Virion composition and genomics of White Spot Syndrome Virus of shrimp

M. C. W. van Hulten

Promotoren:	prof. dr. J. M. Vlak Persoonlijk Hoogleraar bij de Leerstoelgroep Virologie
	prof. dr. R. W. Goldbach Hoogleraar in de Virologie

Samenstelling promotiecommissie:

dr. J. R. Bonami (University of Montpellier II, France)
dr. E. O. Rijke (Intervet International, Nederland)
prof. dr. W. J. Stiekema (Plant Research International, Wageningen)
prof. dr. J. A. Verreth (Wageningen Universiteit)

Virion composition and genomics of White Spot Syndrome Virus of shrimp

M. C. W. van Hulten

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof. dr. ir. L. Speelman,
in het openbaar te verdedigen
op vrijdag 16 november 2001
des namiddags te 13.30 uur in de Aula.

The work presented in this thesis was carried out at the Laboratory of Virology of Wageningen University, The Netherlands. It was financially supported by Intervet International BV, Boxmeer, The Netherlands.

Van Hulten, M. C. W.

Virion composition and genomics of White Spot Syndrome Virus of shrimp
Thesis Wageningen University – with ref. – with summary in Dutch
ISBN 90-5808-516-3
Subject headings: Nimaviridae, *Penaeus monodon*, WSSV genome

Aan mijn ouders

Contents

Chapter 1

General introduction

Penaeus monodon

The black tiger shrimp, *Penaeus monodon* is the largest and commercially most important species among penaeids, reaching 300 mm in body length (Fig. 1.1). *P. monodon* occurs mainly in Southeast Asian waters. Its carapace and abdomen are transversely banded with red and white. However, when cultured in ponds, the color changes to blackish, hence this shrimp species is often referred to as black tiger shrimp (Motoh, 1985). Mating and spawning of *P. monodon* generally takes place at night. The maximum number of eggs spawned at a time can be up to 800,000 per individual. The eggs hatch within 12 hours into nauplii, the first larval stage. The complete larval stage of *P. monodon* consists of 6 nauplius, 3 zoea, 3 mysis, and 3 or 4 postlarva substages (Fig. 1.2) and the time required for each of these substages is about 1.5 days, 5 days, 5 days and 6 to 15 days, respectively. The next stage is the juvenile stage, where brackish water areas serve as nursery grounds and the shrimp develop in subadults in about 4 months. The subadult stage starts at the onset of sexual maturity while the shrimp migrate from nursery to spawning grounds (Fig 1.2). After 4 months the adult stage is reached, which is characterized by the completion of sexual maturity (Motoh, 1985). The food of *P. monodon* consists mainly of crustacea (small crabs and shrimps) and molluscs. Other commercially cultured shrimp species are *P. vannamei* (Pacific white shrimp), *P. japonicus* (kuruma shrimp), *P. stylirostris* (blue shrimp), and *P. chinensis* (Chinese white shrimp) of which the first two species are mainly cultured on the Pacific coast of Central and South America.

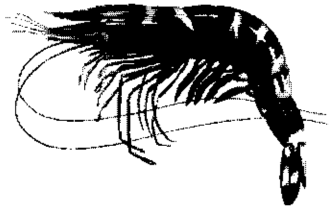


Figure 1.1 Drawing of *Penaeus monodon*, the "black tiger shrimp".

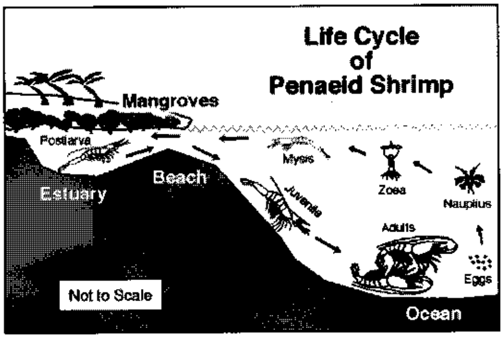


Figure 1.2 Schematic representation of the life cycle of Penaeid shrimp (Rosenberry, 2000).

Shrimp culture and viral diseases

Shrimp farming has its origins in Southeast Asia where for centuries farmers have raised incidental crops of wild shrimp in tidal fish ponds (Rosenberry, 1994). The collection of wild seed in the 1970's and the advances in shrimp reproduction and hatchery technologies in the

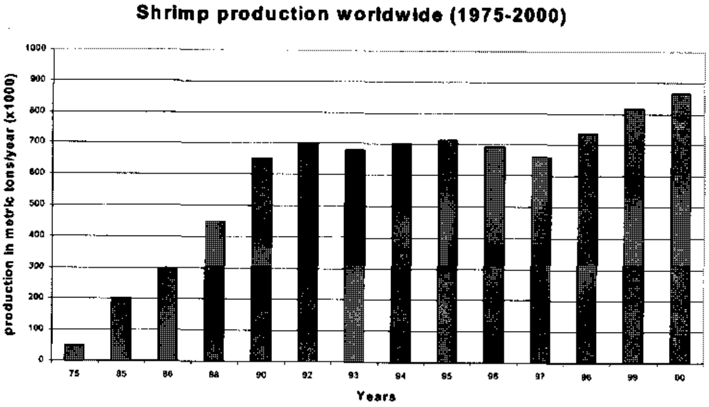


Figure 1.3 Shrimp production from 1975 to 2000.

1980s resulted in a rapid development of shrimp farming. In the 1990s, the world's production of cultured shrimp increased less dramatically and plateaud at around 700,000 tonnes per year (Fig. 1.3) (Rosenberry, 1995, 1996, 1997, 1999, 2000). The growth stagnated despite an increasing number of hectares used for production. This was caused by mass mortalities in pond culture throughout the world, predominantly caused by viruses (Lotz, 1997). In particular in Asia shrimp culture was heavily afflicted, resulting in a decrease of shrimp production three years in a row from 1995 to 1997 (Rosenberry, 1997).

Penaeid shrimp can be affected by many different infectious agents, including viruses, rickettsiae, gram-negative and gram-positive bacteria, fungi and protozoa. The most important diseases of cultured penaeid shrimp have viral or bacterial etiologies. Since 1995 viruses have caused the greatest impact on shrimp culture, as no treatment for these agents is available yet, whereas bacteria can be controlled by antibiotics and vaccines (Teunissen *et al.*, 1998). The major diseases of cultured shrimp caused by viruses are described in table 1.1. Baculovirus penaei (BP) and Monodon baculovirus (MBV) are rod-shaped viruses, which are occluded in large proteinaceous capsules (polyhedra) and are therefore thought to be members of the *Baculoviridae* (Bonami *et al.*, 1995a; Mari *et al.*, 1993). MBV is enzootic in Southeast Asia in both wild and farmed *P. monodon* populations, and has been reported to cause major mortalities in Asia in the late 1980s (Mari *et al.*, 1993). Baculovirus midgut gland necrosis virus (BMNV) has mainly been reported in *P. japonicus* in Japan and generally infects shrimp larvae (Sano *et al.*, 1981; Momoyama and Sano, 1996). Not much research has been performed on this virus and therefore it is not clear what virus family this virus belongs to (Table 1.1). Infectious hypodermal and hematopoietic necrosis virus (IHHNV) is a highly contagious and infectious virus for many penaeid species (Bonami *et al.*, 1990). As the overall genomic organization is similar to that of the mosquito brevidensoviruses, it was concluded that IHHNV is closely related to densoviruses of the genus *Brevidensovirus* in the family *Parvoviridae* (Shike *et al.*, 2000). Hepatopancreatic parvovirus (HPV) causes disease in several species of penaeid shrimp and analysis of its genome indicated that this virus is most likely a member of the family *Parvoviridae* (Sukhumsirichart *et al.*, 1999). The two RNA viruses indicated in table 1.1, Taura syndrome virus (TSV) and Yellow head virus (YHV) both have a great impact on shrimp culture and have been studied extensively. TSV has caused serious mortalities in shrimp (*P. vannamei*) in the Americas (Lightner *et al.*, 1995). Research on the TSV genome organization showed that TSV is similar to insect picornaviruses (Robles-Sikisaka *et al.*, 2001). YHV had it greatest impact in Asian countries (Boonyaratpalin *et al.*, 1993). In Australia the closely related Gill-associated virus (GAV), and its minor variant Lymphoid organ virus (LOV) have been identified (Cowley *et al.*, 1999). Research on the ssRNA genome revealed that these viruses are members of the order *Nidovirales* and further research will be needed to resolve their taxonomic classification within this order (Cowley *et al.*, 2000).

Virus name	**Virus family**	**Reference**	**Detection**
DNA viruses			
Baculovirus penaei (BP)	*Baculoviridae*	Bonami *et al.*, 1995	Bonami *et al.*, 1995; Wang *et al.*, 1996
Monodon baculovirus (MBV)	*Baculoviridae*	Vickers *et al.*, 2000; Mari *et al.*, 1993	Belcher and Young, 1998; Hsu *et al.*, 2000
Baculovirus midgut gland necrosis virus (BMNV)	Unclassified	Momoyama and Sano, 1996	
Infectious hypodermal and hematopoietic necrosis virus (IHHNV)	*Parvoviridae*	Shike *et al.*, 2000	Tang and Lightner, 2001; Nunan *et al.*, 2000
Hepatopancreatic parvovirus (HPV)	*Parvoviridae*	Sukhumsirichart *et al.*, 1999; Bonami *et al.*, 1995	Pantoja and Lightner, 2001; Pantoja and Lightner, 2000
White spot syndrome virus (WSSV)	New virus family	van Hulten *et al.*, 2001a; Wang *et al.*, 1995	Lo *et al.*, 1996b; Tapay *et al.*, 1999; Tang and Lightner, 2000; Shih *et al.*, 2001
RNA viruses			
Taura syndrome virus (TSV)	*Picornaviridae*	Robles-Sikisaka *et al.*, 2001; Bonami *et al.*, 1997	Nunan *et al.*, 1998; Mari *et al.*, 1998
Yellow head virus (YHV)	*Coronaviridae* or *Arteriviridae* (order: *Nidovirales*)	Cowley *et al.*, 2000; Cowley *et al.*, 1999	Tang and Lightner, 1999; Sithigorngul *et al.*, 2000; Nadala and Loh, 2000

Table 1.1 Major viral diseases of cultured Penaeid shrimp. In the first column, the virus name. In the second column, the (putative) virus family and in the third column the reference for this classification. References describing detection techniques in the last column.

White Spot Syndrome Virus

From the viruses mentioned in table 1.1, White spot syndrome virus (WSSV) has had the greatest impact on shrimp culture and remains a major problem up to the present day (Rosenberry, 2000). In cultured shrimp WSSV infection can reach a cumulative mortality of up to 100% within 3-10 days (Lightner, 1996). The virus was first discovered in Taiwan in 1992, from where it quickly spread to other shrimp farming areas in Southeast Asia and the

Indian subcontinents (Cai *et al.*, 1995). WSSV initially appeared to be limited to Asia until it was found in Texas in 1995 (Rosenberry, 1996) (Fig. 1.4). The contamination of the Texas shrimp farm is believed to be caused by untreated wastes from plants processing imported shrimp from Asia (Naylor *et al.*, 2000). In 1997 the virus was detected in South Carolina. In early 1999 WSSV was also reported from Central- and South-America and it continued to cause major losses in 2000, when it was also found in Mexico (Rosenberry, 2000).

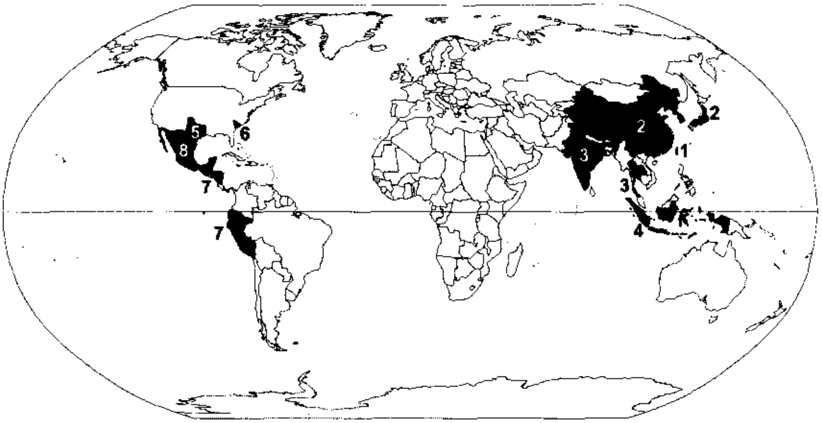


Figure 1.4 Worldwide spread of WSSV. In 1992 (1) the first incidence of WSSV was reported in Taiwan. From there it spread through Asia in 1993 to China, Japan and Korea (2), in 1994 to Thailand and India (3) and in 1995 to Indonesia (4). In 1995 the first outbreak occurred in the Western hemisphere in Texas (5) from where it spread in 1997 to South Carolina (6), in 1999 to Guatemala, Honduras, Nicaragua, Panama, Ecuador and Peru (7) and in 2000 to Mexico (8).

Over the last decade many scientific reports on WSSV appeared, probably describing the same agent, but using several different names. "Systemic ectodermal and mesodermal baculovirus" (SEMBV) was described from Thailand in 1993 (Wongteerasupaya *et al.*, 1995), "Rod-shaped virus of *Penaeus japonicus*" (RV-PJ), or "Penaeid rod-shaped DNA virus" (PRDV) from Japan in 1993 (Inouye *et al.*, 1994; Inouye *et al.*, 1996), "Hypodermal and hematopoietic necrosis baculovirus" (HHNBV) and "Chinese baculovirus" (CBV) in 1993 in China (Huang *et al.*, 1995; Nadala *et al.*, 1998), and "White spot baculovirus" (WSBV) from Taiwan in 1992 (Chou *et al.*, 1995; Wang *et al.*, 1995). Nowadays the name White spot syndrome virus (WSSV) is used by most research groups.

The clinical signs of WSSV are the presence of white spots in the exoskeleton of infected shrimp (Fig. 1.5a,b). Diseased shrimp are lethargic, have a lack of appetite and a reddish body coloration (Chou *et al.*, 1995). WSSV has an exceptional broad host range, as it does not only infect all shrimp species, but also other crustaceans including crab and crayfish (Wang *et al.*, 1998 ; Lo *et al.*, 1996a). For some of these crustaceans WSSV is not lethal and therefore these species can serve as a reservoir for the virus when they are introduced in the shrimp pond along with the sea water, making elimination of WSSV from shrimp ponds very difficult. Intensive shrimp cultivation, inadequate sanitation, poor chain management and worldwide trade further aggravated the disease incidence and enhanced disease dissemination. WSSV can be considered a major threat not only to shrimp, but also to other marine and fresh water crustaceans around the world. No penaeid shrimp species are known to be resistant to WSSV (Lotz, 1997) and for many shrimp species neither age nor size appears to be important for susceptibility to infection with WSSV.

Figure 1.5a Infected *P. monodon* showing white spots.

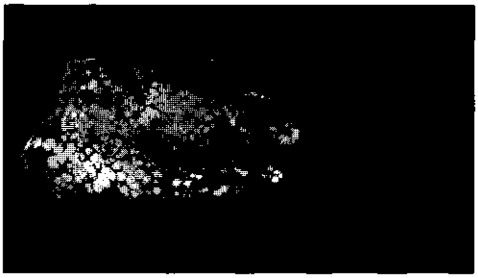


Figure 1.5b Carapace of infected *P. monodon*, showing white spots.

WSSV virion morphology

Electron microscopy studies on virus purified from the haemolymph of infected shrimp revealed that the WSSV virion is an ellipsoid to bacilliform shaped enveloped particle of about 275 nm in length and 120 nm in width. Most characteristic is the tail-like appendage at one end of the virion (Fig. 1.6a). WSSV nucleocapsids are rod-shaped and have a striated appearance and a size of about 300 nm x 70 nm (Fig 1.6b). Electron microscopical examinations revealed that the striations are probably formed by globular subunits (about 10 nm) which are arranged in 2 parallel rows (Durand *et al.*, 1997; Nadala *et al.*, 1998). Each nucleocapsid contains about 14 of these ring-like structures. The WSSV genome consists of a large double stranded DNA molecule of about 290 kilobase pair (kb) (Yang *et al.*, 1997).

Figure 1.6a Electron micrograph of enveloped WSSV virions.

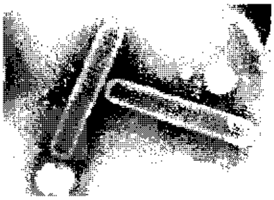


Figure 1.6b Electron micrograph of WSSV nucleocapsids.

WSSV histo- and cytopathology

In situ hybridization studies using WSSV specific probes showed that WSSV can infect shrimp tissues that originate from the ectoderm or the mesoderm (Lo *et al.*, 1997). WSSV targets predominately cuticular epithelium and subsequently connective tissues of some organs (Chang *et al.*, 1996). The virus severely damages the stomach, gills, subcuticular epithelial cells, lymphoid organ, antennal gland and hemocytes. As infection is often initially detected in gills and stomach, these organs are thought to be the entry sites of natural WSSV infection (Chang *et al.*, 1996). Histology on WSSV infected cells using hematoxylin and eosin staining shows hypertrophied nuclei with eosinophilic inclusion and marginated slightly basophilic chromatin in early stages of infection. In a later infection stage the inclusions have a light basophilic, denser appearance and are separated by a transparent zone from the marginated chromatin (Wongteerasupaya *et al.*, 1995; Wang *et al.*, 1999; Lightner, 1996).

Electron microscopy studies on WSSV infected animals show the presence of hypertrophied nuclei filled with virions (Fig. 1.7). WSSV replication takes place in the nucleus and is first indicated by chromatin margination and nuclear hypertrophy (Durand *et al.*, 1997; Wang *et al.*, 1999; Wang *et al.*, 2000a). Virus morphogenesis begins with the formation of fibrillar, viral envelopes which are formed *de novo* in the nucleoplasm. Empty nucleocapsids seem to be formed from segmented, empty, long tubules, which resolve into fragments of 12 to 14 rings in a stacked series. Subsequently, membranes envelop the capsids leaving an open extremity. The nucleoproteins, which have a filamentous appearance, enter the capsid through this open end. When the core is completely formed, the envelope narrows at the open end and forms the apical tail of the mature virion (Durand *et al.*, 1997; Wang *et al.*, 1999 Wang *et al.*, 2000a). It is not clear how the virions are released from the nuclei, but most likely by rupture of the nuclear envelope and cell membrane.

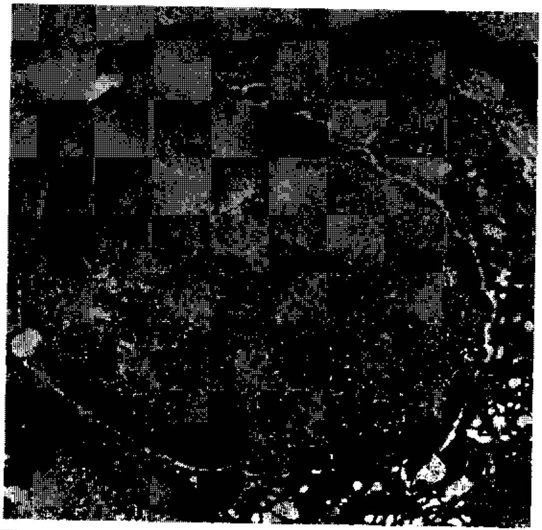


Fig. 1.7 Electron micrograph of WSSV infected cell in the gill of *Procambarus clarkii.*

WSSV diagnosis

Diagnostic methods for detection of shrimp pathogens include the traditional methods of the gross and clinical signs, morphological pathology (light microscopy, routine histopathology and histochemistry). The last decade more techniques have been developed for quick and reliable detection of shrimp viruses. These include *in situ* hybridization, polymerase chain reaction (PCR), and immunological detection. In table 1.1 examples of reported detection techniques are mentioned for the major shrimp viruses. For WSSV several detection techniques have been reported: *in situ* DNA hybridization (Wongteerasupaya *et al.*, 1996; Chang *et al.*, 1996; Durand *et al.*, 1996; Nunan and Lightner, 1997), PCR (Lo *et al.*, 1996b; Takahashi *et al.*, 1996; Kasornchandra *et al.*, 1998) and immunodetection (Nadala and

Loh, 2000; Zhang *et al.*, 2001). The WSSV DNA fragments used for detection (*in situ* hybridization) or amplified (PCR) are a collection of randomly selected segments of the WSSV genome. For the detection of WSSV and other major shrimp viruses, commercial detection kits are available (e.g. several kits from DiagXotics using *in situ* hybridization and immunodetection).

WSSV taxonomy

WSSV has several characteristics in common with members of the family of the *Baculoviridae*. *Baculoviruses* are rod shaped viruses with a large circular DNA genome and they infect insects. Their rod shaped nucleocapsid, nuclear replication and morphogenesis are features shared by WSSV (Federici, 1997; Durand *et al.*, 1997). However, WSSV does not produce occlusion bodies like the nucleopolyhedro- and granuloviruses (*Baculoviridae*). These proteinaceous structures protect the virions from decay in- and outside the host (Funk *et al.*, 1997). The distinctive features of WSSV led to the tentative assignment of WSSV as a member of the genus of the 'non-occluded' rod-shaped viruses (Nudibaculovirinae), a separate genus within the *Baculoviridae* (Franki *et al.*, 1991). In 1995 the non-occluded baculoviruses were eliminated from the *Baculoviridae* and orphaned (Murphy *et al.*, 1995). Only the *nucleopolyhedroviruses* (NPV) and *granuloviruses* (GV) are now recognized genera in this family (van Regenmortel *et al.*, 2000).

With respect to WSSV, the orphan status is justified, as many other characteristics of WSSV and its pathogenesis are distinct from those of baculoviruses. Most notable is the lack of occlusion bodies in WSSV infections. Although the shape and size of the WSSV virions are reminiscent of the baculoviruses, WSSV contains a tail-like appendage at one end of the particle. The WSSV nucleocapsids, although rod-shaped, have a crosshatched appearance and differ in this respect profoundly from the smooth surface of baculovirus nucleocapsids (Funk *et al.*, 1997). The size of the genome of WSSV was estimated to be about 290 kb, which is well out of the range of baculovirus genomes (100 –180 kb) (Yang *et al.*, 1997; Hayakawa *et al.*, 2000). Furthermore, it was not clear whether the DNA genome was linear or circular (Yang *et al.*, 1997). Finally, in contrast to baculoviruses, WSSV has an extremely wide host range including shrimp, crab and crayfish.

Control of WSSV

WSSV is a major threat to shrimp culture, but also to other salt and fresh water crustaceans and thus a threat to natural as well as aquaculture ecosystems. The necessity for early diagnosis of WSSV in cultured and wild populations of shrimp and other crustaceans lead to the development of many WSSV diagnostic probes for *in situ* hybridization and primers for detection by PCR (Wongteerasupaya *et al.*, 1996; Durand *et al.*, 1996; Chang *et al.*, 1996; Lo *et al.*, 1996b; Takahashi *et al.*, 1996; Nunan and Lightner, 1997). No adequate measures to control WSSV other than rigorous sanitation and adequate chain management practices are available yet. Due to the current aquaculture practices this is not a viable option in the foreseeable future. Control of viral diseases is of major importance for future sustainable development of the shrimp farming industry. Therefore, alternative strategies need to be explored, such as the use of vaccination or other invention strategies, to control WSSV.

Aim and outline of the thesis

At the onset of this PhD research (1996) the taxonomic position of WSSV was not resolved due to lack of molecular data. Research on WSSV mainly focussed on the clinical symptoms, histopathology and morphology using electron microscopy on thin-sections of infected animals. No information on the protein composition of the WSSV virion, on the DNA sequence or on the infection mechanism of WSSV in shrimp at the cellular and molecular level was available. A major hindrance in this research has been the lack of relatively simple, well defined systems, such as permissive cell cultures, in which WSSV can be propagated and studied. The goal of this thesis research was to characterize the WSSV virion and its genome in order to gain insight in the structure of the virus and its taxonomic position and to gain feasible starting points for the development of effective vaccines or intervention strategies.

The research described here started with a batch of frozen WSSV infected black tiger shrimps (*Penaeus monodon*) imported from Thailand, where WSSV had a devastating impact on shrimp culture in 1996. The first experiments performed focussed on the development of a robust protocol for the isolation and purification of intact virions from the infected shrimp. This allowed a detailed genetic and structural analysis of WSSV.

As a start the WSSV DNA was fragmented by several restriction enzymes and randomly cloned into bacterial plasmids and terminally sequenced. To study the taxonomical position of WSSV, phylogenetic analysis was performed on genes identified on the viral genome by random sequencing, such as the ribonucleotide reductase large and small subunit genes (**Chapter 2**) and the protein kinase genes (**Chapter 3**). Gene homologs from other large DNA virus families, and of eukaryotic and prokaryotic origin were included in these analyses. The phylogenetic program PAUP (Swofford, 1993) was used to construct unrooted parsimonious trees.

Other genes that are often well conserved amongst related viruses are those coding for viral capsid or envelope proteins (Murphy *et al.*, 1995). These genes are therefore often used to study interviral relatedness (Tidona *et al.*, 1998; Sullivan *et al.*, 1994; Dolja *et al.*, 1991; Dolja and Koonin, 1991). Studies on the WSSV virion performed by other research groups, described the presence of several major proteins located in the envelope and nucleocapsid. Their number and sizes varied depending on the isolate or the techniques used (Nadala *et al.*, 1998; Wang *et al.*, 2000b). The corresponding open reading frames (ORFs) were not yet identified on the WSSV genome. In **Chapter 4** two WSSV virion proteins were identified, the 28 kDa envelope protein (VP28) and the 26 kDa nucleocapsid protein (VP26). The N-terminal amino acid sequence of these proteins was used to locate the respective genes (*vp*28 and *vp*26) in the WSSV genome by direct sequencing (VP26) or by using degenerated primers and colony lifting (VP28).

The subsequent identification of the 24 kDa nucleocapsid protein (VP24) (**Chapter 5**) showed that VP28, VP26 and VP24 might have been derived from a common ancestor gene, and subsequently diverged over evolutionary time to give proteins with different functions (envelope and nucleocapsid). This is an unusual phenomenon in animal DNA viruses.

In **Chapter 6**, parts of the protein sequence of the other major envelope protein of 19 kDa (VP19) and the 15 kDa major nucleocapsid protein (VP15) were determined, leading to the location of their genes on the WSSV genome. Enveloped viruses of vertebrates and invertebrates contain glycoproteins in their viral envelopes and these often play important roles in the interaction between virus and host such as attachment to receptors and fusion with cell membranes (van Regenmortel *et al.*, 2000; Granof & Webster, 1999). Therefore, this chapter also describes the glycosylation status of the five WSSV virion proteins.

As VP28 is the major envelope protein, its role in infection was investigated *in vivo* using a neutralization assay (**Chapter 7**) with VP28 specific antibodies. Neutralization experiments have often been performed to study the role of virion proteins or their domains in the infection process (Burton *et al.*, 2000). These experiment are mostly performed using cell cultures (plaque reduction assays) (e.g. Galmiche *et al.*, 1999; Sunyach *et al.*, 1999; Volkman and Goldsmith, 1985). However, standardized (primary) shrimp cell cultures are not available and therefore an *in vivo* approach was followed to study the role of VP28 in the systemic WSSV infection of shrimp.

Despite the large amount of partial WSSV sequences available from the random sequencing, no homology with other large DNA viruses could be found. To gain insight in WSSV genome structure, the complete sequence of the DNA genome was determined and the genetic information analyzed revealing several unique features of this virus (**Chapter 8**).

In **Chapter 9** the presented results are discussed. To reveal the taxonomic position of WSSV, a consensus phylogenetic tree is constructed, using a set of genes. Furthermore, the WSSV genome is compared to the genome of a second isolate and several features of the virion are discussed.

Chapter 2

Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions

White spot syndrome is a worldwide disease of penaeid shrimp. The disease agent is a bacilliform enveloped virus, called white spot syndrome virus (WSSV), with a double stranded DNA genome of probably well over 200 kb. Analysis of a 12.3 kb segment of WSSV DNA revealed eight open reading frames (ORFs), including the genes for the large (RR1) and small (RR2) subunit of ribonucleotide reductase. The *rr1* and the *rr2* genes were separated by 5760 bp, containing several putative ORFs and two domains with multiple sequence repeats. The first domain contained six direct repeats of 54 bp and is part of a coding region. The second domain had one partial and two complete direct repeats of 253 bp at an intergenic location. This repeat, located immediately upstream of *rr1*, has homologues at several other locations on the WSSV genome. Phylogenetic analysis with RR1 and RR2 indicated that WSSV belongs to the eukaryotic branch of an unrooted parsimonious tree and further seems to suggest that WSSV and baculoviruses most likely do not share an immediate common ancestor. The present analysis of WSSV favors the view that this virus is either a member of a new genus within the Baculoviridae or a member of an entirely new virus family.

This chapter has been published as:

van Hulten, M. C. W., Tsai, M. F., Schipper, C. A., Lo, C. F., Kou, G. H., and Vlak, J. M. (2000). Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions. *Journal of General Virology* **81**, 307-316.

Introduction

White spot syndrome is a worldwide disease of shrimp. The disease has also been observed in other invertebrate aquatic organisms, such as crab and crayfish, and has a major impact on the economy of shrimp farming industry in Southeast Asia (Flegel, 1997). Infected shrimps show strong signs of lethargy and a reddish coloration of the hepatopancreas (Chou *et al.*, 1995). Moribund penaeid shrimps exhibit patchy discolorations or 'white spots' in the exo-mesoderm under the carapace, hence the name 'white spot syndrome'.

The disease agent is a bacilliform, enveloped virus with a rod-shaped nucleocapsid containing double stranded DNA. The virus particles have a tail- or flagellar-like extension at one end. Its morphology, nuclear localization and morphogenesis are reminiscent of baculoviruses in insects (Durand *et al.*, 1997). The virus is known as Hypodermic and Hematopoietic Necrosis BaculoVirus (HHNBV), Rod-shaped nuclear Virus of *Penaeus japonicus* (RV-PJ), Systemic Ectodermal Mesodermal BaculoVirus (SEMBV), White Spot BaculoVirus (WSBV) and White Spot Syndrome Virus (WSSV) (Lightner, 1996). A species name, whispovirus, has been proposed (Vlak *et al.*, 2000). Although WSSV has baculovirus characteristics and was an unassigned member of the Baculoviridae in the past (Francki *et al.*, 1991), at present it is no longer accepted into this family (Murphy *et al.*, 1995) due to the lack of molecular information.

The WSSV genome consists of a double stranded DNA estimated to be well over 200 kb in size (Yang *et al.*, 1997). The structure of the DNA (linear, circular) is not yet known. We have begun to analyze the WSSV genome in order to investigate its taxonomic status and to understand the genetic basis of its pathology in such a wide range of invertebrate aquatic organisms including penaeid shrimp. In this paper, we describe the identification and analysis of a 12.3 kb segment of the WSSV genome. This segment contained two types of tandem repeat sequence regions and encodes several putative open reading frames (ORFs), including two for the large (RR1) and small (RR2) subunit of ribonucleotide reductase. This enzyme is often found encoded in large DNA viruses, including African swine fever virus (Boursnell *et al.*, 1991), herpesviruses (Willoughby *et al.*, 1997), poxviruses (Schmitt & Stunnenberg, 1988), iridoviruses (Tidona & Darai, 1997), phycodnaviruses (Li *et al.*, 1997) and baculoviruses (van Strien *et al.*, 1997). The genes for RR1 and RR2 are the first ORFs in WSSV DNA for which a putative function has been assigned. Both genes were used to investigate the ancestral relationship with other organisms and viruses including baculoviruses.

Results

Analysis of an 11 kb *Bam*HI fragment of WSSV DNA

WSSV DNA was isolated from purified virions and digested with *Bam*HI (Fig. 2.1). As determined from agarose gels the sizes of the fragments ranged from about 22 kb to 3 kb. The size and number of the larger fragments could not be accurately determined due to their poor separation in agarose gels and the possible presence of genetic variation in the WSSV isolate.

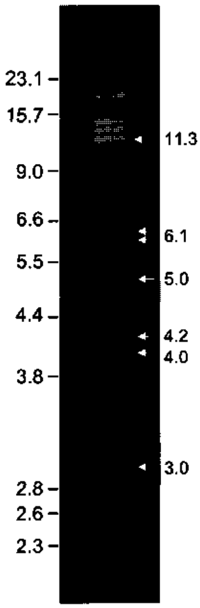


Figure 2.1 Electrophoresis of WSSV DNA isolated from purified virus particles and digested with *Bam*HI in a 0.8% agarose gel. The smaller *Bam*HI fragments are indicated with white arrows. Lambda DNA digested with *Bam*HI, *Hin*dIII, and *Eco*RI and with *Hin*dIII served as size standards. The sizes are shown in kb.

From the gel the total size of the WSSV genome was estimated to be over 200 kb, in agreement with what has been published previously (Yang *et al.*, 1997). *Bam*HI fragments were inserted into bacterial plasmids by shotgun cloning. *In situ* hybridizations were performed on sections of infected and uninfected shrimp tissue (data not shown) using the plasmids with *Bam*HI fragments as probes to confirm the viral origin of the inserts. The enlarged nuclei of cells of infected tissue were heavily stained, whereas sections of uninfected shrimp tissue served as a negative control. A commercial probe and specimen were used as a positive control for infection (DiagXotics, Inc.).

Plasmid inserts of nine hybridization-positive *Bam*HI clones of 18 kb, 14 kb, 11.3 kb, 6.3 kb, 6.0 kb, 5.0 kb, 4.2 kb, 4.0 kb and 3.0 kb (Fig. 2.1) were fine mapped and their termini sequenced. These sequences were subjected to BLAST analysis to identify homologous sequences in databases (GenBank). Most of the terminal sequences tested did not reveal any homology to known sequences. An 11.3 kb *Bam*HI fragment, however, contained a short stretch of terminal sequences homologous to those coding for the small subunit of ribonucleotide reductase (RR2). This enzyme is responsible for the reduction of ribonucleotides into precursors for DNA replication (deoxyribonucleotides) and consists of two subunits, RR1 and RR2 (Elledge *et al.*, 1992). This 11.3 *Bam*HI fragment was sequenced and a detailed physical map was derived (Fig. 2.2).

The WSSV *Bam*HI fragment was 11,319 nucleotides long with a GC content of 42.2%. Seven complete and one partial ORF and two regions with repeated sequences were identified (Fig. 2.2). One ORF showed a high degree of homology to *rr1* genes of prokaryotic

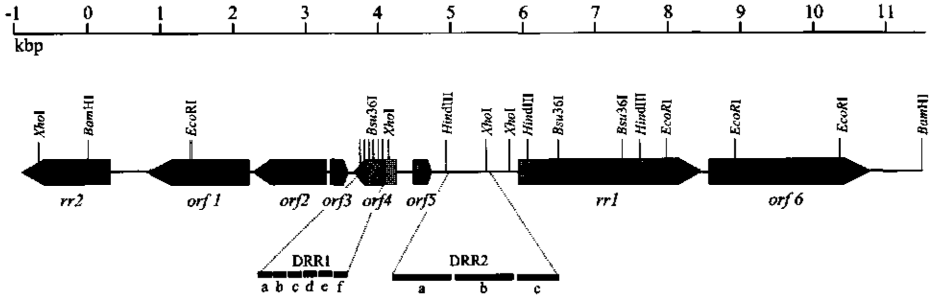


Figure 2.2 Physical map of a 12.3 kb segment of WSSV DNA. The top bar shows the location in kb relative to the *Bam*HI site in the RR2 ORF. The location and direction of transcription of the ORFs are shown below, with arrows. The black boxes represent the repeat units in DRR1 and DRR2.

Name	Location	Amino acids	pI	Mr (kDa)	TATA box	PolyA signal	Kozak context
RR2	- 943 ← 299	413	4.8	47.6	-		F
ORF1	934 ← 2100	388	8.5	45.2	+	+	F
ORF2	2102 ← 3220	372	6.9	43.2	-	-	F
ORF3	3447 → 3671	74	8.5	8.7	-	-	U
ORF4	3707 ← 4291	194	4.5	22.3	+	+	F
DRR1	3710 - 4033						
ORF5	4556 → 4876	106	8.4	12.4	-	-	U
DRR2	4963 - 5649						
RR1	5913 → 8459	848	7.8	95.6	+	+	F
ORF6	8540 → 10768	742	5.1	84.4	+	-	F

Table 2.1 Major predicted ORFs and repeat regions on a 12.3 kb WSSV genome fragment. The presence or absence of a TATA box and a polyadenylation signal are shown by + or -, respectively, and the Kozak context in which the initiation codon is located is shown ad F (favourable) or U (unfavourable).

and eukaryotic organisms, as well as large DNA viruses, and hence was designated WSSV *rr1*. This ORF of 2547 nucleotides, from nucleotide position 5913 (5'end) to 8459 (3' end), potentially encodes a protein of 848 amino acids with a theoretical molecular mass of 96 kDa. The initiation codon of the putative *rr1* ORF is located in a favourable context for translation (Kozak, 1989). A TATAbox is located 113 bp upstream of the putative RR1 translational start site. A polyadenylation signal is present 3 nucleotides downstream of the 3' end of the ORF. At the left-hand end of the *Bam*HI fragment (Fig. 2.2) is the 5' end of an ORF (*rr*2) with high homology to the small subunit of RR.

Six more non-overlapping ORFs ranging in size from 742 amino acids (ORF6) to 74 amino acids (ORF3) were identified (Fig. 2.2). Fifty amino acids were taken as the minimum size for a putative ORF. An overview of the location and direction of transcription of these ORFs and the size and isoelectric point of their proteins is given in Table 2.1. The putative initiation codons of ORF1, ORF2, ORF4 and ORF6 are in a favourable context for translation (Kozak, 1989). A TATA-box upstream of the ATG is present for ORF1, ORF4 and ORF6, whereas, a polyadenylation signal is found only for ORF1 and ORF4 (Fig. 2.2). ORF4 overlaps with direct repeat region one (Fig. 2.2). Its putative product is highly acidic (pI = 4.5) and contains repeated domains in the C-terminal half of the protein. ORF5 is small and has no proximal TATA box or polyadenylation signal; the ATG is in an unfavourable context for translation. ORF6 was identified downstream from the putative WSSV *rr1* gene and has two putative ATG start sites in a favourable context for translation.

Most of the putative proteins encoded by these ORFs have no hydrophobic sequences that might constitute signal sequences or transmembrane domains. The ORF4 protein contains a putative signal sequence, as it has at its N-terminus a positively-charged domain with a central hydrophobic region. A consensus motif for cleavage between amino acid residues 19 and 20 is present. Only proteins from ORFs 1 and 6 contain stretches of amino acids that might form transmembrane domains. ORF 1 has a putative transmembrane region formed by amino acids 4 to 24. ORF6 contains two putative transmembrane regions at positions 17 to 37 and 188 to 204. None of the ORFs contain a canonical nuclear localization signal. Searches against the GenBank, EMBL, PIR and SWISSPROT databases with the amino acid sequences of ORF1-ORF6, using programs BLAST (Altschul *et al.*, 1997) and FASTA (Pearson & Lipman, 1988), did not reveal significant similarities to any known sequence.

Completion of the putative RR2 ORF

Since only the 5' end of *rr2* was located on the 11.3 kb *Bam*HI fragment (Fig. 2.2), a terminal 2.7 kb *Eco*RI-*Xho*I fragment was used to screen a *Xho*I digest of WSSV DNA by Southern hybridization. A 4.8 kb *Xho*I fragment hybridized with this *Eco*RI-*Xho*I fragment and was subsequently cloned and sequenced. This fragment contained a considerable part of the *rr2* gene. The 5' end of the *rr2* sequence in the *Xho*I fragment was identical to the sequence in the 11.3 kb *Bam*HI fragment, confirming the overlap between these two fragments. To identify the 3'end of the *rr2* ORF, a 3' RACE analysis was performed on RNA isolated from infected shrimp tissue (data not shown). The ORF is 1242 nucleotides long and potentially encodes a protein of 413 amino acids with a theoretical molecular mass of 47.6 kDa. Database analysis further confirmed the identity of this ORF as an RR2 homologue. The initiation codon of the putative RR2 protein is located in a favourable context for translation (Kozak, 1989).

Repeat regions

Two direct repeat regions were detected, direct repeat region one (DRR1) and direct repeat region two (DRR2), separated from each other by 929 bp (Fig. 2.2). Both repeat regions consist of direct tandemly repeated DNA stretches with no inverted repeats or palindromic sequences. DRR1 consists of six direct repeats of 54 bp with 98% nucleotide sequence conservation (Fig. 2.3). DRR1 is an intragenic part of ORF4 and results in repeated amino acid stretches in the putative protein.

DRR2 is located 263 bp upstream of the *rr1* gene (Fig. 2.2) and consists of one partial and two complete direct repeat units of 253 bp (Fig. 2.3). The conservation of these repeat units is 87%. Comparison of DRR1 and DRR2 with approximately 147 kb of WSSV sequences (M-F. Tsai & C-F. Lo, unpublished results) showed that DRR2 is present at multiple, non-overlapping fragments of the WSSV genome, but no sequences homologous to DRR1 were found. In the sequences available, six repeat regions similar to DRR2 were found on non-overlapping fragments. The repeat regions consisted of between two and five complete or partial repeat units. The homology between these repeat units and repeat units DRR2a and DRR2b (Fig. 2.3) varies between 70 and 90 %. When all repeat units were aligned one conserved AT-rich (93% A+T) region and two conserved GC-rich (both 75% G+C) regions were found (Fig. 2.3). Comparison of DRR1 and DRR2 with sequences in GenBank did not show any homology to other organisms or viruses.

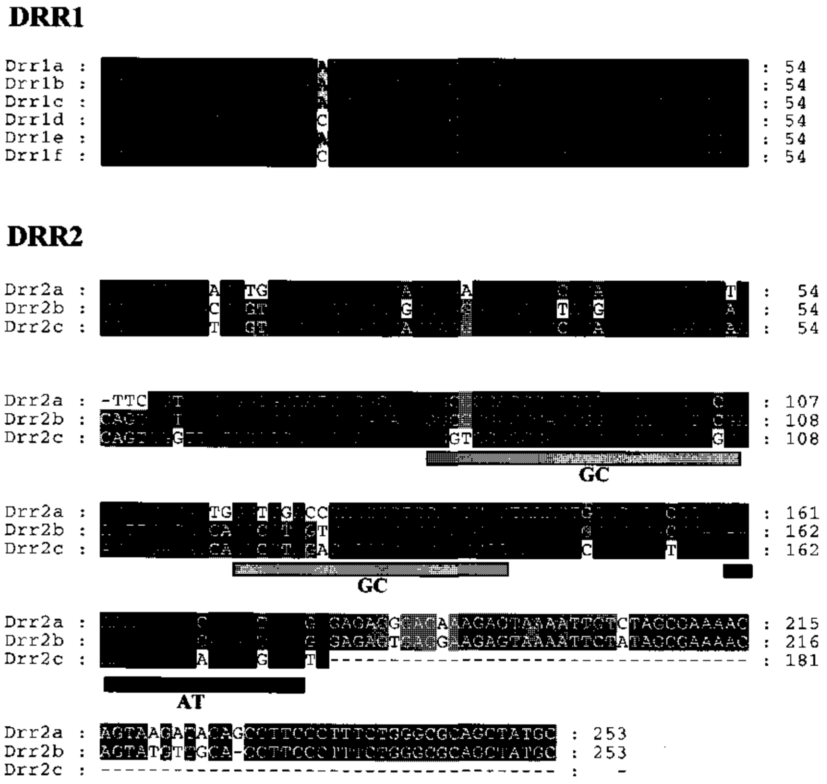


Figure 2.3 Nucleotide sequence alignment of DRR1 and DRR2. Shading is used to indicate the occurrence (black 100%, grey at least 60 %) of identical nucleotides. GC- and AT-rich areas are indicated.

Alignment of WSSV RR proteins

Ribonucleotide reductase is usually composed of two dissimilar subunits, the large subunit (RR1) containing the allosteric regulatory sites and the small subunit (RR2) containing the binuclear iron centre and a tyrosyl-free radical (Elledge *et al.*, 1992). This enzyme is important for DNA replication as a provider of dNTP precursors and hence conserved during evolution. The amino acid sequences of WSSV RR1 and RR2 were compared to those from other organisms and viruses. An alignment of the complete amino acid sequences was made in CLUSTALW and then proofread for inconsistencies (available on request). Most conserved amino acid residues for enzyme activities in the large subunit were also found in WSSV RR1. These include residues involved in the substrate reduction, and residues located in the cavity of the active site. Residues involved in the radical transfer reaction, as well as two cysteines in the C-terminal end with either a Cys-X-X-Cys or Cys-X-X-X-X-Cys motif (Uhlin & Eklund, 1994; Stubbe, 1990) were also conserved. The small subunit (RR2) supplies the reducing capacity of RR and contains an active iron centre and provides electrons via a tyrosyl radical (Schmidt *et al.*, 1998; Liu *et al.*, 1998; Dormeyer *et*

	RR1			RR2		
Source	**WSSV**	**SeMNPV**	**OpMNPV**	**WSSV**	**SeMNPV**	**OpMNPV**
HS	47 / 69	48 / 67	21 / 37	45 / 60	43 / 57	19 / 36
VAC	45 / 64	49 / 67	20 / 38	44 / 57	52 / 69	18 / 33
SlMNPV	45 / 65	49 / 67	21 / 35	-	-	-
SeMNPV	42 / 59	-	21 / 38	34 / 50	-	19 / 36
PBCL	40 / 59	42 / 61	21 / 38	39 / 54	48 / 68	16 / 36
ASFV	34 / 51	35 / 55	22 / 37	31 / 45	34 / 55	37 / 13
EC	25 / 42	24 / 43	19 / 34	18 / 35	21 / 40	17 / 36
HVS	24 / 46	25 / 45	22 / 36	19 / 37	26 / 50	16 / 36
LCDV	22 / 36	23 / 38	19 / 38	19 / 36	25 / 43	14 / 35
CpGV	19 / 35	23 / 37	47 / 64	-	-	-
OpMNPV	19 / 34	21 / 38	-	17 / 34	19 / 36	-
LdMNPV(a)	18 / 34	22 / 36	83 / 91	17 / 35	16 / 33	79 / 84
LdMNPV(b)	-	-	-	43 / 58	48 / 64	18 / 36

Table 2.2 Pairwise amino acid identity and similarity (BLOSUM 35) of RR1 and RR2 proteins. WSSV, SeMNPV and OpMNPV RR1 and RR2 sequences were compared with each other and with other RR genes. Percentage identity / similarity is shown. The RR2 sequences of CpGV and SlMNPV were not available. See Methods for abbreviations. -, not done.

al., 1997). In WSSV RR2 all residues essential in this pathway are conserved.

The alignments showed that the WSSV RRs have a high degree of homology to eukaryotic RRs. The pairwise identity and similarity (BLOSUM 35) of the RR1 and RR2 proteins of WSSV, SeMNPV and OpMNPV to each other, and to other viral RRs, and human and *E. coli* RRs, are given in Table 2.2. The highest homology of WSSV RR1 and RR2 was to human RR1 and RR2, whereas the lowest homology was observed with OpMNPV and LdMNPV RR1 and RR2. Significant homology was found with other baculovirus RRs, such as those from SeMNPV and SlMNPV and also to Chlorella virus. Furthermore, the WSSV RRs showed a relatively high homology to RRs from ASFV. As compared to baculoviruses, WSSV *rr2* has a surprisingly high degree of homology to a second *rr2* gene in LdMNPV. This gene, LdMNPV-*rr2b*, is located distally from LdMNPV *rr1* and *rr2a*, which are juxtaposed (Kuzio *et al.*, 1999). WSSV *rr2* has a relatively high degree of homology with SeMNPV *rr2*, which also does not colocalize with *rr1* (van Strien *et al.*, 1997; IJkel *et al.*, 1999).

Phylogenetic analysis of WSSV RRs

In order to study the relatedness between RRs of WSSV and eukaryotes, prokaryotes, and viruses phylogenetic trees were constructed using the amino acid sequences of WSSV RR1 and RR2 and these of twenty-six other RR1 and twenty-four other RR2 proteins (Fig. 2.4). Alignments without the variable N- and C-terminal sequences were made using CLUSTALW, and used in the phylogeny study. Maximum parsimony phylogenetic trees were obtained using PAUP, followed by bootstrap analysis (100 replicates) to determine the 50% majority-rule consensus tree. A heuristic search was performed, where starting trees were obtained by stepwise addition (starting seed 1), and tree-bisection-reconnection branch-swapping was performed with the MULPARS function.

The phylogenetic trees for RR1 (Fig. 2.4a) and RR2 (Fig. 2.4b) showed considerable resemblance. The bootstrap values of the viruses in the eukaryotic part of the RR2 tree are

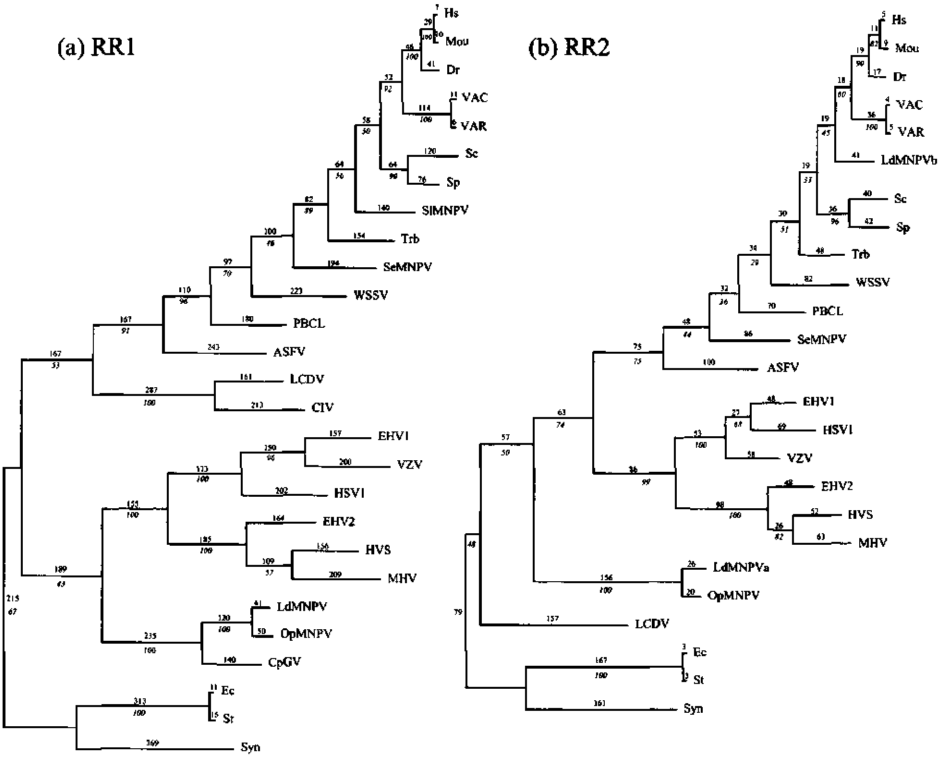


Figure 2.4 Bootstrap analysis (100 replicates) of unrooted phylogenetic trees of RR1 (a) and RR2 (b) proteins constructed with the PAUP heuristic search algorithm. Numbers at the branch indicate branch length (normal type) and frequency of clusters (italics). Outgroups are shown at the bottom. The sequences used were either published or present in GenBank (see Materials and Methods for listing).

rather low, making this part of the tree less informative to show relatedness. Typically for maximum parsimony, bootstrap values of ≥ 70% correspond to a probability of ≥ 95% that the respective clade is a historical lineage. The bootstrap value for vaccinia and variola virus to form a separate group in the eukaryotic part of the tree is 100% in both trees. The position of the herpesviruses in a separate group is also bootstrap supported to almost 100% in both trees, which is also the case for the baculoviruses CpGV, LdMNPV and OpMNPV in the RR1 tree, separating these viruses from SeMNPV and SlMNPV. In the RR2 tree (Fig. 4b) the same separation is found, except that the extra *rr2b* gene in LdMNPV is found in the eukaryotic branch of the tree. In both the RR1 and the RR2 tree the phylogenetic status of WSSV is not resolved. In the RR1 tree, WSSV and SeMNPV branch after ASFV and PBCL and before SlMNPV. However, this branching is not clear in the RR2 tree, because of the very low bootstrap values observed for these viruses in the eukaryotic branch of the tree.

Discussion

White spot syndrome is a devastating viral disease in aquatic organisms, in particular in penaeid shrimp. The virus (WSSV) has not been classified by the ICTV (Murphy *et al.*, 1995), but it resembles baculoviruses on the basis of its morphology, morphogenesis and nuclear localization and replication (Durand *et al.*, 1997). Analysis of the viral DNA with restriction enzymes (Yang *et al.*, 1997) gave no further clue as to its taxonomic status. The virus is obtained from many crustacean species, but limited sequence analysis using a two-step PCR procedure seems to suggest that the isolates are very similar if not identical viruses (Lo *et al.*, 1999). The present report is the first providing functional information on WSSV by analyzing a genomic segment of about 12.3 kb. This segment appeared to contain eight ORFs including two for the large and small subunit of ribonucleotide reductase and two direct repeat regions (Fig. 2.2).

Six ORFs, provisionally named ORF1-6 (Fig. 2.2), had no homologues in accessible databases (GenBank) and thus appeared unique to WSSV. Analysis of further *Bam*HI fragments, as well as fragments from *Eco*RI, *Hin*dIII and *Sal*I libraries, by terminal sequencing showed no significant homology with sequences in the databases. This information included the entire sequences of five baculoviruses, AcMNPV (Ayres *et al.*, 1994), *Bombyx mori* (Bm) NPV (Gomi *et al.*, 1999), OpMNPV (Ahrens *et al.*, 1997), LdMNPV (Kuzio *et al.*, 1999) and SeMNPV (IJkel *et al.*, 1999), as well as partial sequence information of other baculoviruses. This is in contrast to baculoviruses, where a large proportion of genes have homologues (Hu *et al.*, 1998). These homologues usually encode proteins that determine the baculovirus character, such as genes involved in DNA replication and transcription (late expression factor genes) or encoding virion structural proteins. On the basis of the limited sequence information obtained, WSSV is distinct from baculoviruses accommodated in the genera Nucleopolyhedrovirus and Granulovirus (Murphy *et al.*, 1995). The function of the six unique WSSV ORFs with no homology in databases (GenBank) will be further investigated by molecular and immunological methods.

Comparison of the promoter regions of the eight ORFs on a WSSV segment including the *Bam*HI fragment indicated the presence of common motifs, such as TATA boxes and transcription termination signals. For some ORFs (ORF3 and ORF5) these signals were not present, and a favourable Kozak consensus sequence was also absent. Possibly, these ORFs are not expressed, and transcriptional analysis could indicate which of them are functional. Promoter motifs typical of baculovirus transcription, such as the 'early' CAGT and 'late' TAAG (Friesen, 1997; Lu & Miller, 1997), are found scattered in the WSSV sequence but not at appropriate promoter locations of the eight ORFs. This suggests that their transcription strategy is different from that of baculoviruses.

Two ORFs had a high degree of homology with the large and small subunit of RR and were amenable to further investigation of the taxonomic status of WSSV by phylogenetic analysis. Ribonucleotide reductase is a key enzyme in the DNA replication process as it catalyses the reduction of ribonucleotides into deoxyribonucleotides (Elledge *et al.*, 1992). Genes involved in nucleotide metabolism are found in pro- and eukaryotic organisms and also in many large DNA viruses. Therefore these enzymes, including RR, are excellent tools to study relatedness and phylogeny. Bootstrap analysis of parsimonious phylogenetic trees of

RR1 and RR2 genes (Fig. 2.4) confirmed the location of WSSV in the eukaryotic branch of the RR trees. This positioning of WSSV is supported by the pairwise alignment of the RR proteins, where the highest homology of the WSSV proteins is found with human and vaccinia virus RR (Table 2.2). Several other viruses are found in the eukaryotic branch of the trees, including ASFV, PBCL, SeMNPV and SlMNPV (Fig. 2.4a). In the RR1 tree ASFV, PBCL and WSSV form separate clades supported by high bootstrap values. The low bootstrap scores of the SeMNPV and the SlMNPV branch indicate that their position of branching remains ambiguous. None of these viruses are located in a separate cluster and hence a common recent ancestor is not likely. The RR2 tree (Fig. 2.4b) is somewhat less informative due to the lower bootstrap values. It is interesting to note that WSSV RR2 is more related to the RR2 of SeMNPV and the RR2b of LdMNPV (Kuzio *et al.*, 1999). The latter two *rr2s* are located distant from *rr1* on the viral genome, as is the case for WSSV *rr2*, whereas WSSV has a low homology with *rr2s*, such as OpMNPV *rr2*, which are juxtaposed to *rr1*. These phylogenetic data indicate that WSSV and those baculoviruses carrying *rr* genes do not share an immediate common ancestor.

The region separating the *rr1* and *rr2* gene also contains two domains with multiple repeats. The first repeat domain (DRR1) has six direct repeats of 54 bp, which provides the putative ORF4 product with six repeated domains of 18 amino acids. This protein is highly acidic, and its biological function and the significance of these domains remain to be determined. A second domain (DRR2) contains one partial and two complete direct repeats of 253 bp and is located in an intergenic region of the WSSV *Bam*HI fragment. Regions with similar repeats have been identified on at least six non-overlapping WSSV DNA fragments (M-F. Tsai & C-F. Lo, unpublished results). These repeats may have a similar function as the homologous regions (*hrs*) in baculoviruses. *Hr*s also occur at multiple locations in the baculovirus genome and function as enhancers of transcription and origins of DNA replication (Cochran & Faulkner 1983; Guarino & Summers, 1986; Kool *et al.*, 1995). However, the repeats found in WSSV do not have inverted or palindromic repeats, which appear to be important for their function in baculovirus DNA replication. It remains to be determined whether DRR2-like repeats have a similar function in WSSV replication and/or transcription.

WSSV represents a virus species for which the taxonomic status remains to be determined. Originally, this virus was an unassigned member of the Baculoviridae (Francki *et al.*, 1991). At present, WSSV is no longer accepted in the baculovirus family (Murphy *et al.*, 1995). Its large DNA size (>200 kb) and the lack of significant gene homology with baculoviruses seems to justify this classification. This taxonomic status is further supported by the phylogenetic information using RR as a tool, where a close ancestral relatedness to baculoviruses could not be demonstrated. The collective information obtained so far including the sequence information presented in this paper supports the view that WSSV could be either a representative of a new genus (Whispovirus) within the Baculoviridae or a representative of an entirely new virus family.

Materials and methods

Virus production and purification

The virus used in this study was isolated from infected *Penaeus monodon* shrimp from Thailand. Infected tissue was homogenized in TN buffer (20 mM Tris-HCl, 400 mM NaCl, pH 7.4). After centrifugation at 1,700 x g for 10 min the supernatant was filtered (0.45 μm filter). The filtrate was injected into healthy *P. monodon* in the lateral area of the fourth abdominal segment to initiate an infection. After 3 or 4 days haemolymph was withdrawn from moribund shrimp and mixed with modified Alsever solution (Rodriguez *et al.*, 1995) as an anticoagulant. After dilution in TNE (20 mM Tris-HCl, 400 mM NaCl, 5 mM EDTA, pH 7.4), the haemolymph was clarified at 1,700 x g for 10 min at 4°C. The virus particles were removed from the supernatant by centrifugation at 45,000 x g at 4°C for 1 h and suspended in TN buffer.

Alternatively, virus was isolated from infected shrimp tissue. The tissue was homogenized in TNE using a mortar, and after clarification at 1,700 x g for 25 min at 4°C the supernatant was subjected to discontinuous sucrose density gradient centrifugation. The virus was obtained from the 20%/45% sucrose interface and after dilution in 3 volumes of TNE sedimented at 80,000 x g for 1 hour and suspended in TN. Virus preparations were stored at -20 °C .

Nucleic acid purification

Viral DNA was isolated from purified virions by treatment with proteinase K (0.2 mg/ml) and sarcosyl (1%) at 45°C for 2 – 4 h, followed by phenol/chloroform extraction and dialysis against TE (Sambrook *et al.*, 1989). The purity and concentration of the DNA was determined by agarose gel electrophoresis.

Cloning and mapping of WSSV DNA fragments

WSSV DNA was digested with restriction enzymes (GIBCO BRL) and the fragments were separated by electrophoresis in 0.7% agarose gels at 45V (1.5V/cm) for 14-20 h. Lambda DNA digested with *Bam*HI-*Eco*RI-*Hin*dIII or with *Hin*dIII was used as a molecular size standard. WSSV DNA fragments were inserted into the plasmid vector, pBluescriptSK+, by shotgun cloning of digested WSSV DNA using standard techniques (Sambrook *et al.*, 1989). Fragments were purified from agarose gels using Glassmax (GIBCO BRL) prior to cloning. Southern hybridization of cloned fragments with digested WSSV DNA was performed using standard techniques (Sambrook *et al.*, 1989).

Restriction enzyme analysis of the cloned fragments with various enzymes, Southern blot hybridization and subcloning were employed to construct a physical map of the 11.3 kb *Bam*HI and the 4.8 kb *Xho*I WSSV DNA fragments. Sequence data of the subcloned fragments were also used to establish and confirm the detailed map.

In situ hybridization

In situ hybridization with cloned fragments was carried out on tissue sections of WSSV-infected shrimp (a gift from Mr. P. Voorthuis, Laboratory of Fisheries, Wageningen Agricultural University) using the ShrimProbe® protocol (DiagXotics, Inc.). Preparations of uninfected shrimp were used as negative control. Cloned fragments were labeled with digoxygenin according to the DIG System User's Guide for Filter Hybridization (Boehringer Mannheim). The hybridization was visualized using alkaline phosphatase-conjugated anti-digoxigenin antibody. Bismarck Brown was used to counterstain the preparations.

DNA sequencing and computer analysis

Plasmid DNA for sequencing was purified via the QIAprep Miniprep System or JETstar Plasmid Purification System (Qiagen, Inc.). Sequencing was done on both ends of the (sub)cloned fragments using the universal pBluescript forward and reverse nucleotide primers. The complete *Bam*HI fragment was sequenced from both strands by the 'sequence walking' method with custom synthesized primers. Automatic sequencing was carried out at the Sequencing Facility in the Department of Molecular Biology of the Wageningen Agricultural University.

The sequences generated were analyzed with UWGCG computer programs (release 9.0). The DNA and the deduced amino acid sequences were compared with the updated GenBank/EMBL, SWISSPROT and PIR databases using FASTA and BLAST. For studying the *rr1* and *rr2* gene homology, the amino acid sequences of the selected homologues were aligned by the multiple sequence alignment program CLUSTALW (Thompson *et al.*, 1994). Phylogenetic analysis was performed with PAUP3.1 program (Swofford, 1993) using CLUSTALW to produce input files of aligned protein sequences. Bootstrap analysis according to Felsenstein (1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny. The nucleotide sequence of the WSSV *Bam*HI fragment and the *rr*2 gene have been deposited in GenBank and were assigned accession numbers: AF099142 and AF144620, respectively.

The following RR1/RR2 sequences available in GenBank were used in the alignment and phylogenetic analysis: *Homo sapiens* (Hs), X59617/X59618; *Mus musculus* (Mou), K02927/M14223; *Danio rerio* (Dr), U57964/U57965; Vaccinia virus (VAC), A28611/M57977; Variola virus (VAR), P32984/X69198; *Saccharomyces cerevisiae* (Sc), P21524/M17789; *Schizosaccharomyces pombe* (Sp), P36602/X65115; *Spodoptera littoralis* nucleopolyhedrovirus (SlMNPV), X98924; *Trypanosoma brucei* (Trb), U80910/Y10768; *Spodoptera exigua* nucleopolyhedrovirus (SeMNPV), X97578; *Paramecium bursaria* Chlorella virus 1 (PBCL), U42580/P49730; African swine fever virus (isolate malawi lil 20/1) (ASFV), P26685/P26713; Lymphocystis disease virus 1 (LCDV), L63545; *Chilo iridescent* virus (CIV), AF003534; Equine herpesvirus type 1 (strain AB4P) (EHV1), P28846/M86664; Varicella-zoster virus (strain dumas) (VZV), P09248/P09247; Human herpesvirus 1 (HSV1), 66401/X14112; Equine herpesvirus 2 (EHV2), U20824; Herpesvirus saimiri (strain 11) (HVS), Q01037; Murine herpesvirus 68 (MHV), U97553; *Lymantria dispar* nucleopolyhedrovirus (isolate:CI5-6) (LdMNPV), AF081810; *Orgyia pseudotsugata* nucleopolyhedrovirus (OpMNPV), U75930; *Escherichia coli* (Ec), 66408/P00453; *Salmonella typhimurium* (St), P37426/P37427; *Synechocystis* sp. (Syn), D90913/D90917. The *Cydia pomonella* granulovirus (CpGV) RR1 sequence was provided by Dr D. Winstanley (Horticulture Research International, Wellesbourne, UK).

Acknowledgements

We are grateful to Dr E. A. van Strien for assistance in the phylogenetic tree analysis. We thank Dr D. Winstanley (Horticulture Research International, Wellesbourne, UK) for providing us with the CpGV RR1 sequence. This research was supported in part by funds from Intervet International bv, Boxmeer, the Netherlands, and by the National Science Council in Taiwan.

Chapter 3

Identification and phylogeny of two protein kinase genes of White Spot Syndrome Virus

White Spot Syndrome Virus (WSSV) is a virus infecting shrimp and other crustaceans, which is unclassified taxonomically. Two open reading frame, encoding putative protein kinases (PK), were found on the WSSV genome. The identified PKs show a high degree of homology to other viral and eukaryotic PKs. Homology in the catalytic domains suggests that the PKs are serine/threonine protein kinases. All of the conserved PK domains are present in the WSSV PK gene products and this allowed the alignment with PK proteins from other large DNA viruses, which encode one or more PK proteins. An unrooted parsonimous phylogenetic tree was constructed and indicated that the PK genes are well conserved in all DNA virus families and hence can be used as phylogenetic markers. The two WSSV PK genes are present in a separate branch of the tree that is well bootstrap supported. As these PK genes shared a common ancestor, they may have evolved by gene duplication. Baculoviruses to date contain only a single PK gene, which is present in a separate well bootstrap-supported branch in the tree. The WSSV PKs are not present in the baculovirus clade and therefore are clearly separated phylogenetically from the baculovirus PK genes. Furthermore, the WSSV PK genes do not share a most recent common ancestor with any known PK gene from other viruses. This provides further and independent evidence for the unique position of WSSV in a newly proposed genus named Whispovirus.

This chapter has been published in a slightly modified version as:
van Hulten, M. C. W. and Vlak, J. M. (2001). Identification and phylogeny of a protein kinase gene of white spot syndrome virus. *Virus Genes* **22,** 201-207.

Introduction

White spot syndrome virus (WSSV) causes a worldwide disease in shrimp. Its virion morphology, nuclear localization and morphogenesis are reminiscent of baculoviruses in insects (Durand *et al.*, 1997). Therefore WSSV was originally classified as an unassigned member of the Baculoviridae (Francki *et al.*, 1991). In 1995 the non-occluded baculoviruses were eliminated from the Baculoviridae family and orphaned (Murphy *et al.*, 1995). The orphan status of WSSV is justified, as many characteristics of this virus and its pathogenesis are distinct from that of baculoviruses. The lack of occlusion bodies in WSSV infections, the tail-like appendix at one end of the enveloped virions and the absence of characteristic peplomer structures on the apical surface of the virions separate WSSV from baculoviruses. Furthermore, WSSV nucleocapsids have a crosshatched appearance in contrast to the smooth surface of baculovirus nucleocapsids (Durand *et al.*, 1997; Funk *et al.*, 1997). The size of the genome of WSSV was estimated to be over 250 kb, which is well out of the range of a baculovirus genome (Yang *et al.*, 1997). In contrast to baculoviruses, which mainly infect insects, WSSV has an extremely wide host range among crustaceans including shrimp, crayfish and crab. Analysis of the genome organization and content of WSSV as well as gene phylogeny may provide information on the taxonomic position of WSSV relative to baculoviruses and other large DNA viruses. Here we describe the identification of two protein kinase (PK) genes in WSSV, which are used in a phylogenetic analysis.

Results and Discussion

Two open reading frames (ORFs) showing high homology to serine/threonine PK genes were found on the WSSV genome. One of the ORFs, PK-1 was located near the gene encoding the major envelope protein (VP28) of WSSV (van Hulten *et al.*, 2000a). VP28 and PK-1 are juxtaposed and are separated by 94 nucleotides. The PK-1 ORF is 2193 nucleotides and encodes a putative protein of 730 amino acids with a theoretical size of 82 kDa. The PK-2 ORF has a total size of 1743 nucleotides, encoding a 580 amino acids protein of 66 kDa. The putative translational start codons of both PK ORFs are in a favorable context for efficient eukaryotic translation initiation (Kozak, 1989), and both ORFs contain a TATA box and a polyA signal.

Serine/threonine- and tyrosine-protein kinases are essential proteins as they have a central role in signal transduction and cellular regulation in eukaryotes (Hanks *et al.*, 1988). The protein kinases are a large family of enzymes. They can roughly be subdivided into two main groups: the serine/threonine protein kinases, which are serine/threonine-specific and the tyrosine protein kinases, which are tyrosine-specific (Hanks *et al.*, 1988). The catalytic domain encompasses ≈ 250 – 300 amino acids, containing 12 subdomains, which fold into a common catalytic core structure (Hanks and Hunter, 1995). These twelve kinase domains are conserved throughout the PK superfamily, and therefore most probably play essential roles in enzyme function. Eukaryotes contain a large number of PK genes. The mammalian genome is estimated to contain about 1000 PK genes (Hanks and Hunter, 1995). Many large DNA viruses have proteins, that are phosphorylated (e.g. baculoviruses, (O'Reilly, 1997)), and therefore they often encode one or more copies of PK genes. In all baculoviruses that have been sequenced completely only one gene for serine/threonine PK has been identified (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Ijkel *et al.*, 1999). The same situation exists in African swine fever virus (Baylis *et al.*, 1993). Multiple copies of PK genes have been identified in poxviruses (Johnson *et al.*, 1993; Afonso *et al.*, 1999), in herpesviruses (McGeoch *et al.*, 1988; Leader and Purves, 1988), in iridoviruses (Tidona and Darai., 1997), and in phycodnaviruses (Lu *et al.*, 1996).

When the WSSV PKs were compared to other viral (AcMNPV (Ayres *et al.*, 1994), XcGV (Hayakawa *et al.*, 1999), Vac (Johnson *et al.*, 1993), EHV1 (Telford *et al.*, 1992), ASFV (Yanez *et al.*, 1995), LCDV (Tidona and Darai, 1997), and PbCl (Lu *et al.*, 1996)) and eukaryotic (Bov-1 (Wiemann *et al.*, 1992) and Hs-1 (Bacher *et al.*, 1991) PK genes, the pairwise amino acid homology ranged between 25 % and 32 % overall. None of the WSSV PKs shows a significantly higher homology with any of these PK genes, whereas they have a pairwise amino acid sequence similarity of 45%.

An alignment was made in ClustalW (Thompson *et al.*, 1994) using the WSSV PKs and several eukaryotic and viral serine/threonine PK proteins (Fig. 3.1). In this alignment all the conserved PK domains are shown, but the variable N-terminal part of some PKs is not included in the figure. The twelve conserved domains were all present in the WSSV PKs and the conserved amino acids are referred to that of the Bov-1 PK (Hanks *et al.*, 1988; Hanks and Hunter, 1995) numbering. The first PK domain, subdomain I, involved in binding of the ATP contains the consensus motif Gly-x-Gly-x-x-Gly-x-Val starting with Gly50 in the bovine PK (Hanks *et al.*, 1988; Hanks and Hunter, 1995). In WSSV PKs this motif is only partially present, as only two of the Gly residues are conserved. In most of the PK proteins shown in this alignment part of the consensus motif is present including the invariant Gly52, however the motif is not present in XcGV. Subdomain II contains an invariant Lys72, which is essential for maximal enzyme activity. This residue is present in all PK proteins in the alignment, which is also the case for the nearly invariant Glu residue (Glu91) in subdomain III that forms a salt bridge with Lys72 from subdomain II. Subdomain IV and V contain no invariant or nearly invariant residues. Subdomain VI contains the catalytic loop of the protein with consensus motif His-Arg-Asp-Leu-Lys-x-x-Asn of which the Asp166 and Asn171 residues are invariant. These residues are conserved in all PK proteins. Also Lys168, which is replaced by an Arg in conventional protein-tyrosine kinases, is conserved in all PK proteins. Subdomain VII contains the highly conserved Asp184-Phe185-Gly186 (DFG triplet) which is also conserved in WSSV and the other PK proteins in the alignment, except for LCDV, where the Asp has been replaced. However, the APE motif (Ala206-Pro207-Glu208) in subdomain VIII that is highly conserved is not present in all viral PKs. In WSSV, PbCl and EHV1 this APE motif is present, however in the other viral PKs only part of this motif is present. Subdomain IX contains the nearly invariant Asp220, which is also present in all PKs,

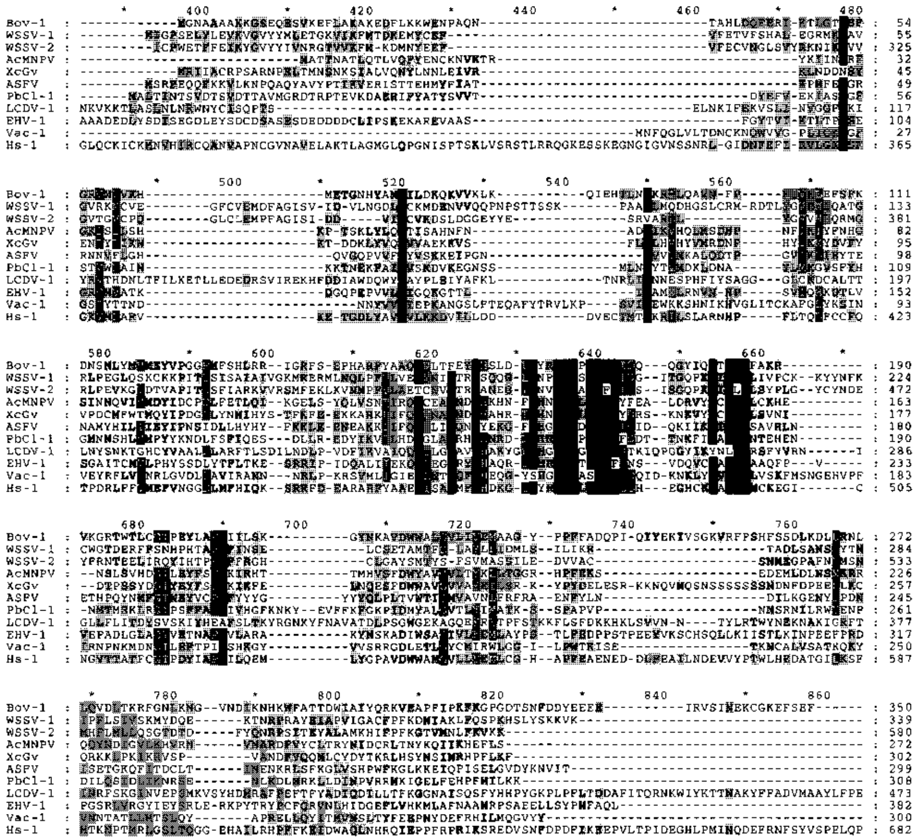


Figure 3.1 Amino acid sequence alignment of eukaryotic and viral PK genes. Shading is used to indicate the occurrence (black is over 90 %, dark gray over 70 % and light gray over 30 %).

except in WSSV, ASFV, and LCDV. Subdomain X is the most poorly conserved subdomain and its function is obscure. In the alignment this domain is not clearly present. In subdomain XI it is not clear if the nearly invariant Arginine (Arg280) is well conserved, as the overall homology in this domain is low. However for most PKs an Arg can be found in this domain at an approximate position.

Phylogenetic analysis was performed using the twelve subdomains of the alignment (Fig. 3.1: position 465 to 800) and 22 additional eukaryotic and viral PK proteins. Maximum parsimony phylogenetic trees were obtained using the PAUP3.1 program (Swofford, 1993), followed by bootstrap analysis (Felsenstein, 1993) (100 replicates) to determine the 50% majority-rule consensus tree. A heuristic search was performed, where starting trees were obtained by stepwise addition (starting seed 1), and tree-bisection-reconnection branch-swapping was performed with the MULPARS function. Genes encoding proteins homologous to eukaryotic serine/threonine protein kinases have been identified in several bacteria (Leonard *et al.*, 1998). An eukaryotic-type protein kinase of the ABC1 family

(Leonard *et al.*, 1998) of *Synechocystis* sp. (Syn) which is only distantly related to eukaryotic PK proteins was included in the alignment and could serve as an outgroup in the phylogenetic analysis (Fig. 3.2).

In this phylogenetic tree (Fig. 3.2) the eukaryotes and several virus group are well conserved. Typically for maximum parsimony, bootstrap values of ≥ 70% correspond to a probability of ≥ 95% that the respective clade is a historical lineage. High bootstrap values are obtained for the human and bovine PK genes. These PKs are all from the PK-A-C-G group and Hs-2 and Bov-1 are from the PK-A subfamily whereas Hs-1 and Bov-2 both are members of the PK-C subfamily (Hanks and Hunter, 1995). Therefore these PK genes share a most recent common ancestor and are divided in two branches of the PK-A and PK-C subfamily, respectively. None of the viral PKs are located in clusters of the eukaryotic PKs, even when other groups of eukaryotic PK genes are included (data not shown). Most probably the viruses have obtained their PK genes early in evolution and these have subsequently diverged. The PK genes of the baculoviruses, including several nucleopolyhedroviruses and one granulovirus, share a common ancestor, which is supported by a high bootstrap value of 96 %. XcGV is separated from the nucleopolyhedroviruses with a bootstrap value of 52%, in accordance with its classification in the granulovirus genus.

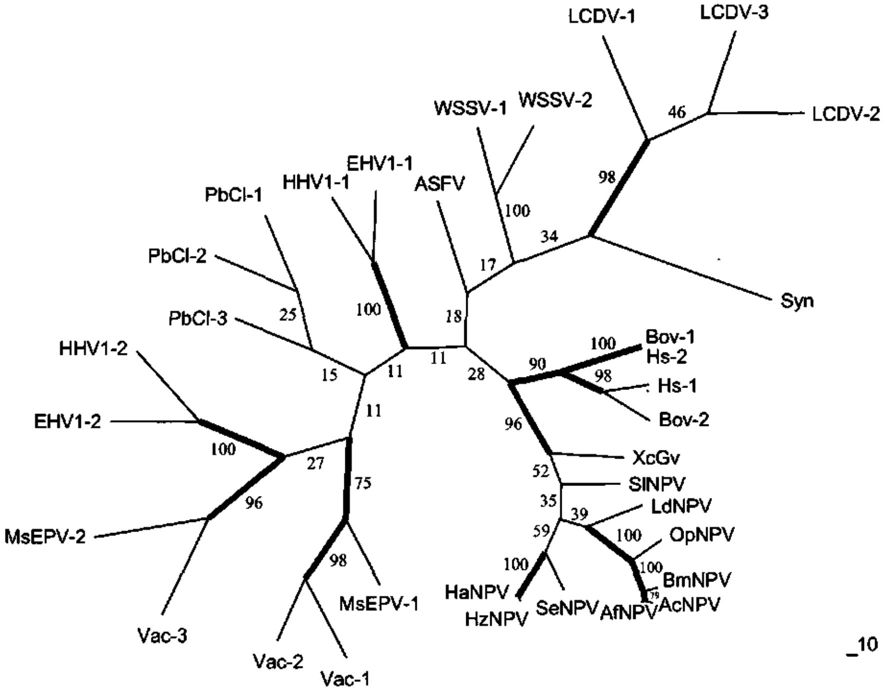


Figure 3.2 Bootstrap analysis (100 replicates) of unrooted phylogenetic trees of PK proteins constructed with the PAUP heuristic search algorithm. Numbers at the branches indicate frequency of clusters and frequencies over 70% are indicated by thick lines. The bar at the bottom equals a branch length of 10. Abbreviations and accession numbers can be found in the materials and methods section.

Based on the analysis of these PK proteins, WSSV does not belong to the baculovirus family. However, it remains possible that even more PK genes are present in WSSV. To date, the presence of more than one PK gene has not been observed in any of the completely sequenced baculovirus genomes yet (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Ijkel *et al.*, 1999), whereas this situation exists in many other large DNA viruses. Multiple copies of PK genes have been identified in herpesvirus, poxvirus, iridovirus, and phycodnavirus genomes. In the α-herpesviruses two PK genes (HHV1: HHV1-1 and HHV1-2 (McGeoch *et al.*, 1988; Leader and Purves, 1988)) are present of which one is virion-associated (HHV1-2). These two PK genes do not share a common ancestor, but form two clades with PK genes from other α-herpesviruses. The same situation exists in vaccinia, where three copies of a PK gene are present (24), which are separated in two clades in the tree. Two PK genes (Vac-1 and Vac-2) share a common recent ancestor and can be the result from a gene duplication event. In MsEPV (Afonso *et al.*, 1999) two PK genes have been identified which have a homologue in the vaccinia genome and therefore are in the same branch as their homologues. The situation in LCDV is different as in total six PK genes are present in the genome, which show significant homology to members of the serine/threonine and tyrosine PK gene families (Tidona and Darai, 1997). Three PK genes, which showed high homology with the serine/threonine PKs, were included in the tree. These are present in a separate branch of the tree that is well bootstrap supported (98%) (Fig. 3.2). As these PK genes share a common ancestor, they may have evolved by gene duplication. A similar situation exists for PbCl where three PK genes are present in the genome (Lu *et al.*, 1996). These are also in a separate branch in the tree, but the bootstrap value is very low. Here a situation may exist where one PK gene was obtained early in evolution, duplicated and diverged. The two WSSV PK genes share a common ancestor as they are presented in a 100% bootstrap supported branch of the tree and therefore may have evolved from gene duplication.

From the analysis of a single gene no definite conclusions on the taxonomic position of a virus can be made. Previous analysis of three major structural proteins of WSSV (VP28, VP26, VP24), which have been identified on the WSSV genome, showed no homology to known baculovirus proteins or to other proteins in databases (van Hulten *et al.*, 2000a; van Hulten *et al.*, 2000b), whereas genes coding for viral capsid or envelope proteins are highly conserved in related viruses (Murphy *et al.*, 1995). Phylogenetic analysis using the putative genes for the large and the small subunit of ribonucleotide reductase (RR1 and RR2) of WSSV (van Hulten *et al.*, 2000c) and RR genes of other large DNA viruses, eukaryotes, and prokaryotes suggested that the WSSV RRs belong to the eukaryotic branch of an unrooted parsimonious tree and further indicated that WSSV RRs do not share an immediate common ancestor with any of the baculovirus RRs available in the GenBank (van Hulten *et al.*, 2000c). The analysis of the PKs of WSSV provides further independent support for our proposition that WSSV belongs to a new virus genus with the proposed genus name Whispovirus (van Hulten *et al.*, 2000c). The phylogenetic analysis of the WSSV PK genes separates WSSV from baculoviruses as the latter share a most recent common ancestor and have obtained their PK gene before the separation into the genera granulovirus and nucleopolyhedrovirus. Although the number or suitable WSSV genes for phylogenetic analysis is limited (only an estimated 5% of the open reading frames identified so far have homologues in GenBank), analysis of additional genes is required to further substantiate the unique position of WSSV.

Materials and methods

Virus production and purification

A low concentration White Spot Syndrome Virus (WSSV) solution was injected intramuscular into healthy *Procambarus clarkii* in the lateral area of the fourth abdominal segment to initiate infection. After one week haemolymph was collected from the infected *P. clarkii*, and virus was purified over a continuous sucrose gradient (55–20% w/w sucrose in TN (20 mM Tris-HCl, 400 mM NaCl, pH 7.4)). Centrifugation was performed at 80000 x g for 1.5 hour at 4°C. The virus band was removed from the gradient, diluted in TN, pelleted by centrifugation at 30000 x g and dissolved in TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5).

Nucleic acid purification and cloning of WSSV DNA fragments

Viral DNA was isolated from purified virions as described in (van Hulten *et al.*, 2000c). Cloning of WSSV DNA fragments in pBluescript SK+ was performed as described in (van Hulten *et al.*, 2000c).

DNA sequencing and computer analysis

WSSV plasmid DNA was sequenced using the universal pBluescript forward and reverse nucleotide primers followed by the 'sequence walking' method with custom synthesized primers. Automatic sequencing was carried out at the Sequencing Facility in the Department of Molecular Biology of the Wageningen University.

The generated sequences were analyzed with UWGCG computer programs (release 10.0). The DNA and the deduced amino acids sequences were compared with GenBank/EMBL, SWISSPORT and PIR databases using the programs FASTA, TFASTA (Pearson and Lipman, 1988) and BLAST (Altschul *et al.*, 1997). For studying the PK gene homology, the amino acid sequences of the selected homologues were aligned by the multiple sequence alignment program CLUSTALW (Thompson *et al.*, 1994). Phylogenetic analysis was performed with PAUP3.1 program (Swofford, 1993) using CLUSTALW to produce input files of aligned protein sequences. Bootstrap analysis according to Felsenstein (1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny. The nucleotide sequence of the PK fragment has been deposited in GenBank and was assigned accession number: AY008843.

The following PK sequences available in GenBank were used in the alignment and phylogenetic analysis: *Synechocystis* sp. (Syn), BAA17617; bovine (Bov-1), P00517; human (Hs-2), NP_002721; human (Hs-1), NP_006246; bovine (Bov-2); P05128, *Xestia c-nigrum* granulovirus (XcGV), AAF05117; *Spodoptera littoralis* nucleopolyhedroviruses (SlNPV), CAA68048; *Lymantria dispar* NPV (LdNPV), P41720; *Orgyia pseudotsugata* NPV (OpNPV), AAC5900; *Bombyx mori* NPV (BmNPV), AAC63685; *Autographa californica* NPV (AcNPV), P41415; *Anagrapha falcifera* NPV (AfNPV), AAB53339; *Spodoptera exigua* NPV (SeNPV), AF169823; *Heliothis zea* NPV (HzNPV), P41719; *Helicoverpa armigera* NPV (HaNPV), AAB53738; *Melanoplus sanguinipes* entomopoxvirus (MsEPV), MsEPV-1, AAC97783, MsEPV-2, AAC97687; Vaccinia virus (VAC), Vac-1, P20505; Vac-2, P21098; Vac-3, P21095; Equine herpesvirus 1 (EHV1), EHV1-1 and EHV1-2, NC_001491; Herpes simplex virus type 1 (HHV1), HHV1-1, P04413; HHV1-2, WMBE71; *Paramecium bursaria* Chlorella virus 1 (PbCl), PbCl-1, AAC96616; PbCl-2, AAC96645; PbCl-3, AAC96657; African swine fever virus (isolate malawi lil 20/1) (ASFV), P34206; Lymphocystis disease virus (LCDV), LCDV-1, 2, and 3, L63545.

Acknowledgement

This research was supported by Intervet International BV, Boxmeer, The Netherlands.

Chapter 4

Identification of two major virion protein genes of white spot syndrome virus of shrimp

Abstract

White Spot Syndrome Virus (WSSV) is an invertebrate virus, causing considerable mortality in shrimp. Two structural proteins of WSSV were identified and expressed in the baculovirus expression system. WSSV virions are enveloped nucleocapsids with a bacilliform morphology, with an approximate size of 275 nm x 120 nm, and a tail-like extension at one end. The double stranded viral DNA has a size greater than 200 kb. WSSV virions, isolated from infected shrimps, contained four major proteins: 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24), and 19 kDa (VP19) in size respectively. VP26 and VP24 were found associated with nucleocapsids; the others were associated with the envelope. N-terminal amino acid sequences of nucleocapsid protein VP26 and the envelope protein VP28 were obtained by protein sequencing and used to identify the respective genes (*vp*26 and *vp*28) in the WSSV genome. To confirm that the open reading frames of WSSV *vp26* (612) and *vp28* (612) are coding for the putative major virion proteins, they were expressed in insect cells using baculovirus vectors and analyzed by Western analysis. A polyclonal antiserum against total WSSV virions confirmed the virion origin of VP26 and VP28. Both proteins contained a putative transmembrane domain at their N-terminus and many putative N- and O-glycosylation sites. These major viral proteins showed no homology to baculovirus structural proteins, suggesting, together with the lack of DNA sequence homology to other viruses, that WSSV may be a representative of a new virus family.

This chapter has been published in a slightly modified version as:
van Hulten, M. C. W., Westenberg, M., Goodall, S. D., and Vlak, J. M. (2000). Identification of two major virion protein genes of White Spot Syndrome virus of shrimp. *Virology* **266,** 227-236.

Introduction

White Spot Syndrome Virus (WSSV) is a major viral disease agent in shrimp in large coastal areas of Southeast Asia and North America. The virus has a wide host range among crustaceans (Flegel, 1997) and distinctive clinical signs (white spots) in penaeid shrimps. There is little genetic variation among WSSV isolates from around the world (Lo *et al.*, 1999). Electron microscopy (EM) studies showed that the virions are enveloped and have a bacilliform shape of about 275 nm in length and 120 nm in width with a tail-like appendage at one end (Wongteerasupaya *et al.*, 1995). Nucleocapsids, which have lost their envelope, have a cross-hatched appearance and a size of about 300 nm x 70 nm (Wongteerasupaya *et al.*, 1995). This virion morphology, its nuclear localization and its morphogenesis are reminiscent of baculoviruses in insects (Durand *et al.*, 1997). Originally, WSSV was classified as an unassigned member of the Baculoviridae (Francki *et al.*, 1991). At present WSSV is no longer accepted into this family (Murphy *et al.*, 1995) due to its unusually wide host range and lack of molecular information. The double stranded viral DNA is larger than 200 kb, as determined from restriction endonuclease analysis (Yang *et al.*, 1997), but sequence information is almost entirely lacking.

Recent analysis of the WSSV DNA revealed the presence of putative genes for the large and the small subunit of ribonucleotide reductase (RR1 and RR2) (van Hulten *et al.*, 2000c). This enzyme is often found encoded by large DNA viruses, including baculoviruses (van Strien *et al.*, 1997). The genes for RR1 and RR2 are the first ORFs in the WSSV genome for which a putative function could be assigned. In order to study the relationship of WSSV to other large DNA viruses, eukaryotes, and prokaryotes, phylogenetic trees were constructed using the amino acid sequences of RR1 and RR2. This analysis showed that WSSV belongs to the eukaryotic branch of an unrooted parsimonious tree and further indicated that WSSV and baculovirus RRs do not share an immediate common ancestor (van Hulten *et al.*, 2000c).

Other genes that are highly conserved in related viruses are those coding for viral capsid or envelope proteins (Murphy *et al.*, 1995). These genes are therefore often used to study virus relatedness. The major capsid proteins were found to be highly conserved among members of the family *Iridoviridae*, *Phycodnaviridae*, and African swine fever virus (Tidona *et al.*, 1998), and were found to be a suitable target for the study of viral evolution of these DNA viruses (Tidona *et al.*, 1998). This is also the case in Poxviridae (Sullivan *et al.*, 1994). In Baculoviridae 70% of the structural virion proteins are well conserved (Ahrens *et al.*, 1997; Gomi *et al.*, 1999) and many baculovirus structural proteins share antigenic determinants (Smith and Summers, 1981). In plant RNA viruses the phylogeny for several different families can also be elucidated based on the viral capsid genes (Dolja *et al.*, 1991; Dolja and Koonin, 1991).

Considering the conserved nature of viral capsid proteins and their use in virus taxonomy and phylogeny, virion proteins of WSSV were analyzed and two genes were identified on the WSSV genome. Their viral identity was further investigated by overexpression of these proteins in the baculovirus-insect cell expression system (Summers and Smith, 1987) and by western analysis using a polyclonal against purified WSSV virions. The two WSSV virion proteins showed no homology to known baculovirus proteins or to other proteins in databases. This result gives further support to the proposition that WSSV might be a representative of a new virus family (van Hulten *et al.*, 2000c).

Results

Identification of WSSV virion proteins

Penaeus monodon shrimp were infected with WSSV by intramuscular injection of a purified virus preparation. Four days after infection virus was isolated from the haemolymph of the infected animals. As a negative control, haemolymph was also taken from uninfected shrimps. These preparations were analyzed by electron microscopy for the presence and purity of WSSV virions. In the samples of uninfected animals no virus particles were observed, but in samples of the infected animals many mainly enveloped virions were observed (Fig. 4.1a).

When the viral envelope was removed from the virus particles after treatment with NP40, the purified nucleocapsids (Fig. 4.1b) had a cross-hatched appearance characteristic for WSSV nucleocapsids (Durand *et al.*, 1997). The proteins of the enveloped virions, the nucleocapsids and the envelope fraction were separated by SDS-PAGE (Fig. 4.1c) and for the nucleocapsid and the evelope fraction, a Western blot using WSSV antiserum was prepared (Fig. 4.1d). Three major protein bands in the range of 67 kDa to 78 kDa were present in the shrimp haemolymph and co-purified with the virions. In the purified WSSV virions (Fig. 4.1c), four major peptides were identified with an apparent molecular mass of 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24) and 19 kDa (VP19), respectively. Several less prominent bands were also observed, approximately six of which are located in the range of 30 to 65 kDa, and at least seven weak protein bands range from 86 kDa to 130 kDa in size. No bands of the major WSSV proteins VP28 and VP19 were present in the lane containing WSSV nucleocapsids, but were weakly visible in the lane containing the envelope fraction (Fig. 4.1c) and clearly visible in the western blot (Fig. 4.1d). They therefore seemed to be derived from the viral envelope or tegument. VP26 and VP24 were present, both in the nucleocapsids and in the virions (Fig. 4.1c), suggesting that they are derived from the nucleocapsid. In the western blot VP26 is clearly visible, however VP24 only weakly as this protein does not react very well with the WSSV antiserum. A low background is visible from the envelope-derived proteins, as these react very strongly with the WSSV antiserum.

The protein from the SDS-PAGE gel was transferred to a polyvinylidene difluoride membrane by semi dry-blotting, and the major viral protein bands were excised and sequenced. More than 40 amino acids were sequenced from the N-terminus of VP28 and VP26 (bold-faced in Fig. 4.2a and 4.2b, respectively). The VP26 N-terminal sequence was M E F G N L T N L D V A I I A I L S I A I I A L I V I M V I M I V F N T R V G R S V V A

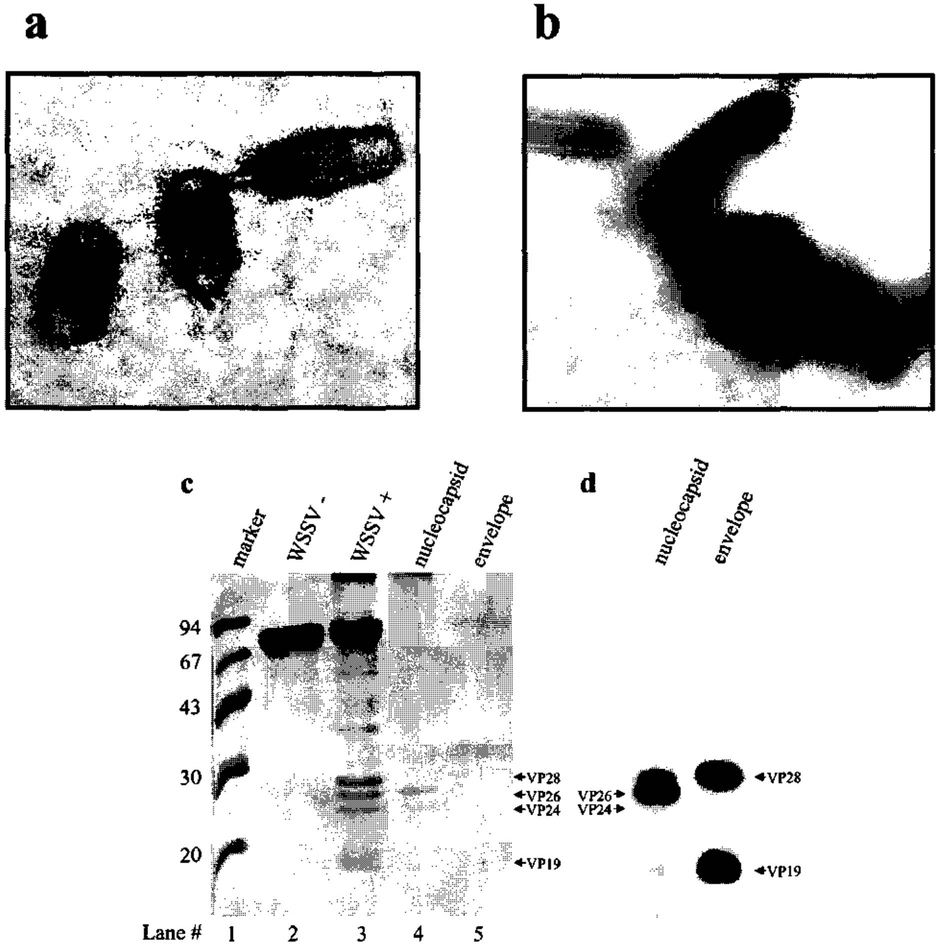


Figure 4.1 (A) Electron microscopic view of negatively stained intact WSSV virions, (B) of negatively stained WSSV nucleocapsids. (C) 15% Coomassie Brilliant Blue-stained SDS PAGE gel of purified WSSV. Lane 1: Low molecular weight protein marker. Lane 2: mock purification from uninfected shrimps. Lane 3: purified WSSV particles. Lane 4: purified WSSV nucleocapsids. Lane 5: WSSV envelope fraction. (D) Western blot of the nucleocapsid and envelope fraction of C. WSSV polyclonal antiserum is used and detection is performed with the ECL kit.

N. N-terminal sequencing of VP28 gave the amino acid sequence M D L S F T L S V V S A I L A I T A V I A V F I V I F R Y H N T V T K T I E t H s D, of which the threonine at position 39 and the serine at position 41 are uncertain. Both N-terminal sequences are hydrophobic (Fig. 4.3).

Localization and sequence of the 26 kDa protein gene

Partial WSSV genomic libraries of *Hin*dIII, and *Bam*HI were constructed in pBluescript-SK+ (van Hulten *et al.*, 2000c), and terminal nucleotide sequences were obtained from many WSSV fragments. The nucleotide sequence encoding the N-terminal amino acid sequence of

VP26 is present near a terminus of a 6 kb *Bam*HI fragment (Fig. 4.2c). The sequence surrounding the methionine start codon (AAA*ATG*G) conformed with the Kozak rule for efficient eukaryotic translation initiation (Kozak, 1989). Only 49 nucleotides of the untranslated leader of *vp26* could be determined, extending towards the terminal *Bam*HI site (Fig. 4.2c).

The 6 kb *Bam*HI fragment contained an open reading frame of 615 nt, including those encoding the N-terminal amino acids of VP26 (Fig. 4.2a). A polyadenylation consensus (poly A) signal is present 34 nt downstream of the translational stop codon of *vp26*. The *vp26* encodes a protein of 204 amino acids with a theoretical size of 22 kDa. The putative protein is basic, with an isoelectric point of 9.3. Three potential sites for N-linked glycosylation (N-{P}-[ST]-{P}) are present, and three putative O-glycosylation sites were predicted using the program NetOglyc (Hansen *et al.*, 1998) (fig. 4.2a). Fourteen possible phosphorylation sites ([ST]-X-X-[DE] or [ST-X-[RK]) were found, but no other motifs present in the PROSITE database.

Hydrophobicity analysis of the amino acid sequence of VP26 showed that a strong hydrophobic region was present at the N-terminus of the protein (Fig. 4.3a). This region contained a putative transmembrane anchor consisting of an α-helix formed by amino acids 12 through 34. The anchor was followed by a positively charged region with two arginines, suggesting that the C-terminal part of the protein is on the cytoplasmic side (Sonnhammer *et al.*, 1998). Besides the transmembrane-spanning α-helix, a potential β-sheet was found at position 127 through 141 using the algorithm of Garnier *et al.* (1978).

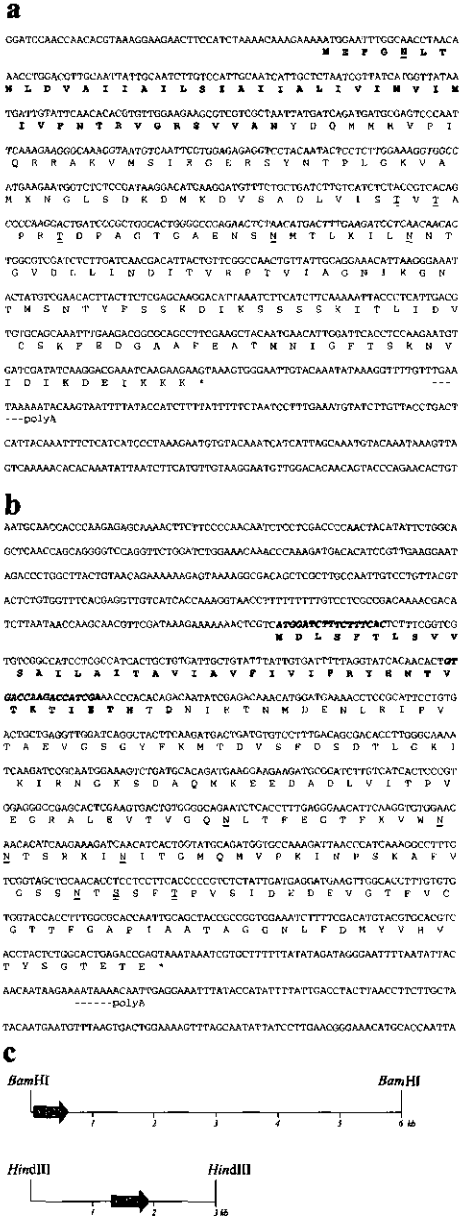


Figure 4.2 Sequence of WSSV VP26 and VP28. (A) Nucleotide and protein sequence of VP26 and (B) of VP28. The N-terminal sequenced amino acids are bold-faced; the location of putative N-glycosylation sites is underlined and of O-glycosylation sites is double underlined. The nucleotide sequence of degenerated primer positions on VP28 is in bold and italics. (C) Location of VP26 and VP28 on WSSV genomic fragments.

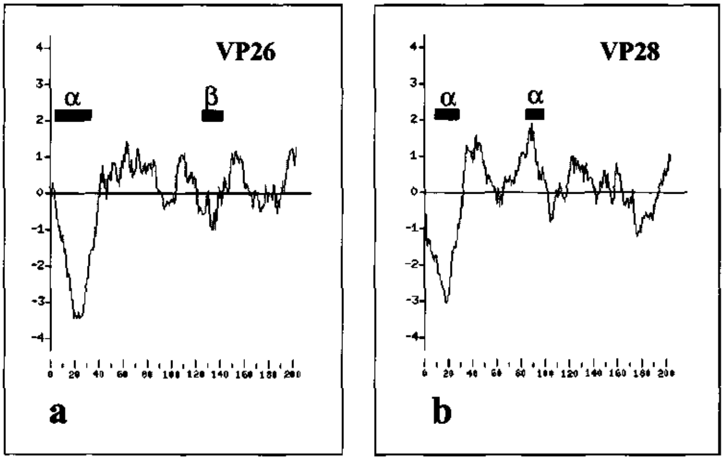


Figure 4.3 Hydrophobicity plots of VP26 (A) and VP28 (B). The amino acid number is on the abcis; the hydrophobicity value on the ordinate. α helices (-) and ß sheets(-) are indicated.

Only one cysteine was present in the protein, indicating that no intra-protein disulfide cross-links can be formed. This cysteine was located in the C-terminal part of the protein.

Localization and sequence of the 28 kDa protein gene

The coding sequence of VP28 could not be determined from sequence analysis of the WSSV DNA fragment termini. Based on the N-terminal protein sequence of VP28 a set of degenerated primers was developed. The forward primer was 5' CA*GAATTC*TCDATNGTY TTNGTNAC 3' and the reverse primer was 5' CA*GAATTC*ATGGAYYTNWSNTTYAC 3'. *Eco*RI sites (*italics*) were incorporated into the primers. The location of the primers in the final sequence is indicated in Fig. 4.2b. PCR was performed using WSSV genomic DNA as template. A 128 bp-long fragment was obtained and, after purification from a 2.5% agarose gel, cloned into pBluescript SK+ and sequenced. The nucleotide sequence of this PCR product encoded the N-terminal protein sequence of WSSV VP28, and this 128 bp fragment was used in a colony lift assay (Sambrook *et al.*, 1989) on several WSSV plasmid libraries. A 3 kb *Hin*dIII fragment hybridized with this fragment and so was further analyzed.

The complete *vp28* ORF and a promoter region of this gene was found on this 3 kb *Hin*dIII fragment (Fig. 4.2b). The methionine start codon (GTC*ATG*G) was in a favorable context for efficient eukaryotic translation initiation (Kozak, 1989). In the promoter region no consensus TATA box was found, but stretches of A/T rich regions were present. A polyA signal was observed 55 nt downstream of the translation stop codon. The *vp28* ORF coded for a putative protein of 204 amino acids, including the N-terminally sequenced amino acids of VP28. The theoretical size of this protein was 22 kDa, and it had an isoelectric point of 4.6. Five potential sites for N-linked glycosylation (N-{P}-[ST]-{P}), two sites for O-glycosylation (Hansen *et al.*, 1998) (Fig. 4.2b) and 9 possible phosphorylation sites ([ST]-X-X-[DE] or [ST-X-[RK]) were found within VP28. No other motifs present in the PROSITE database were found in VP28.

Computer analysis of the 204 amino acids showed that a strong hydrophobic region was present at the N-terminus of VP28 (Fig. 4.3b), including a putative transmembrane α-helix formed by amino acid 9 through 27. As in VP26, this transmembrane anchor sequence is followed by a positively charged region suggesting that the protein might have an outside to inside orientation. At the C-terminal part of the sequence, another hydrophobic region, which might constitute a transmembrane sequence, was found. However, the algorithm of Garnier *et al.* (1978) did not predict an α-helix at this position in VP28. The algorithm predicted another α-helix at position 89 to 99, but no β-sheets along the protein. As in VP26 only one cysteine was present in VP28. This cysteine was also located in the C-terminal part of the protein.

Expression and analysis of recombinant *vp26* and *vp28*.

*Vp*26 and *vp*28 were expressed in the baculovirus insect cell system to determine if the *orfs* represent the major WSSV structural virion proteins and to check the immunoreactivity of these proteins with a WSSV- and *Ac*MNPV-specific polyclonal antisera, respectively. The Bac-to-Bac system (GIBCO BRL) was used to generate recombinant baculoviruses expressing the putative WSSV virion proteins in insect cells. The VP26 and VP28 genes were cloned downstream of the polyhedrin promoter in the plasmid pFastBac-D/GFP, which contains a GFP gene downstream of the p10 promoter. The recombinant viruses generated from pFastBac-D/GFP (control), and the plasmids with VP26 and VP28 were designated *Ac*MNPV-GFP, *Ac*MNPV-WSSVvp26 and *Ac*MNPV-WSSVvp28, respectively. All recombinant viruses expressed GFP off the p10 promoter to facilitate detection and titration; the latter two also expressed VP26 and VP28, respectively, off the polyhedrin promoter.

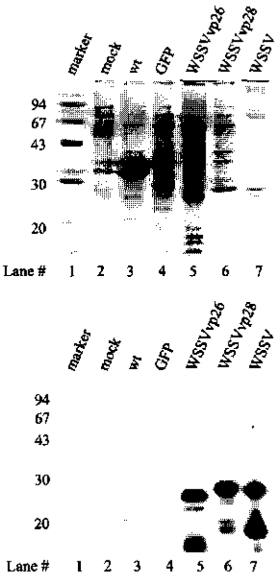


Figure 4 Baculovirus expression of WSSV structural proteins in insect cells. (A) Coomassie Brilliant Blue-stained 15% SDS PAGE gel with low molecular weight protein markers (lane 1), with extracts of *Sf*21 cells (lane 2), infected with wild type *Ac*MNPV infection (lane 3), with *Ac*MNPV-GFP (lane 4), with *Ac*MNPV-WSSVvp26 (lane 5) and *Ac*MNPV-WSSVvp28 (lane 6), and with purified WSSV-virions (lane 7). (B) Western blot of the SDS-PAGE of A. WSSV polyclonal antiserum is used and detection is performed with the ECL kit.

Extracts of *Sf*21 cells infected with *Ac*MNPV-wt, *Ac*MNPV-GFP, *Ac*MNPV-WSSVvp26, and *Ac*MNPV-WSSVvp28 were analyzed on a 15% SDS-PAGE gel. In cells infected with wild type *Ac*MNPV (Fig. 4.4a, lane 3), a 32 kDa band which represented polyhedrin was visible. In the lanes containing extracts of *Ac*MNPV-GFP infected cells (lane 4) and cells infected with the recombinants expressing WSSV proteins (lanes 5 and 6), a GFP protein band was observed at approximately 29 kDa. The GFP expression in the cells infected with *Ac*MNPV-GFP was stronger than the GFP expression in recombinant virus producing WSSV proteins from the polyhedrin promoter (lanes 5 and 6). This was also readily observed after UV illumination of cells infected with the various *Ac*MNPV recombinants, where the fluorescence of GFP in *Ac*MNPV-GFP infected cells was the strongest (not shown). The expression of the WSSV proteins from the polyhedrin promoter was significantly higher than the expression of GFP from the p10 promoter (lane 5 and 6). Strong expression of a 26 kDa protein was observed in extracts of *Ac*MNPV-WSSVvp26 infected cells, most likely representing WSSV VP26 (lane 5). Cells infected with *Ac*MNPV-WSSVvp28 (lane 6) showed a strong expression of a 28 kDa protein. The position of GFP in these gels was confirmed by Western analysis using anti-GFP antiserum (data not shown).

Western analysis was performed on samples from *Sf*21 cells infected with wild-type and recombinant *Ac*MNPV. A polyclonal antibody against WSSV virions was used to detect recombinant VP26 and VP28 (Fig. 4.4b). Both VP26 and VP28 were detected in these cell extracts. VP26 was detected at 26 kDa, in conformity with the Coomassie Brilliant Blue-stained gel (fig. 4.4a, lane 5; fig. 4.4b, lane 5). This is identical to its size when directly extracted from WSSV virions, but lower compared to its theoretical size of 22 kDa. VP28 migrated at the same position as VP28 from WSSV virions, which is significant higher than the theoretical size of 22 kDa for this protein. The polyclonal antibody did not show major cross reactivity with insect cells (lane 2) or baculovirus proteins (lanes 3 and 4), as observed from the very low background reaction in these samples.

Relatedness of VP26 and VP28 to other viruses

Homology searches with WSSV VP26 and VP28 were performed against GenBank/EMBL, SWISSPROT and PIR databases using FASTA, TFASTA and BLAST. No significant homology with baculovirus envelope or capsid proteins, or with structural proteins from other large DNA viruses could be found with the sequences in the GenBank.

Discussion

Four major and several less prominent protein bands were observed, when purified WSSV virions were analyzed in SDS-PAGE (Fig. 4.1). The major protein bands had approximate sizes of 19 kDa (VP19), 24 kDa (VP24), 26 kDa (VP26), and 28 kDa (VP28), respectively. Only two of these, VP24 and VP26, were observed in the nucleocapsid preparation, suggesting that these are the major components of the nucleocapsid of WSSV. The other major virion proteins, VP19 and VP28, most likely are constituents of the virion envelope or the tegument, as they were stripped off the virion by NP40 treatment. The sizes found for the major virion proteins were similar to those described by Hameed *et al.* (1998) and Nadala *et al.* (1998). A major difference is the number of protein bands observed. Hameed *et al.* (1998) and Nadala *et al.* (1998) observed three major proteins in the range of

18 kDa to 28 kDa, whereas we observed four major proteins in this size range (Fig. 4.1). Whether this difference is due to variation in WSSV isolates or the result of the WSSV purification procedure remains to be investigated. The 19 kDa and a 27.5 kDa protein which Nadala *et al.* (1998) observe in the envelope fraction have the same size as VP19 and VP28 from our envelope fraction and may therefore be the same proteins. Furthermore these authors observed a 23.5 kDa protein in the nucleocapsid fraction, which may correspond to VP24 or VP26, both of which are found in the nucleocapsid fraction.

Two of the major structural WSSV proteins, one from the envelope fraction (VP28) and one from the nucleocapsid preparation (VP26), were selected for further analysis, in particular to study their relatedness to structural proteins of other viruses including baculoviruses. The N-terminal amino acid sequence of these proteins was used to locate the ORFs coding for these proteins on the WSSV genome by direct sequencing (VP26) or by using degenerated primers and colony lifting (VP28). As such these are the first WSSV virion proteins, whose genes have been identified. Antibodies against VP26 and/or VP28 are being generated and may serve as a specific diagnostic reagent to detect WSSV infection in shrimp. Furthermore, these antibodies could be used in immunogold labeling of virus particles. Such a study would provide visualized evidence that the proteins are indeed structural virion proteins and exclude the possibility that they are fortuitously associated with the virus particles.

The ORFs of *vp26* and *vp28* coded for proteins with theoretical sizes of 22 kDa. The authentic N-terminal amino acid sequences of VP26 and VP28 were confirmed by direct protein sequencing. Both protein sequences contain multiple putative N- and O-glycosylation sites as well as phosphorylation sites (Fig 4.2a and 4.2b). These theoretical sizes differed by approximately 4 and 6 kDa, respectively, from their mobility in Coomassie Brilliant Blue-stained SDS-PAGE gels and can be explained either by post-translational modifications such as glycosylation and phosphorylation, or by splicing. Expression in the baculovirus insect cell system was performed to determine whether the ORFs encoded the major WSSV structural virion proteins and to confirm the identity and coding capacity of these virion proteins. Furthermore the immunoreacitivity of these proteins with a WSSV- and AcMNPV-specific polyclonal antisera was investigated. Both VP26 and VP28 were highly expressed from the polyhedrin promoter in cells infected with recombinant *Ac*MNPV (Fig. 4.4). The theoretical sizes of both proteins encoded by the *vp28* and *vp26* ORF, respectively was 22 kDa. Expression in insect cells resulted in a protein bands of 28 kDa and 26 kDa, respectively, which is the same size as the proteins from WSSV isolated from infected shrimp. This suggests that posttranslational modifications of these proteins may have been correctly performed in insect cells. The recombinant WSSV VP26 and VP28 can be used to generate monospecific polyclonal antibodies for studies on the structure and morphogenesis of WSSV virions.

Linear and circular membranous structures appear to be formed *de novo* in the nucleoplasm of cells of WSSV infected shrimps (Durand *et al.*, 1997). Similar virus induced intranuclear microvesicles are also found in the nucleus of baculovirus infected cells, where *de novo* synthesis of viral envelopes occurs (Williams and Faulkner, 1997). The viral envelope protein encoded by *Ac*MNPV-*orf46* contains a hydrophobic domain on the extreme N-terminus, which is sufficient to direct this protein into these intranuclear microvesicles (Hong *et al*, 1997). Both VP26 and VP28 of WSSV virions have a hydrophobic domain on their N-terminus, which might have a similar signal function for transport and also may be responsible for the membranous structures associated with WSSV infection *in vivo* (Durand *et al.*, 1997). VP26, the major protein of the nucleocapsid, appears to be a very basic protein with a theoretical isoelectric point of 9.4. Because of its basic character it may have a close association with the viral DNA.

The objective of this study was also to reveal possible homology of WSSV virion proteins with structural proteins of other viruses including baculovirus. Western analysis confirmed the WSSV origin of the expressed VP26 and VP28. Both peptides gave a strong reaction with the polyclonal antiserum against purified WSSV (Fig. 4.4b). The absence of any reaction of this WSSV antiserum with proteins of wild type *Ac*MNPV infected cells also indicated that there is no serological relationship between WSSV virion proteins and baculovirus proteins. Baculovirus structural protein do share antigenic determinants (Smith and Summers, 1981). When the WSSV nucleocapsid protein (VP26) and the envelope or tegument protein (VP28) sequences are compared with proteins known to be present in the baculovirus virion (e.g. *Ac*MNPV-gp41, *Ac*MNPV-vp39), or envelope (e.g. *Ac*MNPV-gp64, *Ac*MNPV-p25) and other sequences available in the GenBank, no significant homologies were found. From this analysis we conclude that the WSSV virion structural proteins described in this paper, do not have any homology with baculovirus structural proteins. However, it cannot be excluded that other WSSV virion proteins, such as VP24 and VP19, may show homology to baculovirus proteins.

In the latest taxonomic revision of viruses by the International Committee on Taxonomy of Viruses (Murphy *et al.*, 1995), WSSV has not been classified due to the lack of molecular information. In a previous report (Francki *et al.*, 1991), WSSV was classified as a non-occluded baculovirus based on the rod to bacilliform shape of the WSSV virion. Ongoing sequence analysis (35,000 bp) of the WSSV genome so far indicated that WSSV has only very limited (*rr* genes), if any, homology with baculovirus sequences to date, or with any other known virus or organism. In contrast, baculoviruses do exhibit homology of a high proportion (>50%) of their genes, including those encoding virion structural proteins (Kuzio *et al.*, 1999; Hu *et al.*, 1998). The two major WSSV structural proteins investigated here do not share any homology with baculovirus virion proteins. This supports the view that WSSV may be a representative of a new virus family (proposed name Whispoviridae). The recent results on the phylogenetic analysis of *rr* genes of WSSV (van Hulten *et al.*, 2000c) give further credence to the latter view.

Materials and methods

White Spot Syndrome Virus production and purification

The virus used in this study was isolated from infected *Penaeus monodon* shrimps from Thailand. Infected tissue was homogenized in TN buffer (20 mM Tris-HCl, 400 mM NaCl, pH 7.4). After centrifugation at 1,700 x g for 10 min the supernatant was filtered (0.45 µm filter) and injected intramuscularly into healthy *P. monodon* in the lateral area of the fourth abdominal segment to initiate infection. After 4 days haemolymph was withdrawn from moribund shrimps and mixed with modified Alsever solution (Rodriguez *et al.*, 1995) as an anticoagulant. After dilution in TNE (20 mM Tris-HCl, 400 mM NaCl, 5 mM EDTA, pH 7.4) the haemolymph was clarified from haemocytes at 1,700 x g for 10 min at 4°C. The virus particles were then sedimented by centrifugation at 45,000 x g at 4°C for 1 h and resuspended in TN-buffer.

The virus envelope was removed from the virus particles by treatment with Nonidet P40 (NP40). One percent NP40 was added to the virus solution and incubated for 30 min at room temperature with gentle rocking. The nucleocapsids were sedimented at 80,000 x g for 30 min at 4°C. The pellet was dissolved in TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). The proteins in the supernatant were acetone-precipitated and resuspended in TE.

Electron microscopy

For transmission electron microscopy (TEM), the virus suspension was mounted on formvar-coated, carbon-stabilized nickel grids (400 mesh) and negatively stained with phosphotungstic acid (2% PTA). The specimens were examined in a Philips CM12 electron microscope.

Nucleic acid purification

Viral DNA was isolated from purified virions by treatment with proteinase K (0.2 mg/ml) and sarcosyl (1%) at 45°C for 3 h, followed by phenol/chloroform extraction and dialysis against TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). The purity and concentration of the DNA was determined by agarose gel electrophoresis.

Plasmid constructions

WSSV subgenomic fragments were generated by restriction enzyme analysis, cloned into pBluescript SK+ (Stratagene) and transformed into *E. coli* DH5α using standard techniques (Sambrook *et al.*, 1989). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis and colony lifting were carried out according to standard protocols (Sambrook *et al.*, 1989). PCR was performed using custom designed and synthesized primers. DNA encoding the N-terminus of *vp28* was amplified by PCR from total WSSV DNA using degenerated primers based on the N-terminal amino acid sequence of VP28. The forward primer used was 5' CA*GAATTC*TCDATNGTYT TNGTNAC 3' and the reverse primer was 5' CA*GAATTC*ATGGAYYTNWSNTT YAC 3' with *Eco*RI sites (italics) (D = A, T or G; N = A, C, G, or T; Y = C or T; W = A or T; S = C or G).

DNA sequencing and computer analysis

Plasmid DNA for sequencing was purified using the QIAprep Miniprep System or JETstar Plasmid Purification System (Qiagen, Inc.). Sequencing was performed using the universal pBluescript forward and reverse primers and custom primers for both strands. Automatic sequencing was carried out using an Applied Biosystems automated DNA sequencer (Eurogentec, Belgium).

The generated sequences were analyzed with UWGCG computer programs (release 10.0). The DNA and the deduced amino acids sequences were compared with GenBank/EMBL, SWISSPORT and PIR databases using the programs FASTA, TFASTA (Pearson & Lipman, 1988) and BLAST (Altschul *et al.*, 1997).

The nucleotide sequence of the WSSV 26 kDa (VP26) and the WSSV 28 kDa (VP28) protein genes have been deposited in GenBank and were assigned accession number AF173992 and AF173993, respectively.

Cells and viruses

Spodoptera frugiperda (*Sf*-AE-21) cells (Vaughn *et al.*, 1977) were cultured in Grace's insect medium (GIBCO BRL) supplemented with 10% foetal calf serum (FCS). The E2 strain of *Autographa californica* multicapsid nucleopolyhedrovirus (*Ac*MNPV) (Summers and Smith, 1978) was used as wild type (wt) virus. Routine cell culture maintenance and virus infection procedures were carried out according to published procedures (Smith and Summers, 1987; King and Possee, 1992).

Engineering of recombinants

The Bac-to-Bac system (GIBCO BRL) was employed to overexpress WSSV VP26 and VP28 in insect cells. To facilitate detection and titration of Bac-to-Bac recombinants upon infection of insect cells, the Green Fluorescent Protein (GFP) gene was introduced into the pFastBac-DUAL vector downstream of the p10 promoter. The GFP gene was removed from plasmid pVL92GFP (Reilander *et al.,* 1996) after digestion of this plasmid with *Xba*I and *Kpn*I. The 700 bp GFP-containing fragment was isolated by agarose gel electrophoresis and GlassMAX purification (GIBCO BRL), blunt-ended using DNA polymerase and inserted into the *Sma*I site of multiple cloning region II of pFastBac-DUAL downstream of the p10 promoter. The resulting plasmid was named pFastBac-D/GFP, and contained region I for insertion of a foreign gene downstream of the polyhedrin promoter. Recombinant virus, expressing only the GFP from the p10 promoter, was constructed according to the Bac-to-Bac system protocol (GIBCO BRL) and the resulting virus was designated *Ac*MNPV-GFP.

PCR was performed on the WSSV plasmids containing the putative complete open reading frames (ORFs) of *vp26* and *vp28*. A *Bam*HI site was introduced at the 3' end of the ORFs and a *Hin*dIII site at the 5' end. *Vp26* and *vp28* were first cloned into the pET28a vector (Novagen), excised with *Bam*HI and *Not*I, and inserted downstream of the polyhedrin promoter of plasmid pFastBac-D/GFP. The resulting plasmids were named pFastBac-D/G-vp26 and pFastBac-D/G-vp28, respectively. Recombinant viruses expressing the GFP from the p10 promoter and VP26 or VP28 from the polyhedrin promoter were constructed according to the Bac-to-Bac system protocol (GIBCO BRL), and the viruses were designated *Ac*MNPV-WSSVvp26 and *Ac*MNPV-WSSVvp28, respectively.

SDS-PAGE, protein sequencing and immunoblotting

Insect cells infected with wild type *Ac*MNPV, and insect cells infected with recombinant *Ac*MNPV expressing heterologous proteins (GFP, VP26, VP28), and purified WSSV were analyzed in 15% SDS-PAGE gels as described in Laemmli (1970). Proteins were visualized using Coomassie Brilliant Blue staining. Semi-dry blotting was performed onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a CAPS buffer (10 mM CAPS in 10% Methanol), or onto an Immobilon™-P (Millipore) using a Tris-Glycine buffer (25 mM Tris base, 192 mM glycine, 10% (v/v) methanol, pH 8.3). Proteins were visualized on the PVDF membrane using Coomassie Brilliant Blue staining. Major protein bands from WSSV virion preparations were excised from the filter and N-terminal sequenced (ProSeq. Inc., Massachusetts).

Immobilon-P membranes were blocked in 2% low-fat milk powder (Campina, the Netherlands) in TBS (0.2 M NaCl, 50 mM Tris-HCl, pH 7.4). Immunodetection was performed by incubation of the blot in a polyclonal rabbit anti-WSSV serum (a gift from Prof. P.C. Loh, University of Honolulu, Hawaii) diluted 1:2000 in TBS with 0.2% low-fat milk powder for 1 h at room temperature. Subsequently, anti-rabbit antibody conjugated with horseradish peroxidase (Amersham) was used at a concentration of 1:2000 and detection was performed with an Enhanced Chemiluminescent-light Detection Kit (Amersham).

Acknowledgment

This research was supported by Intervet International BV, Boxmeer, The Netherlands, through financial aid from the Ministry of Economic Affairs. The polyclonal rabbit anti-WSSV serum was a generous gift from Dr. Ph. Loh, University of Honolulu, Hawaii. This paper is dedicated to dr. Jan H. Boon (✝ 1997), Laboratory of Fisheries, Agricultural University, who initiated this study.

Chapter 5

Three functional diverged major envelope proteins of white spot syndrome virus evolved by gene duplication

White Spot Syndrome Virus (WSSV) is an invertebrate virus causing considerable mortality in penaeid shrimp. Virions are enveloped nucleocapsids with an oval to bacilliform morphology of approximately 275 nm x 120 nm, and a tail-like extension. WSSV virions, isolated from infected shrimp (*Penaeus monodon*), contain four major proteins: VP28 (28 kDa), VP26 (26 kDa), VP24 (24 kDa) and VP19 (19 kDa). VP26 and VP24 are associated with nucleocapsids and the remaining two with the envelope. Forty-one N-terminal amino acids of VP24 were obtained, allowing identification of its gene (*vp24*) in the WSSV genome and comparison with databases. Computer-assisted analysis revealed a striking sequence similarity between WSSV VP24, VP26 and VP28 at the amino acid level, as well at the on nucleotide level. This strongly suggests that their genes have evolved by gene duplication and subsequently diverged into proteins with different functions in the WSSV virion, i.e. envelope and nucleocapsid. Several conserved domains were identified in the three WSSV structural proteins, which suggests that the proteins may have a similar structure. None of the three structural WSSV proteins analysed showed homology to proteins of other viruses, underscoring the distinct taxonomic position of WSSV among other invertebrate viruses.

This chapter has been published as:
van Hulten, M. C. W., Goldbach, R. W., and Vlak, J. M. (2000). Three functionally diverged major structural proteins of white spot syndrome virus evolved by gene duplication. *Journal of General Virology* **81**, 2525-2529.

White Spot Syndrome Virus (WSSV) is a rapidly emerging viral disease agent in shrimp in Southeast Asia and the Americas. The virus has a wide host range among crustaceans (Flegel, 1997) and induces distinctive clinical signs (white spots) under the carapace of penaeid shrimps. WSSV has a double stranded DNA genome with a size exceeding 250 kb (Yang *et al.*, 1997) and may be a representative of a new floating genus Whispovirus (van Hulten *et al.*, 2000c, a; van Hulten and Vlak, 2001b). Electron microscopy studies revealed that WSSV virions are enveloped rod-shaped nucleocapsids with a bacilliform to ovoid shape of about 275 nm in length and 120 nm in width. Most characteristic is the tail-like appendage at one end of the virion (Wongteerasupaya *et al.*, 1995; Durand *et al.*, 1997). WSSV nucleocapsids have a striated appearance and a size of about 300 nm x 70 nm (Wongteerasupaya *et al.*, 1995). The striations are probably the result of stacked ring-like structures consisting of rows of globular subunits of about 10 nm in diameter (Durand *et al.*, 1997; Nadala *et al.*, 1998).

The characterisation of the structural proteins and their genomic sequence is of major importance to determine the taxonomic position of the virus. Furthermore, the structure and interaction of the WSSV virion proteins may explain the unique morphological features of this virus. Finally, diagnostic tests could be designed based on one or more of these structural proteins. VP28 and VP26, present in the envelope and nucleocapsid, respectively, were identified previously and showed no homology with sequences available in GenBank (van Hulten *et al.*, 2000a). Here we report the identification of a third major structural protein (VP24) and the surprising relatedness of this protein to the previously identified WSSV structural proteins.

Purified WSSV was used to infect *P. monodon* by intramuscular injections in the lateral area of the fourth abdominal segment. Virions were purified from haemolymph of infected *P. monodon* as described by van Hulten *et al.* (2000a). As a negative control, haemolymph was taken from uninfected *P. monodon*. The preparations were analysed by electron microscopy

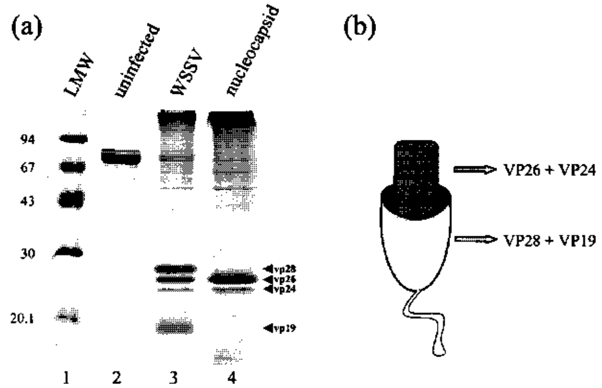


Figure 5.1 (a) Coomassie Brilliant Blue-stained SDS-15% PAGE gel of purified WSSV. Lane 1: Low molecular weight protein marker. Lane 2: mock purification. Lane 3: purified WSSV particles. Lane 4: purified WSSV nucleocapsids. (b) Schematic presentation of the WSSV virion.

for the presence and purity of WSSV virions (not shown). The viral envelope was removed from the nucleocapsid by treatment with 1% Nonidet-P40 (van Hulten *et al.*, 2000a). In the intact WSSV virions purified from *P. monodon* (Fig. 5.1a, lane 3), four major polypeptide species were identified with an apparent molecular mass of 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24), and 19 kDa (VP19). From the SDS-PAGE analysis (Fig. 5.1a) it can be seen that VP26 and VP24 are the major proteins present in the purified nucleocapsids (Fig 5.1a, lane 4). VP28 and VP19 are removed by the NP40 treatment and therefore associated with the viral envelope or tegument (van Hulten *et al.*, 2000a). In Fig. 5.1b a schematic presentation of the WSSV virion is shown.

The sizes found for the major virion proteins were similar to those described by Hameed *et al.* (1998) and Nadala *et al.* (1998), but somewhat different from those described by Wang *et al.* (2000b). The latter authors described the presence of three major proteins in WSSV isolates from different origins with slightly different sizes of 25 kDa, 23 kDa and 19 kDa, respectively. The N-terminal sequences of these proteins demonstrate, however, that the 25 kDa and 23 kDa proteins correspond to our VP28 and VP26, respectively. In our WSSV isolate a protein of 24 kDa (VP24) is clearly a major component of the nucleocapsid. Here we describe the amino acid and genomic sequence of WSSV VP24 and some characteristics of this protein.

VP24 isolated from *P. monodon* was transferred from a SDS-PAGE gel onto a polyvinylidene difluoride membrane by semi dry-blotting. The 24 kDa band was excised from two WSSV preparations and each was sequenced by Edman degradation as described previously (van Hulten *et al.*, 2000a). The first N-terminal sequence obtained was MHMWGVYAAILAGLTLILVVIdI, of which the aspartic acid at position 22 was uncertain. From the second VP24 band more than 40 residues were sequenced (bold-faced in Fig. 5.2) giving the sequence MHMWGVYAAILAGLTLILVVISIVVTNIELNKKLDKKDKdA, in which a serine residue was found at position 22 and an uncertain aspartic acid at position 40.

Based on this partial VP24 sequence a set of degenerate PCR primers was developed, with 5' CA*GAATTC*ATGCAYATGTGGGGNGT 3' as forward primer, and 5' CA*GAATTC*YTTRTCYTTYTTRTCIARYTT 3' as reverse primer, both containing *Eco*RI sites (*italics*) for cloning purposes. The location of the primers in the final sequence is indicated in Fig. 5.2. PCR was performed using WSSV genomic DNA as template. A 133 bp-long fragment was obtained and, after purification from a 2% agarose gel, cloned into pBluescript SK+ and sequenced. The sequence of this PCR product corresponded with the N-terminal protein sequence of WSSV VP24 and was used as probe in a colony lift assay (Sambrook *et al.*, 1989) on WSSV plasmid libraries (van Hulten *et al.*, 2000c) to identify the complete ORF for VP24. An 18 kb *Bam*HI fragment hybridising with this fragment was selected for further analysis.

The complete *vp24* ORF, encompassing 627 nucleotides, and the promoter region of this gene were found on the 18 kb *Bam*HI fragment (Fig. 5.2). The translational start codon was in a favourable context (AAA*ATG*C) for efficient eukaryotic translation initiation (Kozak,

AAAGGCTATCCAGTACTTCAACAAAACTAGCAGAAATAATACTGCACATCATTTCAAGATGCCGGCTTCA

AAGGATCGCAACTATTCTTCTTTCGAGTATGCTGAAACGGCAGTTGCTGCTCACAATATATCTCAGTGGT

GTTTTCTACGACTAAATAAAATATAACAAACTTTAATATCTGTTTTTTTCTCTCATGACCTTTGTACAAC

TTTCTTCAAA***ATGCACATGTGGGGGGTT***TACGCCGCTATACTGGCGGGTTTGACATTGATACTCGTGGTT
M H M W G V Y A A I L A G L T L I L V V

ATATCTATAGTTGTAACCAACATAGAACTTAACAAG***AAATTGGACAAGAAGGATAAA***GACGCCTACCCTG
I S I V V T N I E L N K K L D K K D K D A Y P V

TTGAATCTGAAATAATAAACTTGACCATTAACGGTGTTGCTAGAGGAAACCACTTTAACTTTGTAAACGG
E S E I I N L T I N G V A R G N H F N F V N G

CACATTACAAACCAGGAACTATGGAAAGGTATATGTAGCTGGCCAAGGAACGTCCGATTCTGAACTGGTA
T L Q T R N Y G K V Y V A G Q G T S D S E L V

AAAAAGAAAGGAGACATAATCCTCACATCTTTACTTGGAGACGGAGACCACACACTAAATGTAAACAAAG
K K K G D I I L T S L L G D G D H T L N V N K A

CCGAATCTAAAGAATTAGAATTGTATGCAAGAGTATACAATAATACAAAGAGGGATATAACAGTGGACTC
E S K E L E L Y A R V Y N N T K R D I T V D S

TGTTTCACTGTCTCCAGGTCTAAATGCTACAGGAAGGGAATTTTCAGCTAACAAATTTGTATTATATTTC
V S L S P G L N A T G R E F S A N K F V L Y F

AAACCAACAGTTTTGAAGAAAAATAGGATCAACACACTTGTGTTTGGAGCAACGTTTGACGAAGACATCG
K P T V L K K N R I N T L V F G A T F D E D I D

ATGATACAAATAGGCATTATCTGTTAAGTATGCGATTTTCTCCTGGCAATGATCTGTTTAAGGTTGGGGA
D T N R H Y L L S M R F S P G N D L F K V G E

AAAATAAACACATTTTTTAAAAATATACCTTTATTTTATTAACCAATTATTTACATTGTAGCTTGAAATT
K *

CATATATATACCTTTCATTGTTAGGGAGAGATTTATTGTAGTCAAACCCAG

Figure 5.2 Nucleotide and protein sequence of WSSV VP24. The N-terminal sequenced amino acids are bold-faced; the location of putative N-glycosylation sites is underlined and of O-glycosylation sites is double underlined. The nucleotide sequence of degenerate primer positions is in bold and italics.

1989). In the promoter region stretches of A/T rich sequence, but no consensus TATA box, were found. A polyA signal overlapped the translation stop codon. The *vp24* ORF encoded a putative protein of 208 amino acids with an amino acid sequence containing the experimentally determined N-terminal sequence of VP24. VP24 has a theoretical size of 23 kDa and an isoelectric point of 8.7. Four potential sites for N-linked glycosylation (N-{P}-[ST]-{P}), one site for O-glycosylation (Hansen *et al.,* 1998) (Fig. 5.2) and 9 possible phosphorylation sites ([ST]-X-X-[DE] or [ST-X-[RK]) were found within VP24, but it is not known whether any of these modifications do occur. No other motifs present in the PROSITE database were found in VP24. Computer analysis of the 208 amino acids showed that a strong hydrophobic region was present at the N-terminus of VP24 (Fig. 5.3a), including a putative transmembrane α-helix formed by amino acid 6 through 25. The algorithm of Garnier *et al.* (1978) predicted several other α-helices and β-sheets along the protein. It is remarkable that

VP28, VP26 and VP24 roughly have the same size (~ 206 amino acids) but have distinct electrophoretic mobilities. This may be due to differences in isoelectric points, conformational differences or post-translational modifications.

Homology searches with WSSV VP24 were performed against GenBank/EMBL, SWISSPROT and PIR databases using FASTA, TFASTA and BLAST, but no significant homology with structural proteins from other large DNA viruses could be found. Surprisingly, statistically significant similarity was found with the sequence of two other WSSV virion structural proteins, VP26 and VP28 (van Hulten *et al.*, 2000a), with 41% and 46% amino acid similarity, respectively. Also VP28 and VP26 showed a similarity of 41%. An alignment was made of the three WSSV proteins using ClustalW (Thompson *et al.*, 1994), and revealed several conserved regions (Fig. 5.3b). In the N-terminal region a well-conserved stretch of amino acids is observed at positions 15 – 30. A strong hydrophobic region with an α-helix is observed for all three proteins in the hydrophilicity plots (Fig. 5.3a). These residues might represent a transmembrane region, or be involved in the interaction of the structural proteins to form homo- or heteromultimers. In two other conserved regions, around positions 88 – 102 and positions 138 – 148, the algorithm of Garnier *et al.* (1978) predicted an α-helix and a β-sheet for all three proteins (Fig. 5.3a, 5.3b). Two-thirds of the

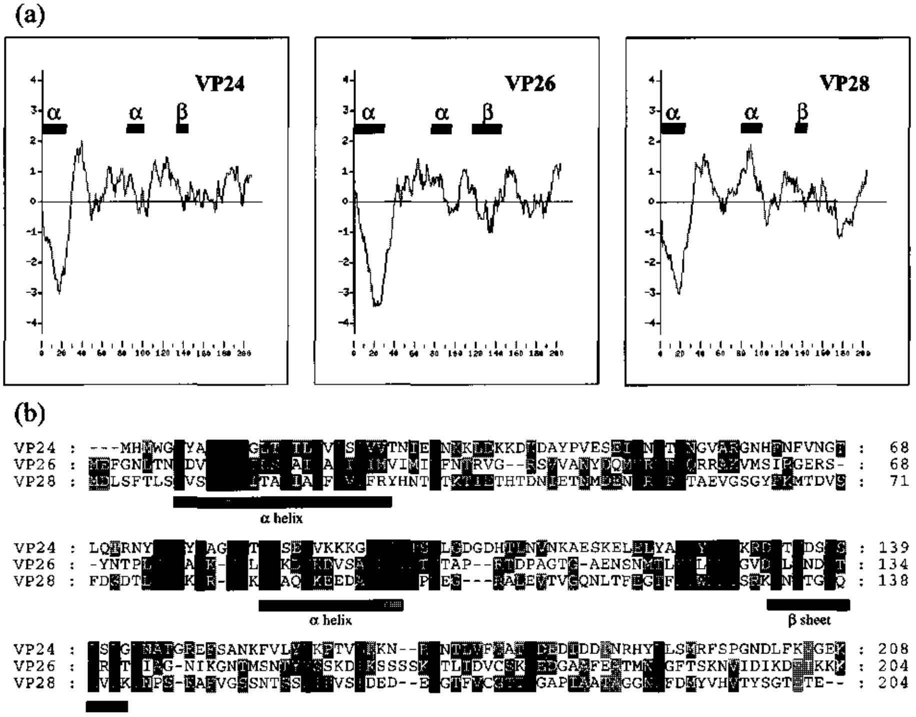


Figure 5.3 (a) Hydrophilicity plots of WSSV VP24, VP26 and VP28. The amino acid number is on the abscissa , and the hydrophilicity value on the ordinate. α-helices and β-sheets are indicated (b) Amino acid sequence alignment of VP24, VP26 and VP28. Shading is used to indicate the occurrence (black 100%, grey 67 %) of similar amino acids. Conserved α-helices and β-sheets are indicated.

residues conserved in the three proteins were hydrophobic and might be involved in the folding of the proteins, giving them a similar structure. The nucleocapsid and envelope proteins differed in their isoelectric points. The two nucleocapsid proteins (VP26 and VP24) both had a basic character with isoelectric points of 9.3 and 8.7, respectively, and might therefore have a close association with the viral DNA, whereas the envelope protein (VP28) was more acidic with an isoelectric point of 4.6.

As there is a high homology at the amino acid level among the three structural WSSV proteins, and conserved domains are present, there is reason to believe that their structures are similar. The presence of the hydrophobic domain indicates that these proteins most probably are capable of forming homo- and heteromultimers. Studies on the interaction of these proteins and their location in the virion are required to substantiate this hypothesis.

A way to explain the high degree of amino acid similarity of the three structural WSSV proteins is to assume that these genes have evolved by gene duplication and divergence. Nucleotide comparisons supported this hypothesis, as significant homology was found. Alignment of *vp*24, *vp*26 and *vp*28, revealed that *vp*24 has 40% nucleotide identity with *vp*26 and 43% with *vp*28, whereas *vp*26 has 48% nucleotide identity with *vp*28. The data presented here strongly suggest that these three WSSV structural protein genes share a common ancestor.

The most surprising observation might be that these proteins have evolved to give proteins with different functions in the WSSV virion, i.e. in the nucleocapsid and the envelope. Such a situation is unusual in animal DNA viruses, although a parallel may exist for the virion glycoproteins of alpha-herpesviruses as their genes might have evolved by duplication and divergence (McGeoch, 1990). However, the homology of these genes is considerably lower than the homology among the WSSV virion genes. The function of the alpha-herpesvirus genes also has not diverged. Gene duplication and functional divergence, however, can be observed in the plant-infecting Closteroviruses, where a minor regulatory protein VP24 appears to be a diverged copy of the coat protein (Boyko *et al.*, 1992). Also in the animal Rhabdoviruses a structural and a non-structural glycoprotein may have evolved from a common ancestral gene (Wang and Walker, 1993).

Structural proteins are well conserved within virus families. However, the three WSSV structural proteins identified so far have no homology to structural proteins of other viruses. The unique feature of the homologous structural virion proteins further supports the proposition that WSSV might be a representative of a new virus genus (Whispovirus) or perhaps a new family (Whispoviridae) (van Hulten *et al.*, 2000c, a; van Hulten and Vlak, 2001b).

Acknowledgement

This research was supported by Intervet International BV, Boxmeer, The Netherlands.

Chapter 6

Identification of VP19 and VP15 and glycosylation status of the white spot syndrome virus major structural proteins

White spot syndrome virus (WSSV) is a virus infecting Penaeid shrimp and other crustaceans. The WSSV virion consists of an enveloped rod-shaped nucleocapsid enclosing a large circular dsDNA genome of 293 kb. The virion envelope contains two major proteins of 28 (VP28) and 19 kDa (VP19) and the nucleocapsid consists of three major proteins of 26 (VP26), 24 (VP24), and 15 kDa (VP15). Study on the morphogenesis of the WSSV particle requires the genomic identification and chemical characterization of these WSSV virion proteins. An internal amino acid sequence of envelope protein VP19 was obtained by amino acid sequencing and used to locate the VP19 open reading frame of this protein on the genome, as WSSV ORF182. VP19 contained two putative transmembrane domains, which may anchor this protein in the WSSV envelope. Similarly, the gene for VP15 was located on the WSSV genome, as ORF109. N-terminal amino acid sequencing on VP15 suggested that this protein was expressed from the second ATG of its ORF and the first methionine is lost by N-terminal protein processing. The 15 kDa protein is very basic and a candidate DNA binding protein in the WSSV nucleocapsid. None of the five major structural WSSV proteins appear to be glycosylated, which is an unusual feature among enveloped animal viruses.

This chapter has been accepted for publication in *Journal of General Virology* as:
van Hulten, M. C. W., Reijns, M., Vermeesch, A. M. G., Zandbergen, F., and Vlak, J. M. (2001). Identification of VP19 and VP15 and glycosylation status of the white spot syndrome virus major structural proteins.

Introduction

White Spot Syndrome Virus (WSSV) is a large enveloped virus infecting shrimp and other crustaceans. The WSSV virion consists of an enveloped nucleocapsid containing a circular double-stranded DNA genome of 292,967 bp (van Hulten *et al.*, 2001a). The enveloped virions are symmetrical particles ellipsoid to bacilliform in shape, and measure about 120 – 150 nm in diameter at the center by 270 – 290 nm in length. Most notable is a tail-like appendage at one end of the virion. The isolated nucleocapsids measure 65 – 70 nm in diameter and 300 - 350 nm in length and have a cross-hatched appearance. The virions contain five major and an unknown number of minor polypeptides. WSSV structural proteins VP28 and VP19 are located in the envelope and VP26, VP24 and VP15 are located in the nucleocapsid (van Hulten *et al.*, 2001a). The open reading frames (ORFs) coding for VP28, VP26 and VP24 have been located on the WSSV genome by N-terminal amino acid sequencing of these proteins (van Hulten *et al.*, 2000a, b). Further studies on the major envelope protein VP28 have shown that this protein plays a key role in the systemic infection of WSSV in shrimp (van Hulten *et al.*, 2001b).

VP28, VP26 and VP24 may have evolved by gene duplication (van Hulten *et al.*, 2000b). They all are encoded by ORFs of roughly the same size (~206 amino acids), but their proteins have distinct electrophoretic mobilities. The theoretical size of the envelope protein VP28 (22.1 kDa) differs 6 kDa from its apparent size in SDS-PAGE (28 kDa), which may be the result of post-translational modifications (e.g. glycosylation, phosphorylation) (van Hulten *et al.*, 2000a). A size difference between the theoretical size and the size determined by SDS-PAGE was also identified for the other major virion proteins VP19 and VP15 in the present study. The major envelope protein VP19 and major nucleocapsid protein VP15 were identified by protein sequencing and further characterized after identification on the WSSV genome.

Enveloped viruses of vertebrates and invertebrates contain glycoproteins in their viral envelopes and these often play important roles in the interaction between virus and host such as attachment to receptors and fusion with cell membranes (van Regenmortel *et al.*, 2000; Granof & Webster, 1999). The glycosylation status of the five major WSSV virion proteins, VP28, VP26, VP24, VP19 and VP15 was determined.

Results

Identification of WSSV virion proteins VP19 and VP15

WSSV was purified from infected *Procambarus clarkii* (crayfish) and healthy control crayfish were used for a mock-virus isolation. To identify and isolate additional major virion proteins the enveloped virions and nucleocapsids were obtained and their proteins separated by SDS-PAGE (Fig. 6.1a). In the nucleocapsid fraction three major bands were observed of 26 kDa (VP26), 24 kDa (VP24) and 15 kDa (VP15), respectively. The 28 kDa (VP28) and 19 kDa (VP19) proteins were present in the NP40 soluble fraction and are therefore located in

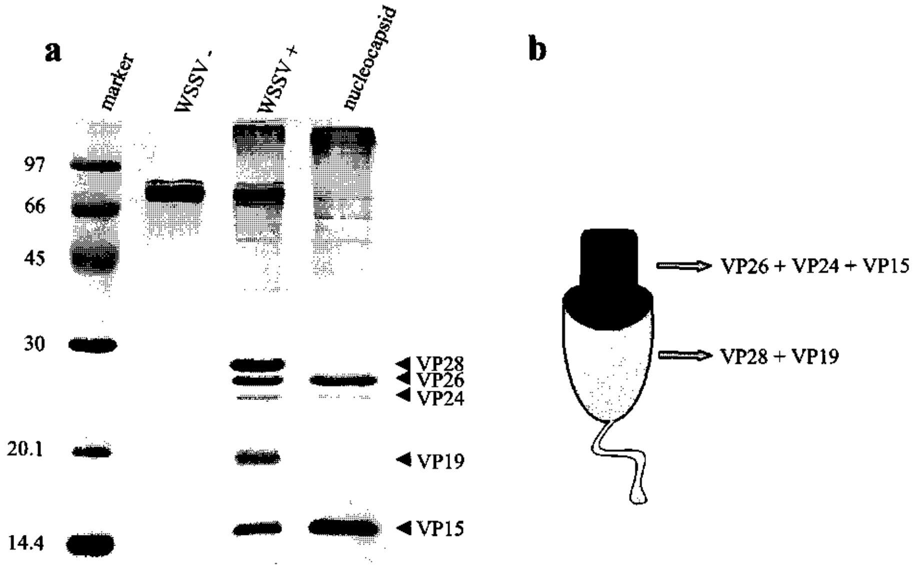


Figure 6.1 (A) 15% Coomassie Brilliant Blue-stained SDS PAGE gel of purified WSSV. Marker: protein marker (in kDa). WSSV-: mock purification from uninfected crayfish. WSSV+: purified WSSV particles. Nucleocapsid: purified WSSV nucleocapsids. (B) Simplified model of the WSSV virion.

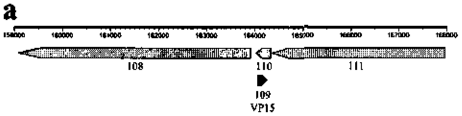
the envelope or tegument of the virion (Fig. 6.1b). As compared to previous publications (van Hulten *et al.*, 2000a, b) in the nucleocapsid fraction an additional major band was observed (VP15), which was not observed previously as the present gel better resolves proteins of about 18 kDa and smaller.

The proteins from the SDS-PAGE gel were transferred to a polyvinylidene difluoride membrane by semi dry-blotting, and the viral protein bands of VP15 and VP19 were excised from the filter and sequenced. Twenty-nine amino acids were obtained from the N-terminus of VP15 (bold-faced in Fig. 6.2b): V A R S S K T K S R R G S K K R S T T A G R I S K R R S P. No N-terminal amino acid sequence of VP19 could be obtained as this protein appeared N-terminally blocked in the analysis. After cyanogen bromide cleavage, an internal VP19 sequence of 24 amino acids was obtained (bold-faced in Fig. 6.3b): I V L I S I (G/V) I L V L A V M N V x M G P K K D S, of which x was undetermined.

Sequence and location of the 15 kDa protein gene on the WSSV genome

The complete sequence of WSSV has recently been determined (van Hulten *et al.*, 2001a) and the N-terminal amino acid sequence determined for VP15 was found to be encoded by ORF109. This ORF is located as a single-copy gene from nucleotide position 163,996 to 164,238 on the WSSV genome. It partially overlaps with a small putative ORF of 11 kDa, ORF110, which is in the reversed orientation on the WSSV genome (Fig. 6.2a). Both ORFs are flanked by two large ORFs (ORF108 and ORF111), encoding putative proteins of 174 and 132 kDa, respectively. The experimentally determined N-terminus of VP15 is not located at the putative start of ORF109, but at the valine at position 21, which is the amino acid following the methionine encoded by the second ATG (Fig. 6.2b). The first ATG of this ORF is apparently not used. This ATG is in an unfavorable Kozak (1989) context (TTC**ATG**A) for efficient translation initiation, whereas the second ATG is in a favorable Kozak context (AAA**ATG**G). Furthermore, the TATA-box, 87 nucleotides upstream of the second ATG, has a position relative to a putative transcriptional start, which also favors this second ATG. As no signal peptide was predicted for the first nor for the second ATG, the most likely explanation for the presence of the valine at the N-terminus of the protein, is the removal of the N-terminal methionine by N-terminal protein processing as is observed for more than 60% of eukaryotic and prokaryotic proteins (Giglione *et al.*, 2000). When the second residue is a valine, the removal of the N-terminal methionine is favored (Flinta *et al.*, 1986). A polyA signal is present 62 nucleotides downstream of the translation stop codon (Fig. 6.2b).

a



b

TTTTCTTTTAATCCTTTGACGCGTAAAGAGTTTCGAGACATGGCTCCTGTATCTACAAAAGAGCTCGCTT

AACCACAGTTGCACTCACCCGGGTACAATAAGACAAAAAT**<u>TATAAA</u>**TTGGGAAGACCGATACAGTCTTTC

ATGACAAAATACCCCGAGAACAAAAGATTGTTGTCTAGGAACAAAGAAACATTAAAA**ATG**GTTGCCCGAA
V A R S

GCTCCAAGACCAAATCCCGCCGTGGAAGCAAGAAGAGGTCCACCACTGCTGGACGCATCTCCAAGCGGAG
S K T K S R R G S K K R S T T A G R I S K R R

GAGCCCATCAATGAAGAAGCCGTGCAGGAAAGAAGAGCTCCACTGTCCGTCGCCGTTCCTCAAAGAGCGGA
S P S M K K R A G K K S S T V R R R S S K S G

AAGAAGTCTGGAGCCCGCAAGTCAAGGCGTTAATTCTTCCCTGTACAACAACTATGTTATTTAATTGATT
K K S G A R K S R R *

TTTTTTCTTCTGAATAATTGGAAAT<u>AATAAA</u>ACATCCATTGAAACTTATGCAGTATTTTTATTCAATTTT

TTAACCAACTATAAATCCACATGTGGTATAAAAGTTTAAGGGTACAGATATATTATAAGATGATGTAACA

GATGCACGGTCAATAACAACAGGGCCTGAAGAAGTGTCTATGAATTTCTTGTAGAATTCTTCCCTGTCTG

Figure 6.2 Location of WSSV VP15 on the WSSV genome (A) and nucleotide and protein sequence of VP15 (B). The N-terminal sequenced amino acids are bold-faced. The location of the putative TATA-box is underlined and bold-faced and the putative polyA signal is underlined.

The ORF109 starting at the preferred (second) ATG encodes 61 amino acids. The theoretical molecular mass of the 60 amino acid protein after removal of the N-terminal methionine, is only 6.7 kDa, but it migrates as a 15 kDa protein in SDS-PAGE. Its aberrant migration may be caused by the fact that this viral protein has an extreme basic pI value of 13.2. The protein has an amino acid composition rich in lysine (22%), arginine (23%) and serine (25%). No potential sites for N-linked glycosylation (N-{P}-[ST]-{P}), or for O-glycosylation (Hansen *et al.*, 1998) could be identified. Eight possible phosphorylation sites ([ST]-X-X-[DE] or [ST]-X-[RK]) were found within VP15. No cysteine residues are present, and therefore VP15 is not able to form any disulfide bridges. No hydrophobic regions are present in this protein (Fig. 6.4a) and it has a strong positive charge along the whole length of the polypeptide. A putative function for this highly basic protein is that of a histone-like, DNA-associated protein. Its small size, positive charge at physiological pH (+26.9 at pH 7.0) and its presence within the nucleocapsid fraction are in agreement with this hypothesis.

Homology searches using BLAST programs (Altschul *et al.*, 1997) showed that VP15 has homology with several DNA binding proteins of eukaryotic origin. Comparison of VP15 with DNA binding proteins of baculoviruses, e.g. with AcMNPV DNA binding protein p6.9 (Wilson & Price, 1988) and *Lymantria dispar* MNPV (LdMNPV) DNA binding protein p6.9 (Kuzio *et al.*, 1999), revealed an overall pairwise amino acid sequence similarity/identity of 47%/26% with AcMNPV and 45%/27% with LdMNPV. This homology with other DNA binding proteins is the result of the high content of basic amino acids (e. g. arginines) and serines, characteristic for these proteins. This further substantiates the putative function of VP15 as a DNA binding protein.

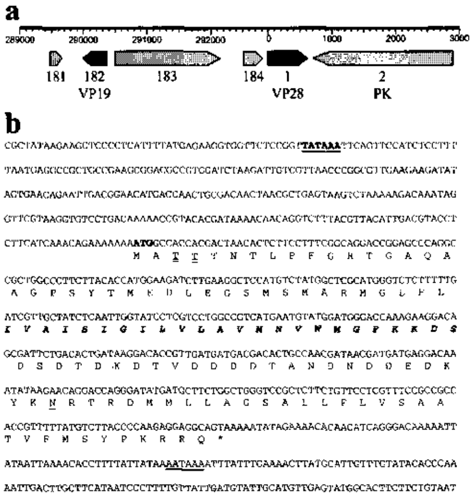


Figure 6.3 Location of WSSV VP19 on the WSSV genome (a) and nucleotide and protein sequence of VP19 (b). The sequenced amino acids are bold-faced; the location of putative N-glycosylation sites is underlined and of O-glycosylation sites is double underlined. The putative TATA-box is underlined and bold-faced and the putative polyA signal underlined.

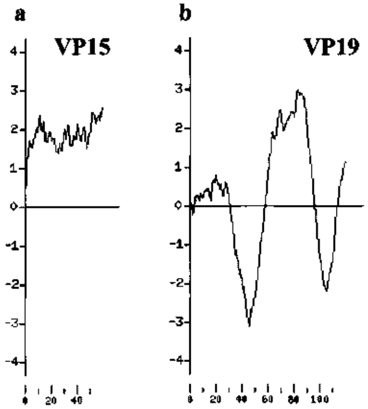


Figure 6.4 Hydrophilicity plots of VP15 (A) and VP19 (B). The amino acid number is on the abcis; the hydrophilicity value on the ordinate.

Sequence and location of the 19 kDa protein gene on the WSSV genome

The amino acid sequence for the internal peptide of VP19 was found to be encoded in WSSV ORF182, located at nucleotide position 290363 to 289998 on the WSSV genome sequence (van Hulten *et al.*, 2001a). The third amino acid (leucine) determined by the amino acid sequencing was found to be incorrect and is predicted to be an alanine. The seventh amino acid (glycine or valine) appeared to be a glycine and the undetermined x was a tryptophan. *Vp19* (ORF182) has a distance of only 2604 nucleotides to *vp28* (WSSV ORF1) on the circular DNA genome. The two major WSSV envelope proteins are only separated by two other WSSV ORFs (ORF183 and ORF184) of which the (putative) functions are unknown (Fig. 6.3a). The sequence surrounding the methionine start codon (AAAATGG) conformed to the Kozak rule for efficient eukaryotic translation initiation (Kozak, 1989). A TATA-box consensus sequence is present 254 nucleotides upstream of the ATG and a consensus polyadenylation (poly A) signal is present 60 nt downstream of the translational stop codon of *vp19*. *Vp19* encodes a protein of 121 amino acids with a theoretical size of 13 kDa. The putative protein is acidic, with an isoelectric point of 4.2. One potential site for N-linked glycosylation (N-{P}-[ST]-{P}) is present, and two putative O-glycosylation sites were predicted using the program NetOglyc (Hansen *et al.,* 1998) and seven possible phosphorylation sites ([ST]-X-X-[DE] or [ST]-X-[RK]) were found (Fig. 6.3b).

Hydrophobicity analysis of the amino acid sequence of VP19 showed that two strong hydrophobic regions were present at two positions in the protein (Fig. 6.4b). In both hydrophobic regions putative transmembrane domains formed by amino acids 37 through 59 and amino acids 95 through 117, respectively, were predicted (Sonnhammer *et al.*, 1998). The second putative transmembrane region has a preferred orientation of outside to inside, as it was followed by a positively charged region starting with a lysine and two arginines (Sonnhammer *et al.*, 1998). The putative transmembrane regions are most likely involved in the anchoring of this protein in the WSSV envelope. No cysteines were present in the protein, indicating that no intra- and inter-protein disulfide cross-links can be formed. For VP19 no amino acid sequence homology was found with sequences in GenBank.

Expression of WSSV structural proteins in insect cells and N-glycosylation

The VP19 and VP15 ORFs encode proteins with theoretical molecular masses of 13.2 and 6.7 kDa, respectively, which differ substantially from their apparent mobility as 19 kDa and 15 kDa proteins in SDS-PAGE. This situation also exists for VP28 and VP26, which have been identified previously (van Hulten *et al.*, 2000a). To confirm that ORF109 and ORF182 are indeed encoding the 15 kDa (VP15) and 19 kDa (VP19) virion protein, these ORFs were expressed in insect cells using a baculovirus vector. The Bac-to-Bac system (GIBCO BRL) was used to generate recombinant baculoviruses (AcMNPV) expressing the putative WSSV virion proteins VP19 and VP15 from the baculovirus polyhedrin promoter in

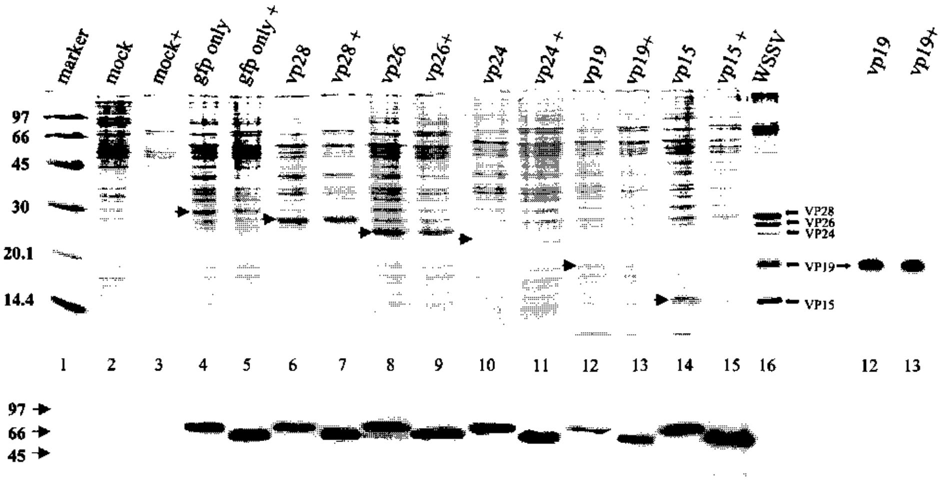


Figure 5 Baculovirus expression of WSSV structural proteins in *Sf*21 insect cells in presence and absence of tunicamycin. (A) Coomassie Brilliant Blue-stained 15% SDS PAGE gel with protein marker (lane 1), with extracts of *Sf*21 cells (lane 2 and 3), infected with AcMNPV-GFP (lane 4, 5), with AcMNPV-WSSVvp28 (lane 6, 7), with AcMNPV-WSSVvp26 (lane 8, 9), AcMNPV-WSSVvp24 (lane 10, 11), AcMNPV-WSSVvp19 (lane 12, 13), and AcMNPV-WSSVvp15 (lane 14, 15), and with purified WSSV virions (lane 16). + indicates the addition of tunicamycin. The location of the WSSV proteins are indicated with arrowheads. (B) Western blot of lanes 12 and 13 of the SDS-PAGE of A. WSSV polyclonal antiserum is used. (C) Western blot of the SDS-PAGE of A. A monoclonal antiserum against GP64 (AcV1) is used and detection is performed with the ECL kit.

insect cells as described previously (van Hulten *et al.*, 2000a). The same procedure was followed for VP24 and a new construct was prepared for VP26 as a cloning artifact resulted in a truncated protein in a previous report (van Hulten *et al.*, 2000a). Recombinant viruses were generated expressing the Green Fluorescent Protein (GFP) from the p10 promoter and each of the WSSV structural proteins from the baculovirus polyhedrin promoter.

*Sf*21 insect cells were infected with AcMNPV-GFP, AcMNPV-WSSVvp28, and AcMNPV-WSSVvp26, AcMNPV-WSSVvp24, AcMNPV-WSSVvp19, and AcMNPV-WSSVvp15 with a MOI of 5 and harvested at 72 h. post infection. Extracts of infected *Sf*21 cells were analyzed in a 15% SDS-PAGE gel (Fig. 6.5a). Clear expression products can be observed for VP28, VP26, VP24 and VP15 (Fig. 6.5a, lanes 6, 8, 10 and 14, respectively), which have the same electrophoretic mobility as their authentic counterparts in the WSSV virion (Fig. 6.5a, lane 16). No clear expression product could be observed for VP19 (Fig. 6.5a, lane 12). Therefore, a western analysis was carried out using a polyclonal antiserum raised against purified WSSV. This analysis showed that VP19 was expressed at the expected position (Fig. 6.5b, lane 12), and hence that the *vp19* ORF encoded a WSSV virion protein.

All major WSSV virion proteins, except VP15, contain multiple sites for N-glycosylation, which could partly or completely explain their large difference in predicted versus apparent molecular weight by SDS-PAGE. Therefore a second infection of Sf21 cells with the recombinant baculoviruses expressing the WSSV proteins was performed in the presence of tunicamycin, which blocks N-glycosylation. The WSSV expression products produced in the presence of tunicamycin were all located at the same position in the gel (Fig. 6.5a, lanes 7, 9, 11, 13, 15) as the expression products in the absence of tunicamycin (Fig. 6.5a, lanes 6, 8, 10, 12, 14). To verify that the tunicamycin treatment had been effective, a polyclonal serum against the AcMNPV envelope glycoprotein GP64 (Stiles & Wood, 1983) was used to identify the location of this baculovirus glycoprotein on a western blot (Fig. 6.5c). A clear downward shift in the position of GP64 of about 8 kDa was observed, confirming the effectiveness of the tunicamycin treatment. This experiment shows that the five major structural WSSV proteins are not N-glycosylated.

Glycosylation of WSSV structural proteins in the virion

Besides N-linked carbohydrate modification, O-linked carbohydrate modifications could exist on the WSSV virion proteins, resulting in larger molecular masses of the proteins as compared to their theoretical masses. Consensus sites for O-linked glycosylation are not well defined, and therefore it is difficult to predict the potential for O-glycosylation. Using the program NetOGlyc (Hansen *et al.*, 1998) putative glycosylation sites for

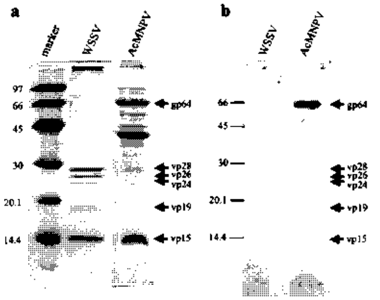


Figure 6 Staining of virion glycoproteins. (A) 15% Coomassie Brilliant Blue-stained SDS PAGE gel. Marker: protein marker. WSSV: purified WSSV virions. AcMNPV: AcMNPV budded virus. (B) PAS glycostain on Western blot of A.

all VPs (VP28, VP26, VP24, VP19) except VP15 were predicted. The O-glycosylation status of the major WSSV structural proteins was investigated using a non-discriminating chemical staining of glycoproteins (PAS) immobilized on a membrane. Figure 6.6a shows a Coomassie brilliant blue-stained gel containing WSSV virions and AcMNPV budded virus. The GP64 glycoprotein present at 64 kDa in AcMNPV budded virus served as a positive control (Stiles & Wood, 1983). After blotting of the proteins PAS staining was performed, resulting in a clear band for GP64. The WSSV structural proteins VP28, VP26, VP24, VP19 and VP15 are not stained and therefore are not glycosylated.

Discussion

The five major WSSV virion proteins have now been identified on the WSSV genome (van Hulten *et al.*, 2000a, b; this publication). These proteins are located in the nucleocapsid (VP26, VP24 and VP15) or in the envelope / tegument (VP28 and VP19) of the virion (Fig 6.1b). When observed in electron microscopy the WSSV nucleocapsid appears to be formed by stacks of rings, which are composed of two rows of regularly spaced subunits (Durand *et al.*, 1997). The assembly of the nucleocapsid proteins VP26, VP24 and VP15, as well as other minor proteins in this structure, has to be elucidated to explain these morphological features. As VP15 is a very basic protein and resembles histone proteins, it is very likely that this protein also functions as a DNA binding protein in the WSSV nucleocapsid. N-terminal amino acid sequencing showed that the second ATG of the *vp15* ORF (ORF109) is used for translation initiation. The N-terminal methionine is most probably removed from the nascent VP15 polypeptide by N-terminal protein processing (Giglione *et al.*, 2000) and this explains the start of VP15 with a valine. A 14.5 kDa protein has been described by Wang *et al.* (2000b) in the WSSV virion and N-terminal sequencing revealed the sequence V A R G G K T K G R R G, which is for 75% identical to the VP15 sequence described here. Assuming that the same protein has been sequenced, the dissimilarity could be explained by the use of a different WSSV isolate. ORF110 is overlapping the *vp15* ORF on the opposite strand of the WSSV genome (Fig. 6.2a) (van Hulten *et al.*, 2001a). No TATA box or polyA signal was found for this ORF and the initiation codon is in an unfavorable Kozak context for translation suggesting that this ORF is probably not functional.

We also identified the ORF encoding the major WSSV envelope protein VP19. VP28 and VP19 are both located in the WSSV envelope (Fig. 6.1b). These proteins contain hydrophobic regions, which may have a function in anchoring these proteins in the envelope. Alternatively, these hydrophobic domains may be involved in the formation of homo- or heterodimers. For VP28 it has been demonstrated using an *in vivo* neutralization assay that this protein has an important function in the systemic infection of WSSV in shrimp (van Hulten *et al.*, 2001b). It remains to be investigated whether VP19 is involved in a similar way in WSSV infection and how it interacts with VP28. Several minor proteins are present in the WSSV envelope and nucleocapsid. These proteins have not been identified on the WSSV genome yet, but may also have important functions in the virion structure, in virion scaffolding or in the WSSV infection.

The theoretical sizes determined for the polypeptides encoded by the *vp19* ORF182 (13.2 kDa) and the *vp15* ORF109 (6.7 kDa) differ significantly from their apparent molecular weight in a SDS-PAGE gel (19 kDa and 15 kDa, respectively). Expression in insect cells using baculovirus vectors resulted in expression products of these ORFs with a similar size as present in the purified virions (Fig. 6.5). Furthermore, VP19 in insect cells reacted well with the WSSV antibody. This confirmed that these ORFs indeed encode the WSSV 19 and 15 kDa virion proteins and showed that if post-translational modifications occur, it is performed in insect cells in a similar way as in *P. monodon.*

In a previous study it was suggested that four prominent WSSV structural proteins identified on a Western blot were not glycosylated (Nadala *et al.*, 1998), but experimental evidence was lacking. It is unclear which of these four bands correspond to any of the five major virion proteins described in this study. As the theoretical sizes of VP28, VP26, VP19 and VP15 differ significantly from the sizes estimated from SDS PAGE, the presence of N- and O-glycosylation was investigated. The results of these experiments showed that none of these proteins is glycosylated. It is possible that the difference in molecular weight is due to an unusual SDS-binding capacity of the proteins or other post-translational modifications. Potential post-translational modifications of the WSSV structural proteins by other events such as phosphorylation, acylation, or isoprenylation are currently being investigated.

In animal viruses it is very common that one or more proteins in viral envelopes or nucleocapsids are glycosylated (van Regenmortel *et al.*, 2000). Cell attachment and fusion are properties conferred upon each virus by one (or in some cases more) of these envelope glycoproteins (Granof & Webster, 1999). For WSSV, however, none of the major virion proteins is glycosylated, which is a feature WSSV has in common with african swine fever virus (ASFV), which also lacks major glycoproteins in the envelope. In ASFV several minor glycoproteins were detected in the virion (del Val *et al.*, 1986; Granof & Webster, 1999). Further research will be performed to identify if there are minor glycoproteins present in the WSSV envelope.

Materials and Methods

White Spot Syndrome Virus virion and nucleocapsid production

The virus isolate used in this study originates from infected *P. monodon* shrimps imported from Thailand in 1996 and was obtained as described before (van Hulten *et al.*, 2000a). Crayfish *Procambarus clarkii* were injected intramuscularly with a lethal dose of WSSV using a 26-gauge needle. After one week the haemolymph was withdrawn and virus was purified as described previously (van Hulten *et al.*, 2000a). The integrity of the purified virus was checked by electron microscopy. The virus envelope was removed from the virus particles by treatment with Nonidet-P40 to obtain nucleocapsids as described by van Hulten *et al.*, 2000a.

Cell culture and baculoviruses

Spodoptera frugiperda (*Sf*-AE-21) cells (Vaughn *et al.*, 1977) were cultured in Grace's insect medium (GIBCO BRL) supplemented with 10% foetal calf serum (FCS). Routine cell culture maintenance, virus infection procedures, and budded virus production were carried out according to published procedures (Smith & Summers, 1987; King & Possee, 1992). To inhibit N-glycosylation, 20 μg/ml tunicamycin (Sigma) was added to the medium at the start of the infection. Recombinant virus *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV)-GFP, was used as a control virus and expressed the Green Fluorescent Protein (GFP) from the p10 promoter (van Hulten *et al.*, 2000a). Recombinant virus AcMNPV-WSSVvp28, was used for the expression of WSSV VP28 from the polyhedrin promoter and GFP from the p10 promoter (van Hulten *et al.*, 2000a).

Engineering of recombinants

The Bac-to-Bac system (GIBCO BRL) was employed to overexpress WSSV VP26, VP24, VP19 and VP15 in insect cells. To facilitate detection and titration of these AcMNPV recombinants upon infection of insect cells, the GFP gene was introduced into the pFastBac-DUAL vector downstream of the p10 promoter into the *Sma*I site of multiple cloning region II of pFastBac-DUAL downstream of the p10 promoter as described previously (van Hulten *et al.*, 2000a). The resulting plasmid was named pFastBac-D/GFP, and maintained region I for insertion of a foreign (WSSV) gene downstream of the polyhedrin promoter.

PCR was performed on the WSSV genome, introducing a *Bam*HI site at the 3' end of the ORFs coding for VP26, VP24, VP19 and VP15 and a *Eco*RI site at the 5' ends of *vp26* and *vp24* and a *Not*I site at the 5' ends of *vp19* and *vp15*. *Vp26, vp24, vp19* and *vp15* were first cloned into the pGEM-Teasy vector (Promega), excised with *Bam*HI and *Eco*RI (*vp26* and *vp24*) or with *Bam*HI and *Not*I (*vp19* and *vp15*), and inserted downstream of the polyhedrin promoter of plasmid pFastBac-D/GFP. The resulting plasmids were named pFastBac-D/G-vp26, pFastBac-D/G-vp24, pFastBac-D/G-vp19 and pFastBac-D/G-vp15, respectively. Recombinant AcMNPV expressing the GFP from the p10 promoter and VP26, VP24, VP19 or VP15 from the polyhedrin promoter were constructed according to the Bac-to-Bac system protocol (GIBCO BRL), and the viruses were designated AcMNPV-WSSVvp26, AcMNPV-WSSVvp24, AcMNPV-WSSVvp19, and AcMNPV-WSSVvp15, respectively.

SDS-PAGE, protein sequencing and immunodetection

Protein samples were analyzed in 15% SDS-PAGE gels as described in Laemmli (1970). Proteins were visualized using Coomassie Brilliant Blue staining and protein weight determined using the Low molecular weight protein marker (Amersham Pharmacia Biotech). Semi-dry blotting was performed onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a CAPS buffer (10 mM CAPS in 10% methanol), or onto an Immobilon™-P (Millipore) using a Tris-Glycine buffer (25 mM Tris base, 192 mM glycine, 10% (v/v) methanol, pH 8.3). Proteins were visualized on the PVDF membrane using Coomassie Brilliant Blue staining. Major protein bands from WSSV virion preparations were excised from the filter and sequenced (ProSeq. Inc., Massachusetts).

Immobilon-P membranes were blocked in 2% low-fat milk powder (Campina, the Netherlands) in TBS (0.2 M NaCl, 50 mM Tris-HCl, pH 7.4). Immunodetection on WSSV proteins was performed by incubation of the blot in a polyclonal rabbit anti-WSSV serum diluted 1:5000 in TBS with 0.2% low-fat milk powder for 1 h at room temperature. Subsequently, anti-rabbit antibody conjugated with horseradish peroxidase (Amersham) was used at a concentration of 1:5000 and detection was performed with an Enhanced Chemiluminescent-light Detection Kit (Amersham). To raise the WSSV polyclonal antiserum, 200 µg purified WSSV virions were injected into a rabbit. The rabbit was boosted with 500 µg WSSV virions after 6 weeks and the antiserum was prepared 2 weeks thereafter.

Immunodetection on AcMNPV GP64 was performed in a similar way using monoclonal antibody AcV1 (Hohmann & Faulkner, 1983) at a concentration of 1:500, and using anti-mouse antibody conjugated with horseradish peroxidase (Amersham) at a concentration of 1:3000.

Staining immobilized glycoproteins

The glycosylation status of the WSSV structural proteins was analyzed by the periodic acid/Schiff method (PAS), a non-discriminating chemical staining of all carbohydrates as described by Rosenberg (1996). After separation by SDS-PAGE, the proteins were transferred to a PVDF membrane. After the blot was washed in water incubation was performed in PAS (1% periodic acid and 3% acetic acid) for 15 min. The blot was washed three times in water and subsequently transferred to Schiff's reagent (Fuchsin-sulfite reagent) for 15 min. After incubation in 0.5% (w/v) sodium bisulfite, the blot was rinsed with water and left to airdry.

Acknowledgements

This research was supported by Intervet International BV, Boxmeer, The Netherlands. We thank Marcel Westenberg for the supply of AcMNPV budded virus.

Chapter 7

White Spot Syndrome Virus envelope protein VP28 is involved in the systemic infection of shrimp

White Spot Syndrome Virus (WSSV) is a large DNA virus infecting shrimp and other crustaceans. The virus particles contain at least five major virion proteins, of which three (VP26, VP24 and VP15) are present in the rod-shaped nucleocapsid and two (VP28 and VP19) reside in the envelope. The mode of entry and systemic infection of WSSV in the black tiger shrimp, *Penaeus monodon* and the role of these proteins in these processes are not known. A specific polyclonal antibody was generated against the major envelope protein VP28 using a baculovirus expression vector system. The VP28 antiserum was able to neutralize WSSV infection of *P. monodon* in a concentration-dependent manner upon intramuscular injection. This result suggests that VP28 is located on the surface of the virus particle and is likely to play a key role in the initial steps of the systemic WSSV infection in shrimp.

This chapter has been published as:
van Hulten, M. C. W., Witteveldt, J., Snippe, M., and Vlak, J. M. (2001). White Spot Syndrome Virus envelope protein VP28 is involved in the systemic infection of shrimp. *Virology* **285**, 228-233.

Introduction

White Spot Syndrome Virus (WSSV) is a major disease agent of penaeid shrimp in Southeast Asia, the Indian continent, and in South and Central America (Rosenberg, 2000). The disease is caused by an ovoid-to-bacilliform virus with a rod-shaped nucleocapsid and a tail-like appendix at one end of the virion (Durand *et al.*, 1997; Nadala *et al.*, 1998). The virus contains a double-stranded DNA with an estimated size of 290 kb (Yang *et al.*, 1997). Genetic analysis indicates that WSSV is a representative of a new virus group provisionally named whispovirus (van Hulten *et al.*, 2000c; Tsai *et al.*, 2000b).

WSSV has a broad host range, infecting several crustacean species, like shrimp, crab, and crayfish (Wang *et al.*, 1998). Little is known about WSSV infection and morphogenesis *in vivo*. Upon infection per os, infected cells are observed first in the stomach, gill and cuticular epidermis of the shrimp. The infection subsequently spreads systemically in the shrimp to other tissues of mesodermal and ectodermal origin (Chang *et al.*, 1996). Research on virus replication and virion morphogenesis shows that DNA replication and *de novo* envelope formation take place in the nucleus (Durand *et al.*, 1997; Wang *et al.*, 2000a). The mechanism of virus entry into the shrimp and of the spread of the virus in the crustacean body is not known.

The virus particle consists of at least five major proteins with estimated sizes of 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24), 19 kDa (VP19) and 15 kDa (VP15). VP28 and VP19 are associated with the virion envelope and VP26, VP24, and VP15 with the nucleocapsid (van Hulten *et al.*, 2000a). Amino acid analysis of VP28, VP26 and VP24 indicated that these proteins have about 40% amino acid identity and that their genes may have evolved from a common ancestral gene (van Hulten *et al.*, 2000b). The role of the envelope and its proteins in the establishment of the systemic infection process has not been determined.

Neutralization experiments have often been performed to study the role of virion proteins or their domains in the infection process. Neutralizing antibodies bind to envelope spikes on the virion and prevent attachment of the virus to the cell surface, cell entry or virus uncoating (Burton *et al.*, 2000). For many vertebrate viruses like poxviruses (Galmiche *et al.*, 1999) and hepadnaviruses (Sunyach *et al.*, 1999) *in vitro* neutralization experiments involving cell cultures (plaque reduction assays) have been used for this purpose. For invertebrate baculoviruses, *in vitro* neutralization experiments have been exploited to show that *Autographa californica* nucleopolyhedrovirus (AcMNPV) can be neutralized by complexing the budded virions with specific antibodies against the viral envelope protein (GP64) (Volkman and Goldsmith, 1985). These authors showed that the mechanism of neutralization is by inhibition of virus entry and adsorptive endocytosis. However, standardized (primary) shrimp cell cultures are not available and therefore an *in vivo* approach is followed.

In vivo neutralization experiments have been widely used for many vertebrate viruses and have even led to passive immunization strategies. When combined with the use of monoclonal antibodies, this strategy has been used to identify the virion protein epitope(s) involved in the neutralization (e.g. Schofield *et al.*, 2000). *In vivo* neutralization assays have also been successfully used in insects, e.g. in inhibiting infection of larvae of the Douglas fir tussock moth *Orgyia pseudotsugata* with its nucleopolyhedrovirus (OpMNPV) using OpMNPV antiserum (Martignoni *et al.*, 1980). This strategy has now been applied to shrimp with the added advantage that it is as close as possible to the *in vivo* situation. In this paper we provide evidence that VP28 is directly involved in the systemic infection of the shrimp *P. monodon* by WSSV.

Results

Virus titration

A WSSV virus stock was produced in the crayfish *Procambarus clarkii* by intramuscular injection of purified WSSV. In order to determine the dilution resulting in 90-100% mortality in the black tiger shrimp *P. monodon*, an *in vivo* virus titration was performed using animals of approximately 1 gram in weight. The virus stock was diluted in steps from 1 x 10^5 to 5 x 10^{11} times in 330 mM NaCl as indicated (Fig. 7.1) and for each dilution 10 μl was injected intramuscularly into 10 shrimps. Shrimps that were injected with 330 mM NaCl, served as negative control for the infection. All shrimps serving as negative control (not shown) and those having received the 5 x 10^{11} virus dilution survived, whereas mortality due to virus infection occurred in all groups with a lower virus dilution (Fig. 7.1). Administration of virus dilutions of 1 x 10^5 and 1 x 10^7 resulted in almost 100% mortality in a period of 20 days. A delay in mortality was observed when virus dilutions of 1 x 10^8 and 5 x 10^9 were used. The 1

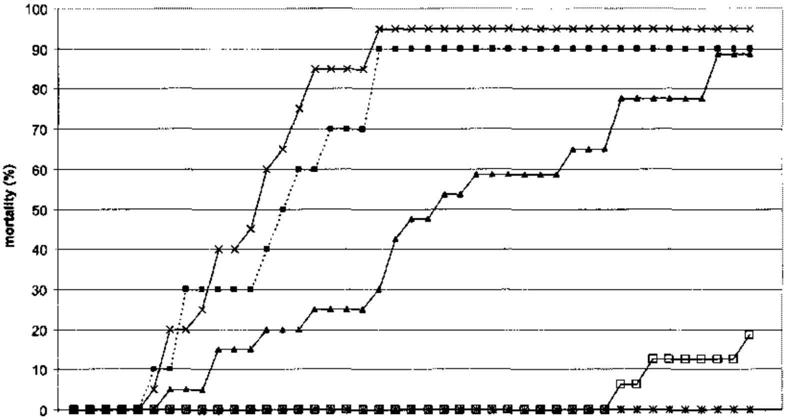


Figure 7.1 Titration of WSSV in *P. monodon* shrimp. Days post injection of the virus are shown on the abcis and the accumulated mortality (in %) on the ordinate. Ten microliters of a 1 x 10^5, 1 x 10^7, 1 x 10^8, 5 x 10^9 and 5 x 10^{11} diluted virus stock was injected.

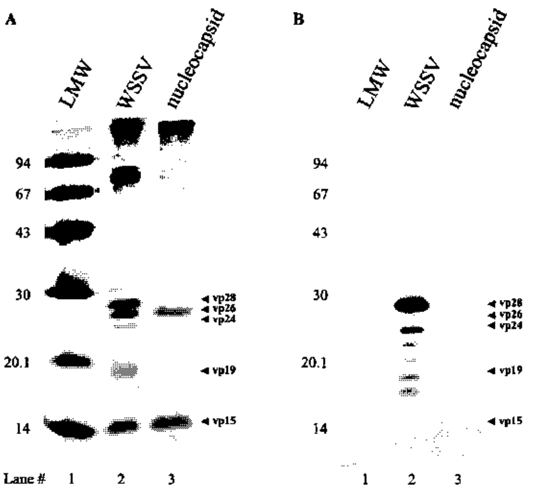


Figure 7.2 (A) Fifteen percent Coomassie Brilliant Blue-stained SDS PAGE gel of purified WSSV. Lane 1: Low molecular weight protein marker. Lane 2: purified WSSV virions. Lane 3: purified WSSV nucleocapsids. (B) Western blot of the virions (lane 2), and nucleocapsid (lane 3) of A. VP28 polyclonal antiserum is used and detection is performed with the ECL kit.

x 10^8 dilution resulted in 90% final mortality, but the time of mortality was delayed and spanned a period of 40 days. The experiment was repeated with the 1×10^7, the 1×10^8, and the 5×10^9 dilution yielding essentially the same results. The dilution of 1×10^8 was chosen as the virus dose for further experiments as this condition was expected to give the optimal response to the neutralization in terms of mortality reduction.

Antibody against recombinant VP28

The major WSSV envelope protein VP28 was expressed under control of the polyhedrin promoter in insect cells using recombinant baculovirus AcMNPV-WSSVvp28 (van Hulten *et al.*, 2000a) and was used after purification to raise specific polyclonal antibodies in rabbits. The VP28 polyclonal antiserum reacted strongly with baculovirus expressed as well as with bacterial expressed VP28 (data not shown).

As VP28 shows a considerable degree of amino acid homology with nucleocapsid proteins VP26 and VP24 (van Hulten *et al.*, 2000b), the specificity of the polyclonal was tested against purified WSSV virions and WSSV nucleocapsids (Fig. 7.2). All major proteins are present in the virion fraction (Fig. 7.2a, lane 2) and only VP26, VP24 and VP15 are present in the nucleocapsid fraction (Fig. 7.2a, lane 3). In the Western analysis (Fig. 7.2b) the VP28 antiserum (1:5000 dilution) shows a clear reaction with the VP28 present in the WSSV virion (Fig. 7.2b, lane 2). A minor reaction was observed with smaller products, most likely VP28 breakdown products. There was no reaction with proteins of the WSSV nucleocapsids (Fig. 7.2b, lane 3). This shows that there is no cross-reactivity of the VP28 polyclonal antiserum with VP26 or VP24, despite the notable degree of amino acid homology.

WSSV neutralization *in vivo*

The VP28 polyclonal antiserum was used in an *in vivo* neutralization assay in *P.*

monodon. A constant amount of WSSV was incubated with various antiserum concentrations (Table 1) and injected into shrimps. No shrimp died in the negative control injected only with 330 mM NaCl (group 6). The shrimp in the positive control, which were injected with WSSV only (group 1), showed a 100% mortality at day 23 (Fig. 7.3). Addition of the pre-immune serum (group 2) resulted in a small initial delay in shrimp mortality, which reached 100% at day 25 (Fig. 7.3). When the virus was pre-incubated with a 10-fold dilution of the VP28 antiserum (group 3), shrimp mortality was 100% at day 22. Apparently the VP28 antiserum at this dilution is not able to neutralize the virus. When the VP28 serum is diluted only 2 times (group 4) or used undiluted (group 5) none of the shrimps died, indicating that WSSV can be neutralized by the VP28 antibodies in a dose dependent manner (Fig. 7.3).

#	Type	Injection (10 μl total)	No. shrimp
1	Positive control	WSSV	10
2	Pre-immune serum control	WSSV + 9 μl pre-immune serum	15
3	10 x dilution of VP28 antiserum	WSSV + 1 μl VP28 antiserum	15
4	2 x dilution of VP28 antiserum	WSSV + 5 μl VP28 antiserum	15
5	VP28 antiserum	WSSV + 9 μl VP28 antiserum	15
6	Negative control	330 mM NaCl	10

Table 7.1 Constitution of injection solutions. The first column shows the group numbers, the second the treatment, the third the content of the 10 μl injected volume and the last column the number of shrimps used in each group. The total amount of WSSV injected is the same for group 1 to 5.

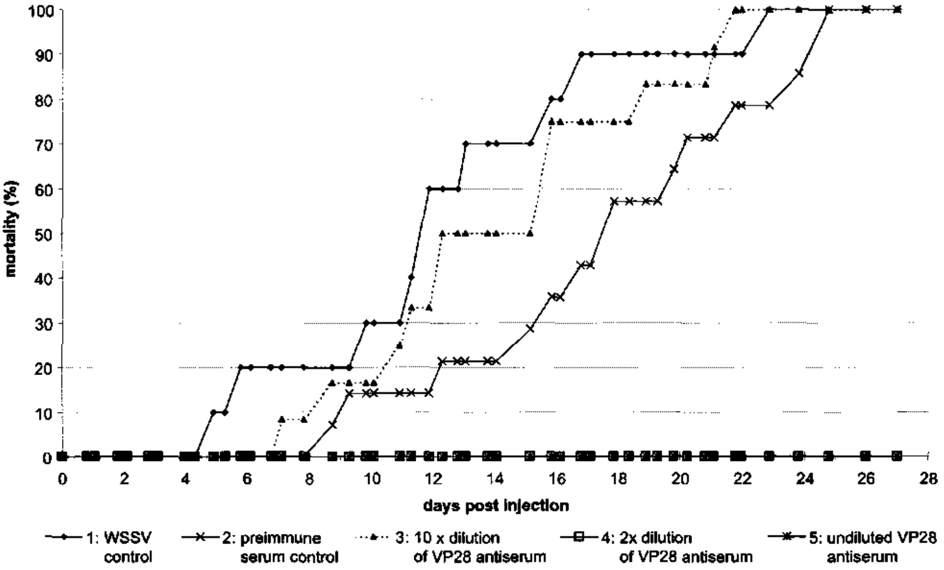


Figure 7.3 Neutralization of WSSV infection in *P. monodon* using VP28 polyclonal antiserum. The days post injection are shown on the abcis and % mortality on the ordinate. The treatments of the five groups used are described in table 1.

Discussion

To study the role of the major WSSV envelope protein VP28 in WSSV infection, a specific antiserum was produced against this protein. As there is a significant degree of homology between the envelope protein VP28 and the nucleocapsid proteins VP26 (41% amino acid similarity) and VP24 (46% amino acid similarity) (van Hulten *et al.*, 2000b), a Western analysis (Fig. 7.2) was performed, confirming the specificity of the VP28 polyclonal antiserum for VP28.

A WSSV virus stock was produced from infected *Procambarus clarkii*. Since an absolute measure of virus activity in infectious units cannot be given, we have determined the highest virus dilution of the virus stock that results in almost 100% mortality upon injection of a fixed volume (10 μl) into shrimp. Shrimp were used for the titration of the virus stock, as no reliable cell culture system is available to measure WSSV infection and the effect of neutralization. Since the 1×10^8 dilution of the virus stock was the lowest dose still resulting in almost 100% final mortality (Fig. 7.1), this dilution was used for neutralization experiments. In the latter experiments (Fig. 7.3) the dilution of 1×10^8 resulted in a somewhat quicker mortality (Fig. 7.3, group 1: WSSV) than in the titration experiment (Fig. 7.1). This difference in response might be the consequence of the use of a different batch of shrimp. However, this does not influence the results of the neutralization experiment, as a control (no antiserum) was included.

To study the role of VP28 in WSSV infection in shrimps, an *in vivo* neutralization test was performed. This test showed that WSSV infection was neutralized by the VP28 polyclonal antiserum (Fig. 7.3) and that VP28 is involved in this process. The pre-immune serum control resulted in a small delay of shrimp mortality. This could be due to compounds present in the serum stimulating the shrimp defense system. WSSV neutralization using the VP28 polyclonal antiserum was concentration dependent; only the two highest antibody concentrations used in this study resulted in neutralization (Fig. 7.3).

VP28 is the major protein in the WSSV envelope, but its location in this structure is not known. The neutralizing activity of the VP28 antiserum shown here might depend on the relative abundance of this protein on the virion envelope. However, mere binding of the antibody to the surface of the virus does not automatically result in virus neutralization. The existence of non-neutralizing antibodies, which bind to virus without diminishing infectivity, has long been recognized (Dimmock, 1984). In other virus systems only anti-envelope antibodies binding to the envelope spike on the virion will be neutralizing or show antiviral activity (Burton *et al.*, 2000). Therefore we postulate that VP28 or its neutralization domain is located in the envelope spike of WSSV virions.

Further research is required to reveal the exact role of VP28 in WSSV infection. Neutralization of viral infectivity by antibodies is a complex and, as yet, poorly understood phenomenon. Studies on the functional domains of proteins suggest that neutralization sites and virus attachment sites are often distinct (Ramsey *et al.*, 1998). Neutralizing antibodies often inhibit a subsequent stage of infection, which is then responsible for the loss of infectivity. There are only a few examples of residues within neutralization epitopes that are also involved in the attachment of the virus to its cellular receptor (Sunyach *et al.*, 1999). For VP28 a similar situation could exist as for the major envelope protein of the budded viruses of AcMNPV, where the virus can be neutralized using antibodies to GP64 (Volkman *et al.*, 1984). Further studies with AcMNPV showed that the mechanism of neutralization is not by inhibition of adsorption, but by inhibition of the fusion of the viral envelope with the cell membrane (Volkman and Goldsmith, 1985). However, alternative mechanisms of neutralization are possible. Inhibition can also take place during uncoating of virus or transport of DNA in the nucleus. Furthermore, binding of antibodies can induce conformational changes in virus proteins and these may be relevant for the neutralization process.

The *in vivo* neutralization experiments on WSSV in *P. monodon* with VP28 antibodies suggest that VP28 is located in the 'spikes' of the WSSV envelope and this protein may thus be involved in the systemic infection of WSSV in shrimps. It cannot be excluded that other WSSV envelope proteins, such as VP19, are also involved in this process, either alone or in concert with VP28. Antibodies against VP19 will assist in the elucidation of this point. Future experiments using *in vivo* neutralization will demonstrate which part of VP28 is involved in the neutralization process and what the role of VP28 in WSSV attachment and entry in the systemic infection is.

Materials and Methods

Shrimp culture

Cultures of healthy shrimp were performed in a recirculation system at the Laboratory of Fish Culture and Fisheries at Wageningen University. For the experiments shrimp were transferred to an experimental system located at the Laboratory of Virology, Wageningen University, and kept in groups of 10-15 individuals in 60 liter aquariums with an individual filter (Eheim, Germany) and heating (Schego, Germany) at 28°C. *P. monodon* shrimps of approximately 1 gram were used in the titration and neutralization experiments.

White Spot Syndrome Virus stock production

The virus isolate used in this study originates from infected *P. monodon* shrimps imported from Thailand in 1996 and was obtained as described before (van Hulten *et al.*, 2000c). Crayfish *Procambarus clarkii* were injected intramuscularly with a lethal dose of WSSV using a 26-gauge needle. After one week the haemolymph was withdrawn from moribund crayfish and mixed with modified Alsever solution (Rodriguez *et al.*, 1995) as an anticoagulant. The virus was purified by centrifugation at 80,000 x g for 1.5 h at 4°C on a 20-45% continuous sucrose gradient in TN (20 mM Tris, 400 mM NaCl, pH 7.4). The visible virus bands were removed and the virus particles were subsequently sedimented by centrifugation at 45,000 x g at 4°C for 1 h. The virus pellet was resuspended in TE (pH 7.5) and the virus integrity was checked by electron microscopy. The virus stock was stored at -80°C until use in the experiments.

VP28 polyclonal antibody

The major WSSV structural envelope protein VP28 was expressed in insect cells using baculovirus *Ac*MNPV-WSSVvp28 (van Hulten *et al.*, 2000a). The protein band containing the VP28 was purified using a Model 491 PrepCell (Biorad) according to the instruction manual. Fractions were collected using a Model 2110 fraction collector (Biorad) and analyzed in a silver-stained SDS PAGE gel. Western blotting using a polyclonal WSSV antibody was employed to determine the VP28-containing fractions. These fractions were pooled and SDS was removed by dialysis against several volumes of 0.1 x TE. The protein was subsequently concentrated by freeze-drying and resuspended in 0.1 x TE (pH 7.5). The purified VP28 protein (100 μg) was injected into a rabbit to produce a polyclonal antibody. The rabbit was boosted with 300 μg of VP28 after six weeks and the antiserum was prepared 2 weeks thereafter.

In vivo injection

P. monodon shrimp of approximately 1 gram were injected intramuscularly with 10 μl of virus solution in 330 mM NaCl in the 4th or 5th tail segment of the shrimp with a 29 gauge needle (Microfine B&D). The shrimps were subsequently cultured for a period of 40 days and the mortality was monitored twice daily. For each group 10 – 15 shrimps were used. Deceased shrimps were monitored for WSSV infection by viewing haemolymph extracts in the electron microscope.

Neutralization assay

Shrimps of 1 gram were injected with WSSV in the presence or absence of VP28 antibody. A negative (330 mM NaCl) and a positive control (virus only) were included (Table 1). The total amount of virus administered per shrimp is constant in all groups and is equivalent to 10 μl of the 1 x 10^8 dilution of the virus stock. The pre-immune serum was included as a control for the effect of the serum on shrimp mortality. Several dilutions of the antiserum were incubated with the virus for 1 h at room temperature, prior to injection in shrimp. After injection the shrimps were monitored for 28 days and dead shrimp were examined for the presence of WSSV by electron microscopy.

Acknowledgement

This research was supported by Intervet International BV, Boxmeer, The Netherlands.

Chapter 8

The white spot syndrome virus DNA genome sequence

White spot syndrome virus (WSSV) is at present a major scourge to worldwide shrimp cultivation. We have determined the entire sequence of the double-stranded, circular DNA genome of WSSV, which contains 292,967 nucleotides encompassing 184 major open reading frames (ORFs). Only 6% of the WSSV ORFs have putative homologues in databases, mainly representing genes encoding enzymes for nucleotide metabolism, DNA replication and protein modification. The remaining ORFs are mostly unassigned except for five, which encode structural virion proteins. Unique features of WSSV are the presence of a very long ORF of 18,234 nucleotides, with unknown function, a collagen-like ORF, and nine regions, dispersed along the genome, each containing a variable number of 250-bp tandem repeats. The collective information on WSSV and the phylogenetic analysis on the viral DNA polymerase suggest that WSSV differs profoundly from all presently known viruses and that it is a representative of a new virus family.

This chapter has been published as:
van Hulten, M. C. W., Witteveldt, J., Peters, S., Kloosterboer, N., Tarchini, R., Fiers, M., Sandbrink, H., Klein Lankhorst, R., and Vlak, J. M. (2001). The white spot syndrome virus DNA genome sequence. *Virology* **286,** 7-22.

Introduction

White spot syndrome virus (WSSV) is a pathogen of major economic importance in cultured penaeid shrimp. The virus is not only present in shrimp but also occurs in other freshwater and marine crustaceans including crabs and crayfish (Lo *et al.*, 1996a). In cultured shrimp WSSV infection can reach a cumulative mortality of up to 100% within 3-10 days (Lightner, 1996) and can cause large economic losses to the shrimp culture industry. The virus was first discovered in Taiwan, from where it quickly spread to other shrimp farming areas in Southeast Asia (Cai *et al.*, 1995). WSSV initially appeared to be limited to Asia until it was found in Texas and South-Carolina in November 1995 (Rosenberry, 1996). In early 1999 WSSV was also reported from Central- and South-America and it has now also been detected in Europe and Australia (Rosenberry, 2000). Intensive shrimp cultivation, inadequate sanitation and worldwide trade has aggravated the disease incidence in crustaceans and enhanced disease dissemination. As such WSS has become an epizootic disease and is not only a major threat to shrimp culture but also to marine ecology (Flegel, 1997).

WSSV virions are ovoid-to-bacilliform in shape with a tail-like appendage at one end. They circulate ubiquitously in the haemolymph of infected shrimp. The virions contain a rod-shaped nucleocapsid, typically measuring 65 - 70 nm in diameter and 300 - 350 nm in length. The nucleocapsids, which contain a DNA-protein core bounded by a distinctive capsid layer giving it a cross-hatched appearance, are wrapped singly into an envelope to shape the virion (Durand *et al.*, 1997; Nadala *et al.*, 1998). The virus contains a large double stranded DNA of about 290 kb as evidenced from restriction enzyme analysis (Yang *et al.*, 1997). Based on the analysis of WSSV-specific sequences it can be concluded that there is genetic variation among WSSV isolates (Wang *et al.*, 2000c; Lo *et al.*, 1999). This was further confirmed by analysis of WSSV structural proteins from different geographical isolates which showed differential profiles (Wang *et al.*, 2000b).

The shape of WSSV virions and nucleocapsids resemble baculoviruses (van Regenmortel *et al.*, 2000), but the size of the viral DNA of about 300 kb is well above the range (100 - 180 kb) of baculovirus genomes (Hayakawa *et al.*, 2000). Random terminal sequencing of WSSV DNA inserts of plasmid libraries indicated surprisingly that less than 5% of the translated sequences had homologues in sequence databases (van Hulten and Vlak, 2001a). A few genes though were identified with homology to other genes in databases, including those encoding for the large and small subunit of ribonucleotide reductase (van Hulten *et al.*, 2000c), a thymidine-thymidylate kinase (Tsai *et al.*, 2000b) and a protein kinase (van Hulten and Vlak, 2001a). Phylogenetic analysis of these genes indicated that WSSV and baculoviruses are not closely related. Three major structural WSSV virion protein genes have been identified and their translated proteins showed no relationship with baculovirus structural proteins (van Hulten *et al.*, 2000b, a). Based on the limited amount of genomic information available it was postulated that WSSV may be a member of an entirely new virus family (van Hulten and Vlak, 2001b).

To further study the taxonomic position of WSSV and to allow a detailed understanding of the pathology of this virus in shrimp we have determined the entire nucleotide sequence of the WSSV genome. Analysis of the 293 kb circular genome revealed 184 open reading frames (ORFs) of 50 amino acids or more, an unusual long ORF (18 kb) and nine regions along the genome with tandem repeat sequences. Many of the predicted proteins have no homology with other viral or cellular genes and hitherto unknown properties. Although the *Paramecium bursaria* Chlorella virus of the Phycodnaviridae with a genome of 330 kb (Li *et al.*, 1997) is the largest virus sequenced, the WSSV genome of 293 kb is at present the largest animal virus genome that has been entirely sequenced.

Results and discussion

Organization of the WSSV genome

The complete DNA sequence of the WSSV genome was assembled into a circular sequence of 292,967 base pairs (bp) in size. This is close to the 290-kb estimated by restriction digestion (Yang *et al.*, 1997), but smaller than the 305-kb reported for a putative WSSV genome of another source (Anonymous, 1999). Although the WSSV sequence was not determined from a clonal WSSV isolate, the sequence heterogeneity was minimal (less than 0.01%). The adenine residue at the translation initiation codon of the major structural virion envelope protein VP28, of which the coding capacity has been confirmed by amino acid sequencing (van Hulten *et al.*, 2000a), was designated as the starting point of the physical map of the WSSV genome (Fig. 8.1, Table 1).

The WSSV genome has an A+T content of 58.9% uniformly distributed over the genome. The frequency of occurrence of the start codon ATG (1.9%) and stop codon TGA (1.9%) was not different from the expected random distribution (1.8% for both codons). However, a paucity of the stop codons TAA (1.8%) and TAG (1.3%), occurring less often than the expected random distribution (2.6% and 1.8%, respectively), was found. The transcription of WSSV genes has not been studied extensively and therefore few WSSV specific promoter motifs have been identified. A transcription initiation sequence (TCAc/tTC) has been identified for the large and small subunit of ribonucleotide reductase by 5'RACE (Tsai *et al.*, 2000a), and this sequence was present almost 50% less frequently in the WSSV genome sequence than expected based on a random distribution.

In total 684 ORFs starting with an ATG initiation codon and 50 amino acids or larger were located on both strands of the WSSV genome. From these ORFs, 184 ORFs of 51 to 6077 amino acids in size with minimal overlap were selected (Fig. 8.1). These 184 predicted ORFs account for 92% of the genetic information in the WSSV genome. Twenty-five of the 184 ORFs have an overlap of 1 – 365 bp (Fig. 8.1). The average distance between the 159 non-overlapping ORFs is 155 bp with a smallest distance of 1 bp and a maximum distance of

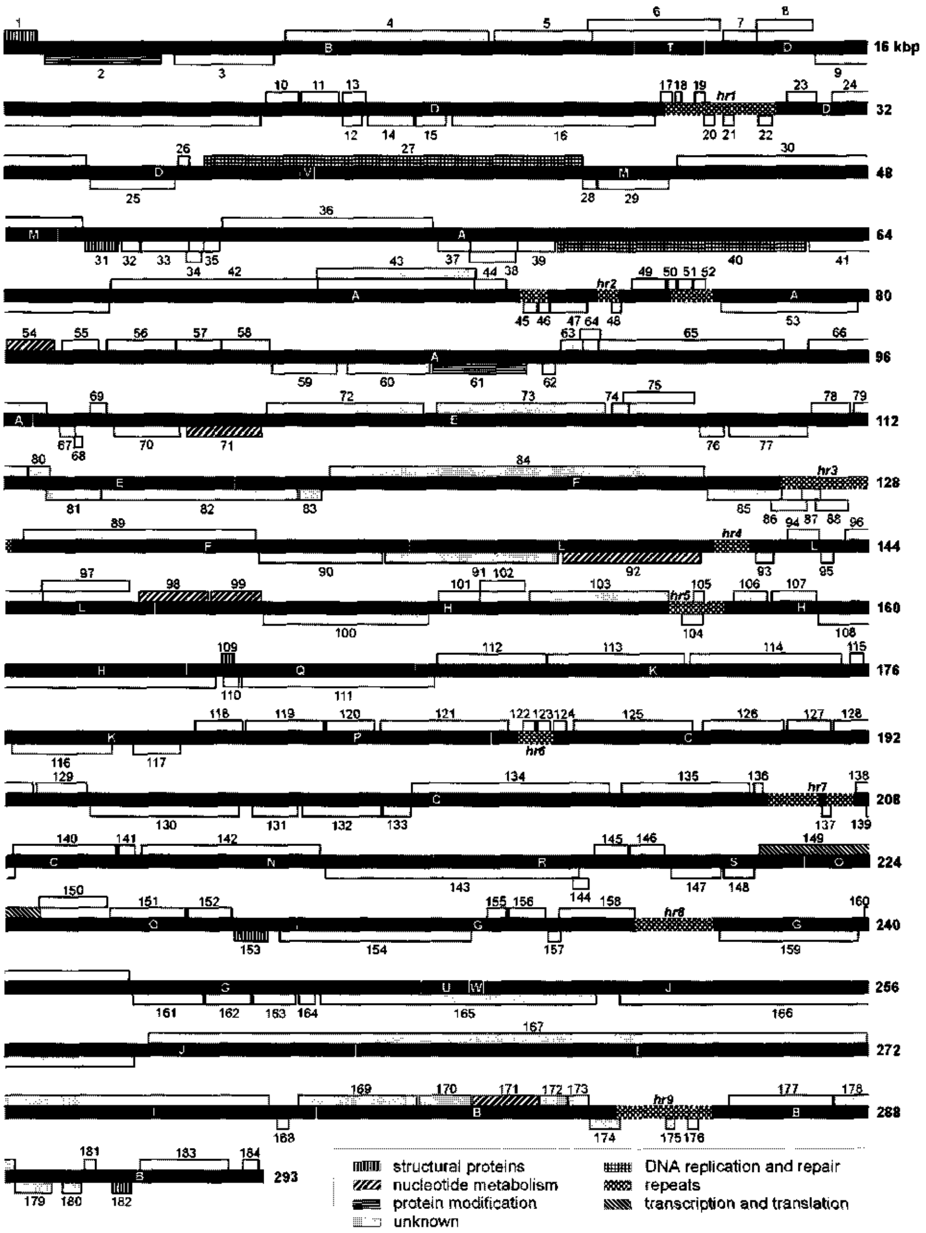


Figure 8.1 Linearized map of the circular double-stranded WSSV genome showing the genomic organization. The A of the ATG initiation codon of VP28 (ORF1) has been arbitrarily designated position 1. Restriction *Bam*HI sites are shown in the black central bar; fragments are indicated A to W according to size from the largest (A) to the smallest (W). ORFs are numbered from left to right. ORFs transcribed forward are located above the genome; ORFs transcribed in the reverse orientation are located below. Genes with similar functions are indicated according to the figure key. *Hrs* are presented according to the figure key and numbered (1-9). Numbers on the right indicate number of nucleotides in kilobase pairs.

Table 8.1

	Position^a		Size^b				
ORF	**Start**	**Stop**	**aa**	**Mr**	**pI^c**	**Characteristics^d**	**Predicted function^e**
1	1 →	615	204	22	4.6	TM; Gene family 1	VP28; envelope protein* (van Hulten *et al.*, 2000a)
2	710 ←	2902	730	82	9.3	Similar to *Homo sapiens* protein kinase (NP_055311); Gene family 2	Protein kinase (van Hulten and Vlak, 2001a)
3	3118 ←	4989	623	70	7.5	EF-hand calcium-binding domain [PS00018]	
4	5185 →	8970	1261	142	9		
5	9056 →	10879	607	67	7.6		
6	10834 →	13236	800	89	6.5		
7	13311 →	13982	223	25	6.4	SP; ATP/GTP-binding site motif A [PS00017]	
8	13979 →	14890	303	35	6.3	TM	
9	14923 ←	20733	1936	216	7	ATP/GTP-binding site motif A [PS00017]; Eukaryotic and viral aspartyl proteases signature and profile [PS00141]; Prenyl group binding site [PS00274]	
10	20837 →	21358	173	19	11.9	SP	
11	21364 →	22161	265	30	5	TM	
12	22201 ←	22596	131	13	11.2	TM; Gene family 3	
13	22232 →	22648	138	15	10.3	SP; 2 TMs	
14	22685 ←	23581	298	34	5.9		
15	23591 ←	24157	188	21	6.1		
16	24265 ←	27996	1243	138	6		
17	28024 →	28296	90	11	9.9		
Hr1	28250	30320					
18	28366 →	28530	54	6	9.7		
19	28760 →	28960	66	8	9.7		
20	28957 ←	29142	61	7	9.1		
21	29283 ←	29468	61	7	9.1		
22	29934 ←	30149	71	9	9.4		
23	30426 →	31052	208	24	6.2	ATP/GTP-binding site motif A [PS00017]	
24	31320 →	33485	721	81	7.2	ATP/GTP-binding site motif A [PS00017]; Cell attachment sequence [PS00016]	
25	33532 ←	35148	538	62	8.3	Gene family 4	
26	35172 →	35402	76	9	4.3	TM	
27	35571 →	42626	2351	262	7.2	DNA polymerase family B signature [PS00116]; Similar to DNA polymerase of *Saccharomyces cerevisiae* (X61920)	DNA polymerase
28	42667 ←	42882	71	8	4.7	SP	
29	42935 ←	44281	448	50	5.2		
30	44350 →	49404	1684	186	9.4	TM; Similar to several collagen types	Collagen
31	49448 ←	50074	208	23	8.7	TM; Gene family 1	VP24 nucleocapsid protein* (van Hulten *et al.*, 2000b)
32	50129 ←	50467	112	13	9.7	Microbodies C-terminal targeting signal [PS00342]	
33	50494 ←	51381	295	33	4.2	TM	
34	51341 ←	51628	95	11	4.6		

	Position[a]			Size[b]				
ORF	**Start**		**Stop**	**aa**	**Mr**	**pI[c]**	**Characteristics[d]**	**Predicted function[e]**
35	51659	←	51952	97	11	4.2		
36	52007	→	55912	1301	144	5.5	3 TMs	
37	55999	←	56601	200	23	9.4	TM	
38	56598	←	57458	286	31	4.8		
39	57509	←	58204	231	26	9.1	TM	
40	58285	←	62892	1535	172	6.2	TM; Similar to *Sno* gene of *Drosophila melanogaster* (U95760)	
41	63021	←	65939	972	108	7	TM; Cell attachment sequence [PS00016]	
42	65956	→	69795	1279	143	5.2	TM	
43	69737	→	72682	981	109	5.7	ATP/GTP-binding site motif A [PS00017]	
44	72663	→	73253	196	23	4.9		
Hr2	73550		77150					
45	73614	←	73859	81	9	9.8		
46	73915	←	74106	63	7	9.8		
47	74151	←	74831	226	26	4.5	Gene family 5	
48	75246	←	75422	58	6	12.1		
49	75584	→	76210	208	25	8.8	Gene family 6	
50	76237	→	76401	54	7	10.8	Microbodies C-terminal targeting signal [PS00342]	
51	76463	→	76714	83	10	9.7	Microbodies C-terminal targeting signal [PS00342]	
52	76776	→	77000	74	9	9		
53	77284	←	79815	843	96	6.4	Gene family 7	
54	80046	→	80915	289	33	7.1	Thymidylate synthase active site [PS00091]; Similar to *Homo sapiens* thymidylate synthase (NP_001062) and other thymidylate synthases	Thymidylate synthase
55	81077	→	81751	224	25	4.8	Gene family 5	
56	81900	→	83168	422	47	4.8	TM; Gene family 8	
57	83170	→	84000	276	32	8.6		
58	84026	→	84919	297	33	4.7	Cell attachment sequence [PS00016]	
59	85001	←	86197	398	45	9.6		
60	86334	←	87869	511	57	4.2	EF-hand calcium-binding domain [PS00018]	
61	87925	←	89667	580	66	6.9	Similar to *Homo sapiens* protein kinase (NP_009202); Gene family 2	Protein kinase
62	89955	←	90197	80	9	8.5	Cell attachment sequence [PS00016]	
63	90298	→	90744	148	17	10.6	SP	
64	90669	→	91046	125	14	8.7		
65	91003	→	94443	1146	126	4.8	TM; Cell attachment sequence [PS00016]	
66	94903	→	96777	624	69	5.1		
67	97012	←	97242	76	9	10.0		
68	97239	←	97394	51	6	4.8		
69	97587	→	97898	103	12	4.4	SP	
70	98032	←	99252	406	44	8.9		
71	99376	←	100761	461	52	5.4	Similar to fowl adenovirus dUTPase (NP_043869), and other viral and eukaryotic dUTPases	dUTPase
72	100959	→	103865	968	108	6.3	4 TMs	
73	104007	→	107141	1044	118	6.4		
74	107265	→	107570	101	12	10		

	Position^{a}		**Size^{b}**				
ORF	**Start**	**Stop**	**aa**	**Mr**	**pI^{c}**	**Characteristics^{d}**	**Predicted function^{e}**
75	107467 →	108789	440	48	10.2	TM	
76	108889 ←	109341	150	17	8		
77	109433 ←	110887	484	53	4.8		
78	110964 →	111779	271	31	4.8		
79	111751 →	112419	222	25	6.5		
80	112426 →	112812	128	15	9	TM	
81	112771 ←	113784	337	38	7.5	TM	
82	113793 ←	117419	1208	138	6		
83	117465 ←	117878	137	16	8	Glycosyl hydrolases family 5 signature [PS00659]	
84	118025 →	124969	2314	289	5.1		
85	125037 ←	126416	459	52	7.7	Glucagon / GIP / secretin / VIP family signature [PS00260]	
86	126211 ←	126876	221	26	9.8		
Hr3	126388	128112					
87	126782 ←	127129	115	13	9.6		
88	127035 ←	127634	199	24	9.6		
89	128334 →	132644	1436	161	5.4	Nt-dnaJ domain signature [PS00636]	
90	132697 ←	134976	759	85	5.6		
91	135031 ←	138249	1072	122	5.5	2 TMs	
92	138330 ←	140876	848	96	7.8	Similar to ribonucleotide reductase large subunits	Ribonucleotide reductase (large subunit) (van Hulten *et al.*, 2000c)
Hr4	141139	141827					
93	141913 ←	142233	106	12	8.4		
94	142498 →	143082	194	22	4.5		
95	143118 ←	143342	74	9	8.5		
96	143569 →	144687	372	43	7.1		
97	144689 →	146314	541	63	9	TM	
98	146492 →	147733	413	48	4.8	Ribonucleotide reductase small subunit signature [PS00368]; Similar to viral and eukaryotic ribonucleotide reductase small subunits	Ribonucleotide reductase small subunit (van Hulten *et al.*, 2000c)
99	147798 →	148733	311	36	8.8	SP; similar to *Penaeus japonicus* deoxyribonuclease I (CAB55635); similar to eukaryotic endonucleases	Endonuclease
100	148770 ←	151829	1019	117	8.1	Eukaryotic RNA Recognition Motif (RRM) RNP-1 region signature [PS00030]	
101	152015 →	152788	257	29	6.7	TM	
102	152788 →	153624	278	31	6.7		
103	153704 →	156274	856	98	8.2	Ribosomal protein L35 signature [PS00936]; Gene family 7	
Hr5	156319	157366					
104	156538 ←	156927	129	14	9.6		
105	156746 →	156955	69	8	11.1		
106	157493 →	158107	204	23	9.6	Gene family 6	
107	158204 →	159031	275	32	7.9	TM	
108	159076 ←	163896	1606	174	6.3	SP	
109	163996 →	164238	80	9	12.6		VP15 nucleocapsid protein* (van Hulten *et al.*, 2001c)
110	164030 ←	164314	94	11	9.3		

	Position[a]		Size[b]				
ORF	Start	Stop	aa	Mr	pI[c]	Characteristics[d]	Predicted function[e]
111	164346 ←	167930	1194	132	5.8	TM; Protein splicing signature [PS00881]; Soybean trypsin inhibitor (Kunitz) protease inhibitors family signature [PS00283]	
112	168000 →	170024	674	76	5.6	Long hematopoietin receptor, gp130 family signature [PS01353]	Class I cytokine receptor
113	170043 →	172577	844	97	6.3	ATP/GTP-binding site motif A [PS00017]; Gene family 9	
114	172701 →	175511	936	108	7.4	TM	
115	175716 →	175964	82	9	4		
116	176120 ←	177967	615	71	7.1	TM; Gene family 4	
117	178367 ←	179251	294	34	5.5	Gene family 4	
118	179527 →	180405	292	33	4.5	Gene family 10	
119	180442 →	181884	480	51	4.6	SP; Gene family 8	
120	181937 →	182839	300	34	5.8	Gene family 3	
121	182911 →	185286	791	90	8.9	TM	
Hr6	185500	186155					
122	185588 →	185818	76	9	10.5	Microbodies C-terminal targeting signal [PS00342]	
123	185843 →	186073	76	9	11.1	Microbodies C-terminal targeting signal [PS00342]	
124	186135 →	186374	79	9	9.8	Microbodies C-terminal targeting signal [PS00342]	
125	186534 →	188747	737	84	8.0	Vitamin K-dependent carboxylation domain (PS00011); Gene family 9	
126	188918 →	190420	500	56	5.2	Prenyl group binding site [PS00274]; Cell attachment sequence [PS00016]; Gene family 4	
127	190500 →	191345	281	32	4.6	Cell attachment sequence [PS00016]; Gene family 10	
128	191349 →	192503	384	43	4.6	SP; Gene family 8	
129	192564 →	193493	309	35	4.6	Gene family 3	
130	193553 ←	196321	922	103	4.4		
131	196571 ←	197416	281	31	6.1		
132	197480 ←	198949	489	56	8.8		
133	198967 ←	199479	170	20	9.1		
134	199492 →	203151	1219	135	7.8	Gram-positive cocci surface proteins 'anchoring' hexapeptide [PS00343]; Cell attachment sequence [PS00016]	
135	203364 →	205739	791	87	6.3		
136	205865 →	206029	54	6	10.9		
Hr7	206140	207726					
137	207118 ←	207279	53	6	10		
138	207790 →	207999	69	7	7		
139	207992 ←	208159	55	6	9.2	TM	
140	208153 →	210057	634	69	5.5		
141	210064 →	210366	100	11	4.8	TM	
142	210519 →	213821	1100	123	5.1	ATP/GTP-binding site motif A [PS00017]; Cell attachment sequence [PS00016]	
143	213918 ←	218612	1564	174	6.6	FGGY family of carbohydrate kinases signature 1 [PS00933]; Aminoacyl-transfer RNA synthetases class-II signature 2 [PS00339]	

	Position[a]			**Size**[b]				
ORF	**Start**		**Stop**	**aa**	**Mr**	**pI**[c]	**Characteristics**[d]	**Predicted function**[e]
144	218566	←	218859	97	11	9.5	TM	
145	218912	→	219532	206	23	5.5		
146	219631	→	220260	209	22	4		
147	220309	←	221238	309	34	8.6	TM	
148	221305	←	221874	189	21	5.1		
149	221977	→	224652	891	100	9.2	TM; similar to *Aspergillus nidulans* TATA-box binding protein (AAB57874)	TATA box binding protein
150	224639	→	225898	419	47	5.5		
151	225923	→	227323	466	52	7.2		
152	227329	→	228147	272	31	7.8		
153	228221	←	228835	204	22	9.3	TM; Gene family 1	VP26; nucleocapsid protein* (van Hulten *et al.*, 2000a)
154	229074	←	232613	1179	132	4.2		
155	232928	→	233281	117	13	9.4	TM	
156	233295	→	233978	227	26	8.8	ABC transporters family signature [PS00211]	
157	233982	←	234230	82	9	9.2		
158	234229	→	235626	465	51	8.5	TM; Gram-positive cocci surface proteins 'anchoring' hexapeptide [PS00343]	
Hr8	235672		237156					
159	237222	←	239792	856	96	9	Cell attachment sequence [PS00016]	
160	239925	→	242285	786	88	6.4	Immunoglobulins and major histocompatibility complex proteins signature [PS00290]	
161	242377	←	243678	433	48	4.6		
162	243701	←	244552	283	32	4.9		
163	244556	←	245341	261	30	6.9	Cell attachment sequence [PS00016]	
164	245444	←	245746	100	12	11.3		
165	245849	←	250966	1705	190	7.6		
166	251400	←	258392	2330	261	5.7		
167	258666	→	276899	6077	664	6.7	Cell attachment sequence [PS00016]; Leucine zipper pattern [PS00029]	
168	277040	←	277246	68	7	8.2	2 TMs	
169	277425	→	279614	729	85	8.3		
170	279667	→	280632	321	36	5		
171	280683	→	281849	388	43	6.3	ATP/GTP-binding site motif A [PS00017]; Thymidine kinase cellular-type signature [PS00603]; Thymidylate kinase signature [PS01331]	Chimeric Thymidine kinase-Thymidylate kinase (Tsai *et al.*, 2000b)
172	281869	→	282384	171	20	4.9	Prenyl group binding site [PS00274]	
173	282433	→	282816	127	14	9.1	SP	
174	282829	←	283380	183	22	9.1		
Hr9	283323		285125					
175	284246	←	284401	51	6	8.6		
176	284646	←	284843	65	7	9.4		
177	285406	→	287331	641	74	6.7	Gene family 9	
178	287386	→	288165	259	30	6.6		
179	288183	←	288866	227	26	6.2		
180	289149	←	289343	64	7	8.5	Leucine zipper pattern [PS00029]	

	Position[a]		Size[b]				
ORF	Start	Stop	aa	Mr	pI[c]	Characteristics[d]	Predicted function[e]
181	289474 →	289680	68	8	11.7		
182	289998 ←	290363	121	13	4.2	2 TMs	VP19, envelope protein* (van Hulten *et al.*, 2001c)
183	290501 →	292135	544	62	7.1	MIP family signature [PS00221]	
184	292511 →	292804	97	11	8.4	TM	

Table 8.1 WSSV ORFs

[a]Position and orientation of the ORFs in the WSSV genome

[b]Size of ORFs in amino acids (aa) and predicted molecular mass in kDa (Mr)

[c]Predicted isoelectric point (pI)

[d]The presence of transmembrane domains (TM) and signal peptides (SP) are indicated. Presence of motifs in the PROSITE databank is indicated and PROSITE accession numbers are shown in between square brackets. Similarity with proteins in GenBank, including accession number between brackets, is indicated.

[e]Predicted function; empirically demonstrated functions are indicated with an *.

1595 bp. ORFs are present on both strands in almost equal proportions (54% forward, 46% reverse), and ORFs frequently (60%) occur in head-to-tail tandem arrays (Fig. 8.1). The largest cluster of consecutive genes with the same transcriptional orientation contains 12 ORFs (118-129). Based on homologies with other viral or cellular genes in GenBank only 11 of the 184 WSSV ORFs have been assigned a putative function or have similarity with known genes (Table 1). In contrast, baculoviruses share about 50% of their genes (Hayakawa *et al.*, 2000) and this clearly separates WSSV from this group of viruses. Computer analysis using the minor ORFs overlapping the 184 WSSV ORFs showed no relevant homologies to data in GenBank.

Homologous regions

The WSSV genome was analyzed for the presence of repeats using the repeat finder program REPuter (Kurtz and Schleiermacher, 1999). The complete genome was compared to itself to identify perfect direct repeats of minimally 15 bp and subsequently a circular representation of the genome was generated where repeat regions of 30 bp or longer were connected by a line (Fig. 8.2a). Nine direct repeat regions with different sizes were found dispersed throughout the genome (Fig. 8.2a; Table 1). Analysis of these regions revealed that they all consisted of identical repeat units of 250 bp or parts thereof. In accordance with homologous regions in baculoviruses (Cochran and Faulkner, 1983) the nine repeat regions were designated homologous region (*hr)* 1 to *hr*9. One of these repeats (*hr*4), has previously been described by van Hulten *et al.* (2000c).

The repeat units of the identified *hrs* were found in both orientations on the WSSV genome (Fig. 8.2b). Four of the *hrs* consisted of repeat units all in a forward orientation (*hr*1, *hr*3, *hr*5, and *hr*9), three consisted of repeat units all in the reverse orientation (*hr*4, *hr*6 and *hr*8), and two *hrs* contained repeat units in both orientations (*hr*2 and *hr*7) (Fig. 8.2b). This is also shown in figure 8.2a, where repeat units in the same orientation are connected by lines.

The *hrs* all contain 3 to 8 repeat units of about 250 bp (Fig. 8.2b), with a total of 53 repeat units for the WSSV genome. The 53 repeat units were aligned and part of this alignment is depicted in figure 8.2c with one representative repeat unit from each *hr*. A highly conserved domain of 115 bp is present in the center of all the repeat units and is flanked by two more variable domains ("variable domain I", and "variable domain II") of approximately 70 bp each. Based on homology in "variable region I" two types of repeats were distinguished. *Hrs* 1, 2 and 9 belong to type A and *hrs* 3, 4, 5, 6, 7, and 8 belong to type B. The highly conserved central domain contains an imperfect palindrome of 21 bp, which mainly consists of A+T (Fig. 8.2c).

The *hrs* are largely located in intergenic regions (Fig. 8.1), although several short ORFs are present. The WSSV repeat regions resemble baculovirus *hrs*, which also occur dispersed in all baculovirus genomes sequenced to date (Hayakawa *et al.*, 2000). However, the WSSV *hr* repeat units (250 bp) are much larger as compared to the repeat unit (about 70 bp) in the nucleopolyhedroviruses (NPVs) *hrs*. Furthermore, a key structural feature of the NPV *hrs* is a conserved 30-bp imperfect palindrome located in the center of the 70-bp repeat unit (Hayakawa *et al.*, 2000), whereas the WSSV *hrs* only have a 21-bp imperfect palindrome which is not in the center of the repeat units. The WSSV *hrs* have some resemblance to the *hrs* of granulovirus *Plutella xylostella* (PxGV), as the *hrs* of this virus are larger (105 bp) and only contain a small 15-bp palindromic region in the center of the repeat (Hashimoto *et al.*, 2000).

As WSSV is not closely related to baculoviruses based on genome content and is clearly phylogenetically separated from the baculoviruses on the basis of gene phylogeny (van Hulten *et al.*, 2000c; Tsai *et al.*, 2000b; van Hulten and Vlak, 2001a), the presence of *hrs* could be a general feature of large circular viral DNA genomes. A possible essential function of these *hrs* might be their involvement in the replication of viral DNA (Kool *et al.*, 1995), or in enhancement of transcription (Guarino and Summers, 1986) as was shown for baculovirus *hrs*.

Comparison to other WSSV isolates

WSSV sequence data, available in GenBank, were compared to the complete genome sequence presented here and most sequences showed a high degree of homology. Ninety-eight to 100% homology was found with sequences from WSSV isolated from *P. chinesis* (accession numbers: U92007 (2424 bp) and U89843 (420bp)), with sequence data from WSSV isolated from *P. monodon* from Vietnam (accession number: AJ297947 (941 bp)), and with sequence data from a Taiwan isolate of WSSV (Lo *et al.*, 1999) (accession numbers: AF272669 (1400 bp), AF272979 (1250 bp), and AF272980 (1450bp)). Wang *et al.* (2000c) analyzed three fragments of WSSV DNA of which two are present in GenBank, (C42, accession number AF29524 and A6, accession number AF295123). The third fragment (LN4) partly overlaps with a sequence present in GenBank (accession number AF178573). Compared to our WSSV complete genome sequence, a 100% nt homology was found with the A6 fragment (1416 bp), but the C42 fragment (510 bp) and the LN4 overlapping fragment (2833 bp) were, surprisingly, not present in the complete WSSV genome sequence. A possible explanation is the naturally occurring genetic heterogeneity between WSSV isolates of different origin. The observation of restriction fragment length polymorphisms in different

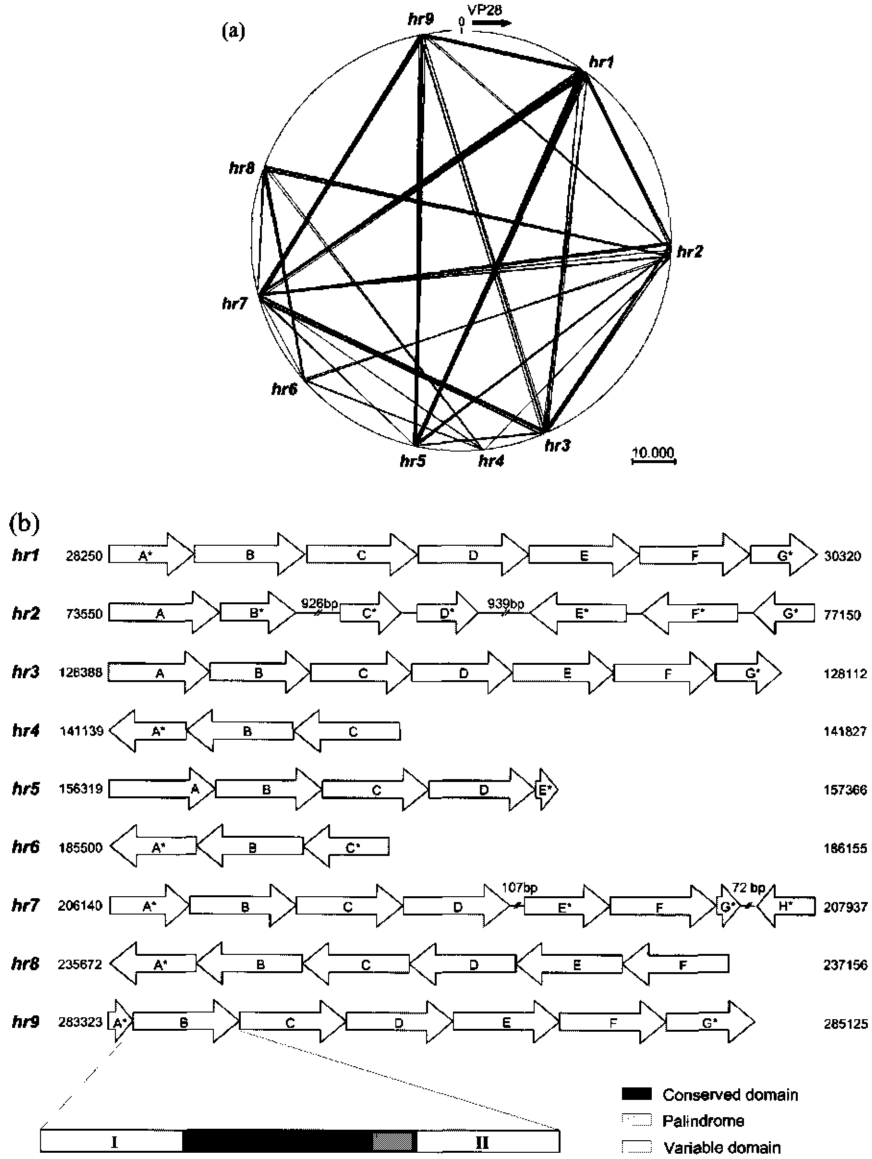


Figure 8.2 (a) Circular display of the WSSV genome showing direct repeat regions predicted by REPuter (Kurtz and Schleiermacher, 1999). Lines connect regions with minimal 30 bp homology. The location and orientation of VP28 is shown as an arrow above the circular genome presentation. A bar at the bottom indicates the scale. The homologous regions (*hrs*) identified on the genome are numbered 1-9. (b) Schematic representation of the repeat structure of the WSSV *hrs*. The repeat units are depicted as arrows, indicating their respective orientation on the genome. Partial repeats are shown by a shorter arrow and an asterisk (*) following its letter. At the bottom of the figure a schematic representation of the repeat domains is shown as a linear bar with the conserved domain including the perfect palindrome and variable domains indicated as shown in the legend and detailed in (c). (c) Nucleotide sequence alignment of one representative repeat unit from each of the nine *hrs* of the WSSV genome. Shading is used to indicate the occurrence (black: 90%, dark grey: 70%, light grey: 30%) of identical nucleotides. The conserved domain (grey bar), palindrome (black bar) and variable domains (white bar) are indicated underneath the alignment.

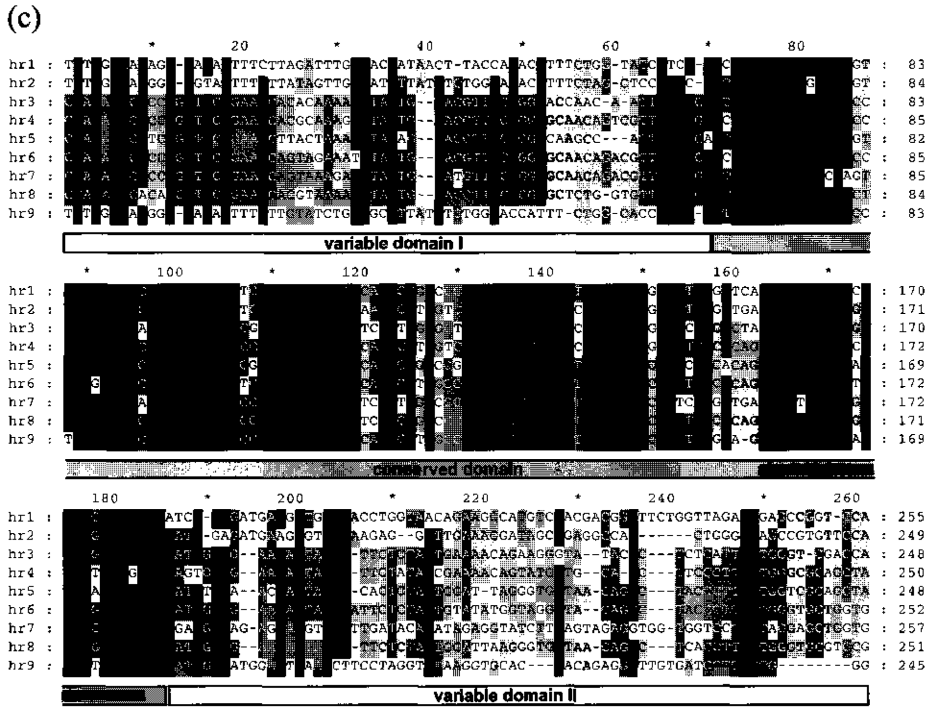


Figure 2 - continued

WSSV isolates supports this view (Lo *et al.*, 1999; Wang *et al.*, 2000c).

To exclude the possibility that the absence of these sequences is the consequence of a sequencing artifact, we have tested the primers used by Wang *et al.* (2000c) (A6, C42, and LN4 primer sets) to amplify the three fragments (1128 bp, 425 bp, and 750 bp, respectively). Furthermore, we tested a different set of primers (control) used for WSSV detection by PCR (Lo *et al.*, 1996b). The results of this PCR showed that the WSSV isolate used in this study does not contain fragments LN4 and C42, whereas the A6 and the control PCR fragments were present (data not shown). Assuming that C42 and LN4 are WSSV specific this result suggests the existence of WSSV variants or isolates with different genetic complexity.

Gene expression

The complete WSSV genome sequence was searched for transcriptional and translational motifs. Seventy-two percent of the ORFs selected have an ATG in a favorable Kozak context (Kozak, 1989). From the 27 ORFs located in the *hrs*, only 4 have an ATG in a favorable Kozak context. Furthermore, these ORFs have a small size (average of 89 aa). A TATA box sequence was found in the promoter regions of 46% of the WSSV ORFs. Early transcribed genes, like the ribonucleotide reductase large and small subunit homologues (Tsai *et al.*, 2000a), contain a TATA box, which is also the case for other potential early transcribed genes like the thymidine-thymidylate (ORF171, Table 1) and dUTPase (ORF73, Table 1) homologue. From the structural proteins, which have been identified by N-terminal sequencing (VP28 and VP26, van Hulten *et al.*, 2000a; VP24, van Hulten *et al.*, 2000b; VP19 and VP15, van Hulten *et al.*, 2001c), only VP15 and VP19 contain a TATA box sequence, indicating that this sequence is not essential in WSSV for efficient transcription of these putative late genes. No putative late promoter elements have been identified in WSSV yet. The late promoter element "RTAAG", canonical in baculoviruses, was not found in putative WSSV ORF promoter regions and occurs at an average frequency in the WSSV genome sequence. Consensus polyA signal sequences are found located in or after the termination codon for 54% of the ORFs, indicating that the WSSV transcripts of these WSSV genes are most probably poly-adenylated.

Sequence similarities to proteins in the databases

Homology searches were performed with the major ORFs of the WSSV sequence (Fig. 8.1, Table 8.1). The deduced translation products of the 184 ORFs were compared to amino acid sequences in GenBank. ORFs which produced a significant BLASTp score and ORFs with lower, but interesting similarities and PROSITE motifs are listed in Table 1. For only 6% of the ORFs a putative function could be assigned based on homology with GenBank sequences. Three percent of the ORFs were identified as major structural proteins in the WSSV virion confirmed by N-terminal amino acid sequencing (van Hulten *et al.*, 2000b, a; van Hulten *et al.*, 2001c).

DNA replication

Genes involved in DNA replication and repair (such as DNA polymerase, DNA helicase and DNA binding proteins) are often found in the genome of large DNA viruses. However, for WSSV we could only identify one of these genes. The presence of a DNA polymerase family B signature (PROSITE entry: PS00116) in ORF27 led us to the identification of a putative DNA polymerase gene. In the BLAST homology search only low homology (maximum BLASTp score 52) was found with several eukaryotic DNA polymerase genes. Alignment of the putative WSSV polymerase gene and several viral and eukaryotic DNA polymerases showed that all seven conserved DNA polymerase sequence motifs (Bernad *et al.*, 1987) were present on the polypeptide encoded by WSSV ORF27. Furthermore, the three conserved regions implicated in DNA polymerase 3'-5' exonuclease activity (Bernad *et al.*, 1989) were also conserved. Despite the presence of these conserved domains the overall homology of WSSV to several viral and eukaryotic DNA polymerases was only maximally 22%. A notable difference of the WSSV polymerase gene in comparison with other DNA polymerases is its size (2351 amino acids), which is about twice the size of an average DNA polymerase. The additional amino acids of the WSSV polymerase gene are located at both the N- and C-terminus as well as in between the conserved DNA polymerase motifs. Except for the DNA polymerase no other genes involved in DNA replication could be identified based on homologies with such genes in GenBank or based on the presence of conserved domain, present in the PROSITE databank.

Nucleotide metabolism

Most large DNA viruses encode a set of genes involved in nucleotide metabolism, enabling their efficient replication in non-dividing cells (Reichard, 1988). WSSV encodes three key enzymes for the synthesis of deoxynucleotide precursors for DNA replication: ribonucleotide reductase, thymidine kinase and thymidylate kinase. The large and the small subunit of ribonucleotide reductase (RR1 and RR2, respectively) were already identified previously (van Hulten *et al.*, 2000c). ORF92 (RR1) and ORF98 (RR2) are located in proximity on the WSSV genome, separated only by 5615 bp, including *hr*4. A chimeric protein consisting of a thymidine kinase and thymidylate kinase (TK-TMK) (Tsai *et al.*, 2000b) is encoded by ORF171. This chimeric protein is a unique feature of WSSV, as these genes are normally encoded by separate ORFs in other large DNA viruses.

WSSV ORF71 contains a putative homologue of dUTP pyrophosphatase (dUTPase). This enzyme is encoded by many large DNA viruses and is responsible for regulating cellular levels of dUTP (Baldo and McClure, 1999). High homology was found with other viral and eukaryotic dUTPase genes. The highest BLASTp score (84) was found with fowl adenovirus dUTPase (accession: NP_043869), which showed a 46% amino acid similarity over a stretch of 200 amino acids.

The WSSV genome also contains a highly conserved gene (ORF54) for thymidylate synthase (TSY). Such a gene is, until now, only observed in *Melanoplus sanguinipes* entomopoxvirus (MsEPV), several herpesviruses and bacteriophages (Afonso *et al.*, 1999). Homodimeric TSY catalyzes the methylation of dUMP to the nucleotide precursor dTMP, thus representing an important part of the *de novo* pathway of pyrimidine biosynthesis (Carreras *et al.*, 1995). The polypeptide encoded by ORF54 contains the PROSITE motif for the TSY active site (PS00091) and is very similar to TSY from eukaryotes and large DNA viruses. The highest BLASTp score (392) was found with the human TSY, which had a 74% overall amino acid similarity (61% identity) to the WSSV TSY.

ORF99 encodes a putative non-specific endonuclease and can be translated as a 311 amino acid polypeptide, which includes a putative hydrophobic signal peptide. DNases are encoded by several other large DNA viruses (Afonso *et al.*, 2000; Li *et al.*, 1997). The function of an endonuclease in the viral replication cycle is unknown, but it may serve in DNA catabolism during apoptosis (Krieser and Eastman, 1998). The WSSV endonuclease homologue has the highest homology (34% overall amino acid similarity and 17% identity) with the DNase I gene of *Penaeus japonicus* (GenBank accession: CAB55635), suggesting that this WSSV gene may have been obtained from a crustacean host.

Transcription and mRNA biosynthesis

The N-terminal part of ORF149 encodes a polypeptide with a high similarity to a transcription initiation factor (TATA-box binding protein (TBP)) of eukaryotes (Berk, 2000). The highest BLASTp score (41) was found with the *Aspergillus nidulans* TBP (accession: AAB57874), where a 150 amino acid part of the N-terminal region of the WSSV ORF149 has a 40% similarity (22% identity) with this protein. Despite this homology, the internal repeat that is conserved in TBPs (Kim *et al.*, 1993) is not present in WSSV ORF149. Therefore it is not clear if ORF149 could have a similar function as eukaryotic TBPs, which play a major role in the activation of eukaryotic genes transcribed by RNA polymerase II and binds to the TATA box promoter element (Berk, 2000).

Many viruses encode RNA polymerase subunits, which are involved in mRNA transcription, initiation, elongation and termination. However, no homologues of these enzymes have yet been identified in the WSSV genome. Furthermore, no RNA helicase, poly(A)polymerase or other genes involved in transcription and mRNA biogenesis were found and may therefore be absent or too diverged from known homologues to be found based on amino acid homology.

Protein modification

A gene (ORF2) coding for a serine/threonine protein kinase (PK) has recently been identified (van Hulten and Vlak, 2001a). Such enzymes are responsible for the phosphorylation of proteins. Phylogenetic analysis using this gene underscored the unique taxonomic position of WSSV relative to baculoviruses and other large DNA viruses (van Hulten and Vlak, 2001a). All twelve conserved domains of a PK were present in the polypeptide encoding this ORF. A second PK gene homologue has been identified as ORF61. For both proteins the highest homology in the BLASTp search was found with a *Homo sapiens* PK gene. ORF2 had 30% amino acid similarity (*H. sapiens* PK: NP_055311) and ORF61 29% (*H. sapiens* PK: NP_009202). These two WSSV PK genes have a pairwise amino acid sequence similarity of 45% (27% identity). When included in the unrooted parsonimous phylogenetic tree of PK described by van Hulten and Vlak (2001a), the two genes have a most recent common ancestor and could therefore be the result of gene duplication.

Immune evasion functions

ORF43 has 43% similarity (22% identity) in a 220 amino acids-long overlap with a *sno* gene of *Drosophila melanogaster*. The *sno* product is part of a complex, which negatively regulates transforming growth factor-β (TGF-β) signaling. This process is important in mediating inflammatory and cytotoxic reactions (Shinagawa *et al.*, 2000). As not much is known about the shrimp immune system the presence of a putative *sno* gene in the WSSV genome cannot be fully explained, but might be involved in abrogating the host defense response.

The polypeptide encoded by ORF112 contains a 'long hematopoietin receptor, gp130 family signature' (PROSITE: PS01353). Genes containing this motif all belong to the class-I cytokine family of receptors in higher eukaryotes (Hibi *et al.,* 1996). ORF112 contains sequences similar to a signal, an immunoglobulin-like C2-type domain and a number of fibronectin type III-like modules. Compared to other cytokine receptor genes ORF112 is somewhat shorter and lacks a transmembrane region. Absence of the transmembrane region may suggest that the protein is produced in a soluble form. Such forms normally arise via both proteolytic processing and alternative splicing (Taga and Kishimoto, 1997). Because of the homology with the class-I cytokine genes and the presence of the motifs typical for cytokine receptors, it is possible that the ORF112 product is involved in signal transductions related to the defense response system in shrimp.

Structural WSSV virion proteins

Five structural WSSV virion proteins have been identified so far by amino acid sequencing of the individual proteins and reverse genetics. The major envelope protein, VP28 (ORF1) and two major nucleocapsid proteins, VP26 (ORF153) and VP24 (ORF31), have been described before (van Hulten *et al.*, 2000a, b). Nucleotide and amino acid comparison revealed that these three proteins are homologues and that they may be the result of gene duplication and divergence into proteins with different functions in the nucleocapsid (VP24, VP26) and the envelope (VP28) (van Hulten *et al.*, 2000b). All three ORFs have an initiation codon in a favorable Kozak (1989) context. Their promoter regions contain stretches of A/T rich sequences but no consensus TATA box sequence. As a polyadenylation consensus (poly A) signal is present for all ORFs, these transcripts are most probably poly-adenylated. Nucleotide sequencing of WSSV confirmed that these three virion structural protein genes are present in single copies.

Internal amino acid sequencing was performed on the envelope protein of 19 kDa (VP19) and N-terminal sequencing on the major nucleocapsid protein of 15 kDa (VP15) (van Hulten *et al.*, 2001c). The ORF encoding VP19 is ORF182. The initiation codon of this major envelope protein ORF is in a favorable Kozak context (AAA**ATG**G), a TATA box was identified 254 nucleotides upstream of the ATG and a polyA signal sequence is located 59 nucleotides downstream of the termination codon. Two putative transmembrane domains were identified in the amino acid sequence, which could anchor VP19 in the virion envelope.

The amino acid sequence obtained for VP15 showed that this nucleocapsid protein is encoded by ORF109. The first initiation codon in this ORF is not in a favorable Kozak context (position 163996, TTC**ATG**A), whereas the second ATG (position 164052), 57 nt downstream of this ATG, is in a favorable Kozak context AAA**ATG**A for efficient translation. The N-terminal amino acid sequence data suggest that the second ATG is used for translation of this ORF. The TATA box is present 87 nucleotides upstream of the second ATG, and has a preferred location for this ATG. A polyA signal is present 62 nucleotides downstream of the translation stop codon. The very basic nature of the ORF109 product (pI = 12.6) and its association with the nucleocapsid of the virion may suggest that it is a basic DNA binding protein.

Putative membrane-associated protein and secreted proteins

The ORFs were analyzed for the presence of putative transmembrane domains (TMs) and signal peptide (SP) sequences. One or more putative TMs were found in 45 ORFs and putative SPs were located in 14 ORFs (Table 1). The proteins containing a putative TM may be associated with membrane structures. Of the structural proteins, the envelope proteins both contain one (VP28) or two (VP19) TMs, which is expected as these proteins are present in the WSSV virion envelope. Also for VP26 and VP24 a TM was identified, although these proteins are not located in a membrane, but in the nucleocapsid of the WSSV virion. The presence of these hydrophobic domains may well be involved in protein-protein interactions which are necessary for the formation of the nucleocapsid which consists of globular subunits (Durand *et al.*, 1997; Nadala *et al.*, 1998).

Other genes with interesting properties

ORF3 and ORF60 contain an EF-hand calcium-binding domain (PROSITE accession: PS00018), suggesting that their products may belong to the class of the calcium-binding proteins. A further function cannot be assigned to these proteins as no homologues were found in GenBank.

ORF30 encodes a large putative protein (168 kDa) from the collagen family, as the collagen family GXY repeat motif G-x(2)-G-x(2)-G-x(2)-G-x(2)-G-x(2)-G (van der Rest and Garrone, 1991) is present from amino acid position 161 to 1326 in the 1684 amino acid long polypeptide. At the N-terminal side a predicted transmembrane region is found at position 54 – 70. No N-terminal signal peptide was found. The function of this collagen homologue in the WSSV genome is not clear, but it is interesting to note that in viruses only lymphocystis virus (iridovirus) has a homologue of this protein (Tidona and Darai, 1997).

An extremely large ORF (ORF167) of 18234 bp coding for a polypeptide of 6077 aa with a theoretical mass of 664 kDa, was found on the WSSV genome. The ATG of this ORF (ATC*ATG*G) is in a favorable Kozak context and a polyA signal is present 87 nucleotides downstream of this ORF. No consensus sequence for a TATA box and only one polyA signal is located in the coding region of this ORF. Several methionines encoded in this ORF are in a favorable Kozak context. Additional analysis will have to prove if this giant protein encoded by this ORF is indeed expressed. This is the largest ORF to date in viruses. ORFs of about half the size have been identified in herpesviruses and the proteins encoded by these ORFs are located in the tegument. ORFs of similar sizes as WSSV ORF167 are found in eukaryotes and are members of the family of giant actin-binding/cytoskeletal cross-linking proteins (Sun *et al.*, 1999). No homology with sequences in the GenBank was found and therefore the function of this exceptionally large ORF remains unresolved.

Gene families

The FASTA3 program package (Pearson, 2000) was used to identify gene families in the WSSV genome. All ORFs, except for those in the *hr* regions, were compared to find genes with homology to each other. Alignments were made of ten groups of ORFs belonging to the same putative gene family, and which are possibly duplicated in the WSSV genome as they showed pairwise similarities of 40% or higher (Table 1). Gene family 1 consists of three duplicated ORFs (ORF1, 31 and 153) with an amino acid similarity of about 42%, which are the major envelope protein VP28 and two nucleocapsid proteins VP26 and VP24 (van Hulten *et al.*, 2000b). Family 2 contains both putative protein kinase genes (ORFs 2 and 61) which have a 46% similarity. All other families contained ORFs with unknown functions. The highest homology (55% identity, 73% similarity) was found for gene family 7.

Comparison of the WSSV genome with other virus families

WSSV resembles baculoviruses in overall genome structure based on the large circular DNA genome and presence of *hrs*. Baculoviruses share about 50% of their genes, which

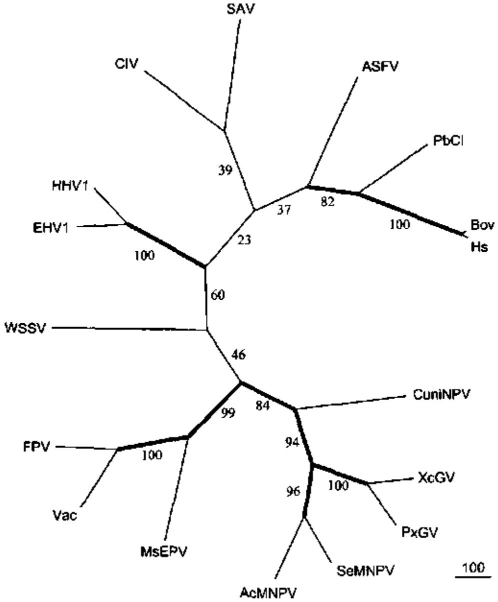


Figure 8.3 Bootstrap analysis (100 replicates) of an unrooted phylogenetic tree of DNA polymerase proteins constructed with the PAUP heuristic search algorithm. Numbers at the branches indicate frequency of clusters and frequencies over 70% are indicated by thick lines. The bar at the bottom equals a branch length of 100. DNA polymerase genes used and their accession numbers in brackets: SAV: *Spodoptera frugiperda* ascovirus 1 (CAC19170), EHV1: Equine herpesvirus 1 (NP_041039), HHV1: Human herpesvirus 1 (NP_044632), PbCl: *Paramecium bursaria Chlorella* virus 1 (NP_048532), Bov: *Bos taurus* (P28339), Hs: *Homo sapiens* (S35455), CuniNPV: *Culex nigripalpus* baculovirus (AF274291), XcGV: *Xestia c-nigrum* GV (NP_059280), PxGV: (NP_068312), SeMNPV: *Spodoptera exigua* MNPV (NP_037853), AcMNPV: *Autographa californica* MNPV, (NP_054095), MsEPV: (NP_048107), Vac: Vaccinia virus (NP_063712), FPV: Fowlpox virus (NP_039057), ASFV: African swine fever virus (NP_042783), CIV: Chilo iridescent virus (AAD48150).

separates WSSV from baculoviruses, as for WSSV only 6% of its genes have a viral or cellular homologue in GenBank. Furthermore, no specific similarity with other viruses was observed based on gene content. The genes for RR1, RR2, TK, TMK and PK were used in phylogenetic analysis to compare the position of WSSV relative to other viruses (van Hulten *et al.*, 2000c; Tsai *et al.*, 2000b; van Hulten and Vlak, 2001a). Here we report the analysis of WSSV DNA polymerase, an enzyme that is the prototype for phylogenetic analysis.

Phylogeny of DNA polymerase

The putative DNA polymerase gene (ORF27) was used in an alignment with 14 other viral and 2 eukaryotic polymerases. All seven conserved DNA polymerase sequence motifs and the three conserved regions implicated in DNA polymerase 3'-5' exonuclease activity (Bernad *et al.*, 1987; Bernad *et al.*, 1989) were identified. Phylogenetic analysis was performed using the region containing the conserved DNA polymerase motifs. Maximum parsimony phylogenetic trees were obtained using PAUP, followed by 100 bootstrap replicates to determine the 50% majority-rule consensus tree. Typically for maximum parsimony, bootstrap values of ≥ 70% correspond to a probability of ≥ 95% that the respective clade is a historical lineage.

In the DNA polymerase tree (Fig. 8.3) the different virus families are all present in clades which are high bootstrap-supported. The herpesviruses included in the tree are present in a branch, which is 100% bootstrap-supported. The same strong support exists for the poxviruses (99%), which are further separated into a chordopoxvirus and an entomopoxvirus branch. The baculoviruses are present in a well bootstrap-supported branch (84%) of the tree and further divided in NPVs and GVs. The *Culex nigripalpus* baculovirus (CuniNPV) was also included, but is not located in the NPV branch as has been shown before (Moser *et al.*, 2001). The remaining viral DNA polymerase genes, including those from WSSV and other viruses from different virus families, all have a unique position in the tree and do not share a most recent common ancestor. This DNA polymerase tree strengthens the proposition that WSSV is a member of a new virus family.

Conclusions

With a size of 293 kb WSSV is the largest animal DNA virus sequenced to date and second in size overall after *Chlorella* virus PBCV-1 (331 kb) (Li *et al.*, 1997). The largest animal DNA virus so far sequenced has been the fowlpox virus (FPV), infecting chickens and turkeys, with a size of 288 kb (Afonso *et al.*, 2000). Known large DNA viruses, such as herpesviruses (108-229 kb) (Montague and Hutchinson III, 2000; Mar Albà *et al.*, 2001), iridoviruses (102 kb) (Tidona and Darai, 1997), baculoviruses (100-180 kb) (Hayakawa *et al.*, 2000) and poxviruses (145-288 kb) (Accession: NC_002642; Afonso *et al.*, 2000) have genomes of considerable size and genetic complexity. The most remarkable property of WSSV is the lack of significant gene sequence homology to any member of these recognized virus families. The presence of an extremely large gene (ORF167), encoding a putative 664 kDa protein adds to the unique character of this virus. The available data including the sequence and phylogeny on DNA polymerase strongly suggest that WSSV is a member of a new virus family. The presence of hrs dispersed along the WSSV genome, a property shared with baculoviruses, may "supergroup" the large circular DNA viruses of arthropods. The analysis of the WSSV genome provides the first complete information of such a large DNA virus of crustaceans, and shows that this virus is distinct from previously identified DNA viruses. An improved understanding of the structure of this virus and its replication, its pathology and gene functions may permit the development of novel intervention strategies.

Materials and methods

WSSV isolation

The virus isolate used in this study originates from WSSV-infected *Penaeus monodon* shrimps imported from Thailand in 1996 and was obtained as described before (van Hulten *et al.*, 2000c). Crayfish *Procambarus clarkii* were injected intramuscularly with a lethal dose of WSSV. After one week the haemolymph was withdrawn from moribund crayfish and mixed with modified Alsever solution (Rodriguez *et al.*, 1995) as anticoagulant. The virus was purified by centrifugation at 80,000 x g for 1.5 h at 4°C on a 20-45% continuous sucrose gradient in TN (20 mM Tris, 400 mM NaCl, pH 7.4). The visible virus bands were removed and the virus particles were subsequently sedimented by centrifugation at 45,000 x g at 4°C for 1 h after dilution with TN. The virus pellet was resuspended in TE (pH 7.5).

WSSV DNA isolation, cloning, and sequence determination

The WSSV DNA was sequenced to a 6-fold genomic coverage using a shotgun approach essentially as described by Chen *et al.* (2001) for baculovirus *Helicoverpa armigera* NPV. The viral DNA was purified as described in van Hulten *et al.* (2000a) and sheared by nebulization into fragments with an average size of 1,200 bp. Blunt repair of the ends was performed with *Pfu* DNA polymerase (Stratagene) according to the manufacturer's directions. DNA fragments were size-fractionated by gel electrophoresis and cloned into the dephosphorylated *Eco*RV site of pBluescriptSK (Stratagene). After transformation into XL2 blue competent cells (Stratagene) 1510 recombinant colonies were picked randomly. DNA templates for sequencing were isolated using QIAprep Turbo kits (Qiagen) on a QIAGEN BioRobot 9600. Sequencing was performed using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready reaction kit with FS AmpliTaq DNA polymerase (Perkin Elmer) and analyzed on an ABI 3700 DNA Analyzer.

Sequences were base-called by the PRED basecaller and assembled with the PHRAP assembler (Ewing *et al.*, 1998; Ewing and Green, 1998). Using the PREGAP4 interface, PHRAP-assembled data were stored in the GAP4 assembly database (Bonfield *et al.*, 1995). The GAP4 interface and its features were then used for editing and sequence finishing. Consensus calculations with a quality cutoff value of 40 were performed from within GAP4 using a probabilistic consensus algorithm based on expected error rates output by PHRED. Sequencing PCR products bridging the ends of existing contigs filled remaining gaps in the sequence.

DNA sequence analysis

Genomic DNA composition, structure, and restriction enzyme pattern were analyzed with DNASTAR™ (Lasergene). Open reading frames (ORFs) encoding more than 50 amino acids were considered to be protein encoding and hence designated putative genes. DNA and protein comparisons with entries in the sequence databases were performed with FASTA and BLAST programs (Pearson, 1990; Altschul *et al.*, 1997). Multiple sequence alignments were performed with the ClustalX computer program (Thompson *et al.*, 1997). Phylogenetic analysis was performed with PAUP3.1 program (Swofford, 1993) using ClustalX to produce input files of aligned protein sequences. A heuristic search was performed, where starting trees were obtained by stepwise addition (starting seed 1), and tree-bisection-reconnection branch-swapping was performed with the MULPARS function. Bootstrap analysis according to Felsenstein (1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny.

Prediction of signal sequences and transmembrane domains was accomplished using the PSORT II prediction program which uses the McGeoch's method (McGeoch, 1985) and Von Heijne's method (Von Heijne, 1986) for the signal sequence recognition and the Klein *et al.*'s method (Klein *et al.*, 1985) to detect potential transmembrane domains. The program REPuter (Kurtz and Schleiermacher, 1999) was used to identify direct, reversed, and palindromic repeat families.

PCR primers

PCR was performed using three primer pairs, which amplify the C42, A6 and LN4 fragment described by Wang *et al.* (2000c). The 146F and 146R primer pair described by Lo *et al.* (1996) was used as a control in the PCR reaction. Purified WSSV DNA, used for sequencing, was used as template in the PCR reactions. The PCR products were separated in 0.8% agarose gels (Sambrook *et al.*, 1989).

Nucleotide sequence accession number

The WSSV genome sequence has been deposited in GenBank under accession no. AF369029.

Acknowledgements

This research was supported by Intervet International BV, Boxmeer, The Netherlands. Marleen Abma-Henkens, Paul Mooyman, and Joost de Groot are thanked for their skilful technical assistance. We are grateful to Dr. Douwe Zuidema for reading the manuscript.

Chapter 9

General discussion

Since the early 1990s, White spot syndrome virus (WSSV) has had a devastating impact on shrimp culture (Rosenberry, 2000; Naylor *et al.*, 2000). Measures to control this virus are urgently needed and characterization of this virus is therefore pivotal. In this thesis research the complete dsDNA genome sequence of WSSV has been determined and analyzed. Furthermore, the major structural proteins present in the virion have been identified and the involvement of one of these proteins in the infection confirmed. This opens the way to the design of vaccination strategies and the taxonomical classification of this virus.

WSSV taxonomy

Research on WSSV initially focussed on the clinical symptoms, histopathology and virion morphology in order to obtain tools for virus diagnosis. As only limited molecular information was available, it was unclear to which virus family WSSV belonged. Because of its virion morphology and nuclear localization WSSV was originally thought to be a member of the genus of the 'non-occluded' rod-shaped viruses (Nudibaculovirinae) (Wongteerasupaya *et al.*, 1995; Huang *et al.*, 1995; Wang, *et al.*, 1995), a separate genus within the *Baculoviridae* (Franki *et al.*, 1991), which was in 1995 eliminated from the *Baculoviridae* and orphaned (Murphy *et al.*, 1995). Unlike members of the baculoviruses, these viruses do not produce occlusion bodies. The two most well characterized viruses are *Oryctes* virus (Or-1V), which infects *Oryctes rhinoceros* and Hz-1V, which was found in *Helicoverpa zea*. Both viruses have rod-shaped enveloped particles with supercoiled, double-stranded DNA genomes of 127 kb (Or-1V, Crawford *et al.*, 1985) and 225 kb (Hz-1V, Huang *et al.*, 1982). The Or-1V enveloped virus particle is 220 x 120 nm and a tail-like structure has been reported at one end. The Hz-1V virion is longer and narrower, measuring 414 x 80 nm (Granoff and Webster, 1999). Both viruses have several features in common with WSSV, but as only for a few genes sequence data is available for these viruses, a comparison is difficult.

To reveal the taxonomic position of WSSV phylogeny was performed on several genes identified on the WSSV genome, including the ribonucleotide reductase large (*rr1*) and small (*rr2*) subunits (Chapter 2), the protein kinase (*pk*) genes (Chapter 3) and the DNA polymerase (*pol*) gene (Chapter 8). The unrooted parsimonious trees constructed for these genes showed that WSSV does not share a recent common ancestor with members of established virus families. Two additional genes, encoding thymidine kinase (*tk*) and thymidylate kinase (*tmk*) have been used in phylogenetic analyses, which also revealed a unique position for WSSV (Tsai *et al.*, 2000b). The position of WSSV relative to other virus families appears not to be conserved when these trees are compared. However, results based on single gene phylogenies are seldom uniform (McGeoch and Cook, 1994; Herniou *et al.*, 2001). The conflicts in such trees may reflect the unequal rate of gene evolution or the occurrence of horizontal gene transfer. Genes can be acquired at different times in evolution as has been described for the baculovirus RR genes (van Strien *et al.*, 1997; van Hulten *et al.*, 2000c).

Whole genome phylogeny

With more complete viral genomes sequences becoming available, whole genome phylogeny is more often used. "Gene order phylogeny" is a method, which uses the gene order of conserved genes in the genomes as a phylogenetic marker. This method can only be used to study closely related viral genomes within a virus family (Herniou *et al.*, 2001; Hannanhalli *et al.*, 1995). "Gene content phylogeny" uses the presence/absence of genes in completely sequenced genomes as a phylogenetic marker (Herniou *et al.*, 2001; Montaque and Hutchison III, 2000). Compared to the gene order phylogeny, this method can be used to study phylogeny for more diverged genomes. As this method is based on the identification of gene homologues between viruses, it is not suitable for WSSV phylogeny, as for only few ORFs (6%) a gene homologue could be identified (van Hulten *et al.*, 2001a).

Combined protein data sets of multiple genes have been used successfully for the generation of well-resolved phylogenetic trees. This method does not require completed genome sequences and can be used for unrelated genomes when multiple gene homologues are present. Universal trees of life using a wide variety of species, including Archaea (archaebacteria), Bacteria (eubacteria) and Eucarya (eukaryotes) were constructed using this method (Brown *et al.*, 2001). The relationships among completely sequenced baculovirus genomes could also be resolved using their conserved genes and this method was superior to the gene order phylogeny or gene content phylogeny (Herniou *et al.*, 2001). As combined gene phylogeny can be used for unrelated genomes, this method is suitable to resolve the phylogenetic position of WSSV compared to other large DNA viruses. However, the amount of genes conserved among large DNA viruses is limited and may not be enough to obtain a completely resolved tree.

Combined gene phylogeny for WSSV

The WSSV genome was analyzed for the presence of genes conserved in large DNA viruses, in order to select a set of genes which can be used in combined phylogenetic analysis. Eight such genes were identified: two *pks* (ORF2 and ORF61), *pol* (ORF27), *dUTP pyrophosphatase* (*dUTPase*) (ORF71), *rr1* (ORF92), *rr2* (ORF98), *tk* and *tmk* (ORF171). Subsequently, the presence of these genes on the genomes of members of *Poxviridae, Asfarviridae, Iridoviridae, Phycodnaviridae, Baculoviridae, Herpesviridae* and *Ascoviridae* was analyzed. As only limited sequence information is available in GenBank for members of the *Ascoviridae*, it was not possible to include this virus family in the analysis. Members of the *Baculoviridae* and *Phycodnaviridae* have no TK homologue, for TMK no homologues were present in members of the *Phycodnaviridae, Baculoviridae* and *Herpesviridae* and dUTPase was missing for the *Iridoviridae.* Therefore, a consensus tree for WSSV and the large DNA viruses (excluding the *Ascoviridae*) was constructed using the amino acid

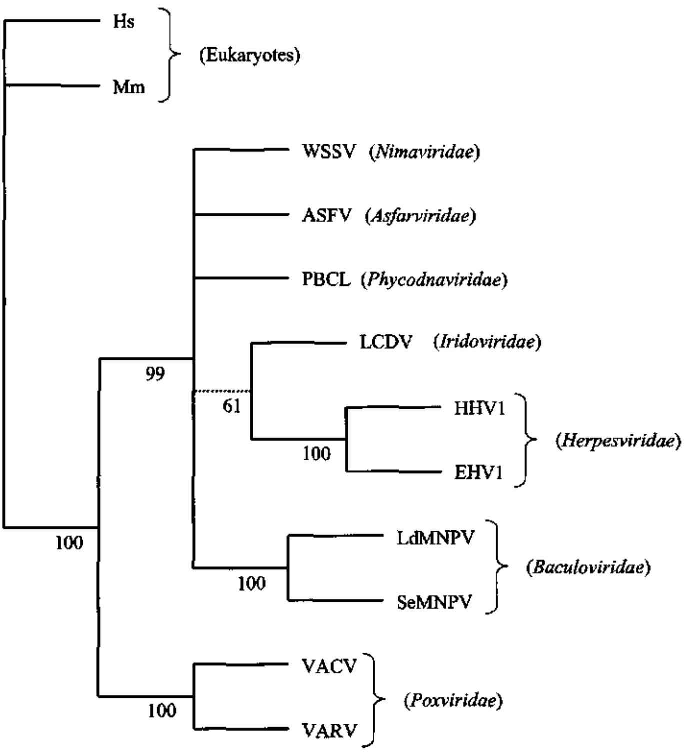


Figure 9.1 Bootstrap analysis of an unrooted phylogenetic tree of RR1, RR2, PK and POL constructed with the PAUP heuristic search algorithm. Number at the branches indicate frequency of clusters. The dotted lines represent bootstrap values below 70%. Hs: *Homo sapiens*; Mm: *Mus musculus*; ASFV: *African swine fever virus*; PBCL: *Paramecium bursaria Chlorella virus 1*; LCDV: *Lymphocystis disease virus*; HHV1: *Human herpesvirus 1*; EHV1: *Equid herpesvirus 1*; LdMNPV: *Lymantria dispar nucleopolyhedrovirus*; SeMNPV: *Spodoptera exigua MNPV*; VACV: *Vaccinia virus*; VARV: *Variola virus*.

sequences of the RR1, RR2, PK and POL proteins.

A phylogenetic tree was obtained from the combined alignments of the conserved regions of these proteins using PAUP (version 4.0), followed by 500 bootstrap replicates (Fig. 9.1). In this tree, the two poxviruses branch first after the eukaryotes, which suggests that compared to herpesviruses these viruses are more closely related to their eukaryotic hosts. All members of established virus families are present in well bootstrap supported branches of this tree. Remarkably, LCDV is present in a clade with the herpesviruses, which may suggest an ancient relationship of these viruses. As the bootstrap value is only 61%, no indisputable conclusions can be made. Based on this tree no recent relationship was revealed for WSSV with any of the large DNA virus families. Therefore, it is justified to accommodate WSSV in a separate virus family.

For the presented consensus tree only four genes were included, which resulted in a well bootstrap supported tree. However, to obtain a more representative tree, more genes from preferably more taxa should be included in the analysis. This will only be possible to some extent, as the number of genes shared among large DNA viruses is limited. In WSSV homologues of genes involved in DNA replication, nucleotide metabolism, mRNA biogenesis and protein modification (e.g. helicase, RNA polymerase subunits) are expected to be present, but have not been identified yet, because of their low homology.

Horizontal gene transfer

In the WSSV genome, a putative endonuclease gene was identified which showed highest homology (34% overall amino acid similarity and 17% identity) with the DNase I gene of *Penaeus japonicus* (GenBank accession: CAB55635). Alignment of the viral and eukaryotic endonuclease ORFs revealed that most catalytically and structurally important amino acid residues of the DNA/RNA non-specific endonuclease motif were present in the putative WSSV homologue (Witteveldt *et al.*, 2001). An unrooted parsimonious tree indicated that the non-specific endonuclease of WSSV was located in a well bootstrap supported clade containing only arthropods (Witteveldt *et al.*, 2001). Therefore, it is tempting to conclude that WSSV obtained this gene from one of its hosts. Further sequence information on WSSV hosts may reveal if more WSSV genes are acquired by horizontal gene transfer.

Aquatic vs. terrestrial virus family members

Knowing now that WSSV, despite earlier references (Wongteerasupaya *et al.*, 1995; Huang *et al.*, 1995; Wang, *et al.*, 1995), does not represent a baculovirus, but rather is a representative of a new, separate taxon, the question is justified if baculoviruses do occur in Penaeid shrimp. Two putative members of the *Baculoviridae* have been identified so far, Monodon Baculovirus (MBV) (Mari *et al.*, 1993) and Baculovirus Penaei (BP) (Bonami *et al.*, 1995b). Both viruses produce occlusion bodies, characteristic for the *nucleopolyhedroviruses*. As no sequence data are available for these viruses yet, their definitive classification still awaits authorization by the ICTV.

Large DNA viruses most probably co-evolved with their hosts. Thus aquatic viruses can differ substantially from their non-aquatic counterparts, as demonstrated for Channel Catfish Virus (CVV), a member of the *Herpesviridae* (Davison, 1992). This virus has diverged so extensively from other members of the *Herpesviridae*, that even the sequences of conserved proteins, like the structural proteins, are no longer conserved. Also the *Baculoviridae* contain an outgroup derived from an aquatic environment, the *Culex nigripalpus* baculovirus (CuniNPV) (Moser *et al.*, 2001). This virus though has not diverged as extensively as CVV, since many genes, unique for baculoviruses, including the structural proteins genes, are still identifiable in its genome (Moser *et al.*, 2001).

One could speculate that WSSV is an extremely diverged member of the *Baculoviridae*, as CVV is for the *Herpesviridae*. However, despite the fact that there is no sequence homology between CVV and other herpesvirus proteins, the major features of their virions are remarkably similar. Beside the conserved architecture of the nucleocapsid structure, the protein composition of CCV basically mirrors that of HHV1 (Booy *et al.*, 1996). The design of the virion structure is more and more recognized as a well-defined taxonomical marker

(Strauss and Strauss, 2001; Yan *et al.*, 2000). Even though the morphology of the WSSV virion is reminiscent of baculoviruses, the tail-like appendage of the envelope and the crosshatched appearance of the nucleocapsid of WSSV, are different from the morphology observed for the *Baculoviridae.* Furthermore, the protein composition of baculoviruses, exemplified by *Autographa californica* MNPV and SeMNPV, is very different to that of WSSV (Fig. 9.2). In baculoviruses approximately 10 polypeptide species can be distinguished by SDS-PAGE analysis, whereas in WSSV only 5 prominent bands can be identified, besides an unknown number of minor polypeptides. Furthermore, the presence of the putative Penaeid baculoviruses MBV and BP demonstrates that genuine baculoviruses in shrimp appear to exist with a morphology similar to that of insect baculoviruses, producing occlusion bodies in infected cells.

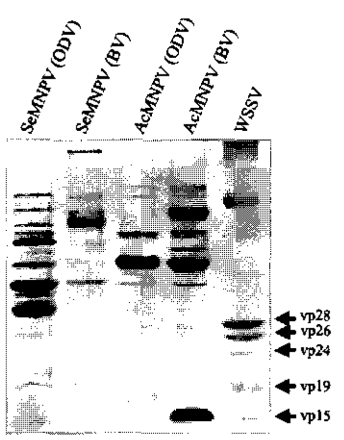


Figure 9.2 15% Coomassie Brilliant Blue-stained SDS-PAGE gel of purified virus proteins. Lane 1: SeMNPV occlusion derived virions (ODVs); 2: SeMNPV budded virions (BVs); 3: AcMNPV ODVs; 4: AcMNPV BVs; 5: WSSV.

WSSV genome organization

WSSV has a circular genome configuration, a characteristic it has in common with the *Baculoviridae* and the *Ascoviridae.* In the WSSV genome nine regions, each containing 3 to 8 direct repeat units of about 250 bp, were identified dispersed throughout the genome (Fig. 8.2). Similar regions have also been identified in baculovirus genomes and these regions were, in accordance to baculovirus teminology (Cochran and Faulkner, 1983), designated homologous region *(hr)* 1 to *hr*9 (van Hulten *et al.*, 2001a). Besides WSSV and baculoviruses, repeat regions dispersed throughout the genome have also been found in ascoviruses (Bigot *et al.*, 2000). In baculoviruses it was demonstrated that these *hrs* function as enhancers of transcription (Guarino and Summers, 1986) and putative origins of replication (Pearson *et al.*, 1992; Kool *et al.*, 1993). Based on the observation that all large DNA viruses with a circular genome configuration (WSSV, baculoviruses and ascoviruses) contain *hrs* dispersed in their genomes, it is very well possible that these all have a function in virus replication. Their presence may imply that these viruses share a similar replication strategy and this may cluster WSSV, the baculoviruses and the ascoviruses into a superfamily of large circular DNA viruses. The function of the WSSV *hrs* in either replication or transcription remains to be investigated.

Conclusion

The unrooted parsimonious tree constructed from the combined RR1, RR2, PK and POL genes of several large DNA viruses, is a well bootstrap supported tree and the different viruses included in these trees are present in clades corresponding to the virus families, according to the current taxonomy (van Regenmortel *et al.*, 2000). Based on this extensive phylogenetic analysis and the distinct morphology of its virion, WSSV has a unique position among the large DNA viruses and it is therefore justified to consider this virus as a member of a new virus family. A proposal to this end has been submitted to the International Committee on Taxonomy of Viruses for approval (Vlak *et al.*, 2001). Further research on the replication strategy of WSSV, baculoviruses and ascoviruses may reveal whether these viruses have a similar genome replication strategy, which may indicate an ancient relationship.

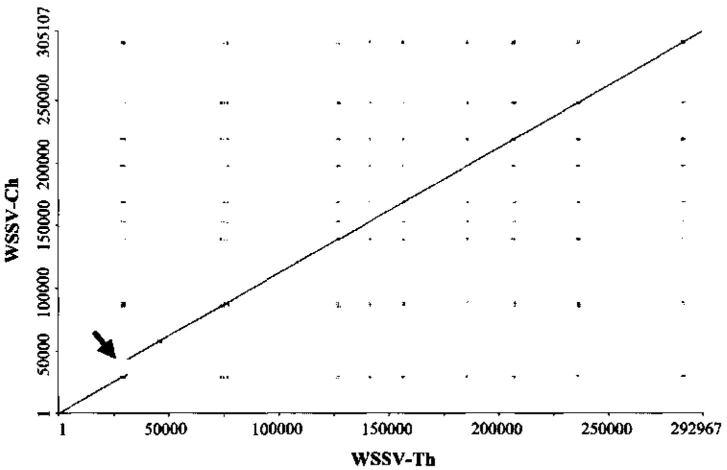


Figure 9.3 Dotplot comparison of WSSV-Th and WSSV-Ch. The comparison was performed with a window size of 30 and a stringency of 98%. Arrow indicates 12049 bp deletion in WSSV-Th.

Potential variation within the species WSSV

WSSV was first discovered in Taiwan in 1992 and subsequently spread to other shrimp farming areas in Southeast Asia, the Indian subcontinents and the Americas (Fig. 1.4). We recently determined the complete WSSV circular dsDNA genome sequence of 292,967 nucleotides from an isolate originating from Thailand in 1996 (WSSV-Th) (van Hulten *et al.*, 2001a). Although the WSSV-Th sequence was not determined from a clonal WSSV isolate, the sequence heterogeneity was minimal (less than 0.01%). Taken into account the wide host range and geographic distribution of the virus, it can be questioned if WSSV is an invariant species. Indeed some restriction fragment length polymorphism has been reported for WSSV DNA originating from different geographical isolates (Nadala and Loh, 1998; Lo *et al.*, 1999; Wang *et al.*, 2000c).

Recently, the complete sequence of a second isolate of WSSV has become available (GenBank accession number: AX151396) originating from China. To make a distinction between both completed genome sequences the latter sequence will be referred to as WSSV-Ch. The total size of the WSSV-Ch genome is 305,107 bp, which is 12,140 bp longer than the WSSV-Th genome. To analyze the genetic differences between the WSSV-Th and WSSV-Ch the sequences were compared using dotplot analysis (Fig. 9.3). This revealed that the two genome sequences were very similar (over 99% identity) and that the size difference could be explained by a 12 kb segment present in the WSSV-Ch genome and absent in the WSSV-Th genome (Fig. 9.3).

The 12,048 bp insertion of in the WSSV-Ch genome was located at position 31135 in the intergenic region of ORF23 and ORF24 in the WSSV-Th genome, and contains 14 ORFs (Fig. 9.4), of which two ORFs (ORF-A and ORF-L) belong to gene family 4 of the WSSV-Th genome (van Hulten *et al.*, 2001a). ORF-M showed homology with WSSV-Th ORF23, which is flanking the deletion in WSSV-Th. For the other ORFs no homologues could be identified in GenBank. The presence of ORFs belonging to WSSV gene families in the 12,049 bp WSSV-Ch specific segment, strongly suggests that this is an authentic part of the WSSV genome and that the shorter size of the WSSV-Th genome is due to a deletion event. This deletion could have been present in the original virus isolate and thus represent geographic variation or been generated during serial passage in laboratory animals. In baculoviruses it has been shown that genomes with smaller size and an increased density of *hrs*, which function as origins of DNA replication (ori's) may have a replication advantage (Krell, 1996). If the *hrs* on the WSSV genome function as ori's, the WSSV-Th genome may have a replication advantage compared to the WSSV-Ch, as it has an increased density of *hrs* and this may result in the predominance of this genotype after several viral passages in *P. monodon* and *Procambarus clarkii*. The function of the genes present on the 12048 bp segment is not known. However, these genes are apparently dispensable for infection and replication in *P. monodon* and *Procambarus clarkii* (van Hulten *et al.*, 2001b).

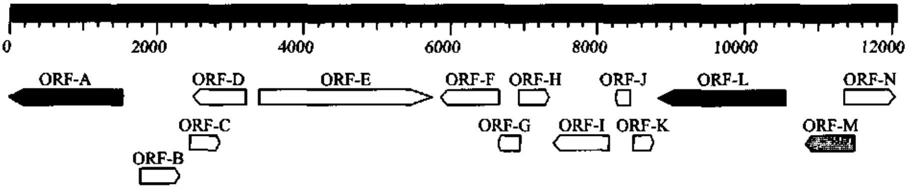


Figure 9.4 Physical map of the 12.0 kb segment of WSSV-Ch. The top bar shows the size in kb. The location and direction of transcription of the ORFs are shown below, with arrows. The black arrows indicate members of WSSV gene family 4 and the grey arrow the homologue of WSSV-Th ORF23.

It should be noted that two WSSV fragments (C42 and LN4) described by Wang *et al.* (2000c) for WSSV detection are lacking in the WSSV-Th genome sequence (van Hulten *et al.*, 2001a), as they are part of the 12049 bp insertion fragment of the WSSV-Ch genome. Therefore, these fragments are not suitable for reliable detection of WSSV, as they are present in a variable part of the genome, not essential for WSSV infection of shrimp, and absent in the infectious WSSV-Th isolate.

Further comparison of the two genomes revealed several less prominent differences, which fall into two categories: (1) point mutations, (2) rearrangements. An alignment was made with the two isolates and when the inserted/deleted regions were omitted, an overall sequence homology was found of about 99.97 %. In the WSSV-Th genome nine *hrs* have been identified (van Hulten *et al.*, 2001a), which were also present in WSSV-Ch. The location and sequence of these *hrs* were conserved, whereas the number of repeat units within the *hrs* varied between the isolates. In *hr1*, *hr3* and *hr8* extra repeat units are present in WSSV-Th. In *hr9* the opposite situation is found, resulting in a longer repeat region in the WSSV-Ch sequence. Several ORFs containing small repeated domains, WSSV-Th ORFs 75, 94 and 125, were also different in both isolates, as the number of repeated domains varied. Finally, a small region of 399 bp containing 362 bp of the 5'end of WSSV-Th ORF14 and 27 bp of the 3'end of ORF15 was absent in the WSSV-Ch sequence and replaced with a 584 bp region with no homology with data in GenBank.

It can be concluded that even though the WSSV-Th and WSSV-Ch genomes are very similar (over 99 % sequence identity), several genomic rearrangements have occurred, of which the 12 kb deletion is the most drastic one. This raises the question whether other isolates from geographically different locations will also exhibit genomic variations. Furthermore, this comparison shows that care should be taken when using WSSV detection tests, based on genomic fragments. Multiple fragments located at various position of the WSSV genome may be included in a DNA probe or for detection using PCR to improve WSSV detection.

The WSSV virion

WSSV virion proteins

Five major structural proteins were identified in the WSSV virion and their genes subsequently located on the WSSV genome. The 28 kDa (VP28) and 19 kDa (VP19) proteins were found to be located in the WSSV envelope, whereas the 26 kDa (VP26), the 24 kDa (VP24) and 15 kDa (VP15) proteins were found to be associated with the envelope. A number of papers describing the structural proteins of WSSV report on proteins with different apparent sizes, obviously depending on the electrophoretic conditions and possibly the protein isolation procedures used. Hence, some confusion has been introduced in WSSV literature concerning the WSSV virion proteins. To obtain a consensus view on occurrence of the major virion proteins their reported sizes were compared (Table 9.1). In some studies larger proteins (65 – 75 kDa) were described (Nadala *et al.* 1998; Huang *et al.*, 2001) (not shown in table 9.1). These proteins most likely represent haemolymph contaminations as they

Theoretical size		**Apparent size**	**Nadala and Loh., 1998**	**Hameed *et al.*, 1998**	**Wang *et al.*, 2000**	**Shih *et al.*, 2001**	**Huang *et al.*, 2001**
						35 kDa	
204 aa	22 kDa	28 kDa: **VP28**	27.5 kDa	27 kDa	25 kDa (VP28)	28 kDa	28 kDa
204 aa	22 kDa	26 kDa: **VP26**			23 kDa (VP26)		
208 aa	23 kDa	24 kDa: **VP24**	23.5 kDa	22 kDa		24 kDa	23 kDa
121 aa	13 kDa	19 kDa: **VP19**	19 kDa	18 kDa	19 kDa	19 kDa	19 kDa
61 aa	7 kDa	15 kDa: **VP15**	?		15 kDa (VP15?)		16 kDa

Table 9.1 WSSV structural proteins described in the literature. First three columns describe the proteins characteristics determined in this thesis.

can also be observed in control purifications of uninfected animals (van Hulten *et al.*, 2000a; Shih *et al.*, 2001; Huang *et al.*, 2001). The apparent protein sizes described by Wang *et al.* (2000b) are most diverged, however the N-terminal sequencing performed by the authors revealed that the 25 kDa polypeptide band corresponds to VP28 and the 23 kDa band to VP26. The 15 kDa band most probably corresponds to VP15, despite the fact that the N-terminal sequence data described only has a 75% identity to the VP15 polypeptide (van Hulten *et al.*, 2001c). VP15 has not been observed by all authors, as in some gel systems this band has been eluted from the gel. Nadala *et al.* (1998) reported a small protein associated with the viral DNA (indicated with the question mark). Shih *et al.* (2001) reported an extra band of 35 kDa that has not been identified in other reports. The most striking difference observed, is the absence of VP26 in most analyses (Table 9.1). The 28 kDa (VP28) and 19 kDa (VP19) proteins have been identified in all studies and were described to be present in the viral envelope (Nadala *et al.*, 1998; van Hulten *et al.*, 2000a, 2001b).

As poly- and monoclonal antibodies against the different virion proteins are available now (van Hulten *et al.*, 2001b, unpublished results; Shih *et al.*, 2001), it will be possible to reveal possible differences in protein composition among different isolates.

WSSV virion structure

The WSSV nucleocapsid appears to be formed by stacks of rings, which are composed of two rows of regularly spaced subunits (Durand *et al.*, 1997). Three major proteins were identified in the nucleocapsid, VP15, VP24 and VP26. VP15 has a high homology to histon proteins and therefore binds most likely to the WSSV DNA, forming the nucleoprotein core. The capsid is mainly composed of VP26 and VP24 of which VP26 is the major component. Both VP26 and VP24 contain a hydrophobic domain, which may have a function in the formation of homo- and multimers. To reveal the structure, composition and assembly of the WSSV nucleocapsid the interactions between various proteins has to be determined. Expression of these proteins in insect cells using baculovirus vectors resulted in proteins with identical sizes as in the purified virions (van Hulten *et al.*, 2001c). The baculovirus-insect cell expression system has proven to be an effective system to study capsid formation of complex viruses like herpesviruses (Tatman *et al.*, 1994). Along this line, the baculovirus-expressed VP26 and VP24 may be suitable to study WSSV nucleocapsid formation.

In the envelope two major proteins have been identified, VP28 and VP19. VP28 is involved in the systemic infection of WSSV in shrimps, as was demonstrated by *in vivo* neutralization experiment using VP28 polyclonal antiserum (van Hulten *et al.*, 2001b). Future experiments using *in vivo* neutralization will demonstrate which part of VP28 is involved in the neutralization process and further reveal the role of VP28 in WSSV attachment and entry into the shrimp. Viruses may use different envelope proteins for initial entry and systemic spread. The involvement of VP28 in the initial entry in the shrimp still has to be demonstrated. It cannot be excluded that additional WSSV envelope proteins, such as VP19, are involved in this process, either alone or in concert with VP28. Antibodies against VP19 will assist in testing this hypothesis.

The host range and tissue tropism of enveloped viruses is often dependent on cell attachment and fusion (Schneider-Schaulies, 2000), properties conferred upon each virus by one or sometimes more envelope glycoproteins (Granof & Webster, 1999). However, neither VP28 nor any of the other major WSSV virion proteins is glycosylated (van Hulten *et al.*, 2001c). As WSSV has a large host range and a broad tissue tropism, WSSV may recognize a relatively common determinant (e.g. a sugar moiety of a membrane glycoprotein) conserved on the surface of crustacean species as a receptor. The nature of the WSSV receptor is not known yet, however identification of this receptor may open the way for the design of WSSV intervention strategies.

WSSV vaccines or intervention strategies

WSSV has become a global problem in aquaculture and its economic impact is enormous (Naylor *et al.*, 2000). Therefore, the design of intervention strategies is urgently needed and vaccination should be considered. The potential of shrimp vaccination is equivocal as the immune system of crustaceans has not been studied in great detail. Research showed that there are many parallels between the innate immunity -the ancient form of defence against microbial infection- in vertebrates and invertebrates (Khush and Lemaitre, 2000). However, the adaptive immunity using somatic gene rearrangements to generate lymphocytes expressing random, unique antigen receptors is restricted to vertebrates (Khush and Lemaitre, 2000; Arala-Chaves and Sequeira, 2000). Some studies suggest that there is a different form of adaptive immune response in invertebrates and bacterial vaccines have proven to be effective for shrimp (Arala-Chaves and Sequeira, 2000; Teunissen *et al.*, 1998).

In this thesis the complete genome sequence and the major structural proteins of WSSV have been determined. This solid information has created a good starting point for further studies on the replication strategy and infection mechanism of the virus, and last but not least, will open the way for the design of novel strategies to control this devastating pathogen.

References

Afonso, C. L., Tulman, E. R., Lu, Z., Oma, E., Kutish, G. F., and Rock, D. L. (1999). The genome of *Melanoplus sanguinipes* entomopoxvirus. *Journal of Virology* **73**, 533-552.

Afonso, C. L., Tulman, E. R., Lu, Z., Zsak, L., Kutish, G. F., and Rock, D. L. (2000). The genome of fowlpox virus. *Journal of Virology* **74**, 3815-3831.

Ahrens, C. H., Russell, R. L. Q., Funk, C. J., Evans, J. T., Harwood, S. H., and Rohrmann, G. F. (1997). The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* **229**, 381-399.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Anonymous. (1999). Genome project aims to combat prawn scourge. *Nature* **397**, 465.

Arala-Chaves, M., and Sequeira, T. (2000). Is there any kind of adaptive immunity in invertebrates? *Aquaculture* **191**, 247-258.

Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M., and Possee, R. D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**, 586-605.

Bacher, N., Zisman, Y., Berent, E., and Livneh, E. (1991). Isolation and characterization of PKC-L, a new member of the protein kinase C-related gene family specifically expressed in lung, skin, and heart. *Molecular Cell Biology* **11**, 126-133.

Baldo, A. M., and McClure, M. A. (1999). Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. *Journal of Virology* **73**, 7710-7721.

Baylis, S. A., Banham, A. H., Vydelingum, S., Dixon, L. K., and Smith, G. L. (1993). African swine fever virus encodes a serine protein kinase which is packaged into virions. *Journal of General Virology* **67**, 4549-4556.

Belcher, C. R., and Young, P. R. (1998). Colourimetric PCR-based detection of monodon baculovirus in whole *Penaeus monodon* postlarvae. *Journal of Virological Methods* **74**, 21-29.

Berk, A. J. (2000). TBP-like factors come into focus. *Cell* **103**, 5-8.

Bernad, A., Blanco, L., Lazaro, J. M., Marin, G., and Salas, M. (1989). A conserved 3'-5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59**, 219-228.

Bernad, A., Zaballos, Z., Salas, M., and Blanco, L. (1987). Structural and functional relationships between prokaryotic and eukaryotic DNA polymerases. *EMBO Journal* **6**, 4219-4225.

Bigot, Y., Stasiak, K., Rouleux-Bonnin, F., and Federici, B. A. (2000). Characterization of repetitive DNA regions and methylated DNA in ascovirus genomes. *Journal of General Virology* **81**, 3073-3082.

Bonami, J. R., Bruce, L. D., Poulos, B. T., Mari, J., and Lightner, D. V. (1995b). Partial characterization and cloning of the genome of PvSNPV (=BP-type virus) pathogenic for *Penaeus vannamei. Diseases of Aquatic Organisms* **23**, 59-66.

Bonami, J. R., Hasson, K. W., Mari, J., Poulos, B. T., and Lightner, D. V. (1997). Taura syndrome of marine penaeid shrimp: characterization of the viral agent. *Journal of General Virology* **78**, 313-319.

Bonami, J. R., Mari, J., Poulos, B. T., and Lightner, D. V. (1995a). Characterization of hepatopancreatic parvo-like virus, a second unusual parvovirus pathogenic for penaeid shrimps. *Journal of General Virology* **76**, 813-817.

Bonami, J. R., Trumper, B., Mari, J., Brehelin, M., and Lightner, D. V. (1990). Purification and characterization of the infectious hypodermal and haematopoietic necrosis virus of penaeid shrimps. *Journal of General Virology* **71**, 2657-2664.

Bonfield, J. K., Smith, K. F., and Staden, R. (1995). A new DNA sequence assembly program. *Nucleic Acids Research* **23**, 4992-4999.

Boonyaratpalin, S., Supamattaya, K., Kasornchandra, J., Direkbusaracom, S., Aekpanithanpong, U., and Chantanachooklin, C. (1993). Non-occluded baculo-like virus, the causative agent of yellow head disease in the black tiger shrimp (*Penaeus monodon*). *Fish Pathology* **28**, 103-109.

Booy, F. P., Trus, B. L., Davison, A. J., and Steven, A. C. (1996). The capsid architecture of channel catfish virus, an evolutionarily distant herpesvirus, is largely conserved in the absence of discernible sequence homology with herpes simplex virus. *Virology* **215**, 134-141.

Boursnell, M., Shaw, K., Yanez, R. J., Vinuela, E., Dixon, L., and Tekensopletters, I. (1991). The sequences of the ribonucleotide reductase genes from African swine fever virus show considerable homology with those of the orthopoxvirus, vaccinia virus. *Virology* **184**, 411-416.

Boyko, V. P., Karasev, A. V., Agranovsky, A. A., Koonin, E. V., and Dolja, V. V. (1992). Coat protein gene duplication in a filamentous RNA virus of plants. *Proceedings of the National Academy of Sciences USA* **89**, 9156-60.

Brown, J. R., Douady, C. J., Italia, M. J., Marshall, W. E., and Stanhope, M. J. (2001). Universal trees based on large combined protein sequence data sets. *Nature Genetics* **28**, 281-285.

Burton, D. R., Williamson, R. A., and Parren, P. W. H. I. (2000). Antibody and virus: binding and neutralization. *Virology* **270**, 1-3.

Cai, S. L., Huang, J., Wang, C. M., Song, X. L., Sun, X., Yu, J., Zhang, Y., and Yang, C. H. (1995). Epidemiological studies on the explosive epidemic disease of prawn in 1993-1994. *Journal of China Fish* **19**, 112-117.

Carreras, C. W., and Santi, D. V. (1995). The catalytic mechanism and structure of thymidylate synthase. *Annual Review of Biochemistry*, 64721-64762.

Chang, P., Lo, C., Wang, Y., and Kou, G. (1996). Identification of white spot syndrome associated baculovirus (WSBV) target organs in the shrimp *Penaeus monodon* by in situ hybridization. *Diseases of Aquatic Organisms* **27**, 131-139.

Chen, X. W., Ijkel, W. F. J., Tarchini, R., Sun, X. L., Sandbrink, H., Wang, H. L., Peters, S., Zuidema, D., Lankhorst, R. K., Vlak, J. M., and Hu, Z. H. (2001). The sequence of the Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus genome. *Journal of General Virology* **82**, 241-257.

Chou, H. Y., Huang, C. Y., Wang, C. H., Chiang, H. C., and Lo, C. F. (1995). Pathogenicity of a baculovirus infection causing white spot syndrome in cultured penaeid shrimp in Taiwan. *Diseases of Aquatic Organisms* **23**, 165-173.

Cochran, M. A., and Faulkner, P. (1983). Localization of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *Journal of Virology* **45**, 961-970.

Cowley, J. A., C.M., D., Spann, K. M., and Walker, P. J. (2000). Gill-associated virus of *Penaeus monodon* prawns: an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. *Journal of General Virology* **81**, 1473-1484.

Cowley, J. A., Dimmock, C. M., Wongteerasupaya, C., Boonsaeng, V., Panyim, S., and Walker, P. J. (1999). Yellow head virus from Thailand and gill-associated virus from Australia are closely related but distinct prawn viruses. *Diseases of Aquatic Organisms* **36**, 153-157.

Crawford, A. M., Ashbridge, K., Sheehan, C., and Faulkner, P. (1985). A physical map of *oryctes* baculovirus genome. *Journal of general virology* **66**, 2649-2658.

Davison, A. J. (1992). Channel catfish virus: A new type of herpesvirus. *Virology* **186**, 9-14.

del Val, M., Carrascosa, J. L., and Vinuela, E. (1986). Glycosylated components of African swine fever virus particles. *Virology* **152**, 39-49.

Dimmock, N. J. (1984). Mechanisms of neutralization of animal viruses. *Journal of General Virology* **65**, 1015-1022.

Dolja, V. V., Boyko, V. P., Agranovsky, A. A., and Koonin, E. V. (1991). Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. *Virology* **184**, 79-86.

Dolja, V. V., and Koonin, E. V. (1991). Phylogeny of capsid proteins of small icosahedral RNA plant viruses. *Journal of General Virology* **72**, 1481-1486.

Dormeyer, M., Schöneck, R., Dittmar, G. A., and Krauth-Siegel, R. L. (1997). Cloning, sequencing and expression of ribonucleotide reductase R2 from Trypanosoma brucei. *FEBS Letters* **414**, 449-453.

Durand, S., Lightner, D., Nunan, L., Redman, R., Mari, J., and Bonami, J. (1996). Application of gene probes as diagnostic tools for White Spot Baculovirus (WSBV) of penaeid shrimp. *Diseases of Aquatic Organisms* **27**, 59-66.

Durand, S., Lightner, D. V., Redman, R. M., and Bonami, J. R. (1997). Ultrastructure and morphogenesis of white spot syndrome baculovirus (WSSV). *Diseases of aquatic organisms* **29,** 205-211.

Elledge, S. J., Zhou, Z., and Allen, J. B. (1992). Ribonucleotide reductase: regulation, regulation, regulation. *Trends in Biochemical Science* **17,** 119-123.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* **8,** 186-194.

Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8,** 175-185.

Federici, B. A. (1997). Baculovirus pathogenesis. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 33-59. Plenum Press, New York.

Felsenstein, J. (1993). PHYLIP (Phylogeny Interference Package) version 3.5. Seattle: Department of Genetics, University of Washington

Flegel, T. W. (1997). Major viral diseases of the black tiger prawn (Penaeus monodon) in Thailand. *World Journal of Microbiology and Biotechnology* **13,** 433-442.

Flinta, C., Persson, B., Jornvall, H., and von Heijne, G. (1986). Sequence determinants of cytosolic N-terminal protein processing. *European Journal of Biochemistry* **154,** 193-196.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L., and Brown, F., Eds. (1991). Virus Taxonomy. Fifth Report of the International Committee on Taxonomy of Viruses. New York: Springer-Verlag.

Friesen, P. D. (1997). Regulation of Baculovirus Early Gene Expression. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 141-170. Plenum Press, New York.

Funk, C. J., Braunagel, S. C., and Rohrmann, G. F. (1997). Baculovirus structure. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 7-32. Plenum Press, New York.

Galmiche, M. C., Goenaga, J., Wittek, R., and Rindisbacher, L. (1999). Neutralizing and protective antibodies directed against vaccinia virus envelope antigens. *Virology* **254,** 71-80.

Garnier, J., Osguthorpe, D. J., and Robson, B. (1978). Analysis of the accuracy and implications of simple method for predicting the secondary structure of globular proteins. *Journal of Molecular Biology* **120,** 97-120.

Giglione, C., Serero, A., Pierre, M., Boisson, B., and Meinnel, T. (2000). Identification of eukaryotic peptide deformylases reveals universality of N-terminal protein processing mechanisms. *EMBO Journal* **19,** 5916-5929.

Gomi, S., Majima, K., and Maeda, S. (1999). Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus. *Journal of General Virology* **80,** 1323-1337.

Granoff, A., and Webster, R. G., Eds. (1999). Encyclopedia of Virology. Second edition ed. San Diego: Academic Press.

Guarino, L. A., and Summers, M. D. (1986). Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed early gene expression. *Journal of Virology* **60,** 215-223.

Hameed, A. S., Anilkumar, M., Raj, M. L. S., and Jayaraman, K. (1998). Studies on the pathogenicity of systemic ectodermal and mesodermal baculovirus and its detection in shrimp by immunological methods. *Aquaculture* **160,** 31-45.

Hanks, S. K., and Hunter, T. (1995). The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB Journal* **9,** 576-96.

Hanks, S. K., Quinn, A. M., and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241,** 42-52.

Hannenhalli, S., Chappey, C., Koonin, E. V., and Pevzner, P. A. (1995). Genome sequence comparison and scenarios for gene rearrangements: a test case. *Genomics* **30,** 299-311.

Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L., and Brunak, S. (1998). NetOglyc: Prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconjugate Journal* **15,** 115-130.

Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y., and Matsumoto, T. (2000). Sequence analysis of the Plutella xylostella granulovirus genome. *Virology* **275,** 358-372.

Hayakawa, T., Ko, R., Okano, K., Seong, S. I., Goto, C., and Maeda, S. (1999). Sequence analysis of the Xestia c-nigrum granulovirus genome. *Virology* **262,** 277-297.

Hayakawa, T., Rohrmann, G. F., and Hashimoto, Y. (2000). Patterns of genome organization and content in lepidopteran baculoviruses. *Virology* **278,** 1-12.

Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., S., C. J., and O'Reilly, D. R. (2001). Use of whole genome sequence data to infer baculovirus phylogeny. *Journal of Virology* **75,** 8117-8126.

Hibi, M., Nakajima, K., and Hirano, T. (1996). IL-6 cytokine family and signal transduction: a model of the cytokine system. *Journal of Molecular Medicine* **74,** 1-12.

Hohmann, A. W., and Faulkner, P. (1983). Monoclonal antibodies to baculovirus structural proteins: determination of specificities by Western blot analysis. *Virology* **125,** 432-444.

Hong, T., Summers, M. D., and Braunagel, S. C. (1997). N-terminal sequences from Autographa californica nuclear polyhedrosis virus envelope proteins ODV-E66 and ODV-E25 are sufficient to direct reporter proteins to the nuclear envelope, intranuclear microvesicles and the envelope of occlusion derived virus. *Proceedings of the National Academy of Sciences of the USA* **94,** 4050-4055.

Hsu, Y. L., Wang, K. H., Yang, Y. H., Tung, M. C., Hu, C. H., Lo, C. F., Wang, C. H., and Hsu, T. (2000). Diagnosis of *Penaeus monodon*-type baculovirus by PCR and by ELISA of occlusion bodies. *Diseases of Aquatic Organisms* **40,** 93-99.

Hu, Z. H., Arif, B. M., Sun, J. S., Chen, X. W., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1998). Genetic organization of the HindIII-I region of the single-nucleocapsid nucleopolyhedrovirus of Buzura suppressaria. *Virus Research* **55,** 71-82.

Huang, C. H., Zhang, L. R., Zhang, J. H., Xiao, L. C., Wu, Q. J., Chen, D. H., and Li, J. K. K. (2001). Purification and characterization of White Spot Syndrome Virus (WSSV) produced in an alternate host: crayfish, *Cambarus clarkii*. *Virus Research* **76,** 115-125.

Huang, J., Song, X. L., Yu, J., and Yang, C. H. (1995). Baculoviral hypodermal and hematopoietic necrosis - study on the pathogen and pathology of the explosive epidemic disease of shrimp. *Marine Fish Research* **16,** 11-23.

Huang, Y. S., Hedberg, M., and Kawanishi, C. Y. (1982). Characterization of the DNA of a nonoccluded baculovirus Hz-1V. *Journal of Virology* **43,** 174-181.

IJkel, W. F. J., van Strien, E. A., Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1999). Sequence and organization of the Spodoptera exigua multicapsid nucleopolyhedrovirus genome. *Journal of General Virology* **12,** 3289-3304.

Inouye, K., Miwa, S., Oseko, N., Nakano, H., Kimura, T., Momoyama, K., and Hiraoka, M. (1994). Mass mortalities of cultured kuruma shrimp *Penaeus japonicus* in Japan in 1993: electron microscopic evidence of the causitive virus. *Fish Pathology* **29,** 149-158.

Inouye, K., Yamano, K., Ikeda, N., Kimura, T., Nakano, H., Momoyama, K., Kobayashi, J., and Miyajima, S. (1996). The penaeid rod-shaped DNA virus (PRDV), which causes penaeid acute viremia (PAV). *Fish Pathology* **31,** 39-45.

Johnson, G. P., Goebel, S. J., and Paoletti, E. (1993). An update on the vaccinia virus genome. *Virology* **196,** 381-401.

Kasornchandra, J., Boonyaratpalin, S., and Itami, T. (1998). Detection of white-spot syndrome in cultured penaeid shrimp in Asia: Microscopic observation and polymerase chain reaction. *Aquaculture* **164,** 243-251.

Khush, R. S., and Lemaitre, B. (2000). Genes that fight infection: what the *Drosophila* genome says about animal immunity. *Trends in Genetics* **16,** 442-448.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365,** 520-527.

King, L. A., and Possee, R. D. (1992). "The baculovirus expression system." Chapman & Hall, London.

Klein, P., Kanehisa, M., and DeLisi, C. (1985). The detection and classification of membrane-spanning proteins. *Biochemical and Biophysical Acta* **815,** 468-476.

Kool, M., Ahrens, C. H., Vlak, J. M., and Rohrmann, G. F. (1995). Replication of baculovirus DNA. *Journal of General Virology* **76,** 2103-2118.

Kool, M., Voeten, J. T. M., Goldbach, R. W., Tramper, J., and Vlak, J. M. (1993). Identification of seven putative origins of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus DNA replication. *Journal of General Virology* **74,** 2661-2668.

Kozak, M. (1989). The scanning model for translation: an update. *Journal of Cell Biology* **108,** 229-41.

Krell, P. J. (1996). Passage effect of virus infection in insect cells. *Cytotechnology* **20,** 125-137.

Krieser, R. J., and Eastman, A. (1998). The cloning and expression of human deoxyribonuclease II. A possible role in apoptosis. *Journal of Biological Chemistry* **273,** 30909-30914.

Kurtz, S., and Schleiermacher, C. (1999). REPuter: fast computation of maximal repeats in complete genomes.

Bioinformatics **15,** 426-427.

Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, C. J., Evans, J. T., Slavicek, J. M., and Rohrmann, G. F. (1999). Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria dispar. *Virology* **253,** 17-34.

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

Leader, D. P., and Purves, F. C. (1988). The herpesvirus protein kinase: a new departure in protein phosphorylation? *Trends in Biochemical Science* **13,** 244-246.

Leonard, C. J., Aravind, L., and Koonin, E. V. (1998). Novel families of putative protein kinases in bacteria and archaea: Evolution of the "eukaryotic" protein kinase superfamily. *Genome Research. Oct* **8,** 1038-1047.

Li, Y., Lu, Z., Sun, L., Ropp, S., Kutish, G. F., Rock, D. L., and Van Etten, J. L. (1997). Analysis of 74 kb of DNA located at the right end of the 330-kb chlorella virus PBCV-1 genome. *Virology* **237,** 360-377.

Lightner, D. V. (1996). "A handbook of pathology and diagnostic procedures for diseases of penaeid shrimp.". Special publication of the World Aquaculture Society. Baton Rouge, LA.

Lightner, D. V., Redman, R. M., Hasson, K. W., and Pantoja, C. R. (1995). Taura syndrome in *Penaeus vannamei* (Crustacea: Decapoda): gross signs, histopathology and ultrastructure. *Diseases of Aquatic Organisms* **21,** 53-59.

Liu, A., Sahlin, M., Potsch, S., Sjoberg, B. M., and Graslund, A. (1998). New paramagnetic species formed at the expense of the transient tyrosyl radical in mutant protein R2 F208Y of Escherichia coli ribonucleotide reductase. *Biochemical and Biophysical Research Communications* **246,** 740-745.

Lo, C. F., Ho, C. H., Chen, C. H., Liu, K. F., Chiu, Y. L., Yeh, P. Y., Peng, S. E., Hsu, H. C., Liu, H. C., Chang, C. F., Su, M. S., Wang, C. H., and Kou, G. H. (1997). Detection and tissue tropism of white spot syndrome baculovirus (WSBV) in captured brooders of *Penaeus monodon* with a special emphasis on reproductive organs. *Diseases of Aquatic Organisms* **30,** 53-72.

Lo, C. F., Ho, C. H., Peng, S. E., Chen, C. H., Hsu, H. C., Chiu, Y. L., Chang, C. F., Liu, K. F., Su, M. S., Wang, C. H., and Kou, G. H. (1996a). White spot syndrome baculovirus (WSBV) detected in cultured and captured shrimp, crabs and other arthropods. *Diseases of Aquatic Organisms* **27,** 215-225.

Lo, C. F., Hsu, H. C., Tsai, M. F., Ho, C. H., Peng, S. E., Kou, G. H., and Lightner, D. V. (1999). Specific genomic DNA fragment analysis of different geographical clinical samples of shrimp white spot syndrome virus. *Diseases of Aquatic Organisms* **35,** 175-185.

Lo, C. F., Leu, J. H., Ho, C. H., Chen, C. H., Peng, S. E., Chen, Y. T., Chou, C. M., Yeh, P. Y., Huang, C. J., Chou, H. Y., Wang, C. H., and Kou, G. H. (1996b). Detection of baculovirus associated with white spot syndrome (WSBV) in penaeid shrimps using polymerase chain reaction. *Diseases of Aquatic Organisms* **25,** 133-141.

Lotz, J. M. (1997). Viruses, Biosecurity and Specific Pathogen Free Stocks In Shrimp Aquaculture. *World Journal of Microbiology and Biotechnology* **13,** 405-413.

Lu, A., and Miller, L. K. (1997). Regulation of baculovirus late and very late gene expression. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 193 - 216. Plenum Press, New York.

Lu, Z., Li, Y., Que, Q., Kutish, G. F., Rock, D. L., and Van Etten, J. L. (1996). Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map positions 88 to 182. *Virology* **216,** 102-23.

Mar Albà, M., Rhiju Das, C., Orengo, A., and Kellam, P. (2001). Genomewide function conservation and phylogeny in the Herpesviridae. *Genome Research* **11,** 43-54.

Mari, J., Bonami, J. R., and Lightner, D. V. (1998). Taura syndrome of penaeid shrimp: cloning of viral genome fragments and development of specific gene probes. *Diseases of Aquatic Organisms* **33,** 11-17.

Mari, J., Bonami, J. R., Poulos, B., and Lightner, D. (1993). Preliminary characterization and partial cloning of the genome of a baculovirus from *Penaeus monodon* (PmSNPV = MBV). *Diseases of Aquatic Organisms* **16,** 207-215.

Martignoni, M. E., Iwai, P. J., and Rohrmann, G. F. (1980). Serum neutralization of nucleopolyhedrosis viruses (baculovirus subgroup A) pathogenic for Orgyia pseudotsugata. *Journal of Invertebrate Pathology* **36,** 12-20.

McGeoch, D. J. (1985). On the predictive recognition of signal peptide sequences. *Virus Research* **3,** 271-286.

McGeoch, D. J. (1990). Evolutionary relationships of virion glycoprotein genes in the S regions of alphaherpesvirus genomes. *Journal of General Virology* **71,** 2361-2368.

McGeoch, D. J., and Cook, S. (1994). Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *Journal of Molecular Biology* **238,** 9-22.

McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E., and Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *Journal of General Virology* **69,** 1531-1574.

Momoyama, K., and Sano, T. (1996). Infectivity of baculoviral mid-gut gland necrosis virus (BMNV) to larvae of five crustacean species. *Fish Pathology* **31,** 81-85.

Montague, M. G., and Hutchison, C. A., 3rd. (2000). Gene content phylogeny of herpesviruses. *Proceedings of the National Academy of Sciences USA* **97,** 5334-9.

Moser, B., Becnel, J., White, S., Afonso, C., Kutish, G., Shanker, S., and Almira, E. (2001). Morphological and molecular evidence that Culex nigripalpus baculovirus is an unusual member of the family Baculoviridae. *Journal of General Virology* **82,** 283-97.

Motoh, H. (1985). *Proceedings of the first international conference on the culture of penaeik prawns/shrimp, Hoilo city, Philippines.*

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D., Eds. (1995). Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses. New York: Springer-Verlag.

Nadala, E. C., Jr., and Loh, P. C. (2000). Dot-blot nitrocellulose enzyme immunoassays for the detection of white-spot virus and yellow-head virus of penaeid shrimp. *Journal of Virological Methods* **84,** 175-179.

Nadala, E. C. B., Jr., and Loh, P. C. (1998). A comparative study of three different isolates of white spot virus. *Diseases of Aquatic Organisms* **33,** 231-234.

Nadala, E. C. B., Jr., Tapay, L. M., and Loh, P. C. (1998). Characterization of a non-occluded baculovirus-like agent pathogenic to penaeid shrimp. *Diseases of Aquatic Organisms* **33,** 221-229.

Naylor, R. L., Goldburg, R. J., Primavera, J. H., Kautsky, N., Beveridge, M. C., Clay, J., Folke, C., Lubchenco, J., Mooney, H., and Troell, M. (2000). Effect of aquaculture on world fish supplies. *Nature* **405,** 1017-1024.

Nunan, L. M., and Lightner, D. V. (1997). Development of a non radioactive gene probe by PCR for detection of white spot syndrome virus (WSSV). *Journal of Virological Methods* **63,** 193-201.

Nunan, L. M., Poulos, B. T., and Lightner, D. V. (1998). Reverse transcription polymerase chain reaction (RT-PCR) used for the detection of Taura Syndrome Virus (TSV) in experimentally infected shrimp. *Diseases of Aquatic Organisms* **34,** 87-91.

Nunan, L. M., Poulos, B. T., and Lightner, D. V. (2000). Use of polymerase chain reaction for the detection of infectious hypodermal and hematopoietic necrosis virus in penaeid shrimp. *Marine Biotechnology* **2,** 319-328.

O'Reilly, D. R. (1997). Auxiliary genes of baculoviruses. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 267-300. Plenum Press, New York.

Pantoja, C. R., and Lightner, D. V. (2000). A non-destructive method based on the polymerase chain reaction for detection of hepatopancreatic parvovirus (HPV) of penaeid shrimp. *Diseases of Aquatic Organisms* **39,** 177-182.

Pantoja, C. R., and Lightner, D. V. (2001). Detection of hepatopancreatic parvovirus (HPV) of penaeid shrimp by in situ hybridization at the electron microscope level. *Diseases of Aquatic Organisms* **44,** 87-96.

Pearson, M., Bjornson, R., Pearson, G., and Rohrmann, G. (1992). The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* **257,** 1382-1384.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in Enzymology*, 18363-18398.

Pearson, W. R. (2000). Flexible sequence similarity searching with the FASTA3 program package. *Methods in Molecular Biology*, 132185-132219.

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence analysis. *Proceedings of the National Academy of Sciences USA* **85,** 2444-2448.

Ramsey, I. K., Spibey, N., and Jarrett, O. (1998). The receptor binding site of feline leukemia virus surface glycoprotein is distinct from the site involved in virus neutralization. *Journal of Virology* **72,** 3268-77.

Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA sysnthesis. *Annual Review of Biochemistry* **57,** 349-374.

Reilander, H., Haase, W., and Maul, G. (1996). Functional expression of the Aequorea victoria green fluorescent protein in insect cells using the baculovirus expression system. *Biochem-Biophys-Res-Commun*

219, 14-20.

Robles Sikisaka, R., Garcia, D. K., Klimpel, K. R., and Dhar, A. K. (2001). Nucleotide sequence of 3 '-end of the genome of Taura syndrome virus of shrimp suggests that it is related to insect picornaviruses. *Archives of Virology* **146,** 941-952.

Rodriguez, J., Boulo, V., Mialhe, E., and Bachere, E. (1995). Characterisation of shrimp haemocytes and plasma components by monoclonal antibodies. *Journal of Cell Science* **108,** 1043-1050.

Rosenberg, I. M. (1996). "Protein analysis and purification: benchtop techniques." Birkhäuser, Boston.

Rosenberry, B. (1994). "World shrimp farming.". Shrimp News International, San Diego, California, USA.

Rosenberry, B. (1995). "World shrimp farming 1995.". Shrimp News International, San Diego, California, USA.

Rosenberry, B. (1996). "World shrimp farming 1996.". Shrimp News International, San Diego, California, USA.

Rosenberry, B. (1997). "World shrimp farming 1997.". Shrimp News International, San Diego, California, USA.

Rosenberry, B. (1999). "World shrimp farming 1999.". Shrimp News International, San Diego, California, USA.

Rosenberry, B. (2000). "World shrimp farming 2000.". Shrimp News International, San Diego, California, USA.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A laboratory Manual." 2 ed. Cold Spring Harbor Laboratory, New York.

Sano, T. T., Nishimura, K., Oguma, K., Momoyama, K., and Takeno, N. (1981). Baculovirus infection of kuruma shrimp, *Penaeus japonicus* in Japan. *Fish Pathology* **15,** 185-191.

Schmidt, P. P., Rova, U., Katterle, B., Thelander, L., and Graslund, A. (1998). Kinetic evidence that a radical transfer pathway in protein R2 of mouse ribonucleotide reductase is involved in generation of the tyrosyl free radical. *Journal of Biological Chemistry* **273,** 21463-21472.

Schmitt, J. F., and Stunnenberg, H. G. (1988). Sequence and transcriptional analysis of the vaccinia virus HindIII I fragment. *Journal of Virology* **62,** 1889-97.

Schneider-Schaulies, J. (2000). Cellular receptors for viruses: links to tropism and pathogenesis. *Journal of General Virology* **81,** 1413-1429.

Schofield, D. J., Glamann, J., Emerson, U., and Purchell, R. H. (2000). Identification by phage display and characterization of two neutralizing chimpanzee monoclonal antibodies to the hepatitis E virus capsid protein. *Journal of Virology* **74,** 5548-5555.

Shih, H. H., Wang, C. S., Tan, L. F., and Chen, S. N. (2001). Characterization and application of monoclonal antibodies against white spot syndrome virus. *Journal of Fish Diseases* **24,** 143-150.

Shike, H., Dhar, A. K., Burns, J. C., Shimizu, C., Jousset, F. X., Klimpel, K. R., and Bergoin, M. (2000). Infectious hypodermal and hematopoietic necrosis virus of shrimp is related to mosquito Brevidensoviruses. *Virology* **277,** 167-177.

Shinagawa, T., Dong, H. D., Xu, M., Maekawa, T., and Ishii, S. (2000). The *sno* gene, which encodes a component of the histone deacetylase complex, acts as a tumor suppressor in mice. *EMBO Journal* **19,** 2280-2291.

Sithigorngul, P., Chauychuwong, P., Sithigorngul, W., Longyant, S., Chaivisuthangkura, P., and Menasveta, P. (2000). Development of a monoclonal antibody specific to yellow head virus (YHV) from *Penaeus monodon*. *Diseases of Aquatic Organisms* **42,** 27-34.

Smith, G. E., and Summers, M. D. (1978). Analysis of baculovirus genomes with restriction endonucleases. *Virology* **89,** 517-527.

Smith, G. E., and Summers, M. D. (1981). Application of a novel radio-immunoassay to identify baculovirus structural proteins that share interspecies antigenic determinants. *Journal of Virology* **39,** 125-137.

Sonnhammer, E. L. L., von Heijne, G., and Krogh, A. (1998). *Sixth international conference on intelligent systems for molecular biology*.

Stiles, B., and Wood, H. A. (1983). A study of the glycoproteins of Autographa californica nuclear polyhedrosis virus (AcNPV). *Virology* **131,** 230-241.

Strauss, J. H., and Strauss, E. G. (2001). Virus evolution: how does an enveloped virus make a regular structure? *Cell* **105,** 5-8.

Stubbe, J. (1990). Ribonucleotide reductases: amazing and confusing. *J Biol Chem* **265,** 5329-32.

Sukhumsirichart, W., Wongteerasupaya, C., Boonsaeng, V., Panyim, S., Sriurairatana, S., Withyachumnarnkul, B., and Flegel, T. W. (1999). Characterization and PCR detection of hepatopancreatic parvovirus (HPV) from *Penaeus monodon* in Thailand. *Diseases of Aquatic Organisms* **38,** 1-10.

Sullivan, J. T., Mercer, A. A., Fleming, S. B., and Robinson, A. J. (1994). Identification and characterization of an orf virus homologue of the vaccinia virus gene encoding the major envelope antigen p37K. *Virology* **202**, 968-973.

Summers, M. D., and Smith, G. E. (1987). "A manual of methods for baculovirus vectors and insect cell culture procedures.". Texas agricultural experiment station bulletin, no 1555.

Sun, Y., Zhang, J., Kraeft, S. K., Auclair, D., Chang, M. S., Liu, Y., Sutherland, R., Salgia, R., Griffin, J. D., Ferland, L. H., and Chen, L. B. (1999). Molecular cloning and characterization of human trabeculin-alpha, a giant protein defining a new family of actin-binding proteins. *Journal of Biological Chemistry* **274**, 33522-33530.

Sunyach, C., Rollier, C., Robaczewska, M., Borel, C., Barraud, L., Kay, A., Trepo, C., Will, H., and Cova, L. (1999). Residues critical for duck hepatitis B virus neutralization are involved in host cell interaction. *Journal of Virology* **73**, 2569-2575.

PAUP: Phylogenetic analysis using parsimony. Version 3.1. Champaign, IL: Illinois Natural History Survey

Taga, T., and Kishimoto, T. (1997). Gp130 and the interleukin-6 family of cytokines. *Annual Review of Immunology*, 15797-15819.

Takahashi, Y., Itami, T., Maeda, M., Suzuki, N., Kasornchandra, J., Supamattaya, K., Khongpradit, R., Boonyaratpalin, S., Kondo, M., Kawai, K., Kusuda, R., Hirono, I., and Aoki, T. (1996). Polymerase chain reaction (PCR) amplification of bacilliform virus (RV-PJ) DNA in *Penaeus japonicus* Bate and systemic ectodermal and mesodermal baculovirus (SEMBV) DNA in *Penaeus monodon* Fabricius. *Journal of Fish Diseases* **19**, 399-403.

Tang, K. F. J., and Lightner, D. V. (1999). A yellow head virus gene probe: nucleotide sequence and application for in situ hybridization. *Diseases of Aquatic Organisms* **35**, 165-173.

Tang, K. F. J., and Lightner, D. V. (2000). Quantification of white spot syndrome virus DNA through a competitive polymerase chain reaction. *Aquaculture* **189**, 11-21.

Tapay, L. M., Nadala, E. C. B., and Loh, P. C. (1999). A polymerase chain reaction protocol for the detection of various geographical isolates of white spot virus. *Journal of Virological Methods* **82**, 39-43.

Tatman, J. D., Preston, V. G., Nicholson, P., Elliot, R. M., and Rixon, F. J. (1994). Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. *Journal of General Virology* **75**, 1101-1113.

Telford, E. A., Watson, M. S., McBride, K., and Davison, A. J. (1992). The DNA sequence of equine herpesvirus-1. *Virology* **189**, 304-316.

Teunissen, O. S. P., Faber, R., Booms, G. H. R., Latscha, T., and Boon, J. H. (1998). Influence of vaccination on vibriosis resistance of the giant black tiger shrimp Penaeus monodon (Fabricius). *Aquaculture* **164**, 359-366.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876-4882.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Tidona, C. A., and Darai, G. (1997). The complete DNA sequence of lymphocystis disease virus. *Virology* **230**, 207-16.

Tidona, C. A., Schnitzler, P., Kehm, R., and Darai, G. (1998). Is the major capsid protein of iridoviruses a suitable target for the study of viral evolution? *Virus Genes* **16**, 59-66.

Tsai, M. F., Lo, C. F., van Hulten, M. C. W., Tzeng, H. F., Chou, C. M., Huang, C. J., Wang, C. H., Lin, J. Y., Vlak, J. M., and Kou, G. H. (2000a). Transcriptional analysis of the ribonucleotide reductase genes of shrimp white spot syndrome virus. *Virology* **277**, 92-99.

Tsai, M. F., Yu, H. T., Tzeng, H. F., Leu, J. H., Chou, C. M., Huang, C. J., Wang, C. H., Lin, J. Y., Kou, G. H., and Lo, C. F. (2000b). Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase. *Virology* **277**, 100-110.

Uhlin, U., and Eklund, H. (1994). Structure of ribonucleotide reductase protein R1 [see comments]. *Nature* **370**, 533-9.

van der Rest, M., and Garrone, R. (1991). Collagen family of proteins. *FASEB Journal* **5**, 2814-2823.

van Hulten, M. C. W., Goldbach, R. W., and Vlak, J. M. (2000b). Three functionally diverged major structural proteins of white spot syndrome virus evolved by gene duplication. *Journal of General Virology* **81**, 2525-

2529.

van Hulten, M. C. W., Reijns, M., Vermeesch, A. M. G., Zandbergen, F., and Vlak, J. M. (2001c). Identification of VP19 and VP15 and glycosylation status of the white spot syndrome virus major structural proteins. *Journal of General Virology* **in press.**

van Hulten, M. C. W., Tsai, M. F., Schipper, C. A., Lo, C. F., Kou, G. H., and Vlak, J. M. (2000c). Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions. *Journal of General Virology* **81,** 307-316.

van Hulten, M. C. W., and Vlak, J. M. (2001a). Identification and phylogeny of a protein kinase gene of white spot syndrome virus. *Virus Genes* **22,** 201-207.

van Hulten, M. C. W., and Vlak, J. M. (2001b). Genetic evidence for a unique taxonomic position of white spot syndrome virus of shrimp: genus whispovirus. *Proceedings of the Fourth Symposium on Diseases in Asian Aquaculture (C. Lavilla-Pitogo et al., eds).* **in press.**

van Hulten, M. C. W., Westenberg, M., Goodall, S. D., and Vlak, J. M. (2000a). Identification of two major virion protein genes of white spot syndrome virus of shrimp. *Virology* **266,** 227-236.

van Hulten, M. C. W., Witteveldt, J., Peters, S., Kloosterboer, N., Tarchini, R., Fiers, M., Sandbrink, H., Klein Lankhorst, R., and Vlak, J. M. (2001a). The white spot syndrome virus DNA genome sequence. *Virology* **286,** 7-22.

van Hulten, M. C. W., Witteveldt, J., Snippe, M., and Vlak, J. M. (2001b). White spot syndrome virus envelope protein VP28 is involved in the systemic infection of shrimp. *Virology* **285,** 228-233.

van Regelmortel, 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., Eds. (2000). Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press.

van Strien, E. A., Faktor, O., Hu, Z. H., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1997). Baculoviruses contain a gene for the large subunit of ribonucleotide reductase. *Journal of General Virology* **78,** 2365-2377.

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J., and McCawley, P. (1977). The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). *In Vitro* **13,** 213-217.

Vickers, J. E., Webb, R., and Young, P. R. (2000). Monodon baculovirus from Australia: ultrastructural observations. *Diseases of Aquatic Organisms* **39,** 169-176.

Vlak, J. M., Bonami, J. R., Flegel, T. W., Kou, G. H., Lightner, D. V., Lo, C. F., Loh, P. C., and Walker, P. J. (2001). Family *Nimaviridae*. Submitted to the International Committee on Taxonomy of Viruses for approval.

Volkman, L. E., and Goldsmith, P. A. (1985). Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology* **143,** 185-195.

Volkman, L. E., Goldsmith, P. A., Hess, R. T., and Faulkner, P. (1984). Neutralization of budded *Autographa californica* NPV by a monoclonal antibody: identification of the target antigen. *Virology* **133,** 354-362.

Von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14,** 4683-4690.

Wang, C. H., Lo, C. F., Leu, J. H., Chou, C. M., Yeh, P. Y., Chou, H. Y., Tung, M. C., Chang, C. F., Su, M. S., and Kou, G. H. (1995). Purification and genomic analysis of baculovirus associated with white spot syndrome (WSBV) of *Penaeus monodon*. *Diseases of Aquatic Organisms* **23,** 239-242.

Wang, C. H., Yang, H. N., Tang, C. Y., Lu, C. H., Kou, G. H., and Lo, C. F. (2000a). Ultrastructure of white spot syndrome virus development in primary lymphoid organ cell cultures. *Diseases of Aquatic Organisms* **41,** 91-104.

Wang, Q., Nunan, L. M., and Lightner, D. V. (2000c). Identification of genomic variations among geographic isolates of white spot syndrome virus using restriction analysis and Southern blot hybridization. *Diseases of Aquatic Organisms* **43,** 175-181.

Wang, Q., Poulos, B. T., and Lightner, D. V. (2000b). Protein analysis of geographic isolates of shrimp white spot syndrome virus. *Archives of Virology* **145,** 263-274.

Wang, S. Y., Hong, C., and Lotz, J. M. (1996). Development of a PCR procedure for the detection of Baculovirus penaei in shrimp. *Diseases of Aquatic Organisms* **25,** 123-131.

Wang, Y. C., Lo, C. F., Chang, P. S., and Kou, G. H. (1998). Experimental infection of white spot baculovirus in some cultured and wild decapods in Taiwan. *Aquaculture* **164,** 221-231.

Wang, Y. G., Hassan, M. D., Shariff, M., Zamri, S. M., and Chen, X. (1999a). Histopathology and cytopathology of white spot syndrome virus (WSSV) in cultured *Penaeus monodon* from peninsular Malaysia with emphasis on pathogenesis and the mechanism of white spot formation. *Diseases of Aquatic Organisms* **39**, 1-11.

Wiemann, S., Kinzel, V., and Pyerin, W. (1992). Cloning of the C alpha catalytic subunit of the bovine cAMP-dependent protein kinase. *Biochemical and Biophysical Acta* **1171**, 93-96.

Williams, G. V., and Faulkner, P. (1997). Cytological changes and viral morphogenesis during Baculovirus infection. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 61-107. Plenum Press, New York.

Willoughby, K., Bennett, M., Williams, R. A., McCracken, C., and Gaskell, R. M. (1997). Sequences of the ribonucleotide reductase-encoding genes of feline herpesvirus 1 and molecular phylogenetic analysis. *Virus Genes* **15**, 203-218.

Wilson, M. E., and Price, K. H. (1988). Association of Autographa californica nuclear polyhedrosis virus (AcMNPV) with the nuclear matrix. *Virology* **167**, 233-241.

Witteveldt, J., van Hulten, M. C. W., and Vlak, J. M. (2001). Identification and phylogeny of a non-specific endonuclease gene of White Spot Syndrome Virus of shrimp. *Virus Genes* **23**, in press.

Wongteerasupaya, C., Vickers, J. E., Sriurairatana, S., Nash, G. L., Akarajamorn, A., Boonsaeng, V., Panyim, S., Tassanakajon, A., Withyachumnarnkul, B., and Flegel, T. W. (1995). A non-occluded, systemic baculovirus that occurs in cells of ectodermal and mesodermal origin and causes high mortality in the black tiger prawn *Penaeus monodon*. *Diseases of Aquatic Organisms* **21**, 69-77.

Wongteerasupaya, C., Wongwisansri, S., Boonsaeng, V., Panyim, S., Pratanpipat, P., Nash, G. L., Withyachumnarnkul, B., and Flegel, T. W. (1996). DNA fragment of *Penaeus monodon* baculovirus PmNOBII gives positive in situ hybridization with white-spot viral infections in six penaeid shrimp species. *Aquaculture* **143**, 23-32.

Yan, X., Olson, N. H., Van Etten, J. L., Bergoin, M., Rossmann, M. G., and Baker, T. S. (2000). Structure and assembly of large lipid-containing dsDNA viruses. *Nature Structural Biology* **7**, 101-103.

Yanez, R. J., Rodriguez, J. M., Nogal, M. L., Yuste, L., Enriquez, C., Rodriguez, J. F., and Vinuela, E. (1995). Analysis of the Complete Nucleotide Sequence of African Swine Fever Virus. *Virology* **208**, 249-278.

Yang, F., Wang, W., Chen, R. Z., and Xu, X. (1997). A simple and efficient method for purification of prawn baculovirus DNA. *Journal of Virological Methods* **67**, 1-4.

Zhang, X., Xu, L., and Xu, X. (2001). Detection of prawn white spot bacilliform virus by immunoassay with recombinant antigen. *Journal of Virological Methods* **92**, 193-197.

Summary

Since its first discovery in Taiwan in 1992, *White spot syndrome virus* (WSSV) has caused major economic damage to shrimp culture. The virus has spread rapidly through Asia and reached the Western Hemisphere in 1995 (Texas), where it continued its devastating effect further into Central- and South-America. In cultured shrimp WSSV infection can reach a cumulative mortality of up to 100% within 3 to 10 days. One of the clinical signs of WSSV is the appearance of white spots in the exoskeleton of infected shrimp, hence its name. WSSV has a remarkably broad host range, it not only infects all known shrimp species, but also many other marine and freshwater crustaceans, including crab and crayfish. Therefore, WSSV can be considered a major threat not only to shrimp, but also to other crustaceans around the world.

The WSSV virion is a large enveloped particle of about 275 nm in length and 120 nm in width with an ellipsoid to bacilliform shape and a tail-like extension on one end. The nucleocapsid is rod-shaped with a striated appearance and has a size of about 300 nm x 70 nm. Its virion morphology, nuclear localization and morphogenesis are reminiscent of baculoviruses in insects. Therefore, WSSV was originally thought to be a member of the *Baculoviridae*.

At the onset of the research presented in this thesis, only limited molecular information was available for WSSV, hampering its definitive classification as well as profound studies of the viral infection mechanism. As the first step towards unraveling the molecular biology of WSSV, terminal sequencing was performed on constructed genomic libraries of its genome. This led to the identification of genes for the large (*rr1*) and small (*rr2*) subunit of ribonucleotide reductase, which were present on a 12.3 kb genomic fragment (Chapter 2). Phylogenetic analyses using the RR1 and RR2 proteins indicated that WSSV belongs to the eukaryotic branch of an unrooted parsimonious tree and further showed that WSSV and baculoviruses do not share a recent common ancestor.

Subsequently two protein kinase (*pk*) genes were located on the WSSV genome, showing low homology to other viral and eukaryotic *pk* genes (Chapter 3). The presence of conserved domains, suggested that these PKs are serine/threonine protein kinases. A considerable number of large DNA viruses contains one or more *pk* genes and these were used to construct an unrooted parsimonious phylogenetic tree. This tree indicated that the two WSSV *pk* genes originated most likely by gene duplication. Furthermore, the tree provided strong evidence that WSSV takes a unique position among large DNA virus families and was clearly separated from the *Baculoviridae*.

As a further step to analyze WSSV in more detail, its major virion proteins were analyzed. In general, structural proteins are well conserved within virus families and therefore represent good phylogenetic markers. Furthermore, knowledge on these proteins can lead to better insight in the viral infection mechanism. Five major proteins of 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24), 19 kDa (VP19), and 15 kDa (VP15) in size were identified (Chapter 4, 5 and 6). VP26, VP24 and VP15 were found associated with the nucleocapsid, while VP28 and VP19 were found associated with the viral envelope. Partial amino acid sequencing was performed on these proteins to identify their respective genes in the WSSV genome.

The first structural genes to be identified on the WSSV genome were those coding for VP28 and VP26, which are most abundant in the virion (Chapter 4). The correct identification of these genes was confirmed by heterologous expression in the baculovirus insect cell expression system and detection by Western analysis using a polyclonal antiserum against total WSSV virions. Subsequently, VP24 was characterized (Chapter 5) and computer-assisted analysis revealed a striking amino acid and nucleotide similarity between VP24, VP26 and VP28 and their genes, respectively. This strongly suggests that these genes have evolved by gene duplication and subsequently diverged into proteins with different functions within the virion, i.e. envelope and nucleocapsid. All three proteins contained a putative transmembrane domain at their N-terminus and multiple putative N- and O-glycosylation sites. The putative transmembrane sequence in VP28 may anchor this protein in the viral envelope. The hydrophobic sequences may also be involved in the interaction of the structural proteins to form homo- or heteromultimers.

In Chapter 6 the identification of the structural proteins VP19 and VP15 is described. The VP19 polypeptide contained two putative transmembrane domains, which may anchor this protein in the WSSV envelope. Also this protein contained multiple putative glycosylation sites. N-terminal sequencing on VP15 showed that this protein was expressed from the second translational start codon within its gene and that the first methionine was cleaved off. As VP15 is a very basic protein and resembles histone proteins, it is tempting to assume that this protein functions as a DNA binding protein within the viral nucleocapsid. None of the identified structural proteins showed homology to viral proteins in other viruses, which further supports the proposition that WSSV has a unique taxonomical position.

As the theoretical sizes determined of the various structural proteins, as derived from their genes, were smaller than the apparent sizes on SDS-PAGE, it was suspected that some of these proteins were glycosylated (Chapter 6). All five identified proteins were expressed in insect cells using baculovirus vectors, resulting in expression products of similar sizes as in the WSSV virion. The glycosylation status of the proteins was analyzed and this indicated that none of the five major structural proteins was glycosylated. This is a very unusual feature of WSSV, as enveloped viruses of vertebrates and invertebrates contain glycoproteins in their viral envelopes, which often play important roles in the interaction between virus and host, such as attachment to receptors and fusion with cell membranes.

To study the mode of entry and systemic infection of WSSV in the black tiger shrimp, *Penaeus monodon,* the role of the major envelope protein VP28 in the systemic infection in shrimp was studied (Chapter 7). An *in vivo* neutralization assay was performed in *P.*

monodon, using a specific polyclonal antibody generated against VP28. The VP28 antiserum was able to neutralize WSSV infection of *P. monodon* in a concentration-dependent manner upon intramuscular injection. This result suggests that VP28 is located on the surface of the virus particle and is likely to play a key role in the initial steps of the systemic infection of shrimp.

To analyze the genome structure and composition, the entire sequence of the double-stranded, circular DNA genome of WSSV was determined (Chapter 8). On the 292,967 nucleotide genome 184 open reading frames (ORFs) of 50 amino acids or larger were identified. Only 6% of the WSSV ORFs had putative homologues in databases, mainly representing genes encoding enzymes for nucleotide metabolism, DNA replication and protein modification. The remaining ORFs were mostly unassigned except for the five encoding the structural proteins. Unique features of the WSSV genome are the presence of an extremely long ORF of 18,234 nucleotides with unknown function, a collagen-like ORF, and nine regions, dispersed along the genome, each containing a variable number of 250-bp tandem repeats. When this WSSV genome sequence was compared to that of a second isolate from a different geographic location, the isolates were found to be remarkably similar (over 99% homology) (Chapter 9). The major difference was a 12 kbp deletion in the WSSV isolate, described here, which is apparently dispensable for virus infectivity.

To complete the taxonomic research on WSSV, its DNA polymerase gene was used in a phylogenetic study (Chapter 8), confirming the results of the phylogeny performed on PK. To obtain a consensus tree, combined gene phylogeny analysis was performed using the *rr1*, *rr2*, *pk* and *pol* genes, which were also present in other large dsDNA virus families (Chapter 9). Based on this consensus tree no relationship was revealed for WSSV with any of the established families of large DNA viruses. The collective information on WSSV and the phylogenetic analysis suggest that WSSV differs profoundly from all presently known viruses and is a representative of a new virus family, with the proposed name 'Nimaviridae' (nima = thread).

The present knowledge on the WSSV genome and its major structural proteins, has created a good starting point for further studies on the replication strategy and infection mechanism of the virus, and last but not least, will open the way for the design of novel strategies to control this devastating pathogen.

Samenvatting

Het Witte-vlekken-syndroom-virus (Engelse naam: White spot syndrome virus, afkorting: WSSV) is een belangrijk garnalenvirus, dat voor het eerst in 1992 is aangetroffen nabij Taiwan. Het virus verspreidde zich sindsdien snel door Zuidoost-Azië en in 1995 werd het voor het eerst in een garnalenkwekerij op het westelijk halfrond, in Texas, aangetroffen. Vervolgens verspreidde het virus zich over Centraal- en Zuid-Amerika en richtte ook hier veel schade aan. In gekweekte garnalen kan WSSV binnen 3 tot 10 dagen een cumulatieve sterfte van 100% veroorzaken. Eén van de symptomen van een geïnfecteerde garnaal is het verschijnen van witte stippen onder het exoskelet, waaraan het virus zijn naam dankt. WSSV heeft een opvallend groot aantal gastheren, aangezien het behalve garnalen ook andere zoet- en zoutwaterkreeftachtigen, zoals krabben en kreeften, kan infecteren. Daarom is het virus niet alleen een bedreiging voor de commerciële garnalenteelt, maar wereldwijd ook voor andere kreeftachtigen.

Het WSSV-virion bestaat uit een nucleocapside, omgeven door een lipidemembraan ("envelop") en meet ongeveer 275 nm bij 120 nm. Het virion heeft een ellips-staafvormig uiterlijk en bezit een staartje aan één van de uiteinden. Het nucleocapside is staafvormig, meet 300 nm bij 70 nm, en bevat een dubbelstrengs DNA-genoom. Aangezien het virus in de celkern gevormd wordt en de vorm van het virion enigszins overeenkomt met die van baculovirussen, werd oorspronkelijk gedacht dat WSSV tot de familie der *Baculoviridae* behoorde.

Bij aanvang van het in dit proefschrift beschreven onderzoek was de moleculair-biologische kennis over WSSV zeer beperkt. Hierdoor konden geen gedetailleerde studies worden gedaan naar het infectiemechanisme van het virus en was een definitieve taxonomische classificering van het virus via moleculaire technieken onmogelijk. Om inzicht in de moleculaire biologie van WSSV te verkrijgen, werden allereerst de nucleotidenvolgorden aan de uiteinden van gekloneerde fragmenten van het virale genoom bepaald. Dit leidde tot de identificatie van genen, coderend voor de kleine (RR2) en grote (RR1) subeenheid van ribonucleotide reductase op een genomisch fragment van 12.3 kilobasenparen (kb) (Hoofdstuk 2). Fylogenetische analyse op basis van deze genen toonde aan dat de RR-eiwitten van WSSV het meest verwant zijn met de RR1- en RR2-eiwitten van eukaryote organismen. Er bleek geen directe verwantschap te zijn tussen WSSV en de baculovirussen.

Verder werden twee genen, coderend voor het enzym protein kinase (*pk*), gelokaliseerd op het WSSV-genoom. Beide hadden een lage homologie met andere virale en eukaryote *pk*-genen (Hoofdstuk 3). Op grond van geconserveerde domeinen in het eiwit bleken beide genen waarschijnlijk te behoren tot de klasse van de serine/threonine *pk*-genen. Ook andere grote virale DNA-genomen bevatten één of meerdere *pk*-genen en deze werden gebruikt voor een fylogenetische analyse. Hieruit kwam naar voren dat de twee *pk*-genen op het WSSV-genoom waarschijnlijk door genduplicatie zijn ontstaan. Tevens kon op basis van deze *pk*-genen opnieuw vastgesteld worden dat WSSV een unieke positie inneemt ten aanzien van andere grote DNA-virussen, waaronder de *Baculoviridae*.

Naast analyse van het virale genoom werden de eiwitten in het virion geanalyseerd. Aangezien structurele eiwitten vaak sterk geconserveerd zijn binnen een virusfamilie, was de verwachting dat deze analyse ook verder inzicht zou geven in de taxonomische positie van WSSV. Vijf eiwitten van 28 kDa (VP28), 26 kDa (VP26), 24 kDa (VP24), 19 kDa (VP19) en 15 kDa (VP15) kwamen het meest prominent voor in het virion (Hoofdstukken 4, 5 en 6), waarbij VP26, VP24 en VP15 in de nucleocapside, en VP28 en VP19 in de virale envelop konden worden gelokaliseerd. Met behulp van gedeeltelijk bepaalde aminozuurvolgorden werden de genen, coderend voor deze eiwitten, op het WSSV-genoom gevonden.

De *vp28*- en *vp26*-genen werden als eerste op het WSSV-genoom geïdentificeerd (Hoofdstuk 4). Deze genen werden vervolgens in insectencellen tot expressie gebracht m.b.v. baculovirusvectoren. D.m.v. immunologische detectie kon worden bevestigd dat deze inderdaad codeerden voor de respectievelijke virioneiwitten. Het volgende eiwit wat onderzocht werd, was VP24 (Hoofdstuk 5). Een opmerkelijke sequentie-homologie werd gevonden tussen de eiwitten VP28, VP26 en VP24 en ook tussen hun genen, waaruit bleek dat deze naar alle waarschijnlijkheid ontstaan zijn door genduplicatie. Aan de amino-terminus van deze drie eiwitten bevond zich een mogelijk transmembraandomein en verder werden vele mogelijke glycosyleringsplaatsen aangetroffen in deze eiwitten. In VP28 zou dit transmembraandomein kunnen zorgen voor membraanverankering en in de capside eiwitten VP26 en VP24 voor de vorming van homo- en heteromultimeren.

In Hoofstuk 6 wordt de identificatie van de structurele eiwitten VP19 en VP15 beschreven. In het VP19-eiwit werden twee potentiële transmembraandomeinen gevonden, welke dit eiwit in de envelop zouden kunnen verankeren en vele mogelijke N- en O-glycosyleringsplaatsen. De aminozuurvolgorde van VP15 vertoonde overeenkomst met die van histoneiwitten, hetgeen erop wijst dat VP15 ook als een DNA-bindend eiwit zou kunnen functioneren in de virale nucleocapside. Geen van de vijf eiwitten uit het WSSV-virion vertoonde homologie met eiwitten van andere virussen, hetgeen de unieke taxonomische positie van WSSV nogmaals onderstreept.

De theoretische groottes die gevonden werden op basis van de aminozuursamenstelling van deze structurele eiwitten zijn aanzienlijk kleiner dan de molecuulmassa in eiwitgels. Een mogelijke verklaring hiervoor zou kunnen zijn dat deze eiwitten geglycosyleerd zijn. Om dit nader te onderzoeken werden de vijf structurele WSSV-eiwitten tot expressie gebracht in het baculovirus-insektencel-expressiesysteem en dit resulteerde in eiwitten van dezelfde grootte als in het virusdeeltje (Hoofdstuk 6). Vervolgens werd aangetoond dat geen van de eiwitten geglycosyleerd is. Dit is een bijzondere waarneming aangezien de meeste andere membraanvirussen geglycosyleerde envelopeiwitten bezitten.

De rol van VP28 in het infectieproces en in de verspreiding van WSSV in de zwarte tijgergarnaal, *Penaeus monodon*, werd nader bestudeerd (Hoofdstuk 7). Daartoe werd een *in vivo* neutralisatietest uitgevoerd met een polyclonaal antiserum, opgewekt tegen dit eiwit. Het VP28-antiserum was in staat om de virale infectie concentratieafhankelijk te neutraliseren wanneer virusdeeltjes in de garnaal werden ingespoten. Dit suggereert dat VP28 zich aan het oppervlak van het virion bevindt en een essentiële rol speelt in de initiële stappen van systemische infectie in garnalen.

Om de genetische organisatie van het WSSV-genoom in meer detail te bestuderen werd de volledige nucleotidenvolgorde ervan bepaald (Hoofdstuk 8). Het WSSV-genoom bleek 292.967 basenparen (bp) groot te zijn en 184 potentiële genen, ook wel 'open reading frames' (ORFs) genoemd, te bevatten. Voor slechts 6% van deze ORFs kon een mogelijke homoloog worden geïdentificeerd in de GenBank en deze codeerden voornamelijk voor enzymen, betrokken in het nucleotide-metabolisme, DNA-replicatie of eiwitmodificatie. Voor de meeste overige ORFs kon geen mogelijke functie worden gevonden. Unieke eigenschappen van het WSSV-genoom bleken te zijn: de aanwezigheid van een wel zeer groot ORF (18.234 bp) met nog onbekende functie, een collageen-ORF en negen gebieden, verspreid in het genoom, bestaande uit aaneengesloten, repeterende 250 bp-sequenties. Na publicatie van de volledige WSSV-genoomsequentie, kwam een tweede sequentie van WSSV beschikbaar. Beide WSSV-genomen bleken grotendeels identiek te zijn (meer dan 99% nucleotiden-overeenkomst). Het grootste verschil was echter de al dan niet aanwezigheid van een 12 kb-fragment, dat niet nodig lijkt te zijn voor virale infectiositeit (Hoofdstuk 9).

Om het taxonomische onderzoek voor WSSV af te sluiten, werd het virale DNA-polymerase (POL) gebruikt voor fylogenetische analyse. Op basis van dit gen bleek WSSV ook geen verwantschap te hebben met andere virussen. Tenslotte werd een gecombineerde fylogenetische analyse uitgevoerd met de gezamenlijke *rr1*-, *rr2*-, *pk*- en *pol*-genen, aangezien deze voorkomen in alle grote DNA-virusfamilies. Dit resulteerde in een betrouwbare consensusstamboom, waaruit duidelijk naar voren komt dat WSSV geen verwantschap heeft met enige gevestigde virusfamilie en derhalve in een nieuwe virusfamilie geplaatst dient te worden. Hiervoor is de familienaam 'Nimaviridae' (nima =draad) voorgesteld.

De huidige kennis van zowel het genoom als de belangrijkste structurele eiwitten van WSSV vormen een goede uitgangspositie voor verder onderzoek aan het replicatie- en infectiemechanisme van het virus en opent de weg voor het ontwerpen van nieuwe strategieën om dit voor de garnalenkweek uiterst schadelijke virus te bestrijden.

Nawoord

Eindelijk is het dan zover; het nawoord kan geschreven worden. De afgelopen jaren in nabeschouwing nemend, kan ik concluderen dat ik met veel plezier op virologie heb gewerkt en dat ik de laatste maanden met stress voor de afronding van het boekje al ongeveer vergeten ben. Hier wil ik graag de mensen bedanken die een directe of indirecte bijdrage aan het tot stand komen van dit boekje hebben geleverd. Als eerste wil ik graag mijn begeleider Just Vlak bedanken. Just, je bent een geweldige begeleider geweest en je hebt me de afgelopen vier jaar op het rechte spoor gehouden, zodat de al te wilde experimenten vervangen werden door het schrijven van artikelen. Rob Goldbach en Douwe Zuidema wil ik bedanken voor de goede discussies en vele kritische opmerkingen tijdens de werkbesprekingen of daarbuiten. Rob, nog een extra woord van dank voor je goede begeleiding tijdens de afronding van mijn proefschrift.

Verder wil ik alle virologen bedanken die door hun aanwezigheid virologie tot een zeer gezellige werkplek maken of hebben gemaakt. De wekelijkse kroegsessie met de collegae heeft zeker ook een bijdrage geleverd tot de goede werksfeer, want hoe kun je beter werkstress kwijtraken dan met mensen die precies weten waar je het over hebt. Ingeborg, Gorben, Marcel, Theo, Danny, Sandra, Xinwen, en Monique, jullie wil ik hier even persoonlijk bedanken voor de gezellige uren in de zaaier. Hier wil ik graag ook Jacco, Elisabeth en René, mijn baculovirus collegae in het begin van mijn promotietijd bedanken, voor de aangename RBBS uurtjes (toen was nog niet bekend, dat WSSV niet tot de *Baculoviridae* behoorde). Wilfred en Marcel, jullie waren gezellige kamergenoten, bedankt. Els en Magda, bedankt voor jullie adviezen op het lab, Thea voor de snoepjes, Wout voor je gezelligheid en vele hulp en Danny voor zijn vele tips. Verder wil ik Joop bedanken voor zijn gezelligheid en goede gesprekken bij (en over) de elektronenmicroscoop; ik had graag gewild dat je dit nog mee had kunnen maken.

Ik wil graag alle mensen betrokken in het project met Intervet hier ook bedanken. Eric Rijke en Sjo Koumans voor al hun goede ideeën, Margriet te Winkel voor de gezellige uurtjes op het lab, Karin van de Braak voor de leuke wetenschappelijke en niet wetenschappelijke discussies en Bert-Jan Roosendaal voor het helpen bij het opzetten van de experimenten in de beginfase van mijn onderzoek. Ik heb veel plezier beleefd aan deze samenwerking en er ook veel van opgestoken. Zo ook van Ronald Keus en Claudia Ogilvie bij het schrijven van de patenten.

Dit proefschrift zou nooit in zijn huidige vorm tot stand zijn gekomen zonder de inspanning van mijn studenten, Christel Schipper, Marcel Westenberg, Dimitri Diavatopoulos, Joris Dhont, Marjolein Snippe, Fokko Zandbergen, Martin Reijns en Nico Kloosterboer. Jeroen Witteveldt wil ik bij deze ook bedanken voor zijn bijdrage aan het laatste deel van mijn promotie onderzoek.

Verder wil ik mijn huidige AIO's Hendrik Marks, Jeroen Witteveldt en Cesar Escobedo bedanken voor hun ondersteuning van mijn studenten, die ik de laatste maanden een beetje heb verwaarloosd en ook mijn studenten Melanie Mennens, Marjolein van Esschoten en Ruben Martherus voor hun grote zelfstandigheid. Toch zou het lab niet zo goed lopen als Angela Vermeesch er niet zou zijn. Angela, bedankt voor al je steun in deze voor mij drukke tijd.

Een promotie komt echter niet tot een goed einde als er niet de mensen zijn, die de zaken goed weten te relativeren. Zoals je ziet Eyke, als er één schaap over de dam is.... Suzanne, zet 'm op, jij bent de volgende! Verder wil ik mijn ouders bedanken voor hun steun en ik hoop dat dit boekje ze een beter idee geeft wat ik de laatste jaren heb gedaan...... Verder wil ik mijn paranimfen, Connie van Hulten en Danny Duijsings bij deze al bedanken.

Gorben, zonder jou zouden de laatste jaren met de drukke tijden een stuk minder leuk zijn geweest. Heel erg bedankt voor je begrip en geduld en reken maar, als het voor jou zo ver is, kook ik iedere dag!

Curriculum vitae

Maria Cornelia Wilhelmina van Hulten werd geboren op 23 september 1973 in Drunen. In 1991 heeft zij het gymnasium-β diploma aan het Dr. Mollercollege in Waalwijk behaald. In datzelfde jaar werd aangevangen met de studie Bioprocestechnologie aan de Wageningen Universiteit (WU). In 1996 behaalde zij het ingenieursdiploma. Tijdens de doctoraalfase heeft zij afstudeeronderzoek verricht bij achtereenvolgens het laboratorium voor Virologie aan de WU (Dr. D. Zuidema), the Commonwealth Scientific and Industrial Research Organisation, Division of Entomology in Australië (Dr. K.H.J. Gordon and Dr. T.N. Hanzlik) en leerstoelgroep Celbiologie en Immunologie aan de WU (Dr. R.J.M. Stet).

Vervolgens heeft zij 6 maanden gewerkt aan een haalbaarheidsstudie voor onderzoek aan garnalenvirussen bij het laboratorium van Virologie en aansluitend hierop was zij van november 1996 tot december 2000 werkzaam als Assistent in Opleiding bij het laboratorium voor Virologie van de WU onder begeleiding van Prof.dr. J.M. Vlak en Prof.dr. R.W. Goldbach. Van het daar uitgevoerde onderzoek, dat gefinancierd werd door Intervet International BV, staan de resultaten beschreven in dit proefschrift.

Vanaf december 2000 is zij werkzaam als Postdoc bij het laboratorium voor Virologie van de WU, waar zij de coördinatie van het Whispovac project op zich heeft genomen.

Account

van Hulten, M. C. W., Westenberg, M., Goodall, S. D., and Vlak, J. M. (2000). Identification of two major virion protein genes of white spot syndrome virus of shrimp. *Virology* **266,** 227-236.

van Hulten, M. C. W., Tsai, M. F., Schipper, C. A., Lo, C. F., Kou, G. H., and Vlak, J. M. (2000). Analysis of a genomic segment of white spot syndrome virus of shrimp containing ribonucleotide reductase genes and repeat regions. *Journal of General Virology* **81,** 307-316.

van Hulten, M. C. W., Goldbach, R. W., and Vlak, J. M. (2000). Three functionally diverged major structural proteins of white spot syndrome virus evolved by gene duplication. *Journal of General Virology* **81,** 2525-2529.

Tsai, M. F., Lo, C. F., van Hulten, M. C. W., Tzeng, H. F., Chou, C. M., Huang, C. J., Wang, C. H., Lin, J. Y., Vlak, J. M., and Kou, G. H. (2000). Transcriptional analysis of the ribonucleotide reductase genes of shrimp white spot syndrome virus. *Virology* **277,** 92-99.

van Hulten, M. C. W., and Vlak, J. M. (2001). Identification and phylogeny of a protein kinase gene of white spot syndrome virus. *Virus Genes* **22,** 201-207.

van Hulten, M. C. W., Witteveldt, J., Snippe, M., and Vlak, J. M. (2001). White spot syndrome virus envelope protein VP28 is involved in the systemic infection of shrimp. *Virology* **285,** 228-233.

van Hulten, M. C. W., Witteveldt, J., Peters, S., Kloosterboer, N., Tarchini, R., Fiers, M., Sandbrink, H., Klein Lankhorst, R., and Vlak, J. M. (2001). The white spot syndrome virus DNA genome sequence. *Virology* **286,** 7-22.

van Hulten, M. C. W., Reijns, M., Vermeesch, A. M. G., Zandbergen, F., and Vlak, J. M. (2001). Identification of VP19 and VP15 and glycosylation status of the white spot syndrome virus major structural proteins. *Journal of General Virology* **in press.**

Witteveldt, J., van Hulten, M. C. W., and Vlak, J. M. (2001). Identification and phylogeny of a non-specific endonuclease gene of white spot syndrome virus of shrimp. *Virus Genes* **23,** 331-337.

van Hulten, M. C. W., and Vlak, J. M. (2001). Genetic evidence for a unique taxonomic position of white spot syndrome virus of shrimp: genus whispovirus. *Proceedings of the Fourth Symposium on Diseases in Asian Aquaculture (C. Lavilla-Pitogo et al., eds).* **in press.**