ROLE OF A18 IN VACCINIA VIRUS POST-REPLICATIVE GENE TRANSCRIPTION TERMINATION

By

APARNA MANOHARAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010
To my brother, Aroon, who inspires me to reach for everything I want; my dad, who shows me everyday what it is to learn and grow; and my mom, who humbles me with her quiet strength and fortitude
ACKNOWLEDGMENTS

To begin with, I thank the many people I've worked with in the Condit lab - Baron McFadden, Brad Dilling, Carson Rodeffer, Desyree Jesus, Dr. Hendrik Nollens, Nicole Kay, Olga Boyd, Ricky Antonia, Dr. Sayuri Kato, Dr. Steve Cresawn, Travis Bainbridge and Dr. Amber Shatzer – for the science, the good music and great times. They have been my home away from home. I also thank Dr. Cari Lackner from whom I inherited this project for her ready answers to all my questions, Dr. Cindy Prins for holding my hands and helping me get started on this project, and especially Dr. Susan D'Costa and Dr. Nissin Moussatche for their constant help with my experiments, for letting me pick their brains incessantly, their kindness and generosity.

I owe my deepest gratitude to my mentor, Dr. Richard Condit, without his help this dissertation would not have been possible. He has been an incredible source of support and guidance through the ups and downs of this project. I have learnt, as much from his words as his actions, valuable lessons in science, work ethics, conduct, temperament and life. I thank him for the opportunity to learn and grow. I am greatly indebted to the members of my dissertation committee - Dr. David Bloom, Dr. Jorg Bungert and Dr. James B. Flanegan for their valuable time and guidance and most importantly the helpful and interactive discussions.

I would like to acknowledge the efforts of the support staff at the Department of Molecular Genetics and Microbiology - Joyce Conners, Michele Ramsey, Julie Dillard, Dorcas Ortiz, Steve Howard, Kristyn Minkoff and Deborah Burgess.

It is my pleasure to thank the various friends and family, far too many to name, who have been a source of great strength and encouragement through the difficult phases of this journey.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ............................................................................................................. 4

LIST OF FIGURES ...................................................................................................................... 7

ABSTRACT .................................................................................................................................. 8

CHAPTER

1 INTRODUCTION ................................................................................................................... 10

Transcription & Gene Expression .............................................................................................. 10
Prokaryotic Transcription ........................................................................................................ 11
   RNA Polymerase .................................................................................................................. 12
   Initiation ............................................................................................................................... 13
   Elongation ............................................................................................................................. 14
   Termination .......................................................................................................................... 15
      Rho-dependent termination ............................................................................................. 15
      Intrinsic termination ........................................................................................................ 17
      Mfd-mediated termination ............................................................................................... 18
Eukaryotic Transcription .......................................................................................................... 20
   RNA Polymerase .................................................................................................................. 21
   Initiation ............................................................................................................................... 22
   Elongation ............................................................................................................................. 26
   Termination .......................................................................................................................... 28
Vaccinia Virus Biology ........................................................................................................... 32
   Vaccinia Life Cycle ................................................................................................................ 33
   Vaccinia Virus Transcription ............................................................................................ 35
      Early gene expression ........................................................................................................ 38
      Post-replicative gene expression ...................................................................................... 44
      A18 .................................................................................................................................. 48
Significance of this Study ......................................................................................................... 50

2 MATERIALS AND METHODS ............................................................................................ 53

Eukaryotic Cells, Prokaryotic Hosts, Viruses and Extracts .......................................................... 53
Plasmids ...................................................................................................................................... 53
Transcription Competent Lysolecithin Extracts ........................................................................ 54
HeLa Cytoplasmic Extracts ....................................................................................................... 54
Preparation of His-A18 ............................................................................................................... 55
   E.coli Lysate Preparation ..................................................................................................... 55
   Purification on a HisTrap Column ....................................................................................... 56
Templates Used in In Vitro Transcription Assays .................................................................... 56
   Preparation of pG8G Template ........................................................................................... 56
   Preparation of pG8GU Template ........................................................................................ 57
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Termination of stalled ternary complexes (STC)</td>
</tr>
<tr>
<td>3-2</td>
<td>Non-specific substitution of host factor activity</td>
</tr>
<tr>
<td>3-3</td>
<td>Detergent stimulation of A18 mediated termination</td>
</tr>
<tr>
<td>3-4</td>
<td>Heat stability of the host factor activity</td>
</tr>
<tr>
<td>3-5</td>
<td>Host factor is a protein</td>
</tr>
<tr>
<td>3-6</td>
<td>Host factor is a ubiquitous eukaryotic protein</td>
</tr>
<tr>
<td>4-1</td>
<td>Kinetics of A18 mediated termination</td>
</tr>
<tr>
<td>4-2</td>
<td>Salt optima for A18 mediated termination</td>
</tr>
<tr>
<td>4-3</td>
<td>Metal ion optima for termination</td>
</tr>
<tr>
<td>4-4</td>
<td>Energy requirement for termination</td>
</tr>
<tr>
<td>4-5</td>
<td>Transcription of paused complexes</td>
</tr>
<tr>
<td>4-6</td>
<td>Termination of paused complexes</td>
</tr>
<tr>
<td>4-7</td>
<td>Early time course of elongation</td>
</tr>
<tr>
<td>4-8</td>
<td>A18 specific termination</td>
</tr>
<tr>
<td>5-1</td>
<td>Model for A18 mediated termination</td>
</tr>
<tr>
<td>5-2</td>
<td>Schematic representation of the A18 protein</td>
</tr>
</tbody>
</table>
Vaccinia virus is a large double stranded DNA virus, unique among DNA viruses because it replicates entirely in the host cytoplasm and hence encodes for its own transcriptional machinery. Elucidation of the process of transcription in vaccinia has served as an excellent model to understand different aspects of eukaryotic transcription biology. Transcription in vaccinia virus is temporally regulated, with the genome being transcribed as early, intermediate and late genes. Intermediate and late gene expressions require the onset of genome replication and hence are referred to as post-replicative gene expression. Our lab is specifically interested in the process of transcription regulation of post-replicative genes in vaccinia virus. Several viral proteins have been identified to play a role in post-replicative transcription regulation. This dissertation is based on work done with one such protein, A18, and its role in viral transcription termination.

The vaccinia virus (VV) protein A18 was identified in our lab to play a role in late gene transcription regulation, from studies done with temperature sensitive mutant viruses. The mutant viruses made late transcripts that were longer than wildtype virus transcripts, implying that the mutant gene product, A18, functioned as a negative
elongation factor. Purified histidine-tagged A18 shows DNA-dependent ATPase activity and a weak helicase activity for dsDNA. An in vitro transcription assay shows that A18 functions as a transcript release factor in the presence of additional host factors from uninfected HeLa extracts.

This dissertation involves studying A18 in the context of an in vitro transcription assay to understand the role of A18 as a transcription termination factor. Attempts made to characterize the nature of the host factor indicated that a specific, ubiquitous eukaryotic protein component of the HeLa extracts aided in A18 mediated termination. Stalled transcription ternary complexes generated from a vaccinia intermediate gene promoter that were isolated and used to study A18 mediated termination showed that neither A18 nor the host factor were required to interact with the elongating RNA polymerase complex in order to mediate termination.

These stalled ternary complexes proved an important tool to delve into the nature of A18 mediated termination and study the reaction kinetics, define salt and divalent metal ion optima and energy requirements for this enzymatic process. An important aspect of the in vitro studies with A18 involves looking at the role of the enzyme in the context of transcriptional pausing. We have been able to show that the enzyme mediates termination of ternary complexes paused at natural pause sites, proving that pausing is a requirement for A18 mediated termination. In the context of transcriptional pausing, we have also identified that A18 has to a certain degree host factor independent termination potential.
CHAPTER 1
INTRODUCTION

Transcription & Gene Expression

The primary step in the process of gene expression is transcription, which can be simplistically defined as a process where a DNA molecule is used as a template by enzymes to synthesize the complementary RNA product. The central molecule that drives the process of transcription is the DNA-dependent RNA polymerase. RNA polymerases vary in complexity from a single subunit enzyme used by the bacteriophages to the multisubunit enzyme used by prokaryotes, archea and eukaryotes, possessing up to 12 subunits and an overall molecular mass of 500 KDa. Despite the differences in the complexity of the transcription apparatus across the different realms of life, the process of transcription has generally been divided into three stages – initiation, elongation and termination. Transcription initiation begins with the formation of a binary complex which consists of the double-stranded DNA template and the DNA-dependent RNA polymerase. It involves, along with the RNAP and the DNA template; other associated transcription factors. The process is initiated by the recognition of promoter sequences that drive transcription from a specific gene by transcription factors. These factors also recruit the RNA polymerase to the DNA template and help initiate RNA synthesis. Elongation involves the productive synthesis of the newly formed RNA transcript along the length of the gene and involves regulation by transcription factors. The final stage of transcription, termination, is the least understood of the stages. In the simplest sense it involves the cessation of transcription, mediated by the dissociation of the ternary complex which consists of the DNA template, the RNA transcript and the RNA polymerase. Our lab is involved in
understanding the process of transcription regulation, especially with respect to termination, in the late stages of vaccinia virus transcription. In the following sections of the introduction, the process of transcription is discussed in prokaryotes, eukaryotes and vaccinia virus, with an emphasis on the various mechanisms of termination employed in the different organisms.

**Prokaryotic Transcription**

Prokaryotic genomes are organized differently and are smaller than even the smallest eukaryotic genome. While in *E.coli*, which has been historically studied extensively as a prototypical representative of the prokaryotic world, the genome is organized as a unipartite, circular DNA molecule, some prokaryotes also exhibit multipartite genomes. The individual genes in prokaryotes are transcribed, in most cases, as multi-gene transcription units, also referred to as operons, which were identified by Jacob and Monod and described as a bacterial phenomenon where expression of functionally related genes was co-regulated. Evidence accrued over the years suggests that operons no longer necessarily include functionally related genes and any given operon is a dynamic unit that can either form or disintegrate or reconfigure over many generations either due to sequence or gene deletion, insertion, translocation or transfer (Osbourn & Field, 2009).

Given the high degree of conservation in their gene products, the variations among the different prokaryotic species are mediated mainly through the regulation of gene expression. The expression of a gene or an operon is highly regulated at the level of transcription, specifically during initiation, by external factors like growth conditions. The specific mechanism through which gene expression is regulated involves three
main components, the RNA polymerase complex, cis-acting promoter sequences and transcription factors, which are discussed in the following paragraphs.

**RNA Polymerase**

The prokaryotic RNA polymerase (RNAP) is a multisubunit enzyme complex that can be isolated as either a core complex or a holoenzyme made up of the core complex in association with additional subunits or factors. The *E.coli* core enzyme subunits were identified by purifying the complex on a phosphocellulose column followed by size exclusion on a sephadex G200 column. The enzyme was shown to be made up of five subunits, two α subunits, and one each of β, β’, ω subunit, with a total mass of about 400 KDa (Burgess, 1969). The α subunits that serve as a scaffold on which the rest of the polymerase complex is assembled, have identical amino acid sequences but functionally one binds the β subunit while the other binds β’. The α subunits have two domains: an amino terminal domain (NTD) that interacts with the rest of RNAP complex, and a carboxyl terminal domain (CTD) that interacts with transcription factors and the promoter DNA, separated by a flexible linker region (Ebright & Busby, 1995). The β and β’ subunits which are the catalytic centers are arranged like the pincers of a crab claw.

The ω subunit, though identified with the rest of the subunits was largely ignored as an associated impurity of purification that was not necessary for the functional RNAP. Subsequent work done in the late 1990s validated its presence as part of the core complex and indicated a requirement for the ω subunit in the assembly of the RNA polymerase core by sequestering the β’ and recruiting it to the α₂β subassembly (Mukherjee & Chatterji, 1997; Ghosh et al., 2001; Mathew & Chatterji, 2006). The RNAP holoenzyme comprises of the core polymerase in conjunction with one of the many
available σ factors, which determine the specificity of the polymerase in promoter binding.

**Initiation**

The promoter region is a major determinant of gene regulation and was identified initially as cis-acting DNA sequences which when mutated affects the rate of initiation of transcription (Scaife & Beckwith, 1966). Classical mutational analyses of the DNA template, biochemical analyses using techniques such as primer extension, s1 nuclease, *in vitro* transcription, gel shift assays and eventually sequence comparison helped define recognition elements in the promoter region that were crucial in transcription initiation (Hawley & McClure, 1983; Mulligan et al., 1984; Lisser & Margalit, 1993). The promoter region has been determined to be a 70-80 bp region spanning from -60 to +20 with respect to the initiation site, +1. The most common and conserved elements of the promoter are the -35 TTGACA and -10 TATAAT hexamers. Additional elements include the UP element, an AT rich region with A and T tracts found between -57 and-38, an extended -10 element (ext) immediately upstream of the -10 and a discriminator element (dis) downstream from the -10 (Shultzaberger et al., 2007; Haugen et al., 2008; Ross & Gourse, 2009).

Basal transcription, defined as the level of transcription seen at a given promoter in the absence of any trans-activating factors, is initiated when the holoenzyme is recruited to and positioned at the promoter by the presence of the multi-domain σ factor. The holoenzyme makes sequence specific contacts with the promoter with the α-CTDs binding the UP element and the σ factor binding the -10, -35 and ext elements, thus forming a transcriptionally inactive closed complex that is positioned at the promoter. An isomerization of the σ subunit unwinds ~12-15 nucleotides of the double-stranded DNA
at the start site, thus forming an open complex which initiates transcription. Additional stability is provided to the promoter-RNA polymerase complex by the β and β’ subunits, which can nonspecifically bind the promoter downstream from the initiation site and also in the spacer region between -10 and -35 elements (Haugen et al., 2008). Additional promoter proximal and distant DNA binding regions also serve as binding sites for transcription associated factors.

A final layer of regulation is added to the basal interaction between the promoter and the RNA polymerase complex by the presence of transcription factors. The factors can function as regulators by either binding DNA elements or by interacting with the holoenzyme. Transcriptional activators bind to regions upstream of the promoter or to the α-CTD and σ subunits and recruit the holoenzyme to the promoter. In contrast, the repressors inhibit either RNAP-DNA interaction by binding regions surrounding the -35 and -10 regions or by competing for specific elements, or inhibit various intermediates during initiation. Other repressors bind and sequester either the sigma or alpha subunits and prevent RNAP holoenzyme formation or RNAP-DNA binding (Browning & Busby, 2004; Beck et al., 2007).

**Elongation**

The transition from an initiation complex to a competent elongation complex involves conformation changes in the RNA polymerase, associated with the process of abortive initiation, where the holoenzyme undergoes a repetitive cycle of RNA chain synthesis and release. When the RNA chain length reaches ~12 nucleotides the polymerase forms a stable ternary complex and clears the promoter (Hsu et al., 2003). This transition is accompanied by but does not require the release of the σ subunit. The elongating polymerase maintains a constant size with a 12-13 nucleotide bubble.
housing an 8–9 nucleotide RNA-DNA hybrid at its catalytic center (Borukhov & Nudler, 2008). The ternary elongation complex is also subject to regulation in the form of template sequence or sequence-directed nucleic acid structures or DNA binding proteins that force the RNA polymerase to pause or backtrack. Productive re-elongation of the ternary complex is mediated by elongation or antitermination factors such as the bacterial GreA, GreB, Nus factors or the bacteriophage N and Q antiterminators through transient interaction with the ternary complex (Borukhov et al., 2005).

Termination

As with both initiation and elongation the process of transcription termination is a regulated process that has been the focus of extensive research. The process is complex and seems to be mediated through several different mechanisms. The prokaryotic system has served as an excellent model to decipher the final stage of transcription. Evidence accrued over the years suggests three major mechanisms of transcription termination that are conserved among the prokaryotes. The following subsections are reviews of our current understanding of the various mechanisms.

Rho-dependent termination

The rho-dependent termination mechanism, which accounts for half the termination events seen in prokaryotes, involves a multisubunit bacterial protein rho (ρ). The rho factor has the distinction of being the first known protein to play a role in the phenomenon of transcription termination. It was discovered in 1969 by Jeffrey Roberts when a partially purified fraction of crude *E.coli* extracts seemed to depress the net RNA synthesis in an *in vitro* synthesis reaction using phage λ DNA as template. The multisubunit factor seemed not only to produce discrete transcripts, indicating an
involvement of a specific site or sequence on the template, but also released these transcripts from the DNA template RNA polymerase complex (Roberts, 1969).

Over the span of the next forty years emphasis was laid on the biochemical and structural analysis of rho in order to determine a mechanism for its termination activity. The enzyme binds nucleic acid in the absence of NTP (Richardson, 1970). An essential component of its function in termination is its RNA-dependent beta-gamma phosphate hydrolyzing ATPase activity with a marked specificity towards poly-cytidine substrates (Lowery-Goldhammer & Richardson, 1974; Howard & de Crombrugghe, 1976; Lowery & Richardson, 1977a; Lowery & Richardson, 1977b). Rho-dependent terminator sequences identified in both bacterial and phage transcript RNA has two specific regions, a 5’ rho utilization (rut) site where rho is thought to bind and load onto the RNA and downstream transcription stop (tsp) region that has a cluster of sites where the elongating RNAP can pause and is made to terminate (Richardson & Richardson, 1996). When provided RNA with the appropriate loading sequence the enzyme has been shown in vitro to translocate along the RNA in the 5’-3’ direction and utilize energy derived from ATP and dATP hydrolysis to unwind downstream RNA-RNA and RNA-DNA hybrids (Brennan et al., 1987), and also displace streptavidin molecules placed at the 3’ end of the RNA (Schwartz et al., 2007).

The biochemical data are well supported by the structural characterization of the enzyme. Rho is a homohexamer with individual protomers shaped and assembled like the wedges of an orange. Each protomer has 419 residues and a molecular weight of 46.8 KDa with three functional domains; an N-terminal primary RNA binding domain that can also bind single stranded DNA, a central ATP binding domain and a C-terminal
secondary RNA binding domain (Dolan et al., 1990; Wei & Richardson, 2001). Rho exists predominantly as homohexamer, but can be found in different states of assembly depending on the ionic environment and the presence of additional cofactors (Geiselmann et al., 1992). The hexameric structure has an open, five subunit conformation that upon RNA binding forms a closed six unit ring conformation.

A classical model for the mechanism of rho mediated termination was proposed based on the above observations and also evidence that while rho-mediated termination is kinetically linked to the elongating polymerase (Jin et al., 1992), rho did not directly bind the core RNAP complex (Schmidt & Chamberlin, 1984). The enzyme was postulated to scan for and load onto the rut site of the newly transcribed mRNA, and track along the transcript in search of the paused ternary complex and dissociate the complex. An alternate model for rho mediated termination has been proposed based on recent evidence (Epshtein et al., 2010) of rho interacting early on with the RNAP and conformational changes to the RNAP paused for dissociation. This model has the enzyme persistently bound to the transcribing RNAP, loading onto the extruding RNA forming a loop that tightens as it pulls on the RNA thereby trapping the elongation complex and ultimately dissociating it by invading the main channel and unwinding the RNA:DNA hybrid.

**Intrinsic termination**

Intrinsic termination is a factor independent mechanism of termination seen in prokaryotes. *In vitro* studies using rho in transcription termination showed that a portion of the transcripts terminated in the absence of additional factors (Roberts, 1969). Termination by this mechanism is thought to be mediated by a conformational change in the transcribing polymerase induced by regions in the DNA template. Analysis of the
rho-dependent and rho-independent or intrinsic termination sites within a given DNA template revealed that the transcribing polymerase pauses at either of these sites and terminates only the intrinsic sites in the absence of rho (Adhya & Gottesman, 1978).

Two key features were identified in the intrinsic transcription termination sites, an RNA:RNA interaction region transcribed from CG rich regions of the DNA with a dyad symmetry and downstream from it an RNA:DNA interaction region that is AT rich. It was surmised that the dyad CG-rich region gave rise to an RNA stem loop structure that in conjunction with the instability of the DNA:RNA hybrid in the AT-rich region caused the elongating ternary complex to pause, alter its conformation and ultimately dissociate (Farnham & Platt, 1981). These features are seen in all rho-independent and some rho-dependent sites. Intrinsic termination sites have been identified at the ends of bacterial and phage operons and are sometimes also seen within operons between two cistrons. The orientation of the stem loop followed by the AT rich region is important and termination occurs heterogeneously past the stem loop region resulting in RNAs with 3' udridylate residues (Holmes et al., 1983).

**Mfd-mediated termination**

Mutation frequency decline (mfd) is a protein that has emerged as an alternate mechanism of termination in bacterial systems. Along with rho-dependent release and intrinsic terminator mediated release, mfd mediated release is one of the three main termination mechanisms in bacterial cells. Mfd is highly conserved across the bacterial genome. Mfd was originally discovered as a protein that decreased the mutations in cells that were subjected to UV irradiation; Mfd mutant cells had an increased sensitivity to UV damage. Mfd is involved in the process of transcription coupled repair (TCR) and hence is also referred to as the transcription repair coupling factor (TRCF). It is capable
of identifying transcription complexes that have stalled due to damaged DNA and recruiting DNA repair enzymes to the site of damage while causing the dissociation of the stalled complex.

Mfd is a monomeric 130 KDa protein. Its structure has been resolved into various functional domains (Roberts & Park, 2004). In the C-terminal end it has seven helicase motifs, a translocase domain and the C-terminal domain. This region of the protein bears a strong homology to the superfamily II helicases, especially to RecG which is involved in DNA repair. Based on the helicase domains the protein was tested for helicase activity and found to contain none. However the helicase motifs account for the ATPase activity and DNA binding activity. Adjacent to the helicase domain towards the C-terminus is a motif homologous to the RecG translocase, the translocase (TRG) motif that might provide the translocase motor functions. The N-terminal region of the protein is involved in recruiting DNA repair enzymes to the site of DNA damage and shares homology with the UvrB region that is responsible for recruiting the repair enzyme UvrA. The central domain of the protein is the RNA polymerase interacting domain (RID). This domain has been shown to bind the β subunit of the RNA polymerase that is present in the upstream region of the ternary complex. Consistent with these are experimental results that show that mfd needs access to about 25 base pairs of DNA upstream of the ternary complex in order to bind it and moreover mfd does not bind the $\sigma^{70}$ subunit containing polymerase complex (Selby & Sancar, 1993).

Not only is mfd involved in DNA damage repair but it can bind stalled elongation complexes and depending on the environment can either promote productive elongation or cause dissociation of the stalled complex. Mfd binds and dissociates from the
elongation complex repeatedly in an ATP dependent fashion. While its binding does not affect normally transcribing elongation complexes, it targets slow moving or stalled, backtracked complexes. Mfd is recruited to the ternary complex by the RID domain and in the presence of ATP binds dsDNA. The translocase and helicase domains help in the forward translocation of the enzyme. The enzyme pushes backtracked complexes into its active state by applying a force on the DNA and the polymerase in opposite directions. If the active complex is no longer hindered it results in productive re-elongation of the complex in vitro. However, in the presence of either a roadblock or insufficient nucleotides or DNA damage, the complex is incapable of further elongation and the force generated by the mfd action results in dissociation of the ternary complex (Park et al., 2002).

**Eukaryotic Transcription**

The eukaryotic genome is transcribed to produce a variety of different RNA products; mRNA, tRNA, rRNA, etc. The protein coding sequences are embedded in the mRNA, which as opposed to the prokaryotic mRNA is monocistronic, each mRNA coding for a single protein. Eukaryotic mRNA undergo extensive processing: splicing to produce variations of the primary transcript as dictated by the environment and 5` and 3` end processing to maintain stability of the primary transcripts within the nucleus. Eukaryotic transcription is carried out by three species of related RNA polymerases. RNA polymerase II (RNAPII) is responsible for transcription of protein-coding genes and many noncoding RNAs, including spliceosomal small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), microRNA (miRNA) precursors, and cryptic unstable transcripts (CUTs). RNA polymerase I (RNAPI) transcribes the abundant ribosomal RNAs (rRNAs), and RNA polymerase III (RNAPIII) transcribes noncoding RNAs such as
RNA Polymerase

The DNA-dependent RNA polymerase II (Pol II) in eukaryotic cells is a 12 subunit, 514 KDa enzyme highly conserved between yeast, human and drosophila which are the model eukaryotic organisms. The yeast core enzyme has 10 subunits: Rpb1, Rpb2, Rpb3, Rpb5, Rpb6, Rpb8, Rpb9, Rpb10, Rpb11 and Rpb12. The peripheral heterodimer is made up of Rpb4 and Rpb7. Core bacterial polymerase shares sequence, structure and functional homology with subunits of the Pol II polymerase in that the largest and second largest prokaryotic subunits, β’ and β, that are involved in catalysis are similar to the Pol II Rpb1 and Rpb2 respectively, the two α subunits involved in polymerase assembly and regulation are homologous to Rpb3 and Rpb11 and the ω subunit involved in assembly has a counterpart in Rpb6 (Allison et al., 1985; Sweetser et al., 1987; Larkin & Guilfoyle, 1997; Minakhin et al., 2001). The rest of the core subunits are either shared between or have homologues within eukaryotic Pol I and Pol III. The Rpb4/7 subcomplex is peripherally bound to the core enzyme, required for promoter-specific initiation and capable of dissociating from and not required for but associated with the elongating polymerase (Edwards et al., 1991; Jasiak et al., 2008). Pol II is also unique among the polymerases in that the largest subunit contains a flexibly linked C-terminal domain (CTD) heptapeptide sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser, varying between 27 repeats in yeast to 52 repeats in humans, whose phosphorylation status determines the position of the polymerase during transcription, by interacting with various binding partners (Buratowski, 2003).
Initiation

Eukaryotic class II genes’ core promoter region can either be focused with a single transcription start site or dispersed with multiple start sites, with the latter being more predominant in vertebrate systems. The core promoter region spans a length of 80 base pairs between -40 and +40 nucleotides with respect to the transcription start site (TSS). Core promoters contain distinct elements, only a subset of which is seen in any given promoter. The most common element is the initiator (Inr) motif which starts at -2 and contains an A that often is the transcription start. The earliest identified core element includes the TATA box consensus sequence TATAWAAR at the -31 position with respect to transcription start site. B recognition elements (BRE) are found either upstream or downstream from the TATA box. A conserved downstream promoter element (DPE) found between +28 to +33 plays an important role in basal transcription. A motif ten element (MTE) found between +18 and +27 acts in synergy with the TATA box and DPE (Yang et al., 2007; Juven-Gershon et al., 2008; Juven-Gershon & Kadonaga, 2010). There are promoter proximal regions and distant gene regions, called enhancers, sometimes several kilobases away that also regulate transcription from a specific promoter.

Pol II by itself is capable only of promoter non-specific transcription. While the eukaryotic polymerase, Pol II, lacks a bacterial σ factor homolog, the promoter specific recruitment and regulation of transcription initiation of Pol II is mediated by a host of transcription factors. Five distinct activities purified from yeast transcription extracts necessary for promoter specific transcription of Pol II (Sayre et al., 1992), were attributed to a conserved group of eukaryotic factors referred to as basal or general transcription factors (GTF), comprising of TFIIB, TFIID, TFIIE, TFIIF and TFIIH. The
GTFs in association with the Pol II can contain as many as 30 subunits with a total mass over 2 MDa. TFIID is a multisubunit complex comprising of a TATA-binding protein (TBP) and 14 TBP associated factors (TAFs). In promoters containing TATA elements, transcription is initiated by the recognition and binding of the element by the saddle-shapped molecule, TBP. Similarly, each member of the GTF and other co-activators such as TFIIA and TBP related factors (TRFs) function to assemble the polymerase at the core promoter to form a pre-initiation complex (PIC), similar to the closed complex in prokaryotes, by binding specific core elements, each other and binding pol II (Dvir et al., 2001; Hahn, 2004).

The process of eukaryotic transcription initiation is also regulated by a host of proteins that function as activators or repressors, with some factors capable of either function depending on the transcriptional environment. Most regulators bind promoter proximal DNA or distant enhancers or bind other regulators, but some have functional kinase, helicase or acetyl transferase activities (Hahn, 2004). In order for the regulators to interact with the GTFs or the polymerase, a higher order complex called mediator is required. The mediator complex, first identified in yeast, is a large multisubunit complex with more than 20 subunits which also acts as a GTF by interacting directly with CTD of the polymerase (Kelleher et al., 1990; Kim et al., 1994; Kornberg, 2005; Malik & Roeder, 2005). A final dimension is added to the complexity of transcription regulation in eukaryotes by genome packaging. DNA is packaged as nucleosomes, which are basic units of the chromatin with ~147 nucleotides of DNA wrapped around a histone octamer. Nucleosomes act as natural repressors of transcription, making the promoter region inaccessible for DNA binding by the GTFs and the polymerase. Depletion of a histone
molecule in yeast has been shown to deplete nucleosomes and activate several genes (Han et al., 1988). Constitutively active genes have open promoters, where the region upstream from the transcription start site (TSS) is depleted of nucleosomes. Highly regulated genes have closed promoters where the region upstream of the TSS is covered with nucleosomes. Transcriptional activation of these covered promoters involves chromatin modifications or chromatin remodeling mediated by transcription activators that recruit modifiers like histone acetyltransferases or remodelers like the SWI/SNF family of proteins (Boeger et al., 2003; Kornberg, 2007; Cairns, 2009).

The combined activity of the GTFs and other activators results in the up regulation of assembly of stable pre-initiation complexes (PIC). PIC assembly occurs in an ordered fashion beginning with the binding of TFIID to the DNA and ends with the binding of TFIIE and TFIIH to form a completed PIC. TFIIH is a multifunctional, 10 subunit enzyme that catalyzes the unwinding of the ~11-15 bp of downstream DNA in an ATP-dependent manner by virtue of its helicase domain, to form an open complex and allow for transcription initiation (Svejstrup et al., 1996). Transcription initiates with the formation of the first phosphodiester bond and proceeds to a process of abortive initiation within 3-10 nucleotides due to the instability of the newly formed DNA-RNA hybrid and the inability of the polymerase to dissociate from the GTFs. TFIIH in addition to other TFs aids the polymerase in promoter clearance by unwinding downstream DNA and phosphorylating the Pol II CTD at the Ser5 residue, thus dissociating the polymerase from the promoter associated factors (Conaway et al., 2000; Fuda et al., 2009). Hyperphosphorylation of the CTD by factors such as P-TEFb pushes the ternary complex into productive elongation.
While mutational and biochemical analyses set the stage for examining the transcription complex, an overall understanding of the topology of the transcription complex and the catalytic mechanisms involved in the various stages of transcription, comes from structural studies. Initial structural studies of both the bacterial and yeast polymerases involved analyses of electron micrographs of negatively stained 2D crystals constrained on a lipid layer, at a relatively low resolution of 30 Å (Darst et al., 1988; Edwards et al., 1990). Subsequent advances that allowed for formation of 3D crystals and heavy atom derivatives helped solve the structure of the polymerase at higher resolutions (Darst et al., 1991; Fu et al., 1999; Cramer et al., 2001). The core prokaryotic and eukaryotic polymerases were found to have similar structural basis. Analyses of the polymerases in various functional states, alone and in conjunction with transcription factors, inhibitors, nucleic acid template and NTP substrates helped determine the architecture of the transcription complex.

An atomic model for the evolutionarily conserved core polymerase involves a positively charged active center contained in a cleft formed by the interaction of the two largest subunits, β' and β in prokaryotes and Rpb1 and Rpb2 in eukaryotes. These flanking subunits resemble the pincers of a crab-claw and have subdomain regions and flanking structures with different functional roles. The largest subunit forms a mobile clamp that can swing over the cleft during transcription to form a closed structure in the presence of DNA and RNA. This subunit also has a bridge helix along the length of this main channel or cleft. An additional helix forms a wall-like structure to bifurcate the main channel (27 Å) into secondary channels (11 Å) near the downstream end. The second largest subunit forms a protein wall that blocks upstream end of the cleft and also
contains a lid and rudder loop regions involved in DNA-RNA strand separation. These loops interact with each other and other protein elements to form exit channels for the newly transcribed RNA and exiting DNA (Westover et al., 2004; Chen et al., 2009). The other subunits bind on the external surface of the catalytic core resulting in an overall size of about 100-150 Å. Transcribing complexes have a two-metal-ion requirement, with one metal ion held in the active center by three aspartate side chains, bound to the 3` end of the growing RNA and another that is associated with the incoming NTP. The active center has two NTP binding sites; i and i+1, a newly formed 3` end binds at i while the incoming NTP binds at i+1 to initiate phosphodiester bond formation between the residues at i and i+1 (Zhang et al., 1999; Gnatt et al., 2001).

During initiation, promoter DNA binds outside the cleft and upon DNA melting to form open complexes, the unwound template strand enters the cleft and is positioned at the active site to initiate RNA synthesis, leading to the formation of a ~15 nucleotide transcription bubble containing a 1-3 nucleotide unwound DNA strand in the leading edge followed by an 8-9 nucleotide DNA-RNA hybrid at the active center. Footprinting analyses have concluded that an overall 14 base region of the newly transcribed mRNA is protected by the elongating complex; the 9 nucleotide hybrid region in the main channel and an upstream 5 nucleotide protected from nucleases within the exit channel (Kettenberger et al., 2004; Cramer et al., 2008).

Elongation

Eukaryotic pol II elongation complexes have been shown to transcribe at the rate of 20-70 nucleotides per second (Darzacq et al., 2007). However, the movement of the elongating Pol II complex along the genome is regulated to be discontinuous due to several factors that are discussed in this paragraph. The ratchet like movement of the
elongating complex in conjunction with template sequence, transcript or protein factors cause the elongation complex (EC) to form paused complexes. Paused complexes are formed when the elongation complex is unstable and at times backtracks leading to the extrusion of one or two nucleotides of the 3' end of the RNA transcript past the catalytic site. Paused complexes resume normal elongation by forward translocation on their own or due to the direct binding or action of elongation factors, without any catalytic modification to the elongation complex. Transcription of AT rich regions leading to unstable hybrids, DNA lesions, nucleosomes and DNA binding road block proteins have been shown to result in paused complexes. Genome wide analysis of density of transcribing complexes has revealed a phenomenon of promoter proximal pausing in eukaryotic genes, where transcriptionally active complexes that have cleared the promoter form paused complexes 20-50 nucleotides downstream from the initiation site (Price, 2008; Core & Lis, 2008). In the presence of a nucleotide analog repressor of transcription, DRB, the activity of negative elongation factors like DSIF which interacts with pol II and NELF which interacts with the pol II-DSIF complex have been shown to coincide with promoter proximally paused complexes (Yamaguchi et al., 1999). Release from pausing is regulated by elongation factors like TFIIF, PTEF-b, Fcp1, ELL proteins, CSB or the elongins that are seen in higher eukaryotes. Although these factors promote re-elongation, depending upon the transcription requirement or the environment, some of these factors can target the paused polymerase for degradation (Sims et al., 2004).

Paused complexes can over time decay into arrested complexes, where many more nucleotides of the 3’ end of the nascent transcript extend past the catalytic site and the polymerase can no longer resume elongation. Re-elongation requires the
cleavage of the extruding 3’ end of the mRNA. This cleavage function is inherent to the polymerase but additional stimulation in the presence of TFIIS which functions similar to the prokaryotic Gre factors has shown to be critical for cell viability (Sigurdsson et al., 2010). An important component of regulation of the eukaryotic elongation complex is the CTD of the largest subunit, Rpb1. The CTD is hypophosphorylated in the PIC and its hyperphosphorylation coincides with promoter clearance. Promoter proximally paused complexes are mostly phosphorylated at Ser5 and as the polymerase transcribes through the gene the CTD becomes dually phosphorylated at Ser5 and Ser2. The phosphorylation status of the CTD determines its binding partners and plays a major role in recruiting positive and negative elongation factors to the polymerase. Since mRNA processing activities such as mRNA capping and splicing has been shown to occur co-transcriptionally in eukaryotes, the CTD in promoter proximally paused complexes serves to recruit binding partners involved in mRNA processing (Glover-Cutter et al., 2008). They also recruit histone modification enzymes such as methylases and acetyltransferases and chromatin remodeling complexes to regulate elongation (Sims et al., 2004).

**Termination**

As the elongating ternary complex progresses towards the final stage of transcription, namely termination, the CTD of the large subunit of RNAP acquires a different signature. These elongation complexes are predominantly phosphorylated in their Ser2 position and lose Ser5 phosphorylation and thereby recruit new regulatory molecules. Most of these molecules play a role in 3’ end processing. In addition to the changes in the ternary complex mediated by the CTD, another facet of transcription of the 3` ends of genes that plays an important role in termination is the polyA signal. The
role of the polyA signal was clearly elucidated from experiments that show that when 3´ ends of the eukaryotic mouse beta-globin gene were ligated into an in vitro template, termination only occurred in cases where the 3´ regions had the polyA signal site included. In addition, experiments where the polyA sequence was altered in vitro showed inactivation of their 3´ end processing, clearly linking both 3´ end processing and termination to the polyA signal (Proudfoot, 1989). However, given the 3´end processing and relative instability of the primary transcript within the nucleus, mapping the 3´ ends of the primary transcript to distinct termination sites, as with prokaryotes, is not possible. In eukaryotes, normal termination of genes occurs at regions downstream of the polyA site and occurs with great heterogeneity, in some cases terminating at various sites over a kilobase region downstream from the polyA signal. While this may hint at termination being a random mechanism, evidence for the presence of various factors that regulate termination have accrued over the years and in the following paragraphs the models postulated based on these data will be discussed.

The primary model for termination among eukaryotes is the allosteric or anti-terminator mechanism, which is based upon two key evidences; the change in the phosphorylation status of the CTD and the polyA signal. The polyA signal in eukaryotes consists of a conserved hexanucleotide, AAUAAA followed downstream by a GU rich region, which are recognized by multisubunit cleavage and polyadenylation specificity factors (CPSF) and cleavage stimulation factors (CstF), respectively (Venkataraman et al., 2005). Recognition of the polyA signal by these factors results in the recruitment of other 3´ end processing factors, some among which are the polyA polymerase (PAP), polyA polymerase binding protein (PAPB) and Pcf11. Pcf11 is an important
polyadenylation factor, shown to be necessary for cleavage and polyadenylation and termination. Pcf11 also has been shown to be recruited by the Ser2 phosphorylated CTD, thereby providing a bridge between the processing factors and the ternary complex. Cleavage and polyadenylation of the mRNA transcript occurs at a region between the hexanucleotide and the GU region. The allosteric or anti-terminator model suggests that as the elongating ternary complex transcribes past the polyA signal in the template, conformational changes to the complex lead to the release of elongation or processivity factors and an association of termination factors, leading to termination of the complex in regions downstream of the polyA site (Richard & Manley, 2009).

Cleavage of the mRNA transcript has been shown to occur co-transcriptionally (West et al., 2008), resulting in two molecules of RNA, one that has its 5` end capped and 3` end associated with polyA factors and another with an uncapped 5` end that is still associated with the transcribing polymerase. The second model for transcription termination among eukaryotes, torpedo model, is based on the observation that these 5` uncapped RNA species were being degraded. Purification of this factor, Rat1, identified a 5`-3` exonuclease seen in abundance in the 3` ends of genes (Kim et al., 2004). Rat1 has not been shown to interact with the CTD and is suspected to interact via the 3` processing factors. While deletion of Rat1 does not affect cleavage at the polyA site, it does inhibit termination. Rat1 chews up the uncapped RNA and in association with Sen1, an associated RNA/DNA helicase, mediates dissociation of the ternary complex (Kawauchi et al., 2008).

An emerging idea in the field is one where the process of eukaryotic termination is a combination of both these models involving changes that occur to the ternary
complex on transcription past the polyA site, thereby mediating dissociation of the ternary complex (Rosonina et al., 2006). While in the prokaryotic world termination has been attributed to three major mechanisms, the mechanism of termination in eukaryotes is still debated upon. The mechanistic details of termination remain unclear and possibly involve a complex interaction between several multi-functional factors. Among the various factors investigated in the context of eukaryotic termination only three have been shown to be capable of dissociating the ternary complex in vitro. TTF2, the best characterized of the three proteins, was identified as a eukaryotic helicase capable of terminating the ternary complex in vitro. It shares similarities with mfd in that its activity is based on energy derived from dsDNA dependent ATPase activity (Xie & Price, 1997). TTF2 has ATP independent nucleic acid binding capacity and no in vitro helicase activity. TTF2 regulates transcription during mitotic repression and is not affected by the phosphorylation status of the CTD. Sen1, another eukaryotic helicase, has been shown to dissociate ternary complexes in vitro. The protein is a superfamily I RNA helicase shown to unwind nucleic acid hybrids. It is a multifunctional protein involved in the termination of snRNA and snoRNA and capable of multiple protein-protein interactions, notably with the CTD of the RNA polymerase (Finkel et al., 2010). Pcf11 the only other eukaryotic protein with the ability to dismantle the ternary complex, has no helicase activity, but is part of the 3’ mRNA end processing factors. Its ability to mediate termination is independent of nucleotides and dependent of its interaction with the CTD (Zhang & Gilmour, 2006).

As with initiation and elongation, the process of termination is also regulated by the presence of pause sites and chromatin remodeling. Even though these mechanisms
have been proposed based on general observations in different genes, across different eukaryotic model organisms, there exist exceptions to both these mechanisms. For example, in the beta-globin gene, cleavage of the mRNA transcript to generate a new uncapped 5' end does not take place at the polyA site but at a specific cleavage site downstream from the polyA signal. Another example of the variations is the 3' end processing of histone mRNAs (Richard & Manley, 2009). These mRNAs lack a polyA tail and are processed by an alternate mechanism. Recent evidence also suggests the possibility of a third mechanism of termination for mRNAs involving complexes involved in the termination of snRNAs and snoRNAs (Rondon et al., 2009).

**Vaccinia Virus Biology**

Vaccinia virus is the prototypic member of the Poxviridae family of large DNA viruses. Poxviruses belong to a group of nucleocytoplasmic large DNA viruses (NCLDV) and replicate entirely in the host cytoplasm. The family can be subdivided into two subfamilies, seven genera, and a dozen species of poxviruses infecting a variety of host species including mammals, birds, and insects. Vaccinia is subgrouped under the Chordopoxvirinae which have a vertebrate host range. The most notorious member of this group is variola, the causative agent for smallpox. Vaccinia, due to its serological cross reactivity to variola and limited pathology in humans was the primary active component in the live vaccine used in the long campaign by the WHO for the eradication of smallpox. It has been used as the common lab workhorse in the efforts to study the various aspects of poxvirus life cycle. In addition to this, vaccinia enjoys significance as a model used to elucidate various aspects of gene expression in higher eukaryotes, as a recombinant viral vector in vaccine therapy and as an oncolytic viral agent.
The infectious virus is a brick-shaped structure with slightly rounded edges and recently reported dimensions of about 360 x 270 x 250 nm (Cyrklaff et al., 2005). Intact virions are presumed to have surface tubule elements as seen from various imaging studies (Heuser, 2005). The virus particle is encapsidated by a lipid bilayer that upon controlled degradation in the presence of a non-ionic detergent and reducing agent leaves behind intact internal structures; a barrel shaped core flanked on either side by a lateral body. The bilayered core wall encloses a tightly packed nucleo-protein genome that is 192 Kb long and condensed in the form of a tubular structure. The lateral bodies are proteinaceous in nature, however their significance is yet to be determined (Condit et al., 2006).

**Vaccinia Life Cycle**

Viral life cycles begin with the entry of the infectious particles into the host cell using one among a variety of mechanisms. Vaccinia viral entry into host cells has been shown to differ based on the type of cell, type of infectious viral particle and even the strain of the virus (Whitbeck et al., 2009; Bengali et al., 2009). The major infectious form of the virus has been shown to enter cells by direct fusion with the plasma membrane, mediated by a group of eight or more proteins that form part of the entry fusion complex (EFC), by direct interactions with either the cell surface glycoseaminoglycans (GAGs) or laminins (Carter et al., 2005). Infectious virions have also been shown to enter through a low-pH dependent endosomal pathway (Townsley et al., 2006). A novel mechanism of entry has been defined via macropinocytosis, where the virus particles by virtue of interactions mediated by the phospholipid molecules on their membrane mimic apoptotic debris and are therefore engulfed into the cell (Mercer & Helenius, 2008).
Upon entry into the cytoplasm, early gene expression within the viral cores is initiated by the virally encoded transcription enzymes packaged within the core and peaks within 1-3 hours post infection, resulting in the transcription of half the viral genome. The newly synthesized viral mRNA is extruded in an ATP-dependent manner into the cytoplasm where it is translated. Early genes trigger the expression of intermediate class of genes which in turn help transcribe late genes leading to a temporal pattern of gene expression. Transcription of the viral genome is described in detail in the following section. A key trigger in the transfer from early to intermediate gene expression is genome replication. Early proteins mediate the uncoating and release of the viral genome, which is then replicated by viral factors in distinct membrane-bound cytoplasmic viral factories. The viral genome is a linear, double-stranded molecule with covalently closed ends and terminal inverted repeat regions.

The final stage of the viral life cycle, virus assembly, is initiated within viral factories by the late gene products at about 5 hours post infection. The earliest identifiable components of viral assembly are crescent shaped structures which grow with time to form closed circular forms called immature virions (IV), with enclosed viroplasm. The viroplasm condenses to form internal nucleoid subdomains to transition from IVs to particles referred to as IVNs. Morphogenesis of IVNs into mature viral particles MVs is associated with the proteolytic cleavage of several viral proteins. MVs are mostly found outside of the viral factories and are released upon cell lysis. Some MVs acquire golgi derived outer membranes to become wrapped virions (WV) and eventually exit the cell via exocytosis losing one layer of the wrapped membrane in the process and keeping one additional membrane more than the MVs to form particles
referred to as enveloped virions (EVs). MVs and EVs are antigenically different as they contain different sets of viral proteins on their surface membrane and re-enter viral life cycle via different pathways. MVs are the most abundant and most stable of the two infectious forms and hence the most widely studied (Condit et al., 2006).

**Vaccinia Virus Transcription**

Vaccinia virus transcription is interlinked with the genome organization. Two key features of the viral transcription, the bidirectionally transcribed viral genome and the temporal regulation of transcription aids in the efficient packaging of the viral genes. The genes are tightly packed within the genome, for example in the HindIII D fragment there is less than 42 bp between adjacent genes (Lee-Chen & Niles, 1988), leading to scenarios where the promoter regions of adjacent genes overlap or the termination signal of one gene is present within the coding region of the adjacent gene. These problems are addressed at the level of transcription by dividing the genes into different temporal classes and transcribing both strands of the genome, and at the level of genome organization by placing different gene classes adjacent to each other. The temporal regulation of gene expression also helps decrease incidences of dsRNA formed from converging transcripts.

The temporal regulation of viral gene expression is commonly used among viruses to regulate the various stages of the viral life cycle. In vaccinia virus, the genome is transcribed as early, intermediate and late classes of genes or in some cases throughout the viral life cycle. The early genes are transcribed within the host cytoplasm inside the infectious viral cores and encode for proteins involved in host immune evasion, viral DNA replication and intermediate gene transcription. Intermediate and late gene transcriptions occur post initiation of and concurrent with genome
replication and are classified as post-replicative gene transcription. Intermediate gene classes encode proteins involved in regulation of late gene transcription and also viral morphogenesis and late transcription in turn encodes for factors involved in morphogenesis and also early transcription factors that are packaged into the newly formed virions.

Transcription among the different classes share many common facets that are described in the succeeding paragraphs, the most important of them being the virus encoded RNA polymerase complex. While the cytoplasm had been recognized as the site of vaccinia virus infection in 1960, it was not until 1967 the existence of the viral RNA polymerase was demonstrated by purifying the virus cores and studying the core-associated RNA synthesis and NTP incorporation activities \textit{in vitro} (Munyon et al., 1967; Kates & McAuslan, 1967). The viral polymerase complex was isolated in 1977 from the cytoplasm of infected HeLa cells as a multisubunit enzyme with distinct chromatographic properties from the host polymerase (Nevins & Joklik, 1977b). It also differed from the bacterial and host polymerases in that unlike the bacterial and mammalian polymerases the viral counterpart was resistant to rifampicin and $\alpha$-amanitin, respectively (Costanzo et al., 1970). The virion associated RNA polymerase was purified and determined to have eight distinct subunits with an overall weight of 500 KDa (Spencer et al., 1980; Baroudy & Moss, 1980). Immunoprecipitaion and mapping studies identified that the RNA polymerase subunits were encoded in the central region of the viral genome (Jones et al., 1987). Sequence analysis of the subunits indicated that the two largest subunits of the viral RNA polymerase, J6 and A24, 147 KDa and 132 KDa respectively, shared sequence similarity with both the prokaryotic and
eukaryotic two largest subunits (Broyles & Moss, 1986; Patel & Pickup, 1989). Among the other subunits, the smallest G5.5 was shown to be homologous to a eukaryotic RNA polymerase II subunit and E4 shared homology with the mammalian transcription factor TFIIS. Structural and biochemical analyses of the functional polymerase complex were done in the context of the early gene specific template and are discussed later on in this section.

Viral mRNAs behave like eukaryotic mRNA in that they have polyadenylated 3’ ends. However the viral polyA polymerase (PAP) bears very little structural and mechanistic resemblance, if any to the cellular PAP. Viral mRNAs containing 3’ polyA tails were seen in 1970 and was determined to be a template independent phenomenon (Kates & Beeson, 1970). The activity was attributed to a stable viral heterodimeric enzyme purified both from cores and infected cytoplasm, with a molecular weight of 80 KDa (Brakel & Kates, 1974a; Brakel & Kates, 1974b; Moss et al., 1975; Nevins & Joklik, 1977a). In vitro activity studies concluded that the heterodimer was capable of adding polyA stretches to oligonucleotide primers at a rate of 1000 adenylate residues per primer at 37°C, the activity being inhibited by dATP and ddATP (Shuman & Moss, 1988). The larger 55 KDa subunit, VP55 was determined to be the polyA polymerase capable of polymerizing 30-35 nucleotide stretches of polyA beyond which its processivity decreases (Gershon et al., 1991; Gershon & Moss, 1992). The smaller 39 KDa subunit, J3 was determined to act as the processivity factor, increasing the processivity of the polyA polymerase by up to 40 fold (Gershon & Moss, 1993). J3 which is present in molar excess and can be isolated as both heterodimers and monomers, plays other roles in viral transcription regulation as listed below. Although the
heterodimer purifies independently of the RNA polymerase, it could interact with the RNAP through associations with other polymerase associated factors (Mohamed et al., 2001).

The 5` ends of viral mRNA are also processed similar to eukaryotic mRNA. The viral capping enzyme was purified from disrupted viral cores and identified by its ability to modify the 5` ends of uncapped mRNA (Ensinger et al., 1975; Martin et al., 1975). The enzyme has been shown to have three catalytic functions in mRNA capping that have been attributed to individual domains of this heterodimeric protein, encoded by the 95 KDa, D1 subunit and the 33 KDa, D12 subunit. The amino terminus of D1 has been shown to be responsible for an RNA triphosphatase function that acts on the 5` triphosphate of the nascent mRNA to produce a diphosphate end (Shuman & Morham, 1990). Also contained in the amino terminus of D1 is the guanylyltransferase activity that forms a bridge between the 5` end of guanosine monophosphate with the previously formed 5` diphoaphate end (Cong & Shuman, 1995). The carboxyl terminus of D1 in conjunction with D12 has the methyltransferase activity that methylates the N7 position of the terminal guanine residue (Cong & Shuman, 1992; Higman et al., 1994).

**Early gene expression**

With the purification of the vaccinia RNA polymerase in 1980, experiments conducted in the following decade helped decipher the cis-acting template sequences important for early transcription. Mapping of the 5` termini of several early genes, both in the terminal repeat and the central regions of the genome, by S1 nuclease analysis and primer extension revealed that regions about 60 bp upstream from the transcription start sites were 88% AT rich and the transcribed early mRNAs had a 50 bp 5` UTR sequence (Venkatesan et al., 1981; Venkatesan et al., 1982; Cochran et al., 1985;
Rosel et al., 1986). Recombinant viruses generated by linking reporter cassettes with different lengths of the upstream regions of the early gene TK helped define the 5' and 3' ends of the early promoters as being between -31 and -8 bp respectively (Coupar et al., 1987; Wei & Richardson, 2001). Mutational analysis of a 33 bp segment of the promoter region was carried out by creating 99 mutations in the 33 bp region, annealed to a lacZ reporter. The consensus drawn from these results and other studies indicate that vaccinia early promoter is highly flexible and can be divided into three elements; an A rich critical core region separated from a 7 bp initiation region by an 11 bp T rich spacer region. The high level of flexibility seen in early promoters is defined by the lack of a consensus sequence and the most conserved nucleotides seem to be at the A on -20,-19 and -13. The initiator is defined at position +1 and requires a purine nucleotide. The spacer region is less important for transcription and is an 11 bp long T rich region. The core region, which can be compared to the TATA box in eukaryotic pol II transcription, is 15 bp long and critical for recognition and even single substitution mutations in this region can have a drastic effect on transcription. Multiple copies of the critical region in a given promoter, even when they overlap are capable of initiating transcription from the appropriate independent downstream sites leading to a hypothesis where the critical region serves as a recognition site to bind TFs and RNAP and initiation begins at a downstream region as dictated by the stereochemistry of the ternary complex at that site (Davison & Moss, 1989a).

The transcription factor associated with early initiation in the virus is the virally encoded heterodimer VETF, vaccinia early transcription factor, identified during the purification of the polymerase from vaccinia virus cores (Broyles et al., 1988). VETF
comprises of an 82 KDa subunit encoded by the late gene A8L and a 70 KDa subunit encoded by the late gene D6R, both subunits sharing no homology with any known prokaryotic or eukaryotic transcription factor (Broyles & Fesler, 1990; Gershon & Moss, 1990). The D6 subunit has a helicase motif with a functional ATPase domain in the amino terminus, contributing to the DNA-dependent ATPase function of VETF (Broyles & Moss, 1988). VETF has been shown to bind early promoters independent of its ATPase activity, with A8 interacting with regions -12 to -29 in a sequence specific manner and D6 interacting sequence-independently with regions +8 to +10, relative to the initiation site thereby inducing a structural alteration in the promoter region that promotes transcription (Broyles et al., 1991; Cassetti & Moss, 1996). The ATPase activity of VETF although not required for DNA binding is required for its dissociation from the promoter region and also transcriptional activation (Broyles, 1991; Li & Broyles, 1993; Li et al., 1994). VETF has been shown to recruit the viral early RNA polymerase to the early promoter region forming a functionally stable ternary complex with the formation of a tetrameric RNA (Hagler & Shuman, 1992d; Baldick, Jr. et al., 1994).

The interaction between the VETF and the early vaccinia polymerase is mediated by an important component of the early vaccinia polymerase that sets it apart from the post-replicative complex, the presence of the RNA polymerase associated protein, RAP94, the 94 KDa product of the late gene H4L (Ahn & Moss, 1992). RAP94 is found associated with 30-40% of all polymerase complexes isolated from the core and only this fraction of the polymerase can interact with VETF and is active in transcription (Ahn et al., 1994). Recent evidence for the RAP94 mediated interaction of the RNA polymerase and VETF demonstrated that both the subunits of VETF were involved in an
interaction with the amino terminal end of RAP94, while the polymerase complex was associated with the carboxyl terminus (Yang & Moss, 2009).

Initiation of early gene transcription during vaccinia virus infection seems to involve a step-wise efficient mechanism that involves the formation of the preinitiation complex (PIC) and promoter clearance. VETF, the promoter recognition unit of the PIC binds the DNA and recruits early polymerase holoenzyme by virtue of its interaction with RAP94 to form a functional PIC. Although ATP is not required for the formation of the PIC, binding of the holoenzyme to form the PIC has a divalent metal ion requirement (Baldick, Jr. et al., 1994). Promoter clearance is achieved with the ATP-dependent dissociation of the VETF from DNA both in the promoter region and downstream from the initiator, allowing for the holoenzyme to proceed into elongation.

The well defined components of the early core vaccinia transcription initiation have allowed for \textit{in vitro} experiments involving purified factors. \textit{In vitro} transcription assays with paramagnetic bead bound templates containing viral early promoters followed by a G-less stretch can be transcribed in the presence purified early vaccinia polymerase, VETF and added NTPs to form an elongation competent ternary complex. The ternary elongation complex (TEC) once formed is stable in the presence of salt, sarkosyl and heparin concentrations that inhibit complex formation during initiation (Hagler & Shuman, 1992d). Elongation is accompanied by co-transcriptional mRNA processing, as shown by the presence of 5’ caps on mRNAs greater than 30 nucleotides (Hagler & Shuman, 1992a). The capping enzyme associates with the elongation complex via interaction with the carboxyl terminus of RAP94 (Yang & Moss, 2009). Structural analysis of the TEC by DNase I and chemical footprinting revealed an
overall footprint of 48-49 nucleotides on the non-template strand and 40-41 nucleotides on the template strand. The leading edge of the polymerase was 17 nucleotide from the catalytic site on the template strand and the concurrent lagging edge about 23-24 nucleotides. The transcription bubble was identified to stretch to 17-18 nucleotides, with a leading edge 3 nucleotides downstream from the catalytic start site and a 9 nucleotide long DNA-RNA hybrid (Hagler & Shuman, 1992c). RNA footprinting analysis of the elongation complex reveal 18 nucleotides of mRNA from the 3`end protected within the ternary complex (Hagler & Shuman, 1992a). Elongation complexes have been shown to transcribe at the rate of 20-50 nucleotides per second (Hagler & Shuman, 1992b), however like their counterparts in the prokaryote and eukaryote systems the rate of elongation is regulated by template and NTP concentration driven pausing (Deng & Shuman, 1997). The vaccinia RNA polymerase also possesses inherent cleavage activity to help regulate backtracking during elongation (Hagler & Shuman, 1993).

Termination of early vaccinia transcription takes place in a sequence and factor dependent manner in the presence of energy derived from ATP or dATP hydrolysis. The TTTTTNT sequence, present about 50 nucleotides upstream from the 3` ends, functioning as the minimum sequence required for termination was identified by analyzing 3` ends of early mRNA and confirmed by mutagenesis (Yuen & Moss, 1987; Lee-Chen et al., 1988). Modification of the T5NT DNA sequence and its cognate U5NU RNA sequence determined that termination was mediated by the RNA (Shuman & Moss, 1989). The two factors involved in early termination are the Vaccinia Termination factor (VTF) and the Nucleosidetriphosphate Phosphohydrolase I (NPH I). VTF was identified to be the viral capping enzyme (Shuman et al., 1987). Although both D1 and
D12 subunits of the capping enzyme were required for termination, the termination function of the enzyme is independent of the capping activities including its ATPase activity (Luo et al., 1995; Yu & Shuman, 1996). However VTF was shown to recognize and bind the U5NU sequence provided it was 30 nucleotides away from the 3` end. Recognition of the termination signal and energy dependent termination of the ternary complex is kinetically linked by the presence of NPH I, a ssDNA-dependent ATPase (Deng & Shuman, 1998). NPH I was found to be activated by ssDNA binding, suggesting the ssDNA in the transcription bubble is an appropriate binding area for the NPH I (Condit & Niles, 2002). NPH I associates with the transcribing early RNA polymerase by virtue of its carboxyl terminus interaction with the amino terminus of RAP94 (Mohamed & Niles, 2000).

Based on the work done on the individual factors involved in vaccinia early transcription termination and their interactions with each other the model proposed for early termination is discussed below. The vaccinia early elongation complex consists of the early polymerase in association with the RAP94 subunit. RAP94 recruits NPH I to the polymerase via its amino terminal end. RAP94 has been shown capable of recruiting mRNA processing factors such as the capping enzyme and the J3 polyA polymerase via its amino terminal end, but neither of these factors is required during elongation (Mohamed et al., 2001). Termination of early transcription requires both VTF and NPH I and in the absence of VTF NPH I has been shown to stimulate elongation (Deng & Shuman, 1998). Efficient termination has been shown to be inversely proportional to the rate of elongation and the distance between the 3` end and the termination sequence is dependent on the NTP concentration. Mechanistically,
termination is initiated by the recognition of the U5NU motif in the nascent mRNA about
30-50 nucleotides from the catalytic site as it exits the transcription complex, by RAP94
and VTF. This interaction serves to signal the NPH I molecule associated with the
ternary complex to mediate termination via energy acquired through ATP hydrolysis
(Christen et al., 2008; Piacente et al., 2008).

**Post-replicative gene expression**

Intermediate and late promoters were identified in similar fashion as the early
promoters by mapping the 5` ends of intermediate and late gene products. The
upstream regions were compared and characterized by mutational analysis to reveal
the key functional elements. Intermediate promoters appear similar to early promoters in
size and % AT content. The promoter regions of five intermediate genes annealed to a
reporter cassette were analyzed to identify the three key elements; a 14 bp core
element separated from a short initiator by a 10-11 by spacer that is equivalent to one
turn of the DNA helix. The core and spacer regions did not show any sequence
specificity, however the initiator region had a common TAAA element between -1 to +3
with the first A residue serving as the transcription start site (Hirschmann et al., 1990;
Baldick, Jr. et al., 1992). Some but not all intermediate promoters had GG residues
downstream from the initiator that seemed to play a role in transcription factor binding
(Broyles, 2003). Late promoters also possessed the same three key elements, however
the core region was found to be a 20bp AT rich region with T-tracts crucial in deciding
the promoter strength, the spacer element was 6 bp and the initiator had a TAAAT
element containing the transcription start site (Hanggi et al., 1986; Davison & Moss,
1989b).
Post-replicative gene transcription takes place in the host cytoplasm and hence is complicated by its requirement for host proteins in addition to viral factors. This has made it difficult to study in detail the post-replicative ternary complexes and their characteristics. While the viral proteins involved in intermediate and late transcription initiation have been identified, it is still unclear if the list of host proteins identified so far is a comprehensive one. The first attempt made to reconstitute intermediate gene specific transcription in vitro identified the need for two fractions of virus infected cell extracts, in addition to the eight subunit viral RNA polymerase. Both these fractions were absent in the core extracts indicating that they were early viral proteins or host proteins (Vos et al., 1991a). Subsequent purifications helped characterize the factors involved in intermediate gene transcription to be the viral capping enzyme (Vos et al., 1991b), VITF-1 which is encoded by the viral protein E4 (Rosales et al., 1994), VITF-3 which is a heterodimer of the viral proteins A8 and A23 (Sanz & Moss, 1999) and VITF-2 which is a host heterodimer (Katsafanas & Moss, 2004). The viral capping enzyme was found to be associated with an affinity tagged purified RNA polymerase used in these assays. The E4 protein also functions as a subunit of the viral RNA polymerase and shares homology with the eukaryotic cleavage factor, TFIIS. The A8, A23 heterodimer are conserved among poxviruses but have no cellular homologs. VITF-2 is a heterodimer of p137 and G3BP, involved in cell activation but have no identified role in transcription.

Late gene transcription factors were identified by co-transfecting candidate intermediate viral genes along with DNA containing a late promoter under infection conditions where DNA replication was inhibited (Keck et al., 1990). The three viral gene
products G8, A1 and A2 were and later a cellular factor VLTF-X was identified to be required for late gene transcription. VLTF-X activity was identified as belonging to heterogeneous nuclear ribonucleoproteins (hnRNP) A2 and RBM3 (Wright et al., 2001). In addition, the TAAA sequence in the initiator of intermediate and late gene promoters have been reported to be targets of host TATA-binding protein (TBP) and YY1 (Knutson et al., 2006; Knutson et al., 2009). While a clear mechanism for intermediate and late transcription initiation is yet to be defined, the body of evidence proves that transcription in both these classes initiates at the A residue of the TAAA initiator and undergoes abortive initiation at this site resulting in the addition of a 30 nucleotide polyA segment at the 5` ends of the mRNA transcripts.

Post-replicative transcription elongation is regulated in a sequence and factor dependent manner. Elongation of the post-replicative ternary complex has been shown to be affected by sequence or NTP concentration dependent pausing (Prins et al., 2004). Factors involved in regulating post-replicative transcription complexes were identified through genetic analysis and interaction studies. Phenotypic characterization of temperature sensitive mutants revealed an interesting post-replicative transcription associated phenotype at non-permissive temperatures in a virus mutant in the A18R gene. The mutants synthesized normal early transcripts but longer than normal intermediate and late transcripts implicating a role for the protein as a negative elongation factor (Pacha & Condit, 1985). A similar phenotype was seen when infections were carried out in the presence of the anti-poxviral drug IBT, indicating the drug targeted a factor(s) associated with transcription regulation, which was later identified to be the second largest subunit of the viral RNA polymerase, A24.
Subsequent work done with IBT resistant and dependent mutants revealed viral mutants in two different genes G2R, which encodes for a 26 KDa protein and J3R, the viral polyA polymerase, with the same elongation defect phenotype in postreplicative gene transcription (Black & Condit, 1996; Latner et al., 2000). These observations suggest a model of a post-replicative ternary complex that is regulated by positive elongation factors such as G2 and J3 and negative elongation factors such as A18. This model is further strengthened by evidence that these factors interact with each other. It has been shown that mutations in the A18R gene can rescue G2R mutants (Condit et al., 1996b). *In vitro* interaction studies have shown that A18 immunoprecipitates with G2. Moreover, G2 co-purifies and interacts *in vivo* with a viral protein, H5, which has been shown to stimulate late gene transcription. *In vitro* interaction of H5 with A18 and evidence that H5 mutants were IBT-resistant includes this factor among those that are presumed to regulate post-replicative ternary complexes (Black et al., 1998; Cresawn et al., 2007).

Vaccinia post-replicative transcripts possess 5` caps and 3` polyA tails much like their early counterparts. However, they differ from the early transcripts in that they have polyA heads at their 5` ends and are highly heterogeneous in their 3` ends, resulting in transcripts from each gene that vary in length by a few kilobases (Mahr & Roberts, 1984). This indicates the lack of a discrete termination sequence for post-replicative termination. In the event that termination does involve DNA sequence, it must involve an element that is ubiquitous or degenerate or both. Two viral factors have been implicated to play a role in transcription termination; A18 and H5. A18 is a 56 KDa viral protein that has been shown to release nascent mRNA from stalled elongation
complexes, *in vitro*. The characteristics of the protein will be discussed in detail in the following section. Evidence for the role of host protein(s) in post-replicative transcription termination have been reported in studies with A18 (Lackner & Condit, 2000). Despite the apparent heterogeneity among post-replicative transcripts, there exists evidence for distinct 3’ end formation by site-specific cleavage among a few genes; A24R, F17R, ATI (Patel & Pickup, 1989; D’Costa et al., 2004). The cleavage factor responsible for cleavage of F17 transcripts was identified by conventional purification to be H5. H5 is a multifunctional tetramer with an apparent molecular mass of 440 KDa on a gel filtration column (D’Costa et al., 2008). The protein has been implicated in DNA replication, virion morphogenesis and post-replicative transcription regulation (D’Costa et al., 2010). Given their association with each other and their role in post-replicative transcription regulation, the model for post-replicative transcription termination will likely involve a concerted mechanism of action between both H5, in its role as a transcript cleavage factor and A18, in its role as a transcript release factor.

**A18**

The vaccinia virus A18R gene product is a 56KDa protein that has been shown to function *in vitro* as a transcript release factor. Based on its sequence homology to the DEXH subfamily of DNA and RNA superfamily II helicases, the protein was tested for ATPase activity (Bayliss & Condit, 1995). A18 was shown to have a DNA-dependent ATPase activity. It has the capacity to cleave the β-γ phosphodiester bond in ATP, dATP and to a lesser extent in the other ribonucleotides using ssDNA, dsDNA or DNA:RNA hybrids, as short as 17 nucleotides, as cofactors. However the ATPase activity is not stimulated in the presence of either ssRNA or dsRNA. Similarly the purified histidine tagged protein product was tested for helicase activity and is known to
be capable of effectively unwinding in the 3’-5’ direction a DNA:DNA hybrid no greater than 20-22nt in length in the presence of ATP or dATP (Simpson & Condit, 1995).

A18 was initially identified during the characterization of temperature sensitive mutants of the virus. Infections with the three non-complementing ts mutants – cts22, cts23, cts4 – which produce thermolabile A18, revealed an interesting abortive late phenotype where infections with the mutants at non-permissive temperatures have protein synthesis comparable to the wild-type infections up to 8 hours post infection. Between 8-10 hours post-infection all protein synthesis comes to an abrupt irreversible stop in the mutant infections at non-permissive temperatures. Viral mRNA and host rRNA was found degraded under these conditions. Since the pattern of RNA degradation was similar to that seen during RNase L degradation in viral infections (Pacha & Condit, 1985; Cohrs et al., 1989), the role of A18 in maintaining RNA stability was investigated. RNase L degradation is the downstream function of the 2-5A pathway in which dsRNA-activated-2-5A synthetase uses ATP to form 2’-5’-linked adenylate oligomers which activate RNase L. Mutant infections at nonpermissive temperatures showed a higher level of both dsRNA and 2-5A during late infection. However, since A18 does not possess RNA helicase activity and does not form a stable complex with RNA (Simpson & Condit, 1995) it seems unlikely that A18 by itself functions to inhibit the 2-5A pathway (Xiang et al., 1998).

The A18 ts mutants also showed a promiscuous transcription phenotype where early genes that are transcriptionally silent become activated late during infection. This along with A18’s shared homology with members of the helicase family especially human EERC3 and yeast RAD25 which are involved in transcription, prompted
investigations into the role of A18 in vaccinia transcription. The protein is expressed throughout the virus lifecycle and also found packaged in the virions (Simpson & Condit, 1994). *In vivo* studies show that the A18 mutants have normal early transcripts and defective post-replicative transcription phenotype. The mutants produced longer than normal later transcripts that explained the activation of silent genes since these genes were present downstream of transcribed late genes. Moreover mutations in an early gene G2R that produced shorter-than-normal transcripts were shown to suppress the A18R ts mutants. *In vivo* immunoprecipitation data showed the interactions between A18 and G2 and also with H5 (Black et al., 1998), yet another protein involved in late gene transcription. These results implied a possible role played by A18 in post-replicative transcription termination, the mechanism of which is not completely understood.

An *in vitro* assay with immobilized templates containing vaccinia virus promoters concluded that A18 is a transcript release factor (Lackner & Condit, 2000). A18 required the presence of uninfected cell extracts and ATP hydrolysis to release nascent transcripts from stalled elongation complexes. A18 mediated release of transcripts apparently required the protein to be present during transcription elongation. Release was mediated from templates having early, intermediate or late promoters. However, the exact mechanism by which A18 induces release *in vitro* or its role in post-replicative transcription termination *in vivo* is unknown.

**Significance of this Study**

Termination of post-replicative genes in vaccinia virus is distinct from early gene termination in the virus in that it lacks a distinct termination sequence leading to the formation of post-replicative transcripts with heterogeneous 3’ ends. However, this does
not mean that the process of post-replicative transcription termination is a spontaneous process. It is in fact a factor-mediated process, regulated by the presence of the viral post-replicative transcription termination factor, A18. Although the protein has been shown to mediate release of nascent mRNA transcript *in vitro*, the mechanism by which termination is mediated by this enzyme is still unknown.

Transcription termination in prokaryotes is a mechanistically easier process to study given the simpler experimental system and lack of 3′ end processing. Eukaryotic termination is a more complicated scenario that is poorly understood, complicated by the 3′ end processing reactions. Moreover eukaryotic termination shares similarities such as the lack of cis-acting termination signals and heterogeneous 3′ ends with vaccinia post-replicative transcription. Vaccinia post-replicative transcription termination proves a relatively simpler model to study termination with the overall goal of understanding transcription regulation in eukaryotes. This project was aimed at studying A18 mediated termination in the context of its *in vitro* function as a transcript release factor. Experiments were proposed to investigate the termination properties of the enzyme that would lead to an understanding of the mechanism of vaccinia virus post-replicative termination. In addition we aimed to probe the nature of the host factor requirement in A18 mediated termination. The experiments described in chapter three of this dissertation describe the efforts at a partial characterization of the host factor component. As part of this chapter, is described the isolation of stalled ternary complexes which give an insight into the role of A18 during elongation. Chapter four contains the results from experiments using these stalled complexes to investigate the biochemical properties of A18 mediated termination. In addition this chapter also
contains experiments studying A18 mediated termination in the context of transcriptional pausing.
CHAPTER 2
MATERIALS AND METHODS

Eukaryotic Cells, Prokaryotic Hosts, Viruses and Extracts

A549 cells, human lung carcinoma adherent cultures were maintained in Dulbecos’s Modified Eagle’s Medium with 10% FBS as described earlier (Condit et al., 1996a). Frozen HeLa cell pellets were obtained from NCCC. Conditions for wild-type vaccinia virus strain WR growth, infection and plaque titration have been described (Condit & Motyczka, 1981). Bacterial expression host used was BL21 pLysS. Drosophila nuclear extracts were a gift from Dr. David Price, University of Iowa. Yeast extracts were obtained through Dr. Paul Gollnick, University of Buffalo.

Plasmids

The templates used in the in vitro transcription assays were from the plasmids pG8G and pG8GU. Both plasmids were derivatives of the pC2AT19 described previously (Sawadogo & Roeder, 1985) to have been constructed from an insertion of a 375 bp G-less segment with an average composition of bases 2C,A,T, into pUC13. G-less cassettes are a common feature of transcription templates; transcription through a G-less cassette in the absence of GTP or the presence of 3’OMeGTP, a GTP analogue, stalls the ternary complex at the end of the cassette. pG8G had a vaccinia virus intermediate promoter, G8, with a 3’ A in place of its 3’ G, inserted 5’ of the G-less cassette (Condit et al., 1996a). pG8U plasmid was derived from pG8G and contains the G8 promoter followed by a 37 nucleotide G-less cassette and then a 41 nucleotide A less cassette and finally by a stretch of 9 Ts (Condit et al., 1996a; Prins et al., 2004). Both plasmids were extracted using Qiagen plasmid megaprep kits.
The plasmid p16A18 contains a amino terminally polyhistidine tagged A18 inserted into the pET16b vector (Bayliss & Condit, 1995).

**Transcription Competent Lysolecithin Extracts**

Intermediate gene transcription competent extracts are made using lysolecithin as described previously (Condit et al., 1996a). A549 cells grown to confluence in 100 mm dishes in DME media containing 10% FBS were infected with wild-type vaccinia virus strain WR at a multiplicity of infection of 10-15 plaque forming units in PBS/AM at 37°C. Infection was allowed to proceed for 14-16 hours in the presence of 10 mM hydroxyurea at 37°C. The cells were washed and permeabilized on ice with lysolecithin in a buffer containing 150 mM sucrose, 30 mM Hepes pH 7.4, 50 mM KoAc, 4.5 mM Mg(OAc)_2. The cells were then harvested and proteins extracted in a buffer containing 25 mM Hepes pH 7.4, 50 mM KoAc, 1 mM dithiothreitol, 7.5% glycerol by scraping the cells off the dish, triturating them with a Pasteur pipette followed by micrococcal nuclease (Worthington) treatment and collecting the clarified supernatant from a centrifugal spin designed to mimic the removal of the nuclei. The total protein concentration was determined for each preparation of extracts and aliquots of this supernatant were stored at -80°C and used in the transcription assays described below.

**HeLa Cytoplasmic Extracts**

A ten liter volume of HeLa cells, obtained from the National Cell Culture Center, Minneapolis as frozen cell pellets, were resuspended in two volume of buffer A (25 mM Tris-HCl, pH 8, 1 mM MgCl₂, 0.1 mM EDTA, 2 mM dithiothreitol, protease inhibitor cocktail (Sigma)) containing 10 mM NaCl, and allowed to swell for 10 minutes on ice. Cells were dounce homogenized and checked for cell lysis, and the nuclei were separated from the cytoplasmic fractions by centrifugation at 32000 x g for 30 minutes.
at 4°C. The cytoplasmic fractions in the supernatant were resuspended with an equal volume of Buffer A containing 800 mM NaCl and 20 % glycerol was added drop wise with stirring such that the cytoplasmic fractions were in a solution containing Buffer A with 400 mM NaCl and 10% glycerol. The extracts were centrifuged at 150,000 x g for 2 hours in a swinging bucket rotor to derive a clarified supernatant of the HeLa Cytoplasmic extracts.

The clarified supernatant was applied to a 5 ml DEAE column that had been equilibrated with Buffer A containing 400 mM NaCl, 10 % glycerol in the same buffer. The DEAE flow-through was collected and dialyzed overnight in the cold into a buffer containing 25 mM Hepes pH 7.5, 80 mM KoAc, 1 mM dithiothreitol, 10 % glycerol and protease inhibitor cocktail (Sigma). The total protein concentration was determined using the Bradford protein assay and the aliquots stored at -80°C without loss of activity.

**Preparation of His-A18**

*E.coli* Lysate Preparation

BL21 pLysS cells containing pET16b A18 vector, expressing A18 tagged with polyhistidine at its amino terminal end, was used to inoculate a 100 ml of Luria Bertini (LB) broth containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol and incubated overnight in a shaker at 37°C. The starter culture was used to inoculate 5 L of LB media with ampicilin and chloramphenicol and incubator in a shaker at 37°C with frequent OD measurements taken at 600 nm. At OD 0.6-0.8, the cultures were induced with 1 mM IPTG and allowed to incubate for an additional four hours. The cells were pelleted at 5000 x g at 4°C and the pellets stored overnight at -80°C. The following day the pellets were thawed and resuspended in 250 ml of lysis buffer containing 25 mM Tris pH 7.0, 1 mM MgCl₂, 500 mM NaCl, 20 mM immidazole and 10 % glycerol. Fresh lysoyme and
triton-x 100 were added to a final concentration of 50 µg/ml and 0.1 %, respectively. The mixture was allowed to incubate on ice for 30 minutes and then sonicated on ice for ten sequences with a microtip with 15 seconds on and 45 seconds off. The insoluble material was removed by centrifuging at 39,000 x g for 30 minutes at 4°C. The supernatant was filtered through a 0.45 µm filter.

**Purification on a HisTrap Column**

A 2 ml HisTrap column prepacked with nickel (GE Healthcare) was equilibrated and the lysates were applied to the column at a flowrate of 2 mls per minute. The column was washed with 20 column volumes of Buffer I (25 mM Tris pH 7.0, 1 mM MgCl₂, 500 mM NaCl, 1 mM dithiothreitol, 10 % glycerol and protease inhibitor cocktail) containing 20 mM imidazole, followed by a second wash with 20 column volumes of Buffer I containing 50 mM imidazole. The protein was eluted in 20 column volumes of Buffer I, using a 50 mM – 500 mM imidazole gradient and fractions collected. The peak fractions were collected and dialyzed overnight in the cold into an A18 dilution buffer containing 40 mM Tris pH 7.0, 20 mM KCl, 1 mM dithiothreitol, 40 % glycerol and protease inhibitor cocktail. The fractions were assayed by SDS PAGE and western blot analysis to identify the protein and concentration determined by Bradford analysis. A total of 4.2 mg of protein was stored at -80°C.

**Templates Used in In Vitro Transcription Assays**

**Preparation of pG8G Template**

Templates were prepared by linearizing 100 µg of pG8G plasmid with the restriction enzyme, Nde1 (NEB), resulting in a 220 bp non-specific DNA region upstream from the G8R promoter. The templates were then end filled in the presence of klenow, dATP, dGTP, dCTP and biotin-16-dUTP. The products were purified using
Roche high pure PCR purification kit to separate out the biotinylated DNA molecule from the dNTPs. The template was then bound to 1.2 ml of M280 streptavidin-coated paramagnetic beads (Invitrogen). The average template per µl of bead was calculated to be at 50 fmoles. Prior to binding the beads were washed in equal volume PBS containing 0.1 % BSA, followed by an equal volume to Tris-EDTA. Binding is carried out in Tris-EDTA buffer containing 1 M NaCl, at 37°C for 30 minutes.

**Preparation of pG8GU Template**

Templates were prepared by PCR amplification of a 522 bp region from the pG8GU plasmid. The forward primer, TCGACACCTGCGGTGTGAA, was biotinylated at the 5` end and binds a region 229 bp upstream from the G8R promoter. The reverse primer, GCACCCCAGGCTTTACACTTT, binds a region 169 bp downstream from the 9T site. The biotinylated PCR fragment was purified from the unincorporated free biotin using Roche high pure PCR purification kit. The template was then bound to an appropriate volume of M280 streptavidin-coated paramagnetic beads (Invitrogen). Binding was carried out as described above.

**In Vitro Transcription Release of Stalled Ternary Complexes**

Transcription was initiated by incubating 5 µl of beadbound pG8G template with 10 µl of A549 extracts in the presence of transcription buffer (25 mM Hepes pH 7.4, 80 mM KoAc, 5 mM MgCl₂, 1 mM DTT) containing 1 mM ATP, 1 mM UTP, 6 µCi (80 µM) α-³²P labeled CTP in a final reaction volume of 25 µl at 30°C for 10 minutes. This generated ternary complexes with 5` radiolabeled transcripts. The ternary complexes were isolated on ice by applying the reaction tube to a magnetic stand and removing the clear supernatant. The beadbound ternary complexes were washed with 1.2 volumes of high stringency HSSW (25 mM Hepes pH 7.4, 1 M KoAc, 0.1 % sarkosyl, 5 mM MgCl₂,
1.6 mM DTT and 7.5 % glycerol) followed by three washes with 1.2 volumes of LSW (25 mM Hepes pH 7.4, 80 mM KoAc, 5 mM MgCl₂, 200 μg/ml acetylated BSA, 1.6 mM DTT and 7.5 % glycerol). The complexes were resuspended and allowed to chase at 30°C for 10 minutes in transcription buffer containing 1 mM ATP, 1 mM CTP, 1 mM UTP, 200 μM 3’OMe-GTP, 40 U/μl RNasin (Promega). The ternary complexes elongate to the end of the G-less cassette in the template where they are blocked from further elongation in the absence of GTP, to form stalled ternary complexes (STC). STC were isolated on ice and washed once with 1.2 volumes of LSW. The complex was resuspended into transcription buffer containing 40U/μl RNasin. For formation of STC, the reactions were carried out in multiples depending on the requirements of each assay, generating a master mix. 18 μl of the STC containing master mix was aliquoted into individual reaction tubes containing reagents as required by each experiment, such that the final volume remained 25 μl and the final buffer concentration was that of the transcription buffer. The reactions were incubated at 30°C for 10 minutes. The reaction tubes were applied to a magnet on ice and the components separated into beadbound fractions and supernatant fractions. The fractions were resuspended in 25 μl formamide loading dye, denatured at 96°C for 2 minutes and loaded onto a 6 % denaturing urea-PAGE. Gels were imaged and quantified by autoradiography and phosphorimager analysis.

**In Vitro Transcription Release of Paused Elongation Complexes**

Transcription was initiated by incubating 5 μl of beadbound pG8GU template with 20 μl of A549 extracts in the presence of transcription buffer (25 mM Hepes pH 7.4, 80 mM KoAc, 5 mM MgCl₂, 1 mM DTT) containing 1 mM ATP, 1 mM UTP, 1.25 μM CTP, 6 μCi (80 μM) α-³²P labeled CTP, 200 μM 3’OMe-GTP in a final reaction volume of 25 μl.
at 30°C for 20 minutes. This generated ternary complexes with 5’ radiolabeled transcripts that were stalled at the end of the 37 nucleotide G-less cassette. The ternary complexes were isolated as above and washed with 1.2 volumes of HSSW followed by three washes with 1.2 volumes of LSW. The complex was resuspended in transcription buffer containing 40U/µl RNasin. The reactions were carried out in multiples depending on the requirements of each assay, generating a master mix. 18 µl of the master mix was aliquoted into individual reaction tubes containing NTPs and other reagents as required by each experiment, such that the final volume remained 25 µl and the final buffer concentration was that of the transcription buffer. The reactions were allowed to chase at 30°C for 10 minutes. The reaction tubes were applied to a magnet on ice and the components split into beadbound fractions and supernatant fractions. The fractions were resuspended with 25 µl formamide loading dye, denatured at 96°C for 2 minutes and loaded onto a 10 % denaturing urea-PAGE. Gels were imaged and quantified by autoradiography and phosphorimager analysis.
CHAPTER 3
PROPERTIES OF THE RELEASE ASSAY AND HOST FACTOR

A18 was identified during the characterization of temperature sensitive mutants to function in vivo as a negative elongation factor during post-replicative vaccinia transcription. An in vitro function for the purified protein was described using an in vitro transcription assay which demonstrated the release of nascent mRNA transcripts during transcription in the presence of A18 and additional uninfected cellular extracts (Lackner & Condit, 2000). In this chapter of the dissertation, I will describe the specific aims that pertain to experiments done to understand the basic components of the in vitro assay: the role of A18 and the host extracts in mediating release of the mRNA transcripts and the nature of the host activity that aids in transcript release.

Specific Aim 1: Is A18 Mediated Termination Functionally Dependent on Elongation?

The in vitro assay that was previously used to study A18 mediated release activity was based upon in vitro transcription assays. The template used in these assays is a streptavidin coated paramagnetic bead bound DNA containing a vaccinia intermediate gene specific promoter, followed by a 375 nucleotide G-less cassette. Transcription was initiated by incubating the templates in the presence of vaccinia virus infected cell extracts and limiting radiolabeled nucleotide to form ternary elongation complexes containing transcripts radiolabeled in the 5' region. These complexes were washed at a maximum stringency of 1M salt to remove unincorporated nucleotides and accessory factors and allowed to re-elongate in the presence of ATP, CTP, UTP and 3'OMeGTP and varying combinations of A18 and host extracts. The reactions were then separated as bead bound and supernatant fractions and assessed for the presence of mRNA transcripts and classified as ternary complex associated (bound) transcripts or
transcripts released into the supernatant, respectively. Based on this assay it was previously determined that A18 and additional host extracts were required for the release of RNA from bead bound templates implying the dissociation of the vaccinia ternary complexes. This release of the nascent mRNA transcripts required energy in the form of ATP and the activity was independent of the presence of A18 during initiation of transcription. However the requirement for A18 during elongation in order to mediate termination of the ternary complex, or the optimal reaction conditions for termination independent of elongation could not be determined.

**Isolation of the Stalled Ternary Complex**

The study of the properties of termination of a ternary complex independent of its elongation function involves the isolation of a stable stalled ternary complex (STC). Previous attempts at isolating such a complex proved unsuccessful due to the high background release of transcripts, due possibly to the inherent complex instability or the presence of termination factors still associated with the elongating ternary complex. Our efforts at isolating a stable stalled ternary complex began with experiments that sought to strip the ternary complex by processing it under high stringency wash conditions.

Stalled ternary complexes are isolated using *in vitro* transcription assays. The DNA template, pG8G, was bound to streptavidin coated paramagnetic beads. This allowed for easy isolation of the template and any complex associated with it at various points during transcription by applying the reaction tube to a magnetic stand; the beads stuck to the walls of the tube leaving behind a clear, bead free supernatant. The pG8G template possesses a vaccinia virus intermediate promoter, G8, followed by a 375 nucleotide G-less cassette. As described in the materials and methods, transcription through this cassette in the absence of GTP or presence of 3’OMeGTP stalled the
ternary complex at the end of the G-less cassette. Transcription was initiated by incubating the templates with extracts prepared from infection of A549 cells with wild type vaccinia virus in the presence of hydroxyurea to enrich for intermediate transcription factors. The reactions were incubated with saturating amounts of ATP and UTP and limiting \( \alpha^{-32}\text{P} \) labeled CTP at 30\(^{\circ}\)C for 10 minutes to generate transcription complexes with mRNA transcripts labeled in the 5` regions. The complexes were washed under high stringency wash conditions in buffers containing 1 M salt in the presence of 0.1 % sarkosyl. The complexes were then further elongated to the end of the G-less cassette in the presence of ATP, UTP, CTP and 3’OMeGTP. This stalled ternary complex could be successfully isolated and washed under transcription conditions to recover a stable complex that could be used in further experiments (Fig. 3-1A).

**Are A18 and the Host Extracts Required During Elongation In Order to Mediate Termination?**

The stalled ternary complex was used to determine an important feature of A18 mediated termination of vaccinia intermediate gene specific ternary complexes. A18 has been shown to be unnecessary during transcription initiation. The nature of its interaction with the elongating ternary complex and the importance of such an interaction for its function during termination were undetermined. We incubated the STC either by itself as a control to test its stability or in the presence of A18 or the uninfected cellular extracts or both at 30\(^{\circ}\)C. The reactions were then applied to a magnetic stand on ice and separated into beadbound fractions or released fractions and probed for the presence of the radiolabeled transcript by electrophoresis of the fractions on a denaturing polyacrylamide gel (Fig. 3-1B). The autoradiogram shows that in the
absence of any additional factors, the STC is stable under the reaction conditions. The percent of released transcripts when quantified and graphed indicate that while the termination of a stalled ternary complex requires the presence of both purified A18 and uninfected cellular extracts, neither of these termination components are not required to associate with the ternary complex during elongation in order to mediate termination.

Specific Aim 2: Characterization of the Host Factor Activity

A18 mediated termination in vitro requires a factor present in uninfected host cells. Initial experiments identifying A18’s role in termination involved comparison of activity in wild-type virus infected extracts to that of extracts infected with a temperature-sensitive virus defective in the A18R gene product. The purified protein although active in in vitro nucleic acid binding, ATPase and helicase assays could not mimic the termination properties of the wild-type virus infected extracts. Termination by the purified protein required the presence of additional uninfected cellular extracts. Attempts at purifying the factor(s) showed that while the activity had defined chromatographic binding properties, in that it bound anion exchange columns and hydroxyapatite columns but not cation exchange columns, however the activity eluted under different conditions with a broad profile and no apparent increase in specific activity thus preventing purification of the factor. The host factor(s) are present in uninfected HeLa cellular extracts, both nuclear and cytoplasmic. In the experiments described in this dissertation we use as “host factor” a high speed supernatant of HeLa cytoplasmic extracts that have been passed over a DEAE column in 0.4 M salt to remove nucleic acids. The experiments in this specific aim describe our attempts at understanding and defining the nature of the host factor activity in the cellular extracts.
Non-specific Substitution of Host Factor Activity

Given that the host activity that aided A18 mediated termination could not be fractionated to homogeneity under different investigator non-specific purification protocols, we questioned the validity of host factor activity in the *in vitro* termination assay. In an effort to determine if the host factor activity was due to a non-specific component of the uninfected extracts that did not specifically affect termination, but affected the *in vitro* assay artifactually we attempted to substitute for the host activity in the *in vitro* termination assays. Stalled ternary complexes were isolated as previously described and incubated in the presence of ATP with A18 and host factor extracts to serve as a positive control. Incubation of STC under these conditions with the absence of the host factor served as a negative control. Termination was tested by substituting the host factor with non-specific nucleic acids (poly dI/dC, tRNA), nucleic acid binding proteins (TRAP - a bacterial RNA binding protein, H5 – a vaccinia viral multi functional protein with nucleic acid binding properties, ssDNA binding protein), non-specific proteins (BSA, actin), a polyamine (spermidine), polysaccharide (glycogen, heparin), a phospholipid (lysolecithin), a non-ionic detergent (NP-40 substitute), phosphocreatinine and translation initiation factor fractions, The host factor activity could be partially substituted by the presence of NP-40 substitute (NP-40 s) and completely substituted by initiation factor (IF) fractions (Fig. 3-2). Initiation factor fractions (a gift from Dr. James B. Flanegan, University of Florida) are ribosome associated translation initiation factors isolated from uninfected HeLa cellular extracts. Substitution with IF was not unexpected as attempts at fractionating the host factor activity by differential centrifugation had shown that the activity did not partition with a specific cellular compartment and was
present in both soluble and insoluble fractions. Partial substitution of the host activity by NP-40 substitute suggested a possible non-specific effect and was further investigated.

**Detergent Stimulation of A18-mediated Termination**

We investigated the ability of detergents to substitute for the host factor activity during *in vitro* A18 mediated termination. Based on the evidence that 0.1 % NP-40 substitute (NP-40s) partly substituted for the host factor activity, we titrated the detergent into the termination assay in place of the host factor(s) (HF) (data not shown). The assay was conducted by incubating STC with varying amounts of NP-40s between 0.01 % - 0.2 % NP-40 substitute in the presence of A18 and ATP. The same reactions were carried out in the absence of A18 to exclude the possibility that detergent affected the stability of the complex thus resulting in dissociation of the complex. Although the detergent effect on A18 mediated termination was specific, there was no discernible difference in the activity between the different concentrations (data not shown). We investigated the effect of other detergents such as tween 20, triton x-100 and CHAPS on the termination assay. STCs were allowed to terminate under the four control reaction conditions, in the absence of any factors and in the presence of A18 and HF added alone and in combination. The four conditions were tested also in the presence of each of the detergent (Fig. 3-3). The results indicated that while the detergents stimulated A18 termination activity, it could neither completely substitute for the host factor extracts nor synergistically affect the termination potential of A18 in combination with the host factor(s). The spike in termination potential of A18 in the presence of the detergents could possibly be attributed to a stabilizing effect of the detergent on the enzyme.
Heat Stability of the Host Factor Activity

The heat stability of the host factor activity was investigated. Host factor extracts were heated at 37, 45, 60 or 90°C for up to 1.5 hours. The extracts were tested for their ability to aid A18 mediated termination by incubating them with the STC in the presence of A18 and ATP at 30°C (Fig. 3-4). Extended incubations at 37 or 45°C did not compromise termination activity. However at temperatures above 60°C, the host factor activity was rapidly inactivated.

Host Factor is a Protein

The heat sensitivity of the host factor activity suggested that the active component in the cellular extracts could be a protein factor(s). However as the factor has proved inaccessible to conventional purification protocols, we investigated the nature of the active components of the cellular extracts. The extracts were treated with immobilized TPCK treated trypsin (Piercenet). Briefly, extracts were incubated in an equal volume of 20 mM Tris pH 8.3 containing no trypsin or a minimum of 7.5-75 U of immobilized trypsin. The reactions were incubated at 37°C for an hour, centrifuged to remove trypsin and the supernatant applied to a STC in the presence of A18 and ATP (Fig. 3-5A). The trypsin treated extracts when analyzed by SDS gel electrophoresis showed a decrease in the protein profile in comparison to untreated controls (Fig. 3-5B), demonstrating successful degradation. Termination of the STC decreased with increasing concentrations of trypsin treatment. The mock treated reaction containing no trypsin retained its activity indicating that incubation did not compromise the extracts. In order to confirm that the reactions were unaffected by any residual trypsin in the extracts we added trypsin treated extracts to the positive control and noted no changes to the
activity. These data indicate that a protein component of the extracts was responsible for the host factor activity.

The host factor activity was also tested for its sensitivity to nuclease treatment. Briefly, extracts were treated with 0-48 ug/ml of micrococcal nuclease in the presence of calcium chloride at 30°C for 15 minutes. The samples were chilled and treated with EGTA, the precipitates were removed by centrifugation and the supernatant applied to STC in the presence of A18 and ATP (Fig. 3-5C). Effective nuclease treatment of the samples was confirmed by agarose gel analysis (Fig. 3-5D). The results showed an absence of nucleic acid specific decrease in termination activity, indicating the host factor activity in the extracts is nuclease resistant.

Host Factor is a Ubiquitous Eukaryotic Protein

The host protein(s) required for A18 mediated termination in the in vitro assays was ubiquitous in that upon fractionation of the cellular components, the activity was present in both soluble and insoluble fractions. Moreover the inability to substitute for the factors indicated that specific cellular factors were involved in A18 mediated termination. We were interested in testing if these cellular factors were specific to mammalian extracts. The presence of the host proteins required for A18 mediated termination, in yeast and drosophila extracts were determined by titrating in these extracts in reactions containing the STC and A18 in the presence of ATP and assessing the percentage of terminated complexes (Fig. 3-6). The results indicate that the host protein(s) involved in A18 mediated termination is a ubiquitous eukaryotic protein(s). We tested bacterial lysates to determine if the host factors were present in prokaryotic extracts. The results were inconclusive due to inherent nucleases in the extracts. In
reactions containing the bacterial extracts the labeled transcripts were either degraded or an increased background release of transcripts in the absence of A18 was seen.
Figure 3-1. Termination of stalled ternary complexes (STC). A) pG8G template has a vaccinia intermediate gene G8R promoter, shown as the yellow box, followed by a G-less cassette, shown as the blue box. The template was bound to a streptavidin coated paramagnetic bead, depicted by the brown circle. Transcription of the beadbound template in the presence of vaccinia infected cell extracts and radiolabeled nucleotide, followed by high stringency washes generated a pulse labeled complex. This complex was then chased to the end of the G-less cassette to form a functionally stable stalled ternary complex that can be further manipulated to study termination. B) STC was incubated in the presence of A18 and host factor (HF), as indicated in the autoradiogram of the separated bead bound (▼) and released (▲) fractions. The results were quantified and graphed as percentage released.
Figure 3-2. Non-specific substitution of host factor activity. STC were incubated for 10 minutes at 30°C in the presence of A18 and ATP either alone (negative control), or host factor extracts (positive control) or other non-specific components such as poly dl/dC (500 ng), tRNA (5 μg), TRAP (5 μg), H5 (3.5 μg), ssDBP (5 μg), Actin (5 μg), BSA (10 μg), glycogen (10 μg), spermidine (25 μg), heparin (25 μg), NP – 40 s (0.1 %), lyssolecithin (5 μg), IF (3.4 μg), p-creatinine (5 μg). The reactions were then separated and the presence of the labeled transcript in the bound and released fractions quantified and represented here as the percentage released transcripts.
Figure 3-3. Detergent stimulation of A18 mediated termination. The termination potential of STCs was assayed by incubating the complexes under the four test reaction conditions with no added detergents as a control. Similar reactions were carried out in the presence of a detergent. The results are represented as the percentage of the total transcripts that were released into the supernatant.
Figure 3-4. Heat stability of the host factor activity. Host factor extracts were heat treated to different degrees for increasing lengths of time as indicated. The heat treated extracts were assayed for their activity in the termination assay by incubating STCs in the presence of untreated or heat treated host factor extracts and A18 and ATP.
Figure 3-5. Host factor is a protein. A & B. Trypsin treatment of host factor extracts. Host extracts were treated with zero (M), 7.5 U, 15 U, 38 U or 75 U of trypsin and loss of activity was analyzed by the termination assay (A) and SDS PAGE (B). C & D. Micrococcal nuclease sensitivity of the host extracts. Host extracts were treated with 0 (M), 12, 24 or 48 µg/ml micrococcal nuclease and tested for loss of activity in both the termination assay (c) and agarose gel electrophoresis (D).
Figure 3-6. Host factor is a ubiquitous eukaryotic protein. HeLa cytoplasmic extracts (HCE), drosophila nuclear extracts (DNE), yeast extracts (YE) and bovine serum albumin (BSA) controls were titrated into termination reactions containing STC, A18 and ATP. The percentage of total transcripts released into the supernatant was quantified and is graphically represented.
CHAPTER 4
BIOCHEMICAL CHARACTERISTICS OF A18 MEDIATED TERMINATION

The isolation of the stalled ternary complex (STC) made it possible to investigate the properties of the termination of vaccinia post-replicative ternary complex in the presence of A18 independent of the elongation properties of the transcription ternary complex. The biochemical characteristics of A18 mediated termination are detailed in this chapter. An important pre-requisite for termination of ternary complexes is the ability of ternary complexes to pause. The role of pausing of the ternary complex during A18 mediated termination is also investigated in this chapter.

Specific Aim 3: Properties of A18 Mediated Termination of Stalled Ternary Complexes

The optimal conditions for A18 mediated termination were determined by terminating the STC in the presence of A18 and HF under different reaction conditions. These studies helped assess not only the reaction optima for A18 mediated termination, such as salt and divalent metal ion requirements, but also the enzyme kinetics and energy requirements for the process of termination.

Kinetics of Termination of STC

The kinetics of the termination reaction was examined. The concentration dependence of A18 mediated termination was determined by incubating the STCs with increasing concentrations of purified histidine tagged A18 in the presence of host factor and ATP. The reactions when separated into bound and released fractions and analyzed for the release of mRNA transcripts show a linear increase in termination with increasing concentrations of A18. Release of 50 % of the transcripts was achieved at 20 ng of the purified protein and saturation levels of release under 60 % were achieved.
with 100 ng of the protein (Fig. 4-1A). Higher concentrations of A18 when tested mediated no further release of transcripts (data not shown).

The time course of A18 mediated termination was determined by assessing the percentage of nascent mRNA transcripts released from a STC when such a complex was incubated at 30°C in the presence of A18, host factor and ATP, for varying lengths of time. The results when quantified show that the rate of termination increases linearly with time and reaches maximum by 10 minutes, beyond which length of time no further increase in termination is noted (Fig. 4-1B). These data indicate that termination of the STC occurs in a concentration and time dependent manner.

**Salt Optima for Termination**

Ternary complexes in the various systems, prokaryotic, eukaryotic and even early vaccinia early complexes have been shown to have distinct salt sensitivities. The salt sensitivity of the ternary complex also varies along with the stage of transcription. Termination of vaccinia early polymerase has been shown to be highly salt sensitive, with a marked inhibition of termination seen in the presence of as much as 50 mM NaCl. Transcription initiation, elongation and hence termination under our reaction conditions were carried out at 80 mM KOAc. In order to determine the salt optima for and sensitivity of the process of termination of STC generated from vaccinia intermediate gene specific promoters, we titrated in varying concentrations of the different salts during A18 mediated termination reactions. Briefly, STC were washed in transcription buffer in the absence of salt. The complexes were then terminated in the presence of A18, host factor and ATP in buffers containing varying concentrations of salt. A18 and the host factor were added prior to the indicated concentrations of salt.
Minimal levels of termination (30 %) over the background (10 %) seen regularly in the absence of any factors is seen, indicating that termination of STC occurs independent of salt (Fig. 4-2). Increase in termination is seen with increasing salt concentrations with optima between 100 – 150 mM salt, beyond which concentration we see a gradual inhibition of termination. Concentrations over 200 mM salt could inhibit interaction of A18 and host factor with the STC or each other, thereby inhibiting termination. Of the different salts tests, the STC is most stable in reactions containing KoAc, as seen from the autoradiogram in figure 4-2. In the presence of KoAc, destabilization of the ternary complex, indicated by the presence of lower molecular weight transcripts is seen only in concentrations of 0.5 M salt. In the presence of the other salts, lower molecular weight transcripts indicating unstable complexes appear at salt concentrations as low as 200 – 250 mM salt. At higher salt concentrations (0.5 M) of these salts, an increase in the terminated products was observed which can be attributed to the destabilization and dissociation of the complex in the presence of high salt.

**Divalent Metal Ion Optima for Termination**

The divalent metal ion optima for termination was tested by incubating STCs that were washed and resuspended in transcription buffer lacking magnesium chloride, in reactions containing 100 ng A18, 8.5 μg HF and increasing concentrations of magnesium chloride between 0 – 20 mM. These magnesium titrations were carried out under different ATP concentrations, 0.2 mM, 1 mM and 5 mM. Termination was quantified as the percentage of transcripts released into the supernatant for each reaction and represented graphically (Fig. 4-3). A18 mediated termination of vaccinia intermediate ternary complexes is dependent on magnesium ion. This is not surprising
given that A18 belongs to the superfamily II helicases whose true substrate is the Mg:ATP complex. The reaction is optimal at 5 mM magnesium beyond which concentration the reaction is inhibited possibly due to the formation of dimagnesium-ATP complexes. The percentage of terminated products increases in an ATP dependent manner.

**Energy Requirement for Termination**

A requirement for hydrolysable ATP during termination was initially identified; however the nature the energy requirement for A18 mediated termination could not be examined since the process of termination was coupled to transcription elongation. Isolation of STC and hence the uncoupling to the two processes allowed us to study the energy requirement for A18 mediated termination. STCs were terminated in the presence of A18 and HF and 1 mM concentrations of different NTPs and dNTP (Fig. 4-4A). The results show that in the absence of ATP or the presence of a non-hydrolysable ATP analog, AMP-PNP, termination is inhibited. Termination is mediated in the presence of ATP and dATP but none of the other NTP/dNTP. The energy requirement for termination mimics the energy requirement described for A18 in an *in vitro* ATPase assay, indicating the significance of the enzyme’s ATPase function during termination.

Based on the energy requirement of the enzyme, we determined the optimal ATP concentrations for the termination process. STC were allowed to termination in presence of A18 and host factor under increasing concentrations of ATP. The reactions were incubated at 30°C for 2, 5 or 10 minutes. The saturation curves derived for these reactions indicate that for our reaction conditions the ATP maxima is achieved at 0.5 mM ATP (Fig. 4-4B). Based on the initial velocity curves of the reaction we were able to
determine an approximate Km for the enzyme in the termination process of 200 µM ATP.

**Specific Aim 4: A18 Mediated Termination of Paused Complexes**

Transcriptional pausing, which allows for a temporary halt of the ternary complex during the process of elongation, forms a key intermediate between elongation and termination. Not all paused complexes terminate, since paused complexes are capable of subsequent re-elongation. However, in the various termination mechanisms examined, some of which were described in the introduction to this thesis, we see that pausing of the ternary complex is necessary during termination. In our lab the role of A18 has been examined both during elongation and on artificially stalled ternary complexes. The experiments described in the following section of the thesis examine the role of A18 in termination with respect to ternary complexes that are capable of pausing.

In order to examine A18 mediated termination in the context of paused complexes, we utilized a previously discussed template system. A schematic representation of the transcription scenarios with this system is presented in figure 4-5. The bead bound pG8GU template, as described in the materials and methods, contains the vaccinia intermediate gene G8R promoter, followed sequentially by a 37 nucleotide G-less cassette, a 41 nucleotide T-less cassette, a 9 T region and 169 nucleotide non-specific downstream template (Fig. 4-5). Transcription through the G-less cassette in the presence of radiolabeled nucleotide and 3'oMeGTP generates a pulse labeled transcript (P) stalled at the end of the 37 nucleotide cassette. These complexes are washed under high stringency conditions and allowed to resume elongation under different nucleotide conditions to derive different end products. When the ternary complex is elongated with
1 mM NTPs in the absence of UTP, the complex transcribes through the T-less cassette and pauses at the 9 T site, giving rise to transcripts referred to as T9 transcripts. Upon elongation in the presence of a full complement of NTPs, the ternary complex transcribes to the end of the template to generate a full length (FL) transcript. Elongation in the presence of 1 mM NTPs and limiting concentrations of UTP generates a mixed population of T9 and FL transcripts.

**Termination of Paused Elongation Complexes**

We investigated the effect of A18 mediated termination on paused complexes by adding A18 and HF to elongation complexes under increasing concentrations of UTP. The reactions were then separated into beadbound and released fractions and analyzed on a 10 % urea polyacrylamide gel (Fig. 4-6). Heterogeneity of the pulse labeled transcripts can be attributed to the 5' poly A heads that are representative of vaccinia post-replicative transcripts. Elongation in the absence of UTP generated T9 transcripts that fractionated with the bound fractions in the absence of A18 and in the released fractions in the presence of A18. Elongation in the presence of saturating (1 mM) UTP generated run-off FL transcripts that were split between the bound and released fractions in the absence of A18 and were completely released in the presence of A18. Intermediate concentrations of UTP generated a mixed population of T9 and FL transcripts that were also released in the presence of A18. Interestingly, under these elongation conditions, discrete intermediate size transcripts between 150 – 200 bases on the ladder were apparent. These transcripts appear only in the presence of A18 and are released into the supernatant. A similar but diminished effect is noted in the presence of 1 mM UTP indicated by the appearance of a single intermediate sized transcript in the released lane in the presence of A18. The results imply that the
presence of both A18 and the HF during elongation reveals the presence of cryptic pause sites of the ternary complex by mediating termination at these sites. Analysis of the template sequence downstream from the 9 T site revealed several stretches of T-rich regions which could act as pause sites in the presence of limiting concentrations of UTP.

**Early Time Course of Elongation**

The results from the experiment in Figure 4-6 indicate that cryptic pause sites are revealed in the presence of A18 and HF, but not in the presence of HF alone. The absence of these paused polymerase complexes when A18 is absent could be attributed to the extended incubation time (20 minutes) to which reactions are subjected. Polymerases have been shown to form transient pauses, lasting between 1-6 seconds, during the normal course of elongation. The ternary complexes used in our assays were generated from infected cell extracts and then stripped under high stringency wash conditions. We wanted to confirm the ability of the polymerase in our reactions to pause during elongation by examining the early time points during elongation. Pulse-labeled complexes (P) were allowed to elongate under limiting (40 μM) UTP conditions for 110 seconds. Time points were collected every ten seconds and analyzed on a 10 % urea polyacrylamide gel (Fig. 4-7). The results indicate the appearance of the T9 transcript within 10 seconds of elongation, increasing in intensity with time. The T9 transcripts form the major pause product and most of the transcripts still remain as pulse labeled transcript. This is due to short span of the elongation time and indicates only the early events during elongation. Even within this time span it is clear from the appearance of the gradient of increasing transcript lengths with
increasing time that the ternary complexes that are being generated and investigated under these conditions form transient pauses during elongation.

**A18-specific Termination of Paused Complexes**

A18 mediated termination has shown a dependence on host factor(s) in the context of the stalled ternary complex. In this section we investigate the host factor dependence of A18 mediated termination of paused complexes. Pulse labeled complexes were generated as described in the preceding sections and elongated in the presence of limiting (40 μM) UTP. Elongation was carried out under increasing concentrations of A18, both in the presence and absence of the host factor extracts (Fig. 4-8). Elongation in the absence of A18 and the host factor under these conditions results in the formation of a mixed population of T9 and FL transcripts in the bound fractions. The presence of the host factor by itself shows similar results. An overall decrease in transcripts seen in the presence of the host factor could be due to nucleases present in the extracts. Addition of A18 to these reactions results in the appearance of the intermediate length paused complexes. Increasing concentrations of A18 mediates a decrease in the amount of FL and higher molecular weight paused complexes, indicating that A18 negatively affects transcription. Importantly, even in the absence of the host factor A18 seems to mediate termination in a population of the paused complexes as evidenced by the presence of transcripts in both the bound and released lanes in the absence of the host factor. However, complete termination of the paused complexes does require the presence of the host factor.
Figure 4-1. Kinetics of A18 mediated termination. A) STCs were incubated in the presence of 8.5 μg HF and 1 mM ATP with increasing concentrations of A18. B) STCs were incubated with 100 ng A18, 8.5 μg HF and 1 mM ATP at 30°C for increasing lengths of time as indicated. Termination of ternary complex in both these experiments is represented by the percentage of total transcripts released into the supernatant is represented.
Figure 4-2. Salt optima for A18 mediated termination. STCs were terminated under different salt conditions, separated as bound and released fractions and run on a 6 % denaturing urea polyacrylamide gel. Each row of the autoradiogram represents a different salt assayed in the termination reaction. In each gel, the lanes below ▼ represent the bound fractions and the lanes below ▼ represent the released fractions. The results from the autoradiogram are also represented graphically in the panel below.
Figure 4-3. Metal ion optima for termination. STCs were terminated under increasing concentrations of magnesium ion, in the presence of A18 and HF. Termination reactions were carried out under different ATP conditions. Part A represents the titration between 0 – 20 mM magnesium chloride. Part B is a view of the finer concentrations of the titration curves.
Figure 4-4. Energy requirement for termination. A) SECs were terminated in the presence of A18 and HF. The reactions were separated as bound and released and termination represented as the percentage of released transcripts. The NTP or dNTP used as the energy source for each individual reaction is as indicated. B) SECs incubated with A18 and HF were titrated into reactions containing increasing concentrations of ATP. Successful termination at 2, 5 or 10 minutes was assessed by quantifying the percentage of total transcripts that were released into the supernatant and is represented graphically.
Figure 4-5. Transcription of paused complexes. The beadbound (brown circle) pG8GU template has a vaccinia G8R promoter (yellow box), followed sequentially by a 37 nt G-less cassette (blue box), a 41 nt T-less cassette (green box), a 9 T site (red box) and downstream non-specific DNA (grey box). Transcription through the G-less cassette with radiolabeled nucleotide and in the absence of GTP generates complexes with a pulse labeled transcript (P). Elongation of these ternary complexes in the absence of UTP pauses the complexes at the 9T site generating T9 transcripts. Elongation in the presence of saturating UTP generates complexes with full length (FL) transcripts. Elongation under limiting UTP conditions generates mixed populations of complexes.
Figure 4-6. Termination of paused complexes. Transcription of beadbound pG8GU with radiolabeled nucleotide and 3’oMeGTP generated pulse labeled transcripts (P), which were washed and elongated under increasing concentrations of UTP, either in the presence of HF alone or HF and A18 as indicated above each autoradiogram. The reactions were separated into bound ▽ and released ▼ fractions and run on a 10 % urea polyacrylamide gel. T9 indicates transcripts formed from complexes paused at the 9 T site and FL indicates read-through full length transcripts.
Figure 4-7. Early time course of elongation. Pulse labeled complexes (P) that had been washed under high stringency conditions and allowed to elongate in the presence of 1 mM ATP, 1 mM CTP, 1 mM GTP and 40 μM UTP at 30°C for 110 seconds. Time points were collected every 10 seconds and analyzed on a 10% urea polyacrylamide gel. (T9 – 9 T paused transcript; FL – full length transcript; L – ladder)
Figure 4-8. A18 specific termination. Pulse labeled complexes (P) were generated and elongated with 1mM ATP, 1 mM CTP, 1 mM GTP and 40μM UTP, increasing concentrations of A18, either in the presence or the absence of HF as indicated above. The reactions were separated into bound ▼ and released ▽ fractions and run on a 10% urea polyacrylamide gel. (T9 – 9 T paused transcript; FL – full length transcript; L – ladder)
Transcription of vaccinia intermediate and late classes of genes is carried out in the host cytoplasm. The generation of intermediate promoter specific transcription complexes \textit{in vitro} by incubating virus infected cell extracts with beadbound templates is an important tool to study the process of viral post-replicative transcription. In this study we have shown that stable forms of the ternary complex can be isolated by stripping the intermediate transcription complexes in high stringency conditions containing 0.1 % sarkosyl in the presence of 1 M salt. We have shown that these complexes are functionally active, capable of elongation and termination. These active complexes can be used to probe the roles of the both viral and host proteins in the regulation of vaccinia intermediate and presumably late transcription. In our study, we have manipulated these complexes to study the process of A18 mediated termination of vaccinia intermediate genes.

**Host Factor Requirement for A18 Mediated Termination**

In contrast to early gene expression, vaccinia post-replicative gene expression is regulated by cellular factors due to the cytoplasmic site of gene expression. As described in the introduction, host factors have been shown to play a role in both intermediate and late transcription initiation. Host factors presumably regulate the process of transcription elongation. It has been shown previously that the process of post-replicative termination also has a host factor requirement. Although the cellular component that aids termination has proved resistant to our attempts at purification, we have in our study shown that the host factor(s) is a protein(s) component of the host extracts, due to the trypsin sensitivity and heat sensitivity of the active components. The
active component is a specific yet ubiquitously present eukaryotic protein or protein complex. Our experience with the purification and identification of the host factor suggests that the factor could be part of a dynamic complex involved in the regulation of transcription. Purification of such a multi-subunit complex to homogeneity would prove difficult through conventional chromatography. Eukaryotic transcription is regulated by several such multi-subunit complexes which could be involved in regulation of viral transcription. Vaccinia post-replicative ternary complex could, like its eukaryotic counterpart, be present as a holoenzyme consisting of the viral core subunits and additional host subunits. The host factor could be a component of the viral holoenzyme that gets detached \textit{in vitro} and requires to be added back in to make the polymerase termination efficient. The presence of the protein in prokaryotic extracts has been difficult to determine due to the presence of nucleases.

**Properties of A18 Mediated Termination of STC**

Termination assays involving STC have demonstrated that while both A18 and the HF are required to terminate the stalled ternary complex, neither factor is required to be part of the elongating polymerase in order to mediate termination. Vaccinia stalled intermediate ternary complexes are most stable in potassium salt. Termination of these complexes is insensitive to the absence of salt, but has definite salt optima at 100 mM salt, conditions which have been shown to be inhibitory to the termination of vaccinia early ternary complexes. Inhibition of termination at higher salts where the STC is still stable hints at the disruption of protein-nucleic acid or protein-protein interactions that may be essential for A18 mediated termination. The termination assay has a divalent metal dependence with optima at 5 mM, higher concentrations than which are inhibitory to termination. The energy requirement profile of A18 in the termination assay mimics
the energy requirement profile of the purified enzyme in an in vitro ATPase assay, capable of hydrolyzing only ATP and dATP. These data suggest that the enzyme utilizes its ATPase and translocase potential to bind the template DNA and translocate along the template to function as remodeling the ternary complex. ATPase titrations of the termination assay provide an approximate Km for the enzyme at substrate concentrations of 200 μM ATP, which is a more physiologically relevant value when compared to 2.4 mM derived for the enzyme in an in vitro ATPase assay. This indicates that the function of the enzyme in the context of the ternary complex is energetically more favorable.

**A18 Mediated Termination and Transcriptional Pausing**

We investigated the relationship between A18 mediated termination and transcriptional pausing by introducing A18 to actively elongating ternary complexes, capable of pausing at an artificial pause site introduced into the template. In the presence of the HF, A18 terminates the paused complexes. A18 revealed the presence of natural or intrinsic pause sites on the template by terminating the ternary complexes at these sites, under limiting nucleotide conditions. At saturating concentrations of nucleotides, when the ternary complex has an increased rate of elongation and a decreased propensity to pause, we observe a reduction in the termination of intrinsically paused complexes. This proves that pausing is a key requirement for A18-mediated termination. In the context of the actively elongating ternary complex capable of pausing, the presence of A18 by itself negatively affects the transcription potential of the elongating complex as evidenced by a decrease in the amount of full length transcripts and the increase in the amount of paused complexes with increasing concentrations of A18 in bound fractions. Moreover, A18 alone seems to terminate a subset of the above
paused complexes independent of the host factor. This is direct contrast to the STC where termination is solely dependent on the presence of both A18 and HF. The vaccinia genome is AT rich with several T-rich regions that could serve as cryptic pause sites from where elongating polymerases can by terminated by A18 leading to an overall heterogeneity of the 3’ ends \textit{in vivo}.

\textbf{Model for A18 Mediated Termination}

A comparison of the different mechanisms of termination in both prokaryotic and eukaryotic systems shows two common features that are associated with transcription termination. In factor dependent termination mechanisms the factors are either RNA or DNA helicases that have translocase activity. The other common feature involves the state of the RNA polymerase. Pausing or stalling of elongating RNA polymerases is a key intermediate between active elongation and complete termination. In bacteria, termination is mediated by rho, a hexameric helicase, capable of dissociating paused elongation complexes utilizing an RNA-dependent ATPase activity to translocate along the nascent mRNA transcript. The other factor dependent mechanism involves mfd, a DNA helicase that translocates along the DNA template, capable of terminating a transcription complex stalled at an area containing a DNA lesion in an energy dependent manner. In addition to these factor-mediated mechanisms of termination, there also exists a factor independent mechanism, intrinsic termination, where a stable RNA hairpin loop structure formed when the polymerase pauses, further destabilizes and terminates the complex. In eukaryotic system, although the mechanism for termination is not clearly understood, the importance of pausing in termination and the helicase nature of the transcription factors involved in termination has been well documented. The following paragraphs contain a description of these common features.
Pausing is a key regulatory mechanism employed by elongating RNA polymerases. Transcriptional complexes move in an asynchronous or discontinuous manner along a DNA template. At some sequence specific regions of the template the polymerase is capable of transient pausing. These pauses are not considered a normal state of the elongating polymerase but an “off-pathway” state of the polymerase, because for any given pause sequence in the DNA not all polymerases undergo pausing; a subset of the polymerase population pauses at the site while the rest of them continue their normal elongation. Bacterial RNA polymerases have been shown *in vitro* to be capable of pausing ubiquitously once every 100-200 base pairs of the DNA template for as short as 1-6 seconds. Ubiquitous pauses can be further stabilized or prolonged by additional mechanisms that include i) backtracking of the RNA polymerase, leading to a displacement and extrusion of the 3’ end of the RNA transcript from the active site within the ternary complex, ii) protein interactions with the paused polymerase, especially in the ssDNA region of the non-template strand, iii) stable RNA hairpin loop structures, iv) nature of the downstream DNA template (Landick, 2006). In their role as regulatory components of transcription, paused polymerases can either be acted upon by positive elongation factors and resume productive elongation or backtrack excessively or decay into arrested complexes that are associated with the template DNA but are incapable of elongation on their own or be terminated by appropriate termination factors.

Helicases play a major role in many cellular processes including transcription. Historically helicases were defined as nucleic acid dependent ATPases that couple energy from NTP hydrolysis to unwinding duplex nucleic acid structures. Based on the
sequence similarities among these proteins, common motifs were identified and used to classify new members of the family. However not all helicases are functionally capable of unwinding double stranded nucleic acid. Currently, all helicases are considered motor proteins capable of translocating along nucleic acids in an energy dependent manner. A subset of these proteins is capable of unwinding duplex DNA or RNA molecules. Another function associated with translocases/helicases is dismantling or remodeling nucleoprotein complexes, such as nucleosomes, splicesosomes, transcription complexes (Singleton et al., 2007).

A18 shares these features that are a hallmark of different mechanisms of termination. Previous studies have identified A18 as a DNA helicase capable in vitro of only a weak helicase activity activated by a ssDNA, dsDNA or DNA:RNA dependent ATPase activity. We postulate that A18, by virtue of its helicase motifs, functions to translocate along nucleic acid and disrupt or dissociate the ternary complex. In this dissertation we have shown the relevance of pausing during A18 mediated termination. The pause sites revealed during elongation in the presence of A18 indicates that A18 possibly stabilizes the RNAP at certain pause sites and ultimately causes its termination.

Our overall hypothesis based on these data is that A18 binds and stabilizes ubiquitously paused polymerases and utilizes ATPase derived energy to dissociate the polymerase. Since the ATPase activity is required for termination and is activated by either ssDNA, dsDNA or DNA:RNA hybrids, in the context of a paused polymerase there are three possible regions for A18 to interact with the ternary complex, A1, A2 or A3 (Fig. 5-1). In binding to position A1, which is the downstream DNA template, A18
could act as a roadblock protein to stabilize the paused polymerase and by virtue of its translocase function exert an opposing force on the ternary complex and mediate its disruption or dissociation. In binding to position A2, which is the exposed ssDNA region of the non-template strand of the ternary complex, A18 could stabilize the ubiquitously paused polymerase, one of the mechanisms suggested to stabilize pauses (Landick, 2006). Termination by this mechanism could be possible in a manner similar to early vaccinia polymerases, where NPHI binds one of the polymerase subunits and couples ssDNA dependent ATPase energy to mediate termination (Christen et al., 2008; Piacente et al., 2008). This would involve in addition to A18’s 3'-5' translocase activity and ssDNA dependent ATPase activity, a possible interaction of A18 with the vaccinia RNA polymerase, no such interaction has been identified to date. The third possibility involves the binding of A18 to position A3, which is the upstream DNA sequence. Termination in this scenario could be mediated by A18’s binding directly to the RNA polymerase and stabilizing the pause. DNA and polymerase bound A18 could dissociate the ternary complex by translocating in the direction against the direction of normal polymerase movement. This mechanism would share similarities with mfd mediated termination, in that mfd also interacts with both the RNA polymerase and the upstream DNA. Mfd terminates only stalled ternary complexes incapable of elongation by inducing a forward translocation of the polymerase and ultimately dissociating the polymerase. However, in our scenario, A18 can mediate termination of elongation competent polymerases by stabilizing their pausing. This would indicate that A18 upon binding the upstream DNA and RNA polymerase would have to dissociate the complex by translocating away from the polymerase rather than towards it as in the case of mfd.
There are two possible roles for the host factor during termination. In the first scenario, the host protein or protein complex would interact with A18 and enhance or activate its termination functions, which would be 1) the ability to bind DNA and catalyze ATP hydrolysis, 2) ATPase catalyzed translocase activity, 3) interaction with the RNA polymerase. Our data with paused polymerases suggest that A18 can mediate termination of a subset of paused complexes independent of the host factor, indicating that A18 does not require the host factor to activate its termination functions. A second possible role for the host factor involves interaction and subsequent enhancement of the ternary complex by the host protein or protein complex making the ternary complex sensitive to A18 mediated termination. The RNA polymerase ternary complex can exist in multiple conformational states that are dictated by the nature of its current interactions. In our studies we examined termination from two sets of ternary complexes; a stalled ternary complex which was completely resistant to A18 mediated termination in the absence of the host factor and an actively elongating ternary complex which was partially sensitive to A18 mediated termination in the absence of the host factor. These data indicate that the active state of the polymerase plays a distinguishing role between the two scenarios. The possible role of the host factor could be inducing a termination sensitive state of the ternary complex. Moreover, under normal circumstances during viral replication, the host protein could be a RNA polymerase associated factor that was separated from the complex during \textit{in vitro} manipulations.

\textbf{Role of A18 in Viral Post-replicative Transcription Regulation}

Post-replicative transcription termination in vaccinia virus lacks a distinct cis-acting termination signal and results in the formation of transcripts of variable lengths with ragged 3’ ends. In A18 defective mutants, infections result in normal early transcripts.
but much longer than normal intermediate and late transcripts proving its importance in transcription termination. Our model for A18 mediated termination can be applied to the viral infection scenario where ubiquitous pauses of the elongating RNA polymerase are stabilized by the presence of A18 leading to the dissociation of the complex. The AT-rich (60 %) vaccinia genome serves as an ideal template to induce ubiquitous pausing of an elongating RNA polymerase. Termination of these polymerases would result in transcripts with heterogeneous 3’ ends. Owing to the cytoplasmic site of gene expression the vaccinia post-replicative ternary complex is possibly regulated by both viral and host factors. With the exception of the host factor requirement during A18 mediated termination in vitro, no other host requirement during elongation and termination has been described. The role of viral transcription regulators like G2, J3 – positive elongation factors and a multifunctional factor, H5 were identified during genetic characterization of temperature sensitive mutants. Evidence from in vitro interaction studies have indicated at a possible interaction of A18 with these factors. Although A18 does not need to interact with an elongating polymerase in order to terminate it, in the context of the viral infection, the post-replicative elongation complex is probably regulated by both positive regulators like J3 and G2 and negative regulators like A18 and the host factor, which bias the ternary complex either into productive elongation or termination.

In addition, in some post-replicative scenarios that are exceptions to the rule, H5 has been shown to mediate cleavage of transcripts resulting in homogeneous transcripts, indicating at a possibly more complex process of transcription regulation, similar to those seen in eukaryotic systems. Transcription termination in eukaryotic
system involves transcription past the poly A signal which leads to the recruitment of 3’ end processing factors. Among the 3’ end processing factors are the multisubunit cleavage and polyadenylation specificity factors (CPSF) and cleavage stimulation factors (CstF) that cleave the nascent mRNA to generate two RNA molecules, a capped mRNA that is subsequently polyadenylated and a polymerase associated RNA with an uncapped 5’end. This uncapped molecule serves as a substrate for cellular exonucleases that function in tandem with other factors to mediate termination of the still elongating polymerase. The role of the vaccinia cleavage factor, H5, in the formation of a few homogeneous post-replicative transcripts implies that the process of transcription termination in vaccinia post-replicative genes might involve a mechanism similar to eukaryotic co-transcriptional cleavage, where nascent mRNA transcripts are generated by the cleavage of RNA at cryptic cleavage sites, leaving behind a functionally elongating polymerase with an uncapped 5’ end that is terminated by A18.

**Future Directions**

A18 was identified during phenotypic studies of temperature sensitive mutants. A role for the purified protein was described by a series of biochemical experiments, ultimately leading to investigations of its properties in the context of a transcription complex. A detailed understanding of the protein’s mechanistic role during termination and macroscopic function in viral transcription regulation can be achieved through multipronged studies involving the protein.

The properties of A18 mediated termination of a stalled ternary complex were investigated in this dissertation. In the context of actively elongating polymerases under limiting nucleotide conditions, we have shown that A18 is capable of i) negatively affecting elongation, 2) stabilizing pauses and 3) terminating a subset of the
polymerases on its own. It will be interesting to analyze the properties of these three functions, with respect to their salt sensitivity and hydrolysable energy requirement to identify conditions to distinguish the three functions.

The ATPase activity associated with the protein was shown to be activated in the presence of ssDNA, dsDNA and DNA:RNA hybrids. The levels of activation of A18's ATPase function in an *in vitro* assay were seen to vary depending on the nature of the oligomer used. In addition A18 has ATP independent DNA binding property. It would prove useful to investigate the nature of A18's binding to nucleic acids in order to determine if the enzyme has a sequence preference or specificity. Eukaryotic pol I termination mechanism involves a two factor mechanism, where one of the factors binds sequence specifically to downstream DNA regions to serve as a roadblock to the polymerase complex. The nature of A18’s nucleic acid binding can be investigated either by performing filter binding assays or non-denaturing gel analysis of A18 bound labeled nucleic acids.

A18 is classified as belonging to the superfamily II helicase family of proteins and has a weak DNA helicase activity *in vitro*. Our hypothesis suggests that A18 functions as a translocase protein that ultimately mediates nucleoprotein remodeling or disruption. The translocase function of A18 can be tested by performing the triple helix forming oligonucleotide (TFO) assay (Whitehouse et al., 2003; Smith et al., 2007). Pyrimidine-rich TFOs functions as a sequence specific dsDNA binding ligand. TFOs recognize the complementary purine containing dsDNA and bind the major groove of dsDNA utilizing Hoogsten hydrogen bonds to form T-AT and C-GC triplets. These triple helices form only in low pH conditions and not under neutral pH. By incubating a radiolabeled TFO
with the target DNA at low pH triplexes are formed. The TFO can be dissociated from
the dsDNA in the presence of a translocase. Performing translocase assays under pH
conditions where re-binding of the TFO to the dsDNA is inhibited, we can identify the
ability of A18 to function as a translocase.

The role of A18 in the context of the viral transcription regulators is an important
area of investigation. This would involve transcription in the presence of purified
vaccinia polymerase. While purified vaccinia polymerase has been used to study early
vaccinia termination, post-replicative termination studies have been carried using
extract generated ternary complexes. In post-replicative vaccinia transcription, purified
polymerases have been re-constituted in vitro to study transcription initiation.
Characterization of this system to study elongation and termination would make it
possible to perform mechanistic studies involving transcription regulation. Using such a
re-constituted system, not only can we investigate the effect of individual transcription
factors but also the effect these factors have on each other and in combination on the
ternary complex.

A reconstituted post-replicative transcription system using purified polymerases
and the minimal required factors to initiate transcription in vitro can also be used to
investigate the structural details and the stability of the transcription complex, as has
been described for early vaccinia polymerases (Hagler & Shuman, 1992a; Hagler &
Shuman, 1992b; Hagler & Shuman, 1992c). Structural studies by DNA footprinting
analyses will help define parameters of the post-replicative ternary complexes such as
the length of the ternary complex, distance between the active site and the edges of the
ternary complex and length of the DNA-RNA hybrid. Such a well-defined ternary
complex will be suited to investigate the mechanistic details of A18 mediated
termination. Using this system we will be able to define the region of A18’s interaction
with the ternary complex. A mutant version of the EcoR1 enzyme (Gln111) that can bind
the recognition site but is catalytically inactive has been used in studies with mfd to
define the template region required for mfd binding during termination (Park et al.,
2002). This involves using templates where the elongating polymerase is stalled at
specific positions either downstream or upstream from the EcoR1 binding sites. Mutant
EcoR1 is allowed to bind templates with the stalled polymerases prior to A18 mediated
termination. Termination will be abrogated in those constructs where the EcoR1 mutant
is bound to A18 binding regions, thus identifying A18’s site of interaction with the ternary
complex.

This experimental setup will also help address a fundamental mechanistic aspect
of A18 mediated termination. Mfd and rho mediated transcription termination has been
shown to occur by a mechanism of forward translocation where the RNA polymerase
terminates when it is forced into forward translocation in the absence of RNA synthesis
(Park et al., 2002; Park & Roberts, 2006). A similar mechanism was proposed for
intrinsic terminators as well (Yarnell & Roberts, 1999). In recent years an alternate
allosteric model which does not require forward translocation to mediate termination has
been demonstrated for intrinsic terminators as well as rho dependent terminators
(Epshtein et al., 2007; Epshtein et al, 2010). Given the basic structural conservation
among the different RNA polymerases, A18 mediated termination would also be
mediated by either of these two processes. Evidence from the Gollnick lab suggests
that vaccinia early polymerase termination might involve forward translocation. Of the
three models for A18 mediated termination, binding of A18 to positions A1 and A3 (Fig. 5-1) and subsequent dissociation of the ternary would inhibit forward translocation. The dependence of A18 mediated termination on forward translocation can be determined by comparing A18’s ability to terminate under two different reaction scenarios. In one case the ternary complex would be stalled at a region directly upstream from the mutant EcoR1 protein allowing no space for the forward translocation of the polymerase. In another setup the ternary complex would be stalled a few bases upstream from the mutant EcoR1, allowing for forward translocation of the polymerase. Preferential termination of one of these complexes over the other would help us gain an insight into the mechanistic aspects of A18 mediated termination.

A different experimental approach to investigate the properties and function of A18 involves deletion analysis of the purified protein to test its \textit{in vitro} functions. Most superfamily II helicase proteins have, in addition to the helicase motifs, accessory domains that confers specificity for each protein. These accessory domains can posses catalytic function or be involved in protein interaction. A18 is a 56 KDa protein belonging to the superfamily II helicases consisting of 493 amino acids. As seen in figure 5-2, the helicase motifs occupy the central regions of the protein, leaving behind uncharacterized 100 amino acid regions on both the amino and carboxyl terminal ends which could function as accessory domains. We possess a number of assays to test for different aspects of the enzyme’s function; ATPase assays, nucleic acid binding assay, stabilization of paused complexes, host factor dependent and independent termination. Deletion constructs of the protein can be expressed, and tested in these assays to identify if these regions are required for normal ATPase function or additional nucleic
acid binding, possible host factor interaction. These deletion mutants can also be tested for their ability to bind other viral transcription factors.

In conclusion, the work described in this dissertation has refined our understanding of the vaccinia termination factor; in that it has identified pausing to be a key mechanistic requirement for A18 mediated termination. It has also described a possible role for the protein in stabilizing ubiquitously paused polymerases. We have been able to show for the first time host factor independent termination by A18, leading us to a possible hypothesis that the active conformation state of the polymerase is a requirement for A18 mediated termination. Moreover, we have shown that the host factor is a cellular protein or protein complex and suggest that it functions to modulate the ternary complex to make it termination sensitive. Further investigation of the mechanistic aspects of vaccinia post-replicative transcription termination and A18’s role in terminating the ternary complex would further the understanding of the process of eukaryotic termination.
Figure 5-1. Model for A18 mediated termination. An ubiquitously paused ternary elongation complex, represented here by the RNA polymerase in grey, DNA template in blue and RNA transcript in orange, are stably paused and ultimately terminated by the action of A18, shown here as the purple circles. A18 has three possible binding sites. A1, downstream DNA template region. A2, ssDNA in the ternary bubble. A3, upstream DNA template region. Grey arrow indicates direction of RNA polymerase transcription. Green arrows indicate direction of A18 translocation.
Figure 5-2. Schematic representation of the A18 protein. The protein A18 is depicted as the grey bar containing in purple the different helicase signature motifs that classify it as a superfamily 2 helicase. (N-amino terminus; C-carboxyl terminus)
LIST OF REFERENCES


Gershon, P.D., Moss, B., 1992. Transition from rapid processive to slow nonprocessive polyadenylation by vaccinia virus poly(A) polymerase catalytic subunit is regulated by the net length of the poly(A) tail. Genes Dev. 6, 1575-1586.


Katsafanas, G.C., Moss, B., 2004. Vaccinia virus intermediate stage transcription is complemented by Ras-GTPase-activating protein SH3 domain-binding protein (G3BP) and cytoplasmic activation/proliferation-associated protein (p137) individually or as a heterodimer. J. Biol. Chem. 279, 52210-52217.


Rosales, R., Harris, N., Ahn, B.Y., Moss, B., 1994. Purification and identification of a vaccinia virus-encoded intermediate stage promoter-specific transcription factor that has homology to eukaryotic transcription factor SII (TFIIS) and an additional role as a viral RNA polymerase subunit. J. Biol. Chem. 269, 14260-14267.


BIOGRAPHICAL SKETCH

Aparna Manoharan was born in Madras, India in the year 1982. She enrolled for undergraduate studies in 1999 at the Birla Institute of Technology and Science (BITS), Pilani, India where she graduated in 2003 with a master’s (Hons.) in biological sciences. She was admitted into the Interdisciplinary Program in Biomedical Sciences at the University of Florida in 2003. She joined the Condit lab in the summer of 2004 where she started work on her dissertation. In the summer of 2010 she receives her Ph.D. in biomedical sciences.