Transcriptional analysis of *Penaeus stylirostris* densovirus genes

Arun K. Dhar a,b,⁎,1, Krista N. Kaizer a,b,1, Dilip K. Lakshman c

a Viracine Therapeutics Corporation, 7155-H Columbia Gateway Dr., Columbia, MD 21046, USA
b Advanced BioNutrition Corporation, Columbia, MD 21046, USA
c USDA-ARS, Sustainable Agricultural Systems Laboratory, Beltsville, MD 20705, USA

A R T I C L E   I N F O

Article history:
Received 22 December 2009
Returned to author for revision 3 February 2010
Accepted 1 March 2010
Available online 8 April 2010

Keywords:
*Penaeus stylirostris* densovirus
PstDNV
IHHNV
Infectious hypodermal and hematopoietic necrosis virus
Shrimp

A B S T R A C T

*Penaeus stylirostris* densovirus (PstDNV) genome contains three open reading frames (ORFs), left, middle, and right, which encode a non-structural (NS) protein, an unknown protein, and a capsid protein (CP), respectively. Transcription mapping revealed that P2, P11 and P61 promoters transcribe the left, middle and right ORFs. NS transcript uses the D1/A1 donor/acceptor sites for splicing and has two alternate transcription termination sites (TTS) that were different from the previously predicted TTS. The transcription initiation site (TIS) and the TTS for the middle and the right ORFs conform to predicted sites. PstDNV transcript quantification in infected shrimp revealed that the NS and CP transcripts were expressed at an equivalent level and significantly higher than the middle ORF transcript. *In vitro* assay showed that P2 had the highest promoter activity followed by P11 and P61. Transcription mapping data provided new insights into PstDNV gene expression strategy.

© 2010 Elsevier Inc. All rights reserved.

Introduction

*Penaeus stylirostris* densovirus (PstDNV), previously known as the infectious hypodermal and hematopoietic necrosis virus (IHHNV), is one of the major viruses infecting penaeid shrimp worldwide (Lightner et al., 1996). PstDNV has been classified as a tentative species in the genus *Breviadensoivirus*, sub-family *Densovirinae*, family *Parvoviridae*. The virus was first detected in Hawaii in 1981, causing up to 90% mortality in juvenile blue shrimp (*Penaeus stylirostris*) (Lightner et al., 1983; Bell and Lightner, 1984). By 1987, the virus spread to the cultured *P. stylirostris* shrimp populations in northwest Mexico (Lightner et al., 1992) and wild population of penaeid shrimp in the Sea of Cortez, Mexico (Morales-Covarrubias et al., 1999). By 1990, the virus caused mass mortalities in farmed and wild populations of *P. stylirostris* shrimp in the Gulf of California, Mexico (Pantoja and Lightner, 1999). In a recent survey, it was observed that PstDNV prevalence has reached 95–100% across the entire Gulf of California, Mexico by 2004 (R. Robles-Sikisaka and A. K. Dhar, unpublished). In spite of wide prevalence of PstDNV, wild *P. stylirostris* fishery landing has increased steadily since the initial epizootics in 1990 (Morales-Covarrubias et al., 1999). It has been hypothesized that the virus has reached equilibrium with the host (Tang and Lightner, 2002). Elucidating the PstDNV genome expression strategies and determining the factors governing PstDNV virulence will be helpful in developing strategies to better manage the PstDNV-caused diseases in penaeid shrimp farming.

PstDNV infects most penaeid shrimp (*Penaeus* sp.) (Lightner et al., 1996). In *P. vannamei* it causes “Runt Deformity Syndrome” (RDS), which is characterized by reduced growth rates and a variety of cuticular deformities of the rostrum, antenna, thoracic and abdominal areas (Kalagayan et al., 1991; Primavera and Quinitio, 2000). Depending on the level of infection, the losses due to RDS may vary from 10% to 30% (Lightner et al., 1996).

PstDNV virus consists of non-enveloped icosahedral particles measuring 22 nm in diameter (Bonami et al., 1990). The genome is composed of a single-stranded, linear DNA of 4.1 kb with three large open reading frames (ORFs) (Shike et al., 2000). The left ORF most likely represents the non-structural protein since it contains replication initiator motifs, NTP-binding and helicase domains (Shike et al., 2000). The right ORF encodes the capsid protein, while the function of the middle ORF is unknown. In addition to three large ORFs, PstDNV genome contains a small ORF upstream of the left ORF. The small ORF contains one putative 5’ donor site (D1) and three putative acceptor sites (A1, A2 and A3) are present at the 5’-end of the left ORF (Shike et al., 2000: Fig. 1A). All three major ORFs and the small ORF preceding the left ORF are present on the plus strand. In addition to these ORFs, there is a potential ORF in the minus strand. The overall genome organization of PstDNV is similar to the densoviruses of the genus *Breviadensoivirus* in the family *Parvoviridae* (Shike et al., 2000). *In silico* analysis of the P2 and P61 promoters located upstream of left and right ORFs, respectively, and the functional characterization of these two promoters were carried out in a variety of cell types (Dhar et al., 2007).
Gene expression in mammalian parvoviruses is regulated through a combination of several mechanisms, such as transcriptional initiation site usage, transcript splicing, translation initiation site usage, protein processing, leaky ribosomal scanning, temporal gene expression, and transactivation (Cotmore and Tattersall, 2006). However, much less is known about Brevidensovirus gene expression. The type species, Aedes aegypti densonucleosisvirus (AaeDNV) genome encodes two overlapping ORFs the nonstructural proteins, NS1 and NS2 (Afanasiev et al., 1991). NS1 transactivates the structural gene promoter (Afanasiev et al., 1994). It is not known whether NS1 and NS2 originate from alternative splicing events as in some Densoviruses (Yamagishi et al., 1999) or from initiation at an alternative AUG codon (Bergoin and Tijssen, 2000; Kimmick et al., 1998).

In order to harness more insight into PstDNV gene expression, we report in this study the structural and functional promoter analysis of three PstDNV ORFs and the possible role of splicing in the gene expression strategy of this shrimp virus. The transcription mapping data revealed that a hitherto unknown promoter transcribes the PstDNV middle ORF. Quantitative RT-PCR was used to determine the levels of PstDNV transcripts in infected shrimp. PstDNV promoters were cloned into a promoter assay vector, pGL3Basic, and the activities of these promoters were assayed in SF9 cells.

**Results**

**Northern analysis of PstDNV gene expression**

Shrimp injected with PstDNV started dying by day 9 post-injection. Seven out of 36 PstDNV injected shrimp died by day 12 post-injection. There was no mortality in the control treatment. At day 16, all surviving animals were sacrificed and tissue archived at −80 °C. Total RNA was isolated from tail tissue of a subset of healthy and PstDNV-infected animals and taken for northern blot analysis. Hybridization of total RNA of PstDNV-infected shrimp tail muscle tissues with probe encompassing almost the entire PstDNV genome revealed five bands of 4.1 kb, 2.6 kb, 1.9 kb, 1.3 kb and 0.9 kb (Fig. 1C). Based on the predicted size, the 2.6 kb, 1.9 kb and 1.3 kb bands correspond to the transcripts encoded by the left, middle and the right ORFs, respectively. The 0.9 kb transcript might represent the small ORF present in the minus strand of the PstDNV genome, but this needs to be further verified. The origin of the remaining 4.1 kb transcript cannot be predicted based on the published genome sequence analysis (Shike et al., 2000). However, the genome size of PstDNV being 4.1 kb, the largest northern band may represent the full-length transcript of the virus.
Mapping of transcription initiation sites (TIS) of PstDVN transcripts in *Penaeus vannamei* shrimp

To capture the transcription initiation sites (TIS) of PstDVN transcripts, 5'RACE RT-PCR was performed. When RT-PCR was performed using the primer P2A2 559R, a ~150 bp amplicon was obtained (Fig. 2, Panel A, lane 1). The cDNA was gel purified, cloned into pCR-TOPO vector and four clones were sequenced (Fig. 3, Panel A). The nucleotide sequence data of all the clones represented the transcript encoded by the left ORF. Of the four left ORF 5’ RACE-cDNA clones, two clones showed transcription initiation at base 101, in conformity with our previously reported bioinformatics analysis (Dhar et al., 2007). The other two mRNAs were transcribed three bases upstream, from base 98 of PstDVN (GenBank accession AF272215). The left ORF also matches 7 of 7 consensus Kozak sequence (RNNAUGG), indicating it is a strongly expressed gene (Kozak, 2002). The nucleotide sequence data of the left ORF 5’ RACE clones, therefore, revealed that PstDVN pre-mRNA transcribed by the P2 promoter undergoes splicing at D1/A1 sites (Figs. 1B and 3, Panel A).

None of the transcripts initiated by the P2 promoter represented transcript for the middle ORF. Subsequently, RT-PCR was performed using a newly designed primer, P2A3 768R. Two amplicons, ~550 bp and ~350 bp, were obtained (Fig. 2, Panel B). These two cDNAs were gel purified, cloned into pCR-TOPO vector and sequenced. The nucleotide sequence of the recombinant clones representing the ~550 bp cDNA represented the transcript encoded by the left ORF. The transcription initiation site and the splicing sites (D1/A1) were the same as shown in Fig. 1, Panel B. However, the nucleotide sequence data of the cDNA clones for the ~350 bp product represented the middle ORF. Eight of seventeen 5’-end RACE-cDNA clones of the middle ORF have transcription initiation at nucleotide position 466 (C) and the nine other clones have initiation at 467 (A) base of PstDVN (Fig. 3, Panel B). Only 5 of the seven nucleotides near translation initiation matched with the Kozak consensus sequence (Kozak, 2002), indicating that it is a weakly expressed gene (Sakai et al., 2001).

When RT-PCR was performed using the adapter primer from the 5’ RACE kit and the primer P61 2648R, a 300-bp cDNA was amplified (Fig. 2, Panel A Lane 2). The amplified cDNA was gel purified, cloned into pCR-TOPO vector, and sequenced (Fig. 3, Panel C). Of the eleven right ORF 5’-end RACE-cDNA clones, 6 clones showed transcription initiation at base 2441 and the remaining five clones started at base 2442 of PstDVN, supporting our previous bioinformatic prediction (Dhar et al., 2007). The nucleotide sequence of the recombinant clones showed that the transcription initiation occurred at 20 (C) and 21 (A) nucleotide downstream of the TATA-box of P61 promoter (Fig. 3, Panel C). Only 6 of the 7 nucleotides at the predicted translation initiation site match with the consensus Kozak sequence, indicating the ORF as an “adequately expressed” gene (Kozak, 2002).

Mapping of transcription termination (TTS) sites of PstDVN transcripts in *Penaeus vannamei* shrimp

A 3’ RACE RT-PCR performed using an adapter primer and the primer L2443F designed based on the left ORF provided two amplicons of approximately 300 bp and 430 bp (Fig. 4, Panel A). The RT-PCR results indicated that the transcription of the PstDVN left ORF is terminated at two different sites. These cDNAs were cloned and sequenced to determine the transcription termination sites (TTS) discussed in further detail below. To capture the 3’-end of the transcript representing the middle ORF, RT-PCR was performed using the adapter primer and the primer M1374F. A cDNA of approximately 900 bp was amplified (Fig. 4, Panel B Lane 1). Cloning and sequencing of this cDNA revealed that the 900 bp product contained two amplicons differing in size by 53 bases. The 3’-RACE RT-PCR performed to capture the 3’-end of the right ORF provided a cDNA of 450 bp (Fig. 4, Panel B Lane 2). The nucleotide sequence data of the 450 bp cloned cDNA amplicons revealed that it represents only a single class of transcript as discussed below.

The mature transcripts of PstDVN left ORF polyadenylated in two different sites in the transcript (Fig. 5, Panel A). The shorter transcripts polyadenylate after base positions 2375 to 2379 of PstDVN. The longer transcripts polyadenylate after 2510 base and are 132–136 bases larger. Although there is no canonical positioning AAUAAA signal present upstream of the shorter left ORF transcripts, there is an AACAAA element (nucleotide positions 2323–2328, Fig. 5A) in the corresponding site. However, there is an AAUAAA element (2410–2415, Fig. 5A) upstream of polyadenylation for the longer transcripts. A-rich sequences are present immediately downstream of both transcript sizes in the PstDVN DNA. No U or G and U rich site could be located downstream of either size-transcripts on the PstDVN DNA.

We also detected two different sizes of mature transcripts for the middle ORF (Fig. 5, Panel B). The shorter mRNA ends in 1867 nucleotide position, whereas the longer mRNA polyadenylates at nucleotide position 1920. Unlike in left ORF, however, there is only one AAUAAA element (from nucleotide position 1769 to 1774) upstream of both transcripts. Similar to left ORF transcripts, middle ORF transcripts also polyadenylate right before A-rich sequences of the PstDVN genome. Also similar to left ORF transcripts, no T or GT-rich site could be located downstream of the left ORF polyadenylation site.

We could detect only one size class of transcript for the right ORF; all the mRNA being polyadenylated from the nucleotide position 3595 in PstDVN genome with the exception that clone 3 had an extra G before polyadenylation (Fig. 5, Panel C). There is an AAUAAA (nucleotide positions 3577–3582) element in the proximity of the polyadenylation site. In addition, there were two more hexanucleotide elements (AACAAA, from nucleotide positions 3501 to 3506 and ACUAAA, from nucleotide positions 3526 to 3531) upstream of the AAUAAA element. Nucleotide polymorphism was noted (U, C or G) at the last two nucleotides of polyadenylation. In lieu of A-rich sequences, there is a CT-rich site (3661–3688 b) in PstDVN, downstream of the right ORF polyadenylation site.

**Identifying the promoter region of the middle ORF**

The nucleotide sequence region upstream of the middle ORF (GenBank accession AF273215) was analyzed using the both Neural Network Promoter Prediction and TRANSFAC Promoter signal scan programs to identify any potential sequence motif that showed similarity to the promoter elements of eukaryotic origin. We detected two tandem repeat CTTC elements (406–410 and 416–420 bases), a downstream TATA-like box (AAATATCC, 437–445 bases), an initiation
of transcription signal (Inr or TIS) CATT at 466–469 base, three Gs at position +22 to +24 relative to A of the Inr and a downstream promoter element (DPE) AGACC at 494–498 base (Fig. 6), all conforming to the rules of an eukaryotic promoter region (Kutach and Kadonaga, 2000). Since this promoter region was located near map unit 11, this promoter has been referred as P11.

Determining the functionalities of PstDNV promoters in insect cells

To compare the functional activities, PstDNV promoters (P2, P11, and P61) were cloned upstream of a firefly luciferase gene in a promoter assay vector, pGL3-Basic (Fig. 7A) and cotransfected with pGL4.75 hRLuc/CMV into Sf9 cells. The luciferase assay revealed that all three PstDNV promoters were functional in the cells of Spodoptera frugiperda, a non-host for PstDNV. The relative expression of firefly luciferase driven by the PstDNV promoters compared to the Renilla luciferase is shown in Fig. 7B. The luciferase expression driven by the P2 promoter was 3.7 fold higher than P11 and over 5-fold higher than P61 promoter. Paired two-t-test showed that the in vitro activity of the P2 promoter was significantly different (p<0.006) from both P11 and P61. The luciferase expression driven by the P11 promoter was 1.4 fold higher than the P61 promoter. Although there was no significant
difference between these two promoter (at 95% confidence level $p = 0.08$), there was a trend in higher luciferase expression driven by P11 promoter compared to the P61 promoter.

Quantification of PstDNV transcripts in infected shrimp

The abundance of PstDNV transcripts in infected shrimp was measured by real-time RT-PCR. The average copy numbers of transcripts representing the NS, middle ORF and the CP genes were $843 \pm 303$, $222 \pm 113$, and $789 \pm 522$ copies/µg of total RNA, respectively (Fig. 8).

There was no significant difference in the copy numbers of transcript between NS and CP genes ($p = 0.669$). However, the transcript copy numbers of both NS and CP genes were significantly higher than the transcript copy number of the gene representing the middle ORF ($p = 0.001$). Overall, there was a high variability in the transcript number for all three genes among the samples tested. This is most likely due to the variation in PstDNV infection level among different animals. Due to the lack of an immortal cell line in shrimp, PstDNV

Fig. 5. The alignments of nucleotide sequences of 3'-RACE cDNA clones representing the left (A), middle (B) and right (C) ORFs of PstDNV. The PstDNV genomic sequence and the nucleotide position (based on GenBank accession AF273215) are indicated at the top. The putative upstream positioning element for polyadenylation are in bold, the A-rich PstDNV sequences at the site of mRNA polyadenylation are in bold and italics, and the GT-rich sequences (GGCATGGTGGAACACTTTTCTTCTTATTC) downstream of polyadenylation in right ORF is indicated by bold letters. The transcription termination site (TTS) in each transcript is indicated by the nucleotide in shade. The numbers on the left of each sequence represent the cDNA clone number. "-" Represents similar nucleotide.

Fig. 6. The nucleotide sequence of PstDNV P11 promoter region. The ASL (Activation Sequence-1 Like) box, TATA-like box, transcription initiation site (TIS), the conserved guanidine (G) residue at 24 nucleotides downstream of TIS (+24) and the downstream promoter element (DPE)-like box are indicated above the nucleotide sequence.
PstDNV, the virus that causes disease, and the economic losses it inflicts on shrimp farming worldwide, diagnostic methods including histology (Lightner et al., 1983), in situ hybridization (Mari et al., 1993), PCR, and real-time PCR (Nunan et al., 2000; Dhar et al., 2001) were developed to detect the virus and management strategies to prevent the introduction of the virus into shrimp facilities have been developed (Lotz, 1997). Recently, PstDNV genome has been sequenced, and the genome organization has been elucidated (Shike et al., 2000). However, no information is available on the gene expression strategies of PstDNV, and how the virus causes the disease. To gain insight on the gene expression strategies of PstDNV, we conducted RT-PCR and northern analysis of PstDNV expressed genes as well as 5'- and 3'- RACE-PCR to delineate transcription initiations and termination sites, respectively, of the three PstDNV encoded genes. We showed that PstDNV middle ORF is transcribed by a novel promoter P11. In addition, the functionality of the P11 promoter was assayed in vitro along with the two other PstDNV promoters, P2 and P61, respectively.

Our northern hybridization data indicated that PstDNV transcribes a 4.1 kb RNA that covers the entire transcription potential of the genome. To our knowledge, this is the first report of profiling the transcripts of PstDNV. The nature of the largest RNA could not be explained at present, but we assume its origin from the P2 promoter and may represent an unprocessed PstDNV transcript. Recently, Qui and colleagues (2006) reevaluated the transcription mapping of Aleutian mink disease parvovirus (AMDV), an autonomous parvovirus, and detected a full-length transcript of AMDV. In addition, the authors found that for functional purposes the internal AMDV P36 promoter is not used to transcribe the two downstream structural proteins, VP1 and VP1. Instead, the full-length (4.1 kb) transcript is processed to translate both nonstructural and structural proteins. Unlike AMDV, however, the detection of the 2.6 kb, 1.9 kb and 1.3 kb transcripts in PstDNV suggest that all three major ORFs present on the plus strand of PstDNV genome are expressed during virus infection in shrimp. The detection of a 0.9 kb transcript, which might represent the small ORF in PstDNV minus strand, suggests that PstDNV genome could be ambisense like other densoviruses (Tijssen et al., 2003; Wang et al., 2005).

In an earlier communication (Dhar et al., 2007), we reported the sequence characterization of the core promoter regions of left ORF and right ORF of PstDNV. In this report, we demonstrated the transcription initiation sites of these two promoters. In addition, we showed that the transcription initiation of the middle ORF occurred at nucleotides 466 (C) and at 467 (A) downstream of the P2 promoter (Fig. 3, Panel B). Transcription of an mRNA generally starts at 25–30 nucleotide of the TATA-box. The initiation sites of these transcripts indicate that the middle ORF is not transcribed by the P2 promoter and is likely to be transcribed by a novel promoter. Therefore, we analyzed the PstDNV genomic region (181 to 540 bases in the GenBank accession AF273215) surrounding the transcription initiation site of the middle ORF and identified several promoter motifs (Fig. 6). Similar to the P61 promoter, P11 promoter does not have a canonical TATA box (Shike et al., 2000; Dhar et al., 2007). However, being in close proximity of the Inr box, the AT-rich site (AAATATCG) most likely substitute TATA element in the middle ORF promoter. The role of the tandem repeat elements CTTTC is not known at present. The Inr element is also strengthened by the presence of G nucleotides at +22 to +24 and the DPE element. The sequence data of the cDNA clones in the present study confirmed that transcription starts both at the C and the T of the Inr element. When transcription starts at nucleotide C, the second G residues among the three downstream Gs, becomes +24. The roles of similar elements in PstDNV P2 and P61 promoter functions had been discussed (Dhar et al., 2007).

The sequences surrounding the translation initiation sites of the transcripts of the left, middle and right ORFs were analyzed for Kozak sequences. The Kozak consensus sequence for initiation of translation in vertebrates is (GCC) GCCRCCATGG, where R is a purine (A or G) and T is a pyrimidine (T or C). The sequences surrounding the start codon are more important than others, particularly the −3 and the +4 nucleotides. In a strongly expressed gene, the −3 (R) and the +4 (G) nucleotide should match exactly. In an adequately expressed gene, only one of those two sites must conform to consensus. In absence of those two matches, a gene is considered as leaky. In some cases, a gene is considered as leaky expressed (Sakai et al., 2001). Ozawa and colleagues (1987) indicated that the Kozak sequence in human parvovirus is RNAATG. A recent reevaluation of translation initiation codons of vertebrates revealed that sequences surrounding the start codon significantly deviate from Kozak sequence (Peri and Pandey, 2001). These authors suggested that some other means, such as leaky scanning, reinitiation, or internal initiation of translation, may play a greater role under such circumstances. Analysis of the PstDNV genes demonstrated that...
only the left ORF follows the perfect consensus Kozak rule, whereas the middle ORF and the right ORF deviate slightly from the consensus. Similar observations have been made in another brividesivirus, hepatopancreatic parovirus (HPV) of Penaues monodon (Sukhumsirichart et al., 2006), an iteravirus infecting Dendrolimus punctatus (Wang et al., 2005), a human pathogenic B19 parovirus (Ozawa et al., 1987) and a simian parovirus (Vashisht et al., 2004).

Unlike in prokaryotes, eukaryotic 3’-ends of transcripts are not processed as precise as the 5’-ends of mature transcripts. Also, heterogeneity of 3’-end is common in eukaryotes (Zhao et al., 1999; Gilmartin, 2009). In general, there are at least two signals required for eukaryotic transcription termination. The mammalian poly-A signal typically consists of a nearly 45 nt long core sequence that may be flanked by diverse auxiliary sequences which enhances cleavage and polyadenylation efficiency (Kim and Martinson, 2003). The core sequence consists of a highly conserved upstream positioning element (AAUAAA) recognized by a cleavage and polyadenylation-specific factor and a poorly defined downstream region rich in U or G/U. Downstream GU or U-rich sequences are known to influence temporal pattern (i.e., early vs. late transcription) of gene expression in SV40 (Connelly and Manley, 1988). The poly (A) cleavage site, usually a dinucleotide CA, is generally located somewhere in between these two elements, (Guo and Sherman, 1996). However, unlike in mammalian systems, in yeast the poly-A signal AAUAAA is not functional and GU-rich sequences are not located in the 3-end yeast ORFs (Guo and Sherman, 1996). Instead, there is also an AU-rich efficiency element upstream of the positioning element, necessary for proper 3’-end formation of yeast pre-mRNA (Guo and Sherman, 1996). Even in human genes, the hexameric sequence is not strictly conserved (Beaudoin et al., 2000) and human pathogenic B19 parovirus is known to utilize variants AAUAAA and AAUAC for polyadenylation and it also demonstrates size heterogeneity of transcripts (Ozawa et al., 1987).

We have observed two distinct size classes of 3’-UTRs for both left and middle ORFs and only one size class for the right ORF. Whereas AATAAA element is present upstream of poly (A) site of the larger transcript of left ORF, AATAAA is absent upstream of the corresponding shorter transcript of left ORF. Instead, AAAAAA sequences are present in the equivalent position. Unlike in left ORF, however, both transcript size groups of the middle ORF (1769–1774 bases) and the transcript of right ORF (3577–3582 bases) contain upstream AATAAA elements. In addition, there are two hexanucleotide elements (AACAAA, 3501–3506 bases and ACTAAA, 3526–3531 bases) upstream of the AATAAA element. It is not clear at this point which of the non-canonical hexanucleotide sequences are necessary for positioning of some of those PstDNV transcripts. More than single transcript size classes and the presence or absence of canonical polyadenylation sites in PstDNV transcripts may reflect general evolutionary diversities of transcription termination in eukaryotic hosts (Gilmartin, 2009).

The promoter assay data showed that PstDNV middle ORF promoter sequences, identified by transcriptional mapping and bioinformatic analysis, is indeed functional and capable of driving the expression of a reporter gene in a heterologous expression system (Fig. 7B). The luciferase assay data, therefore, supports our transcription mapping data to suggest that the middle ORF in PstDNV genome is transcribed by a novel promoter, P11, located upstream of this ORF. Among the three PstDNV promoters (i.e., P2, P11 and P61), the P2 promoter was the strongest in driving the luciferase activity, and there was no significant difference in promoter activities between P11 and P61 promoters (Fig. 7B). It was interesting to note that the luciferase activity driven by the P2 promoter, a promoter driving a non-structural gene, was significantly higher than the promoter activity of a structural gene of PstDNV. However, the PstDNV transcript quantification data showed that both the NS and the CP genes are expressed at equivalent level in PstDNV infected shrimp (Fig. 8). This suggests that the P61 promoter is likely to be trans-activated in vivo by PstDNV NS protein. In parvoviruses transactivation of structural gene promoter by NS protein is well documented (Rhode, 1985; Doering et al., 1988; Storgaard et al., 1993). In addition, in a mammalian parovirus, AMDV, it has been suggested that weakness of the internal promoter may contribute to persistence of the virus in CRFK cells (Christensen et al., 1993; Storgaard et al., 1993, 1997). It remains to be determined how the dynamics of PstDNV transcript abundances modulate viral pathogenesis and if the persistence of PstDNV infection in shrimp is dependent on the abundance of CP transcript.

Materials and methods

PstDNV challenge

Specific Pathogen Free (SPF) shrimp (Penaues vannamei, juvenile with an average weight of 6.5 g) were injected with a PstDNV positive inoculum and maintained in 20 L aquaria following a published protocol (Dhar et al., 2001). Animals injected with 2% saline served as controls. There were four aquaria, each containing 12 animals. Three aquaria contained PstDNV injected animals and one aquarium had healthy control animals. Each aquarium contained artificial seawater (25 ppt) at 27–29 °C and was supplied with continuous aeration through placement of a single air stone. Animals were fed a pelletted ration (SI35 pellets, Ziegler Brothers) ad libitum once per day throughout the study. A 50–70% water exchange was conducted on each tank, following siphoning of organic debris, every 3 days during the course of the bioassay. On day 16 post-injection, animals were sacrificed and frozen at −70 °C until extraction of nucleic acid.

Cloning of PstDNV DNA

Total genomic DNA was isolated from tail muscle of PstDNV-infected shrimp using DNAZol and manufacturer’s protocol (Molecular Research Center, MRC, Cincinnati, OH). A 3573 bp PstDNV DNA was amplified using the primers (IHHNVF and IHHNVR, Table 1) that were designed based on GenBank accession number AF273215. The polymerase chain reaction (PCR) was performed in a 50-µL reaction volume containing 100 ng of DNA, 25 µL of AmpliTaqGold (Applied Biosystems, Foster City, CA) and 300 nM each of forward and reverse primers. The thermal profile for the PCR was 95 °C 10 min followed by 40 cycles of 95 °C 10 s and 60 °C 1 min. The amplified DNA was run in a 1% agarose gel, gel-purified using the QIAquick gel-extraction kit (Qiagen, Valencia, CA), and the PCR product was cloned into a pCR-XL-TOPO vector (Invitrogen, Carlsbad, CA). Three recombinant clones were sequenced.

Isolation of RNA and Northern blot hybridizations

Total RNA was isolated from tail muscle of PstDNV infected shrimp using TRI reagent and manufacturer’s protocol (Molecular Research Center). Twenty micrograms of total RNA were run per well in a 1% formaldehyde agarose gel using MOPS buffer, and blotted onto a nylon membrane following a standard laboratory protocol (Sambrook et al., 1989). Northern blot hybridization was performed using a non-radiolabeled probe (BioPrime DNA labeling System, Invitrogen) and following manufacturer’s protocol (Super Hyb® Kit, Molecular Research Center). In order to make a non-radiolabeled probe, the 3573 bp PstDNV insert was amplified from the plasmid DNA of a sequence verified PstDNV clone by PCR using the primers IHHNVF and IHHNVR (Table 1) using the protocol described above.

Rapid amplification of cDNA end (RACE) technique

To capture the 5’- and 3’-ends of the PstDNV transcripts, the rapid amplification of cDNA ends (RACE) was performed using total RNA isolated from a tail muscle tissue of PstDNV infected shrimp and manufacturer’s protocol (First Choice® RLM RACE Kit, Ambion, Inc.,...
The nucleotide sequence of the primers used for transcriptional mapping of PstDNV genes.

![Table 1](image)

Sequence analysis

The nucleotide sequence of PstDNV promoters were analyzed using both Neural Network Promoter Prediction (FPNN: [http://www.fruityfly.org/seq_tools/promoter.html](http://www.fruityfly.org/seq_tools/promoter.html)) and TRANSFAC Promoter signal scan ([http://bimas.dct.nih.gov:80/molbio/signal/](http://bimas.dct.nih.gov:80/molbio/signal/)) programs. The multiple alignments of the cDNA sequences of PstDNV promoters were performed using CLUSTAL program ([http://align.genome.jp](http://align.genome.jp)).

Cloning of PstDNV promoters

PstDNV promoters P2, P11, and P61 were cloned into luciferase reporter vector pGL-3 Basic (Promega Corp., Madison, WI). This vector contains a modified coding region for firefly (*Photinus pyralis*) luciferase gene that has been optimized for evaluating transcriptional activity in eukaryotic cells. The P2 and P61 promoters were cloned between the *Pst*I sites of the pGL3-Basic vector, upstream of the luciferase gene that has been optimized for evaluating transcriptional activity in eukaryotic cells. The P2 and P61 promoters were cloned into luciferase reporter vector pGL-3 Basic (Promega) ([Fig. 7, Panel A]). The plasmid DNA was isolated from the recombinant clones and sequence verified before using for transfection.

Transfection of Spodoptera frugiperda (Sf9) cells and luciferase assay

The activities of PstDNV promoters in driving the luciferase gene were assayed in insect cell, Sf9. Sf9 cells were grown in BD BaculoGold medium (BD Biosciences) and cells in log phase were cotransfected with the plasmid DNA of P2, P11, or P61 promoter constructs and pGL4.75 hHrLuc/CMV (Renilla luciferase driven by the CMV promoter, Promega) using Cellfectin (Invitrogen) as transfection aid. The plasmid pGL4.75 hHrLuc/CMV was used as an internal control for transfection and to normalize firefly luciferase activity driven by PstDNV promoters. The transfection protocol was essentially same as described earlier ([Dhar et al., 2007]). Briefly, a 24-well tissue plate was seeded with 0.5 mL Sf9 cells per well at 5×10^5 cells/mL and approximately 20 h later the medium was removed, cells were washed once with serum-free TnM-FH, and cells were incubated under 0.5 mL transfection mixture per well. Transfection mixtures consisted of 1 µg plasmid DNA of PstDNV promoter construct, 0.5 µg of pGL4.75 hHrLuc/CMV and 7.5 µL Cellfectin (Invitrogen) according to the manufacturer's protocol. After 2.0 h incubation, the transfection mixture was removed and cells were covered with 0.5 mL BD BaculoGold medium. Forty-eight hours post-transfection, the medium was removed by aspiration, cells were lysed with 100 µL Glo-Lysis Buffer (Promega) and luciferase activities were determined using the Dual-Glo Luciferase assay Kit (Promega) following the manufacturer's recommendations. The in vitro luciferase assay was done three times and the luciferase data (fLuc/rLuc values) were normalized to control treatment (non-transfected Sf9 cells) before running paired two t-test. The normalized data were then taken to plot the bar diagram represented in Fig. 7.

Quantification of PstDNV transcript in shrimp

The quantification of PstDNV transcript was done by real-time RT-PCR following previously published protocol ([Dhar et al., 2001, 2009]). The primers used to perform real-time PCR for the NS, the gene representing the middle ORF and the CP gene are provided in Table 1. To quantify the NS gene, primers were designed in the 5′-end of the gene that does not overlap with the middle ORF. The middle ORF, however, completely overlaps with the left ORF representing the NS gene. So the primers used to amplify the gene representing the middle ORF, also amplified the NS gene. Therefore, to quantify the middle ORF transcript, the copy number of the NS transcript was deducted from the copy number of NS and middle ORF transcript. The 5′-end of the right ORF, that encodes the capsid gene, overlaps with the 3′-end of the left ORF. So to quantify CP gene, the primers were designed from the unique region of the CP gene.

The CDNA synthesis was performed in a 20-µL reaction volume containing 1 µg of DNase I-treated total RNA of PstDNV infected tail muscle tissue, 1× PCR buffer (pH 8.2), 1 mM dNTP (PE Applied Biosystems), 0.75 µM random hexamer primer, 4 units of RNase inhibitor (PE Applied Biosystems), and 5 units of Multiscribe reverse transcriptase (PE Applied Biosystems). The real-time RT-PCR mixture contained 7.1 µl of 2× iQ™ SYBR Green Supermix (Bio-Rad Laboratories, Inc.), 0.24 µM of reverse and forward primer for the corresponding gene, and 1 µl of 1:10 diluted CDNA in a 25 µl reaction volume. Real-time RT-PCR assays were carried out in a MyIQ™ thermocycler (Bio-Rad). Sample reactions were performed in triplicate in a 96 well plate. The thermal profile for the reactions was 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. In order to check the quality of RNA, in each real-time RT-PCR assay, shrimp EF-1alpha gene was amplified as a control gene ([Dhar et al, 2009]). The Ct values for EF-1alpha for all ten animals were very consistent (varied between 21 and 22) indicating that the RNA quality was good.
In order to generate the standard curves for the three PstDNV encoding genes, a ten-fold dilution series (2.8 × 10^{-10} to 2.8 copies) of gel-purified PstDNV of 3573 bp amplicon (see Cloning of PstDNV DNA section above) was used as template for real-time PCR. After a SYBR Green PCR run, data acquisition and subsequent data analyses were performed using the iCycler iQ Real-Time PCR Detection System (Bio-Rad iQ Software Version 1.3). The cycle threshold (Ct) values were exported into a Microsoft Excel spreadsheet. In order to calculate the copy number of the transcript, the Ct values of the PstDNV gene were extrapolated into the standard curve of the corresponding gene. The significance of difference in the transcript abundance of different PstDNV genes were then determined using a Two-Way ANOVA test followed by Fisher’s PLSD test (Mandal, 2007).

Acknowledgments

This project received support from the Defense Advanced Research Projects Agency (DARPA) and from the Defense Threat Reduction Agency (DTRA) Chemical and Biological Technologies Directorate Transformational Medical Technologies Initiative under DTRA contract HDTRA1-07-C-0078. Advanced BioNutrition Corp., Columbia, MD provided partial funding for this project. The authors would like to thank Dr. Nikolai van Beek for technical assistance, and Dr. Kenneth W. Hasson, Texas Veterinary Medical Diagnostic Laboratory, College Station, TX for performing the PstDNV bioassay.

References


