Proc. Indian natn. Sci. Acad.(PINSA) **B66** Nos 2 & 3 pp 83-96 (2000)

Multiple Regulatory Circuits for Transcription from the Polyhedrin Gene Promoter of the *Autographa californica* Multiple Nucleocapsid Nuclear Polyhedrosis Virus

M D BASHYAM, A RAMACHANDRAN, S GHOSH, P VISWANATHAN and S E HASNAIN*
Eukaryotic Gene Expression Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067

(Received on 27 September 1999; Accepted after revision on 3 March 2000)

The baculovirus expression vector system has emerged as the system of choice for the expression of a number of heterologous genes of both prokaryotic and eukaryotic origin. This system utilizes the *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus (*Ac*MNPV) very late hyperactive polyhedrin gene (*polh*) promoter to drive the transcription of foreign genes. The transcription program of the virus is divided into a pre-replicative stage (early) and a post-replicative stage (late and very late) and occurs in a coordinately regulated and highly ordered cascade. In the very late phase, there is hyper-expression of the *polh* and *p10* genes. Late and very late genes are under the control of an α-amanitin and tagetitoxin-resistant RNA polymerase that is induced during infection of the insect host. Despite the immense popularity of the baculovirus expression system, little is known about mechanisms of basal and hyperactivated transcription from the *polh* promoter. This review aims at summarizing the recent developments in understanding the *cis* and *trans* acting elements involved in transcriptional regulation from the *polh* promoter with particular emphasis on the role of insect cell host factors.

Key Words: *Ac*MNPV, PPBP, Sp1

Introduction

The baculovirus expression vector system is arguably the most favored system for the expression of foreign genes. Several features of the baculovirus system are particularly advantageous viz. eukaryotic environment for proper folding, disulfide bond formation, oligomerization, glycosylation, fatty acid acylation, phosphorylation and other post translational modifications; exceptionally high levels of expression; a capacity for large insertions; efficient expression of unspliced genes etc. (Luckow & Summers 1988, Kidd & Emery 1993, Sridhar et al. 1994, Hasnain et al. 1997b, Sudhakar et al. 1999 and Puri et al. 1999). Scale up of insect cells in culture has also been largely perfected, making purification of large quantities of proteins a reality (Van Lier et al. 1992). It has also been possible to express multiple genes simultaneously (Jha et al. 1992, Chatterjee et al. 1996). The biology of the infection cycle of the baculovirus in the insect cell underlies the utility of baculoviruses as expression vectors. The polh gene is non-essential for viral DNA replication

* *Centre for DNA Fingerprinting and Diagnostics, Laboratory Building, CDFD Interim Complex, 7-18, E & F, Nacharam, Hyderabad 500076*

Abbreviations: *AcMNPV, Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus; *IBP*, initiator binding protein; *lef*, late expression factor gene; *ori*, origin of DNA replication; *polh*, gene encoding the polyhedrin protein; PPBP, *polh* promoter-binding protein; *Sf*, *Spodoptera frugiperda*; TSP, transcription start point; *hr*, homologous region sequence(s); *hr1*-BP, *hr1*-binding protein.

and production of infectious virus particles in cell culture, but still is expressed in very high levels during the final phase of infection. The polyhedrin protein is required for generation of occlusion bodies which help the virus to infect a new insect host, but is not required for formation of budded form of the virus that is capable of infecting insect cells in cell culture. The basic theme of the baculovirus expression system is to generate a *polh*⁻ recombinant virus in which the *polh* gene is replaced by the desired foreign gene to be expressed and is placed directly downstream of the hyperactive *polh* promoter. Therefore, recombinant viruses are easily identified by their occlusion negative phenotype (occ⁻). Since the occ⁻ recombinant viruses cannot infect fresh insect larvae efficiently, this expression system also becomes safe from the recombinant DNA safety perspectives. This technique has been perfected to such an extent that with most genes, more than 90% efficiency is achieved. Although, *polh* and *p10* promoters are widely used in the expression system, little is known about the mechanism of basal and hyperactivated transcription from these very late promoters. Hence, the study of cis-acting elements in the very late promoter and trans-acting host and viral proteins interacting with the promoter assumes importance. Elucidation of the complete regulatory components of the *polh* gene shall make it feasible to modulate levels of expression of genes that may need tc be expressed at a lower level. This analysis would also facilitate the development of a cell free gene expression system which has the obvious advantage in ease of purification of the expressed proteins and would be devoid of problems such as secretory load most often encountered in the baculovirus expression system (Nakhai et al. 1991, Sridhar et al. 1993). Cognition of how the virus utilizes host factors for expression of its own genes may provide insights into host-virus interactions in case of human pathogenic viruses. It has been found that cloning of genes in baculo-mediated gene transfer systems for mammalian cells is hindered by presence of the strong *polh* promoter (Barsoum et al. 1997). Therefore, study of regulation of the *polh* promoter would also facilitate generation of constructs to enable cloning of genes during generation of baculo-mediated gene delivery systems. Several reports have implicated the role of viral proteins including a four subunit virally-induced RNA polymerase for transcription from the *polh* promoter. A number of other findings however, suggest an important role for host factors as well. This review aims at providing an update on the recent developments in understanding transcriptional regulation from the *polh* promoter, with particular emphasis on role of host factors in basal and hyperactivated transcription.

A Cascade of Viral Gene Expression

*Ac*MNPV is an insect virus with a double stranded circular, covalently closed DNA genome of 134 kb. The infection cycle of the baculovirus occurs in three phases: early, late and very late and is regulated by a cascade of gene expression (figure 1). The early phase of the life cycle is responsible for programming the virus for DNA replication and subsequently for production of viral progeny. This involves expression of viral genes required for DNA replication and for transcription of late and very late genes. Early genes are transcribed by the host RNA polymerase II, and transcription from the early promoters is abolished in the presence of α-amanitin (Huh & Weaver 1990). The early baculovirus promoters resemble typical eukaryotic RNA polymerase II responsive promoters (Krappa et al. 1992, Kogan & Blissard 1994). Most early

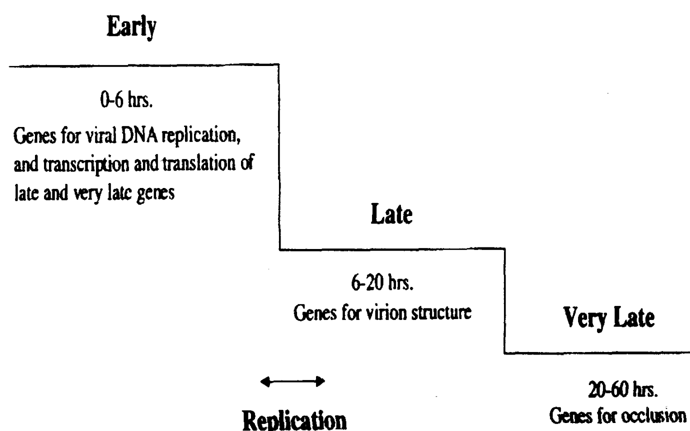

Figure 1 A diagrammatic representation of cascade of gene expression undertaken by the baculovirus *Ac*MNPV upon infection of *Spodoptera frugiperda*.

promoters contain a TATA box, which positions the RNA start site downstream (Dickson & Friesen 1991 and Blissard et al. 1992). Early promoters also contain various *cis*-acting DNA elements present upstream which ensure optimal expression of early genes (Carson et al. 1991, Krappa et al. 1992, and Kogan & Blissard 1994). Four genes of *AcMNPV* viz. *ie-0, ie-1, ie-2,* and *pe-38,* are known to transactivate early baculovirus promoters in transient assays (Reibero et al. 1994). Transcription of the *ie-1* gene is initiated by a CAGT element (Chisholm & Henner 1988, Kovaes et al. 1992). The CAGT element has been shown to function as a typical initiator element (Pullen & Freisen 1995a) and is the first example of a baculovirus initiator element. Using the *cat*-reporter gene, it was shown that maximum activity of the *ie-1* promoter was conferred by sequences that extended from the TATA element through the 5' non-coding leader region (Pullen & Freisen, 1995b). IE-0 is identical to IE-1, but contains extra 543 N-terminal amino acids (Chisholm & Henner 1988). *ie-2* activates *ie-1* promoter approximately 2.5 fold in transient expression assays (Yoo & Guarino 1994a,b). IE2 also blocks cell cycle progression in several cell lines including *Spodoptera frugiperda* 21 (*Sf*21) and *Trichoplusia ni* (Yoo & Guarino 1994b). IE2 in the presence of IE-1 and six other *AcMNPV* genes, augments replication and stability of plasmids bearing Hr (homologous region) sequences (Lu & Carsten 1993). The Hrs act as enhancers for some early promoters in transient expression assays (Guarino & Summers 1987 and Nissen & Friesen 1989). It is possible that some Hrs may play a role in initiation of viral DNA replication. The transcriptional activities of *pe-38* and *ie-2,* that are divergently transcribed, are similar during the course of infection (Krappa et al. 1995). Another early gene *he-65* is transcribed subsequent to the four major early genes (i.e. *ie-1, ie-0, ie-2,* and *pe-38*) (Kremer & Knebel Morsdorf 1998). The *he-65* promoter lies in close proximity to *hr-4a* (Becker & Knebel-Morsdof 1993). Recently, it was shown that *ie-1* can activate the *he-65* promoter in insect *Trichoplusia ni* 368 and mammalian BHK21 cells (Murges et al. 1997). The *p35* gene is another essential baculovirus early gene and its promoter is composed of a TATA element present 30 bp upstream of the transcription start site. The presence of an upstream activating region extending from -30 to -90 with respect to the transcription start site, ensures optimal expression of the *p35* gene. Recent studies have identified another activator of baculovirus early gene expression viz. *orf121. orf121* activates *39k* gene in transient expression studies and this activation is dependent on cotransfection of *ie-1* (Gong et al. 1998). *orf121* also increased *ie-1* expression and therefore *orf121* mediated-stimulation of *39k* gene expression was by an upregulation of *ie-1* gene expression (Gong et al. 1998). Finally, the *39k* promoter is transactivated by viral protein IE-1 and further stimulated by viral enhancer elements and the viral coactivator IE-2.

Early phase is followed by the late phase which includes viral DNA replication, late gene expression and production of the budded form of the baculovirus. The late phase extends from 6 hours postinfection to approximately about 18-24 hr post infection. During this phase, progeny nucleocapsids leave the nucleus and travel to the cytoplasm and usually bud individually from the membrane. Late genes mainly include structural proteins and are under the control of an α-amanitin and tagetitoxin-resistant RNA polymerase that is induced during infection of the insect host (Huh & Weaver 1990, Grula et al. 1981, Fuchs et al. 1983, Glocker et al. 1993). The primary determinant of late promoter activity is the tetranucleotide TAAG, which encompasses the transcriptional start point (Rhormann 1986, Thiem & Miller 1989). The TAAG motif is relatively rare in the baculovirus genome and is found primarily in late and very late promoters. Recently, several laboratories have focused their efforts to unravel the subunit composition of the novel RNA polymerase. Partial purification of the polymerase activity suggested that its protein composition may be different from that of the three known host RNA polymerases (Yang et al. 1991). Nuclear extracts prepared from infected cells were shown to contain late promoter specific RNA polymerase activity as judged by nuclear

run off (Glocker et al. 1993) or primer extension (Xu et al. 1995) analysis. Subsequently, a virus-encoded RNA polymerase was purified from infected insect cells using DEAE-sephadex and Heparin-agarose chromatography, and glycerol gradient centrifugation. SDS-PAGE analysis of the RNA polymerase preparation showed the presence of 9 polypeptides (Beniya et al. 1996). Guarino and coworkers have recently claimed to have purified the viral RNA polymerase to homogeneity. It consists of four subunits, all of viral origin, viz. LEF-4, LEF-8, LEF-9 and p47 (Guarino et al. 1998b). This is the simplest form of eukaryotic RNA polymerase identified so far (Guarino et al. 1998b). The subunit p47 has also been shown to contain guanylyl transferase, RNA triphosphatase, and ATPase activities (Jin et al. 1998, Gross & Shuman 1998 and Guarino et al. 1998a) thus indicating a close relationship between capping and transcription initiation as has been reported in other eukaryotic systems (Ho et al. 1998, Cho et al. 1997, McCracken et al. 1997). The LEF-8 and LEF-9 subunits contain typical RNA polymerase signatures (Passarelli et al. 1994, Lu & Miller 1994).

Eighteen *Ac*MNPV genes have been implicated, by transient expression in co-transfections, to have regulatory roles in transcription from the *polyhedrin* gene promoter (Todd et al. 1995). These are called late gene expression factors (LEFs). Nine of these genes have indirect effect on *polyhedrin* transcription as they are involved in viral DNA replication (Lu & Miller 1995). Four of the rest are components of the viral RNA polymerase (Guarino et al. 1998b). No concrete evidence exists for assigning specific roles to rest of the genes. Many of these genes may be affecting processes upstream to late gene activation, given the cascade nature of viral gene activation (figure 1). Recently, a nineteenth *lef* gene has been identified. A library of 18 lefs under the control of HSP 70 promoter of Drosophila was unable to support transient expression of a viral late promoter in transfection studies (Rapp et al. 1998). This led to the identification of a new lef viz. *lef-12*, located adjacent to the *p47* gene. It is predicted to encode a 21 kDa protein with no significant homology to known proteins in the data base. A set of 19 orfs (including *lef-12*) under the control of the HSP 70 promoter of Drosophila (HSEpiHis *lef* library) could support transient expression from a late viral promoter (Rapp et al. 1998). Recently, Li and coworkers have also shown that the 18 lefs were unable to support transient late gene expression (Li et al. 1999). Their study indicated that *orf41* and *orf69* were also involved in late gene expression (Li et al. 1999). All the factors mentioned in this section are important for transcription of late as well as very late genes.

The *polyhedrin* or *p29* Gene (*polh*)

The very late phase which begins after approximately 20 hr post infection, is the occlusion-specific phase and is non-essential for propagation of the virus in cell culture. The nucleocapsids become enveloped in newly synthesized membrane-like structures, either individually (Single Nucleocapsid Polyhedrosis Virus -- SNPV) or in groups (MNPV) (Bassemir et al. 1983). The envelopment is followed by embodiment within the polyhedrin matrix of an occlusion body. The two major proteins synthesized exclusively during the very late phase are the polyhedrin and p10. As already stated, the *polh* gene is essential for formation of the occluded virus and is transcribed to very high levels during the very late phase of infection. The requirements for transcription of a very late gene is similar to that of a late gene. The *polh* promoter harbours all the essential characteristics of a late/very late promoter. It however exhibits low activity during late phase but becomes highly active beginning about 18 hr post infection and ultimately at about 30 hr post infection, 20% of the total polyadenylated RNA in the cell is *polh* mRNA (Adang & Miller 1982). The temporally regulated and hyperactivated polyhedrin (*polh*) gene promoter of *Ac*MNPV belongs to the class of initiator promoters (Smale & Baltimore 1989). An initiator promoter is the simplest form of a eukaryotic pol II promoter and consists of a core tetranucleotide motif, CAGT, which encompasses the transcriptional start point (TSP). Several initiator containing promoters have been

identified (Chen et al. 1994, Yoo et al. 1991). An analysis of 80 random and mutant initiator elements identified a loose consensus sequence (Py Py A+1 N T/A Py Py) for such initiator promoters (Javahery et al. 1994). Although a weak TATA box-like sequence is present about 30 bp upstream of the transcription start point (TSP), deletion and mutation analysis have shown that an 18 bp region encompassing the TSP is the minimal *polh* promoter containing all the essential *cis*-acting elements for transcription initiation (Morris & Miller 1994). The 18 bp sequence surrounding the *polh* transcription initiation point, represents an initiator-like sequence. The sequences encoding the untranslated mRNA leader may be important for the very late burst of expression (Morris & Miller 1994). There are two important motifs in the *polyhedrin* basal promoter, a hexa motif (AATAAA) present immediately upstream of the TSP, and also in the untranslated leader sequence, and an octa motif **TAAG**TATT, which includes the TSP. The hexa motif present in the very late *p10* promoter, has the sequence GTTAAA instead of AATAAA. The octa motif in *p10* is **TAAG**AATT instead of **TAAG**TATT. As already mentioned, the **TAAG** motif is a characteristic feature of all baculovirus late and very late promoters and is the primary determinant of promoter activity (Rohrmann 1986, Possee & Howard 1987, Thiem & Miller 1989). Linker scan substitutions affecting the TAAG sequence decrease reporter gene expression approximately 2000-fold and lower RNA levels to undetectable levels (Ooi et al. 1989). Mutations downstream of this site did not affect transcription (Ooi et al. 1989). All strong late and very late RNAs have been observed to initiate from A**TAAG**, but substitution of this 5' A immediately upstream of TAAG to T or G have been shown to reduce the level of transcription approximately three fold (Ooi et al. 1989). Insertion of linkers further upstream of TAAG sequence (10-20 nucleotides) increased the level of transcription (Rankin et al. 1988, Ooi et al. 1989).

There is only one factor identified till date that exclusively activates very late gene expression viz. VLF-1 (McLachlin & Miller 1994). *vlf-1* was identified while characterizing a mutant (*ts*B837) of *Ac*MNPV which is temperature sensitive for occluded virus production at the non-permissive temperature. This mutant exhibited reduced transcription of the very late genes like *polh* and *p10* but not of late genes such as for the 603 ORF and *vp39* (McLachlin & Miller 1994). VLF1 has significant homology with the family of integrases and resolvases (McLachlin & Miller 1994). It has been shown to be essential for viral DNA replication and has similarity to lambda phage integrase family (McLachlin & Miller 1994, Yang & Miller 1998). It has also been shown that transactivation from the *polh* promoter does not require the integrase activity of VLF-1 (Yang & Miller 1998). Recently, linker scan mutations in the *polh* and *p10* promoters which abolished or weakened the ability of the promoters to respond to stimulation by VLF-1 have been identified (Yang & Miller 1999). These mutations fall in the previously identified "burst" sequence. Addition of partially purified epitope tagged VLF-1 to DNA encompassing the burst sequence resulted in a shift in electrophoretic mobility shift assay (EMSA) (Yang & Miller 1999). Addition of antibody raised against VLF-1 resulted in a supershift of the VLF-1 DNA complex (Yang & Miller 1999). The binding was further confirmed by DNaseI footprinting studies and linker scan mutations within the binding sequence severely impaired the binding (Yang & Miller 1999). This suggests that VLF-1 transactivates the *polh* promoter by binding to the burst sequence. However, it is doubtful that VLF-1 is solely responsible for the switch from late to very late gene expression.

Role of Host Factors in *polh* Regulation

Inspite of the fact that a very large number of viraly-encoded *lefs* have been identified, none of them have been shown to interact directly with late or very late gene promoters. It is therefore necessary to identify other factors including those of host origin which directly interact with the very late gene promoters and through a possible protein-protein interaction with one or more of the LEFs, are directly responsible for transcription of these genes. Reports on host protein factors important for late and very late gene expression

are limited. It is expected that the four putative RNA polymerase subunits would be recruited by the promoter, but it is not known whether one or more of the four subunits interact directly with the promoter or they are recruited by means of another factor binding directly to the promoter. The second situation is analogous to classical eukaryotic transcription, a positioning or commitment factor (TBP plus TAFs) binds directly to the promoter region and subsequently recruits the RNA polymerase. Therefore, there could be four candidates for direct promoter binding (1) One or more of the RNA pol subunits, (2) another viral factor, (3) TBP or (4) another host factor. The simplest way to answer this question would be to check for direct binding of proteins to the labeled *polh* promoter fragment in EMSA. Such studies have revealed a novel host protein called PPBP (polyhedrin promoter binding protein) that binds directly to the *polh* promoter (Burma et al. 1994).

(i) Characteristics of polh promoter binding by PPBP

Extensive EMSA experiments on synthetic oligonucleotides representing normal as well as mutated versions of several regions of the *polh* promoter have revealed that the hexa-octa motif (AATAAATAAGTATT), which also includes the transcription start site, is the binding sequence for PPBP (Burma et al. 1994). Characterization of this DNA-protein complex revealed that divalent cations were not required for PPBP binding and the PPBP-DNA complex could tolerate a wide range of salt concentrations (Burma et al. 1994). PPBP could also tolerate a 1000-fold excess of polydI-dC and high temperatures upto 65°C. These observations indicate that PPBP is an unusually stable protein (with a high affinity for the *polh* promoter) (Burma et al. 1994). The apparent K_d value was estimated to be 3.7 (+ 0.5) pM. Using UV-crosslinking and south-western blotting analysis the molecular weight of PPBP was found to be about 30 kDa. Other experiments revealed that phosphorylation of PPBP was essential for binding to the promoter (Burma et al. 1994). Similar results were obtained when identical experiments were carried out using the *p10* promoter indicating that there may be a common program for hyperactivation from baculovirus very late promoters. To date PPBP is the only factor described that binds directly to the *polh* promoter. These characteristics make PPBP an unusual DNA-binding protein, perhaps reflecting on the very high expression levels obtained with the *polh* promoter. The fact that PPBP binds to the hexa and octa motifs is consistent with previous observations that showed the importance of sequences present around the TSP (Possee & Howard 1987, Rankin et al. 1988, and Ooi et al. 1989). Examples of DNA binding proteins where phosphorylation is critical for binding is limited (Hunter & Karin 1992). It is possible that phosphorylation helps to keep its DNA binding domain in the correct conformation.

Since PPBP binds to the region that includes the TSP, it is interesting to speculate about the status of its interaction with the promoter after DNA melting. Using EMSAs, it was shown that PPBP could bind to the coding strand of the promoter but not to the non-coding strand (Mukherjee et al. 1995). As seen with double strand binding, both hexa and octa mutations abolished binding. However, unlike double strand binding, PPBP binding to the coding strand may involve ionic interactions. Divalent cations are not required for binding to the coding strand as was the case with binding to duplex promoter DNA (Mukherjee et al. 1995). PPBP displays stronger affinity for the coding strand as compared to the double strand and PPBP-coding strand complex has a longer half life than the PPBP-duplex promoter complex (Mukherjee et al. 1995). The unique structure of initiator promoters ensures a strict regulation of activation and inactivation during cellular differentiation (Smale & Baltimore 1989). Although several initiator binding proteins have been identified (Weis & Reinberg 1992), little is known about the pathway of assembly of active transcription complexes. It is tempting to speculate that initiator binding proteins in other systems may also exhibit such dual binding specificities. Some examples of transcription factors with ss DNA binding ability include the estrogen receptor

which selectively binds to the coding strand of an estrogen receptor element (Lannigan & Notides 1989), muscle factor 3 and MyoD interact with double and ss forms of muscle gene elements (Santoro et al. 1991), sterol regulatory element-binding factor binds to both double and single strand forms of the sterol regulatory element (Stark et al. 1992) etc.

Further experiments revealed that like TBP, PPBP also binds to the minor groove (Ghosh et al. 1998). However, it has been shown convincingly that PPBP and TBP are distinct proteins and that TBP may not have a direct role in *polh* promoter-driven transcription with respect to its ability to directly contact the *polh* promoter (Ghosh et al. 1998).

ii) Role of PPBP in transcription from the polh promoter

The direct role of PPBP in transcription initiation from the *polh* promoter was confirmed by the detection of PPBP in transcription permissive phosphocellulose fractions of nuclear extracts but not in non-permissive fractions. Moreover, using *in vitro* transcription reactions, it was shown that yield of specific transcripts from the *polh* promoter was considerably reduced when PPBP was sequestered by using the B-domain oligo, which was restored when transcription extracts containing PPBP activity was added (Ghosh et al. 1998). Therefore PPBP was not only present in the transcription permissive fraction of the nuclear extract but was recruited during the assembly of transcription complex on the *polh* promoter. Moreover, using transfection studies, it was shown that reporter gene expression driven by the *polh* promoter was reduced by 50% by *in vivo* mopping (Habib & Hasnain 1996, Habib et al. 1996) effected by cotransfection with plasmid containing the PPBP cognate motif (Ghosh et al. 1998). A cartoon depicting the *in vivo* strategy is depicted in figure 2. Mutation of the PPBP cognate motif abolished reporter gene expression indicating the importance of PPBP binding to the *polh* promoter for transcription (Ghosh et al. 1998). Purification through affinity chromatography of PPBP revealed a 30 kDa protein, although several times, three other proteins were also co-purified. Therefore, the initial step in transcription from the *polh* promoter could be binding of PPBP to the core

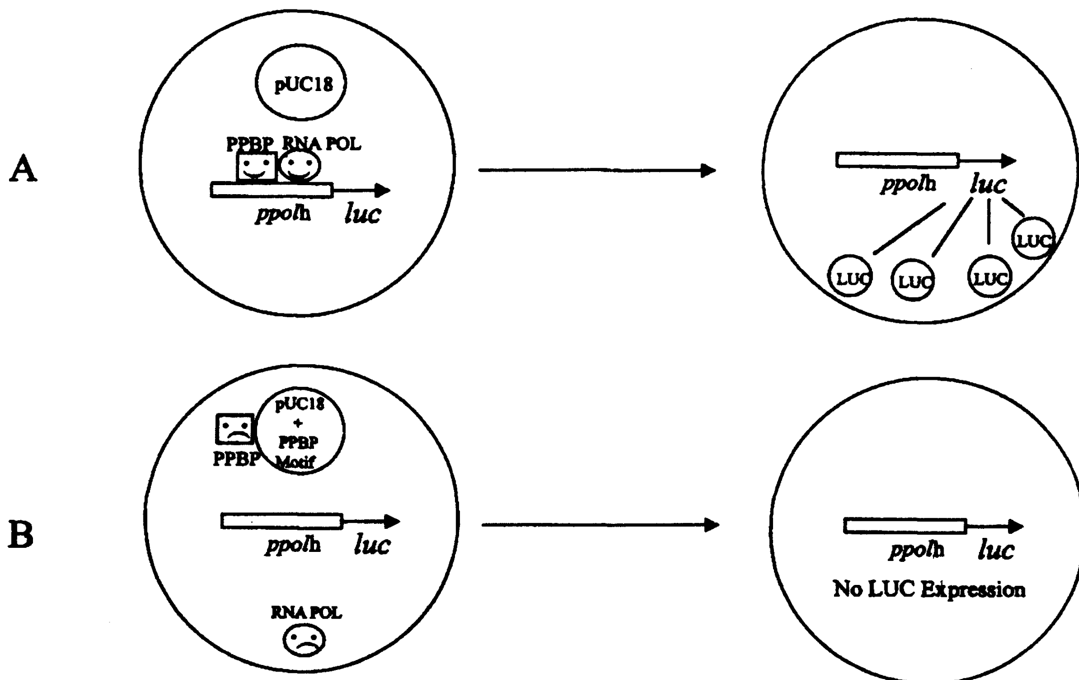

Figure 2 Cartoon depicting the *in vivo* strategy used to confirm the role of PPBP in transcription initiation from the *polh* promoter. The control plasmid pUC18 when cotransfected into *Sf9* cells along with a plasmid containing the PPBP recognition element cloned upstream of the *polh* promoter-luc cassette, does not effect reporter gene expression (panel A). However, when the contransfected pUC18 plasmid also contains the PPBP recognition element, it leads to abrogation of reporter gene expression (panel B).

promoter followed by the recruitment of the viral RNA polymerase (directly or indirectly). Once transcription is initiated and strand melting takes place, PPBP might switch over to the single strand binding mode, thus facilitating repeated rounds of transcription. The observation of ATPase and a helicase activity in partially purified PPBP fraction (Ghosh et al. unpublished data) indicates that other factors may interact with PPBP and actually be involved in strand melting, similar to the basal transcription factor TFIIH (Hoeijmakers et al. 1993).

Another host factor of about 200 kDa has been found to interact with the sequence located between -72 to -86 bp relative to *polh* translational initiation site (Etkin et al. 1994). Specific DNA binding activity was found to decrease with progression of infection. *In vivo* competition of this protein in transient expression assay increased transcription from the polyhedrin promoter implying its negative role in *polh* transcription. Again mutation affecting this region can upregulate polyhedrin expression by 50 % very late in infection (Ooi et al. 1989). The *polh* promoter lacking this region exhibited strong promoter activity (Thiem & Miller 1990). However, this region does not appear to play a major role in regulation of the *polh* promoter (Thiem & Miller 1990).

Bal31 deletions have revealed an essential role of the 4 kb upstream region in transcription from the *polh* promoter (unpulished data). This analysis has identified two regions, region 1 and region 2, spanning map units 0 to 1.5 and 2.5 to 3.12 respectively on the *EcoRI* "I" fragment of the viral genome. Region 2, which contains about 770 bp upstream of the basal promoter, consistently showed about 2-3 fold enhancement in expression of reporter gene in transient expression assays (unpulished data). Moreover, region 2 cloned upstream of a *polh* promoter containing mutated hexa motif, was able to partially rescue reporter gene expression (unpulished data). Analysis of this region revealed sequences resembling cognate Sp binding motifs and EMSAs performed on oligonucleotides representing these sequences called *AcSp*) have identified probable host Sp-like factors(unpublished data). UV cross-linking experiments demonstrated the presence of two proteins of 46 kDa and 90 kDa in the complexes. The *AcSp* oligonucleotide could also enhance reporter gene expression in transient expression studies indicating that this sequence was an important constituent of region 2. Binding of the insect Sp-family proteins and their role in transcription from the *polh* promoter was confirmed by *in vivo* mopping experiments. Moreover, the *AcSp* oligo could compete for TBP binding to a consensus TFIID oligo and vice versa in EMSAs. This could be due to the fact that Sp1 is known to interact with hTAFII130 and dTAFII110, so it is able to indirectly compete for TBP binding to a consensus TFIID oligo and vice versa. These results indicate for the first time the presence of Sp family proteins in insect cells. The importance of Sp-mediated activation of transcription is especially seen in promoters having weak or null affinity for TBP, and therefore assumes importance in case of initiator promoters. The role of insect Sp-family proteins in transcription from the *polh* promoter has recently been confirmed by using recombinant viruses (unpulished data). Recombinant viruses were constructed containing the *AcSp* oligonucleotide cloned upstream of the *polh-luc* cassette. The constructs are depicted in figure 3. The recombinant virus construct containing the *AcSp* oligonucleotide exhibited increase in reporter gene expression compared to a recombinant virus containing just the *polh* basal promoter cloned upstream of the *luc* reporter gene (figure 3; and Jain et al. submitted). These results confirm the role of the Sp-binding motifs present in the upstream sequence element (region 2) of the *polh* promoter in hyperactivation from the promoter.

The role of hr1, present 4 kb upstream of the *polh* promoter, has also been investigated. The hr1 was shown to function as an enhancer for transcription from the *polh* promoter in classical transient enhancer assays (Habib et al. 1996). The enhancement was confirmed by RNase protection assays and hr1 could enhance expression in either orientation and also when placed downstream from the reporter gene (Habib et al. 1996). The transcription in the above case initiated from the authentic TSP of the *polh* gene and was insensitive to α-amanitin

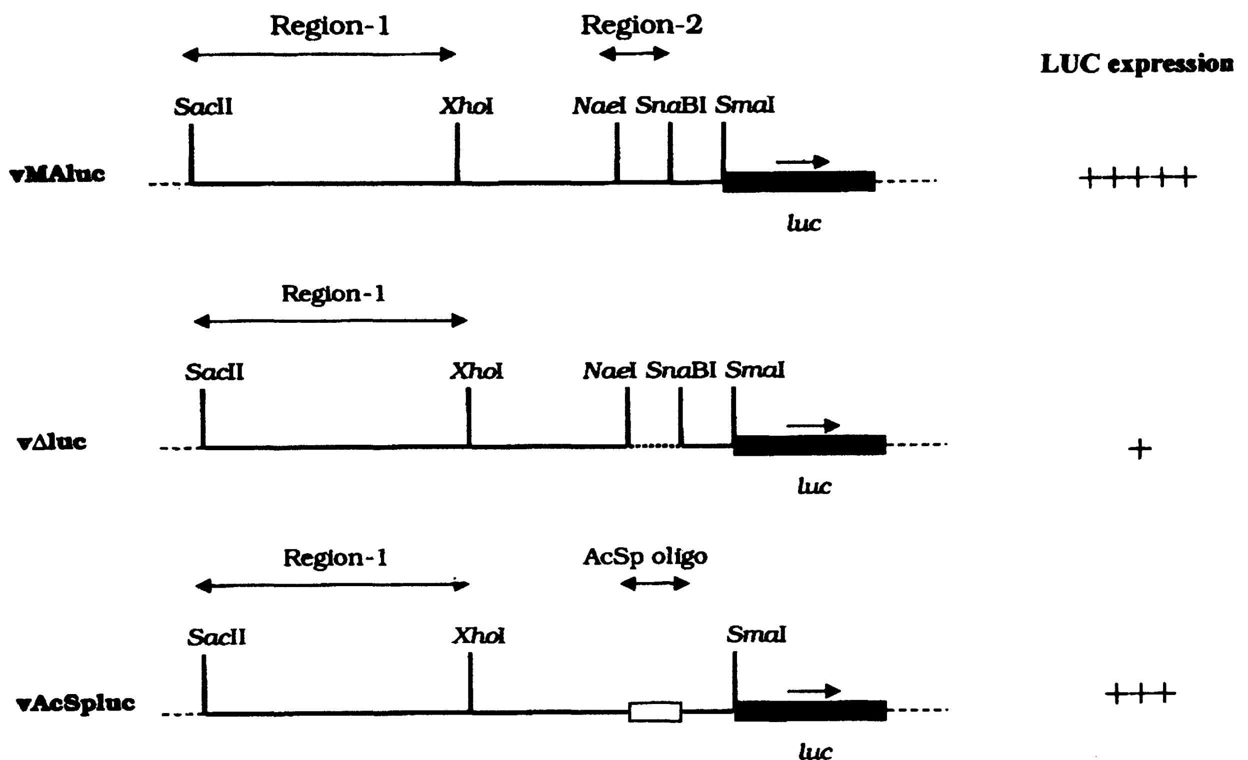

Figure 3 A diagrammatic representation of the recombinant virus constructs made to confirm the role of the *AcSp* sequence in augmenting transcription from the *polh* promoter in the viral context. vMAluc is the wild type virus containing the *luc* gene. vdeltaluc is same as vMAluc but contains a deletion of about 700 bp (region 2) upstream of the *polh* promoter. This deletion removes the *AcSp* sequences. v*AcSp*luc contains the *AcSp* oligo cloned in lieu of the deleted region 2.

indicating that hr1-mediated enhancement followed the typical profile of baculovirus very late gene transcription (Habib et al. 1996). Hr1 has also been described as a putative viral origin of replication, however the enhancer function of hr1 was found to be distinct from the origin function (Habib et al. 1996 and Habib & Hasnain 1997). An hr1 binding host protein (hr1-BP), that binds to multiple sites within the hr1 enhancer element, has been identified (Habib & Hasnain 1996). Also, this binding is important for the enhancer function because sequestration of this protein using a plasmid bearing its binding motif abolishes hr1-mediated enhancement of transcription of a reporter gene under hr1 sequence (Habib & Hasnain 1996).

Therefore, these results show a major role for insect cell host factors (PPBP, Sp-family, and hr1-BP) in regulating transcription from baculovirus very late promoters (Hasnain et al. 1996, 1997a). There are several areas of *polh* transcription still to be worked out. For example, it is still not known what switch triggers the activation of viral RNA polymerase after completion of DNA replication or how the RNA polymerase distinguishes very late promoters from late promoters. It is possible that VLF-1 and PPBP play a major role in this regard. Another area of study could be the elucidation of role of promoter architecture in *polh* transcription. Studies by Mans and Knebel-Morsdorf have revealed the presence of inverted repeats between the transcription start site and the translation initiation codon in the *polh* gene (Mans & Knebel-Morsdorf 1998). This sequence is not conserved in the *polh* genes of different baculoviruses but the repeat *per se* is conserved (Mans & Knebel-Morsdorf 1998). It would be worthwhile to determine how PPBP interacts with the viral RNA polymerase to direct transcription from the *polh* promoter and what is the nature of interaction between the Sp family proteins and the transcription initiation complex. Similarly, role of the region 1 in hyperactivation of transcription from the *polh* gene remains to be elucidated.

References

Adang M J and Miller L K 1982 Molecular cloning of DNA complementary to mRNA of the baculovirus *Autographa californica* nuclear polyhedrosis virus: location and gene products of RNA transcripts found late in infection; *J. Virol.* **44** 782-793

Barsoum J, Brown R, McKee M and Boyce F M 1997 Efficient transduction of mammalian cells by a recombinant baculovirus having the vesicular stomatitis virus G glycoprotein; *Hum. Gene Ther.* **20** 2011-2018

Bassemir U, Miltenburger H and David P 1983 Morphogenesis of nuclear polyhedrosis virus from *Autographa californica* in a cell line from *Mamestra brassicae* (cabbage moth); *Cell Tissue Res.* **228** 587-595

Becker D and Knebel-Morsdorf D 1993 Sequence and temporal appearance of the early transcribed baculovirus gene HE65; *J. Virol.* **67** 5867-5872

Blissard G W, Kogan P H, Wei R and Rohrmann G F 1992 A synthetic early promoter from a baculovirus: roles of the TATA box and conserved start site CAGT sequence in basal levels of transcription; *Virology* **190** 783-793

Beniya H, Funk C J, Rohrmann G F and Weaver R F 1996 Purification of a virus-induced RNA polymerase from *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells that accurately initiates late and very late transcription *in vitro.*; *Virology* **216** 12-19

Burma S, Mukherjee B, Jain A, Habib S and Hasnain S E 1994 An unusual 30 kDa protein binding to the polyhedrin gene promoter of *Autographa californica* nuclear polyhedrosis virus; *J. Biol. Chem.* **269** 2750-2757

Carson D D, Summers M D and Guarino L A 1991 Molecular analysis of a baculovirus regulatory gene; *Virology* **182** 279-286

Chatterji U, Ahmad R, Venkaiah B and Hasnain S E 1996 A recombination-efficient baculovirus vector for simultaneous expression of multiple genes; *Gene* **171** 209-213

Chen H, Vinnakota R and Flint S J 1994 Intragenic activating and repressing elements control transcription from the adenovirus IVa2 initiator; *Mol. Cell. Biol.* **14** 676-685

Chisholm G E and Henner D J 1988 Multiple early transcripts and splicing of the *Autographa californica* nuclear polyhedrosis virus IE1 gene; *J. Virol.* **66** 3193-3200

Cho E, Takagi J T, Moore C R and Buratowski S 1997 mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain; *Genes dev.* **11** 3319-3326

Dickson J A and Friesen P D 1991 Identification of upstream promoter elements mediating early transcription from the 35,000-molecular-weight protein gene of *Autographa californica* nuclear polyhedrosis virus; *J. Virology* **65** 4006-4016

Etkin E, Carb-Seiss L and Levi B Z 1994 *Spodoptera frugiperda Sf-9* cells nuclear factor binds to a specific sequence on the baculovirus polyhedrin gene promoter; *Virus Res.* **31** 343-356

Fuchs L Y, Woods M S and Weaver R F 1983 Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells; *J. Virol.* **66** 3476-3484

Ghosh S, Jain A., Mukherjee B, Habib S and Hasnain S 1998 The host factor polyhedrin promoter binding protein (PPBP) is involved in transcription from the baculovirus polyhedrin gene promoter; *J. Virol.* **72** 7484-7493

Glocker B, Hoopes R R, Hodges L and Rohrmann G .F 1993 *In vitro* transcription from baculovirus late gene promoters: accurate mRNA initiation by nuclear extracts prepared from infected *Spodoptera frugiperda* cells; *J. Virol.* **67** 3771-3776

Gong M, Jin J, Guarino L A 1998 Mapping of ORF121, a factor that activates baculovirus early gene expression; *Virology* **244** 495-503

Gross C H and Shuman S 1998 RNA 5'-triphosphatase, nucleoside triphosphatase, and guanylyltransferase activities of baculovirus LEF-4 protein; *J. Virol.* **72** 10020-10028

Grula M A, Buller P L and Weaver R F1981 α-Amanitin-resistant viral RNA synthesis in nuclei isolated from nuclear polyhedrosis virus-infected *Heliothis zea* larvae and *Spodoptera frugiperda* cells; *J. Virol.* **38** 16-21

Guarino L A, Jin J and Dong W 1998 Guanylyltransferase activity of the LEF-4 subunit of baculovirus RNA polymerase; *J. Virol.* **72** 10003-10010

______ and Summers M D 1987 Nucleotide sequence and temporal expression of a baculovirus regulatory gene; *J. Virol.* **61** 2091-2099

Guarino L A, Xu B, Jin J and Dong W 1998 A virus-encoded RNA polymerase purified from baculovirus-infected cells; *J. Virol.* **72** 7985-7991

Habib S and Hasnain S E 1996 A 38 kDa host factor interacts with functionally important motifs of the *AcMNPV* homologous region 1 (hr1) DNA sequence; *J. Biol. Chem.* **271** 28250-28258

______ and ______ 1997 A bifunctional baculovirus homologous region (hr1) sequence: enhancer and origin of replication functions reside within the same sequence element; *Curr. Sci.* **73** 658

Habib S, Pandey S, Chatterjee U, Burma S, Ahmad R Jain A and Hasnain S E 1996 The bifunctional AcMNPV homologous region sequence (hrl): enhancer and ori functions have different sequence requirements; *DNA Cell Biol.* **15** 737-747

Hasnain S E, Habib S, Jain A, Burma S and Mukherjee B 1996 A host factor with single-stranded DNA-binding activity is involved in transcription from the baculovirus polyhedrin promoter; *Methods Enzymol.* **274** 20-32

______, Jain A, Habib S, Ghosh S, Chatterjee U, Ramachandran A, Das P, Venkiah B, Liang B and Azim A 1997 Involvment of host factors in transcription from baculovirus late gene promoters; *Gene* **190** 113-118

______, Liljestrom P and Roy P 1997 Eukaryotic expression vector systems: biology and applications; *Gene* **190** 1

Ho C K, Sriskanda V, McCracken S, Bentley D, Schwer B and Shuman S 1998 The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II; *J. Biol. Chem.* **273** 9577-9585

Hoeijmakers J H, Chambon P and Egly J M 1993 DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor.; *Science* **260** 58-63

Huh N E and Weaver R F 1990 Identifying the RNA polymerases that synthesize specific transcripts of the *Autographa californica* nuclear polyhedrosis virus; *J. Gen. Virol.* **71** 195-202

Hunter T and Karin M 1992 The regulation of transcription by phosphorylation; *Cell* **70** 375-387

Jain A and Hasnain S E 1996 A 30-kDa host protein binds to two very-late baculovirus promoters; *Eur. J. Biochem.* **239** 384-390

_____, Ramachandran A, Bashyam M D, Ghosh S, Chatterjee U, Azim Ch. A, Burma S and Hasnain S E 1999 Spl-like transcription factors present in insect cells support basal levels of gene expression from the baculovirus polyhedrin initiator promoter (Submitted.)

Javahery R, Khachi A, Lo K, Zenzie-Gregory B and Smale S T 1994 DNA sequence requirements for transcriptional initiator activity in mammalian cells; *Mol. Cell. Biol.* **14** 116-127

Jha P K, Pal R, Nakhai B, Sridhar P and Hasnain S E 1992 Simultaneous synthesis of enzymatically active luciferase and biologically active beta subunit of human chorionic gonadotropin in caterpillars infected with a recombinant baculovirus; *FEBS Lett.* **310** 148-152

Jin J, Dong W and Guarino L A 1998 The LEF-4 subunit of baculovirus RNA polymerase has RNA 5'-triphosphatase and ATPase activities; *J. Virol.* **72** 10011-10019

Kidd M and Emery V C 1993 The use of baculoviruses as expression vectors; *Appl. Biochem. Biotech.* **42** 137-159

Kogan P H and Blissard G W 1994 A baculovirus gp64 early promoter is activated by host transcription factor binding to CACGTG and GATA elements; *J. Virol.* **68** 813-822

Kovaes G R, Choe R J, Gurarino L A, Summers M D 1992 Functional dissection of the *Autographa californica* nuclear polyhedrosis virus immediate-early 1 transcriptional regulatory protein; *J. Virol.* **66** 7429-7437

Krappa R, Behn K A, Jahnel F, Doerfler W and KneBel-Morsdorf D 1992 Differential factor binding at the promoter of early baculovirus gene PE38 during viral infection: GATA motif is recognized by an insect protein; *J. Virol.* **66** 3493-3503

Krappa R, Roncarati K R and KneBel-Morsdorf D 1995 Expression of PE38 and IE2, viral members of the C3HC4 finger family, during baculovirus infection: PE38 and IE2 localize to distinct nuclear regions; *J. Virol.* **69** :5287-5293

Kremer A and Knebel-Morsdorf D 1998 The early baculovirus he65 promoter: On the mechanism of transcriptional activation by IE1; *Virology* **249** 336-351

Lang D and Stamminger T 1994 Minor groove contacts are essential for an interaction of the human cytomegalovirus IE2 protein with its DNA target; *Nucleic Acids Res.* **22** 3331-3338

Lannigan D A and Notides A C 1989 Estrogen receptor selectively binds the "coding strand" of an estrogen responsive element; *Proc. Natl. Acad. Sci. USA.* **86** 863-867

Li L, Harwood S H and Rohrmann G F 1999 Identification of additional genes that influence baculovirus late gene expression; *Virology* **255** 9-19

Lu A and Carsten E B 1993 Immediately-early baculovirus genes transactivate the *p143* gene promoter of *Autographa californica* nuclear polyhedrosis virus; *Virology* **195** 710-718

______ and ______1994 Identification of three late expression factor genes within the 33.8 to 43.4 map-unit region of *Autographa californica* nuclear polyhedrosis virus; *J. Virol.* **68** 6710-6718

______ and ______ 1995 The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication; *J. Virol.* **69** 975-982

Luckow V A and Summers M D 1988 Trends in the development of baculovirus expression vectors; Bio/Technology **6** 47-55

Mans R M W and Knebel-Morsdorf D 1998 *In vitro* transcription of pe38/polyhedrin hybrid promoters reveals sequences essential for recognition by the baculovirus-induced RNA polymerase and for the strength of very late viral promoters; *J. Virol.* **72** 2991-2998

McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S and Bentley D L 1997 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II; *Genes dev.* **11** 3306-3318

McLachlin J R and Miller L K 1994 Identification and characterization of *vlf1*, a baculovirus gene involved in very late gene expression; *J. Virol.* **68** 7746-7756

Morris T D and Miller L K 1994 Mutational analysis of a baculovirus major late promoter; *Gene* **140** 147-153

Murges D, Kremer A and Knebel-Morsdorf D 1997 Baculovirus transactivator IE1 is functional in mammalian cells; *J. Gen. Virol.* **78** 1507-1510

Mukherjee B, Burma S and Hasnain S E 1995 The 30 kDa protein binding to the "initiator" of the baculovirus polyhedrin promoter also binds specifically to the coding strand; *J. Biol. Chem.* **270** 4405-4411

Nakhai B, Sridhar P, Talwar G P and Hasnain S E 1991 Construction, purification and characterization of a recombinant baculovirus containing the gene for alpha subunit of human chorionic gonadotropin; *Indian J. Biochem. Biophys.* **28** 237-242

Nissen M S and Friesen P D 1989 Molecular analysis of the transcriptional regulatory region of an early baculovirus gene; *J. Virol.* **63** 493-503

Ooi B G, Rankin C and Miller L K 1989 Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene; *J. Mol. Biol.* **210** 721-730

Passarelli A L, Todd J W and Miller L K 1994 A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases; *J. Virol.* **68** 4673-4678

Possee R D and Howard S C 1987 Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus; *Nucleic Acids Res.* **15** 10233-10248

Pullen S S and Friesen P D 1995a The CAGT motif functions as an initiator element during early transcription of the baculovirus transregulator *ie-1; J. Virol.* **69** 3573-3583

_____ and _____ 1995b Early transcription of the *ie-1* transregulator gene of *Autographa californica* nuclear polyhedrosis virus is regulated by DNA sequences within its 5' noncoding leader region; *J. Virol.* **69** 156-165

Puri N, Krishnamoorthy S, Habib S, Hasnain S E Goswami S K and Prasad R K 1999 CDR1, a multidrug resistance gene from Candida albicans, contains multiple regulatory domains in its promoter of which a distal AP-1 element mediates its induction by miconazole; *FEMS Microbiol. Lett.* **180** 213-219

Rankin C, Ooi B G and Miller L K 1988 Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression; *Gene* **70** 39-50

Rapp J C, Wilson J A and Miller L K 1998 Nineteen baculovirus open reading frames, including LEF-12, support late gene expression; *J. Virol.* **72** 10197-10206

Reibero B M, Hutchinson K and Miller L K 1994 A mutant baculovirus with a temperature-sensitive IE-1 transregulatory protein; *J. Virol.* **68** 1075-1084

Rohrmann G F 1986 Polyhedrin structure; *J. Gen. Virol.* **67** 1499-1514

Santoro I M, Yi T M and Walsh K 1991 Identification of single-stranded-DNA-binding proteins that interact with muscle gene elements; *Mol. Cell. Biol.* **11** 1944-1953

Smale S T and Baltimore D 1989 The "initiator" as a transcription control element; *Cell* **57** 103-113

Smith G E, Fraser M J and Summers M D 1983 Molecular engineering of the *Autographa californica* nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene; *J. Virol.* **46** 584-593

Solomon M J, Strauss F and Varshavsky A 1986 A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA; *Proc. Natl. Acad. Sci.* **83** 1276-1280

Sridhar P, Awasthi A K, Azim C A, Burma S, Habib S, Jain A, Mukherjee B, Ranjan A and Hasnain S E 1994 Baculovirus vector-mediated expression of heterologous genes in insect cells; *J. Biosci.* **19** 603-614

Sridhar P, Panda A K, Pal R, Talwar G P and Hasnain S E 1993 Temporal nature of the promoter and not relative strength determines the expression of an extensively processed protein in a baculovirus system; *FEBS lett.* **315** 282-286

Stark H C, Weinberger O and Weinberger J 1992 *Proc. Natl. Acad. Sci. U.S.A.* **89** 2180-2184

Sudhakar A, Krishnamoorthi T, Jain A, Chatterjee U, Hasnain S E, Kaufman R J and Ramaiah K V A 1999 Serine 48 in initiation factor 2α(eIF2α) is required for high affinity interaction between eIF2αp and eIF2β; *Biochemistry* 38 15398-15405

Thiem S M and Miller L K 1989 Identification, sequence and transcription mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus; *J. Virol.* **63** 2008-2018

_____ and _____ 1990 Differential gene expression mediated by late, very late and hybrid baculovirus promoters; *Gene* **91** 87-94

Todd J W, Passarelli A L, Lu A and Miller L K 1995 Factors regulating baculovirus late and very late gene expression in transient-expression assays; *J. Virol.* **70** 2307-2317

Van Lier F L J, Vlak J M and Tramper J 1992 Production of baculovirus expressed proteins from suspension cultures of insect cells; *Animal Cell Biotech.* **5** 169-188

Weis L and Reinberg D 1992 Transcription by RNA polymerase II: Initiator-directed formation of transcription-competent complexes; *FASEB J.* **6** 3300-3309

Weyer U, Knight S and Possee R D 1990 Analysis of very late gene expression by *Autographa californica* nuclear polyhedrosis virus and the further development of multiple expression vectors; *J. Gen. Virol.* **71** 1525-1534

Xu B, Yoo S and Guarino L A 1995 Differential transcription of baculovirus late and very late promoters: fractionation of nuclear extracts by phosphocellulose chromatography; *J. Virol.* **69** 2912-2917

Yang S and Miller L K 1998 Control of baculovirus polyhedrin gene expression by very late factor 1.; *Virol.* **248** 131-138

Yang S and Miller L K 1999 Activation of baculovirus very late promoters by interaction with very late factor 1; *J. Virol.* **72** 3403-3409

Yang C L, Stetler D A and Weaver R F 1991 Structural comparison of the *Autographa californica* nuclear polyhedrosis virus-induced RNA polymerase and the three nuclear RNA polymerases from the host, *Spodoptera frugiperda*; *Virus Res.* **20** 251-264

Yoo S and Guarino L A 1994a The *Autographa californica* nuclear polyhedrosis virus *ie2* gene encodes a transcriptional regulator; *Virology* **202** 746-753

_____ and _____ 1994b Functional dissection of the *ie2* gene product of the baculovirus *Autographa californica* nuclear polyhedrosis virus; *Virology* **202** 164-172

Yoo W, Martin M E and Folk W R 1991 PEA1 and PEA3 enhancer elements are primary components of the polyomavirus late transcription initiator element; *J. Virol.* **65** 5391-5400