Characterization of white spot syndrome virus immediate-early gene promoters

Fanyu Lin,1,2 He Huang,1,2 Wei Ke,1,2 Luhong Hou,2 Fang Li2 and Feng Yang2

1School of Life Science, Xiamen University, Xiamen 361005, PR China
2Key Laboratory of Marine Biogenetic Resources, Third Institute of Oceanography, State Oceanic Administration (SOA), Xiamen 361005, PR China

Twenty-one immediate-early (IE) genes of white spot syndrome virus (WSSV) have been identified so far. However, the transcriptional regulation of WSSV IE genes remains largely unknown. In this report, the 5' flanking regions of 18 WSSV IE genes were cloned and eight functional promoter regions were identified. WSSV IE gene promoters normally contained a TATA box approximately 30 bp upstream of the transcriptional initiation site. Also, the cyclic AMP response element (CRE; TGACGTCA) was frequently found within the WSSV IE promoter regions. Mutations of the CREs of WSSV IE promoters P403 and P465 reduced their activity significantly, suggesting that these elements have a role in WSSV IE gene transcription. Our findings provide a more global view of WSSV IE gene promoters and will facilitate the in-depth investigation of viral gene transcriptional regulation.

White spot syndrome virus (WSSV) is a dsDNA virus with a broad host range in crustaceans. It has a genome of approximately 300 kb, encoding about 180 genes (Leu et al., 2009). According to their temporal expression profile, WSSV genes can be classified into three groups: immediate-early (IE), early and late genes (Marks et al., 2005). IE genes are the first class of viral genes to be expressed during infection, and often encode regulatory proteins that are essential for the initiation of viral primary infection or the switch from latency to lysis (Yuan, 2005). Research on the transcriptional regulation of viral IE genes will help to reveal the key events of early viral infection.

In total, 21 IE genes have been identified from WSSV thus far (Li et al., 2009; Lin et al., 2011; Liu et al., 2005). Of these, ie1 is the most-studied gene, having been first identified in 2005, and later shown to encode a transcription factor (Liu et al., 2008). Recent research has revealed that the transcription of ie1 is regulated by PmSTAT (Liu et al., 2007) and NF-κB family proteins LvrRelish (Huang et al., 2010) and Lvdorsal (Wang et al., 2011). These proteins can bind to the promoter region of ie1 and enhance its transcription. The STAT and nuclear factor kappa B (NF-κB) pathways are known to regulate antiviral immunity in mammals and insects, resulting in interferon or small interfering RNA production (Kemp & Imler, 2009; Seth et al., 2006). Therefore, WSSV may take advantage of the host immune pathway to facilitate its own gene expression. In addition to ie1, the 5' flanking regions of two WSSV IE genes, ie2 and ie3, have also been analysed previously (Liu et al., 2005), but these two genes were found to contain no functional promoters. Until now, little is known about the transcriptional regulation of WSSV IE genes apart from ie1.

To investigate whether the remaining 18 WSSV IE genes (wsv051, wsv056, wsv078, wsv079, wsv080, wsv083, wsv091, wsv094, wsv098, wsv099, wsv100, wsv101, wsv103, wsv108, wsv178, wsv249, wsv403 and wsv465) contain functional promoters, we cloned their 5' flanking regions to generate reporter constructs for promoter-activity analysis. The core promoter elements of eukaryotes are normally approximately 35 bp up- or downstream of the transcriptional initiation site (TIS) (Butler & Kadonaga, 2002; Smale & Kadonaga, 2003), and previous research on WSSV gene transcription has revealed that the distance from the TIS to the translational start codon (ATG) is normally 20–85 nt for WSSV early genes and 30–220 nt for late genes (Marks et al., 2006). Therefore, DNA fragments immediately upstream of the start codons of the 18 WSSV IE genes (WSSV China isolate, GenBank accession no. AF332093), which were expected to contain core promoter elements, were cloned into pIZAIE/EGFP (Luo et al., 2007) with primers listed in Table S1 (available in JGV Online). The length and position of these WSSV promoters in WSSV are shown in Table S2.

As no cell lines of WSSV hosts are yet available, an alternative system is required to analyse the putative promoters of WSSV IE genes. It has been reported that WSSV could ‘infect’ insect cells such as Sf9 and S2, resulting in expression of some viral genes and replication.

Three supplementary figures and seven supplementary tables are available with the online version of this paper.
of the viral genome, but no progeny virions could be detected in these cells (Chen et al., 2011; Nupan et al., 2011). A similar phenomenon has been observed in High Five (BTI-TN-5B1-4) cells. Sixteen WSSV IE genes were found to be transcribed in WSSV-infected High Five cells, while the transcripts of the other five IE genes, wsv080, wsv094, wsv099, wsv101 and wsv103, were not detected (Fig. S2). Although WSSV cannot complete its life cycle in insect cells, Sf9, S2 and High Five cells have been used successfully to study WSSV gene function (Liu et al., 2008; Wang et al., 2011) and transcriptional regulation (Huang et al., 2010; Liu et al., 2007, 2008). In the current study, the reporter constructs were transfected into High Five cells for promoter-activity analysis. As shown in Fig. 1(a), 48 h post-transcription, EGFP signal was observed in High Five cells transfected with the constructs containing WSSV promoters P051, P056, P083, P108, P178, P249, P403, P465 and the positive controls (insect virus promoter OpIE2 and WSSV ie1 promoter PIE1), suggesting that the inserted DNA fragments indeed comprise the WSSV IE gene promoter regions. In contrast, no EGFP fluorescence was detected in cells transfected with the other ten constructs containing putative WSSV IE gene promoters, i.e. P078, P079, P080, P091, P094, P098, P099, P100, P101 and P103 (data not shown), or with the vector alone (Fig. 1a).

To compare the activity of the identified WSSV IE gene promoters further, the mean fluorescence intensity of the cells was measured by flow cytometry (FACS Calibur; BD Biosciences). The EGFP intensity of cells transfected with pIZ/EGFP (the positive control containing an OpIE2 promoter) (Luo et al., 2007) was normalized to 100 %, and the relative activity of WSSV promoters compared with OpIE2 was presented as the relative EGFP intensity. Data acquisition was restricted to a total of 50 000 cells, and three independent experiments were performed for each construct. As shown in Fig. 1(b), the relative activities of WSSV IE gene promoters compared with the OpIE2 promoter (pIZ/EGFP) were approximately 131 % (PIE1), 12 % (P051), 29 % (P056), 80 % (P083), 15 % (P108), 3 % (P178), 42 % (P249), 23 % (P403) and 309 % (P465). It was notable that the activity of P465 was about two times

---

Fig. 1. Promoter-activity analysis. (a) High Five cells were transfected with reporter constructs containing WSSV promoters as indicated. The insect virus OpIE2 promoter and WSSV IE1 promoter PIE1 were used as positive controls; vector alone was used as the negative control. Bars, 20 μm. (b) Mean fluorescence intensity of the cells transfected with reporter constructs were measured by flow cytometry. The EGFP intensity of cells transfected with the reporter construct containing the OpIE2 promoter was normalized to 100 %. Error bars indicate SD of three independent experiments.
higher than that of the positive control. Therefore, this promoter might be useful for improving the efficiency of the non-viral expression system in insect cells.

The TIS sequences of WSSV IE genes with functional promoters were determined by 5’ RACE. Total RNA was extracted from hepatopancreas of WSSV-infected crustacean (Procambarus clarkii) with TRIzol (Invitrogen), followed by DNase I (Roche) treatment to remove residual DNA. The 5’ ends of the gene transcripts were obtained using a 5'/3' RACE kit (Roche) with the gene-specific primers listed in Table S3. For each gene, seven or eight random colonies were picked for sequencing. As shown in Fig. S1, the 5’ termini of wsv083, wsv108, wsv178, wsv249, wsv403 and wsv465 were −114, −66, −40, −62 nt upstream of the start codon, respectively. In the case of wsv051, the 5’ terminus was 367 nt upstream of the start codon in four of the seven randomly picked clones and 366 nt upstream in the remaining three clones, suggesting the existence of two TISs for wsv051 at −367 and −366 nt. However, we failed to identify the TIS of wsv056, which might be due to the highly repetitive sequence (Yang et al., 2001) present in the gene. The TISs of wsv083, wsv108, wsv178, wsv249, wsv403 and wsv465 are located 20–30 nt downstream of the putative TATA boxes, which is considered to be a normal distance for host RNA polymerase II to function (Smale & Kadonaga, 2003). Moreover, the initiation sites of wsv051, wsv108 and wsv465 match the consensus motif (A/C/T)CA(G/T)T of arthropod transcriptional initiator elements. This feature is shared by some WSSV IE and early genes, including ie1, DNA polymerase, rrl and rr2 (Liu et al., 2005).

In order to identify key elements involved in transcriptional regulation, deletion mutants of the IE gene promoters, including P051, P056, P083, P108, P249, P403 and P465, were made. The 5’-truncated promoter constructs were generated by inverse PCR (Ochman et al., 1988), using the plasmid containing the full-length promoter as the template. The deletions of each IE gene used the same reverse primer and three to five different forward primers (Table S4). P178 was not used for further analysis, as its original activity was very low. In these experiments, the mean EGFP intensity of cells transfected with each ‘full-length’ promoter was normalized to 100%. The results showed that the TATA boxes approximately 30 bp upstream of the TISs were required for basal activity for all WSSV IE promoters examined (Fig. 2). Deletion of their TATA boxes resulted in a complete loss of the promoter activity to a negative level, which was consistent with the finding for the ie1 promoter (Liu et al., 2007). In addition, computer analysis revealed that nine of the ten 5’ flanking regions that showed no promoter activity in our experiments lacked a TATA box. The requirement for a TATA box in transcriptional initiation suggests that the host RNA polymerase II participates in WSSV IE gene expression.

Moreover, deletion of some fragments of the promoters other than the TATA box-containing regions also led to an obvious reduction in promoter activity. These promoter regions included P051 (−625 to −436), P056 (−419 to −285), P083 (−475 to −384), P403 (−278 to −178) and P465 (−264 to −187) (Fig. 2a, b, c, f, g). In the case of P249, deletions (−554/−331), (−330/−221) and (−220/−151) caused serial reductions in the EGFP intensity from 100 % to 83, 63 and 28 %, respectively (Fig. 2e). Further analysis of these regions using the Transcription Element Search System (TESS) (http://www.cbil.upenn.edu/cgi-bin/tess) revealed the presence of a palindromic sequence ‘TGAGCTCA’ in P051 (−467/−460), P249 (−256/−249), P403 (−212/−205), P465 (−187/−180) and P465 (−111/−104) (Fig. 2, marked with asterisks). This 8 nt palindrome is the core sequence of the cyclic AMP (cAMP)-response element (CRE) that serves as the binding site of a family of transcription factors, the CRE-binding proteins (CREB). The CREB family proteins function as regulators critical for glucose homeostasis, cell survival and long-term memory formation (Mayr & Montminy, 2001), and they are also involved in the regulation of viral gene expression (Murata et al., 2009; Sharma-Walia et al., 2010). Because of the crosstalk between the CREB and AP-1 family proteins on the cis-acting element TGACGCTA (Manna & Stocco, 2007; Pedraza-Alva et al., 1994; Samten et al., 2008), these CREs were also predicted to bind AP-1 family members. To explore the function of CREs in the WSSV IE promoters, the CRE sites in P249, P403 and P465 were replaced by irrelevant sequences by using inverse PCR with the primers listed in Table S5. Flow-cytometric analysis revealed that single or double mutations of CREs in P249 did not reduce the promoter activity, whereas mutation of the CREs in both P403 and P465 caused a significant drop in promoter activity. These findings implied that CRE has a role in WSSV IE gene transcriptional regulation (Fig. 3). Genome-wide analysis showed that there are 16 CREs in WSSV genome; seven of them are located <1 kb upstream of IE genes, whilst six of them are locate <1 kb upstream of non-IE genes (Table S6).

Additionally, five of the IE genes with functional promoters were close to WSSV homologous repeat regions (hrs). Among them, three genes (wsv249, wsv403 and wsv465) are located immediately downstream of hrs, whilst the ORF, as well as the 5’ regulatory region, of wsv056 overlap WSSV hr1 (Table S7). hrs are repeated sequences commonly existing within the genome of large circular DNA virus of arthropods, such as baculovirus (Pearson & Rohrmann, 1995), nudivirus (Wang et al., 2007) and nimavirus (Leu et al., 2009). Research on baculovirus and nudivirus showed that hrs might function as replication origins (Pearson & Rohrmann, 1995) or transcriptional regulation elements (Giarino et al., 1986; Wang et al., 2007). To explore whether the hrs close to WSSV IE promoters work as enhancers, we cloned two fragments of the hrs: the ~250 repeat unit of hr3 (WSSV-CN genome 92324–92558, near wsv178) and the ~250 repeat unit of hr5 (WSSV-CN genome 137035–137300, near wsv249). Each of them was inserted in the sense and antisense directions upstream of the P108 (−696/−1) promoter, whose original activity was relatively low.
Fig. 2. Deletion mutants of WSSV IE promoters. Deletion mutants were made in WSSV promoters and their ability to drive gene expression was measured. The mean EGFP intensity of cells transfected with each ‘full-length’ promoter was normalized to 100%. Relative intensity of each mutant compared with that of the ‘full-length’ promoter was calculated. The plasmid numbers on the left side specify the beginning and end positions (TIS as +1) of deletion mutants of each IE promoter. \( \vee \), TATA boxes; \( \times \), CREs. TISs are marked by arrows. Error bars show SD of three independent experiments.
However, no enhancement of EGFP intensity was observed in the hr-containing constructs compared with the control (Fig. S3). Therefore, these two hr units may not function as enhancers.

In summary, we carried out an overall analysis of the 5' flanking regions of WSSV IE genes. Together with previous work on the ie1 promoter (Liu et al., 2005, 2007), nine out of 21 WSSV IE genes have been shown to contain functional promoters. In addition, deletion mappings on WSSV IE gene promoters revealed that both the TATA box and the CRE are critical for the promoter activity. As transcripts of all of the 21 WSSV IE genes can be detected after WSSV infection (Li et al., 2009; Liu et al., 2005), it raises the question of how viral genes without known functional promoters are transcribed. In this study, we used an insect cell line to analyse WSSV IE gene promoters. Thus, the 10 genes found to contain no functional promoter can be divided into two classes by their expression in WSSV-infected High Five cells. The first class of genes (including wsv080, wsv094, wsv101 and wsv103) could not be transcribed in WSSV-infected High Five cells (Fig. S2), suggesting that they need the cellular proteins of WSSV hosts to activate their expression. The second class of genes (including wsv078, wsv079, wsv091, wsv098 and wsv100) could be transcribed in WSSV-infected High Five cells (Fig. S2). These genes are located within the WSSV IE gene cluster (Li et al., 2009). In the cluster, wsv078 and wsv079 are downstream of promoter P083, while wsv091, wsv098 and wsv100 are downstream of the promoter of ie1 (wsv069). Therefore, they might be transcribed into polycistronic mRNAs under the control of P083 and PIE1. Although what happens in insect cells may not fully represent what happens in WSSV host cells, our findings provide a more global view of WSSV IE gene promoters and build up a base for in-depth investigation of the transcriptional control of WSSV IE genes.

**Acknowledgements**

We thank Professor John van der Meer for proofreading of the manuscript. This work was supported by the National Basic Research Program of China (973 Program, 2012CB114401), the National Natural Science Foundation of China (40976100) and the China Agriculture Research System (CARS-47).

**References**


