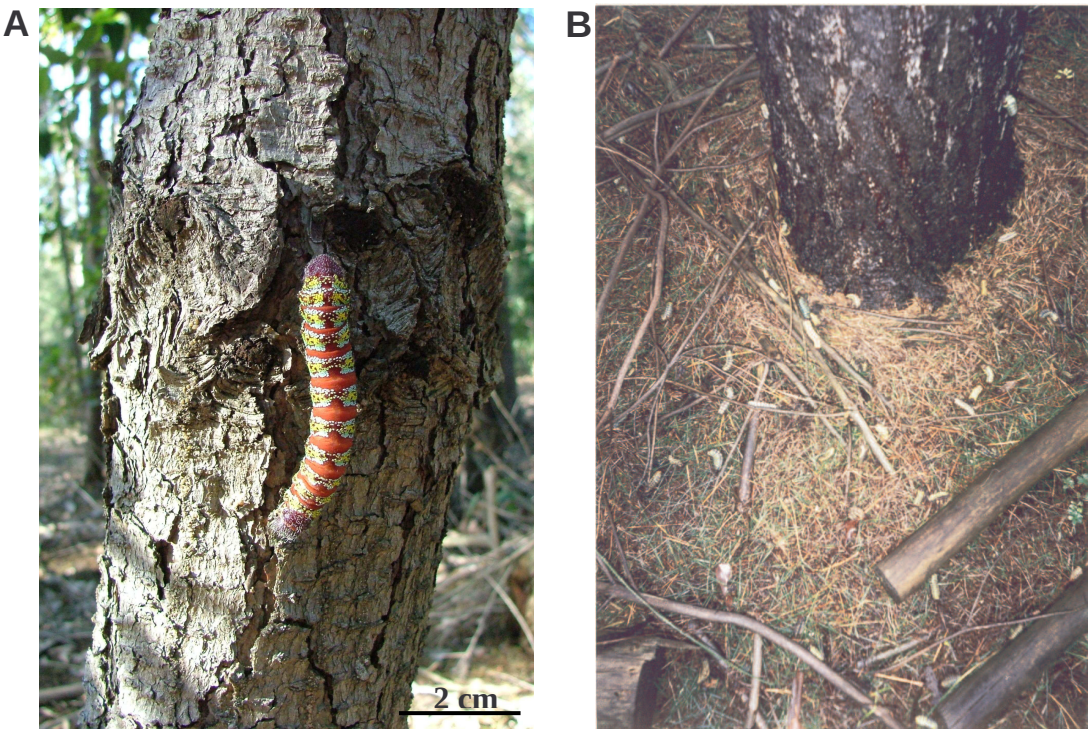


Figure 3.1. ***N. capensis*** **larvae in the pine plantations of the Simonsberg mountains. Panel A.** Larva on the trunk of a pine tree. **Panel B.** Dead and dying larvae around the base of a pine tree.

3.1. Virus purification

Two hundred and fifty grams of frozen larvae was used for each individual virus extraction. The resuspended pellet obtained after centrifugation through a 30% sucrose cushion, was analysed by CsCl buoyant density ultracentrifugation in order to separate out any potential tetraviruses on the basis of buoyant density. This was because tetraviruses have very similar sedimentation coefficients (Hanzlik *et al*., 1995) and sucrose gradients would therefore not be able to separate out extracted viruses.

Fractions collected from each CsCl gradient were assayed by spectrophotometric analysis to select those potentially containing virus. Two peaks at OD_{260} and OD_{280} were observed (Figure 3.2, Fractions 18-19 and 21, respectively).

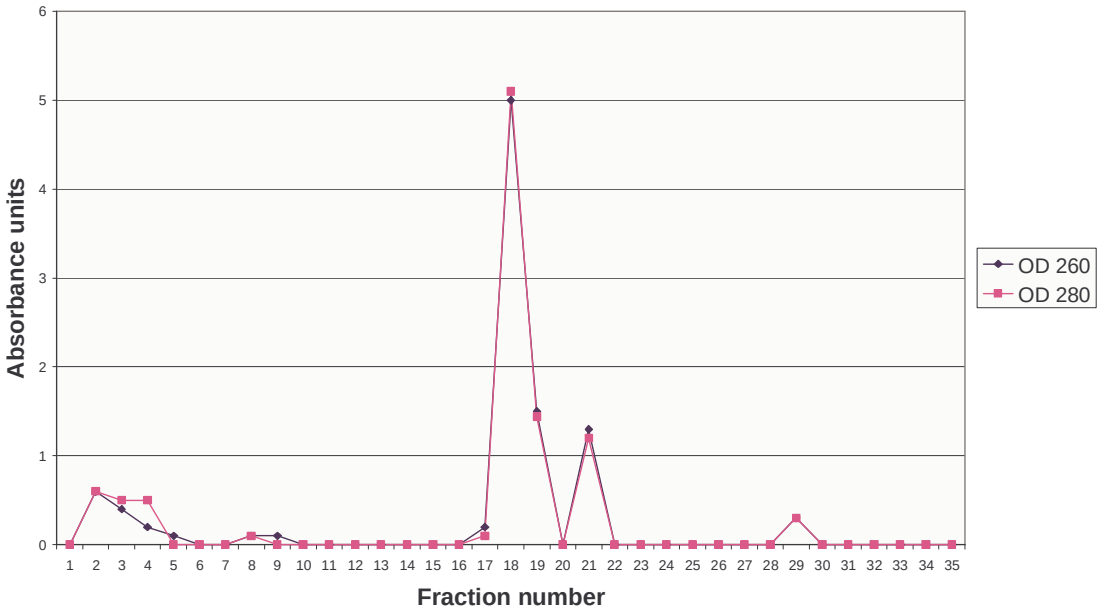


Figure 3.2. Spectrophotometric analysis of CsCl fractions containing purified virus prepared from *N.capensis* larvae collected in 2003 and subjected to CsCl density gradient ultracentrifugation. Thirty-five 200:l fractions were collected and analysed at OD_{260} and OD_{280}. The X-axis depicts fractions collected from the top to the bottom of the gradient.

The refractive index of each fraction was measured and a buoyant density of 1.320 g/cm^3 was calculated for fraction 18, which is close to the buoyant density measured for NβV of 1.330 g/ ml (Hendry *et al*., 1985) as opposed to NωV (1.285g/ ml) or HaSV (1.296 g/ ml) (Hanzlik and Gordon, 1997).

3.2. SDS-PAGE and Western analysis of *Nudaurelia* virus preparations.

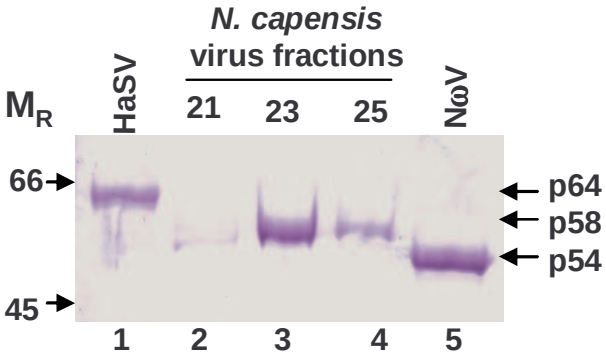


Figure 3.3. SDS-PAGE analysis of selected *Nudaurelia* virus fractions and compared to related tetraviruses. Coomassie-stained 7.5 % SDS-PAGE gel of purified virus samples with HaSV (Lane 1), fractions 21, 23 and 25 from a CsCl gradient with the *Nudaurelia* virus preparation (Lanes 2-4) and NωV (Lane 5). Arrows indicate the molecular weights of the proteins (in kDa).

To confirm the presence of virus particles, virus extractions were performed on fresh larval samples, subjected to CsCl density ultracentrifugation and the selected fractions corresponding to a buoyant density of 1.320 g/ml selected for analysis by SDS-PAGE. In the case of the fractions of the virus preparation analysed by spectrophotometric analysis shown in Figure 3.2, fraction 18 was selected while a second virus preparation yielded the same buoyant density in fractions 21-25 (Figure 3.3). A protein migrating at approximately 58 kDa (p58) was observed in all selected fractions (Figure 3.3 Lanes 2-4). This band did not co-migrate with the capsid proteins of the HaSV (p64) or a wild-type preparation of NωV isolated from the same population in 2002 (DA Hendry unpublished data) (Figure 3.3, Lanes 1 and 5). The size of p58 corresponded to the predicted molecular weight of 58.448 kDa for the mature capsid protein of NβV (Gordon *et al*., 1999). Interestingly, the capsid protein of the NωV isolate migrated at an approximate size of 54 kDa and not the reported 62 kDa (Figure 3.3, Lane 5) (Hendry *et al*., 1985).

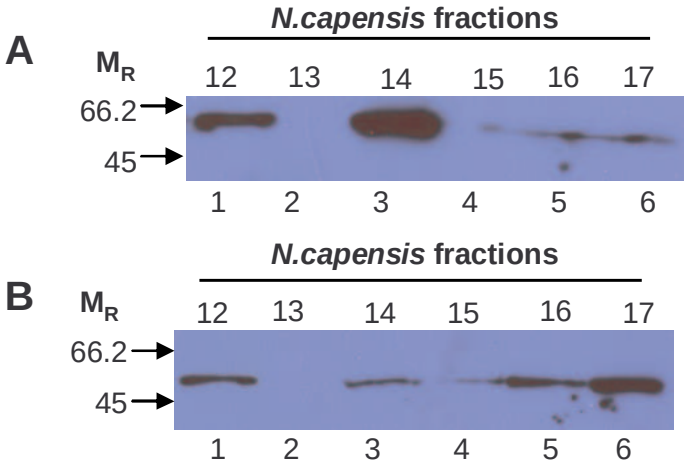


Figure 3.4 Western analysis of purified virus fractions. **Panel A.** Western blot of fractions containing virus probed with anti-NβV polyclonal antibodies. **Lanes 1-6** represent virus- and RNA-containing fractions collected from a CsCl gradient. **Panel B**. Western blot depicted in Panel A that was stripped and reprobed with anti-NωV polyclonal antibodies. Arrows indicate the molecular mass of the proteins (in kDa). Samples were electrophoresed on a 7.5 % SDS-PAGE gel.

In order to confirm that the proteins isolated from the larvae were indeed the capsid proteins of NβV, Western analysis was performed on purified virus using polyclonal antibodies raised against NβV. A virus preparation, using fractions 12-17 corresponding to fractions containing sufficient concentrations of protein and nucleic acid as well as the expected approximate buoyant density of NβV were analysed in this manner.

A protein of approximately 58 kDa was detected with the anti-NβV antibodies (Figure 3.4 Panel A, lanes 1, 3-6), confirming the presence of NβV in the virus preparation. Due to presence of NωV in larvae collected in 2002, the membrane was stripped and reprobed with anti-NωV antibodies. The anti-NβV antibodies also detected proteins of approximately 58 kDa in the same virus samples (Figure 3.4 Panel B, Lanes 1, 3-6). However, the intensity of the bands detected by the anti-NωV antibody in each fraction was not equivalent to those detected by the anti-NβV antibodies (Compare equivalent lanes in Panels A and B of Figure 3.4). This suggested that the antibodies might be recognizing two different proteins of similar size in each of the fractions. A less likely explanation could be that there was some cross-reactivity between the antibodies. The result obtained in this experiment could be explained by the fact that the antibodies were raised against a

mixture of these two viruses or alternatively, an NωV-like virus was present in this particular sample of purified fractions.

Since the molecular mass of the reaction on both of the Western analyses (Panel A and B) was in the range of 55-58 kDa and that the reported buoyant density of NωV is 1.29 g/cm^3 (less dense than NβV), it is likely that an ω-like virus exists in this particular preparation. The proteins observed when the viral proteins were analysed with anti-NβV antibodies migrated to a molecular weight of 58 kDa (Figure 3.4 Panel A) whereas the reaction observed when the proteins were probed with anti-NωV (Figure 3.4 Panel B) was of a molecular weight between 55-57 kDa which indicated the presence of an ω-like virus with a buoyant density of greater than 1.330 g/ cm^3. It was also reported that the complete separation of NβV and NωV using density gradient ultracentrifugation was never successful (V.Hodgson, pers. comm.) suggesting that the antibodies were raised against a mixed virus preparations each virus in this year might have been raised against both NβV and NωV.

Close inspection of the SDS-PAGE gel in Figure 3.3 (Lanes 2-4) suggested the possibility of a second protein band of approximately 56 kDa, in the NβV-containing virus fractions, which had not been well resolved. CsCl gradient fractions containing NβV were re-analysed on a 10 % SDS-PAGE gel and stained with Coomassie to visualize the proteins.

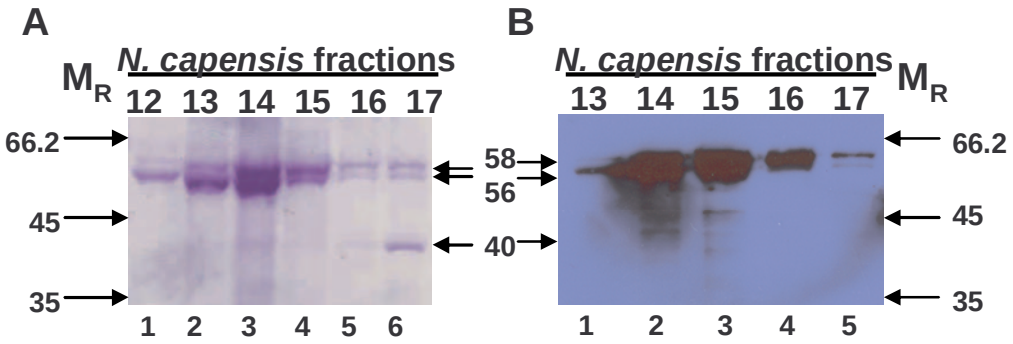


Figure 3.5. SDS-PAGE and Western analysis of concentrated *Nudaurelia* virus fractions. Panel A. Coomassie-stained 10% SDS-PAGE of purified and concentrated virus fractions 12-17 (Lanes 1- 6). **Panel B**. Western analysis of concentrated *Nudaurelia* virus fractions 12-16 (Lanes 1-5) probed with anti-NβV antibodies. Arrows indicate the molecular mass of the proteins (in kDa).

The 58 kDa protein was observed in all the fractions (Figure 3.5 Panel A, Lanes 1-6). In addition, a second protein corresponding to 56 kDa was observed in all but one fraction (Figure 3.5 Panel A, Lanes 2-6). Furthermore, a fraction of protein corresponding to the highest buoyant density also contained a protein with a calculated molecular weight of 40 kDa (p40) (Figure 3.5 Panel A, Lane 6). The presence of more than one protein species in the CsCl fractions supported the possibility that more than one virus was present in the virus preparations. Both p58 and p56 cross-reacted with anti-NβV antibodies, but p40 was not detected (Figure 3.5 Panel B, Lane 5) supporting the hypothesis that the *N. capensis* larvae were infected with NβV.

To confirm this result, the fractions were pooled, concentrated by ultracentrifugation and reanalysed by SDS-PAGE. Both p58 and p56 were again detected by the anti-NβV antibodies (Figure 3.6 Panel B) confirming the presence of two different NβV-like proteins.

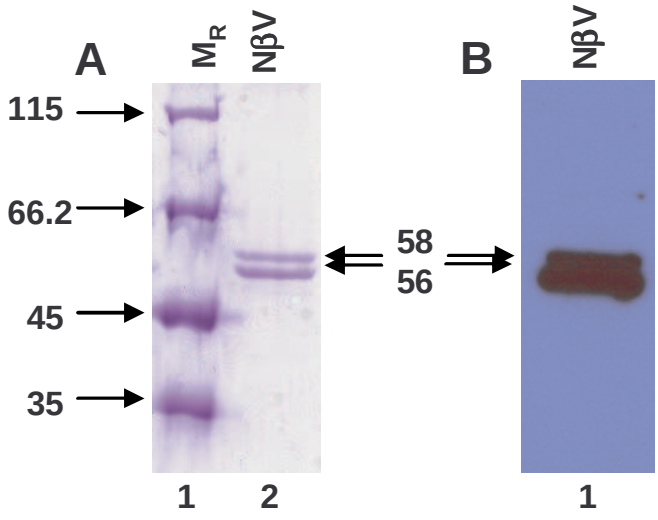


Figure 3.6. SDS-PAGE and Western analysis of purified NβV capsid protein. Panel A. Coomassie-stained 10 % SDS-PAGE gel of purified NβV (5µl which is equivalent to 5 µg of protein). **Lane 1** represents a molecular weight marker and **Lane 2** represents purified *Nudaurelia* β virus capsid protein. **Panel B**. Western analysis of purified NβV probed with anti-NβV antibodies. **Lane 1** represents purified *Nudaurelia* β virus capsid protein. The arrows indicate the molecular mass of the proteins (in kDa).

Taken together, the above data confirmed that the *N. capensis* larvae were infected with NβV. The results with the anti-NωV antibodies and the presence of p40 in the first three virus preparations suggested the presence of other viruses in the population as well. It is interesting to note that nodavirus capsid

proteins are on average approximately 40 kDa in size (van Regenmortel *et al*., 2000) which could suggest that the insects are infected with two families of insect viruses. More work and further isolation of p40 will have to be undertaken.

3.3. Transmission Electron Microscopy

The identification of two proteins (p56 and p58) both cross-reacting with the anti-NβV antibodies suggested the presence of possibly two NβV-like viruses. Transmission electron microscopic examination of the pooled virus preparations showed the presence of icosahedral virus particles (Figure 3.7 Panel A).

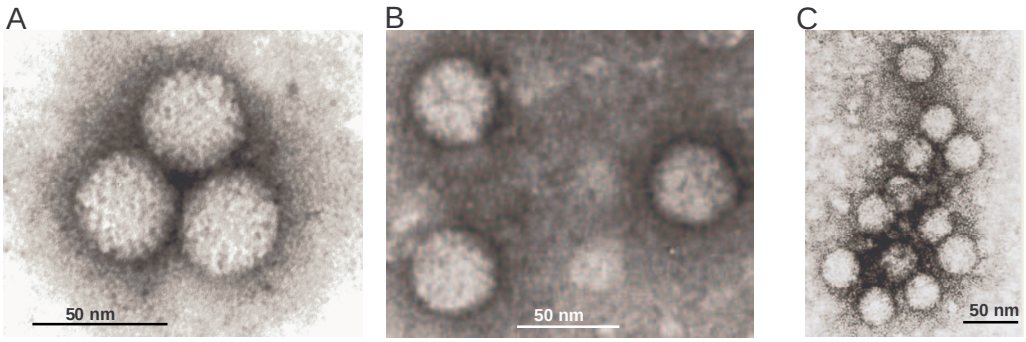


Figure 3.7. Transmission Electron Micrographs of purified *Nudaurelia* β virus, *Nudaurelia* ω virus and *Helicoverpa armigera* stunt virus. Panel A. NβV purified from larvae. **Panel B.** NωV purified in from larvae collected in 2002. **Panel C.** HaSV courtesy of T. Hanzlik. The scale bars are all representative of 50 nm.

These images were compared to images of NωV and HaSV to compare the surface structure and diameter of virus particles. When the particles purified from the larvae (Figure 3.7 Panel A) were compared to NωV (Figure 3.7, Panel B) and HaSV (Figure 3.7 Panel C) a distinct difference in the capsid surface structure was observed. The purified NβV-like particles exhibited the characteristic pitted face and deep grooves between each face observed by Olson *et al*. (1990) but not present on the surface of the Omegatetravirus particles. The diameter of the NβV-like particles was approximately 37 nm which corresponds with that observed by Olson *et al*. (1990) for NβV. The

possibility of the presence of two NβV-like viruses was investigated by measuring the diameter of 1000 virus particles.

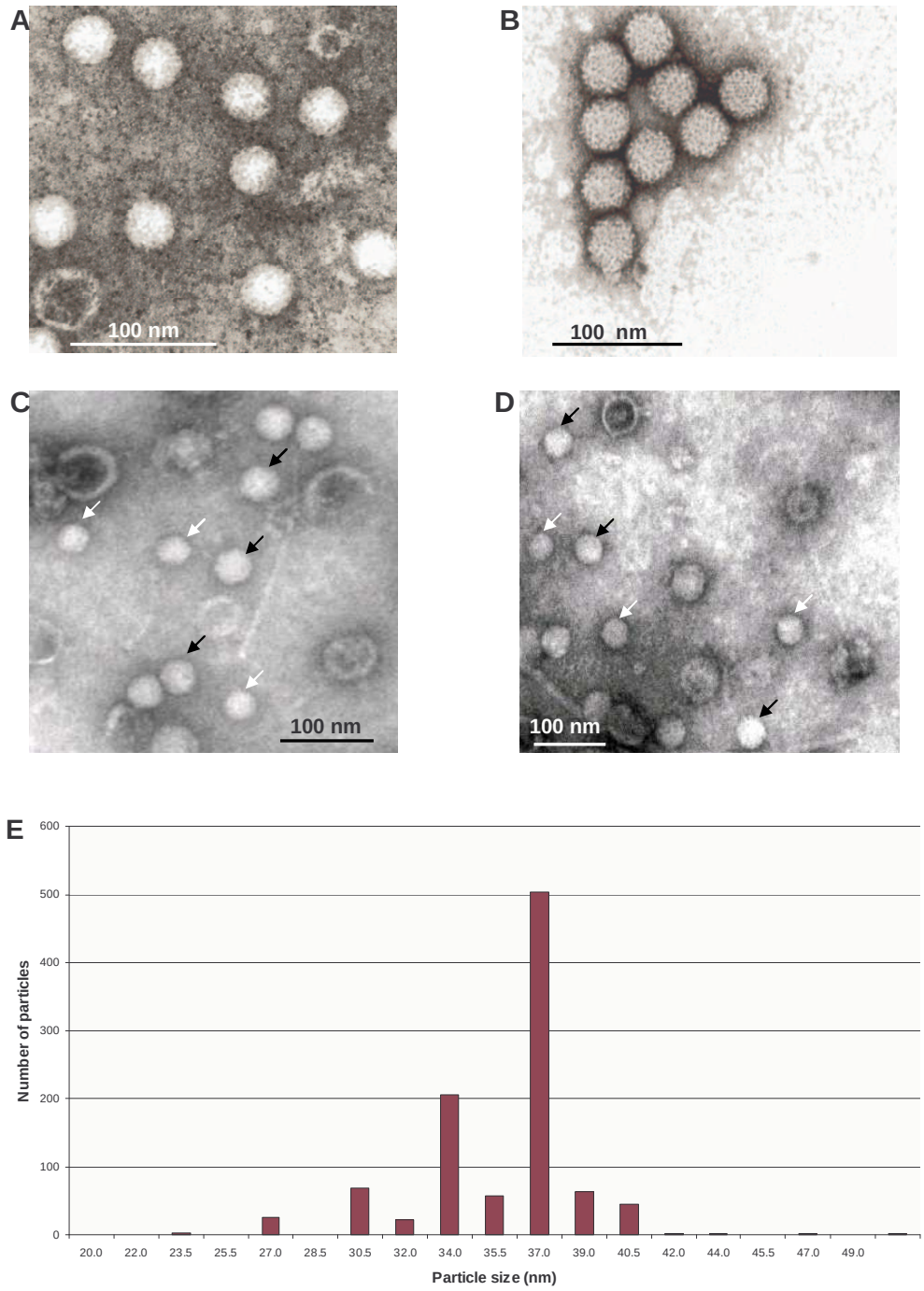


Figure 3.8. Calculation of purified virus particle size by TEM analysis. Panel A and B. Electron Micrograph of purified NβV. **Panel C and D.** Micrographs demonstrating a disparity in virus diameter. **Panel E**. Graph showing distribution of virus diameters. All scale bars represent 100 nm. White arrows indicate viruses with smaller diameters and black arrows indicate viruses with the reported diameter of 37 nm.

The average diameter of the particles that were measured was calculated to be 37.10 nm with a standard deviation of 1.21 nm. The average diameter of these particles (Figure 3.8 Panel A and B) corresponded to the reported diameter of NβV of 37 nm (Gordon *et al*., 1999). When the diameters of the particles were plotted on a graph according to distribution, it was observed that, while the majority of the particles were 37 nm in diameter, a significant proportion had a diameter of 34 nm (Figure 3.8 Panel E). A mixture of 37 and 34 nm particles is shown in Figure 3.8 (Panel C and D). The surface morphology of both these particle types exhibited the characteristic pitted face and deep grooves between each face of NβV. The presence of two distinct sizes of what appears to be NβV could explain the presence of p58 and p56 in the virus preparations and also why both of these proteins reacted to anti-NβV antibodies. These results raised a very interesting question: could p58 represent the 37 nm capsid particles while p56 was found in the 34 nm particles? If this is case, it is interesting to note that there was a more pronounced reaction of p56 to anti-NβV antibodies compared to p58 (Figure 3.6 Panel B) suggesting that there is a larger amount of immunoreactive p56 compared to p58. However TEM analysis of the purified NβV sample shows unequal distribution of two particle sizes. The smaller particle, thought to be of a molecular mass of 56 kDa is not nearly as prevalent as the larger particles, thought to be of a molecular mass of 58 kDa. Perhaps the Western analysis of the purified virus was not indicative of the proportion of each of the two proteins but rather of the immunoreactivity to the particular antibodies used.

3.4. Analysis of viral genomic RNA

In order to confirm the identity of the virus extracted from the insects in 2003 and also to investigate any changes in the sequence of this isolate with the published sequence of NβV, viral RNA was extracted from CsCl fractions containing virus shown by Western analysis and analysed further.

3.4.1 Northern analysis

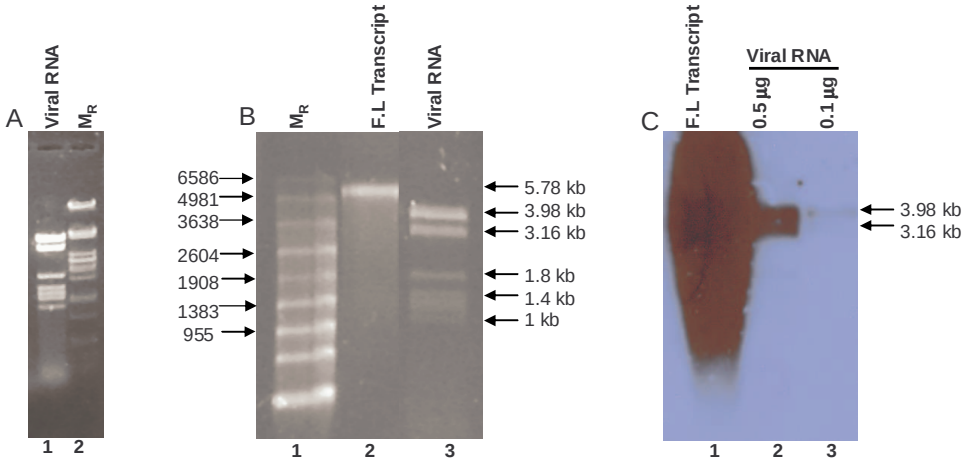


Figure 3.9. Northern analysis of extracted viral RNA. Panel A. Native agarose gel electrophoresis of purified RNA. **Lane 1:** extracted viral RNA, **Lane 2:** Phage lambda DNA digested with *Pst* 1. **Panel B**. Denaturing formaldehyde agarose gel of viral RNA. **Lane 1:** RNA molecular weight marker (Sigma) with sizes denoted in nucleotides **Lane 2:** RNA transcript of pGEM6625 and **Lane 3:** viral RNA. **Panel C**. Northern analysis of viral RNA using radiolabelled antisense transcript of NβV capsid gene. **Lane 1:** RNA transcript of pGEM6625. **Lanes 2-3:** varying concentrations of viral RNA denoted above in µg quantities. F.L. Transcript = full length transcript. Agarose gels stained with 0.01 µg/ml of ethidium bromide.

Initial analysis of extracted RNA by native gel eletrophoresis revealed the presence of at least 7 different species of RNA that were visualised when viral RNA was electrophoresed on a native agarose gel and the gel stained with ethidium bromide (Figure 3.9 Panel A, Lane 1). The sizes of these RNA species determined by denaturing gel electrophoresis, namely 3.98 kb, 3.16 kb, 1.8 kb and 1 kb were observed (Fig 3.9 Panel B, Lane 3). There did not appear to be an RNA band of the approximate size of the full length NβV RNA of 5.78 kb represented by an *in vitro* transcript generated from a cDNA clone of the NβV genome, pGEM6625 (Figure 3.9 Panel B, Lane 2).

To determine whether NβV was present in the sample, extracted RNA was separated by denaturing gel electrophoresis, transferred to nylon membranes and hybridized to an antisense radiolabelled transcript of the NβV capsid gene. Two of the RNA species (corresponding to molecular weights of 3.98 and 3.16 kb) hybridized to this particular transcript (Figure 3.9 Panel C, Lanes 2 and 3). Due to overloading of transcribed pGEM6625 RNA, a lower exposure of this blot (Figure 3.9 Panel C) was not available for interpretation. There appeared to be no band corresponding to the full length NβV genomic RNA at 5.7 kb. To confirm this result, RNA was extracted from a fresh virus preparation, subjected to denaturing gel electrophoresis, transferred to nylon membrane and hybridized to a radiolabelled antisense transcript corresponding to the NβV subgenomic RNA and NωV capsid gene RNA.

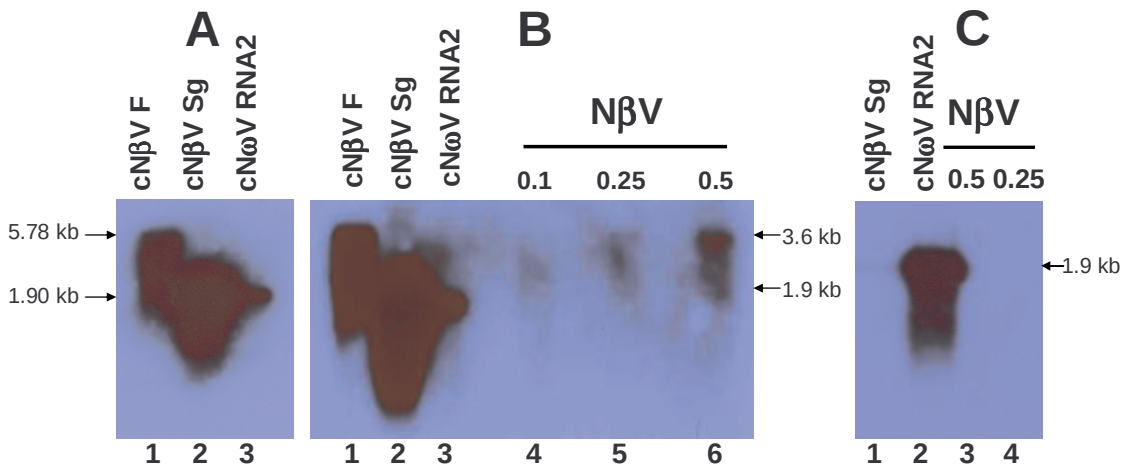


Figure 3.10. Hybridization of NβV and NωV to viral RNA. Panel A. Northern blot of viral RNA after 2 hour exposure probed with radiolabelled NβV capsid transcript. **Lanes 1-3:** RNA transcripts of pGEM6625, pCW4 (NβV) and pCW17R (NωV). **Panel B**. Northern blot of viral RNA after 4 hour exposure probed with radiolabelled NβV capsid transcript. **Lane 1-3:** as for Panel A, Lanes 4-6, various concentrations of viral RNA in μg quantities. **Panel C**. Northern blot of viral RNA after 2 hour exposure probed with radiolabelled NωV capsid transcript. **Lane 1:** full length NβV transcript derived from pCW4 and **Lane 2:** RNA transcript corresponding to NωV capsid coding sequence derived from pCW17R. **Lanes 3** and **4:** 0.5 and 0.25 μ g of viral RNA, respectively.

While the resolution of this gel was not as good as before, a band of approximately 3.6 kb was again observed (Figure 3.10 Panel B, Lanes 4-6) and there was no evidence of a full length genomic RNA. However, a new band co-migrating with a subgenomic transcript of 1.9 kb was observed (Figure 3.10 Panel B, Lane 6). Due to low resolution of bands (Figure 3.10 Panel A and B) and overexposure resulting from the overloading of transcribed plasmids, the sizes discussed were approximated.

Unexpectedly, the NβV probe detected the NωV control RNA transcript (Figure 3.10 Panel A and B, Lane 3). Since there is less than 20 % homology between the NβV and NωV capsid gene sequence, the possibility of cross-reaction was unlikely. It was possible that some of the NβV sample could have leaked into the NωV lane due to very large sample volumes. To determine whether this was the case, viral RNA was hybridized with a radiolabelled NωV capsid coding sequence antisense transcript. No signal was detected in the viral RNA samples (Figure 3.10 Panel C, Lanes 3 and 4 compared with lane 2). As expected, there was no hybridization between the NβV transcript and the NωV radiolabelled probe (Figure 3.10 Panel C, Lane 1), confirming that the signal detected in Figure 3.10, Panel A and B, Lane 3 was due to cross-contamination of the samples during loading of the gels.

3.4.2. RT-PCR and sequence analysis of NβV genomic RNA

In order to confirm the presence of NβV viral RNA in the extracted sample and to compare the sequence of this new isolate with the reported sequence of the NβV genome, analysis was performed on the viral RNA in order to be able to detect any truncated versions of the subgenomic RNA of NβV. Three primer combinations were used to create the cDNA derived from the genomic RNA. The RT-PCR strategy used to amplify sections of the NβV genome is summarised in Figure 3.11. The first primer combination (VCap F1 and VCap R1) hybridize to either end of the NβV capsid ORF and the amplification of

this sequence would result in the capsid gene ORF of NβV (1848 bp). The second set of primers (Rep 1 and Cap 2) bind to a region upstream of the capsid ORF located in the replicase gene of the genome and midway through the capsid gene. The amplification of this product would result in the replicase/capsid gene bridge allowing a study of any significant change in the start of the capsid coding sequence. A third set of primers CWF1 and CWR1 were initially used to amplify the small midsection of the capsid gene from cDNA derived from viral RNA.

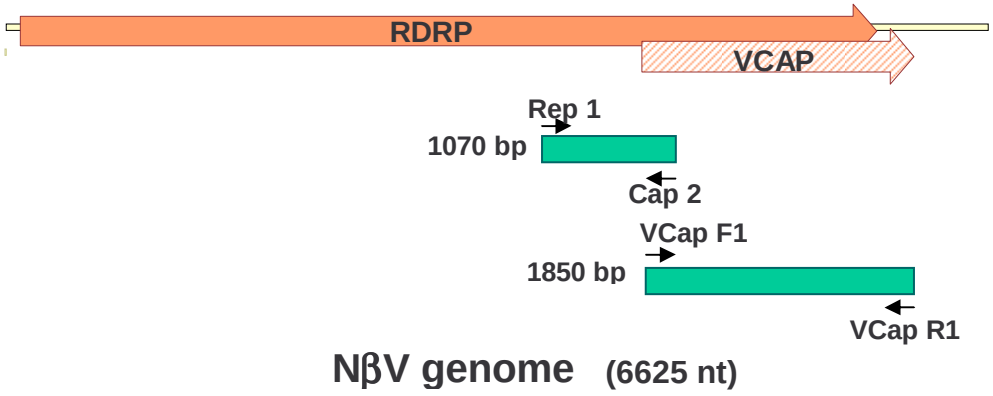


Figure 3.11. Schematic diagram of primer sets, regions of primer binding and molecular mass of amplified products.

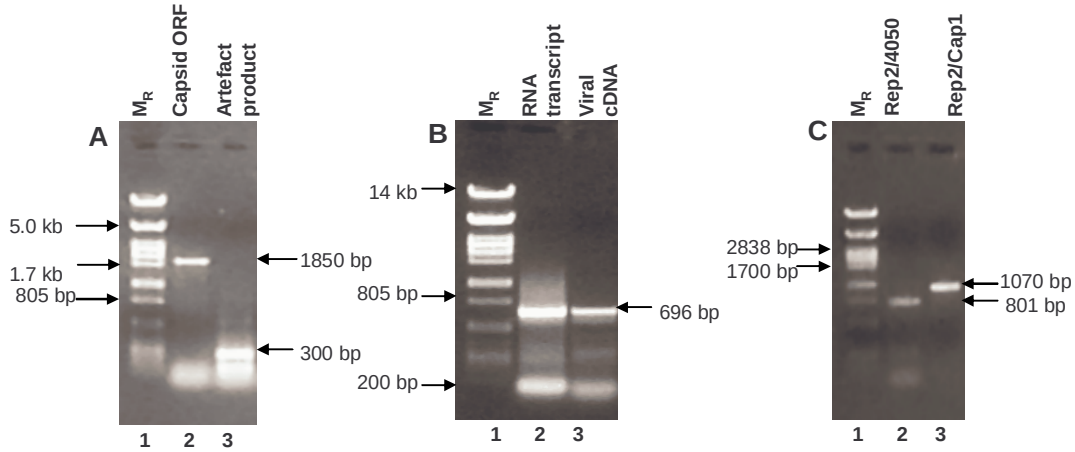


Figure 3.12. RT-PCR analysis of viral RNA. Panel A. Agarose gel electrophoresis of RT-PCR products using primer combinations VCap F1 and VCap R1. **Lane 1**: Phage Lambda DNA digested with *Pst* I. **Lane 2**: NβV capsid gene PCR product. **Lane 3**: Artefact product of 300 bp. **Panel B**. Agarose gel electrophoresis of RT-PCR products using internal capsid primers. **Lane 1**: Phage Lambda DNA digested with *Pst* I. **Lane 2**: RT-PCR product of NβV capsid gene *in vitro* transcript using the primers CWF1 and CWR1. **Lane 3**: RT-PCR product of viral RNA using primers CWR1 and CWF1. **Panel C**. Agarose gel electrophoresis of RT-PCR of viral RNA using the primers Rep2 and Cap1 or NBV4050. **Lane 1**: Phage Lambda DNA digested with *Pst* I. **Lane 2**: RT-PCR product using the primers Rep2 and NBV 4050.

Lane 3: RT-PCR using the primers Rep2 and Cap1. The agarose gels were stained with 0.01 μg/ml of ethidium bromide.

RT-PCR analysis of viral RNA using the primer combinations that amplify the entire capsid ORF of NβV (Figure 3.12 Panel A, Lane 2) as well as the amplification of the internal region of the capsid ORF (Figure 3.12 Panel B, Lane 3) yielded products of 1850 bp and 696 bp respectively. The amplification of the replicase/capsid gene bridge (Figure 3.12. Panel C. Lanes 2 and 3) yielded DNA products of 801bp and 1070 bp respectively. The shorter version of this product created by substituting the primer NβVCap2 with NBV 4050 (binds at the beginning of the capsid gene of NβV) creating a product of 801 bp (Figure 3.12 Panel C, Lane 2) but since this product was merely a shorter version containing no additional sequence, it was excluded from further sequence analysis.

RT-PCR analysis of the viral RNA using NωV capsid-specific primers did not produce an amplification product, showing that the larvae collected in 2003 were not infected with NωV (data not shown). The SDS-PAGE analysis of purified virus from larvae collected in 2002 and thought to be NωV yielded a single protein migrating of 54 kDa. This protein was shown to cross-react with anti-NωV antibodies and was absent from the purified virus purified from larvae in 2003 excluding the possibility that the insects collected in 2003 contained NωV.

3.4.3. Sequence analysis

A total of 2207 nucleotides out of 6625 was sequenced and analysed, representing 33 % of the entire genome of *Nudaurelia* β virus. The amplified cDNA products from the above experiments carrying sequences corresponding to nucleotides 3248-4317 and nucleotides 4039-5887 of the NβV genome were ligated into the pGEM-T Easy vector. Two recombinants containing the NβV capsid ORF (primers RDNBVCAPF1 and RDNBVCAPR) and a further two containing the overlapping sequence between the capsid and replicase gene (primers Rep 2 and Cap1) were obtained and designated

pCW9, pCW10 and pCW11, pCW12, respectively. All four of these constructs were sequenced and the results compared to the published sequence of the NβV genome. The alignments generated from this analysis are shown in Figures 3.13 and 3.14.

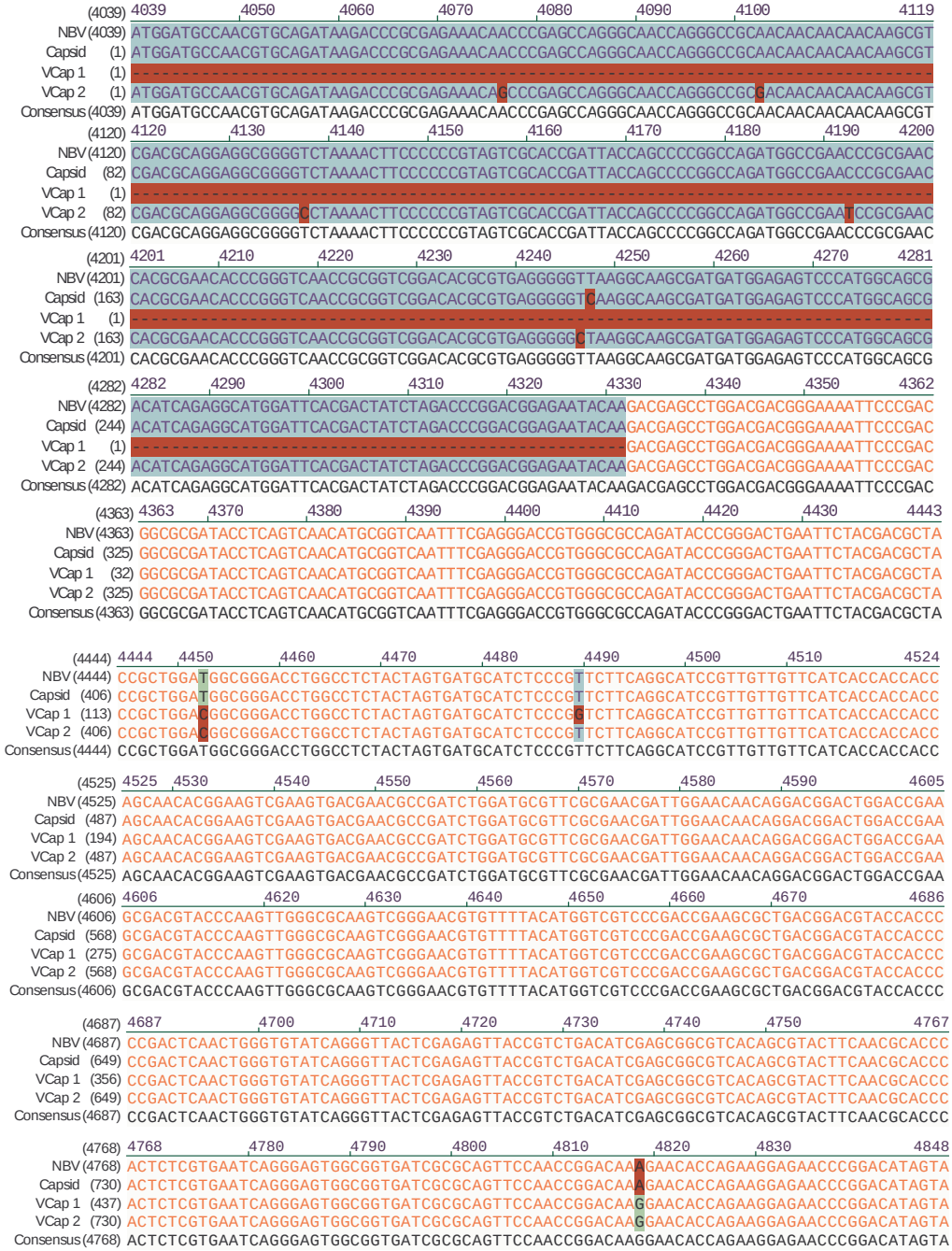


Figure 3.13. Alignment of nucleotide sequence of capsid gene from pCW9 and pCW10 (VCap1, 3 and 2,4 respectively) with the published sequence of NβV and pCW4. Confirmed mutations are highlighted in green. Regions highlighted in blue indicate conserved sequence and regions highlighted in red indicate 100 % conservation between all four sequences.

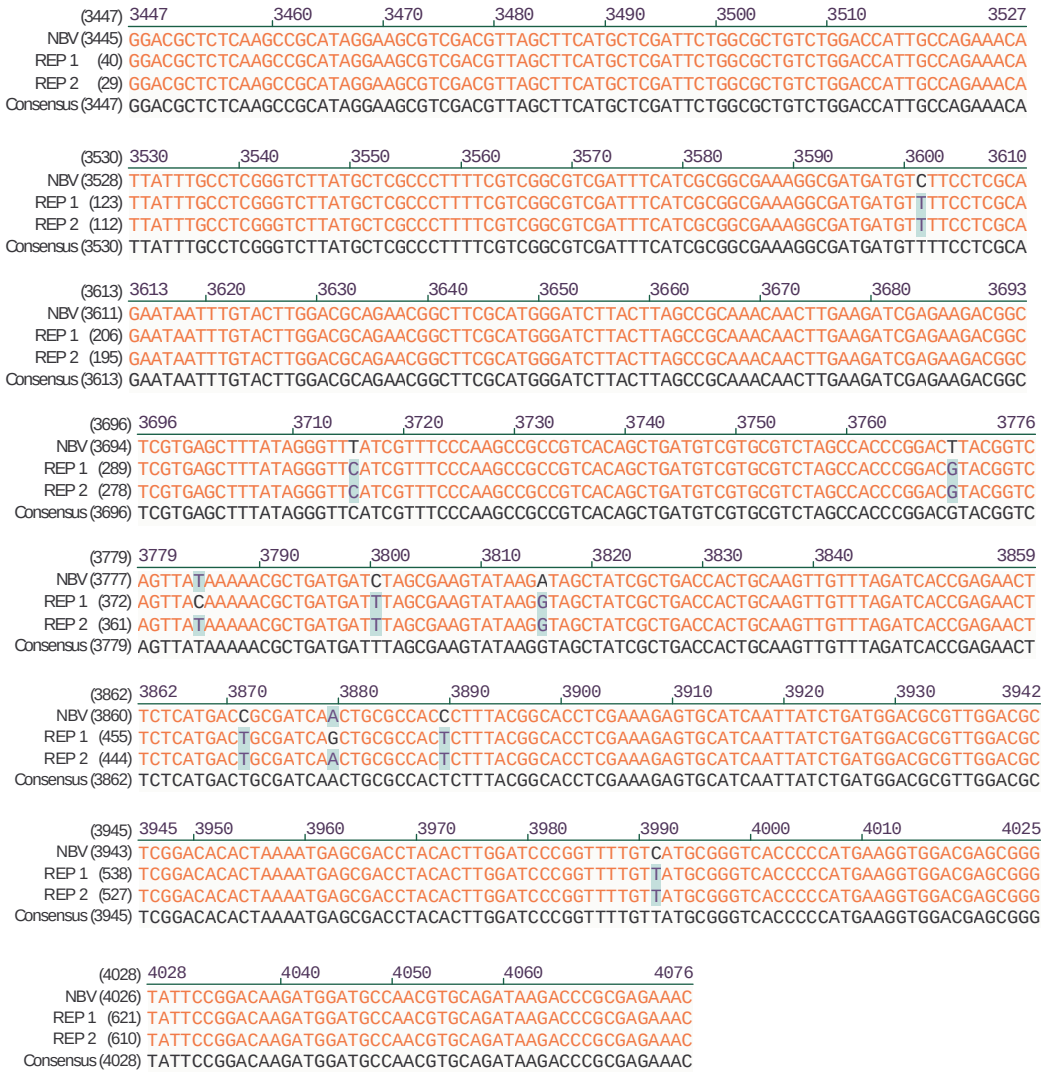


Figure 3.14. Alignment of nucleotide sequence of replicase regions from pCW11 and pCW12 (Rep 1 and 2 respectively) with the published sequence of NβV. Confirmed mutations are highlighted in blue.

The sequences obtained were analysed and separated according to the two genes of the NβV genome. The capsid gene sequence obtained from constructs pCW9 and pCW10 was almost identical to the published sequence of the capsid gene. The sequence exhibited a 99.46 % identity and contained 10 mutations when compared to the published sequence of the NβV capsid gene. These mutations were present in both of the recombinants containing the capsid gene (pCW9 and 10). 13 % of the replicase gene was present in the cDNA amplified from the viral genome, so a general comparison to the published sequence was not possible.

It was however observed that the portion of the replicase gene that was sequenced contained 10 mutations, 8 of which were present in both recombinants (pCW11 and pCW12). When the nucleotide sequence of the NβV capsid gene extracted from the viral RNA was converted to an amino acid residue sequence, it was noted that the sequence had a 99.51 % identity to the published amino acid sequence. A summary of these substitutions can be seen in Table 3.1 below. Three potential amino acid residue substitutions were observed with two out of three of these substitutions in both of the constructs sequenced. These two mutations were found to occur within the β-sandwich domain of the protein (Figure 3.15).

Table 3.1. Summary of amino acid changes found in the capsid gene sequence of two cDNA clones compared to the published sequence.

Position	**Amino acid substitution**			**Trend**
	Published sequence	**pCW9**	**pCW10**	
151	Phe	Cys	Phe	Phe
419	**Phe**	**Ser**	**Ser**	**Ser**
489	**Phe**	**Leu**	**Leu**	**Leu**

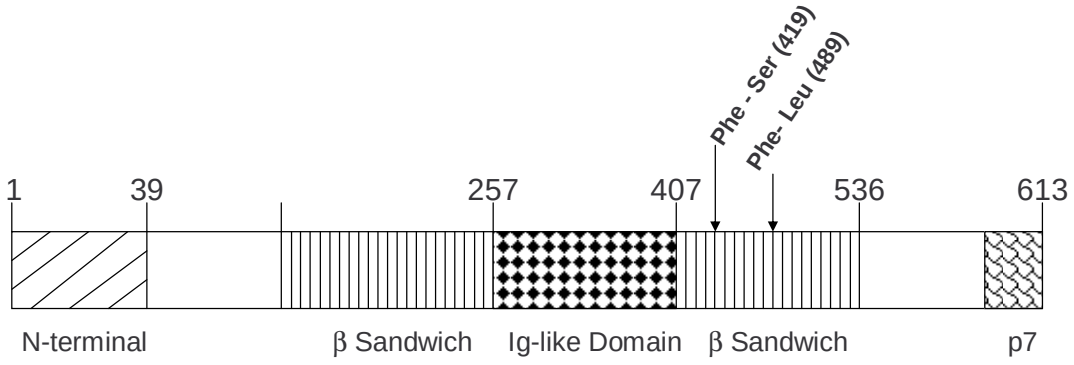


Figure 3.15. Identification of amino acid substitutions within particular protein domains of the NβV capsid protein. The numbers at the start of each domain reflect the amino acid number of the capsid protein.

The authenticity of the substitutions found in the sequence will only be substantiated once more cDNA constructs are sequenced and compared, with each other and with the published sequence of the NβV genome. The

mutations documented could have occurred during the amplification of the cDNA or are legitimate changes in the sequence, which is why further analysis of cDNA obtained from the above experiments must be performed. The two amino acid substitutions documented in both of the capsid cDNA constructs occurred within the β sandwich domain of the capsid protein and one of these substitutions was not conservative. The substitution of phenylalanine (amino acid residue number 419) to serine represents a change from a non-polar hydrophobic to a polar hydrophilic side chain.

It was also discovered that one amino acid change were present in the portion of the replicase gene that was sequenced by both clones - amino acid number 1292 resulting in an substitution of an isoleucine to a valine in the RdRp I domain (Gorbaleyna *et al*., 2002). Another amino acid change was found in pCW11 (amino acid number 1262) resulting in a substitution of an asparagine to a serine. Since only two cDNA clones containing a fragment of the replicase gene were sequenced, more recombinants containing the entire gene will have to sequenced and analysed. The data obtained from the sequence analysis of these cDNA clones is summarized in Table 3.2.

Table 3.2. Overall findings from the sequence analysis of four cDNA clones derived from viral RNA.

	Capsid gene	Replicase gene
Nucleotide changes	10	10
Amino acid changes	3	2
% genome sequenced	100 %	13.8 %
% nucleotide identity	99.46 % (10/1839 bp)	/
% amino acid identity	99.51 %	/
Total bp sequenced (out of 6625)	1839	813 (excl. overlap with capsid gene

The RNA used to generate the cDNA and later the recombinants pCW9-12 originated from a single virus preparation. It has been observed that each preparation of virus differed in terms of the amount of virus present and also the presence of other proteins. In summary, the sequence obtained from the viral RNA of a new isolate of NβV is almost identical the sequence posted in

1999 and derived from cDNA acquired from insects collected in 1985 (Gordon *et al*., 1999). More RT-PCR analysis of different sources of RNA should be performed in future since these results may not necessarily reflect the entire population of viruses infecting the larvae.

Close inspection of the NβV genome sequence revealed a second in-frame methionine residue of the NβV capsid coding sequence. Earlier data had shown the presence of two NβV capsid proteins, p56 and p58. This led to the speculation that the second methionine residue of the capsid coding sequence is also a second translational start for the smaller capsid protein of the wild type virus. The identification of a similar second translational start in the Nodavirus, *Pariacoto virus* was though to initiate the translation of a minor capsid protein 24 amino acids upstream of first translational start of the proposed capsid coding sequence. This truncated form of the capsid coat protein was shown to undergo maturational cleavage and assemble into virus particles (Johnson and Ball, 2003). This observation supports the above-mentioned hypothesis. An alignment of the N-terminal amino acid residues of the capsid protein of NωV, HaSV, NβV, FHV and PaV revealed that the second methionine residues of both PaV and NβV occur in structurally equivalent positions (Figure 3.16).

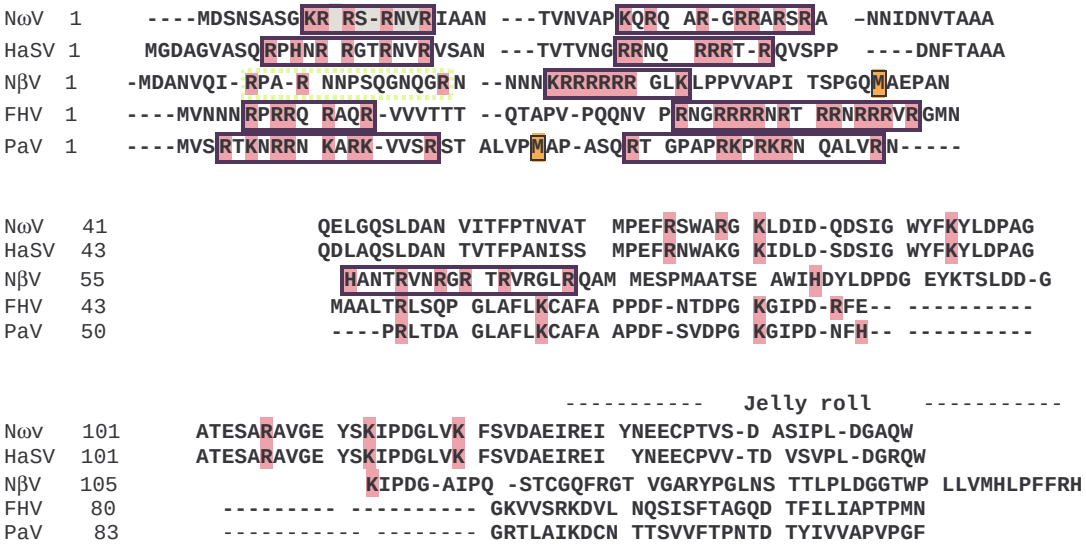


Figure 3.16. Alignment of the N-terminal amino acid residues of tetraviruses and nodaviruses. The second methionine residues speculated as alternative start codons are blocked in orange. PaV: *Pariacoto virus*, NβV: *Nudaurelia β virus*, NωV: *Nudaurelia ω virus*, HaSV: *Helicoverpa armigera stunt virus*, FHV: *Flock House Virus*.

The alignment of the N-terminal amino acid residues of theses proteins also showed two distinct regions of positively charged amino acid residues, such as the amino acid residue arginine, highlighted in purple and the regions blocked (Figure 3.16). It was also interesting to note that both of the second methionine residues of PaV and NβV are situated between these regions. The predicted molecular weight of a capsid protein of NβV translated from the second methionine residue is 53 kDa. This is surprising since the SDS-PAGE analysis of the NβV revealed the presence of what is speculated to be two viral capsid proteins of a molecular weight of 56 and 58 kDa. The discrepancy in mass of 3 kDa of the smaller capsid protein could be resolved by the N-terminal sequencing of the protein in order to accurately asses the size of this protein. In order to substantiate this hypothesis; the full length capsid coding sequence and a truncated (by 50 amino acid residues) capsid coding sequence of NβV were expressed using recombinant baculovirus vectors in insect cells.

3.5 Baculoviral expression of NβV VLPs.

The reasons for including a truncated (by 50 amino acid residues from the N-terminus of the protein) form of the capsid gene were to firstly test the hypothesis that NβV translates its capsid protein from the second methionine residue and secondly to see whether any VLPs produced are smaller in size compared to the wild type particles and if the size of these particles is similar to the smaller particles isolated from diseased larvae.

As an expression vector, pFASTBAC Dual offers the advantage of containing two bi-directional promoters (polyhedrin and p10) from the virus *Autographa californica* nuclear polyhedrosis virus (AcNPV) which are used to produce large quantities of virus protein within an insect cell line expression system. This vector was used in the expression of NβV VLPs and transposed into baculovirus DNA as described below in Secion 3.5.1.

3.5.1. Production and confirmation of recombinant baculovirus vectors.

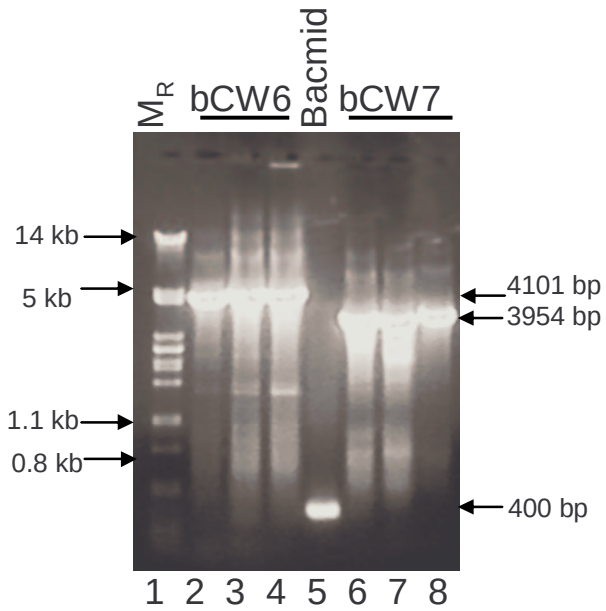


Figure 3.17. PCR analysis of recombinant baculovirus vector. **Lane 1**. Phage lambda DNA digested with *Pst* I. **Lanes 2-4**: PCR product of baculovirus DNA containing bCW6. **Lane 5**: PCR product of unmodified baculovirus DNA. **Lanes 6-8**: PCR product of baculovirus DNA containing bCW7.

The amplification of both full length and truncated versions of NβV capsid ORFs were confirmed by agarose gel electrophoresis and the products ligated into the pGEM-T Easy vector and named pCW4 and 5. The constructs pCW6 and 7 were created by the ligation of the capsid gene from pCW4 and 5 into the vector, pFASTBAC Dual (Figure 2.2). The construct pCW6 represents the full length capsid coding sequence of the capsid gene of NβV and pCW7, the truncated capsid coding sequence. These constructs were used to create recombinant baculovirus DNA in *E.coli* DH10BAC competent cells which contain the bacmid (bMON14272) and a helper plasmid (pMON7124) (Invitrogen). The baculovirus DNA was extracted from these cells and the DNA amplified by PCR using universal primers, pUCF and pUCR (Appendix A). The primers bind at sites flanking the recombinant DNA containing integrated bCW6 and also bCW 7. The sizes that were expected from this amplification reaction were calculated using the information summarised in Figure 2.2.

Bacmid DNA containing the confirmed insert was transfected into confluent insect cell lines the infectious dose was calculated from high titre stocks of the third virus passage using the $TCID_{50}$ (Tissue culture infectious dose assay) and the MOI (multiplicity of infection). The results and calculations of the $TCID_{50}$ assay are summarized below in Table 3.3.

Table 3.3. Infectivity of bCW6 and bCW7 scored using the $TCID_{50}$ Assay. + indicates a positive cytopathic score and – a negative cytopathic score.

Dilution	bCW 6					bCW7				
Dilution	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}
No. “+”	8	8	6	5	4	8	8	6	5	4
No. “-	0	0	2	3	4	0	0	2	3	4
% “+”	100	100	75	62.5	50	100	100	75	62.5	50

Proportionate distance (P.D) of a 50 % response from the response above 50 % is calculated using the following formula:

PD = (A-50)/(A-B),where A is the % response above 50 % and B is the % response below 50 %.

PD = [(% above 50%)-50%]/[(% above 50%)-(%below 50%)]

= (75-50)/(75-50)

= 25/25
= 1

Thus, log $TCID_{50}$= -10 -1 = -11 and equivalent to the $TCID_{50} = 10^{-11}$.
Therefore titre is equal to the reciprocal of this:
1 x 10^{11} $TCID_{50}$/50 µl.

This is equivalent to 2 X 10^{12} pfu/ml.

In order to calculate the viral inoculum required for infections at a multiplicity of infection of 5 to 10, the following formula was used as outlined in the Bac to Bac Manual, GibcoBRl (Invitirogen™ Life Technologies):

$$\text{Inoculum required (ml)} = \frac{\text{desired MOI (pfu/cell) X total number of cells}}{\text{titre of viral inoculum (pfu/cell)}}$$

$$= \frac{5 \times (3X10^{9})}{1 \times 10^{12}}$$

$$= \frac{15 \times 10^{9}}{2 \times 10^{12}}$$

= 0.0075 ml = 7.5 µl.

3.5.2. Production of high titre virus stocks and isolation of VLPs.

Confluent monolayers of *Sf*9 cells were infected with a multiplicity of infection (MOI) of 5 for the expression of the recombinant baculovirus DNA, bCW6 and bCW7. VLPs were purified from the monolayer of cells and supernatant after 10 days. The VLPs were harvested in acidic conditions to ensure maturation (Taylor *et al*., 2002), centrifuged through two sucrose cushions and finally separated out using sucrose gradient ultracentrifugation. Fractions were collected from each sucrose gradient and analysed spectrophotometrically, measuring the absorbance readings at 260 and 280 nm.

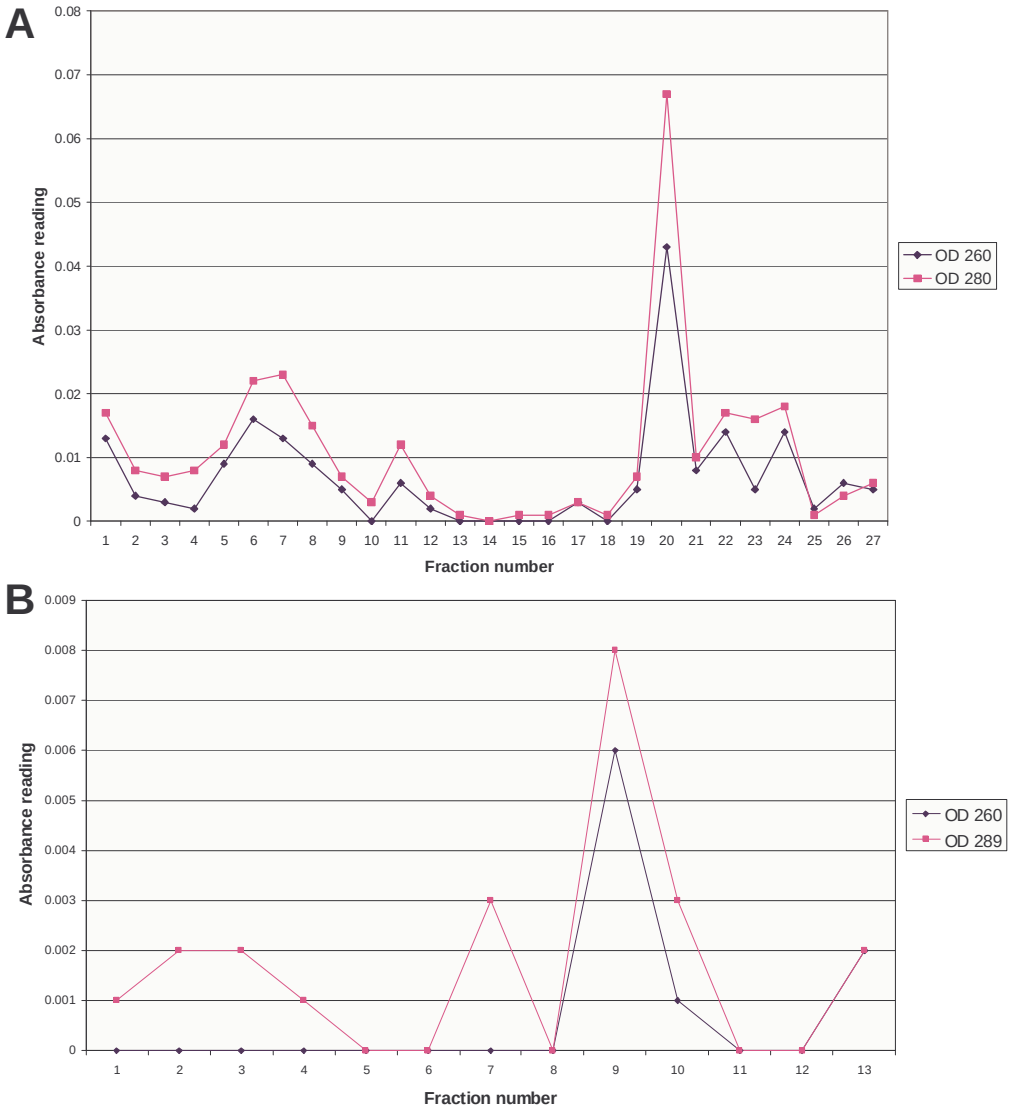


Figure 3.18. Spectrophotometric analysis of fractions containing VLPs extracted from insect cell lines expressing bCW7 (Panel A) and bCW6 (Panel B) and centrifuged through a 10- 40% sucrose gradient. Note: A 1 in 20 dilution was made of each fraction and the absorbance reading on this graph reflects that dilution. The direction of the gradient from top to bottom is represented in the Panels A and B from left to right across each graph.

The spectrophotometric analysis of fractions of VLPs produced from the recombinant vector containing the full length NβV capsid gene (bCW6) and the truncated NβV capsid gene (bCW7) collected from a sucrose gradient revealed a large peak in fractions 9-11 and fractions19-21, suggesting the presence of both nucleic acid and protein (Figure 3.18). These three fractions were analysed by SDS-PAGE, Western and TEM analysis. Western analysis was performed on samples from fractions 19-21 (pCW7) and 9-11 (pCW6) to determine the presence of NβV VLPs.

3.5.3. Western analysis of purified VLPs.

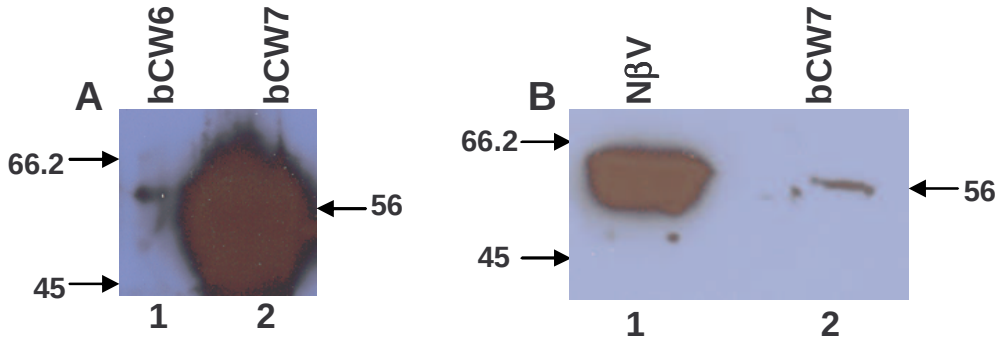


Figure 3.19. Western analysis of expressed VLPs. Panel A. Western blot of 10 %SDS-PAGE gel probed with anti-NβV + NεV antibodies. **Lane 1:** VLPs purified from expressed bCW6. **Lane 2:** VLPs purified from expressed bCW7. **Panel B**. Western blot of 10 % SDS-PAGE gel probed with anti-NβV + NεV antibodies. **Lane 1:** NβV. **Lane 2:** VLPs purified from expressed bCW7. Arrows indicate molecular weight.

Anti-NβV+NεV antibodies were used in this set of experiments due to diminished stocks of anti-NβV antibodies. Although these antibodies were raised against both *Nudaurelia* β virus and *Nudaurelia* ε virus, they were used in the Western analysis of expressed VLPs since little cross-reactivity with other cellular proteins and the absence of other viruses was expected. It was consistently found that bCW6 produced very low amounts of immature VLPs (Figure 3.19 Panel A, Lane 1). A protein of approximately 61 kDa in size reacted to the anti-NβV+NεV antibodies which indicated that the particles formed were not mature capsid particles. The expression of bCW7, containing the truncated capsid coding sequence of NβV yielded a protein which migrated at a molecular weight of 56 kDa when seized according to the molecular weight marker of the Western blot (Figure 3.19 Panel A, Lane 2).

3.5.4. Isolation and analysis of recombinant baculovirus DNA from high titre virus stocks.

No VLPs could be detected in sucrose gradient fractions derived from *Sf*9 cells expressing the full length NβV VCap construct. This was despite the presence of a protein and nucleic acid peak in fractions 9-11 of a completed sucrose gradient (Figure 3.18 Panel B). To determine whether this was due to problems with the recombinant vector construct, baculovirus DNA extracted from the supernatant of high titre socks of both bCW6 and bCW7 were

analysed by PCR. The first PCR reaction amplified the 5` end of each capsid coding sequence and the second amplified the promoter and 5` end of VCap. The primers used in the first amplification reaction were RDNBVCAP F1 (for pCW6), RDNBVCAP F2 (for bCW7) and CWR1 as a reverse primer. The amplified product of the 5` end of each capsid coding sequence was expected to migrate at a molecular weight of 1.1 kb (bCW7) and 1.3 kb (bCW6). These bands were observed as expected (Figure 3.21 Panel A, Lanes 1 and 2).

The second amplification reaction for both expressed constructs used the primer CWF3, which binds to the polyhedrin promoter of the expression vector and the reverse primer, CWR1 which binds at nucleotide 5269-5293 of the NβV genome. The PCR products containing the polyhedrin promoter and amino termini of the NβV capsid was expected to migrate at a molecular mass of 1279 bp (bCW7) and 1429 bp (bCW6) (Figure 3.21 Panel B, Lanes 2 and 3). A summary of the PCR reactions is shown in Figure 3.20.

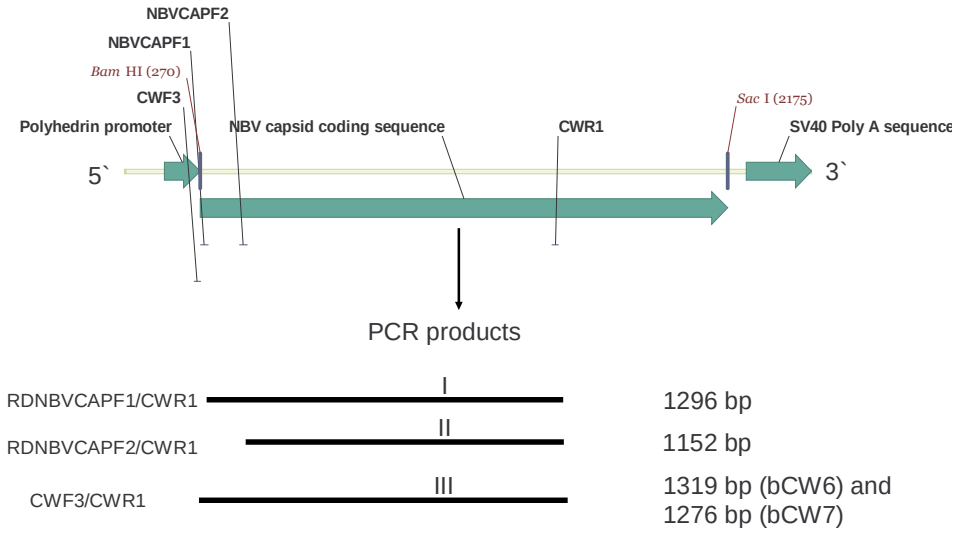


Figure 3.20. Schematic diagram of PCR reactions on recombinant baculovirus DNA. The reactions are numbered I-III and are marked as such above the relevant lane in Figure 3.21

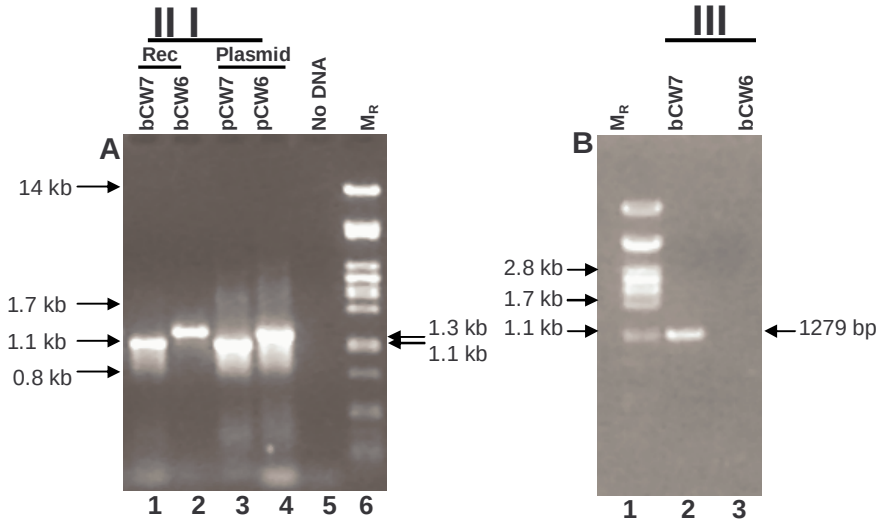


Figure 3.21. Confirmation of inserted genes in baculovirus expression vector. **Panel A**. PCR analysis of extracted baculovirus DNA to amplify the first half of the capsid gene from the expression vector. **Lanes 1 and 2**: PCR product of extracted DNA containing capsid genes of bCW7 and bCW6, respectively. **Lanes 3 and 4**: PCR product from the amplification of capsid genes from the constructs pCW7 and pCW6, respectively. **Lane 5**: No DNA PCR reaction. **Lane 6:** Phage lambda DNA digested with *Pst* I. **Panel B**. PCR analysis of extracted baculovirus DNA to amplify the promoter/capsid gene bridge. **Lane 1**: Phage lambda DNA digested with *Pst* I. **Lane 2**: PCR product of DNA from the expression vector containing the capsid gene from bCW7. **Lane 3**: PCR product of DNA from the expression vector containing the capsid gene from bCW6. Agarose gels were stained with 0.01 μg/ml ethidium bromide.

The PCR analysis of recombinant baculovirus DNA containing the N-terminal region of the capsid genes of bCW6 and bCW7 yielded results that indicated that the amplification of the first half of the capsid coding sequence from both vectors was successful (Figure 3.21 Panel A, Lanes 1 and 2).

The subsequent amplification of the promoter/capsid gene bridge revealed that although this section of the vector was still intact in the vector bCW7 (Figure 3.21 Panel B, Lane 2), this was not the case for the construct bCW6 (Figure 3.21 Panel B, Lane 3). It was suggested that although the full length capsid coding sequence was present in bCW6, the sequence corresponding to the primer CWF3 was not. This would imply that the polyhedrin promoter had been rearranged and not merely a point mutation since no primer binding was thought to have taken place explaining the lack on any amplification product. This result explained the lack of protein expression from the vector bCW6 and lack of detectable VLPs.

3.5.5. TEM analysis of VLPs derived from *Sf*9 cells lines expressing recombinant baculoviruses containing integrated pCW7.

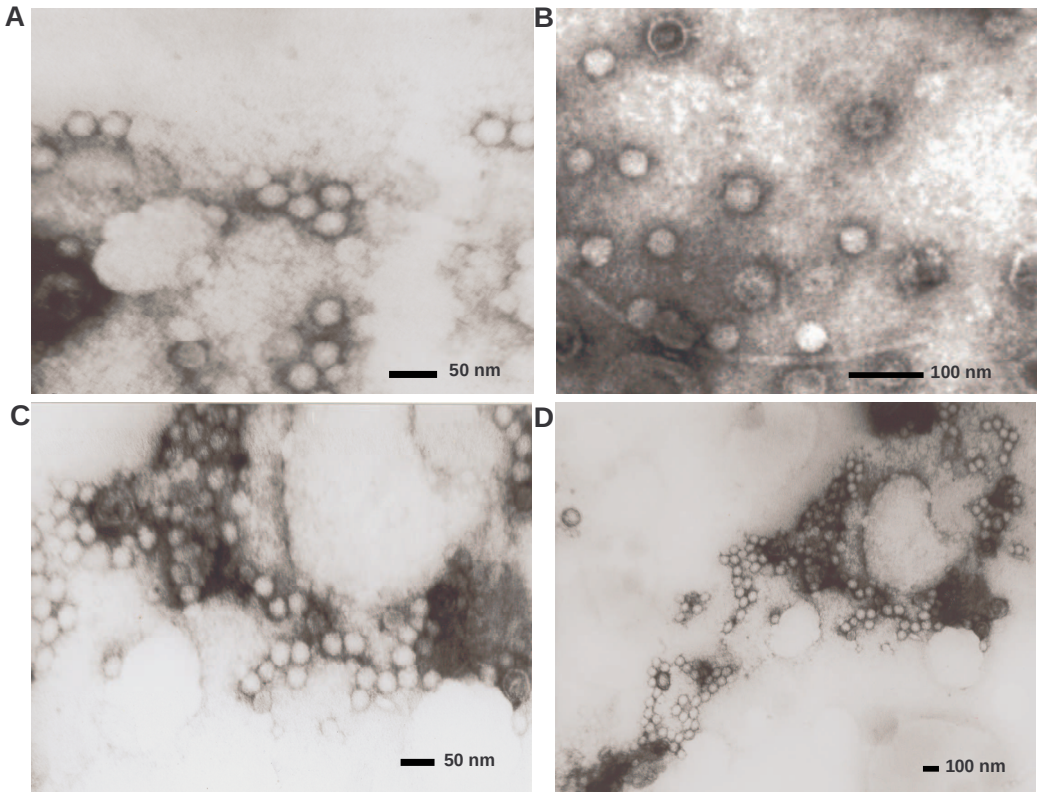


Figure 3.22. Transmission electron microscope images of VLPs after extraction in acidic environment. **Panel A, C and D**: VLPs from bCW7 extracted in an acidic environment and imaged at various magnifications. **Panel B.** NβV purified from infected *N. capensis* larvae.

VLPs were observed in the samples prepared from recombinant baculovirus-infected *Sf*9 cells which had a similar morphology to the wild-type virus particles (Figure 3.22 Panel B). The pits and characteristic triangular arrangement on the face of the particles on the VLPs were the same as those on wild type NβV particles (Figure 3.21, Panel B). The particles were on average smaller than expected, with an average diameter of 33.33 nm based on the measurement of 20 VLPs. This was 4 nm smaller than the purified wild type particles that were measured and an average of 37.10 nm calculated. More of the VLPs will have to be measured and analysed in future in order to compare the distribution in diameters as done for wild-type NβV. The diameter of the VLPs produced is the same size as the population of smaller wild type particles that were observed in the wild type NβV population (Section 3.3). The ability of the truncated capsid coding sequence to

assemble into VLPs with an average diameter of 33 nm supports the hypothesis that the two types of virus particles observed in the wild type population may originate from two translation products of the same NβV capsid gene.

Chapter 4

CONCLUSIONS AND FUTURE PROSPECTS

The discovery of a large population of diseased *N. capensis* in a pine plantation in the Western Cape in 2002 raised the question of whether or not the high mortality rate of larvae was due to a tetravirus infection. The research described in this dissertation entails the isolation of a tetravirus from diseased larvae that was identified as NβV using immunogical and recombinant DNA techniques as well as electron microscopy. In addition, it was shown that the expression of the NβV capsid coding sequence truncated by 50 amino acids at its N-terminal end was able to assemble into VLPs**.**

4.1 Isolation of wild type NβV from diseased *N. capensis* larvae.

NβV was first purified in 1968 by Hendry *et al*. The virus was found to be the dominant virus in a mixture of a least 4 other viruses in diseased *N. capensis* larvae (Juckes, 1974). More than 10 years later, it was found that NωV had become the dominant virus and NβV was last purified at this time (Hendry *et al*., 1985). After this time, NβV was thought to have become extinct. In 2003, the re-isolation of NβV has provided new opportunities for research such as virus localization on infected larvae, interaction of this virus with its host, storage of virus stocks and generation of more antibodies towards this new isolate. The characterization of other viruses such as the ω-like virus extracted from insects in 2002 and considering the fact that the creation of permissive cell lines for the infection by these viruses has yet to be successful, the re-isolation of NβV has opened up many new areas of research towards the *Tetraviridae*. The isolation of NωV and what has now been elucidated as most probably an ω-like tetravirus in 2002 has also contributed towards the study of the omegatetraviruses.

During the preparation of tetraviruses from diseased larvae, it was discovered that denser CsCl fractions with a buoyant density of 1.330 g/cm^3 found in the first few virus preparations contained a protein of 40 kDa in size. It is speculated that this protein is the capsid protein of a nodavirus since this fits the range of the size of a nodavirus capsid protein – between 38 and 40 kDa (van Regenmortel *et al*., 2000). This theory is substantiated by the position of p40 in the latter fractions after density gradient ultracentrifugation. Nodaviruses have an average buoyant density between 1.3 and 1.34 g/cm^3 which is greater than all tetraviruses (van Regenmortel *et al*., 1996) and supports the data generated. This assumption is again substantiated by the fact that Nodaviruses have on average two positive sense strands of RNA of approximately 3.1 kb and 1.4 kb (van Regenmortel *et al*., 2000) supporting the hypothesis that more than one species of viral RNA was extracted from the preparations described in this work.

It can be speculated that p40 might have originated from bacteriophages that were co-purified from batches of particularly decomposed larvae. The presence of bacteriophages in the first three virus preparations was confirmed by TEM analysis of CsCl fractions (data not shown). The 40 kDa protein was consistently found in later fractions of a CsCl gradient after the removal of tetraviruses suggesting that this protein originated from a virus with a higher buoyant density than tetraviruses. However, this is highly unlikely since bacteriophages are mostly composed of more than one protein (van Regenmortel *et al*., 2000).

4.2. Molecular characterization of the 2003 NβV isolate.

SDS-PAGE of fractions collected from both CsCl gradients and sucrose gradients consistently showed the presence of two proteins of 56 and 58 kDa, respectively, which could not be separated sufficiently to obtain fractions

containing separate fractions containing one or the other proteins. It was also established that both p56 and p58 cross-react to anti-NβV antibodies. These results are similar to those reported by Johnson and Ball (2003) whereby it was established that PaV contains a second methionine situated at amino acid residue number 25 at the amino terminus. Upon mutation of this amino acid residue, it was found that the second smaller capsid protein isolated from *G. mellonella* larvae disappeared, suggesting that this protein was produced from a second translational start (Johnson and Ball, 2003). It could also be possible that two β-like tetraviruses existed within the virus preparations. A 50 amino acid truncated capsid coding sequence of NβV was created and expressed in a recombinant baculovirus vector. When VLPs were harvested from *Sf*9 cells expressing a recombinant baculovirus vector containing the coding sequence, it was observed that a capsid protein was produced that assembled into VLPs.

TEM studies of virus preparations containing both p58 and p56 showed that all of the particles visualized were comparable to NβV in appearance. The particles viewed under the microscope exhibited the characteristic *T*=4 symmetry, triangular pitted faces and deep grooves between the faces as described previously for samples of NβV. This observation does not rule out the possibility of a second tetravirus. One thousand particles were imaged and the diameter of each particle measured. An average of 37.10 nm was measured as was expected for NβV yet when the diameters of these particles were plotted according to frequency of occurrence it was discovered that two distinct populations of NβV were present in the sample: those with a diameter of 37 nm and another smaller group with a diameter of 34 nm.

These data suggested the possibility that another tetravirus which is closely related and is morphologically indistinguishable from NβV exists in the sample. An alternative explanation is, as shown for the nodavirus PaV (Johnson and Ball, 2003), that the current *N. capensis* virus population contains two variants of same capsid protein. Baculovirus expression studies

showed that a truncated capsid protein does, in fact, assemble into VLPs as has been shown with the baculoviral expression of a truncated version of NβV capsid. The identity of p58 and p56 could be confirmed the N-terminal sequencing of the proteins, which would establish the nature of the N-terminal start of the capsid protein and whether p56 is the result of translation from the second methionine implying a leaky scanning mechanism or a cap-dependent ribosome as can be the case with PaV (Johnson and Ball, 2003) or the presence of an internal ribosome entry site (IRES) at the second methionine. A less likely alternative is that there might even be a cleavage even occurring during capsid assembly discovered for the Betatetravirus, TaV (Pringle *et al*., 2001).

A problem encountered during the Western analysis of both wild type and VLPs was the lack of a convincing existing sample of *Nudaurelia ωvirus*. The physical appearance of the sample used in the Western and SDS-PAGE analyses in this work when analyzed by TEM identified with the description for NωV in terms of size and outer capsid subunit arrangement but the size of the protein of this sample when analyzed by SDS-PAGE was much smaller than the reported molecular weight of the capsid protein (53-54 kDa) compared to the reported size of 62 kDa in this work. When this purified virus was analyzed using Western blotting techniques, the sample reacted to anti-NωV antibodies. TEM analysis of this same preparation of what was believed to be NωV revealed particles that exhibited a morphology and average diameter the same as that reported for NωV. RT-PCR on RNA extracted from this NωV sample using primers corresponding to NωV capsid sequence was unsuccessful, suggesting some sequence divergence (data not shown). This RNA should be re-analyzed using random hexamer primers in order to analyze the RNA from this virus without bias towards one particular virus. It would be very interesting to ascertain if the *N. capensis* larvae collected in 2002 were infected with another new omegatetravirus.

The analysis of the NβV genomic RNA revealed at least 7 different species. Three of these species with sizes of 3.98, 3.16 and 1.9 kb reacted to a radiolabelled transcript of the capsid gene of NβV. These species were detected by a probe corresponding to the capsid coding sequence of NβV. The reason for the two species of RNA with a molecular weight 3.98 and 3.18 kb hybridizing to the probe remains unclear and one could also speculate that the replicase gene had degraded. This is not thought to be the case since there were two distinct bands and not a smear which would indicate nuclease activity.

The results from limited sequence analysis obtained showed the current NβV isolate had an almost identical nucleotide and amino acid sequence to the original virus isolated in 1985 from which the published sequence of Gordon *et al*. (1999) was derived. The amino acid changes that were documented all appeared in the β-sandwich domains of the capsid protein. These two substitutions were characterized by a change of a phenylalanine to a serine which represents a change of charge of the side group from a non-polar hydrophobic charge to a polar hydrophilic charge and a phenylalanine to a leucine where both side chains are non-polar hydrophobic.

An accurate assessment of sequence divergence between the 1985 and 2003 isolates can only be made with considerably more analysis of the whole genome sequence. Experiments include cDNA synthesis and RT-PCR analysis using oligonucleotides specific for the truncated form of the capsid coding sequence thought to exist within the mixture of RNA isolated from virus particles.

From the above-discussed data, one should consider that the virus identified as NβV might be a ‘quasispecies’ the term used to describe the existence of a discrete level of variation within a population of viruses that share a common origin (NβV) but have distinct genomic sequences as a result of mutation, drift

and the impact of selection (Smith *et al*., 1997). The term has been used to describe heterogenous populations of human immunodeficiency virus (HIV), picornaviruses, birnaviruses as well as many other families (Smith *et al*., 1997). Quasispecies are thought to occur due to the result of the error-prone replication of virus genomes and several features of RNA viruses (antigenic variation, genotypic diversity, immune escape) have been attributed to the lack of proof reading RNA-dependent RNA-polymerases (Smith *et al*., 1997).

Experiments must also be designed to determine if the virus analyzed in this work is indeed a quasispecies or due to natural variation. It can also not be dismissed that this virus isolated might in fact be a newly isolated Betatetravirus. This is due to evidence found in this work that there a number of mutations within the replicase gene of the isolate studied. It is widely known that viral replicase genes along with a few other replicative proteins have been used to identify and classify RNA viruses due to their universal occurrence and more importantly, their exceptional conservation (Gorbalenya, *et al*., 2002). This statement is supported by the distinctive sub-population of virus particles with a smaller than expected diameter and the lack of an identifiable replicase gene found to hybridize to radiolabelled transcribed probes of NβV.

4.3. Assembly of *Nudaurelia β virus* truncated capsid protein into VLPs.

At the start of this research project, it was widely believed that NβV had become extinct. In the absence of an infectious virus, the original aim of this project was to develop experimental systems to study the biology of NβV. Thus the first objective was to construct a recombinant baculovirus expression the NβV capsid coding sequence. Subsequent to the discovery of wild type virus in diseased *N. capensis* larvae, and the observation of the presence of p56 and a population of 34 nm particles, the focus of the baculovirus expression experiment shifted to attempting to explain these observations.

While problems were encountered with the full length capsid coding sequence recombinant vector, bCW6, the successful expression of the truncated capsid coding sequence recombinant vector, bCW7 and subsequent assembly of the protein into VLPS provided some support for the hypothesis that NβV utilizes a second methionine residue as a translational start.

The VLPs that were assembled from the expression of protein from a recombinant baculovirus DNA containing the integrated truncated capsid coding sequence appeared to be mature, based on the protein reacting to anti-NβV+NεV antibodies with a molecular weight of 56 kDa which suggest the cleavage of the gamma protein to produce a mature capsids. This would need to be confirmed by the identification of the gamma peptide by SDS-PAGE analysis. TEM analysis confirmed the maturation of particles, with the appearance of the VLPs resembling mature capsid proteins with trimeric clusters on the surface instead of dumbbell shapes, as was observed on the surface of immature NωV particles (Canady *et al*., 2000). The particles also appeared to be densely packed and impermeable to the stain. Since NωV is the only tetravirus to date whose crystal structure has been elucidated (Munshi *et al*., 1996), one could assume that a deletion of the N-terminal of the capsid protein of NωV might affect RNA encapsidation and recognition as well as particle stability due to its location within the capsid. Due to similarities in the capsid protein structure, the above-mentioned characteristics could be observed in N-terminal deletion mutants of NβV VLPs.

These data support the hypothesis that the first 50 amino acids of the NβV capsid coding sequence are not required for VLP assembly and also substantiates the theory that the population of smaller particles seen from preparations of NβV might be the result of a translated protein from the second methionine residue on the capsid coding sequence. Future work on these experiments requires the analysis of RNA within these smaller particles

– would these particles contain full length genomic RNA? Strategies would need to be devised to study this in the future.

The deletion of the first 31 amino acids in the nodavirus FHV showed a decrease in the packaging of RNA 2 within capsids which suggests that these amino acids play an important role in RNA packaging and selectivity (Krishna *et al*., 2003). It has also been proposed that cleavage of the capsid protein precursor is required in TaV in order to unblock the N-terminal regions of the three proteins produced from this event and that the cleavage of the precursor protein is required for maturation (Pringle *et al*., 1999). These reasons for the cleavage of the capsid precursor protein could explain the presence of a second, shorter capsid protein found for NβV. Instead of this protein being cleaved during maturation, a second protein is translated from the ORF in order to compensate for the lack of an identifiable cleavage site.

Although it appears that the deletion of the first 50 amino acids from the NβV capsid protein has no effect on the NβV VLP exterior morphology and assembly, more studies will have to be performed in order to: a) separate out the two wild type particles during the purification process on the basis of density since the smaller VLPs are most likely also denser and b) analyze the RNA content of each of these particles. As was expected, the average diameter of particles produced was smaller than wild type particles. Purified VLPs should be compared in size to the mixture of proteins isolated from the *N. capensis* larvae in order to confirm the expressed truncated version of the capsid gene is of the same molecular weight as the smaller viral protein.

There are a number of interesting questions that still need to be answered that have arisen from the results of this work. The first question is if there is a protease cleavage site at the N-terminus of the NβV capsid protein or a second translational start? The alignment of the amino acid sequences of the capsid sequence of NωV, HaSV, FHV, PaV, NβV, TaV, EeV and PrV revealed

that a second in-frame methionine residue present between two structurally similar regions was present in both PaV and NβV. This region is the proposed translational start of the NβV capsid gene and has also been found in PaV and can act as the start of the capsid gene (Ball and Johnson, 2003). The second question is whether or not the virus particles extracted contain a mixture of the two proteins, p56 and p58 and if so then this phenomenon must be restricted to just the larger of the two capsid particles since the expression of pCW7 yielded small, uniform particles from a single gene. Would a mixture of p56 and p58 result in the formation of mature VLPs and would these VLPs consist of a mixture in distribution of diameters? Would this distribution be of the same ratio found in wild-type NβV particles and would individual particles be composed of one protein or a mixture of both?

This work has contributed towards the understanding of the betatetraviruses in numerous areas of interest such as presence of NβV in *N. capensis* larvae, the expression of viral genes in a heterologous host, the evolution and occurrence of these viruses in nature and the modes of replication of the genes of these seemingly simple viruses.

Appendices

Appendix A – Oligonucleotides

Table A.1 shows the sequences of the oligonucleotides used in the PCR reactions mentioned including the annealing temperatures of each oligonucleotide in brackets.

Table A.1. The oligonucleotides used in this study.

Primer name	Sequence
RDNBVCAPR (53°C) and RDNBVCAPF1 (53°C)	GGATCCAATATGGATGCCAACGTGCAGATAAC and GGGAACCTTAGCTTGTAGGTTCGAA
pucF (55°C) and pucR (55°C)	CGCCAGGGTTTTCCCAGTCACGAC and TCACACAGGAAACAGCTATGAC
REP2 (45°C)and CAP1 (45°C)	GAATGGGCAGTCAGACCG and CGGGTCTAGATAGTCGTC
REP2 (45°C) and NBV4050R (45°C)	GAATGGGCAGTCAGACCG and TGGCATCCATCTTGTCCG
CW R1 (68°C) and CW F1 (62°)	CACGGGTTGGAACGCCCTCAGGGGC and GTCGGGAACGTGTTTTACATGGTCG
CW 3 (72°C)	CATCTCGCAAATAAATAAGTATTTT
RDNBVCAPF2 and NBVCAP2	GGATCCAATATGGCCGAACCCGCGAACC ATCAACTCGACCCAATCG

Appendix B - Competent cells and transformation thereof.

Preparation of competent cells

5 ml of sterile Luria broth was inoculated with a newly streaked out colony of *E.coli* DH5α or DH10BAC and left to shake at 37°C overnight. The following volumes of culture in mid-log were inoculated into 500 ml flasks with 100 ml of sterile Luria broth: 1.5 ml, 1.0 ml, 0.7 ml and 0.3 ml. These flasks were set shaking at 37°C until an absorbance reading between 0.6 to 0.8 of the 1.5 ml inoculated flask measured, indicating that the cells are in mid to early log phase. The inoculated media was dispensed into sterile JA14 centrifuge bottles and cooled on ice for 5-10 minutes. The bottles were then centrifuged for 10 minutes at 5 000 rpm at 4°C using a JA14 Beckman rotor. The supernatant was discarded and the pellet resuspended in 100 ml of RF 1 solution (100 mM KCL, 50 mM $MnCl_2$, 30 mM CH_3COOK, 10 mM $CaCl_2$ and 15 % glycerol). The bottles were then cooled on ice for 20 minutes, centrifuged at 5 000 rpm for 10 minutes at 4°C and resuspended in 8 ml of

RF2 solution (10 mM MOPS, 10 mM KCL, 75 mM $CaCl_2$ and 15 % glycerol). The resuspended cells were then aliquoted in Eppendorf tubes and stored at -80°C. The competency of the cells was measured by transforming a particular volume with a known concentration of pUC18.

Transformation process (Sambrook *et al*., 1989)

DNA to be transformed was added to 50 µl of thawed competent cells and left on ice for 20 minutes. The cells were then heat shocked at 42°C for 45 seconds and placed back on ice for 5 minutes. 1 ml of Luria Broth was added to each tube and the tubes placed at 37°C for 1 hour. 100 to 200 µl of the transformed cells were plated out onto the appropriate selective medium. A positive control, pUC18 was used to calculate the transformation frequency.

Appendix C - Plasmid miniprep protocols and general cloning solutions.

Smart buffer (Berghammer and Auer, 1993).

0.5 ml10 mM Tris-Cl pH 8.0
0.1 ml1mM EDTA pH 8.0
7.5 g 15 % w/v sucrose
1 ml 200 µg/ml DNAse-free RNAse (10 mg/ml stock)
1 ml 100 mg/ml BSA
Made up to 50 ml, stored in aliquots and kept at -20°C.

Plasmid mini-prep protocol (Berghammer and Auer, 1993).

1.5 ml bacterial culture grown to late log phase was added to an Eppendorf tube. The tube was microfuged for 1 minute at 13 000 rpm, the supernatant discarded and the pellet rigorously resuspended in 50 µl of Smart buffer. The tube was then incubated for 30 minutes to 1 hour at 37°C, boiled for 1 minute and placed on ice for 5 minutes. The tube was microfuged at 13 000 rpm for 10 minutes and the supernatant transferred to a sterile Eppendorf tube. The supernatant contains plasmid DNA.

Media

Luria Broth (1 L)

5 g NaCl
5 g yeast extract
10 g tryptone
Made up to 1 L with ddH_2O and autoclaved

Luria agar (1 L)

5 g NaCl
5 g yeast extract
10 g tryptone
12 g bacteriological agar
Made up to 1 l with ddH_2O and autoclaved

MacConkey agar (1 L)
50 g MacConkey agar
Made up to 1 l with ddH_2O and autoclaved

RNAse A
0.1 g of RNAse A (Roche) was dissolved in 100 µl of Tris-Cl (1 M) pH 7.5 and 30 µl of NaCl (5M) made up to 10 ml in $dddH_2O$. This was boiled for 15 minutes, slowly cooled to room temperature, aliquoted and stored at -20°C.

T.E buffer
121.14 g Tris
372.24 g EDTA
pH 8 with hydrochloric acid and autoclaved

Phenol
Phenol (Sigma) was equilibrated to pH8 with TE buffer (pH8) while stirring at room temperature.

Phenol:Chloroform:isoamyl (P.C.I) [25:24:1]
Phenol (pH8) was mixed with chloroform and isoamyl in the above-mentioned ratios and stored at 4°C away from light.

TAE (50 X) [1 L]
242 g Tris base
57.1 ml glacial acetic acid
37.2 g EDTA
pH 8.5 with glacial acetic acid and autoclaved

X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)
Made up to 20 mg/ml in dimethylformamide and away from light at -20°C.

Appendix D – RNA methodology and Northern blot solutions

DEPC-treated water
1 ml of diethylpyrocarbonate was added to 999 ml of $dddH_2O$, incubated at 37°C shaking and autoclaved for 30 minutes.

2 M $NaPO_4$ (100 ml)
14.2 g Na_2HPO_4 (anhydrous)
13.8 g NaH_2PO_4
Made up to 100 ml with nuclease-free water

50 X Denhardt's solution (500 ml)
5 g Ficoll 400
5 g Polyvinylprrolidone
5 g Bovine Serum Albumin (Fraction V)
Made up to 500 ml with nuclease-free water and store at -20°C

10 mg/ml salmon sperm DNA
1 g of sodium salt type III DNA from salmon testes was dissolved in 100 ml water overnight with rotation. The DNA was then sheared using an 18 g needle, boiled for 5 minutes, aliquoted and stored at -20°C.

De-ionized formamide
1g of Sigma Amberlite MB-1 was weighed out and added to 100 ml of formamide. This was stirred at room temperature for 1 hour and filtered through Whatman 3MM paper. The de-ionized formamide was stored at -20°C.

Hybridization solution (100 ml)
50 ml Formamide
2.5 ml 2 M $NaPO_4$
16 ml 5 M NaCl
20 ml 50 X Denhardt's solution
2.5 ml 10 mg/ml salmon sperm DNA
5 ml 10 % SDS
3.8 ml nuclease-free water
Aliquoted and stored at -20°C away from light

Wash solution (500 ml)
10 ml 10XSSC
1 g SDS
Made up to 500 ml with nuclease-free water.

Appendix E - Protein and Western blot analyses solutions

Running buffer (2 X Q) 1 L
57.6 g glycine
12 g Tris base
4 g Sodium dodecyl sulphate (SDS)
Made up to 1 l and diluted 1 in 2 for use

Destain I (1 L)
500 ml of methanol
100 ml of acetic acid
400 ml ddH_2O

Destain II (1 L)
80ml methanol
70 ml acetic acid
880 ml ddH_2O

Destain III (1 L)
50ml methanol
70 ml acetic acid
30 ml glycerol
850 ml ddH_2O

Stripping buffer (500 ml)
50 ml 1 M Tris pH 7.6
15 ml 5 M NaCl
100 ml 10 % SDS
3.5 ml β-mercaptoethanol
Made up to 500 ml with ddH_2O

Coomassie Stain
1.25 g Coomassie R125 in 500 ml Destain I. Filtered through a Buchner funnel and stored at room temperature.

2 X Sample buffer
1.25 ml of 1 M Tris pH 6.8
4 ml of 10 % SDS
1 ml of β-mercaptoethanol
1 ml of glycerol
0.001 g of bromophenol blue
2.75 ml of ddH_2O
Stored at 4°C away from light.

Transfer buffer (1 L)
3.03 g Tris base
14.48 g glycine
200 ml methanol
Made up to 1 l and kept ice cold.

TBS-TWEEN 20
100 ml 1 M Tris-Cl pH 7.6
30 ml 5 M NaCl
1 ml Tween-20
Made up to 1 L

Blotto solution
5 g of fat free milk powder dissolved in 100 ml of TBS-TWEEN.

Ponceau Stain (10 X)
0.5 ml Ponceau
1 ml glacial acetic acid
Made up to 100 ml in $dddH_2O$

Developer (2 L)
1 L tap water
500 ml solution A
50 ml solution B
50 ml solution C
350 ml water
50 ml starter solution

Stop solution (2 L)
400 ml glacial acetic acid
1.96 L tap water

Fixer solution (2 L)
1.2 L water
400 ml solution A
100 ml solution B
300 ml water

Solution A, B, C and starter solution were all purchased from Agfa™.

Appendix F - Plasmid maps of constructs used in this work.

Figure A.1. **Plasmid maps of pCW4-7**. Restriction endonuclease sites and genes of interest are marked on the plasmids.

Appendix G - Plasmid maps of constructs used in this work.

Figure A.1. **Plasmid maps of pCW4-7**. Restriction endonuclease sites and genes of interest are marked on the plasmids.

References

Agrawal, D.K and Johnson, J.E. 1992. Sequence and analysis of the capsid protein of *Nudaurelia capensis ω virus*, an insect virus with *T*=4 icosahedral symmetry. *Virol*. **190**: 806-814.

Berghammer, H. and Auer, B. 1993. Easypreps: fast and easy plasmid minipreparation for analysis of recombinant clones in *E. coli*. *Biotechniques* **14**: 527-528.

Brooks, K.M., Gordon, K.H.J., Dorrian, S.J., Hines, E.R., Hanzlik, T.N. 2002. Infection of its *Lepidopteran* host by *Helicovepa armigera stunt virus* (*Tetraviridae*). *J. Invert. Path.* **80**: 97-111.

Canady, M.A., Tihava, M., Hanzlik, T.N., Johnson. J.E and Yeager, M. 2000. Large conformational changes in the maturation of a simple RNA virus, *Nudaurelia capensis ω virus* (NωV). *J. Mol. Biol*. **299**: 573-584.

Caspar, R.H and Klug, A. 1962. Physical principles in the construction of regular viruses. Cold Spring Harbour Symp. *Quant. Biol*. **72**: 1-24.

Chao, Y.C., Scott, H.A. and Young, S.Y. 1983. An icosahedral RNA virus of the soybean looper (*Pseudoplusia includens*). *J. Gen. Virol*. **64**: 1835-1838.

Christians, P.D., Dorrian, S.J., Gordon, K.H.J and Hanzlik, T.N. 2001. Pathology and properties of the tetravirus *Helicoverpa armigera stunt virus*. *Biol. Control*. **20**: 65-75.

Dong, X.F., Natarajan, P., Tihova, M., Johnson, J.E. and Schneemann, A. 1998. Particle polymorphism caused by deletion of a peptide molecular switch in a quasi-equivalence virus. *J. Virol*. **72**: 6024-6033.

du Plessis, D.H., G. Mokhosi, and D.A. Hendry. 1991. Cell-free translation and identification of the replicative form of *Nudaurelia β virus* RNA. *J. Gen. Virol*. **72**: 267-273.

Fields, B.N., Knipe, D.M., Howley, P.M., Chanock, J., Melnick, L., Monath, T.P., Roizman, B and Straus, S.E. 1996. Fundamental Virology, Third Edition. Lippincott-Raven. Chapter 13, pages 401-402.

Finch, J.T., R.A. Crowther, D.A. Hendry, and J.K. Struthers. 1974. The structure of *Nudaurelia capensis ω virus*: the First Example of a Capsid with Icosahedral Surface symmetry T=4. *J. Gen. Virol*. **24**: 191-200.

Geertsema, H. and J.H. Giliomee. 1972. Aspects of the Biology of the Pine Tree Emperor moth, *Nudaurelia cytherea* (Fabr.) (*Lepidoptera*: *Saturnidae*). Annals of University of Stellenbosch. **50 A (1**): 1-170.

GibcoBRl (Invitirogen ™ Life Technologies). Instruction manual for Bac-to Bac® Baculovirus Expression Systems. Updated version, 2002.

Gorbalenya, A.E, Pringle, F.M, Zeddam, J-L, Luke, B.T., Cameron, C.E., Kalmakoff, J, Hanzlik, T.N., Gordon, K.H and Ward, V.K. The palm subdomain-based active site is internally permuted in viral RNA-dependent RNA polymerases of an ancient lineage. *J. Mol. Biol.* **324**: 47-62.

Gordon, K.H.J., Williams, M.R., Hendry, D.A., and Hanzlik, T.N. 1999. Sequence of the genomic RNA of *Nudaurelia β virus* (*Tetraviridae*) Defines a Novel Virus Genome Organization. *Virol*. **258**: 42-53.

Grace, T.D.C and Mercer, E.H. 1965. A new virus of the Saturniid *Antheraea eucalypti* Scott. *J. Inverter. Pathol*. **7**: 241-244.

Greenwood, L.K. and Moore, N.F. 1984. Pathogenicity of a small RNA-virus of insects determined by a virulent virus rather than a susceptible host. *Microbiologica*. **7**: 97-102.

Hanzlik, T. N., Dorrian, S .J., Johnson, K. N, Brooks, E .M, Gordon, K H. 1995. Sequence of RNA 2 of the *Helicoverpa armigera stunt virus* (*Tetraviridae*) and bacterial expression of its genes. *J. Gen. Virol*. **76**: 799-811.

Hanzlik, T.N. and Gordon, K.H.J. 1997. The *Tetraviridae*. *Adv. Virus. Res.* **48**: 101-168.

Hanzlik, T.N., Dorrian, S.J., Gordon, K.H.J., and Christian, P.D. 1993. A novel small RNA virus isolated from the cotton boll worm *Helicoverpa armigera*. *J. Gen. Virol*. **74**:1105-1110.

Helgstrand, C., Munshi, S., Johnson, J.E. and Liljas, L. 2004. The refined structure of *Nudaurelia capensis ω virus* reveals control elements for a T=4 capsid maturation. *Virol*. **318**: 192-203.

Hendry, D.A. and D.K. Agrawal. 1994. Tetraviruses. *Encyclopedia of Virology* .Second Edition*:* 1416-1422.

Hendry, D.A., Bekker, M.F. and van Regenmortel, M.H.V. 1968. A non-inclusion virus of the Pine Emperor moth *Nudaurelia cytherea capensis* Stoll. *S.A. Medical Journal* **42**: 117-117.

Hendry, D.A., Hodgson, V, Clark, and Newman, J. 1985. Small RNA viruses coinfecting the Pine Emperor Moth (*Nudaurelia cytherea capensis*). *J. Gen. Virol*. **66**: 627-632.

Johnson, J.E., Munshi, S, Liljas, L, Agrawal, D., Olson, N.H., Reddy, V., Fisher, A., McKinney, B., Schmidt, T., and Baker, T.S. 1994. Comparative studies of *T*=3 and *T*=4 icosahedral RNA insect viruses. *Arch. Virol*. **9**: 497-512.

Johnson, K.N and Ball, L.A. 2003. Virions of *Pariacoto virus* contain a minor protein translated from the second AUG codon of the capsid protein open reading frame. *J. Gen. Virol*. **84**: 2847-2852.

Juckes, I.R.M. 1974. Purification and characterization of viruses of the pine emperor moth *Nudaurelia cytherea capensis* (Fabr.). PhD thesis. University of Cape Town.

Krishna, N.K., Marshall, D and Schneemann, A. 2003. Analysis of RNA Packaging in Wild-Type and Mosaic Protein Capsid of Flock House Virus Using Recombinant Baculovirus Vectors. *Virol*. **305**: 10-24.

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

Lu, M.W. and Lin, C.S. 2003. Involvement of the terminus of grouper betanodavirus capsid protein in virus-like particle assembly. *Arch. Virol*. **148**: 345-355.

Luke, G. J. 2001. Assembly of full-length cDNA, and heterologous expression of, *Nudaurelia capensis β virus* RNA. Masters Thesis, Rhodes University, South Africa.

Marshall, D. and Schneemann, A. 2001. Specific packaging of Nodaviral RNA2 requires the N-terminus of the capsid protein. *Virol*. **285**: 165-175.

Matthews, R.E.F. 1982. *Nudaurelia β virus* group. *Intervirol*. **17**: 135-135.

Mayo, M.A., and Pringle, C.R. 1998. Virus Taxonomy: 1997. *J. Gen. Virol*. **79**: 649-657.

Moore, N.F. 1991*a*. The *Nudaurelia* β family of insect viruses. In Viruses of Invertebrates. Ed. E. Kurstak. New York: Marcel Dekker. Pages 277-285.

Moore, N.F. 1991*b*. Identification, pathology, structure, and replication of insect picornaviruses. *In Viruses of Invertebrates*. Ed. E. Kurstak. New York: Marcel Dekker. Pages 287-299.

Morris, T.J, Hess, R.T and D.E. Pinnock. 1979. Physiochemical characterization of a small RNA virus associated with baculovirus infection in *Trichoplusia ni*. *Intervirol*. **11**: 238-247.

Munshi, S., Liljas, L., Cavarelli, J., Bomu, W., MicKinney, B., Reddy, V., and Johnson, J.E. 1996. The 2.8 Å Structure of a T=4 Animal Virus and its Implications for Membrane Translocation of RNA. *J .Mol. Biol*. **261**: 1-10.

Olson, N.H., T.S. Baker, J.E. Johnson, and D.A. Hendry. 1990. The three-dimensional structure of frozen-hydrated *Nudaurelia capensis β virus*, a T=4 insect virus. *J. Struct. Biol*. **105**: 111-122.

O'Reilly, D.R, Miller, L.K and Luckow, V.A. 1992. *Baculovirus Expression Vectors* – A laboratory manual. New York, W.H. Freeman and Company. 1-10, 27-29, 38-49, 124-141.

Possee, R.D. 1997. Baculoviruses as expression vectors. *Cur. Opin. Biotech*. **8**: 569-572.

Pringle, F.M., Gordon, K.H.J., Hanzlik, T.N., Kalmakoff, J., Scotti, P.D. and Ward. V.K. 1999. A novel capsid expression strategy for *Thosea asigna virus* (*Tetraviridae*). *J. Gen. Virol*. **80**: 1855-1863.

Pringle, F.M., Kalmakoff, J and Ward, V. 2001. Analysis of the capsid processing strategy of *Thosea asigna virus* using baculovirus expression of virus-like particles. *J. Gen. Virol*. **82**: 259-266.

Pringle, F.M., Johnson, K.N., Goodman, C.L., McIntosh, A.H. and Ball, L.A. 2003. Providence virus: a new member of the *Tetraviridae* that infects cultured insect cells. *Virol*. **306**: 359-370.

Reddy, R.S., Natarajan, P., Okerberg, B, Li, K., Damodaran, K.V., Morton, R.T., Brooks, C.L., Johnson, J.E. 2001. Virus Particle Explorer (VIPER), a Website for virus capsid structures and their computational analysis. *J. Virol*. 75:11943-11947

Reinganum, C., Robertson, J.S. and Tinsley, T.W. 1978. A new group of RNA viruses from insects. *J. Gen. Virol*. **40**: 195-202.

Reinganum, C. 1991. *Tetraviridae*. In "Atlas of Invertebrate Viruses". Boca Raton, CRC Press 553-592.

Rossmann, M.G. and Johnson, J.E. 1989. Icosahedral RNA virus structure. *Annu. Rev. Biochem.* **58**: 533-573.

Rossmann, M.G and Tao, Y. 1999. Structural insight into insect viruses. *Nature. Struc. Bio.* **6**: 717-719.

Rowlands, D.J, Mayo, M.A. and Mahy, B.W.J. 1987. The molecular biology of positive strand RNA viruses. FEMS symposium. London Academic Press.

Sambrook, J., Fritsch, E.F and Maniatis, T. 1989. Molecular Cloning – a laboratory manual. 2nd ed. Cold Spring Harbour press, New York.

Smith, D.B., McAllister, J., Casino, C and Simmonds, P. 1997. Virus 'quasispecies': making a mountain out of a molehill? *J. Gen. Virol*. **78**: 1511-1519.

Struthers, J.K. and D.A. Hendry. 1974. Studies of the Protein and Nucleic Acid Components of *Nudaurelia capensis β virus. J. Gen. Virol*. **22**: 355-362.

Taylor, D. J., Krishna, N.K., Canady, M.A., Schneemann, A and Johnson, J.E. 2002. Large-Scale, pH-dependent Quaternary Structure Changes in an RNA Virus Capsid are Reversible in the Absence of Subunit Proteolysis. *J. Virol*. **76**: 9972-9980.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R. and Wickner, R.B. 2000. Virus taxonomy. Seventh report of the International Committee on Taxonomy of Viruses. Academic Press, New York.

Venter, P.A. 2001. Heterologous expression of the *Helicoverpa armigera stunt* v*irus* in *Saccharomyces cerevisiae*. PhD Thesis. Rhodes University, South Africa.

Appendix F – *Nudaurelia β virus* nucleotide and protein sequence.

```
                                                                                                                                M  E  D
.
1     GATTCTCCCA CTTGATTGTA CGGAGAACCC GAAACACGGT GGAGGGGGCC CCGATCCCA AGCCTTGATA TCGCTTGTGT TAGGCACAAG TGATGGAAGA
      CTAAGAGGGT GAACTAACAT GCCTCTTGGG CTTTGTGCCA CCTCCCCCGG TGGCTAGGGT TCGGAACTAT AGCGAACACA ATCCGTGTTC ACTACCTTCT
      · A  S  K   Q  L  R  V   L  D  A   Q  E  R   A  K  A  A   F  Q  L   D  F  I   A  S  V  E   T  L  E   D  A  Q
101   CGCAAGCAAG CAGCTCCGCG TCCTGGATGC CCAGGAGCGC GCGAAGGCCG CCTTCCAACT CGACTTCATA GCCTCTGTCG AGACTTTGGA AGACGCTCAG
      GCGTTCGTTC GTCGAGGCGC AGGACCTACG GGTCCTCGCG CGCTTCCGGC GGAAGGTTGA GCTGAAGTAT CGGAGACAGC TCTGAAACCT TCTGCGAGTC
       E  K  Y  E   G  M  M   F  R  S   G  T  K  L   P  S  T   H  I  K   L  A  I  D   L  R  V   A  E  K   D  L  R  R
.
201   GAGAAGTACG AGGGCATGAT GTTTCGCAGT GGCACGAAAC TGCCATCAAC CCATATTAAG TTGGCAATCG ATCTGAGAGT TGCGGAGAAA GATCTACGCC
      CTCTTCATGC TCCCGTACTA CAAAGCGTCA CCGTGCTTTG ACGGTAGTTG GGTATAATTC AACCGTTAGC TAGACTCTCA ACGCCTCTTT CTAGATGCGG
      · H  V  K   N  V  P   T  V  L  E   I  G  P   S  V  E   S  V  R  Y   A  V  Q   T  R  D   K  E  R  V   H  G  C
.
301   GGCACGTTAA GAATGTACCG ACAGTGCTGG AAATTGGACC CAGTGTTGAG AGCGTGCGTT ACGCTGTGCA GACTCGAGAC AAGGAGAGAG TCCATGGCTG
      CCGTGCAATT CTTACATGGC TGTCACGACC TTTAACCTGG GTCACAACTC TCGCACGCAA TGCGACACGT CTGAGCTCTG TTCCTCTCTC AGGTACCGAC
      · T  F  S   D  A  R  D   N  L  R   H  N  K   I  G  Y  E   A  H  Y   D  R  K   I  G  P  D   A  A  L   L  A  A
401   CACCTTCTCC GACGCGCGTG ATAACCTCCG CCACAATAAG ATCGGTTATG AAGCCCATTA CGACAGAAAG ATTGGACCTG ACGCCGCCCT TCTGGCCGCT
      GTGGAAGAGG CTGCGCGCAC TATTGGAGGC GGTGTTATTC TAGCCAATAC TTCGGGTAAT GCTGTCTTTC TAACCTGGAC TGCGGCGGGA AGACCGGCGA
       G  I  P  T   D  T  F   C  V  D   G  F  S  N   C  E  Y   Q  S  P   L  A  I  A   C  H  S   L  Y  P   D  G  E  S
.
501   GGTATCCCAA CTGACACCTT CTGTGTCGAC GGCTTCTCCA ATTGCGAGTA CCAATCCCCC CTCGCCATTG CCTGCCACTC ACTTTACCCC GATGGGGAAA
      CCATAGGGTT GACTGTGGAA GACACAGCTG CCGAAGAGGT TAACGCTCAT GGTTAGGGGG GAGCGGTAAC GGACGGTGAG TGAAATGGGG CTACCCCTTT
      · N  S  I   M  D  V   A  K  G  M   A  L  H   G  T  H   V  I  Y  A   W  M  H   L  P  V   E  L  L  T   L  T  D
.
601   GTAATAGTAT TATGGACGTG GCTAAAGGCA TGGCTCTCCA CGGCACCCAC GTGATATATG CGTGGATGCA TCTGCCCGTG GAACTGCTAA CGCTCACCGA
      CATTATCATA ATACCTGCAC CGATTTCCGT ACCGAGAGGT GCCGTGGGTG CACTATATAC GCACCTACGT AGACGGGCAC CTTGACGATT GCGAGTGGCT
      · A  D  N   I  F  E  G   Y  S  I   R  F  E   E  T  G  A   L  P  C   T  K  R   R  K  A  I   F  S  G   Y  N  D
701   TGCAGACAAT ATTTTTGAAG GGTATAGCAT TCGGTTTGAG GAGACAGGGG CACTACCCTG CACCAAGAGG AGGAAAGCCA TATTCTCCGG TTATAACGAT
      ACGTCTGTTA TAAAAACTTC CCATATCGTA AGCCAAACTC CTCTGTCCCC GTGATGGGAC GTGGTTCTCC TCCTTTCGGT ATAAGAGGCC AATATTGCTA
       F  G  S  A   Y  V  H   D  A  H   H  W  A  G   W  L  K   H  R  G   V  D  T  P   Y  G  F   S  I  L   I  D  I  Q
.
 801  TTTGGTTCGG CCTATGTGCA CGATGCCCAC CATTGGGCTG GTTGGCTTAA GCATCGGGGA GTAGACACCC CGTATGGCTT CTCCATATTG ATCGACATAC
      AAACCAAGCC GGATACACGT GCTACGGGTG GTAACCCGAC CAACCGAATT CGTAGCCCCT CATCTGTGGG GCATACCGAA GAGGTATAAC TAGCTGTATG
```

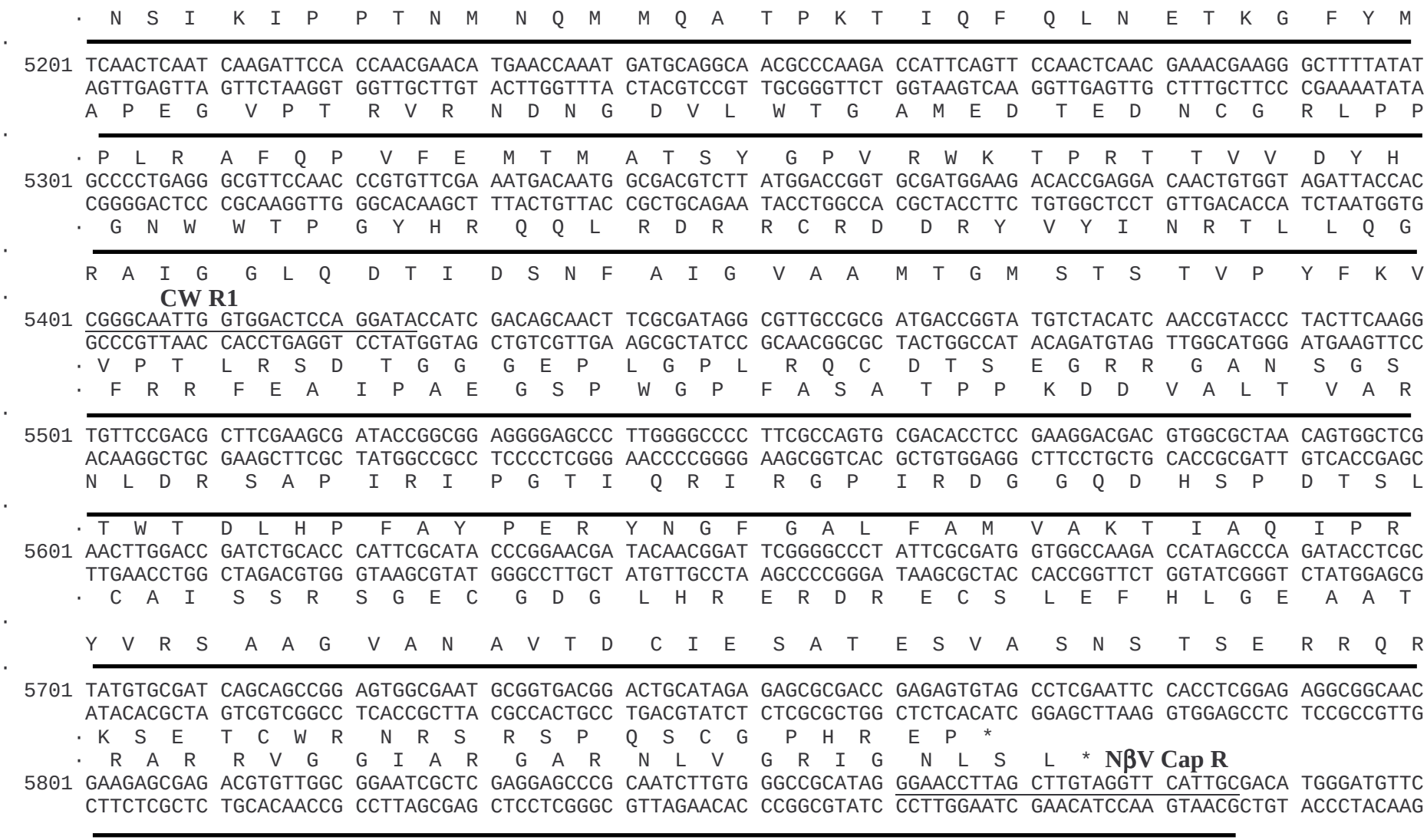
```
     · Q R F  G M H  T K L K  I T R  G H S  S G S I  T T V  F P L  S K L G  L I W
.
  901 AACAGAGGTT CGGTATGCAC ACGAAATTAA AGATCACCCG TGGGCACAGC AGTGGCAGTA TCACCACCGT GTTCCCGTTG TCGAAATTGG GCTTGATCTG
      TTGTCTCCAA GCCATACGTG TGCTTTAATT TCTAGTGGGC ACCCGTGTCG TCACCGTCAT AGTGGTGGCA CAAGGGCAAC AGCTTTAACC CGAACTAGAC
     · V P N  I V K I  M Y P  K A K  H E P E  Y I V  T D K  K K Y E  G V C  V Y V
 1001 GGTGCCGAAC ATAGTCAAAA TAATGTACCC TAAAGCCAAA CACGAGCCGG AGTACATCGT CACGGATAAG AAGAAGTATG AAGGCGTTTG CGTGTACGTC
      CCACGGCTTG TATCAGTTTT ATTACATGGG ATTTCGGTTT GTGCTCGGCC TCATGTAGCA GTGCCTATTC TTCTTCATAC TTCCGCAAAC GCACATGCAG
      G T R V  Q S S  G K S  I T L A  E I V  Q Y I  R T R L  T R I  I L N  G T V H
.
 1101 GGAACGAGGG TGCAAAGTTC CGGCAAGTCT ATTACGCTCG CTGAGATTGT TCAATACATC CGAACAAGAT TAACACGCAT CATTCTGAAT GGCACTGTCC
      CCTTGCTCCC ACGTTTCAAG GCCGTTCAGA TAATGCGAGC GACTCTAACA AGTTATGTAG GCTTGTTCTA ATTGTGCGTA GTAAGACTTA CCGTGACAGG
     · E K T  W T I  A E Q D  I E R  L A V  S I M F  R K N  V E R  A V S E  K A L
.
 1201 ACGAGAAAAC GTGGACCATA GCAGAGCAAG ACATTGAGCG ACTTGCCGTT AGCATTATGT TCCGCAAGAA TGTGGAACGC GCTGTGTCTG AAAAGGCACT
      TGCTCTTTTG CACCTGGTAT CGTCTCGTTC TGTAACTCGC TGAACGGCAA TCGTAATACA AGGCGTTCTT ACACCTTGCG CGACACAGAC TTTTCCGTGA
     · M R A  Q K K C  K S A  E K Q  A L L P  V W M  R R I  A N W F  Q D K  F Q I
 1301 GATGAGAGCG CAGAAGAAGT GCAAGAGCGC TGAAAAACAA GCGCTGCTGC CAGTTTGGAT GCGGAGGATC GCCAATTGGT TCCAAGACAA ATTTCAAATC
      CTACTCTCGC GTCTTCTTCA CGTTCTCGCG ACTTTTTGTT CGCGACGACG GTCAAACCTA CGCCTCCTAG CGGTTAACCA AGGTTCTGTT TAAAGTTTAG
      D E E V  V R K  R Y L  E C L K  A Q P  W I H  A D K V  V N C  E T K  R Y N P
.
 1401 GACGAGGAGG TCGTACGCAA GCGCTACCTT GAGTGTCTCA AGGCGCAACC CTGGATCCAC GCCGATAAAG TGGTGAACTG CGAGACCAAG CGCTATAACC
      CTGCTCCTCC AGCATGCGTT CGCGATGGAA CTCACAGAGT TCCGCGTTGG GACCTAGGTG CGGCTATTTC ACCACTTGAC GCTCTGGTTC GCGATATTGG
     · T V A  E V G  P K N H  L L A  T T G  L R E L  Q R E  I P S  A N E P  Q D R
.
 1501 CTACTGTCGC CGAGGTGGGT CCTAAGAATC ATTTGCTCGC CACTACCGGA TTGCGCGAGC TCCAAAGGGA AATACCCAGT GCTAACGAAC CGCAAGATAG
      GATGACAGCG GCTCCACCCA GGATTCTTAG TAAACGAGCG GTGATGGCCT AACGCGCTCG AGGTTTCCCT TTATGGGTCA CGATTGCTTG GCGTTCTATC
     · G A K  A W H S  A H A  D L D  I Y A E  G L R  L D S  A K E A  A A G  K Q S
 1601 AGGAGCCAAG GCATGGCACT CCGCTCACGC CGATCTCGAC ATTTACGCCG AGGGACTCCG ACTCGACTCC GCTAAAGAGG CAGCAGCGGG TAAGCAGTCC
      TCCTCGGTTC CGTACCGTGA GGCGAGTGCG GCTAGAGCTG TAAATGCGGC TCCCTGAGGC TGAGCTGAGG CGATTTCTCC GTCGTCGCCC ATTCGTCAGG
      L A I T  L Q Q  A F Q  V L G K  T K C  E G C  N N I E  I E Y  W T G  P P G S
.
 1701 CTGGCGATCA CATTGCAGCA AGCTTTCCAA GTCCTAGGGA AGACCAAGTG CGAGGGGTGC AACAACATTG AAATCGAGTA CTGGACCGGA CCCCCCGGTT
      GACCGCTAGT GTAACGTCGT TCGAAAGGTT CAGGATCCCT TCTGGTTCAC GCTCCCCACG TTGTTGTAAC TTTAGCTCAT GACCTGGCCT GGGGGGCCAA
     · G K S  R A A  K P R F  A D L  Q G G  V L Y C  A P T  R T L  R D A L  D E S
.
 1801 CCGGGAAATC CAGGGCCGCT AAGCCGAGAT TTGCAGATTT GCAGGGGGGC GTGTTGTACT GCGCCCCTAC GCGCACGCTG CGCGACGCCC TCGACGAAAG
```

```
     GGCCCTTTAG GTCCCGGCGA TTCGGCTCTA AACGTCTAAA CGTCCCCCCG CACAACATGA CGCGGGGATG CGCGTGCGAC GCGCTGCGGG AGCTGCTTTC
    · V  V  H   P  S  R  V   C  T  Y   H  N  A   L  H  V  A   A  K  E   S  G  N   R  P  F  D   V  I  V   I  D  E
1901 CGTCGTGCAC CCTTCCCGTG TTTGCACTTA CCACAACGCA CTGCATGTCG CTGCCAAAGA GTCTGGCAAT AGGCCTTTTG ACGTTATCGT CATCGATGAA
     GCAGCACGTG GGAAGGGCAC AAACGTGAAT GGTGTTGCGT GACGTACAGC GACGGTTTCT CAGACCGTTA TCCGGAAAAC TGCAATAGCA GTAGCTACTT
     A  E  T  T   P  A  C   Y  V  G   T  M  H  H   A  S  P   S  S  R   I  V  C  L   G  D  P   H  Q  I   G  Y  I  D
.
2001 GCGGAGACGA CGCCGGCTTG CTACGTAGGT ACGATGCATC ATGCATCGCC TAGTAGTAGG ATCGTCTGTC TGGGCGATCC GCACCAGATC GGTTACATCG
     CGCCTCTGCT GCGGCCGAAC GATGCATCCA TGCTACGTAG TACGTAGCGG ATCATCATCC TAGCAGACAG ACCCGCTAGG CGTGGTCTAG CCAATGTAGC
    · F  S  D   R  K  D   D  L  K  P   F  S  I   I  A  A   E  C  R  T   R  R  F   N  T  T   Y  R  C  P   Q  D  V
.
2101 ACTTTTCGGA TCGAAAAGAC GATTTGAAAC CTTTCAGTAT CATAGCAGCC GAATGTCGCA CTCGTAGGTT TAACACCACT TATAGGTGCC CACAAGACGT
     TGAAAAGCCT AGCTTTTCTG CTAAACTTTG GAAAGTCATA GTATCGTCGG CTTACAGCGT GAGCATCCAA ATTGTGGTGA ATATCCACGG GTGTTCTGCA
    · L  N  L   P  I  F  K   T  L  Y   P  D  A   I  S  F  S   K  Q  L   T  S  I   R  Y  L  T   R  A  R   S  V  T
2201 TTTAAACTTG CCCATATTCA AAACTCTATA CCCGGACGCG ATATCGTTCA GCAAACAATT GACTAGCATC CGTTACCTCA CACGGGCAAG ATCAGTTACC
     AAATTTGAAC GGGTATAAGT TTTGAGATAT GGGCCTGCGC TATAGCAAGT CGTTTGTTAA CTGATCGTAG GCAATGGAGT GTGCCCGTTC TAGTCAATGG
     R  T  R  H   A  Q  T   L  T  Q   D  Q  K  P   H  S  E   P  P  V   T  A  H  E   P  Q  A   R  R  T   D  V  I  V
.
2301 CGAACACGCC ACGCTCAGAC CCTGACGCAG GACCAAAAGC CACATTCGGA ACCGCCAGTG ACCGCGCATG AGCCGCAGGC ACGACGTACG GACGTTATAG
     GCTTGTGCGG TGCGAGTCTG GGACTGCGTC CTGGTTTTCG GTGTAAGCCT TGGCGGTCAC TGGCGCGTAC TCGGCGTCCG TGCTGCATGC CTGCAATATC
    · H  Y  A   G  T  L   P  E  R  A   L  L  E   K  V  R   H  I  N  V   A  L  T   R  H  T   N  A  L  Y   I  R  D
.
2401 TGCATTACGC CGGCACTTTA CCCGAAAGGG CACTGTTAGA GAAGGTGCGG CATATAAACG TCGCGTTGAC TCGGCACACA AACGCCCTAT ATATCAGGGA
     ACGTAATGCG GCCGTGAAAT GGGCTTTCCC GTGACAATCT CTTCCACGCC GTATATTTGC AGCGCAACTG AGCCGTGTGT TTGCGGGATA TATAGTCCCT
    · E  S  E   K  G  E  L   V  P  S   L  M  T   P  P  S  W   S  T  Y   R  C  T   P  V  D  K   Q  M  V   P  D  P
2501 CGAAAGTGAA AAAGGAGAGT TGGTACCTTC ATTAATGACA CCGCCAAGCT GGAGCACTTA TCGGTGCACC CCCGTTGACA AGCAAATGGT ACCGGATCCG
     GCTTTCACTT TTTCCTCTCA ACCATGGAAG TAATTACTGT GGCGGTTCGA CCTCGTGAAT AGCCACGTGG GGGCAACTGT TCGTTTACCA TGGCCTAGGC
     V  A  V  E   R  E  N   G  S  S   G  P  C  D   S  H  H   I  G  A   I  T  I  L   Q  E  L   G  K  L   T  D  T  K
.
2601 GTTGCAGTGG AGCGAGAGAA CGGATCGTCT GGTCCGTGTG ACTCCCACCA TATCGGCGCG ATTACTATAT TGCAAGAGCT CGGCAAATTA ACGGATACGA
     CAACGTCACC TCGCTCTCTT GCCTAGCAGA CCAGGCACAC TGAGGGTGGT ATAGCCGCGC TAATGATATA ACGTTCTCGA GCCGTTTAAT TGCCTATGCT
    · G  V  R   V  F  E   S  E  A  V   P  T  A   H  R  R   V  V  L  D   G  N  L   D  S  G   P  D  R  Y   P  M  Y
.
2701 AAGGCGTACG AGTATTTGAA TCCGAAGCCG TCCCAACCGC TCACCGGCGC GTAGTGCTTG ACGGCAACCT CGATTCAGGG CCCGATCGTT ACCCGATGTA
     TTCCGCATGC TCATAAACTT AGGCTTCGGC AGGGTTGGCG AGTGGCCGCG CATCACGAAC TGCCGTTGGA GCTAAGTCCC GGGCTAGCAA TGGGCTACAT
    · Q  F  T   N  L  R  G   T  K  Y   T  N  I   K  D  N  Q   Q  A  L   H  T  L   V  G  R  Y   A  R  K   I  N  S
2801 TCAGTTCACT AACCTCCGCG GGACCAAATA CACGAATATC AAGGACAACC AACAAGCGTT GCATACGCTC GTCGGCCGGT ATGCACGCAA GATAAACAGC
```

```
     AGTCAAGTGA TTGGAGGCGC CCTGGTTTAT GTGCTTATAG TTCCTGTTGG TTGTTCGCAA CGTATGCGAG CAGCCGGCCA TACGTGCGTT CTATTTGTCG
     S  S  R  E   D  A  E   F  D  V   K  R  I  T   A  R  L   K  E  W   I  P  F  R   T  A  E   P  E  Q   V  D  S  C
.
2901 TCGAGCCGAG AGGACGCCGA GTTTGACGTT AAGAGAATCA CAGCCAGGCT CAAGGAATGG ATTCCTTTTA GAACGGCAGA GCCCGAGCAA GTCGACAGTT
     AGCTCGGCTC TCCTGCGGCT CAAACTGCAA TTCTCTTAGT GTCGGTCCGA GTTCCTTACC TAAGGAAAAT CTTGCCGTCT CGGGCTCGTT CAGCTGTCAA
    ·  F  A  D   A  M  Q   K  I  A  E   R  G  H   G  V  D   D  I  E  D   F  W  S   N  E  G   Q  R  I  S   Y  H  L
.
3001 GCTTTGCCGA CGCCATGCAA AAGATCGCCG AACGCGGCCA TGGCGTCGAT GACATCGAGG ACTTCTGGTC GAACGAAGGC CAAAGAATTT CTTACCACCT
     CGAAACGGCT GCGGTACGTT TTCTAGCGGC TTGCGCCGGT ACCGCAGCTA CTGTAGCTCC TGAAGACCAG CTTGCTTCCG GTTTCTTAAA GAATGGTGGA
    · K  G  Q   Q  K  V  M   D  P  T   K  L  K   L  G  Q  G   I  S  A   H  E  K   C  A  N  I   A  L  S   A  W  V
3101 TAAGGGCCAG CAAAAAGTCA TGGACCCCAC CAAACTGAAA CTTGGACAAG GTATCTCCGC GCATGAAAAA TGCGCTAACA TTGCCCTCAG CGCGTGGGTG
     ATTCCCGGTC GTTTTTCAGT ACCTGGGGTG GTTTGACTTT GAACCTGTTC CATAGAGGCG CGTACTTTTT ACGCGATTGT AACGGGAGTC GCGCACCCAC
     R  I  I  Q   D  Q  M   S  T  S   E  K  F  I   F  A  N   G  Q  S   D  R  D  T   M  S  I   I  E  A   R  L  Q  E
.                                                                                                    NβV Rep2
3201 AGGATTATCC AAGATCAGAT GAGCACGTCA GAGAAGTTCA TCTTCGCGAA TGGGCAGTCA GACCGCGATA CCATGTCTAT CATTGAGGCA CGCCTGCAGG
     TCCTAATAGG TTCTAGTCTA CTCGTGCAGT CTCTTCAAGT AGAAGCGCTT ACCCGTCAGT CTGGCGCTAT GGTACAGATA GTAACTCCGT GCGGACGTCC
    · K  A  R   E  F  K   S  I  D  I   K  E  F   D  T  V   H  N  W  V   S  I  L   V  F  S   W  R  C  D   R  G  C
.
3301 AGAAGGCGCG GGAATTCAAA TCTATAGATA TCAAGGAGTT CGATACGGTA CATAACTGGG TCAGTATTCT TGTCTTCTCG TGGCGTTGCG ACCGTGGGTG
     TCTTCCGCGC CCTTAAGTTT AGATATCTAT AGTTCCTCAA GCTATGCCAT GTATTGACCC AGTCATAAGA ACAGAAGAGC ACCGCAACGC TGGCACCCAC
    · P  E  H   L  I  E  Y   F  E  K   R  S  K   S  R  T  L   S  S  R   I  G  S   V  D  V  S   F  M  L   D  S  G
3401 CCCAGAGCAC CTTATCGAGT ATTTCGAGAA ACGCTCGAAA AGCCGGACGC TCTCAAGCCG CATAGGAAGC GTCGACGTTA GCTTCATGCT CGATTCTGGC
     GGGTCTCGTG GAATAGCTCA TAAAGCTCTT TGCGAGCTTT TCGGCCTGCG AGAGTTCGGC GTATCCTTCG CAGCTGCAAT CGAAGTACGA GCTAAGACCG
     A  V  W  T   I  A  R   N  T  L   F  A  S  G   L  M  L   A  L  F   V  G  V  D   F  I  A   A  K  G   D  D  V  F
.
3501 GCTGTCTGGA CCATTGCCAG AAACACCTTA TTTGCCTCGG GTCTTATGCT CGCCCTTTTC GTCGGCGTCG ATTTCATCGC GGCGAAAGGC GATGATGTCT
     CGACAGACCT GGTAACGGTC TTTGTGGAAT AAACGGAGCC CAGAATACGA GCGGGAAAAG CAGCCGCAGC TAAAGTAGCG CCGCTTTCCG CTACTACAGA
    · L  A  G   N  N  L   Y  L  D  A   E  R  L   R  M  G   S  Y  L  A   A  N  N   L  K  I   E  K  T  A   V  V  S
.
3601 TCCTCGCAGG GAATAATTTG TACTTGGACG CAGAACGGCT TCGCATGGGA TCTTACTTAG CCGCAAACAA CTTGAAGATC GAGAAGACGG CGGTCGTGAG
    AGGAGCGTCC CTTATTAAAC ATGAACCTGC GTCTTGCCGA AGCGTACCCT AGAATGAATC GGCGTTTGTT GAACTTCTAG CTCTTCTGCC GCCAGCACTC
    · F  I  G   F  I  V  S   Q  A  A   V  T  A   D  V  V  R   L  A  T   R  T  Y   G  R  S  Y   K  N  A   D  D  L
3701 CTTTATAGGG TTTATCGTTT CCCAAGCCGC CGTCACAGCT GATGTCGTGC GTCTAGCCAC CCGGACTTAC GGTCGAAGTT ATAAAAACGC TGATGATCTA
    GAAATATCCC AAATAGCAAA GGGTTCGGCG GCAGTGTCGA CTACAGCACG CAGATCGGTG GGCCTGAATG CCAGCTTCAA TATTTTTGCG ACTACTAGAT
     A  K  Y  K   I  A  I   A  D  H   C  K  L  F   R  S  P   R  T  R   L  M  T  A   I  N  C   A  T  L   Y  G  T  S
.
```

```
3801 GCGAAGTATA AGATAGCTAT CGCTGACCAC TGCAAGTTGT TTAGATCACC GAGAACTCGT CTCATGACCG CGATCAACTG CGCCACCCTT TACGGCACCT
     CGCTTCATAT TCTATCGATA GCGACTGGTG ACGTTCAACA AATCTAGTGG CTCTTGAGCA GAGTACTGGC GCTAGTTGAC GCGGTGGGAA ATGCCGTGGA
     · K  E  C  I  N  Y  L  M  D  A  L  D  A  F  G  H  T  K  M  S  D  L  H  L  D  P  G  F  V  M  R  V  T

3901 CGAAAGAGTG CATCAATTAT CTGATGGACG CGTTGGACGC ATTCGGACAC ACTAAAATGA GCGACCTACA CTTGGATCCC GGTTTTGTCA TGCGGGTCAC
     GCTTTCTCAC GTAGTTAATA GACTACCTGC GCAACCTGCG TAAGCCTGTG TGATTTTACT CGCTGGATGT GAACCTAGGG CCAAAACAGT ACGCCCAGTG
     · P  M  K  V  D  E  R  V  Y  S  G  Q  D  G  C  Q  R  A  D  K  T  R  E  K  Q  P  E  P  G  Q  P  G  P
                                                M  D  A  N  V  Q  I  R  P  A  R  N  N  P  S  Q  G  N  Q  G  R

                                                NβV Cap F1
4001 CCCCATGAAG GTGGACGAGC GGGTTTATTC CGGACAAGAT GGATGCCAAC GTGCAGATAA GACCCGCGAG AAACAACCCG AGCCAGGGCA ACCAGGGCCG
     GGGGTACTTC CACCTGCTCG CCCAAATAAG GCCTGTTCTA CCTACGGTTG CACGTCTATT CTGGGCGCTC TTTGTTGGGC TCGGTCCCGT TGGTCCCGGC
     Q  Q  Q  Q  Q  A  S  T  Q  E  A  G  S  K  T  S  P  R  S  R  T  D  Y  Q  P  R  P  D  G  R  T  R  E  P

                                                                                NβV Cap F2
     · N  N  N  N  K  R  R  R  R  R  R  G  L  K  L  P  P  V  V  A  P  I  T  S  P  G  Q  M  A  E  P  A  N
4101 CAACAACAAC AACAAGCGTC GACGCAGGAG GCGGGGTCTA AAACTTCCCC CCGTAGTCGC ACCGATTACC AGCCCCGGCC AGATGGCCGA ACCCGCGAAC
     GTTGTTGTTG TTGTTCGCAG CTGCGTCCTC CGCCCCAGAT TTTGAAGGGG GGCATCAGCG TGGCTAATGG TCGGGGCCGG TCTACCGGCT TGGGCGCTTG
     · R  E  H  P  G  Q  P  R  S  D  T  R  E  G  V  K  A  S  D  D  G  E  S  H  G  S  D  I  R  G  M  D  S

     H  A  N  T  R  V  N  R  G  R  T  R  V  R  G  L  R  Q  A  M  M  E  S  P  M  A  A  T  S  E  A  W  I  H

4201 CACGCGAACA CCCGGGTCAA CCGCGGTCGG ACACGCGTGA GGGGGTTAAG GCAAGCGATG ATGGAGAGTC CCATGGCAGC GACATCAGAG GCATGGATTC
     GTGCGCTTGT GGGCCCAGTT GGCGCCAGCC TGTGCGCACT CCCCCAATTC CGTTCGCTAC TACCTCTCAG GGTACCGTCG CTGTAGTCTC CGTACCTAAG
     · R  L  S  R  P  G  R  R  I  Q  D  E  P  G  R  R  E  N  S  R  R  R  D  T  S  V  N  M  R  S  I  S  R
     · D  Y  L  D  P  D  G  E  Y  K  T  S  L  D  D  G  K  I  P  D  G  A  I  P  Q  S  T  C  G  Q  F  R  G

     NβV Cap 1
4301 ACGACTATCT AGACCCGGAC GGAGAATACA AGACGAGCCT GGACGACGGG AAAATTCCCG ACGGCGCGAT ACCTCAGTCA ACATGCGGTC AATTTCGAGG
     TGCTGATAGA TCTGGGCCTG CCTCTTATGT TCTGCTCGGA CCTGCTGCCC TTTTAAGGGC TGCCGCGCTA TGGAGTCAGT TGTACGCCAG TTAAAGCTCC
     D  R  G  R  Q  I  P  G  T  E  F  Y  D  A  T  A  G  W  R  D  L  A  S  T  S  D  A  S  P  V  L  Q  A  S

     · T  V  G  A  R  Y  P  G  L  N  S  T  T  L  P  L  D  G  G  T  W  P  L  L  V  M  H  L  P  F  F  R  H
4401 GACCGTGGGC GCCAGATACC CGGGACTGAA TTCTACGACG CTACCGCTGG ATGGCGGGAC CTGGCCTCTA CTAGTGATGC ATCTCCCGTT CTTCAGGCAT
     CTGGCACCCG CGGTCTATGG GCCCTGACTT AAGATGCTGC GATGGCGACC TACCGCCCTG GACCGGAGAT GATCACTACG TAGAGGGCAA GAAGTCCGTA
     · V  V  V  H  H  H  H  Q  Q  H  G  S  R  S  D  E  R  R  S  G  C  V  R  E  R  L  E  Q  Q  D  G  L  D

     P  L  L  F  I  T  T  T  S  N  T  E  V  E  V  T  N  A  D  L  D  A  F  A  N  D  W  N  N  R  T  D  W  T
```

```
  4501 CCGTTGTTGT TCATCACCAC CACCAGCAAC ACGGAAGTCG AAGTGACGAA CGCCGATCTG GATGCGTTCG CGAACGATTG GAACAACAGG ACGGACTGGA
       GGCAACAACA AGTAGTGGTG GTGGTCGTTG TGCCTTCAGC TTCACTGCTT GCGGCTAGAC CTACGCAAGC GCTTGCTAAC CTTGTTGTCC TGCCTGACCT
       · R  S  D   V  P  K  L   G  A  S   R  E  R   V  L  H  G   R  P  D   R  S  A   D  G  R  T   T  P  D   S  T  G
       · E  A  T   Y  P  S  W   A  Q  V   G  N  V   F  Y  M  V   V  P  T   E  A  L   T  D  V  P   P  P  P   T  Q  L  G
.
  4601 CCGAAGCGAC GTACCCAAGT TGGGCGCAAG TCGGGAACGT GTTTTACATG GTCGTCCCGA CCGAAGCGCT GACGGACGTA CCACCCCCGA CTCAACTGGG
       GGCTTCGCTG CATGGGTTCA ACCCGCGTTC AGCCCTTGCA CAAAATGTAC CAGCAGGGCT GGCTTCGCGA CTGCCTGCAT GGTGGGGGCT GAGTTGACCC
       C  I  R  V   T  R  E   L  P  S   D  I  E  R   R  H  S   V  L  Q   R  T  H  S   R  E  S   G  S  G   G  D  R  A
.
       · V  S  G   L  L  E  S   Y  R  L   T  S  S   G  V  T  A   Y  F  N   A  P  T   L  V  N  Q   G  V  A   V  I  A
  4701 TGTATCAGGG TTACTCGAGA GTTACCGTCT GACATCGAGC GGCGTCACAG CGTACTTCAA CGCACCCACT CTCGTGAATC AGGGAGTGGC GGTGATCGCG
       ACATAGTCCC AATGAGCTCT CAATGGCAGA CTGTAGCTCG CCGCAGTGTC GCATGAAGTT GCGTGGGTGA GAGCACTTAG TCCCTCACCG CCACTAGCGC
       · V  P  T   G  Q  R   T  P  E  G   E  P  G   H  S  S   R  D  H  P   N  G  R   N  V  T   A  R  R  F   R  A  E
.
       Q  F  Q  P   D  K  E   H  Q  K   E  N  P  D   I  V  A   G  T  T   Q  T  G  G   T  L  Q   L  G  G   S  G  P  N
.
  4801 CAGTTCCAAC CGGACAAAGA ACACCAGAAG GAGAACCCGG ACATAGTAGC CGGGACCACC CAAACGGGCG GAACGTTACA GCTCGGCGGT TCAGGGCCGA
       GTCAAGGTTG GCCTGTTTCT TGTGGTCTTC CTCTTGGGCC TGTATCATCG GCCCTGGTGG GTTTGCCCGC CTTGCAATGT CGAGCCGCCA AGTCCCGGCT
       · L  H  I   D  D  D  D   R  G  P   G  R  V   R  G  R  S   N  P  A   T  H  G   V  D  G  A   D  A  G   V  G  A
       · Y  T  L   T  M  T  I   G  D  Q   V  E  F   G  G  A  A   I  P  L   P  T  V   S  M  G   P  M  P  E   S  G  Q
.
  4901 ACTACACATT GACGATGACG ATCGGGGACC AGGTCGAGTT CGGGGGCGCA GCAATCCCGC TACCCACGGT GTCGATGGGG CCGATGCCGG AGTCGGGGCA
       TGATGTGTAA CTGCTACTGC TAGCCCCTGG TCCAGCTCAA GCCCCCGCGT CGTTAGGGCG ATGGGTGCCA CAGCTACCCC GGCTACGGCC TCAGCCCCGT
       A  G  V  P   D  C  E   P  D  I   R  R  R  K   H  N  H   H  H  D   H  A  A  T   R  V  G   D  G  N   V  A  I  H
.
       · L  V  F   Q  T  A  N   L  T  F   D  V  G   N  T  I  T   I  T  T   T  L  P   P  G  S  V   T  G  M   W  Q  F
  5001 GCTGGTGTTC CAGACTGCGA ACCTGACATT CGACGTCGGA AACACAATCA CCATCACGAC CACGCTGCCA CCAGGGTCGG TGACGGGAAT GTGGCAATTC
       CGACCACAAG GTCTGACGCT TGGACTGTAA GCTGCAGCCT TTGTGTTAGT GGTAGTGCTG GTGCGACGGT GGTCCCAGCC ACTGCCCTTA CACCGTTAAG
       · S  Q  Q   R  D  G   H  R  D  R   G  R  G   S  A  T   V  R  V  R   S  E  F   G  R  L   G  T  E  S   A  G  H
.
       T  A  S  N   G  T  D   T  V  T   V  D  A  G   A  R  L   Y  A  F   G  A  N  L   D  A  S   E  L  N   L  Q  D  I
.
  5101 ACAGCCAGCA ACGGGACGGA CACCGTGACC GTGGACGCGG GAGCGCGACT GTACGCGTTC GGAGCGAATT TGGACGCCTC GGAACTGAAT CTGCAGGACA
       TGTCGGTCGT TGCCCTGCCT GTGGCACTGG CACCTGCGCC CTCGCGCTGA CATGCGCAAG CCTCGCTTAA ACCTGCGGAG CCTTGACTTA GACGTCCTGT
       · Q  L  N   Q  D  S  T   N  E  H   E  P  N   D  A  G  N   A  Q  D   H  S  V   P  T  Q  R   N  E  G   L  L  Y
```


CW R1
NβV Cap R

```
5901 TTCCAGTCAG CAGCTCTTCG GTTTCATCTC CACTGACGAC CCTCTGCACG TATTTTCCAT CATCGTGGTG CTGATCATCA TTTGCGGTCT GGTTTGGCTC
     AAGGTCAGTC GTCGAGAAGC CAAAGTAGAG GTGACTGCTG GGAGACGTGC ATAAAAGGTA GTAGCACCAC GACTAGTAGT AAACGCCAGA CCAAACCGAG
6001 GCTTCCTATT GTTTTCATGG CAGATCAACT CGACCCAATC GAAATCTTCA TCGAATTCCT CGGATTTCGA AAGATTTCCG GAGCCATTCG CTGCTTCCAG
     CGAAGGATAA CAAAAGTACC GTCTAGTTGA GCTGGGTTAG CTTTAGAAGT AGCTTAAGGA GCCTAAAGCT TTCTAAAGGC CTCGGTAAGC GACGAAGGTC
6101 TGTAAGCAAG TCTTGGGACC TGCAGGACCC GTGGATTCCC ACCTCACTGA TTGTCAGTGG AAGGCCGCGG TGTTAAATCT TATTAACAAT CAGCTTTACG
     ACATTCGTTC AGAACCCTGG ACGTCCTGGG CACCTAAGGG TGGAGTGACT AACAGTCACC TTCCGGCGCC ACAATTTAGA ATAATTGTTA GTCGAAATGC
6201 ACGTCGATAT CGATGAGACG AATCCGTTTC TTTACGGACC TCACCGCGAC TGAGATGTGG AAGACCACAT TGTTCCTTCA CACATGCCCA GAGTGCGGTT
     TGCAGCTATA GCTACTCTGC TTAGGCAAAG AAATGCCTGG AGTGGCGCTG ACTCTACACC TTCTGGTGTA ACAAGGAAGT GTGTACGGGT CTCACGCCAA
6301 ATTCCACCAG GGACACAGAA ACTACGAGAT CGTGTCCCCG AGATTGCCAA GACGGCAATC TTATGCACGC ATCTTCGGTC GGCTATATTT GTCACAAATG
     TAAGGTGGTC CCTGTGTCTT TGATGCTCTA GCACAGGGGC TCTAACGGTT CTGCCGTTAG AATACGTGCG TAGAAGCCAG CCGATATAAA CAGTGTTTAC
6401 CCGGTTAGAA GCAAACACAT TTTACCACGG TTTATGCTCT CAGTGTCGCG ACCGTGATAA TAAAAAACGA CGCTGAAGAG AGGACTCACA ACTACCTCGA
     GGCCAATCTT CGTTTGTGTA AAATGGTGCC AAATACGAGA GTCACAGCGC TGGCACTATT ATTTTTTGCT GCGACTTCTC TCCTGAGTGT TGATGGAGCT
6501 TCTCGTTTAT CGGACGAGTG ATACAATTGA CCCAGGGTCA TCCTGCAAAA CACGCAGGTT TCCGATAGTG GTGCAAATCC ACCCGCCAGT CGTCGGTGGT
     AGAGCAAATA GCCTGCTCAC TATGTTAACT GGGTCCCAGT AGGACGTTTT GTGCGTCCAA AGGCTATCAC CACGTTTAGG TGGGCGGTCA GCAGCCACCA
6601 CCCTTGCGGG ACCTATACGG TACCA
     GGGAACGCCC TGGATATGCC ATGGT
```

Figure A.1. Nucleotide sequence of NβV genome with amino acid sequence of both the replicase and capsid genes above this sequence. Complementary sites of the oligonucleotides used are underlined and the PCR product consisting of the capsid gene coding sequence is underlined with the thick black line.