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

Mariëlle C. W. van Hulten, Martin Reijns, Angela M. G. Vermeesch, Fokko Zandbergen and Just M. Vlak

Laboratory of Virology, Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

White spot syndrome virus (WSSV) infects penaeid shrimp and other crustaceans. The WSSV virion consists of an enveloped rod-shaped nucleocapsid enclosing a large circular double-stranded DNA genome of 293 kbp. 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 is 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.
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.

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 previously (Van Hulten et al., 2000a). Crayfish (*Procambarus clarkii*) were injected intramuscularly with a lethal dose of WSSV using a 20 gauge needle. After 1 week the haemolymph was withdrawn and virus was purified as described previously (Van Hulten et al., 2000a).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, 1978; King & Possee, 1992). To inhibit N-glycosylation, 20 µg/ml tunicamycin (Sigma) was added to the medium at the start of the infection. Recombinant *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV)–GFP, expressing green fluorescent protein (GFP) from the p10 promoter, was used as a control virus (Van Hulten et al., 2000a). Recombinant virus AcMNPV-WSSVvp28 was used for the expression of WSSV VP28 and expressed GFP from the p10 promoter and VP28 from the polyhedrin 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 pFastBac-DUAL vector downstream of the p10 promoter. The resulting plasmid was named pFastBac-D/GFP, and maintained region 1 for insertion of a foreign (WSSV) gene downstream of the polyhedrin promoter.

  PCR was performed on the WSSV genome, introducing a BamHI site at the 3' end of the ORFs encoding VP26, VP24, VP19 and VP15, an EcoRI site at the 5' ends of vp26 and vp24 and a NolI 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 BamHI and EcoRI (vp26 and vp24) or with BamHI and NolI (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 AcMNPVs expressing 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 analysed in 15% SDS–PAGE gels as described in Laemmli (1970). Proteins were visualized using Coomassie brilliant blue staining and protein molecular mass determined by comparison with Amersham Pharmacia low molecular mass protein marker. Semi-dry blotting was performed onto a PVDF membrane (Bio-Rad) using a CAPS buffer (10 mM CAPS in 10% methanol), or onto an Immobilon-P (Millipore) membrane 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.).

  Immobilon-P membranes were blocked in 2% low-fat milk powder (Camping, 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 antiseraum, 200 µg purified WSSV virions was injected into a rabbit. The rabbit was boosted with 500 µg WSSV virions after 6 weeks and the antisera was prepared 2 weeks thereafter.

  Immunodetection on AcMNPV GP64 was performed in a similar way using monoclonal antibody Acv1 (Hohnmann & 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 analysed by the periodic acid–Schiff method (PAS), a non-discriminatory chemical staining of all carbohydrates as described by Rosenberg (1996). After separation by SDS–PAGE, the proteins were transferred to a PVDF membrane. After washing in water the blot was incubated in PAS (1% periodic acid at 3% acetic acid) for 15 min. The blot was then washed three times in water and subsequently transferred to Schiff’s reagent (fuchsian–sulfite reagent) for 15 min. After incubation in 0.5% (w/v) sodium bisulphite, the blot was rinsed with water and left to air-dry.

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. 1a). In the nucleocapsid fraction three major bands were observed: 26 kDa (VP26), 24 kDa (VP24) and 15 kDa (VP15). The 28 kDa (VP28) and 19 kDa (VP19) proteins were present in the NP40 soluble fraction and are therefore located in the envelope or tegument of the virion (Fig. 1b). As compared to previous publications (Van Hulten et al., 2000a, b) the nucleocapsid fraction gave an additional major band (VP15); this was not observed previously as the present gel gives better resolution of proteins of about 18 kDa and smaller.

The proteins from the SDS–PAGE gel were transferred to a PVDF membrane by semi-dry blottting, and the VP15 and VP19 bands were excised from the filter and sequenced. Twenty-nine amino acids were obtained from the N
WSSV structural proteins

Fig. 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.

terminus of VP15 (bold type in Fig. 2b): VARSSKTKSRRGSKKRSTTAGISKRRSP. 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 type in Fig. 3b): IVLISI(G/V)ILVLAVMNVxMGPKKDS, 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 163996 to 164238 on the WSSV genome. It partially overlaps with a small putative ORF of 11 kDa, ORF110, which is in the reverse orientation on the WSSV genome (Fig. 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 valve at position 21, which is the amino acid following the methionine encoded by the second ATG (Fig. 2b). The first ATG of this ORF is apparently not used. This ATG is in an unfavourable Kozak (1989) context (TTCATGA) for efficient translation initiation, whereas the second ATG is in a favourable Kozak context (AAAAATGG). Furthermore, the TATA-box, 87 nucleotides upstream of the second ATG, has a position relative to a putative transcriptional start, which also favours 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 valve at the N terminus of the protein is 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 valve, removal of the N-terminal methionine is favoured (Flinta et al., 1986). A poly(A) signal is present 62 nucleotides downstream of the translation stop codon (Fig. 2b).

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·[ST]) or for O-glycosylation (Hansen et al., 1998) could be identified. Eight possible phosphorylation sites ([ST]-X-[DE] or [ST]-X-[RK]) were found within VP15. No cysteine residues are present, and therefore VP15 is unable to form any disulfide bridges. No hydrophobic regions are present in this protein (Fig. 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. arginine) and serine, characteristic of these proteins. This further substantiates the putative function of VP15 as a DNA-binding protein.

**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 290363–289998 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 separated by only two other WSSV ORFs (ORF183 and ORF184), of which the (putative) functions are unknown (Fig. 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. 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. 4b). In both hydrophobic regions putative transmembrane domains formed by amino acids 37 through 59 and amino acids 95 through 117, respectively, were predicted (Sonhammer 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 proteins, 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 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 GFP from the p10 promoter and each of the WSSV structural proteins from the baculovirus polyhedrin promoter.

Fig. 4. Hydrophilicity plots of VP15 (A) and VP19 (B). The amino acid number is on the abscissa, the hydrophilicity value on the ordinate.

Sf21 insect cells were infected with AcMNPV-GFP, AcMNPV-WSSV vp28 and AcMNPV-WSSV vp26, AcMNPV-WSSV vp24, AcMNPV-WSSV vp19, and AcMNPV-WSSV vp15 at an m.o.i. of 5 and harvested at 72 h post-infection. Extracts of infected Sf21 cells were analysed in a 15% SDS–PAGE gel (Fig. 5a). Clear expression products can be observed for VP28, VP26, VP24 and VP15 (Fig. 5a, lanes 6, 8, 10 and 14, respectively), which have the same electrophoretic mobility as their authentic counterparts in the WSSV virion (Fig. 5a, lane 16). No clear expression product could be observed for VP19 (Fig. 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. 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 mass 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. 5a, lanes 7, 9, 11, 13, 15) as the expression products in the absence of
tunicamycin (Fig. 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. 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 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. Fig. 6(a) 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. 1b). When observed by 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. (2000) in the WSSV virion and N-terminal sequencing revealed the sequence VARGGKTGRRG, which is 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. 2a) (Van Hulten et al., 2001a). No TATA box or poly(A) signal was found for this ORF and the initiation codon is in an unfavourable 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. 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 shrimp by WSSV (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 infection by WSSV.

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 mass 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 to those present in the purified virions (Fig. 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 not clear whether 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 mass 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 for one or more proteins in viral envelopes or nucleocapsids to be 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, a feature that 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 only (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.

**Note added in proof.** Zhang et al. (Virus Research, 2001, 79, 137–144) recently identified an ORF in the WSSV genome encoding a basic protein (p6.8) with DNA-binding properties in in vitro gel mobility shift assays. This ORF is identical to WSSV ORF109 encoding the WSSV structural protein VP15 described in our present paper.

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

**References**


WSSV structural proteins


Received 11 June 2001; Accepted 19 September 2001