Journal of Animal Research: v.6 n.1, p. 109-113. February 2016

DOI: 10.5958/2277-940X.2016.00018.8

Computational Analysis of Promoters of Immediate Early, Early and Late Genes of Bovine Herpesvirus

Mayank Pokhriyal[1*], Om Prakash Verma[2], Bhaskar Sharma[1], Barkha Ratta[1] and Ajay Kumar[1]

[1]Division of Biochemistry, Indian Veterinary Research Institute, Izatnagar, Bareilly, Uttar Pradesh, INDIA
[2]Department of Molecular and Cellular engineering, Sam Higginbottom Institute of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, INDIA

[1, #]Corresponding author: M Pokhriyal; Email: mayank.pokhriyal@gmail.com

Received: 07 January, 2016 **Accepted:** 19 January, 2016

ABSTRACT

Bovine herpesvirus-1 (BHV-1) is associated with a various respiratory and genital clinical manifestations in cattle. It belongs to *alphaherpesvirinae* family with double stranded DNA genome of 140 kb size. As other herpesviruses, BHV-1 also transcribes viral proteins into three kinetic classes namely immediate early (IE), early (E) and late (L). Transcription of these genes is dependent on several important regulatory motifs like OCT-1 (Ocatmer Binding Proteins), C/EBPα (CAATT enhancer binding protein alpha), TATA box and DPE (Downstream promoter elements). Most of these motifs have been studied in IE genes and very little literature is available for Early and Late genes in BHV-1. Hence the present study is undertaken to computationally identify similar motifs in early and late promoters of BHV-1. Computational analysis of promoters has shown that majority of its promoters uses Downstream Processing Elements (DPE) for the transcription of its genes. The genes which are known to express their proteins in high amount has TATA box along with DPE at the optimum position from the transcription start site (TSS). This study provides an overview of the regulatory motifs important for gene regulation and transcription studies.

Keywords: Immediate early ggenes, downstream processing elements, BHV-1, promoters

Bovine herpesvirus 1 (BHV-1) is a major viral pathogen of cattle that is accountable for a variety of disease conditions. The main clinical manifestations include: infectious bovine rhinotracheitis (IBR), infectious pustular vulvovaginitis (IPV), infectious pustular balanopostitis (IPB), conjunctivitis, abortions, and generalized systemic infections (Muylkens *et al.* 2007).

BHV-1 is classified in the genus *Varicellovirus* within the *alphaherpesvirinae* (Jones, 2003). The viral genome is a linear double-stranded DNA of approximately 140 kilobases (kb). It is composed of a unique long segment (UL; 105 kb) and a unique short segment (Us; 11 kb); the latter is flanked by internal repeat (IR) and terminal repeat (TR) sequences of 12 kb each. (Wirth *et al.* 1989). Like other herpesviruses, BHV-1 displays regulation of viral protein synthesis in a temporal cascade (Misra *et al.* 1981). These viral proteins are classified into three kinetic classes namely immediate early (IE), early (E) and late (L) depending on the temporal order of synthesis. IE gene transcription occurs first and does not require prior accumulation of other viral genes. On the other hand E and L genes are dependent on accumulation of products of IE genes for their transcription.

Transcription is the first step of gene expression, in which a specific segment of DNA is copied into RNA (mRNA). Promoter sequences guides the transcription process of RNAs by acting as target sites for RNA polymerase binding and thus identification of promoter sequences expedites identification of DNA-binding sites and also provides insights into the organization of transcriptional units (Schlüter *et al.* 2013)

In BHV-1 IE genes have been topic of intense research. Wirth *et al* have shown that three major IE genes (*BICP0, BICP4* and *BICP22*) are grouped into two different transcription units i.e., IE transcription unit 1 (IETU1) and IE transcription unit 2 (IETU2). *BICP0* and *BICP4* are transcribed from a single promoter that lies in IETU 1,

while IETU 2 transcribes *BICP22* (Wirth *et al.* 1991) but recently it was found that instead of two transcription units there could be three units that independently transcribes all three major IE genes (Pokhriyal *et al.* 2015). Apart from this several important cellular regulatory factors have been shown to play crucial role in the activation of these transcription units. Octamer binding proteins (OCT 1) and CCAAT enhancer-binding protein alpha (C/EBPα) are such factors that activates IE genes transcription collaboratively with BHV-1 transinducing factor (b-TIF) (Huang and Herr, 1996; Kristie *et al.* 1989; Katan *et al.* 1990; Ludwig and Letchworth, 1987; Meyer and Jones, 2009).

Identification of promoter sequences and various regulatory motifs is crucial to understand the regulation of gene expression in BHV-1. Most of the research has been focused toward the expression studies of IE genes and very limited literature is available for early and late genes promoters of BHV-1. Hence, the present study was done using computational methods for the identification and analysis of putative promoter motifs of early and late genes of BHV-1 and compared them with the 3 known major promoters regions of IE genes (*BICP0, BICP4 and BICP22*) to look for similarities and differences among them.

MATERIALS AND METHODS

Retrieval of genome sequence

The complete genome sequences of BHV-1 were retrieved from biological database such as National Center for Biotechnology Information (NCBI) cited at http://www.ncbi.nlm.nih.gov/genomes/viruses.html.

Computational analysis of Promoters of IE, E and L genes

To delineate the promoter regions of IE genes, we first identified the TSS for IE, E and L genes. We had earlier identified transcription start site (TSS) of IE genes (*BICP0, BICP4* and *BICP22*) experimentally by 5'RLM-RACE (Pokhriyal *et al.* 2015). These genes had a single transcription start. The distance between TSS and putative translation start site (ATG) varied between the three genes. For early (*UL23, UL29, UL30, UL42* and *UL54*) and late genes (*gB, gC, gD*, and *gE*) TSS were computationally identified using Berkeley drosophila neural network promoters prediction online server (http://www.fruitfly.org/seq_tools/promoter.html).TSS search was performed at a stringent cutoff score of 0.85 to 0.90 out of 1.00). Based on the best possible score the TSS was selected. The identified TSS were mapped on the complete genomes of their respective genes. 1000 bp upstream and 100 bp downstream regions from the identified TSS were selected for further analysis of these promoter regions. PATCH online server (http://www.gene-regulation.com/cgi- 301 bin/pub/programs/patch/bin/patch.cgi) associated with TRANSFAC database was searched for specific regulatory motifs in the 1000bp upstream regions while YAPP online servers (http://www.bioinformatics.org/yapp/cgi-bin/yapp. cgi) were used to assess the downstream promoter elements.

Table 1: Distance of TSS from the ATG in BHV-1 genes

Genes	Temporal Class	TSS positions from the ATG
BICP0	Immediate early genes	235 bp
BICP4		156 bp
BICP22		92 bp
UL23	Early Genes	137 bp
UL29		277 bp
UL30		197 bp
UL42		59 bp
UL54		102 bp
gB	Late Genes	122 bp
gC		70 bp
gD		36 bp
gE		255 bp

RESULTS AND DISCUSSION

In this study we identified several important promoter elements both upstream regulatory elements and core (Basal) promoter elements of the IE, E and L genes in BHV-1. We here give a detailed account of comparison of these promoter regions.

Initiation of transcription of a gene is the first step in gene expression where RNA polymerase II (Pol II) plays significant role. It is recruited by various transcription

Table 2: Computational analysis of promoters of IE, E and L genes in BHV-1 (+ present; - Absent)

Genes	Temporal class	Tata BOX	OCT	C/EBPα	DPE	TATA + DPE
BICP 4	Immediate early genes	-	+	+	+	+
BICP 0		-	+	+	+	+
BICP 22		-	+	+	+	+
UL 30	Early Genes	-	-	+	+	-
UL 54		+	-	+	+	+
UL 29		-	-	+	+	-
UL 23		+	-	-	-	-
UL 42		+	-	+	+	+
gD	Late Genes	+	-	+	-	-
gE		-	-	+	-	-
gB		+	-	+	+	+
gC		+	-	-	+	+

factors (TFs) at the TSS within the preinitiation complex (PIC). Hence determining the where about of the TSSs has become crucial for mapping the regulatory elements and further studying the gene regulation machinery. (Wang *et al.* 2007)

We began our investigation by determining position of the experimentally identified TSS of *BICP0, BICP4* and *BICP22* in the complete genome of BHV-1. For *BICP0, BICP4* and *BICP22* the TSS was found to be placed 235 bp, 156 bp and 92 bp (excluding the intron region) respectively from their ATG. After several literature searches it remained unclear whether TSS for late and early genes have been experimentally determined, so we did computational identification of TSS for these genes. The distance of TSS from their respective ATG is given in table 1.

Transcription of promoters by RNA polymerase II requires involvement of various promoter elements and transcription factors. This interaction between RNA polymerase II and other factors regulates the efficiency of transcriptional initiation and also plays additional role of specifying the transcriptional initiation site (Blake *et al.* 1990). Considering this, we next embarked on to identify several important regulatory factors like C/EBPα, OCT-1, TATA box and DPE which assists in transcription of BHV-1 genes. First we computationally identified promoter elements which forms the core promoter of these genes (TATA box and DPE). The core promoter region is generally centered around the TSS, within a length of 100bp (Wang *et al.* 2007) but variations may exists. Several important core promoter elements have been reviewed by Butler *et al*. (Butler 2002). Analysis of core promoters gives insights into chain of events that lead to the initiation of transcription which eventually points at the core promoter region (Burke *et al.* 1998; Struhl 1987; Weis and Reinberg 1997).

Some important elements that make up core promoter are DPE (Downstream promoter elements), INR (initiator) and TATA box (Fig. 1). Computational analysis shows variation in the distribution of TATA box within different BHV-1genes (IE, E and L genes). TATA box were present in *UL54*, *UL23, UL42, gD, gB* and *gC* genes at -30, -30, -29, -31, -29 and -31 bp respectively from the TSS. No traces of TATA box were found in *UL30, UL29* and *gE* genes. In case of *BICP4, BICP0* and *BICP22* TATA box although present but was positioned far upstream from its optimum position (Fig. 2-4).

The optimum position of a TATA box is generally -26 to -31 bp from the TSS (Butler, 2002). Other important regulatory elements include the OCT-1 and C/EBP alpha motif. Oct-1 like sequences were absent from all the Early and Late genes despite extensive search with relaxed parameters (three base mismatch) (Fig. 3-4). The OCT-1 motifs were present in the all the genes of BICP group (Fig. 2). Some earlier research works have suggested that OCT-1 is required for the activation of BHV-1 IE gene

expression. It is known that BHV-1 transinducing factor (bTIF-a homologue of HSV α-TIF) forms a complex with OCT-1 at a specific motif (TAATGrAT – r is purine) which in turn activates expression of IE genes (Kristie *et al.* 1989; Misra *et al.* 1994, 1995). Recently Meyer and Jones have shown that OCT-1 is not the only cellular factor that can transactivate BHV-1 gene expression (Meyer and Jones, 2009). They have suggested role for CCAAT enhancer-binding protein alpha (C/EBPα) similar to OCT-1. Search for C/EBPα binding sites (CCAAT) found it to be present in multiple copies in almost all the BHV-1 genes (IE, E and L genes) except in *UL 29* and gC genes (Fig. 2-4).. The position of C/EBPα from TSS is depicted in Fig. 2-4.

While comparing all the IE, E and L genes we came across some interesting and new findings. To start with we found presence of OCT-1 only in IE genes and not in other E and L genes. This could make this factor exclusive to only IE genes. C/EBPα has been observed in all genes in multiple copies but their role for all these may differ. C/EBPα has crucial role in the activation of IE genes (Meyer and Jones, 2009).

Upon comparing the basal promoter region among IE, E and L genes, we found TATA box to be present far upstream from its usual optimum position (-26 to -31 bp) from the TSS in IE genes (*BICP0, BICP4 and BICP22*). This reduces the probability of these TATA sequences to guide gene expression.

In the absence of TATA box, role of DPE becomes important. The DPE is most commonly seen in promoters that lacks TATA box and in such cases DPE is required for promoter activity (Butler, 2002)

TATA and DPE have been observed in *UL42, gB,* and *gC*. Presence of both TATA and DPE is also interesting as presence of both these factors is rare. Their presence could have some impact on the expression that needs to be investigated. There is one gene which lacks TATA/ DPE elements and codes for minor glycoprotein *gE*. A comparative analysis of all the predicted factors is given in table 2. These findings have not been reported yet in Bovine Herpesvirus 1 and could prove to be insightful for further experimental investigation into the early and late genes. The predicted motifs can help in better understanding their role in gene regulation and transcription studies.

ACKNOWLEDGEMENTS

The authors thank Indian council of Agricultural Research (ICAR) for providing funding to all the authors for carrying out this work.

REFERENCES

Blake, M.C., Jambou, R.C., Swick, AG., Kahn, J.W. and Azizkhan, J.C. 1990. Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol. Cell. Biol.,* **10**: 6632–6641.

Burke, T.W., Willy, P.J., Kutach, A.K., Butler, J.E. andKadonaga, J.T. 1998. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold. Spring. Harb. Symp. Quant. Biol.,* **63**: 75–82.

Butler, J.E.F. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes. Dev.,* **16**: 2583–2592.

Huang, C.C. and Herr, W. 1996. Differential control of transcription by homologous homeodomain coregulators. *Mol. Cell. Biol.,***16**: 2967–76.

Katan, M., Haigh, A., Verrijzer, C.P., Van der vliet, P.C. and O'Hare, P. 1990. Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. *Nucleic. Acids. Res.,***18**: 6871–6880.

Kristie, T.M., Lebowitz, J.H. and Sharp, P.A. 1989. Complexes with the HSV α-TIF regulatory protein. *EMBO. J.,* **8**: 4229–4238.

Ludwig, G.V. and Letchworth, G.J. 1987. Temporal control of bovine herpesvirus 1 glycoprotein synthesis. *J. Virol.,* **61**: 3292–3294.

Meyer, F. and Jones, C.2009. The cellular transcription factor, CCAAT enhancer-binding protein alpha (C/EBP-alpha), has the potential to activate the bovine herpesvirus 1 immediate-early transcription unit 1 promoter. *J. Neurovirol.,* **15**: 123–30.

Misra, V., Blumenthal, R.M. and Babiuk, L.A. 1981. Proteins Specified by bovine herpesvirus 1 (infectious bovine rhinotracheitis virus). *J. Virol.,* **40**: 367–378.

Misra, V., Bratanich, A.C., Carpenter, D. and O'Hare, P.1994**.** Protein and DNA elements involved in transactivation of the promoter of the bovine herpesvirus (BHV) 1 IE-1 transcription unit by the BHV alpha gene trans-inducing factor. *J. Virol.,* **68**: 4898–909.

Misra, V., Walker, S., Hayes, S. and O'Hare, P. 1995. The bovine herpesvirus alpha gene trans-inducing factor activates

transcription by mechanisms different from those of its herpes simplex virus type 1 counterpart VP16. *J. Virol*., **69**(9): 5209-16.

Muylkens, B., Thiry, J., Kirten, P., Schynts, F. and Thiry, E. 2007. Bovine herpesvirus 1 infection and infectious bovine rhinotracheitis. *Vet. Res.,* **38**: 181–209.

Schlüter, J.P., Reinkensmeier, J., Barnett, M.J., Lang, C., Krol, E., Giegerich, R., Long, S.R. and Becker, A. 2013. Global mapping of transcription start sites and promoter motifs in the symbiotic alpha-proteobacterium Sinorhizobium meliloti 1021. *BMC. Genomics.,* **14**: 156.

Struhl, K. 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell.,* **49**: 295–297.

Wang, X., Bandyopadhyay, S., Xuan, Z., Zhao, X., Zhang, M.Q. and Zhang, X. 2007. Prediction of transcription start sites based on feature selection using AMOSA. *Comput Syst Bioinformatics. Conf.,* **6**: 183–93.

Weis, L. and Reinberg, D. 1997. Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. *Mol. Cell. Biol.,* **17**: 2973–2984.

Wirth, U.V., Gunkel, K., Engels, M. and Schwyzer, M. 1989. Spatial and temporal distribution of bovine herpesvirus 1 transcripts. *J. Virol.,* **63**: 4882–9.

Wirth, U., Vogt, B. and Schwyzer, M. 1991. The three major immediate-early transcripts of bovine herpesvirus 1 arise from two divergent and spliced transcription units. *J. Virol.,* **65**: 195–205.