VIROLOGICA SINICA, April. 2007, 22 (2):94-107
CLC number: Q786 Document code: A Article ID: 1003-5125(2007)02-0094-14

Baculovirus RNA Polymerase: Activities, Composition, and Evolution

A. Lorena Passarelli **

(Molecular, Cellular, and Developmental Biology Program, Division of Biology, 116 Ackert Hall, Kansas State University, Manhattan, KS 66506-4901, USA)

Abstract: Baculoviruses are the only nuclear replicating DNA-containing viruses that encode their own DNA-directed RNA polymerase (RNAP). The baculovirus RNAP is specific for the transcription of genes expressed after virus DNA replication. It is composed of four subunits, making it the simplest multisubunit RNAP known. Two subunits contain motifs found at the catalytic center of other RNAPs and a third has capping enzyme functions. The function of the fourth subunit is not known. Structural studies on this unique RNAP will provide new insights into the functions of this enzyme and the regulation of viral genes and may be instrumental to optimize the baculovirus gene expression system.

Key words: Baculoviruses; Virus DNA replication; Regulation of viral genes; RNA polymerase

INTRODUCTION

Baculoviruses contain large double-stranded DNA genomes that potentially encode between 89 and 181 genes (27, 36). Genes are expressed in a step-wise temporally regulated flow and classified in three major phases of viral gene expression referred to as early, late, and very late phases. Host and viral genes synthesized in the prior phase(s) allow the expression of genes in the subsequent phase(s). First, early genes are expressed by host factors and/or viral genes. As the viral replication cycle progresses, host and early genes are shut off and late genes are transcribed. Later, at very late times in the infectious cycle, late genes are shut off and only very late genes are transcribed. The organized temporal choreography of all these genes leads to an efficient regulatory circuit that efficiently fuels virus replication.

Two multisubunit DNA-directed RNA polymerases (RNAPs) are required for this controlled and structured hierarchy of gene expression, the host RNAP II and a virally encoded RNAP. The host RNAP transcribes genes during the early phase and prior to viral genome replication. The promoter sequences of early genes are reminiscent of those present in host genes. On the other hand, the virally encoded RNAP transcribes genes in the later two phases, late and very late, and recognizes compact and specific promoter sequences. The involvement of host and viral factors has been proposed for transcription throughout the infection cycle.

HISTORY

The presence of a novel viral RNAP responsible for late and very late viral gene transcription has been

Received: 2007-02-09, Accepted: 2007-02-12

** Corresponding author. Tel: +1-785-532-3195, Fax: +1-785-532-6653, E-mail: lpassar@ksu.edu

known since the early 1980s. In the past 26 years, its components have been identified and some have been functionally characterized using biochemical and genetic tools. In addition, the multisubunit enzyme has been purified to homogeneity, facilitating its characterization as a complex. Yet, discovering the specific contribution of each component and the role of accessory factors to late gene transcription, its evolutionary lineage, and its three-dimensional structural anatomy, will project well into an exciting fourth decade of research since its discovery.

Early attempts could not detect an RNAP activity in budded or occluded derived virus particles (23). These results suggested that the host RNAP(s) was responsible for transcribing viral genes at early times or at early and late times. However, attempts to confirm this hypothesis led to the discovery of an RNAP with different biochemical properties to the host RNAPs, and initiated a search for a virally encoded RNAP.

α-amanitin resistant RNAP

Late and very late genes were shown to be transcribed by an α-amanitin-resistant RNAP activity (17, 23, 29). In a hallmark study, Weaver and colleagues (23) designed experiments to identify a virally encoded or virally induced RNAP activity. They purified nuclei from either *Spodoptera frugiperda*-derived cells or *Heliothis zea* fourth instar larvae infected with *Autographa californica* M nucleopolyhedrovirus (Ac-*M*NPV) or *Heliothis zea* NPV, respectively. The majority of virus specific RNA biosynthesis from both of these sources was resistant to α-amanitin, a strong inhibitor of host RNAP II, leading to the hypothesis that the virus encoded or induced a novel α-amanitin-resistant enzyme. However, purification of RNAP activity from infected or uninfected cells assayed in the presence or absence α-amanitin did not reveal different chromatographic profiles. Thus, RNAP II did not appear to become resistant to α-amanitin. This result was puzzling but did not exclude the possibility of a novel viral enzyme resistant to α-amanitin; alternative scenarios were also proposed, including a role for the α-amanitin-insensitive RNAP I or III in viral transcription.

Subsequent work revealed that a new RNAP activity resistant to α-amanitin appeared at late times post infection (17). Addition of cycloheximide to nuclear run-on assays showed that the novel RNAP activity was made, at least in part, from newly synthesized proteins between 5 and 8 hours post infection (29). These findings led to the proposal that the host RNAP, an α-amanitin sensitive enzyme available at early times post infection, was responsible for early viral gene synthesis. At late times, an α-amanitin resistant RNAP, which was either virally encoded or virally induced, was carrying out late gene transcription and this RNAP switch was linked to viral DNA replication (17, 29).

Composition of the α-amanitin resistant RNAP

It was suggested that the novel RNAP was composed of at least 8 to 10 polypeptides based on visualization of silver-stained polypeptides from column fractions of partially purified RNAP (3, 77). The RNAP was estimated to have a molecular mass of 396 kilodaltons (kDa) with the assumption that only one molecule of each subunit was present per complex. In addition, the complexity of this structure was in agreement with sedimentation coefficient values of

15S (77). These results are in contrast to our current understanding of the composition of the viral RNAP (see below). The identification of additional polypeptides in earlier studies may be attributed to either partial purification of the enzyme or its purification in the presence of transcription accessory factors.

At this time, it was also proposed that the viral RNAP may be composed of host mitochondrial RNA polymerase subunits. There is some sequence similarity between the late/very late viral transcription initiation site (TAAG) and sequences recognized by the yeast mitochondrial RNAP, and mitochondrial DNA can serve as a promoter for the baculovirus RNAP in vitro (43, 77).

GENES REQUIRED FOR LATE GENE TRANSCRIPTION

Genetic evidence to identify late gene transcription factors

Two major methods were instrumental in identifing genes involved in stimulating late and/or very late gene transcription, a genetic and a functional screen. Initially, viruses with temperature sensitive mutations that affected late processes were screened and characterized (5-8, 14, 37, 47, 57). It took a number of years after these mutants were constructed and the construction of an additional temperature sensitive virus to identify alleles carrying the conditional lethal defects that affected late gene expression (8, 20, 40, 46, 52, 59). Three genes identified by this method encoded RNAP subunits, although this was not clear at the time (8, 52, 62). Three other mutations mapped to genes that affected late/very late gene expression and viral DNA replication (40, 46, 59).

Transient expression assays to determine late gene expression factors

The functional assay relied on a set of 12 fragments that at the time, were thought to overlap and cover the entire Ac*M*NPV genome (54). The genes within several of these fragments were able to specifically stimulate expression of a reporter gene under late or very late promoter control in transient gene expression assays. Only 1 of the 12 fragments was required for stimulation of an early promoter, and early promoter stimulation mapped to two genes (54). Subclones of the 12 fragments identified 19 genes, named late expression factors (*lef*s) that were required for late gene transcription or DNA replication (34, 38, 42, 58, 72). These assay identified genes that were previously mapped by the genetic screen (8, 40) and provided candidate genes for uncharacterized temperature sensitive viruses (52, 59). An additional factor, very late expression factor-1 was specific for very late gene expression (46, 71), however, it may also have a role in packaging viral DNA (73).

INSIGHTS FROM *IN VITRO* TRANSCRIPTION REACTIONS

The existence of a novel viral RNAP activity was further confirmed and the enzyme characterized by in vitro run-off transcription of viral late promoter templates. Extracts or purified RNAP activity prepared from Ac*M*NPV-infected *S. frugiperda* cells collected at late times post infection were able to transcribe templates with a late or very late Ac*M*NPV promoter using relatively low Mg^{2+} concentrations (3, 19). Late transcription was resistant to α-amanitin and tagetitoxin, a eukaryotic RNAP III inhibitor. This implied

that neither RNAP II nor III were responsible for late gene transcription (19). However, it was not confirmed that tagetitoxin was indeed inhibiting RNAP III-directed transcription in this in vitro system, although RNAP III promoter-directed transcription of several other organisms including insects was inhibited by tagetitoxin (19, 65).

Run-off transcription assays are laborious, prone to non-specificity, and contained nuclease contaminants. In order to optimize the characterization of late and very late transcription in vitro, an in vitro transcription system for late and very late gene transcription using a cytidine-free cassette was developed (18, 76). Using this system, it was shown that specific transcription was dependent on viral factors present at late but not at early times post infection (76). In addition, the transcription efficiency from late and very late promoter templates differed when using two nuclear extract fractions, indicating the different requirements for late and very late gene transcription (76).

PURIFICATION AND POLYPEPTIDE IDENTIFICATION OF THE VIRALLY ENCODED RNAP

Prior to the purification of the viral RNAP to homogeneity, partial purifications were able to transcribe late/very late genes (3, 18, 76). Separation of the polypeptides by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) indicated that the complex had less than 10 main polypeptides (3, 77). However, the identification of the subunits forming the RNAP complex remained unknown.

Guarino and colleagues purified the viral RNAP to homogeneity from Ac*M*NPV-infected cells (26). The purified protein complex specifically transcribed a late and a very late promoter template but not an early promoter template in vitro. The complex was estimated to have a molecular mass of 560 kDa. Separation of polypetides by SDS-PAGE identified four polypeptides of relative molecular masses of 98, 55, 53, and 46 kDa present in equimolar concentrations, suggesting that the active complex contained two molecules of each polypeptide.

The composition of this RNAP represents the simplest eukaryotic RNAP described to date (26). In addition, these results are in agreement with *in vitro* transcription system assays in which a preincubation step of template and factors was not required for efficient subunit assembly preceding transcription initiation and elongation (76). Lack of a requirement for a preincubation step suggests that the transcribing complex has a simple composition (76).

In addition to its simple composition, the Ac*M*NPV RNAP is unique to DNA-containing viruses that replicate in the nucleus of cells. In general, viruses with DNA genomes utilize host RNAPs. The only two families of DNA-containing viruses that do not conform are *Poxviridae* and *Asfarviridae*. Members of both of these families encode an RNAP, however, they replicate in the cytoplasm of cells, making it impractical to access the nuclear RNAP II encoded by the host.

The individual subunits of the RNAP complex were identified by sequencing their N termini, peptide fingerprinting, and immunoblotting. Four subunits encoded by previously characterized genes were identified and their predicted sizes roughly corresponded to the sizes of the four purified polypeptides: LEF-8

(98 kDa), LEF-9 (55 kDa), LEF-4 (46 kDa), and P47 (46 kDa). The genes that encode the four RNAP subunits are present in all baculoviruses sequenced to date, and they are the most conserved LEFs (4). The percent similarities between specific RNAP subunits encoded by Group I baculoviruses and those of Group II, granuloviruses, and baculoviruses isolated from non-lepidopteran insects averages 62% (Table 1) in contrast to that of other LEFs which ranges from 36 to 55% and averages 46% (4).

LEF-8

The *lef-8* gene was first identified in a functional transient assay screen for late gene transacting factors and proposed to be a potential RNAP subunit due to the presence of a short invariable motif encoded by well-defined RNAPs (55). The *lef-8* gene was required for expression from two late and two very late promoters but not from an early one (55, 71). Since its identification in 1993, only two functional studies focusing on *lef-8* have been published; one study characterized a temperature sensitive mutation within the *Bombyx mori* NPV (BmNPV) *lef-8* (62), and another described a detailed function-structure analysis (70). Thus, information about the specific mechanisms governing its role in late gene transcription is scant.

Table 1. Average percent identities and similarities among RNAP subunits within different baculovirus groups[a]

RNAP subunit	Average identities (%)	Average similarities (%)
LEF-9	48.2	67.7
LEF-8	45.2	64.6
P47	37.6	59.1
LEF-4	32.6	54.5

[a] Each RNAP subunit sequence from a Group I baculovirus was compared to the corresponding subunits from other baculovirus groups, Group II, granuloviruses, and non-lepidopteran baculoviruses. Subunits predicted from a total of 26 baculovirus genomes were used in pairwise alignments.

Table 2. Known or predicted activities of the baculovirus RNAP subunits

RNAP subunit	Activities	References
LEF-9	Mg^{2+} binding during catalysis[a]	(41)
LEF-8	Catalysis[a]	(55, 70)
P47	Unknown	
	Guanylyltransferase	
LEF-4	RNA 5'-triphosphatase	(22, 25, 30)
	Nucleoside triphosphatase	

[a] Predicted activity based on the presence of motifs

A virus with a temperature sensitive mutation within the BmNPV *lef-8* replicated normally at early times and had normal viral DNA accumulation, but lacked processes that either required late gene expression and/or occurred at late times, including occluded virus production, late gene expression, host protein shutoff, and budded virus production. Furthermore, infection of larvae with this virus at the non-permissive temperature killed only approximately 34% of the infected larvae (62). However, the experiment was not controlled for infection of larvae with wild-type BmNPV at the non-permissive temperature. It is possible that the combination of virus infection and high temperature resulted in different mortality rates.

A detailed functional-structure analysis of *lef-8* was conducted to define the domains of l*ef-8* required for late gene promoter activation (70). The sequence of *lef-8* predicts a polypeptide of 102 kDa, the largest subunit of the viral RNAP and one of the largest predictted Ac*M*NPV gene products. However, its overall sequence does not reveal obvious homology to other RNAPs (55). Towards the C terminus of LEF-8, there is a 13-amino acid motif that is similar to that found in the second largest subunits of both pro-

karyotic (region H) and eukaryotic RNAPs. Mutations within this motif in *lef-8* abolish late gene expression, indicating the importance of this motif in late gene expression (70). In addition, site-directed mutations and deletions throughout *lef-8* resulted in late gene activity reductions, implying that regions throughout the polypeptide were necessary for interactions with other subunits, accessory factors, or nucleic acids (70). Alternatively, alterations or deletions of approximately 50 amino acids within the sequence of *lef-8* had drastic effects on the stability of the polypeptide. It has been suggested that the *lef-8* mRNA may be unstable (1), and we have had difficulty expressing *lef-8* transiently (13).

Fig. 1. Amino acid sequence alignment of a LEF-9 conserved sequence. Thirty-five baculoviruses encoding LEF-9 are listed alphabetically and a sequence upstream of the conserved NADFDGD-like motif is shown. The numbers in parentheses denote the first amino acid in each sequence. def, defective

LEF-9

As with *lef-8*, *lef-9* was also identified in a subtractive complementation transient late gene expression assay (41) but remarkably, 12 years after its identification, no other reports on its functional characterization are available. LEF-9 has an invariable heptapeptide motif also encoded in the largest subunits of other RNAPs (41). Furthermore, this motif has three aspartic acid residues that bind Mg^{2+} at the active center of the enzyme complex, and disruption of these residues obliterates catalytic activity in other RNAPs (78).

The *Escherichia coli* RNAP β' subunit, the largest subunits of eukaryotic RNAP I, II, and III, and the largest RNAP subunits from cytoplasmic DNA viruses all contain the conserved NADFDGD catalytic center motif (63) also present in LEF-9. Moreover, these subunits also have a well-conserved sequence, RQPT/SLH, between 26 and 31 amino acids upstream of the NADFDGD motif. Although the function of this region is not clear, a similar motif, RHPNIS, can be found 32 residues upstream of the Ac*M*NPV NADFDGD-like LEF-9 sequence. Furthermore, this motif upstream of NADFDGD is conserved in all of the LEF-9 homologs from baculoviruses sequenced to date (Fig.1). The conservation of these sequences stresses the evolutionary relationships among different RNAPs.

LEF-4

LEF-4 was identified as a product required for viral late gene expression in transient assays (53) and localizes to the nucleus of infected cells (15, 60). Viruses with mutations within *lef-4* also have defects in late processes (7, 33, 52).

Baculovirus mRNAs are known to be capped (32). The enzyme that may be responsible for processing the 5' ends of viral mRNAs is LEF-4, which is associated with the viral RNAP. Its role in capping was discovered by tracking guanylyltransferase activity of the viral RNAP (25) and by sequence and biochemical analyses (22). LEF-4 also has metal-dependent RNA triphosphatase activity (22, 25, 28). Although its role in capping mRNAs in vivo has not been established, its close association with the transcription machinery suggests a role in capping viral mRNAs *in vivo*.

Baculoviruses also encode a non-essential RNA triphosphatase and diphosphatase, but the enzyme has no associated guanylyltransferase activity (21, 39, 69). However, when fused to a functional guanylyltransferase module, it can form mRNA caps in vivo (44).

Some baculoviruses encode an RNA cap (nucleoside-2'*O*)-methyltransferase, but a mutant virus was viable (75). In transient expression assays, this gene stimulates late gene expression (38), thus, its specific role during virus infection and RNA processing is not clear.

P47

The least characterized of the RNAP subunits is P47. Its sequence does not reveal any obvious motifs or homology to other known proteins. Its role in late gene expression was discovered by characterizing a virus with a temperature sensitive mutation that mapped to *p47* (8, 52) and in transient late gene expression assays (72). The temperature sensitive virus was defective in late gene expression and budded virus production at non-permissive temperatures (52). Its product localizes to the nucleus of infected cells (8), commensurate with its role in transcription. In tran-

sient late gene expression assays, it was found that *p47* was required for late gene expression of viral late and very late gene promoters (71, 72).

ACTIVITIES OF THE BACULOVIRUS RNAP

In vitro transcription studies using the purified RNAP complex resulted in specific transcription of late and very late promoter templates; this implied that the viral RNAP was sufficient to discriminate between early and late/very late promoter regions, specifically bind the DNA template, and transcribe it (26). Which subunit (s) is (are) responsible for each of these functions is not known. Also, the RNAP is potentially able to process nascent mRNAs at both 3' and 5' ends. Description of the role of LEF-4 in modifying the 5' end by addition of a cap is described above. In addition, the RNAP is able to process 3' ends in vitro by terminating transcription as it slips on U-rich sequences and then polyadenylating using non-templated adenosines rather than by the cleavage and polyadenylation mechanism that requires a AAUAAA polyadenylation recognition sequence and a GU-rich region (31). It remains to be determined whether this mechanism is used in vivo.

STUCTURE AND EVOLUTION OF RNAPS

RNAPs and other types of polymerases have evolved conserved sequences and parallel functions. There is extensive homology among the largest subunits of eukaryotic and prokaryotic RNAPs and the DNA binding domain of the *E. coli* DNA polymerase (2, 48, 68). Yeast RNAP I, II, and III share subunits that are immunologically and biochemically indistinguishable, indicating a role in a shared common function rather than a specific catalytic step (74). Furthermore, the three-dimensional structures are reminiscent to each other. Similarly, the crystal structures of the monomeric T7 RNA polymerase, Klenow, and HIV-1 reverse transcriptase share a common structure that led to the identification of polymerase substrate specificity elements (64). The recent structures of yeast RNA polymerase II and the bacterial *Thermus aquaticus* core RNAP also share architectural similarities (11, 79). These parallels imply that the transcription mechanisms and other polymerizing functions have been conserved in unicellular and multicellular organisms. Although polymerases share functionality and positioning of residues at the active center, the architecture of cellular multisubunit RNAPs is different from that of single-subunit RNAPs (50).

Anatomy of the baculovirus RNAP putative active center

The baculovirus RNAP has presumably conserved the modular organization of the catalytic pocket, where residues in different subunits are proximal in space to form the active site. Also, this organization may be important for the reported stretching of the active center during elongation (49). LEF-8 and LEF-9 encode motifs positioned in other RNAPs at the heart of the enzyme. A modular organization may allow regulatory factors the opportunity to interact with a subunit containing a catalytic motif (50). The specific function of six transcription-specific *lefs* (*lef-5, lef-6, lef-10, lef-11, lef-12,* and *pp31*) that are not part of the RNAP has not been defined, and they could potentially affect transcription by interacting with the RNAP. Thus, given the conservation in structure and function among other polymerases, it is

legitimate to compare the novel baculovirus RNAP to other RNAPs.

The structure of the 10-subunit yeast RNAP II (12) and others have two mobile "jaws" that frame a "cleft" into which DNA enters. At the node of these jaws, formed by the interface between the two largest RNAP subunits, lies the catalytic center. The catalytic center contains the metal binding site where Mg^{2+} binds the sequence NADFDGD of the largest subunit, similar to that present in LEF-9, and a "hybrid region" from the second largest subunit that contains the 13-amino acid motif similar to that present in LEF-8. The NADFDGD metal binding site coordinates the growing 3'-OH on the mRNA and the α-phosphate from the substrate NTP, which enters the enzyme complex from a pore at the base of the catalytic center (12). A second metal binding site was identified once the RNAP II was resolved at 2.8 Ångstroms. This second site coordinates the three phosphates on the NTP (12). Both metal binding sites and the 13-amino acid motif are at the node where catalysis takes place. This modular arrangement of the active site may be conserved in the baculovirus RNAP, where the motifs located at the active site are in two separate subunits; however, sequences flanking these motifs are different. Conservation in the overall three-dimensional structure will be important to determine functional analogies.

Implications of the lack of an RNAP C-terminal domain

The C-terminal domain (CTD) is present in the largest subunit of RNAP II of many organisms and consists of a tandemly repeated heptapeptide. However, some organisms (*e.g.*, the protist *Mastigamoeba invertens*, trypanosomids, and some red algae) do not have the canonical heptad repeat but a heptad with a different sequence or have no repeated sequence runs (16, 56, 66). The baculovirus RNAP does not encode this CTD (26). The presence of CTDs can be used to establish evolutionary lineages. It has been determined that all canonical CTDs likely descended from a single ancestor (56). CTDs are present among ancient eukaryotes (*e.g.*, *M. invertens* and microsporidian parasites) and although they are absent in some organisms, the loss of these essential sequences is not likely (67). It would be of interest to analyze the phylogeny of the baculovirus RNAP. It is possible that it diverged prior to CTD-bearing RNAPs. More likely, lacking this C-terminal extension may give the baculovirus RNAP an advantage for transcription of viral late and very late genes over transcription of host genes. The lack of overall sequence conservation between the baculovirus RNAP subunits and other RNAP subunits makes phylogenetic analysis challenging.

The CTD binds a number of nuclear factors associated with processing RNAs at both 3' and 5' ends as they are synthesized (10). The CTD is phosphorylated in a sequential fashion and this regulates and specifies binding of specific factors. Factors that are recruited by the CTD include the capping enzyme, splicing factors, and cleavage and polyadenylating factors (56).

The CTD recruits the capping enzyme to the site of transcription. Since baculovirus RNAP subunits lack a CTD, the virus is presumably able to circumvent this problem by making its capping enzyme, LEF-4, part of its RNAP (26). It is possible that LEF-4 has another function directly related to transcription. There are no

detailed structure-function characterization studies on LEF-4; mutational analysis has been limited to motifs involved in capping enzyme reactions (30, 33, 45, 61) and analysis of other regions may be important to delineate *lef-4* functions. For example, a mutation corresponding to that in the conditional lethal *lef-4* virus (7) has only moderate defects in RNA triphosphatase activity but no defects in ATPase or guanylyltransferase activities (30).

Given that only one baculovirus gene, *ie-0*, is spliced (9,35), the CTD, which is absent in the baculovirus RNAP, would not be necessary for this processing event. Details on the splicing mechanism or timing during transcription by which *ie-0* transcripts are processed have not been studied, although late *ie-0* RNAs are spliced (35). Information on whether the viral RNAP is instrumental in targeting splicing factors during *ie-0* transcription has not been explored.

Since baculovirus late mRNAs are not processed by cleavage/polyadenylation (31), the CTD would also not be required for this process in baculoviruses. Using a different mechanism to process the 3' ends than that of the host may provide the virus a method to selectively process its transcript ends over those of host transcripts.

RNAP subunit interactions

The structure of the baculovirus RNAP or the interactive nature of its four subunits has not been reported. Given the uniqueness of this RNAP in primary sequence and functions, studies on the interrelationship amongst its subunits, culminating in its three-dimensional architecture, may provide an overall picture to compare with available RNAP crystal structures.

Purification of an RNAP activity to homogeneity revealed that there were two molecules of each subunit in the complex, however, the interaction map of these subunits is not known. It is not known whether two complexes of four subunits each assemble into an eight-molecule complex, if each subunit forms a homodimer, or if there is a different arrangement.

There is one study showing the interaction between BmNPV LEF-8 and LEF-9 (1). In this study, a plasmid encoding a Flag-tagged LEF-8 was transfected into insect cells, and at 8 hours post transfection, cells were infected with BmNPV and proteins were immunoprecipitated with FLAG antibody. LEF-9 immunoprecipitated by its association with LEF-8 was detected by immunoblotting using LEF-9 antisera. This experiment was only confirmed in one direction and since all viral proteins were included in the assay, it is not known whether this interaction was direct or if it was bridged by other viral factors. Yet, these results support the interactive nature of the two subunits that encode domains putatively present at the catalytic center as observed for the corresponding subunits in other RNAPs.

Crystal structure determinations of RNAPs have shown that subunits have considerable interactions and the interfacing of each interactive subunit pair differs. For example, the two largest subunits of the *T. aquaticus* core RNAP interact near the active center (79). If this scenario is conserved in the baculovirus RNAP, this would imply that the LEF-8 and LEF-9 subunits interact near the putative catalytic center motifs. We have observed interactions among these

two subunits throughout most of the LEF-8 polypeptide (13). Thus, the vast interfacing between these two subunits and those of the largest subunits of other RNAPs is conserved, but the overall architecture of the enzyme complex may differ.

In multisubunit eukaryotic RNAPs a number of subunits stabilize weak intersubunit interactions (51). We do not yet know the strength of the subunit interactions in the baculovirus RNAP complex. It is possible that these interactions are stabilized *in vivo* by other viral proteins, possibly other *lefs*. *In vitro* transcription assays revealed that only the four-subunit RNAP complex is necessary for transcription of late genes, however, addition of LEF-5 stimulated transcription initiation (24). Therefore, other proteins may aid in forming or stabilizing the complex (26) or optimizing transcription *in vivo*.

It would be of interest to determine the assembly order of the baculovirus RNAP complex. Although subunits may interact in yeast-two hybrid or coimmunoprecipitation assays, the order of assembly could specify the subunit surfaces available for interactions as assembly occurs. In the prokaryotic RNAP, strong interactions between the β subunit-F and-I regions are important for binding to the α subunit (51).

CONCLUSIONS

Baculovirus gene regulation is mainly controlled at the transcriptional level. Thus, studies on the baculovirus RNAP are important to both understand gene regulation in baculoviruses and to shed light into mechanisms used by other viral RNAPs, although they may appear to be distantly related.

The baculovirus RNAP is the simplest eukaryotic multisubunit RNAP known, yet it executes many of the functions of more complex RNAPs. Once we understand all its functions and can assign each function to particular domains, these can be modified to use the RNAP as a molecular biology tool. Structural analysis of the RNAP will be instrumental to understand its functional complexity.

Acknowledgment

I thank Christophe Lehiy for his assistance with sequence alignments. This work was supported in part by NIH award P20 RR16443 and by the NRI Competitive Grants Program, USDA, under agreement number 2001-35302-09983.

References

1. **Acharya A, Gopinathan K P.** 2002. Characterization of late gene expression factors *lef-9* and *lef-8* from *Bombyx mori* nucleopolyhedrovirus. **J Gen Virol,** 83: 2015-2023.
2. **Allison L A, Moyle M, Shales M, *et al.*** 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. **Cell,** 42: 599-610.
3. **Beniya H, Funk C J, Rohrmann G F, *et al.*** 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.
4. **Berretta M F, Passarelli A L.** 2006. Function of *Spodoptera exigua* nucleopolyhedrovirus late gene expression factors in the insect cell line SF-21. **Virology,** 355: 82-93.
5. **Brown M, Crawford A M, Faulkner P.** 1979. Genetic analysis of a baculovirus *Autographa californica* nuclear polyhedrosis virus part 1 isolation of temperature sensitive mutants and assortment into complementation groups. **J Virol,** 31: 190-198.
6. **Brown M, Faulkner P.** 1980. A partial genetic map of the baculovirus *Autographa californica* nuclear poly-

hedrosis virus based on recombination studies with ts mutants. **J Gen Virol,** 48: 247-252.

7. **Carstens E B, Chan H, Yu H, *et al.*** 1994. Genetic analyses of temperature-sensitive mutations in baculovirus late expression factors. **Virology,** 204: 323-337.
8. **Carstens E B, Lu A L, Chan H B.** 1993. Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gene regulating late gene expression. **J Virol,** 67: 2513-2520.
9. **Chisholm G E, Henner D J.** 1988. Multiple early transcripts and splicing of the *Autographa californica* nuclear polyhedrosis virus ie-1 gene. **J Virol,** 62: 3193-3200.
10. **Corden J L.** 1990. Tails of RNA polymerase II. **Trends Biol Sci,** 15: 413-416.
11. **Cramer P, Bushnell D A, Fu J, *et al.*** 2000. Architecture of RNA polymerase II an implications for the transcription mechanism. **Science,** 288: 640-649.
12. **Cramer P, Bushnell D A, Kornberg R D.** 2001. Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. **Science,** 292: 1861-1876.
13. **Crouch E A, Passarelli A L.** Unpublished results.
14. **Duncan R, Faulkner P.** 1982. Bromodeoxy uridine induced mutants of *Autographa californica* nuclear polyhedrosis virus defective in occlusion body formation. **J Gen Virol,** 62: 369-374.
15. **Durantel D, Croizier G, Ravallec M, *et al.*** 1998. Temporal expression of hte AcMNPV *lef-4* gene and subcellular localization of the protein. **Virology,** 241: 276-284.
16. **Evers R, Hammer A, Köck J, *et al.*** 1989. Trypanosoma brucei contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. **Cell,** 56: 585-597.
17. **Fuchs L Y, Woods M S, 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,** 43: 641-646.
18. **Funk C J, Harwood S H, Rohrmann G F.** 1998. Differential stability of baculovirus late transcription complexes during initiation and elongation. **Virology,** 241: 131-140.
19. **Glocker B, Hoopes J, Hodges R R L, *et al.*** 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.
20. **Gordon J D, Carstens E B.** 1984. Phenotypic characterization and physical mapping of a temperature sensitive mutant of *Autographa californica* nuclear polyhedrosis virus defective in DNA synthesis. **Virology,** 138: 69-81.
21. **Gross C H, Shuman S.** 1998. Characterization of a baculovirus-encoded RNA 5'-triphosphatase. **J Virol, 72:** 7057-7063.
22. **Gross C H, Shuman S.** 1998. RNA 5'-triphosphatase, nucleoside triphosphatase, and guanylyltransferase activities of baculovirus LEF-4 protein. **J Virol,** 72: 10020-10028.
23. **Grula M A, Buller P L, Weaver R F.** 1981. α-amanitin-resistant viral RNA synthesis in nuclei isolated from nuclear polyhedrosis virus-infected *Heliothis zea* larvae and *Spodoptera frugiperda* cells. **J Virol,** 38:916-921.
24. **Guarino L A, Dong W, Jin J.** 2002. In vitro activity of the baculovirus late expression factor LEF-5. **J Virol.** 76: 12663-12675.
25. **Guarino L A, Jin J, Dong W.** 1998. Guanylyltransferase activity of the LEF-4 subunit of baculovirus RNA polymerase. **J Virol,** 72: 10003-10010.
26. **Guarino L A, Xu B, Jin J, *et al.*** 1998. A virus-encoded RNA polymerase purified from bacul- ovirus-infected cells. **J Virol,** 72: 7985-7991.
27. **Hayakawa T, Ko R, Okano K, *et al.*** 1999. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. **Virology,** 262: 277-297.
28. **Ho C K, Pie Y, Shuman S.** 1998. Yeast and viral RNA 5' triphosphatases comprise a new nucleoside triphosphatase family. **J Biol Chem,** 273: 34151-34156.
29. **Huh N E, 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.
30. **Jin J, Dong W, Guarino L A.** 1998. The LEF-4 subunit of baculovirus RNA polymerase has RNA 5'-triphosphatase and ATPase activities. **J Virol,** 72: 10011-10019.

31. **Jin J, Guarino L A.** 2000. 3'-end formation of baculovirus late RNAs. **J Virol,** 74: 8930-8937.
32. **Jun-Chuan Q, Weaver R F.** 1982. Capping of viral RNA in cultured *Spodoptera frugiperda* cells infected with *Autographa californica* nuclear polyhedrosis virus. **J Virol,** 43: 234-240.
33. **Knebel-Mörsdorf D, Quadt I, Li Y, *et al.*** 2006. Expression of baculovirus late and very late genes dependes on LEF-4, a component of the viral RNA polymerase whose guanyltransferase function is essential. **J Virol,** 80: 4168-4173.
34. **Kool M, Ahrens C H, Goldbach R W, *et al.*** 1994. Identification of genes involved in DNA replication of the *Autographa californica* baculovirus. **Proc Natl Acad Sci USA,** 91: 11212-11216.
35. **Kovacs G R, Guarino L A, Graham B L, *et al.*** 1991. Identification of spliced baculovirus RNAs expressed late in infection. **Virology,** 185: 633-643.
36. **Lauzon H A M, Lucarotti C J, Krell P J, *et al.*** 2004. Sequence and organization of the *Neodiprion lecontei* nucleopolyhedrovirus genome. **J Virol,** 78: 7023-7035.
37. **Lee H H, Miller L K.** 1979. Isolation, complementation, and initial characterization of temperature-sensitive mutants of the baculovirus *Autographa californica* nuclear polyhedrosis virus. **J Virol,** 31: 240-252.
38. **Li L, Harwood S H, Rohrmann G F.** 1999. Identifi cation of additional genes that influence baculovirus late gene expression. **Virology,** 255: 9-19.
39. **Li Y, Miller L K.** 1995. Properties of a baculovirus mutant defective in the protein phosphatase gene. **J Virol,** 69: 4533-4537.
40. **Lu A, Carstens E B.** 1991. Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus *Autographa californica* nuclear polyhedrosis virus. **Virology,** 181: 336-347.
41. **Lu A, Miller L K.** 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.
42. **Lu A, Miller L K.** 1995. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. **J Virol,** 69: 975-982.
43. **Mans R M W, Knebel-Mörsdorf D.** 1999. Mitochondrial DNA acts as a potential promoter of the baculovirus RNA polymerase. **Biol Chem,** 380: 579-583.
44. **Martins A, Shuman S.** 2002. The domain of mammalian capping enzyme can be inverted and baculovirus phosphatase can function in cap formation *in vivo*. **Virology,** 304: 167-175.
45. **Martins A, Shuman S.** 2001. Mutational analysis of baculovirus capping enzyme Lef4 delineates an autonomous triphosphatase domain and structural determinants of divalent cation specificity. **J Biol Chem,** 276: 45522-45529.
46. **McLachlin J R, Miller L K.** 1994. Identification and characterization of *vlf-1,* a baculovirus gene involved in very late gene expression. **J Virol,** 68: 7746-7756.
47. **Miller L K, Trimarchi R E, Browne D, *et al.*** 1983. A temperature sensitive mutant of the baculovirus *Autographa californica* nuclear polyhedrosis virus defective in an early function required for further gene expression. **Virology,** 126: 376-380.
48. **Mooney R A, Landick R.** 1999. RNA polymerase unveiled. **Cell,** 98:687-690.
49. **Mustaev A, Kashlev M, Zaychikov E, *et al.*** 1993. Active center rearrangement in RNA polymerase initiation complex. **J Biol Chem,** 268: 19185-19187.
50. **Mustaev A, Kozlov M, Markovtsov V, *et al.*** 1997. Modular organization of the catalytic center of RNA polymerase. **Proc Natl Acad Sci USA**, 94: 6641-6645.
51. **Naryshkina T, Rogulja D, Golub L, *et al.*** 2000. Inter- and intrasubunit interactions during the formation of RNA polymerase assembly intermediate. **J Biol Chem,** 275: 31183-31190.
52. **Partington S, Yu H, Lu A, *et al.*** 1990. Isolation of temperature sensitive mutants of *Autographa californica* nuclear polyhedrosis virus phenotype characterization of baculovirus mutants defective in very late gene expression. **Virology,** 175: 91-102.
53. **Passarelli A L, Miller L K.** 1993. Identification of genes encoding late expression factors located between 56.0 and 65.4 map units of the *Autographa californica* nuclear polyhedrosis virus genome. **Virology,** 197: 704-714.
54. **Passarelli A L, Miller L K.** 1993. Three baculovirus genes involved in late and very late gene expression: *ie-1, ie-n,* and *lef-2*. **J Virol,** 67: 2149-2158.

55. **Passarelli A L, Todd J W, 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.
56. **Phatnani H P, Greenleaf A L.** 2006. Phosphorylation and functions of the RNA polymerase II CTD. **Genes & Development,** 20: 2922-2936.
57. **Potter K N, Miller L K.** 1980. Correlating genetic mutations of a baculovirus with the physical map of the DNA genome. In: **Animal virus genetics**: ICN-UCLA Symposia on Molecular and Cellular Biology (Fields B, Jaenisch R, Fox C. F. ed.). New York: Academic Press, p71-80.
58. **Rapp J C, Wilson J A, Miller L K.** 1998. Nineteen baculovirus open reading frames, including LEF-12, support late gene expression. **J Virol,** 72: 10197-10206.
59. **Ribeiro B, Hutchinson K, Miller L K.** 1994. A mutant baculovirus with a temperature sensitive IE-1 transregulatory protein. **J Virol,** 68: 1075-1084.
60. **Sehrawat S, Gopinathan K P.** 2002. Temporal expression profile of *late gene expression factor 4* from *Bombyx mori* nucleopolyhedrovirus. **Gene,** 294: 67-75.
61. **Sehrawat S, Srinivasan N, Gopinathan K P.** 2002. Functional characterization and structural modeling of late expressin factor 4 from *Bombyx mori* nucleopolyhedrovirus. **Biochem J,** 368: 159-169.
62. **Shikata M, Sano Y, Hashimoto Y, *et al.*** 1998. Isolation and characterization of a temperature-sensitive mutant of *Bombyx mori* nucleopolyhedrovirus for a putative RNA polymerase gene. **J Gen Virol,** 79: 2071-2078.
63. **Sonntag K-C, Darai G.** 1996. Evolution of viral DNA-dependent RNA polymerases. **Virus Genes,** 11: 271-284.
64. **Sousa R, Chung Y J, Rose J P, *et al.*** 1993. Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. **Nature,** 364: 593-599.
65. **Steinberg T H, Mathews D E, Durbin R D, *et al.*** 1990. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA polymerase III. **J Biol Chem,** 265: 499-505.
66. **Stiller J W, Duffield E C S, Hall B D.** 1998. Amitochondriate amoebae and the evolution of DNA-dependent RNA polymerase II. **Proc Natl Acad Sci USA,** 95: 11769-11774.
67. **Stiller J W, McConaughy B L, Hall B D.** 2000. Evolutionary complementation for polymerase II CTD function. **Yeast,** 16: 57-64.
68. **Sweetser D, Nonet M, Young R A.** 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. **Proc Natl Acad Sci USA,** 84: 1192-1196.
69. **Takagi T, Taylor G S, Kusakabe T, *et al.*** 1998. A protein tyrosine phasphatase-like protein from baculovirus has RNA 5' triphosphatase and diphosphatase activities. **Proc Natl Acad Sci USA,** 95: 9808-9812.
70. **Titterington J S, Nun T K, Passarelli A L.** 2003. Functional dissection of the baculovirus *late expression factor-8*: Sequence requirements for late gene promoter activation. **J Gen Virol,** 84: 1817-1826.
71. **Todd J W, Passarelli A L, Lu A, *et al.*** 1996. Factors regulating baculovirus late and very late gene expression in transient-expression assays. **J Virol,** 70: 2307-2317.
72. **Todd J W, Passarelli A L, Miller L K.** 1995. Eighteen baculovirus genes, including *lef-11, p35, 39K,* and *p47,* support late gene expression. **J Virol,** 69: 968-974.
73. **Vanarsdall A L, Okano K, Rohrmann G F.** 2006. Characterization of the role of very late expression factor 1 in baculovirus capsid structure and DNA processing. **J Virol,** 80: 1724-1733.
74. **Woychik N A, Liao S-M, Kolodziej P A, *et al.*** 1990. Subunits shared by eukaryotic nuclear RNA polymerases. **Genes & Development,** 4: 313-323.
75. **Wu X, Guarino L A.** 2003. *Autographa californica* nucleopolyhedrovirus *orf69* encodes an RNA cap (nucleoside-2′-*O*)-methyltransferase. **J Virol,** 77: 3430-3440.
76. **Xu B, Yoo S, 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.
77. **Yang C L, Stetler D A, Weaver R. F.** 1991. Structural comparison of the *Autographa californica* nuclear polyhedrosis virus-induced RNA polymerase and the 3 nuclear RNA polymerases from the host, *Spodoptera frugiperda*. **Virus Res,** 20: 251-264.
78. **Zaychikov E, Martin E, Denissova L, *et al.*** 1996. Mapping of catalyltic residues in the RNA polymerase active center. **Science,** 273: 107-109.
79. **Zhang G, Campbell E A, Minakhin L, *et al.*** 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. **Cell,** 98: 811-824.