ENZYMATIC DISASSEMBLY OF PROMOTER BOUND 7SK SNRNP
DRIVES TRANSCRIPTION ELONGATION OF HIV AND CELLULAR GENES

APPROVED BY SUPERVISORY COMMITTEE

Ivan D'Orso, Ph.D., Supervising Professor

Melanie Cobb, Ph.D.

Nicholas Conrad, Ph.D.

Jeffrey Kahn, M.D., Ph.D.

ENZYMATIC DISASSEMBLY OF PROMOTER BOUND 7SK SNRNP DRIVES TRANSCRIPTION ELONGATION OF HIV AND CELLULAR GENES

by

Ryan Philip McNamara

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center

Dallas, Texas

December, 2015

ENZYMATIC DISASSEMBLY OF PROMOTER BOUND 7SK SNRNP
DRIVES TRANSCRIPTION ELONGATION OF HIV AND CELLULAR GENES

RYAN PHILIP MCNAMARA, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2015

IVÁN D'ORSO, Ph.D.

Gene expression of the human immunodeficiency virus (HIV) and cellular primary responsive genes (PRG's) is regulated at the step of transcription elongation. Shortly after transcription initiation, RNA Polymerase II (Pol II) pauses and it only enters into productive elongation after inducible transcription factors (TF's) recruit the P-TEFb kinase to phosphorylate Pol II in response to stimuli. To ensure tight regulation of this process, the majority of P-TEFb is held in a catalytically inactive form, reversibly bound to the 7SK small nuclear ribonucleoprotein (snRNP). In the absence of stimuli, the 7SK snRNP resides in both the nucleoplasm and promoter regions. However, an understanding of how TF's capture P-TEFb from the 7SK snRNP at the promoter and the mechanism and purpose of localizing the 7SK snRNP to promoters has been largely unexplored. It was therefore my goal to biochemically and

functionally characterize this pathway through the use of both the HIV encoded TF Tat and cellular TF's such as nuclear factor kappa b (NF-κB). Detailed throughout this dissertation, I present the novel findings that HIV Tat and NF-κB function to recruit the PPM1G phosphatase to their targeted promoters, which dephosphorylates P-TEFb and triggers its release from the 7SK snRNP. Additionally, this extraction of P-TEFb from the 7SK snRNP occurs at promoters through the transcriptional regulator KAP1, which physically tethers the snRNP to promoters genome wide. Recruitment of the 7SK snRNP complex occurs after transcription initiation, allowing P-TEFb to be directly positioned for rapid extraction by TF's upon stimuli and transferred onto the paused Pol II. The enzymatic uncoupling of P-TEFb from the promoter bound 7SK snRNP enables rapid Pol II elongation and gene expression in response to stimuli (for PRG's) or in the presence of Tat (for HIV). Ultimately, these findings indicate that inducible transcription programs can rapidly respond to environmental cues through the localized positioning of elongation factors at promoters. Moreover, these findings illustrate that HIV has evolved to hijack a cellular gene expression program, thus leading to viral takeover of the host and progression of AIDS.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

***McNamara, R.P.,** *McCann, J.L., Gudipaty, S.A., and D'Orso, I. (2013). Transcription factors mediate enzymatic disassembly of promoter bound 7SK snRNP to locally recruit P-TEFb for transcription elongation. *Cell Reports 5*, 1256-1268.

*Gudipaty, S.A., ***McNamara, R.P**., Morton, E.L., and D'Orso, I. (2015). PPM1G binds 7SK RNA and Hexim1 to block P-TEFb assembly into the 7SK snRNP and sustain transcription elongation. *Molecular and Cellular Biology* August 31 (ahead of print at time of publishing this dissertation)

* Denotes equal authorship

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

a.u. – arbitrary units
AD – activation domain
AP – affinity purification
Arg - arginine
Asp – aspartate
BRD4 – bromodomain containing protein 4
CA – capsid
CDK – cyclin dependent kinase
ChIP – chromatin immunoprecipitation
ChIP-seq – chromatin immunoprecipitation sequencing
CMV - cytomegalovirus
CTD – C-terminal domain
Cyc – cyclin
DNA – deoxyribonucleic acid
ELISA – enzyme linked immunosorbance assay
Env – envelope
Gag – group specific antigen
FFL – firefly luciferase
GFP – green fluorescent protein
GRO-seq – global run-on sequencing
GST – glutathione S-transferase
H3K27ac – histone 3 lysine 27 acetylated
H3K4me1 – histone 3 lysine 4 monomethylated
H3K4me3 – histone 3 lysine 4 trimethylated
H3K9me3 – histone 3 lysine 9 trimethylated
HA – haemagluttanin
HAART – highly active anti-retroviral therapy
HEK – human embryonic kidney
HeLa – Henrietta Lacks
Hexim1 - hexamethylene bis-acetamide inducible 1
HIV – human immunodeficiency virus
hnRNP – heterogenous nuclear ribonucleoprotein
HSF1 – heat shock factor 1
IL - interleukin
IP – immunoprecipitation
KAP1 – krab-associated protein 1
Larp7 – la-related protein 7
LTR – long terminal repeat
MA - matrix

MePCE – methylphosphate capping enzyme
NF-κB- nuclear factor κB
NL4.3 – HIV NL4.3
NLS – nuclear localization sequence
NPM1 - nucleophosmin
P-Cdk9 – phosphorylated Cdk9 at Threonine 186
P-TEFb – positive transcription elongation factor b
PCR – polymerase chain reaction
Pol II – RNA Polymerase II
PPM1A – protein phosphatase metal dependent enzyme 1A
PPM1G – protein phosphatase metal dependent enzyme 1G
Pro – Proline
qPCR – quantitative PCR
qRT-PCR – quantitative reverse transcriptase PCR
RBD – RNA binding domain
RNA – ribonucleic acid
RNAi – RNA interference
RNP - ribonucleoprotein
S2P-CTD – phosphorylated serine 2 of the Pol II c-terminal domain
S5P-CTD – phosphorylated serine 5 of the Pol II c-terminal domain
SEM – standard error of the mean
Ser – Serine
shRNA – short hairpin RNA
snRNP – small nuclear ribonucleoprotein
SV40 T Ag – simian virus 40 T antigen
TAR – transactivating response element
TF – transcription factor
TFIIH – transcription factor II H
Thr – Threonine
TIF – transcription intermediary factor
TNF/TNF-α – tumor necrosis factor α
Trim28 – tripartite-containing motif protein 28
TSS – transcriptional start site
Tyr – Tyrosine
UTR – untranslated region
WT – wild type

CHAPTER I

PART I

LITERATURE REVIEW

Transcription Elongation by RNA Polymerase II

At the majority of metazoan genes, RNA Polymerase II (Pol II) pauses shortly after transcription of ~20-80 nucleotides downstream of the transcriptional start site (TSS) at promoter proximal regions (Adelman and Lis, 2012; Jonkers and Lis, 2015; Muse et al., 2007) (see **Figure 1** below). Upon recruitment of elongation factors, Pol II clears the promoter and enters into the gene body, promoting gene activation. The C-terminal domain (CTD) of Pol II plays a pivotal role in its transition from initiation into productive elongation (Czudnochowski et al., 2012; Guo and Price, 2013; Jonkers and Lis, 2015; Wada et al., 1998; Yik et al., 2003). The CTD of Pol II contains a 52 amino acid stretch of the heptad repeat Tyr-Ser-Pro-Thr-Ser-Pro-Ser, although not every heptad is completely repetitive (McCracken et al., 1997; Zhang and Corden, 1991). Phosphorylation of the Serine residues in the heptad repeats governs Pol II's pause release from the promoter, and this phosphorylation is carried out by several cyclin dependent kinases (CDKs) and ERK 1 and ERK2 (Fisher, 2005; Glover-Cutter et al., 2009; Guo and Price, 2013; Peterlin and Price, 2006; Tee et al., 2014; Zhou et al., 2012).

Phosphorylation of Serine 5, accomplished by the transcription factor II H (TFIIH) subunit CDK7, is the first phosphorylation event of the Pol II CTD (Fisher, 2005; Glover-Cutter et al., 2009; Malumbres, 2014). This initial phosphorylation of Serine 5 is accomplished through the promoter assembled CDK7, although it unknown how many Serine 5's in the CTD repeat

are phosphorylated at this point (Fisher, 2005; Glover-Cutter et al., 2009). At this point, Pol II pauses at the promoter proximal region and awaits further phosphorylation to enter into productive elongation.

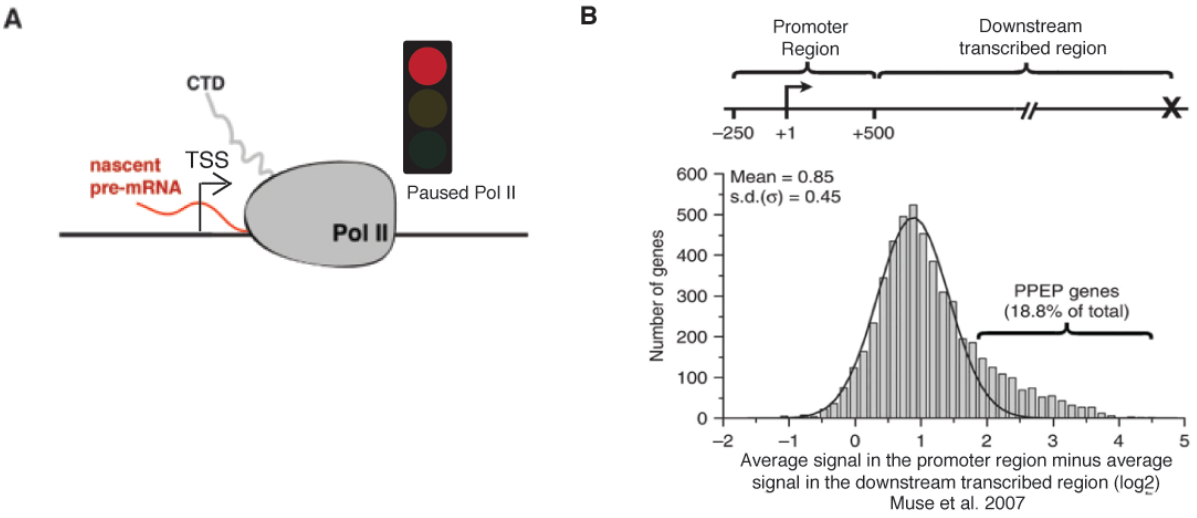

Figure 1. RNA Polymerase II (Pol II) is poised at the promoters of human genes throughout the genome.

(A) Schematic of Pol II pausing after transcription initiation. After transcription of ~20-65 bp, Pol II remains poised at the promoter proximal region of a gene with its C-terminal domain (CTD) in a hypo-phosphorylated form.

(B) Gaussian distribution of Pol II occupancy throughout the human genome as determined by Muse et al. 2007. The majority of Pol II is localized at promoter proximal regions, whereas only a fraction of Pol II is localized in the gene body at steady state.

The transition into productive elongation is controlled by hyper-phosphorylation of Serine 5 and Serine 2 of the Pol II CTD (Adelman and Lis, 2012; Jonkers and Lis, 2015; Lorincz and Schubeler, 2007; Peterlin and Price, 2006; Zhou et al., 2012). The P-TEFb kinase, which is

composed of CDK9 and its regulatory Cyclin (Cyc) subunit, CycT1/T2, is the kinase primarily responsible for this hyper-phosphorylation (Price, 2000; Zhou et al., 2012). Unlike CDK7 of TFIIH, P-TEFb has been proposed to be inducibly recruited by gene specific transcription factors (TFs), which couple with the kinase at their targeted promoters. P-TEFb can then hyper-phosphorylate the Pol II CTD at Serine 5 and Serine 2, resulting in TF mediated gene activation (Zhou et al., 2012). Of note, another kinase, CDK12, has recently been shown to utilize Serine 2 of the CTD as a substrate; however distinguishing its role from that of P-TEFb in transcription elongation is still unclear (Basent et al., 2014; Boesken et al., 2014; Bowman and Kelly, 2014; Blazek et al., 2011; Czudnochowski et al., 2012; Napolitano et al., 2014). The wide array of post-translational modifications to the CTD of Pol II, particularly phosphorylation, likely plays a role in co-transcriptional RNA processing and recruitment of various splicing factors, although there remains much to be understood (Bentley, 2005).

The P-TEFb Kinase

Since its discovery, the P-TEFb kinase has been extensively studied for its regulation of transcription elongation (Marshall and Price, 1995; Peterlin and Price, 2006). P-TEFb has been shown to be utilized by a number of TFs to stimulate gene expression such as p53 (Gomes et al., 2006), nuclear factor κB (NF-κB) (Barboric et al., 2001a), heat shock factor 1 (HSF1) (Lis et al., 2000), bromodomain containing protein 4 (BRD4) (Liu et al., 2013), c-Myc (Lin et al., 2012; Price, 2000), and human immunodeficiency virus (HIV) Tat (Mancebo et al., 1997; Sobhian et al., 2010), just to name a few. The current accepted model is that TF's such as those listed above

act as recruiting scaffolds for P-TEFb, bringing the kinase to their targeted promoters to activate transcriptional programs (see **Figure 2** below).

Like other CDK's, the phosphorylation of the T-loop of CDK9, which contains a specific Thr (at position 186) embedded in the unstructured T-loop, primes the kinase for activation (Fisher, 2005; Malumbres, 2014). In contrast to many of these CDK's, CDK9 has the ability to auto-phosphorylate is its own T-loop, thus bypassing the requirement of another kinase to activate P-TEFb (Baumli et al., 2008). Furthermore, while it had been previously demonstrated that Tat alters the kinase activity/specificity of P-TEFb, the recently solved co-crystal structure of Tat with P-TEFb revealed that Tat directly contacts the T-loop when complexed with the kinase, and potentiates its own auto-phosphorylation rate (Baumli et al., 2008; Chen et al., 2004; Li et al., 2005; McNamara et al., 2013). This auto-phosphorylation of P-TEFb is in contrast to how many other CDK's T-loop's are activated, which is through CDK7, the primary CDK activating kinase (termed the CAK) (Fisher et al. 2005). Although CDK9 can auto-phosphorylate, it has been demonstrated that CDK7 can still use the T-loop of CDK9 as a substrate (Garber et al., 2000). Therefore the cell has evolved multiple mechanisms to ensure that the P-TEFb kinase is readily phosphorylated, maintaining the kinase in an activated form.

As an additional form of regulation, the catalytic activity of CDK9 is regulated by its cyclin counterpart, CycT1/T2/K (Peterlin and Price, 2006). As with various other cyclins within the cell, CycT1 contains a canonical cyclin box, which mediates P-TEFb's interaction with its inhibitor, Hexim1. This region is also used as the Tat-TAR recognition motif (amino acid residues 250-272 of CycT1), which facilitates the loading of the kinase at the HIV promoter (Das et al., 2004; Yik et al., 2003; Peterlin and Price, 2006). The N-terminal region of CycT1 interacts with CDK9 and TF's such as Tat, thus competing with Hexim1 for binding. The

exposed C-terminal end of CycT1 then serves as a guide to direct the kinase complex to its substrates (Mancebo et al., 1997; Barboric et al., 2000; Barboric et al., 2001). The C-terminus of CycT1 contains a histidine-rich region, which has been shown to directly interact with the CTD of Pol II (Peterlin and Price, 2006; Zhou et al., 2012; Kwak and Lis, 2013; Guo and Price, 2013). Therefore, CycT1 mediates the interaction between CDK9 subunit of the complex into proximity of the CTD of Pol II, resulting in its hyper-phosphorylation.

Unlike CDK7 of the TFIIH complex, the P-TEFb kinase travels with Pol II as it transcribes the gene body as part of the super elongation complex (SEC) (Luo et al., 2012). In addition to P-TEFb, the core subunits of the SEC include AFF4 and ELL (Smith et al., 2011; Luo et al., 2012). The SEC travels with Pol II as it is elongating, continuously phosphorylating the CTD and regulating its speed throughout the gene body (Lin et al., 2011; Luo et al., 2012). This event is likely tied to co-transcriptional RNA processing, but a detailed model of how this occurs has yet to be established.

The 7SK snRNP Complex

To ensure that the kinase activity of P-TEFb, and thereby transcriptional integrity of the cell remains controlled, the majority of P-TEFb is inactivated, reversibly bound to the 7SK snRNP (**Figure 2**). The highly abundant nuclear non-coding 7SK RNA, transcribed by RNA Polymerase III, acts as a scaffold for a number of factors that facilitate the regulated assembly of P-TEFb onto the snRNP (Murphy et al., 1987; Wassarman and Steitz, 1991). The 7SK RNA is constitutively bound at the 3' end by the RNA binding protein La related protein 7 (Larp7), which protects the RNA from nuclear ribonucleases (RNases) (Krueger et al., 2008a; Markert et

al., 2008). The 5′ end of the 7SK RNA contains a γ monomethyl cap at the 5’ triphosphate group that is incorporated by the methylphosphate capping enzyme (MePCE). Similar to Larp7, MePCE remains constitutively bound to the 7SK RNA (Jeronimo et al., 2007; Muniz et al., 2013). Upon capping of the 7SK RNA, MePCE and Larp7 are able to physically bind to each other, forming a closed loop of the 7SK RNA, providing further stability to the snRNP (Xue et al., 2010).

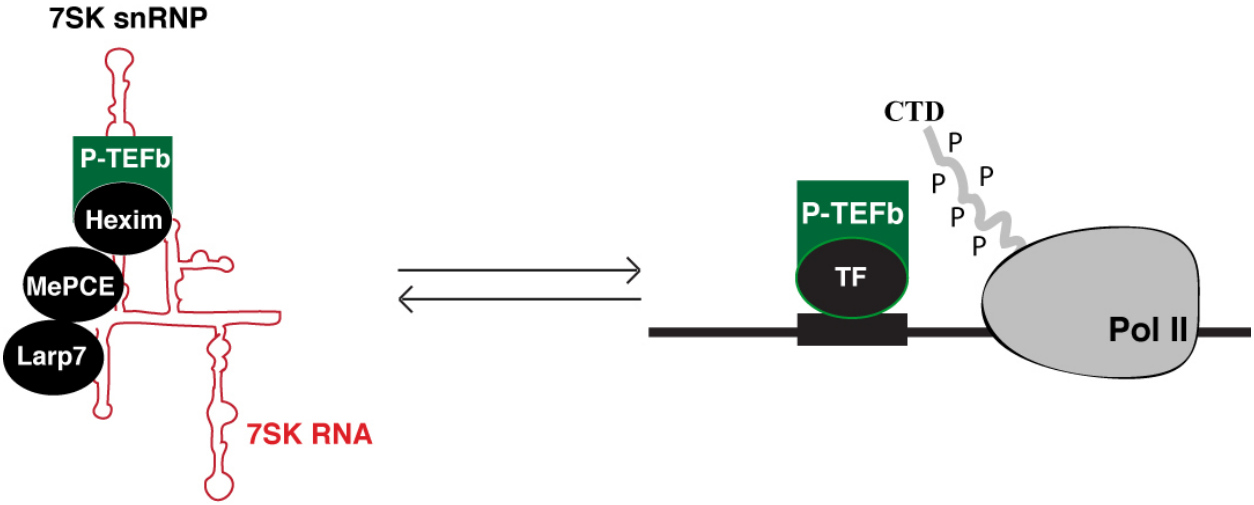

Figure 2. The P-TEFb kinase is extracted from the 7SK snRNP and recruited by TF’s to their targeted promoters. At this point, P-TEFb hyper-phosphorylates the CTD of Pol II, resulting in transcription elongation and gene expression.

A number of reversibly associated components of the 7SK snRNP have been identified such as heterogeneous ribonucleoproteins (hnRNPs) (Barrandon et al., 2007; Hogg and Collins, 2007; Krueger et al., 2008), splicing factors (Barboric et al., 2009; Ji et al., 2013) helicases (Calo et al., 2015), and double stranded (ds) RNA binding proteins (Barrandon et al., 2007b; Krueger et al., 2008a; Marz et al., 2009). One of these dsRNA-binding proteins is termed the hexamethylene bis-acetamide inducible 1 (Hexim1), which physically binds to P-TEFb and

tethers it to the 7SK snRNP by virtue of Hexim1's dsRNA binding of stem loop I of the 7SK RNA (Barrandon et al., 2007; Li et al., 2005; Muniz et al., 2010). This binding event is dependent on the Thr 186 phosphorylation of CDK9, making Hexim1 a physical inhibitor of the primed kinase (Li et al., 2005; Nguyen et al., 2001; Yang et al., 2001; Yik et al., 2003) (**Figure 2**).

The 7SK snRNP exists in nuclear speckles, a region rich in RNA processing activities (Barboric et al., 2009; Barrandon et al., 2007; Diribarne and Bensaude, 2009; Marz et al., 2009). Additionally, the majority of 7SK snRNP can be extracted from nuclear preparations under low salt conditions, signifying that the majority of the complex is found in the nucleoplasm or that the interaction between P-TEFb and core 7SK interactors (Larp7-7SK RNA, MePCE-7SK RNA, Hexim1-7SK RNA, etc.) is of low affinity (Krueger et al., 2008). However ours and other recent reports have demonstrated that a fraction of the 7SK snRNP is reversibly associated with chromatin, both at promoters and enhancers, presumably to directly regulate transcription elongation (Calo et al., 2015; Cherrier et al., 2013; D'Orso and Frankel, 2010a; Eilebrecht et al., 2014; Ji et al., 2013; Liu et al., 2013b; McNamara et al., 2013). This observation was initially demonstrated to occur at the HIV promoter, after synthesis of the TAR RNA (D'Orso and Frankel, 2010a). Subsequent research has shown that placement of the 7SK snRNP at the HIV promoter directly regulates the NF-κB mediated activation of the HIV genome (Morton et al., unpublished data). How this complex influences Tat-mediated transactivation of the HIV genome, though, is still unclear.

Cellular Infection with the Human Immunodeficiency Virus

The human immunodeficiency virus (HIV) belongs to the retroviridae family of viruses, which integrate into the genome of the host cell (Dyda et al., 1994; Ruelas and Greene, 2013). An infectious HIV particle contains two full genomic RNA molecules bound by the HIV nucleocapsid p7 protein, as well as several virally encoded enzymes, encased by the HIV p24 capsid protein (gag). The capsid is then further enveloped by the HIV glycoprotein envelope (env). HIV preferentially infects CD4+T cells containing the CD4 primary receptor and the CXCR4 or CCR5 co-receptors (Klasse, 2012; Sundquist and Kraeusslich, 2012). After viral entry, the virally encoded and packaged reverse transcriptase (RT) converts the viral RNA into dsDNA. Then, the virally encoded and packaged integrase enzyme, along with numerous host cell factors, integrates the dsDNA into the host chromatin (Sundquist and Kraeusslich, 2012). Integration of the HIV DNA genome does not appear to display any flanking sequence specificity of the host DNA, but it is preferentially integrated into transcriptionally active regions of chromatin (Grewe and Uberla, 2010; Jordan et al., 2003; Jordan et al., 2001; Mbonye and Karn, 2011; Schroder et al., 2002)

Upon integration, the virus requires a number of host resources to propagate, including transcriptional machinery (Peterlin and Price, 2006; Zhou et al., 2012; Zhu et al., 1997). The HIV promoter contains many cellular TF binding sites including Sp1, NFAT, and a canonical TATA box (Palmieri et al., 2003). Among these TF binding sites are two well characterized canonical κB binding sites, which makes the virus transcriptionally active upon inflammatory stimuli and/or insults that signal through the secreted chemokine tumor necrosis factor α (TNF)

or through recognition of internal DNA damage (Diamant and Dikstein, 2013; Mbonye and Karn, 2014; Osborn et al., 1989). Upon stimuli, NF-κB binds to the κB elements at the HIV promoter and stimulates transcription elongation (Adelman and Lis, 2012; Amir-Zilberstein et al., 2007; Barboric et al., 2001a; Diamant and Dikstein, 2013).

In the majority of clinical cases, HIV establishes latency through a number of mechanisms (Mbonye and Karn, 2014). Current models propose that the mechanisms that promote viral latency include: occlusion of the viral promoter with restrictive chromatin, absence of the HIV TF Tat, transcriptional silencing of the infected cell, and integration of a replication incompetent virus (Mbonye and Karn, 2014; Pearson et al., 2008; Tyagi et al., 2010). Purging this latent reservoir has become of primary interest to the field, and factors that can mediate HIV's emergence from latency have recently been identified (Jadlowsky et al., 2014; Kim et al., 2006; Mbonye and Karn, 2011, 2014; Pearson et al., 2008; Tyagi et al., 2010). It is proposed that through activation of the latent reservoir of HIV, highly active antiretroviral treatment (HAART) greatly reduces virus transmission to uninfected cells through directly targeting and killing of infecting cells (Mbonye and Karn, 2011, 2014; Ruelas and Greene, 2013; Van Lint et al., 2013).

The HIV Transcription Factor Tat

Like most human genes, Pol II is poised at the HIV promoter and initially transcribes ~ 25-60 nucleotides in the absence of stimulation (Adelman and Lis, 2012; Jadlowsky et al., 2014; Jonkers and Lis, 2015; Muse et al., 2007; Zhang et al., 2007). The nascent RNA contains a highly structured stem loop termed the transactivating response element (TAR). The TAR RNA

serves as a recruiting scaffold for the virally encoded Tat, which binds to TAR along with a number of other cellular co-factors (Baumli et al., 2008; Zhu et al., 1997). One of these co-factors is the P-TEFb kinase, which forms a high affinity complex with Tat on TAR, resulting in hyper-phosphorylation of the CTD of the poised Pol II and activating transcription elongation (Baumli et al., 2008; D'Orso and Frankel, 2010b; Garber et al., 2000; Liu et al., 2002; Muniz et al., 2010; Zhou et al., 2000; Zhu et al., 1997).

Although HIV encodes its own TF Tat, the initial round of transcription of the HIV genome is often owed to cellular TF's such as NF-κB, NFAT, among others (Jordan et al., 2003; Ruelas and Greene, 2013; Van Lint et al., 2013). The Tat gene, along with the Rev and Nef genes, are the first genes to be transcribed, spliced, and exported for translation through the canonical Tap-mediated mRNA export pathway (Frankel and Young, 1998; Guettler and Goerlich, 2011; Van Lint et al., 2013). Upon translation, Tat localizes to the nucleus where it potently activates transcription of the HIV genome, resulting in a positive feedback loop (Karn, 1999; Zhu et al., 1997). Therefore this mechanism of transcriptional activation by Tat allows the virus to quickly transition from relative transcriptional quiescence to viral takeover of the cell.

As with many cellular TF's, HIV Tat requires the P-TEFb kinase for transcriptional activation of the HIV genome (Baumli et al., 2008; D'Orso and Frankel, 2010a; Wei et al., 1998; Zhu et al., 1997). It has been shown that the binding site for the Hexim1-P-TEFb complex on the 7SK RNA is highly homologous to the binding site for the Tat-P-TEFb complex on TAR (Yik et al., 2003). Moreover, the RNA binding domain of Tat bears a striking resemblance to that of Hexim1 (Barboric et al., 2005; Muniz et al., 2010; Sobhian et al., 2010). It was then observed that Tat directly bound the 7SK snRNP, presumably through directly contacting the 7SK RNA. Collectively, this data lead to a model in which HIV Tat can directly extract P-TEFb

from the 7SK snRNP through competitive displacement of Hexim1 (Muniz et al., 2010; Sobhian et al., 2010). At the same time of these publications, it was shown that the 7SK snRNP occupies the HIV promoter, providing evidence that P-TEFb is locally extracted by Tat for activation of the HIV genome (D'Orso and Frankel, 2010a). However competition mediated extraction does not appear to be a physiological mechanism as the steady state levels of Tat produced during infection are low and its predominant localization is, ironically, in the cytoplasm and the inner leaflet of the outer cellular membrane of T cells (Debaisieux et al., 2012; Rayne et al., 2010a; Rayne et al., 2010b).

We recently reported a novel mechanism by which Tat recruits the PPM1G phosphatase to the HIV promoter to dephosphorylate Cdk9 at Thr 186, releasing P-TEFb from the inhibitory 7SK snRNP. This model is supported by other groups' independent findings that phosphatases can release P-TEFb from the 7SK snRNP (Chen et al., 2008; Wang et al., 2008). This dephosphorylated P-TEFb can then undergo a rapid, Tat-facilitated auto-phosphorylation to reactivate the kinase. Upon activation, Tat and P-TEFb then form a high-affinity complex on TAR RNA (Garber et al., 2000), allowing the kinase to hyper-phosphorylate the Pol II CTD and transition it into productive elongation (McNamara et al., 2013) (see **Figure 3** below). The precise mechanism by which Tat reactivates the temporarily dephosphorylated P-TEFb is not clear, but Tat has been shown to directly contact the T-loop of Cdk9 through X-ray crystallography (Baumli et al., 2008). We thus proposed that Tat positions the T-loop into closer proximity to the catalytic cleft of Cdk9, enhancing the rate of the kinase's intrinsic auto-phosphorylation (Garber et al., 2000; McNamara et al. 2013).

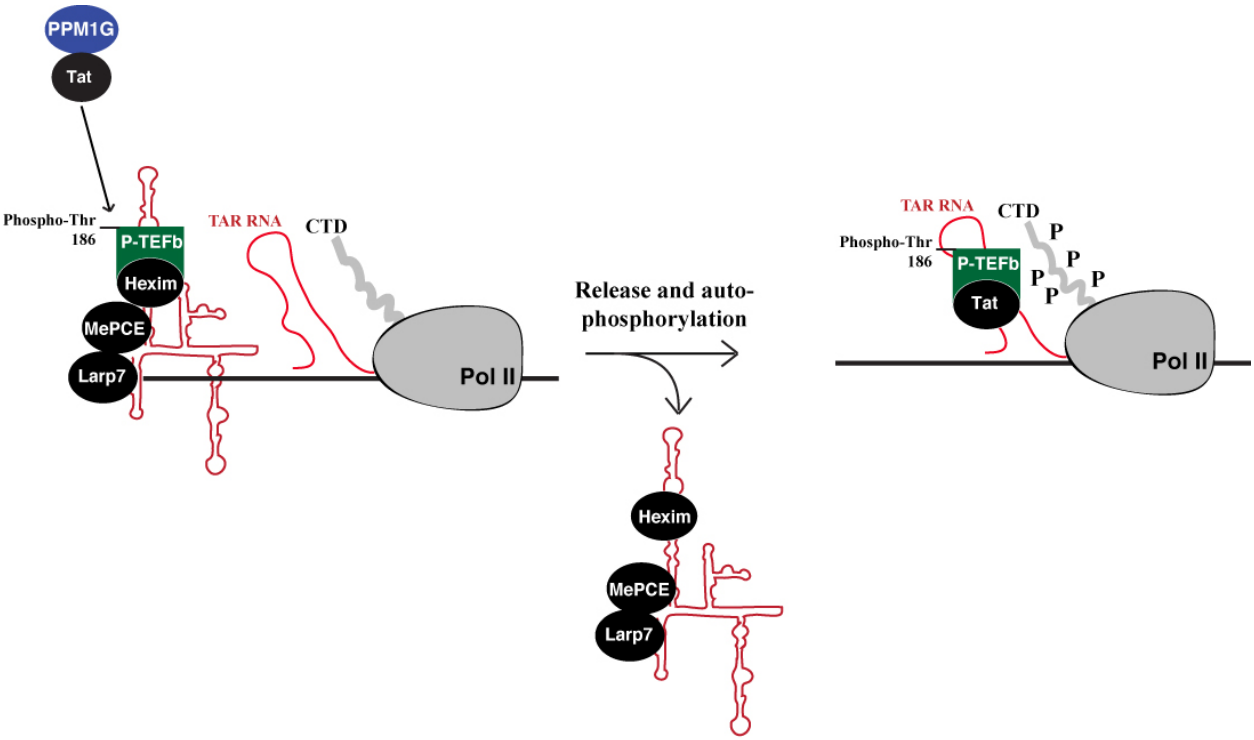

Figure 3. The HIV encoded TF Tat can mediate extraction of P-TEFb from the promoter-assembled 7SK snRNP. Tat recruits the PPM1G phosphatase to the viral promoter, which dephosphorylates and releases P-TEFb from the 7SK snRNP. Upon release, P-TEFb is then rapidly auto-phosphorylated and couples with Tat on the TAR RNA. Here, P-TEFb hyper-phosphorylates the CTD of Pol II at the HIV promoter, leading to transcription elongation.

PART II

RESEARCH INTRODUCTION

The goal of my thesis work was to further characterize how inducible TF's can transition Pol II from its paused state into its actively elongating form at HIV and cellular genes. It is well established that recruitment of the P-TEFb kinase to promoters upon stimuli is a key step in this transition, and TF's such as HIV Tat and NF-κB play a critical role in this recruitment by directly binding the kinase (Barboric et al., 2001; Barboric et al., 2000; Rahl et al., 2010; Zhu et

al., 1997). When not transcriptionally engaged, P-TEFb is sequestered into a catalytically inactive state, reversibly bound to the 7SK snRNP (He et al., 2008; Nguyen et al., 2001). Therefore, TF's must first mediate extraction of P-TEFb from the 7SK snRNP and then recruit the kinase to the targeted promoter. A detailed understanding of how TF's can extract P-TEFb from the 7SK snRNP has been lacking.

The 7SK snRNP serves as the primary reservoir of P-TEFb inside of the nucleus by physically tethering the kinase to the snRNP through the 7SK RNA binding protein Hexim1 (Nguyen et al., 2001; Yang et al., 2001; Yik et al., 2003). When bound to the 7SK snRNP, the CDK9 subunit of P-TEFb is phosphorylated at Thr 186, and thus models of dephosphorylation-mediated release of the kinase have been proposed (Ammosova et al., 2011; Ammosova et al., 2005 Chen et al., 2007; Chen et al., 2008). While several phosphatases have been implicated in dephosphorylating P-TEFb and releasing the kinase from the inhibitory snRNP, no direct TF-mediated phosphatase recruitment has been established.

The HIV TF Tat has been shown to directly bind to 7SK and form a complex with the snRNP (Barboric et al., 2007; Sobhian et al., 2010). Of note, the region of the 7SK RNA that Tat has been implicated in binding is the same region where Hexim1 binds the RNA and tethers P-TEFb to the snRNP.(Sobhian et al., 2010) This finding was further supported by the observation that the RNA binding region of Tat displays incredibly high homology to the RNA binding region of Hexim1 (Barboric et al. 2007). It was therefore proposed that HIV Tat mediates displacement of P-TEFb from the 7SK snRNP by physically releasing P-TEFb from the 7SK snRNP by competing with Hexim1 for binding to the RNA (Barboric et al., 2007; Muniz et al., 2010; Sobhian et al., 2010; Tahirov et al., 2010). From there, Tat then forms a complex with P-TEFb on the TAR RNA, directly adjacent to the paused Pol II, and transition Pol II into its

actively elongating form (Barboric et al., 2000; Liu et al., 2002; Tahirov et al., 2010). However around the same time this model was proposed, it was discovered that the majority of Tat inside of an HIV-infected T-cell localizes to the inner leaflet of the outer membrane. It was observed that only ~20% of Tat would fractionate to the nucleus, yet Tat retained its ability to potently transactivate the HIV genome (Rayne et al., 2010a; Rayne et al., 2010b). This observation supported the notion that there remained an as-yet characterized mechanism of P-TEFb release from the 7SK snRNP by Tat.

It was recently discovered by D'Orso and Frankel that the entire 7SK snRNP was recruited to the HIV promoter in the absence of Tat (D'Orso and Frankel, 2010). This observation led to the model that Tat releases P-TEFb at the viral promoter, allowing the kinase to be directly positioned for rapid extraction and subsequent phosphorylation of the paused Pol II. However a biochemical mechanism for release of P-TEFb from the 7SK snRNP at the HIV promoter was still unclear. Given that only a fraction of Tat inside of an HIV infected T-cell is localized to the nucleus (Rayne et al., 2010a; Rayne et al., 2010b), an enzymatic mechanism of release rather than a competition-mediated mechanism seemed likely.

In this dissertation, I propose a novel mechanism of transcriptional activation by which TF's such as Tat directly recruit the PPM1G phosphatase to the promoter, enzymatically releasing P-TEFb from the 7SK snRNP through dephosphorylation of the kinase. This enzymatic mechanism of P-TEFb extraction allows for rapid hyper-phosphorylation of the CTD of Pol II, leading to transcription elongation. Moreover, in addition to the HIV promoter, we have observed that the 7SK snRNP occupies promoters of Pol II poised genes genome wide. This placement of the 7SK snRNP at promoters, through the protein KAP1, allows TF's to locally extract P-TEFb and activate their targeted genes. Antagonizing 7SK snRNP recruitment

to promoters greatly reduces the magnitude of gene expression, demonstrating that promoter localization of this complex regulates the transcriptional output of cells in response to stimuli. We therefore propose a novel transcriptional signaling pathway in which TF's can enzymatically activate P-TEFb 'on site', leading to robust gene expression.

CHAPTER II – MATERIALS AND METHODS

Materials

5,6-dichloro-1-β-ribofuranosylbenzimidazole (DRB) (Sigma-Aldrich, D1916), and Triptolide (TPL) (Sigma-Aldrich) were diluted in dimethyl sulfoxide (DMSO) and stored at 4˚C. DRB and TPL were each used at a concentration of 1 μM. TNF was obtained in powdered form (Sigma-Aldrich, T6674) and was resuspended in water and maintained at -20˚C. TNF was used at a final concentration of 25-50 ng/mL, depending on the experiment. 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) (Life Technologies) was diluted in water to a final concentration of 5 mg/mL and maintained at -20˚C. DAPI was used at a concentration of 250 ng/mL. Zeocin and blasticidin were each diluted to 100 mg/mL and used at a final concentration of 100 μg/mL.

Plasmids and Antibodies

The plasmids used are shown in **Table 4**. Site-directed mutants were made using 100 ng of the indicated plasmid, 5′-phosphorylated oligonucleotides, and Phusion High-Fidelity Polymerase Chain Reaction (PCR) master mix (New England Biolabs). All antibodies used are shown in **Table 5.**

Cell Culture

HEK 293T, HeLa, HCT116, and U2OS were maintained in Dubellco's Modified Eagle Media (DMEM) with 10% Fetal Bovine Serum (FBS) (Sigma-Aldrich) and 1X Penicillin/Streptomycin at 37˚C with 5% CO_2. HEK 293T-REx cells were maintained as regular HEK 293T cells, but with the addition of 100 µg/mL of blasticidin and zeocin (Invitrogen). Jurkat cells were grown in suspension using Roswell Park Memorial Institute (RPMI) 1640 with 10% FBS and 1X Penicillin/Strepomycin at 37˚C with 5% CO_2.

Plasmid Transfections

HEK 293T cells were tranfected with 1-5 µg (depending on experiment) of supercoiled plasmid encoding gene of interest using 4X to quantify of DNA of Polyjet (SignaGen) (for example, 1 µg of DNA was combined with 3 µL of Polyjet). DNA was diluted in unsupplemented DMEM and Polyjet was diluted in the same volume of unsupplemented DMEM. The DMEM + Polyjet mixture was added to the DMEM + DNA mixture, and entire mixture was briefly vortexed and allowed to incubate at room temperature for 15 minutes. Transfection mixture was added directly to cells at a confluency of ~ 70% and cells were grown for additional 2 days.

Generation of HEK 293 T-REx Stable Cell Lines

HEK 293 T-REx cells were transfected with linearized plasmid as indicated above. 2 days after transfection, stable integrants were selected with the addition of 100 μg/mL of blasticidin and zeocin. Colonies of clonal populations were then grown in fully supplemented DMEM + 100 μg/mL of zeocin and blasticidin.

siRNA Transfections

Short interfering RNA (siRNA) mediated knockdown (KD) was done by transfection of 50 picomoles of siRNA (custom order through Sigma-Aldrich) using Lipofectamine 2000 (LifeTechnology) reagent. siRNA was diluted in unsupplemented DMEM and Lipofectamine was diluted in the same volume of unsupplemented DMEM. The DMEM + Lipofectamine mixture was added to the DMEM + siRNA mixture, and the entire mixture was briefly vortexed and allowed to incubate at room temperature for 15 minutes. Transfection mixture was then added directly to cells at a confluency of ~50% and cells were grown for additional 2 days. KDs were validated either through gene specific quantitative reverse transcriptase PCR (qRT-PCR) or Western blotting (see below). All siRNA's used are shown in **Table 6**.

Lentiviral shRNA Transductions

Short hairpin RNA (shRNA) mediated KD was done by the transduction of lentivirus carrying a self-inactiving pLKO.1 shRNA vector (custom ordered through Sigma-Aldrich, see

Table 4). Transductions were done using a multiplicity of infection (MOI) of 10-500 (depending on the shRNA) into cells using spinoculation combined with Polybrene (Sigma-Aldrich). Lentivirus was diluted in unsupplemented DMEM or RPMI (depending on cell type being transduced – see above) and Polybrene was diluted to 8 μg/mL in unsupplemented DMEM or RPMI. The DMEM/RPMI + Polybrene mixture was added to the DMEM/RPMI + lentivirus mixture, and entire mixture was briefly vortexed. The mixture was then added to the correct number of pelleted cells to achieve correct MOI and briefly vortexed. Cells + lentivirus mixture were then spinoculated at room temperature at 800*g for 2 hours. The supernatant was discarded and the cell pellet was resuspended in fully supplemented media and returned to incubator. 2 days post-transduction, stable Jurkat cells were selected for using 2 μg/mL of puromycin. Cells were maintained in fully supplemented RPMI media + puromycin and KDs were validated either through gene specific qRT-PCR or Western blotting (see below).

RNAi rescue assay

For the rescue assay, HeLa cells were first transfected 50 picomoles of siRNA using Lipofectamine 2000. Subsequently, cells were transfected with 100 ng of an empty pcDNA4/TO (–) or pcDNA4/TO-PPM1G:S (PPM1G) plasmid DNA using PolyJet. One day post-transfection, cells were re-transfected with the indicated siRNAs. Two days post siRNA transfection, cells were induced with TNF and used for RNA extraction and gene expression analysis by qRT-PCR and Western Blot.

Nucleic Acid Detection

To quantify RNA, cells were lysed using the RNease Plus (QIAGEN) and total RNA was extracted. To validate that there was no genomic DNA contamination, DNase treatment was performed directly on column by adding 10 units of DNase I with 1X DNase digestion buffer (Thermo Fisher) at room temperature for 15 minutes. For whole cell lysates, purified RNA was quantified using Nanodrop 2.0 and 500 ng – 2000 ng were reverse transcribed. Reverse transcriptions were conducted using the Maloney Murine Leukemia Virus (M-MuLV) Reverse transcriptase (NEB), oligo dT (Sigma-Aldrich), random decamers (Sigma-Aldrich), and dNTP (Invitrogen). Oligo dT and random decamers were annealed to RNA at 65˚C for 5 minutes in a 1.5 mL microcentrifuge tube and contents were spun down at 10,000 rpm for 30 seconds and placed on ice for 5 minutes. 10 units of enzyme were added to the annealed primer:RNA mixture along with 1X M-MuLV reaction buffer (NEB) and dNTP. Reverse transcription was done at 42˚C for 1 hour and the reaction was stopped by heating samples at 90˚C for 10 minutes. cDNA was then used for gene specific quantitation through qPCR using the ABI 3500. cDNA was added to a mixture containing diluted primer pairs (final concentration of 10 nM), and 1X SYBR green. Ct samples were used to quantify relative cDNA and all samples were standardized to an internal control. To quantify the relative cDNA between multiple samples (ΔCt of X and Y), the following equation was used $(Ct_{control}X - Ct_{control}Y) - (Ct_{gene}X - Ct_{gene}Y)$. To calculate fold changes, the ΔCt of no treatment was subtracted from that of the ΔCt for a specific time point (Tx) and fold change was calculated as $2^{\Delta CtTx}$. A more detailed version of the methods and mathematical formulas, please refer to (Schmittgen and Livak, 2008).

To detect RNA through northern blot hybridization, purified RNA was run on a polyacrylamide gel (Bio-rad) and transferred to a Zeta Probe membrane (Bio-rad). Probes were digoxigenin (DIG) labeled using the PCR DIG Labeling Kit (Roche) and incubated with the Zeta-Probe membrane containing the RNA at 65˚C overnight in hybridization buffer (1% bovine serum albumin (BSA), 7% sodium dodecyl sulfide (SDS), and 600 mM $NaHPO_4$). Hybridized probes were detected using horseradish peroxidase (HRP) labeled anti-DIG (Roche). rRNA was used as a loading control and was detected through ethidium bromide on the polyacrylamide gel prior to transfer.

UV Crosslinking of RNA-Protein Complexes *in vivo*

HEK 293T cells were seeded at $2x10^6$ cells into a 10-cm^2 plate and transfected with 5 µg of a Strep-tagged Tat plasmid 24 hours post seeding. For the experiment, cells were irradiated with 125 mJ/cm^2 using a Spectrolinker XL-1500 UV crosslinker and harvested and lysed using denaturing lysis buffer (150 mM NaCl, 20 mM Tris-HCl pH = 7.0, 5% Glycerol, 1.5 mM $MgCl_2$, 1 mM Dithiothreitol (DTT), 1% Nonidet P 40 (NP-40), 0.25% Deoxycholic acid and 0.1 % SDS) and sonicated using the Bioruptor water bath Sonicator (Diagenode) with 8 cycles (30 seconds on 30 seconds off). Tat complexes 12 were affinity purified using Strep-tactin Superflow Resin (IBA Life Sciences) and washed using denaturing washing buffer (250 mM NaCl, 20 mM Tris-HCl pH = 7.0, 5% Glycerol, 1.5 mM $MgCl_2$, 1 mM DTT, 0.05% NP-40, 0.1 % Deoxycholic acid, 0.1 % SDS). Co-immunoprecipitating RNAs were extracted using Proteinase K digestion [0.5 mg/mL Proteinase K, 20mM Tris-HCl pH = 7.0, 5 mM ethylene diamine tetra-acetic acid (EDTA), 0.5% SDS, 300 pg/µl *in vitro* transcribed β-globin RNA (used as precipitation control,

PC)] (Conrad, 2008) followed by Phenol:Chloroform:Isoamylic alcohol (25:24:1) extraction and ethanol precipitation. The immunoprecipitated material was analyzed by Northern blot using a 7SK probe encompassing stem-loop I were made and labeled as explained in previous section.

In Vitro Transcription and RNA-Binding Gel Shift Assays

7SK RNA was produced by *in vitro* run-off transcription with T7 RNA polymerase (NEB). The transcription reaction mixture contained 0.35 μmol of [α-^{32}P]-uridine triphosphate (UTP) 3000 Ci/mmol (PerkinElmer) along with non-radiolabeled nucleotides (500 μM ATP/cytidine triphosphatase/guanosine triphosphate and 100 μM UTP). RNAs were purified on a 5% polyacrylamide/1X Tris-Glycine gel and eluted overnight at 4°C in elution buffer (0.6 M NaOAc [pH 6.0], 1 mM EDTA, and 0.01% SDS), ethanol precipitated, and resuspended in water. Protein and RNA were incubated for 20 minutes at 4°C in 10 μl binding buffer (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-KOH [pH 7.5], 100 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 10% glycerol, and 50 μg/ml yeast tRNA). Protein-RNA complexes were resolved on a 5% polyacrylamide/1X Tris-Glycine gel at 4°C.

Transcription Reporter Assays

For the reporter assays, a 96 well plate containing HeLa cells were transfected using 0.5 μL of Polyjet with 25 ng of a firefly luciferase (FFL) reporter plasmid, 1 ng of a cytomegalovirus (CMV)-Renilla (RL) luciferase plasmid, and 1 ng of a Tat-expressing plasmid as described (D'Orso et al., 2012). Reporter activities are presented as fold activation relative to reporter

alone and normalized to RL. For the RNAi assays, 24 hours after siRNA transfection, cells were retransfected with a Tat plasmid and luciferase levels were measured using the Dual-Luciferase Reporter Assay (Promega) on a FLUOStar Optima 96-Ill plate reader (BMG Labtech).

Affinity Purification of Proteins from HEK 293T cells

Plasmids encoding a gene of interest were transfected into HEK 293T cells and proteins were purified through affinity purification (AP). Cells removed from plates and spun at 800*g and were washed 1X with cold, 1X PBS and spun down again at 800*g. Cell pellets were lysed with passive lysis buffer (20 mM Tris-HCl pH = 7.0, 150 mM NaCl, 1.5 mM MgCl2, 1 mM DTT, 1 mM PMSF, 1% NP-40, 5% Glycerol, and 1X EDTA-free protease inhibitor (Roche)) while nutating at 4˚C for 20 minutes, vortexing every 5 minutes. Cell lysates were then spun at 13,000*g at 4˚C for 10 minutes. During this time, AP beads (Strep-tactin High Capacity Superflow (IBA Life Sciences) or EZ-View Anti-FLAG M2 agarose (Sigma-Aldrich)) were equilibrated in lysis buffer. Soluble cell lysate was added to the equilibrated beads and allowed to bind at 4˚C for > 1 hour. Beads were then washed in passive wash buffer (20 mM Tris-HCl pH = 7.0, 150 mM NaCl, 1.5 mM MgCl2, 1 mM DTT, 0.5 mM PMSF, 0.05% NP-40, 5% Glycerol, and 0.5X EDTA-free protease inhibitor) or high salt buffer (20 mM Tris-HCl pH = 7.0, 500 mM NaCl, 1.5 mM MgCl2, 1 mM DTT, 1 mM PMSF, 1% NP-40, 5% Glycerol, and 1X EDTA-free protease inhibitor) for a total of 5 washes. For beads washed in high salt buffer, and additional two washes of passive wash buffer was included at the end to equilibrate proteins. Complexes were eluted off from Strep-tactin beads using 2X Strep elution buffer (D-desthiobiotin) diluted in passive wash buffer at 4˚C for 30 minutes. Complexes were eluted off

from FLAG beads using 200 ng/μL of 3X FLAG peptide diluted in wash buffer at 4˚C for 30 minutes.

Protein Purification from Bacterial Cells

PPM1A, PPM1G, and NF-κB were cloned into the pET30a vector (Novagen). BL21 DE3 cells were grown in Luria Broth medium and induced for protein expression with 1 mM isopropyl b-D1-thiogalactopyranoside at 30°C for 4 hours. Cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH = 7.5, 150 mM NaCl, 20 mM imidazole, 0.2 mM EDTA, 1 mM PMSF, 2 mM $MnCl_2$, 1 mM DTT, 5% glycerol, and 1% NP-40). Lysis was carried out by sonication and clarified by centrifugation at 10,000*g. The clarified supernatant was loaded onto a Ni-nitrilotriacetic acid affinity (Ni-NTA) column (Qiagen), washed in buffer (20 mM Tris-HCl pH = 7.5, 250 mM NaCl, 20 mM imidazole, 0.2 mM EDTA, 2 mM $MnCl_2$, 1 mM DTT, 5% glycerol, and 0.1% NP-40), and eluted with imidazole (50 mM, 100 mM, and 150 mM elutions). Recombinant HIV Tat was expressed as an N-terminal glutathione S-transferase (D'Orso and Frankel, 2010), or C-terminal Strep, fusion.

Immunoprecipitation of Endogenous Protein

Endogenous proteins were immunoprecipitated out of cells using antibody-coated Protein G Dynabeads (Lifetechnologies). Cells were lysed in passive lysis buffer and Protein G beads were equilibrated in lysis buffer prior to the binding of 1-5 μg of antibody (depending on the antibody). Antibody coated beads were blocked using 0.5 mg/mL of BSA (Research Products

International Core) and then added to centrifuged lysate and allowed to bind to their targeted antigen for > 2 hours at 4˚C. Beads were then washed using passive wash buffer (conditions listed above) and complexes were removed from beads through either elution buffer (100 mM $NaHCO_3$ and 1% SDS) or directly through heat at 90˚C for 15 minutes.

Tandem Affinity Purification (Strep ➔ FLAG TAP)

HEK 293T cells expressing one protein containing a strep tag (ie. Tat-Strep) and another protein containing a FLAG tag (ie. Larp7-FLAG) were used for tandem affinity purification. The first round of purification (Strep) was done as indicated above in the AP protein purification from HEK 293T cells section. Three elutions from the first AP was done and the Strep AP was then added to equilibrated FLAG resin. Secondary binding was done for > 2 hours at 4˚C and beads were washed using passive wash buffer for a total of 4 washes. Final tandem AP were eluted from the FLAG beads using 200 ng/µL of 3X FLAG peptide.

Preparation of Nuclear Extracts and Chromatin Fractions for Protein Affinity Purification

Nuclei from HEK 293T cells were prepared using the Dignam method (Dignam et al., 1983), except the lysis step in which nuclei were lysed using Buffer C containing 0.20 M NaCl. Nuclei were then resuspended in Buffer C, nutated for 45 minutes at 4ºC, spin down at 10,000*g for 15 minutes, and supernatants (soluble nuclear fraction) removed. Nuclear pellets were resuspended in Buffer C with high salt (0.50 M NaCl) and sonicated for 5 pulses (20 seconds ON, 40 seconds OFF) on the Diagenode Bioruptor. Nuclear debris was eliminated by centrifuging at

10,000*g at 4˚C for 15 minutes and chromatin-enriched fractions directly utilized for affinity purification.

Mass Spectrometry Analysis of Protein Complexes

AP protein complexes, either single or tandem purified, were submitted to the UTSW Proteomics core for protein identification (http://www.utsouthwestern.edu/research/core-facilities/proteomics-core.html). Purified complexes were run out on a polyacrylamide gel and gel slices were submitted for in gel trypsin digestion followed by liquid chromatography (LC) and Mass Spectometry by Mass Spectrometry (MS/MS) analysis. Samples were analyzed using an in-house data analysis pipeline from the UTSW Proteomics Service, with quantitation performed using the Normalized Spectral Index method (SINQ) as previously described (Trudgain et al. 2011). **Table 1** lists the high confidence interactors of the Tat-7SK snRNP and **Table 2** lists the high confidence interactors of the Larp7:S AP. Protein identification was inferred from the identified peptides. Protein inference was performed across samples, using the spectral counts or spectral index/ratio column to locate proteins specific to a certain sample. Protein identification was done using time of flight ionization and % coverage was calculated for confidence intervals.

Protein Detection and Quantitation

Proteins were detected by several independent methods, depending on the assay. To assay for protein purity, protein solutions were run on a polyacrylamide gel and subsequently

silver stain using the Pierce Silver Stain Kit (Pierce) or stained with coomassie blue. To quantify the concentration of the protein, a standard curve of known BSA standards was also run on a polyacrylamide gel ranging from 0.25 to 5 μg and compared to protein of interest. For detection through Western blotting, standard Western blotting techniques were used. Proteins run on a polyacrylamide gel were transferred to a nitrocellulose membrane and membranes were blocked at room temperature with 5% milk diluted in Tris-Boric Acid Saline with 0.1% TIen-20 (TBS-T) for 30 minutes. Membranes were then probed using primary antibodies diluted in 5% milk in TBS-T at dilutions ranging from 1:1000 – 1:10,000 (depending on the antibody). Membranes were washed 3X for in TBS-T and incubated in secondary antibody diluted in TBS-T at dilutions ranging from 1:5000 – 1:20,000. For horseradish peroxidase (HRP) conjugated secondary antibodies, membranes were washed 3X in TBS-T and incubated in ECL substrate (Bio-Rad) and imaged using X-ray film. For fluorescently labeled secondary antibodies, membranes were washed 3X in TBS-T and 2X in PBS and then imaged using the LiCor Odyssey. Band intensity was calculated using Image Studio.

P-TEFb Dephosphorylation and CDK9 T Loop Phosphorylation Assay

P-TEFb (CycT1:Strep + CDK9:FLAG) was purified using Strep-tactin resin from HEK 293T cells. CDK9 was dephosphorylated with PPM1G, and P-TEFb was subsequently purified using FLAG beads. *In vitro* dephosphorylation of P-TEFb (0.5 μg) was carried out in dephosphorylation buffer (20 mM Tris-HCl pH = 7.4, 150 mM NaCl, 5 mM imidazole, 10 mM $MnCl_2$, and 1 mM DTT) at 30°C for 2 hours with 1 μg of the indicated phosphatase. Reactions were stopped by the addition of 2.5 mM EDTA and boiled at 90°C for 5 minutes in laemmli

buffer. For the release assay, dephosphorylation reactions were diluted in half in buffer (20 mM Tris-HCl pH = 7.5, 150 mM NaCl, and 0.05% NP-40) and incubated with FLAG beads overnight. The reaction products were electrophoresed on a SDS-PAGE gel and analyzed by Western blot.

For auto-phosphorylation of P-TEFb, dephosphorylated P-TEFb was then incubated with bacterially synthesized Strep-tagged Tat or a C22G mutant in kinase buffer (10 mM Tris-HCl pH = 7.0, 150 mM KCl, 0.1 mg/ml BSA, 1.5 mM $MgCl_2$, and 0.5 mM ATP) at 30°C for the indicated time points. Kinase reactions were stopped with the addition of laemmli buffer and boiled for 10 minutes at 90°C. Quantification of relative phosphorylation intensities was done using Image J.

Quantitation of TNF-Induced Cell Death

HeLa cells were seeded in 6-well plates and transfected with the indicated siRNAs. At 48 hours post-transfection, cells were treated with TNF and cell death was monitored over time through picture images (n = 3). The graph was plotted by counting the cells within a defined area (three images per time point (cell counts were averaged)). The percent of viable cells expressed as fold change ± TNF was calculated.

***In vivo* RNA Immunoprecipitation**

Cells were cross-linked with formaldehyde to stably trap RNA–protein complexes formed *in vivo*. Briefly, HEK 293T cells were crosslinked in 0.5% formaldehyde for 10 minutes at room temperature, and crosslinking was quenched with addition of 0.15 M Glycine. Cells were pelleted and washed 2X with cold, 1X PBS. Cells pellets were then resuspended in RIPA

buffer (50 mM Tris pH = 7.5, 1% NP-40, 0.5% sodium deoxycholic acid, 0.1% SDS, 1.5 mM $MgCl_2$ and 150 mM NaCl, 2.5 mM EDTA, 2 mM PMSF, 10 U/mL of RNaseIN (Promega), and 1X EDTA-free protease inhibitor) and sonicated for 8 cycles using the Diagenode Bioruptor. Cell lysates were spun at 13,000*g at 4˚C for 10 minutes and lysates were added to equilibrated, antibody bound Protein G Dynabeads. Binding of antibody coated beads to antigen was allowed to take place at 4˚C for > 2 hours. Beads were washed four times for 10 minutes in binding buffer supplemented with 1 M urea and 0.5 M NaCl and then resuspended them in 50 mM Tris, pH = 6.8, 5 mM EDTA, 1% SDS and 10 mM DTT. Cross-links were reversed by heating the samples for 45 minutes at 70 °C, RNA was purified, and first-strand cDNA synthesis was performed using random decamers. 7SK and U6 RNAs were detected by qRT-PCR.

Indirect Immunofluorescence

Cells were plated onto a 6-well plate containing glass coverslips and allowed to adhere to plates and grow overnight. Media was removed and cells were washed 1X with cold, 1X PBS. Cells were then crosslinked using 4% para-formaldehyde (Sigma-Aldrich) diluted in 1X PBS at room temperature for 10 minutes. Slides were then washed 2X with cold, 1X PBS and permeabilized using 0.5% Triton-X100 (Fisher) diluted in 1X PBS at room temperature for 10 minutes. Slides were then washed 1X with cold, 1X PBS and blocked using 10% BSA at 4˚C for 1 hour. Primary antibodies were diluted in 10% BSA at dilutions ranging from 1:200 – 1:2000 (depending on the antibody) and slides were incubated with primary antibody at 4˚C for > 1 hour. Slides were then washed 4X with cold, 1X PBS. Secondary antibodies were diluted in 10% BSA at 1:1000 and slides were incubated with secondary antibodies at 4˚C for > 1 hour.

Slideswerethen washed 4X in cold, 1X PBS and were incubated with DAPI (1:25,000 dilution in PBS) for 2 minutes. Slides were then washed 3X with cold, 1X PBS and were allowed to dry at room temperature for > 20 minutes. Slides were then mounted using Prolong Gold mounting solution (Thermo Fisher).

Confocal Microscopy

Confocal imaging of cells was done using the Zeiss LSM 510 Meta maintained through the UTSW Live Cell Imaging Core. Z-stack images (7-15 slices, depending on the cell) were obtained and de-convoluted using the program AutoquantX3. Images were then assayed using the program Imaris 6.2.2. For co-localization analysis, statistical significance (p-value) was determined through co-localization of pixels compared to that of a random pixel generator. Additionally, Mander's and Pearson's coefficients (Egly and Coin, 2011) were generated using Imaris and recorded for each image. Images of individual slices of a de-convoluted cell were then transferred as high resolution .tiff files. Individual colors, as well as overlaid images, were done through the program ImageJ.

***De Novo* Modeling and Structure Alignment**

PPM1G structure prediction was done using the Robetta server (Kim et al., 2004). Structural similarities were identified after superimposition of the model into the crystal structures using MAMMOTH. Three PPM1G domains (N-terminal, acidic, and C-terminal) were modeled using four different structures. The N-terminal (residues 1–140) was modeled using the

Ser/Thr phosphatase 2C (Protein Data Bank [PDB] ID code 2I0O; chain A). Two regions in the acidic domain (residues 141–231 and 232–324) were modeled with hemagglutinin esterase fusion protein (PDB 1FLC; chain B) and contractile protein (PDB 1C1G; chain A) structures. The C-terminal (residues 325–546) was modeled using PPM1A structure (PDB 1A6Q; chain A). The four models were assembled into a full-length model packed using the Monte-Carlo algorithm with a backbone-dependent side-chain rotamer library. The predicted PPM1G structure (without the acidic domain) and PPM1A structure (PDB 1A6Q; Das et al., 1996) were aligned and superimposed using the Matchmaker option within the University of California, San Francisco Chimera molecular modeling package (Pettersen et al., 2004).

Chromatin Immunoprecipitation (ChIP)

Cells were cross-linked to chromatin with 1% (v/v) formaldehyde for 10 minutes at room temperature and stopped the reaction by adding 0.15 M glycine. Crosslinked cells were then pelleted at 4˚C and washed 2X with cold, 1X PBS. Cell pellets were then resuspended Farham lysis buffer (5 mM piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES) pH = 8.0, 85 mM KCl, 2 mM PMSF, 1X EDTA-free protease inhibitor, 0.5 % NP-40) at $1.0*10^7$ cells/mL and incubated at 4˚C for 30 minutes. Suspension was then spun at 4,000 rpm at 4˚C for 10 minutes. The pelleted nuclei were then resuspended in RIPA buffer (50 mM Tris-HCl, pH = 8.0, 150 mM NaCl, 2.5mM EDTA, 1% NP-40, 0.1% SDS, 0.5% deoxycholic acid, 1X EDTA-free protease inhibitor, 2 mM PMSF) at $2.5*10^7$ cells/mL and homogenized through dounce tissue homogenizer for 10 strokes. Cell lysates were sonicated using the Diagenode Bioruptor to generate samples with ~300bp DNA fragments, which amounted to be 45 cycles for HCT116

and HeLa cells, and 60 for Jurkat cells (30 seconds on and 30 seconds off). Protein G magnetic Dynabeads (Life Technologies) were equilibrated in RIPA buffer prior to addition of 1-5 µg of antibody diluted in RIPA buffer. Antibodies were then allowed to bind to Protein G beads at room temperature for 30 minutes. Beads were then blocked using 0.5 mg/ml BSA and 0.3 mg/ml of salmon sperm DNA at room temperature for 1 hour. Sonicated cell lysates were then spun at 13,000*g at 4˚C for 10 minutes and added to the equilibrated beads. Binding of antigen to the antibody coated beads was done for > 2 hours at 4˚C. Unbound material was discarded and beads were then washed 2X in RIPA buffer followed by subsequent washes using low salt buffer (1% NP-40, 20 mM Tris HCl pH = 8.0, 150 mM NaCl, 0.1 % SDS, 0.5 % deoxycholic acid, 2.5 mM EDTA, 1 mM PMSF), high salt buffer (1% NP-40, 20 mM Tris HCl pH = 8.0, 500 mM NaCl, 0.1 % SDS, 0.5 % deoxycholic acid, 2.5 mM EDTA, 1 mM PMSF), LiCl buffer (1% NP-40, 20 mM Tris HCl pH = 8.0, 250 mM LiCl, 0.1 % SDS, 0.5 % deoxycholic acid, 2.5 mM EDTA, 1 mM PMSF), and twice in 1X Tris-EDTA buffer (TE), all at 4˚C. Complexes were then eluted off from beads using 100 µL of elution buffer (100 mM NaHCO3, 1 % SDS) at 65˚C for 30 minutes, vortexing every 10 minutes. Eluted material was incubate in 100 µL of reverse crosslinking buffer (50 mM Tris-HCl pH = 6.8, 500 mM NaCl, 5 mM EDTA, 0.5mg/mL Proteinase K) at 65˚C for > 4 hours. DNA was then purified using Zymo DNA ChIP Clean and Concentrator Kit (Zymogen) and analyzed by qPCR using conditions listed above. The amount of DNA immunoprecipitated was measured using the ΔCt method, by quantifying the amount of immunoprecipitated DNA/total input.

ChIP-seq Library Preparation and Sequencing

ChIP DNA was quantified on a Qubit® 2.0 Fluorometer (Life Technologies) and ~20 ng DNA were submitted to the McDermott Center Sequencing Core at UT Southwestern Medical Center for library preparation and high-throughput sequencing using the NextSeq 500 kit (Illumina) for the Illumina NextSeq500 instrument. Samples were end repaired, 3'-end adenylated and barcoded with multiplex adapters (Applied Biosystems). After purification with Ampure XP beads (Beckton Coulter), samples were PCR amplified (~15 cycles), size selected with Ampure XP beads, and quantified on the Agilent 2100 Bioanalyzer. Emulsion PCR was then performed and beads enriched on the EZ-Bead system. Each sample yielded 25-40x10^6 50-nt sequencing reads depending on the sample, of which ~70-80% reads mapped uniquely to the human genome (GRCh37/hg19).

Computational Analysis

All scripting was performed using python 2.7.6. All ChIP-seq binding events were loaded into a custom MySQL database to allow for efficient comparisons of multiple factors' binding loci. Co-localizing events were determined if peaks were either within 250 bp of each or, in the case of the broad peaks obtained from Pol II, the summit of the Pol II peak was used for co-localization.

Mapping and Sorting

The output from the ChIP-seq was aligned to the reference human genome (GRCh37/hg19) using Bowtie v1.0.0. (Langmead et al., 2009), allowing 1 mismatch (-v 1). Sequences aligning to multiple locations in the reference genome were discarded. -S was specified to deliver.sam output, which was then sorted and converted to binary (.bam) using Samtools (Li et al., 2009).

ChIP-seq Peak Calling

The .sam files for all ChIP-seq marks were analyzed for binding events using peak calling in the MACS2 software (Zhang et al., 2008). When determining peaks, duplicate read alignments were discarded to avoid amplification artifacts from the library preparation.

ChIP-seq Data Visualization

MACS2 was run in the "call peak" mode with the –B and –SPMR FLAGs set to generate signal pileup tracks in bedGraph format on a per million reads basis. This allows for direct comparison between all samples, regardless of any differences in depth of the sequencing experiment. The pileup tracks were visualized and images were generated using Integrative Genomics Viewer (IGV) (Robinson et al., 2011).

Generation of ChIP-seq Heatmaps

We used the HOMER (Heinz et al., 2010) program annotatePeaks.pl to generate density matrices using normalized bedGraph pileup tracks of ChIP-seq marks as input. The –hist flag was set with a parameter of 50 bins and the –ghist flag was also set to generate a density matrix. The TSS was used as a centering point to generate the heatmaps.

Generation of ChIP-seq Metagene Plots

HOMER (Heinz et al., 2010) annotatePeaks.pl was used to generate metagene plots based on the average profile of a series of ChIP-seq marks in normalized bedGraph pileup tracks. The –hist flag was set using 50 bins for all metagene profiles.

GRO-seq Analysis

The GRO-seq dataset (GSE38140) generated by the Espinosa lab (Galbraith et al., 2013) was used to define the transcriptional status of all human genes in the HCT116 colonic cancer cell line containing transcriptionally engaged Pol II and KEC. Heat maps were generated as above and clustered to Pol II decreasing occupancy for all unique TSS in Refseq.

Data Availability

The following data was been deposited in the NCBI GEO (under accession no.

GSE72622): ChIP-seq raw sequence tags and bigwig files for Pol II, S5P-CTD Pol II, H3K4me3, H3K4me1, H3K27Ac, KAP1, CDK9, Larp7 and Hexim1 from HCT116 colon carcinoma cell line.

CHAPTER III – RESULTS

PART I

7SK RNA Cofactors Assemble with Tat into a Ribonucleoprotein Complex to Control HIV Transcription

To investigate the mechanisms by which transcription factors recruit P-TEFb from the 7SK snRNP during selective activation of transcriptional programs, we used HIV as a model system. Tat promotes the disassembly of the 7SK snRNP at the viral promoter and transfers P-TEFb to TAR, stimulating Pol II phosphorylation and transcription elongation (**Figure 4A**). Because Tat has been shown to complex with the 7SK snRNP (D'Orso and Frankel, 2010a; Faust et al., 2012; Krueger et al., 2010; Muniz et al., 2010) we hypothesized that Tat assembles a Tat-7SK ribonucleoprotein (RNP) complex to regulate the release of P-TEFb from the 7SK snRNP at the HIV promoter during the transition into elongation.

To identify Tat cofactors involved in the release of P-TEFb from the 7SK snRNP, the Tat-7SK RNP complex was isolated through tandem affinity purification (TAP) of Tat and Larp7 (**Figure 4B**). Larp7 was used to select for 7SK-associated fractions of Tat because it is constitutively bound to the RNA (Jeronimo et al., 2007; Krueger et al., 2008). Silver staining and northern blot of the final eluate revealed the presence of Tat, Larp7, 7SK RNA, as well as several bands that were not present upon co-purification of Larp7 without Tat (**Figure 4B**). To reveal the identity of the additional bands, gel slices were subjected to an in-solution complex mixture to mass spectrometry analysis. In addition to Tat and Larp7, eight high-confidence

interactors were identified (**Table 1**). Despite repeated attempts, however, P-TEFb and Hexim1 were not detected in the Tat-7SK RNP complex through Western blot and mass spectrometry, suggesting they are released upon binding of Tat along with the cofactors identified.

The proteomics data set was validated by protein-protein interaction assays by co-transfecting cells with Tat along with each of the identified interactors or negative (GFP) and positive (Larp7, CycT1, and CDK9) controls. We observed that MePCE, Sart3, nucleophosmin 1 (NPM1), PPM1G, and heterogeneous nuclear RNP-F (hnRNP-F) co-purified with Tat, whereas SET and DDX21 did not under the conditions tested (**Figure 4C**). By compiling the protein-protein and protein-RNA interaction data sets, we defined a network of interactions within the Tat-7SK RNP complex (**Figure 4D**). To address if Tat directly contacts the 7SK RNA *in vivo*, we performed a UV-crosslinking assay, which crosslinks proteins to nucleic acids only if there is direct contact. We observed that Tat can indeed crosslink to the 7SK RNA (**Figure 5A**). To identify whether the Tat interactors associate with the 7SK snRNP in cells independently of Tat, we performed RNA-immunoprecipitation (RIP) assays. Purification of the proteins followed by northern blot of associated RNA's, as well as RNA gel shift assays yielded variable association of these identified cofactors with 7SK RNA, with the exception of NPM1 and SET (**Figures 5B–D**). We hypothesized that factors associated with this RNP complex dictate the release of P-TEFb from the 7SK snRNP assembled at the HIV promoter. Here, we focus on the PPM1G phosphatase because of the possible enzymatic disassembly of the snRNP during HIV transcription activation (see below).

Tat Recruits PPM1G to Dephosphorylate the T Loop of CDK9 and Release P-TEFb from the 7SK snRNP Complex

Previous studies have demonstrated that phosphorylation of Thr186 (Phos-T186) at the activating T loop of CDK9 (P-CDK9) promotes the assembly of P-TEFb with Hexim1 and 7SK RNA (Chen et al., 2004; Li et al., 2005). Because the Tat-7SK RNP lacks P-TEFb but contains PPM1G (a member of the PPM/PP2C family of nuclear, metal-dependent Ser/Thr phosphatases (Allemand et al., 2007), we hypothesized that CDK9 dephosphorylation at Thr186 by PPM1G is a major determinant by which Tat induces the disassembly of P-TEFb from the 7SK snRNP. To test this, we initially purified wild-type PPM1G and a catalytically inactive D496A mutant (Allemand et al., 2007) and observed that PPM1G (but not D496A) dephosphorylates CDK9 at Thr186 *in vitro* (**Figure 6A**). To study if CDK9 dephosphorylation triggers P-TEFb release from the snRNP, 7SK-bound P-TEFb was incubated with PPM1G under dephosphorylation conditions and then subsequently P-TEFb to monitor its assembly into the 7SK snRNP (**Figure 6B**). Supporting our hypothesis, CDK9 dephosphorylation by PPM1G (but not D496A) at Thr186 did indeed release P-TEFb from the 7SK snRNP, as shown by the presence of the inhibitory 7SK snRNP components (devoid of P-TEFb) in the supernatant of a P-TEFb purification (**Figure 6B**). Collectively, this data strongly support a functional link between CDK9 T loop dephosphorylation by PPM1G and the release of P-TEFb from the 7SK snRNP *in vitro*.

Because PPM1G and P-TEFb do not stably associate *in vitro* or in cells, we reasoned that Tat mediates the recruitment of PPM1G to the 7SK snRNP to dephosphorylate the T loop of CDK9, thereby releasing P-TEFb for gene activation. To test this hypothesis, we co-transfected Tat along with PPM1G or PPM1A (another family member shown to dephosphorylate P-TEFb

in vitro (Chang et al., 2009; Wang et al., 2008)) and observed that Tat selectively associates with PPM1G, but not PPM1A (**Figure 6C**). The structure of PPM1G is currently unknown, but PPM1A folds with a catalytic domain composed of a central beta sandwich that binds two Mn2+ ions surrounded by α helices (Das et al., 1996) (**Figure 6D**). The phosphatase domains of PPM1A and the structure prediction of PPM1G contain the four Asp residues required for metal coordination and catalysis (**Figures 6D** and **E**), but PPM1G distinguishes itself from the other PPM family members with an internal acidic domain (**Figure 6C**). Co-transfection of Tat along with several PPM1G domains demonstrated that the acidic domain in PPM1G mediates Tat binding (**Figure 6F**), explaining Tat's specificity for PPM1G over PPM1A.

To further define how Tat binds the enzyme, PPM1G was co-transfected with several Tat fragments which demonstrated that the activation domain of Tat mediates its interaction with PPM1G (**Figure 6G**). In agreement, mutations in the activation domain that eliminate Tat transactivation (such as K41A; (D'Orso et al., 2012a)) abolish the Tat-PPM1G interaction (**Figure 6H**).

We next wanted to address if Tat directly contacts PPM1G, or there is need for an intermediary factor. We thus used bacterially expressed Tat and PPM1G, which are devoid of eukaryotic interacting partners as well post-translational modifications added through eukaryotic enzymes. We observed that the bacterially synthesized PPM1G and Tat indeed directly binds *in vitro*, whereas no detectable interaction was observed between PPM1A and Tat (**Figure 7A**). Moreover, we were able to tandem affinity purify of the Tat-PPM1G complex from mammalian cells, further supporting that Tat and PPM1G physically interact (**Figures 7B**). Taken together, we propose that Tat recruits the PPM1G phosphatase to the 7SK snRNP through directly contacting the enzyme.

Recruitment of PPM1G by Tat Activates the Transition into Elongation by Disassembling the HIV-Promoter-Bound 7SK snRNP Complex

The finding that PPM1G associates with the Tat activation domain suggests that Tat may recruit the enzyme to the 7SK snRNP complex during HIV transcription. To determine if PPM1G is required for Tat activation, we utilized a genetic interaction approach. HeLa cells were transfected with short interfering RNAs (siRNAs) to knockdown (KD) PPM1G, CDK9 (used as a positive control), or a non-target control siRNA, and we measured Tat-mediated transcriptional activity on an HIV long terminal repeat (LTR) luciferase reporter (**Figure 8A**). The KD of PPM1G and CDK9 reduced the Tat activation step (~4- to 5-fold) without affecting Tat levels (**Figures 8A**).

To put these findings in the context of HIV, we assayed the amount of infectious particles produced from human embryonic kidney (HEK) 293T cells transfected with a provirus DNA after KD of PPM1G or CDK9. Similar to CDK9 KD, which affects proviral transcription (Ott et al., 2011), loss of PPM1G resulted in a sharp decrease in the production of viral particles, as revealed by a reduction in the levels of the p24 capsid (CA) protein (**Figure 8B**). The processing of the Gag precursor into the smaller products CA and matrix (MA) is required for viral assembly and egress from the infected cell. To evaluate whether the reduction of extracellular p24 is due to improper Gag processing, intracellular Gag levels were assayed by Western blot and observed that the ratio of processed to unprocessed Gag is not affected by the loss of PPM1G (**Figure 8B**). Collectively, these results support a transcriptional, but not a posttranslational, role for PPM1G in the viral life cycle.

Because PPM1G is involved in HIV transcription and Tat functions to relieve the block at

the elongation step, we asked if PPM1G promotes the initiation-to-elongation switch at the viral promoter. To test this, a quantitative elongation assay to monitor levels of promoter-proximal and -distal transcripts was employed, which approximate the extent of initiation and elongation, respectively (**Figure 8C**). Even in the presence of Tat, PPM1G KD markedly decreased (~5-fold) the production of promoter-distal transcripts, thus negatively affecting the transcription elongation step. This elongation blockage by PPM1G KD phenocopies the effect of CDK9 KD or kinase inactivation with the Pol II transcription elongation inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) (Kim et al., 2002; Marshall and Price, 1992). These results functionally link the release of P-TEFb from the 7SK snRNP by Tat and PPM1G with activation of the elongation step.

We then asked whether decreased elongation due to PPM1G KD correlated with reduced P-TEFb recruitment and Pol II phosphorylation at the HIV promoter during Tat activation. To assess this, we performed a chromatin immunoprecipitation (ChIP) assay on the HeLa-LTR cell line by depleting PPM1G in the absence or presence of a functional Tat **(Figure 8D**). Assembly of HIV transcription preinitiation complexes at the promoter (demonstrated by Sp1 occupancy and levels of the initiating Ser-5 CTD-phosphorylated Pol II (S5P-CTD)) were minimally affected by PPM1G KD. However, PPM1G KD markedly reduced the Tat-mediated increase in Pol II S2P-CTD levels and P-TEFb occupancy in downstream regions (lowered by a factor of ~2.5–9-fold), consistent with a role of PPM1G in elongation. This finding correlates well with the inability of Tat to disassemble the 7SK snRNP (Hexim1 and Larp7) at the promoter. Residual levels of the 7SK snRNP were still be detected after transfection with Tat, which we attribute to not all cells expressing equal levels of Tat.

To test if Tat mediates recruitment of PPM1G to the HIV promoter, we performed a ChIP assay on the HeLa-LTR cell line in the absence or presence of Tat (**Figure 8E**). Whereas in the absence of Tat there is no detectable PPM1G, Tat recruits PPM1G to the promoter (~4-fold over no Tat). Together, Tat mediates the enzymatic disassembly of the 7SK snRNP from the viral promoter *in vivo* to release P-TEFb for transcription elongation.

Because PPM1G appears to be implicated in transcription, we next examined if the other components of the Tat-7SK RNP complex identified by mass spectrometry (**Figure 4**) also play a similar role. Using the same approach as we did with PPM1G, we performed siRNA-mediated KD of each of the components identified in the Tat-7SK snRNP (**Figure 9A**). The KD of Sart3, NPM1, and hnRNP-F affect the Tat activation step in a reporter assay, thus implicating these factors, like PPM1G, in HIV transcription (**Figure 9B**). While we cannot conclude precisely what the role of these factors are in Tat-mediated activation of the reporter, the strong antagonizing effects of their KD on LTR activation supports a model in which Tat recruits many cellular cofactors to the viral promoter for full genome transactivation.

Tat Binds Released Dephosphorylated P-TEFb and Stimulates Phosphorylation of the CDK9 T Loop to Activate the Transition into Elongation

The release of P-TEFb from the promoter-assembled 7SK snRNP, upon dephosphorylation of the CDK9 T loop, indicates that Tat must capture dephosphorylated P-TEFb prior to the assembly of Tat:TAR:P-TEFb elongation complexes (**Figure 6A**). To test whether the phosphorylation state of the CDK9 T loop plays any role in Tat binding, we purified P-TEFb and obtained homogeneous T loop phosphorylated and unphosphorylated kinase

preparations that were incubated with Tat in an *in vitro* binding assay. we observed that Tat binds both P-TEFb forms *in vitro* (**Figure 10A**), demonstrating that Tat can capture dephosphorylated P-TEFb upon its release from the promoter-bound 7SK snRNP.

Because CDK9 T loop dephosphorylation renders an inactive kinase, we hypothesized that, by binding dephosphorylated P-TEFb, Tat can stimulate auto-phosphorylation of the CDK9 T loop. To test this, dephosphorylated P-TEFb was incubated in kinase buffer with ATP: (1) alone, (2) with Tat, or (3) with a nonfunctional C22G mutant and measured levels of CDK9 auto-phosphorylation at Thr186. Tat, but not the C22G mutant, stimulates (~5-fold) CDK9 auto-phosphorylation *in vitro* in conditions where the kinase activity is in the linear range (**Figures 10B** and **C**).

Collectively, we propose an enzymatic model of P-TEFb transfer from the promoter-assembled 7SK snRNP to TAR during HIV transcription activation (**Figure 10D**). In a first step, Tat recruits PPM1G to the promoter to dephosphorylate the T loop of CDK9 and release P-TEFb from the 7SK snRNP. In a second step, Tat binds dephosphorylated P-TEFb to stimulate auto-phosphorylation of the CDK9 T loop and transfers the locally released kinase to TAR to assemble transcription elongation complexes. This enzymatic mechanism of 7SK snRNP disassembly takes into consideration the low levels of Tat during infection and is favored over the previously proposed competition model where excess Tat physically captures P-TEFb from the 7SK snRNP (Barboric et al., 2007; D'Orso and Frankel, 2010a; Krueger et al., 2010; Sedore et al., 2007). In fact, we observed that the competitive release of P-TEFb by Tat occurs primarily due to protein reassortment *in vitro*, but not in cells, because Tat and P-TEFb ectopically expressed in separate plates can efficiently associate in a test tube leading to Hexim1 displacement (**Figure 11**).

PPM1G Is a Transcriptional Co-activator of NF-κB in the Inflammatory Pathway

Because P-TEFb is a general elongation factor, we asked whether PPM1G is required in the context of transcriptional programs activated during cellular responses. we examined the inflammatory pathway where the proinflammatory cytokine tumor necrosis factor-alpha (TNF) triggers signaling cascades that converge on the activation of the transcription factor NF-κB (Ghosh and Karin, 2002). Upon TNF stimulation, NF-κB rapidly translocates from the cytoplasm into the nucleus to activate a set of inflammatory-responsive genes in a P-TEFb-dependent manner (Barboric et al., 2001a; Nissen and Yamamoto, 2000). PPM1G KD sharply downregulates (~7–15-fold) the induction of two NF-κB target genes (interleukin (IL)-8 and IκBα) upon TNF treatment without affecting several nonresponsive genes (**Figure 12A**). Although PPM1G KD virtually abolishes the activation of inflammatory-responsive genes, it does not alter NF-κB translocation to the nucleus upon TNF treatment (**Figure 12B**), implying that PPM1G functions as a nuclear regulator of NF-κB activity.

Upon stimulation with proinflammatory cytokines, NF-κB induces the transcription of a number of antiapoptotic genes (Barkett and Gilmore, 1999; Ghosh and Karin, 2002; Sidoti-de Fraisse et al., 1998). To investigate the biological significance of PPM1G in these processes, we analyzed its involvement in the apoptotic response regulated by NF-κB (**Figure 12C**). To this end, we transfected HeLa cells with control and PPM1G siRNAs, and 2 days later, the cells were treated with or without TNF. we assayed for cell death at different time intervals post-treatment and then computed a cumulative cell death index by plotting the effect of TNF on cell viability. Treatment with TNF was not toxic to cells transfected with control siRNA; however, PPM1G KD drastically sensitized cells to TNF-induced apoptosis (**Figure 12C**). This suggests that NF-

κB requires PPM1G, in addition to P-TEFb (Barboric et al., 2001a), to stimulate transcription elongation of prosurvival genes to prevent apoptosis induced by proinflammatory cytokines.

Given the fact that Tat physically recruits PPM1G to stimulate HIV transcription and that PPM1G is required for activation of inflammatory-responsive genes, we asked whether NF-κB binds PPM1G. To test this, we crosslinked cells after TNF treatment to eliminate any *in vitro* protein reassortment during the affinity purification step. We observed that transfected PPM1G stably associates with endogenous NF-κB in a stimulus-dependent manner (**Figure 12D**). Supporting this result, endogenous PPM1G and NF-κB proteins co-localize in the nucleus upon TNF treatment (**Figure 12E**) and bacterially synthesized NF-κB directly binds PPM1G *in vitro* (**Figure 12F**). Taken together, these results **functionally** link PPM1G and NF-κB during activation of the inflammatory transcriptional program.

To rule out potential RNAi off-target effects, we used two siRNAs (**Figures 13A** and **B**) to target a noncoding region in PPM1G and observed that subsequent transfection of a siRNA-resistant PPM1G construct (**Figure 13C**) restored the stimulus-dependent activation of inflammatory-responsive genes (**Figure 13D** and **E**). These results demonstrate that PPM1G is an essential nuclear cofactor of NF-κB for the normal activation of the inflammatory transcriptional program.

The PPM1G Phosphatase Controls NF-κB Signaling through Directly Binding the 7SK snRNP in a TNF-Dependent Manner

The finding that PPM1G co-purifies with the 7SK snRNP prompted us to characterize the interaction of the enzyme with the RNA and core snRNP components. To examine the

interaction with 7SK RNA, we performed a RIP assay from cells transfected with Strep-tagged PPM1G and PPM1A (**Figure 14A**). To test whether PPM1G directly associates with 7SK RNA, we purified recombinant PPM1G and PPM1A proteins from bacteria and incubated them with a radiolabeled 7SK RNA. We observed that PPM1G, but not PPM1A, directly binds 7SK RNA in *in vitro* gel shift assays (**Figure 14B**). In addition to binding 7SK RNA, we observed that PPM1G, but not PPM1A, directly interacts (in the absence of 7SK RNA) with MePCE and Larp7 core snRNP components *in vitro* (**Figure 14C**), demonstrating that PPM1G can make multiple contacts with the 7SK snRNP complex. This data further supports our previous findings that KD of PPM1G, but not PPM1A, affects transcription of inflammatory-responsive genes by NF-κB (**Figure 12A**).

All aspects of transcription and its regulation involve dynamic events where regulatory components are recruited to their target genes (Hager et al., 2009). Thus, although PPM1G has RNA-binding activity *per se*, we envisioned that there is a stimulus-dependent PPM1G-7SK RNA interaction upon activation of the inflammatory transcriptional program. Using a RIP assay, we observed that treatment of HeLa cells with TNF induces the recruitment of endogenous PPM1G to 7SK RNA (**Figure 14D**).

To test whether there is a functional link between NF-κB and the recruitment of PPM1G to the 7SK RNA, we used RNAi to KD NF-κB and compared the levels of PPM1G-7SK RNA association with and without TNF using a RIP assay (**Figure 14E**). Intriguingly, KD of NF-κB reduced (by a factor of ~2-fold) the PPM1G-7SK RNA association in the presence of TNF, which correlates with the decreased expression of NF-κB target genes. Thus, NF-κB, in addition to binding its target genes, plays a key role in the stimulus-dependent recruitment of PPM1G to the 7SK RNA to activate transcription.

To further define the role of PPM1G during activation of the inflammatory transcriptional program, we knocked down PPM1G from cells treated with or without TNF and examined Pol II and cofactor recruitment by ChIP at the IL-8 locus (**Figure 14F**). PPM1G KD reduces (~2–4-fold) the TNF-mediated increase of Pol II and P-TEFb occupancy in the gene body, which correlates with our previous observation in reduced gene expression (**Figure 12A**). Similar to the HIV promoter, we observed that the inhibitory 7SK snRNP components (Larp7 and Hexim1) are recruited to the promoter and TNF treatment promotes their eviction. This TNF-mediated ejection of the 7SK snRNP relies on PPM1G because KD of the enzyme resulted in retention of both Hexim1 and Larp7 at the IL-8 promoter.

To test the model that PPM1G is required to directly activate transcriptional pause release through eviction of the 7SK snRNP at the IL-8 locus, we examined PPM1G recruitment by ChIP (**Figure 14G**). We observed that PPM1G is recruited to the IL-8 promoter, but not to the gene body, upon TNF stimulation, mirroring the occupancy profile of NF-κB. This result demonstrates that the recruitment of PPM1G by NF-κB to the IL-8 promoter is inducible and argues a role for PPM1G in activating the transition into elongation.

Because P-TEFb is a general elongation factor, we examined the requirement of PPM1G for several Pol II-transcribed genes. As has been previously reported, CDK9 KD antagonized the activation of heat-shock-responsive genes (Lis et al., 2000); however, KD of PPM1G had no significant impact on activation of Hsp70 and only a minor impact on Hsp90 activation upon heat shock stress (**Figure 15**). While Hsp90 induction was slightly reduced in PPM1G KD, the far more drastic effects of PPM1G KD in antagonizing induction of TNF responsive genes supports that model that PPM1G is a critical regulator of the NF-κB pathway.

Figures and Figure Legends

Figure 4

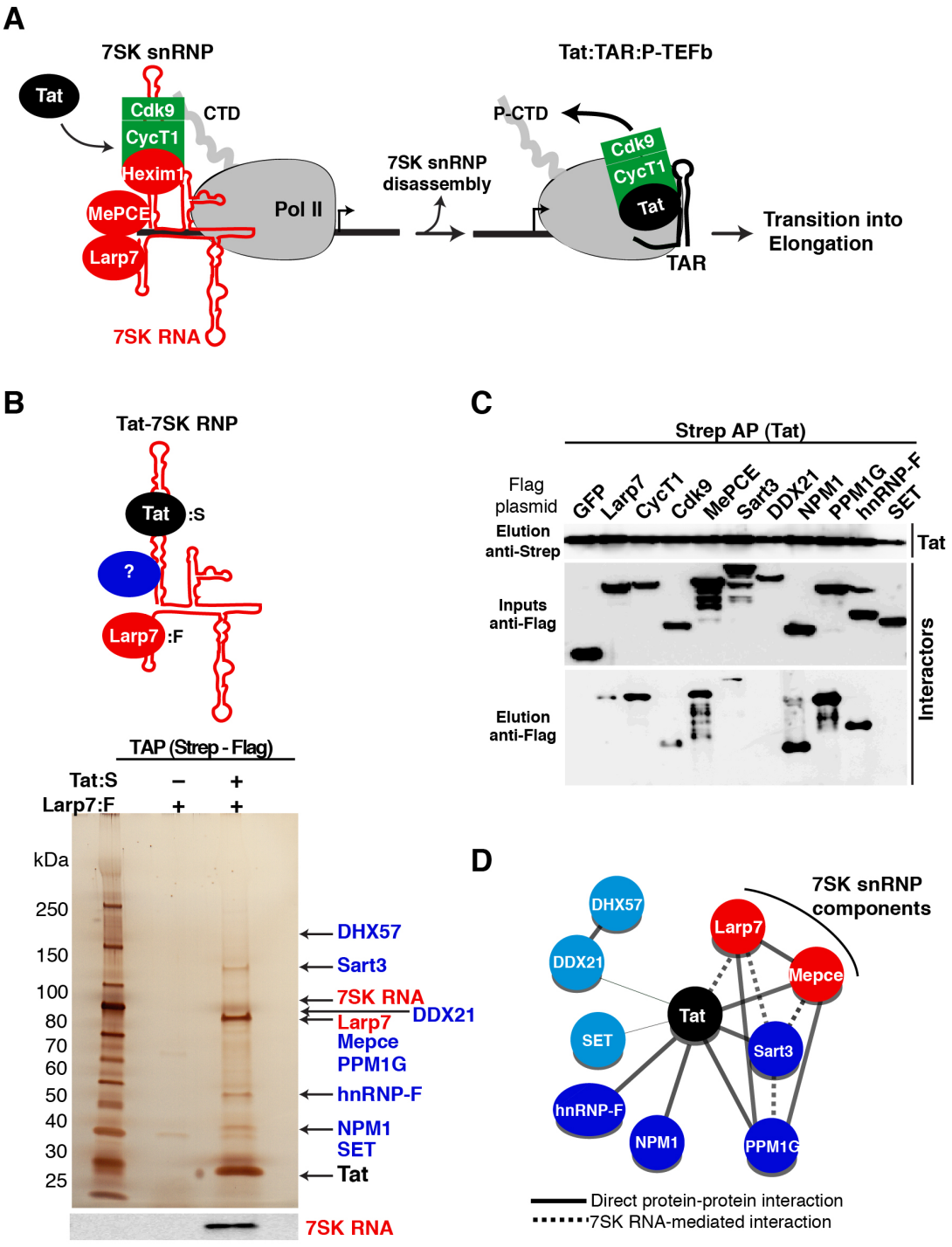

Figure 4. Identification of an HIV Host Tat-7SK Protein-RNA Complex Implicated in Transcriptional Control

(A) Model of HIV transcription activation by Tat. Tat promotes the disassembly of the promoter-bound 7SKsnRNP to transfer P-TEFb (CycT1:CDK9) to the nascent viral RNA (TAR), allowing for Pol II CTD hyper-phosphorylation (P-CTD) and the transition into elongation.

(B) Tandem affinity purification (TAP) of the Tat-7SK ribonucleoprotein (RNP) complex by sequential Strep-tagged Tat (Tat:S) and FLAG-tagged Larp7 (Larp7:FLAG) purifications from HEK 293T cells. High-confidence Tat-7SK RNP Interactors were identified by mass spectrometry analysis (see also **Table 1**) and indicated by arrows on a silver-stained gel. The bottom panel shows a 7SK RNA northern blot.

(C) Validation of the interactors identified. Tat:S was co-expressed into HEK 293T cells along with the identified FLAG-tagged factors. Strep AP followed by Western blot reveals proteins co-immunoprecipitating with Tat.

(D) A network representation of the protein-protein and protein-RNA interactions within the Tat-7SK RNP complex. The thickness of the edges corresponds to the semi-quantitative score of the interactions among nodes. Solid and dashed edges correspond to direct or RNA-mediated interactions. Light-blue nodes represent interactors identified by mass spectrometry, but not validated by protein-protein interaction assays.

Figure 5

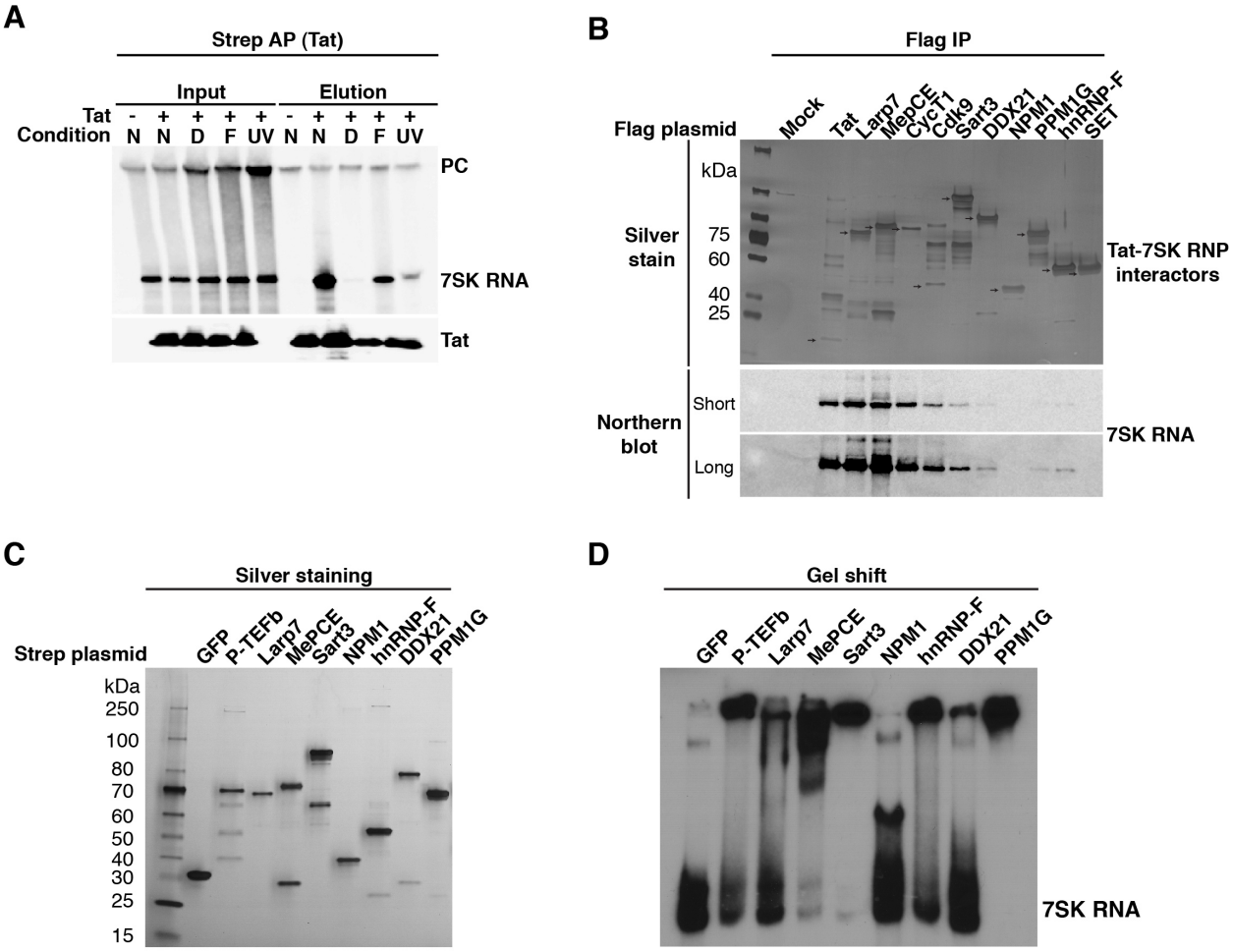

Figure 5. Interactions between Tat and the 7SK RNP interactors with 7SK RNA *in vitro* and in cells

(A) Tat was affinity purified from HEK 293T cells under different conditions: using a Native (N) buffer or from cells previously chemically crosslinked with formaldehyde (F), irradiated with UV (UV) or left untreated using a Denaturing (D) buffer. (D) is a negative control for F and UV samples. PC denotes an RNA used as an internal precipitation control used in the affinity purification (AP) step to control for RNA degradation. Tat was detected with a Strep antibody.

(B) HEK 293T cells were transfected with the indicated FLAG plasmids. Cells were lysed in native conditions and total lysates used to purify the FLAG-tagged factors (FLAG IP). The

eluted samples were directly electrophoresed on SDS-PAGE and silver stained or used to extract co-purifying 3 RNAs to monitor levels of 7SK RNA by Northern blot. Short and long denote two exposure time points.

(C) HEK 293T cells were transfected with the indicated Strep plasmids and proteins were purified in the presence 1 μg/mL of RNase A (to eliminate any co-purifying RNA). A fraction of the elutions was visualized by silver staining.

(D) Gel shift assays between *in vitro* synthesized 7SK RNA and the interactors shown in panel (C). P-TEFb binds 7SK RNA in the presence of co-purifying Hexim1 protein.

Figure 6

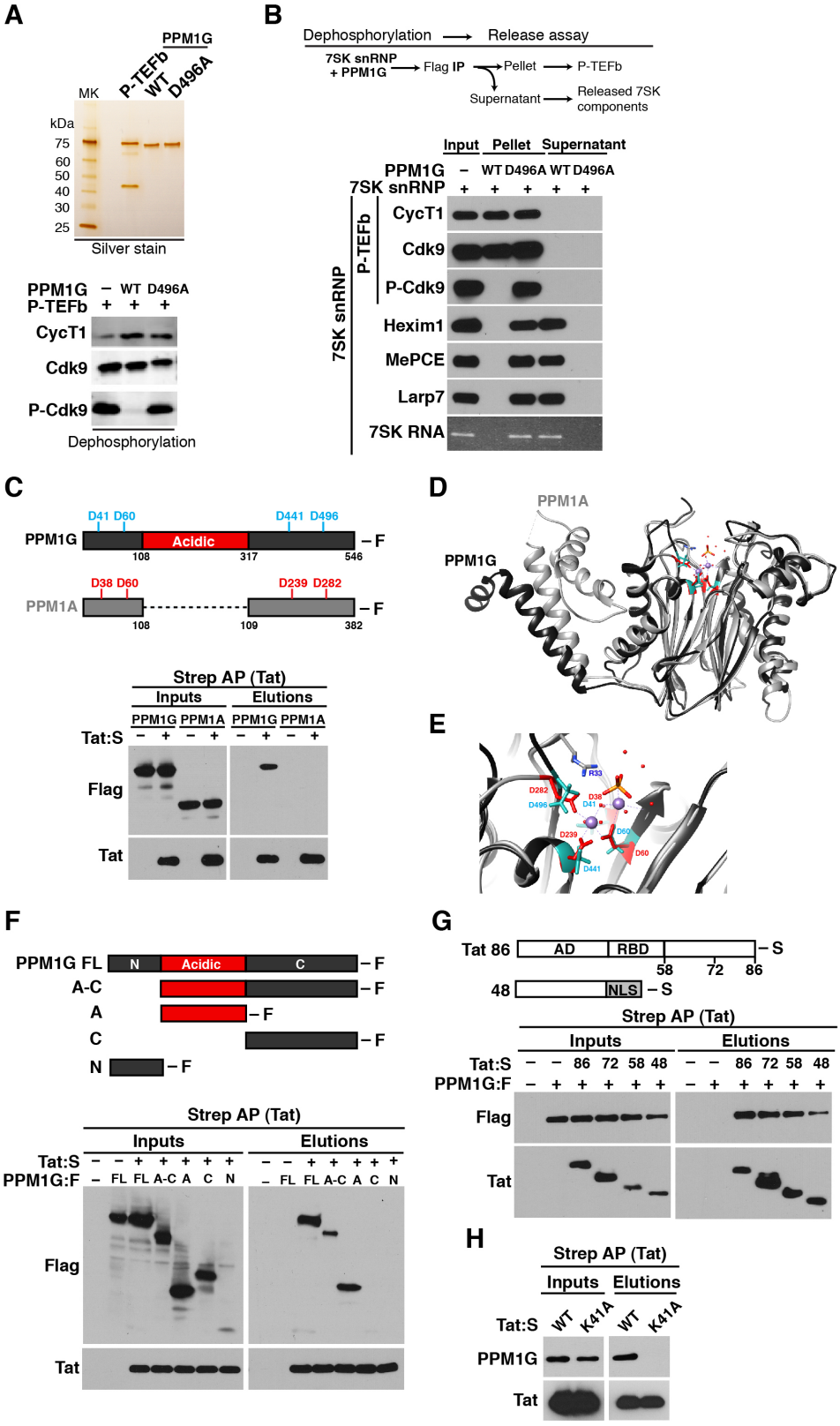
A
PPM1G
P-TEFb
WT
D496A
MK
kDa
75
60
50
40
30
25
Silver stain
PPM1G – WT D496A
P-TEFb + + +
CycT1
Cdk9
P-Cdk9
Dephosphorylation
B
Dephosphorylation → Release assay
7SK snRNP + PPM1G → Flag IP → Pellet → P-TEFb
Supernatant → Released 7SK components
Input Pellet Supernatant
PPM1G – WT D496A WT D496A
7SK snRNP + + + + +
7SK snRNP
P-TEFb
CycT1
Cdk9
P-Cdk9
Hexim1
MePCE
Larp7
7SK RNA
C
D41 D60
D441 D496
PPM1G
Acidic
108
317
546
– F
D38 D60
D239 D282
PPM1A
108
109
382
Strep AP (Tat)
Inputs
Elutions
PPM1G PPM1A PPM1G PPM1A
Tat:S – + – + – + – +
Flag
Tat
D
PPM1A
PPM1G
E
R33
D282
D496
D38
D41
D60
D239
D441
F
PPM1G FL
N
Acidic
C
A-C
A
C
N
– F
Strep AP (Tat)
Inputs
Elutions
Tat:S – – + + + + + – – + + + + +
PPM1G:F – FL FL A-C A C N – FL FL A-C A C N
Flag
Tat
G
Tat 86
AD
RBD
58
72
86
– S
48
NLS
Strep AP (Tat)
Inputs
Elutions
Tat:S – – 86 72 58 48 – – 86 72 58 48
PPM1G:F – + + + + + – + + + + +
Flag
Tat
H
Strep AP (Tat)
Inputs
Elutions
Tat:S WT K41A WT K41A
PPM1G
Tat

Figure 6. Tat Selectively Recruits the PPM1G Phosphatase to Disassemble the 7SK snRNP and Release P-TEFb through CDK9 T Loop Dephosphorylation

(A) Top, P-TEFb and wild-type PPM1G or catalytically inactive mutant (D496A) were purified and visualized by silver stain. Bottom, P-TEFb was incubated with PPM1G or D496A under dephosphorylation conditions, and CDK9 T loop phosphorylation at Thr186 (P-CDK9) was monitored by Western blot. WT, wild-type.

(B) Enzymatic release assay of P-TEFb from the 7SK snRNP. 7SK-bound P-TEFb (CycT1:S-CDK9:FLAG) complexes were purified from mammalian cells and incubated with PPM1G or D496A under dephosphorylation conditions. Subsequent purification of P-TEFb using FLAG beads was done to monitor released (supernatant) and retained (pellet) components by Western blot.

(C) Tat selectively binds PPM1G, but not PPM1A. Top, PPM1G possesses an acidic region intervening the N- and C-terminal phosphatase domain. The position of the four Asp (D) residues required for metal coordination and catalysis are shown above the schemes. Bottom, Strep-tagged Tat and FLAG-tagged PPM1G or PPM1A were co-transfected into HEK 293T cells, affinity purified using Strep beads, and analyzed by western blot. CA, capsid; MA, matrix.

(D) Superimposition of the PPM1A structure (Das et al., 1996) and structure prediction of PPM1G lacking its acidic region.

(E) Close-up view of the PPM1A-PPM1G superimposition showing the catalytic core in the PPM1A structure and structure prediction of PPM1G (light blue). The two Mn+ ions (violet) and the phosphate group (orange) are shown. In PPM1A, Arg33 (R33) side chains form H bonds with the phosphate ion (orange).

(F) Tat binds to the PPM1G acidic domain. Strep-tagged Tat and FLAG-tagged PPM1G (full-length [FL] or domains) were co-transfected into HEK 293T cells, affinity purified using Strep beads, and analyzed by Western blot.

(G) Tat domain mapping. Strep-tagged Tat full-length (Tat86) or domains (AD, activation domain; RBD,RNA-binding domain; NLS, SV40 T-Ag nuclear localization signal) and FLAG-tagged PPM1G were co-transfected into HEK 293T cells, affinity purified using Strep beads, and analyzed by Western blot.

(H) Strep-tagged wild-type (WT) Tat or a nonfunctional K41A mutant and FLAG-tagged PPM1G were co-transfected into HEK 293T cells, affinity purified using Strep beads, and analyzed by Western blot.

Figure 7

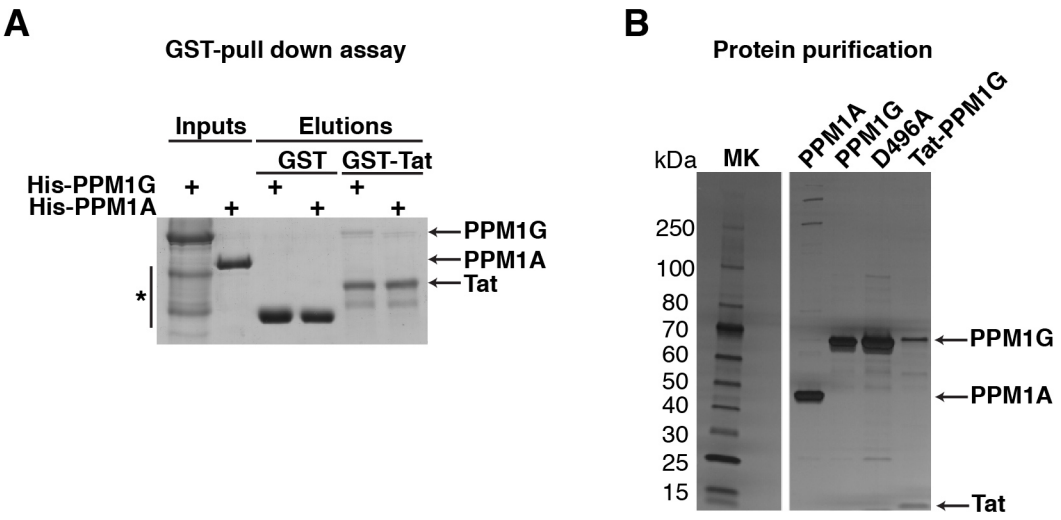

Figure 7. Tat directly binds PPM1G but not PPM1A.

(A) GST-pull down assay of GST or GST-Tat and His-tagged PPM1G or PPM1A. The indicated proteins were incubated *in vitro* and elutions were stained with coomassie blue. The asterisk denotes the presence of PPM1G cleavage products.

(B) HEK 293T cells were transfected with the indicated plasmids (Strep-PPM1A, Strep-PPM1G, Strep-PPM1G D496A mutant or co-transfected with Strep-Tat and FLAG-PPM1G). Cell lysates were used to purify PPM1A, PPM1G and D496A using a single Strep purification step or the Tat-PPM1G complex using a Strep-FLAG tandem affinity purification protocol. A fraction of the elutions was visualized by silver staining. Proteins are indicated with arrows.

Figure 8

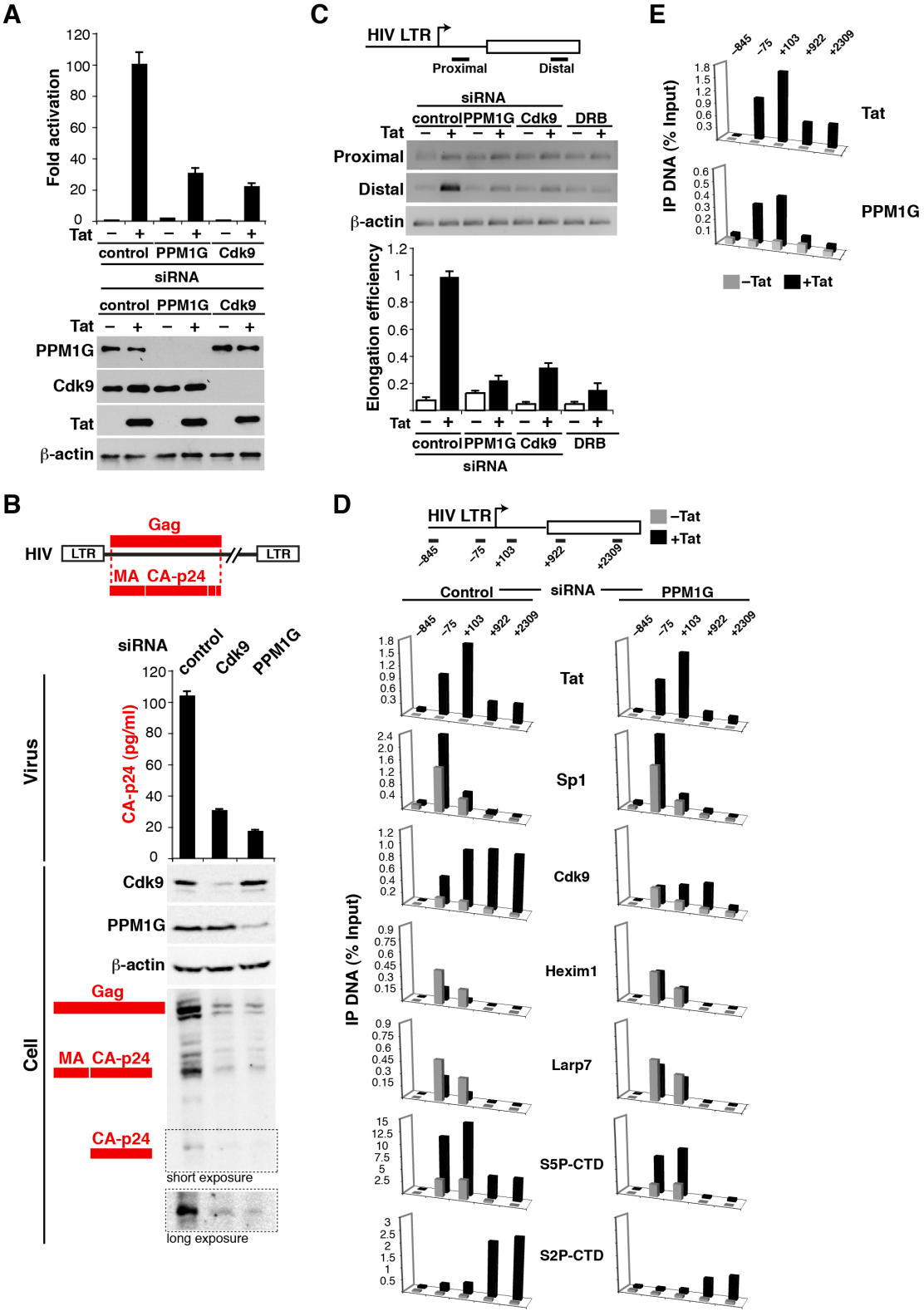
A
Fold activation
Tat
control PPM1G Cdk9
siRNA
PPM1G
Cdk9
Tat
β-actin
B
HIV LTR Gag MA CA-p24 LTR
siRNA control Cdk9 PPM1G
CA-p24 (pg/ml)
Virus
Cell
Cdk9
PPM1G
β-actin
Gag
MA CA-p24
CA-p24
short exposure
long exposure
C
HIV LTR
Proximal
Distal
siRNA
control PPM1G Cdk9 DRB
Tat
Proximal
Distal
β-actin
Elongation efficiency
D
HIV LTR
−Tat
+Tat
−845 −75 +103 +922 +2309
Control
siRNA
PPM1G
IP DNA (% Input)
Tat
Sp1
Cdk9
Hexim1
Larp7
S5P-CTD
S2P-CTD
E
IP DNA (% Input)
Tat
PPM1G
−Tat
+Tat

Figure 8. Recruitment of PPM1G by Tat Mediates Disassembly of the Promoter-Bound 7SK snRNP and Activates the Transition into Elongation

(A) KD of PPM1G antagonizes Tat-mediated HIV transcription activation. Top, HeLa cells were co-transfected with an HIV LTR-Firefly luciferase and FLAG-tagged Tat plasmid after siRNA KD. HIV luciferase activity was normalized to CMV Renilla luciferase (mean ± SEM are shown; n = 3). Bottom, HeLa samples were used for Western blots with the indicated antibodies; β-actin was used as an internal loading control.

(B) Production of viral particles from cells. HEK 293T cells were sequentially transfected with three different siRNAs and a full-length HIV proviral DNA (NL4-3). Only Gag is shown in the scheme. Viral particles released to the supernatant were quantified by p24 ELISA (mean ± SEM are shown; n = 4), and the levels of intracellular CDK9, PPM1G, β-actin, and HIV Gag were detected by Western blot.

(C) Transcription elongation assay with the schematic of RT-PCR products. The gel shows the products resulting from transfection of a HeLa LTR-FFL cell line with three different siRNAs or pretreated with DRB in the presence of an empty (−) or Tat plasmid (+). The graph indicates the calculated elongation efficiencies expressed in arbitrary units (a.u.) standardized to β-actin (mean ± SEM are shown; n = 3).

(D) ChIP assay to analyze the distribution of Tat and cofactors at the HIV locus in a HeLa LTR-FFL cell line transfected with a control or PPM1G siRNA along with a mock (gray bars) or FLAG-tagged Tat (black bars) plasmid. The position of the amplicons used in ChIP-qPCR is shown with the schematic of the HIV locus. Values represent the average of three independent experiments. The SEM is less than 10% and not shown for simplicity.

(E) ChIP assay to analyze the distribution of Tat and PPM1G at the HIV locus in a HeLa LTR-FFL cell line transfected with a mock (gray bars) or Tat (black bars) plasmid. Values represent the average of three independent experiments. The SEM is less than 10% and not shown for simplicity.

Figure 9

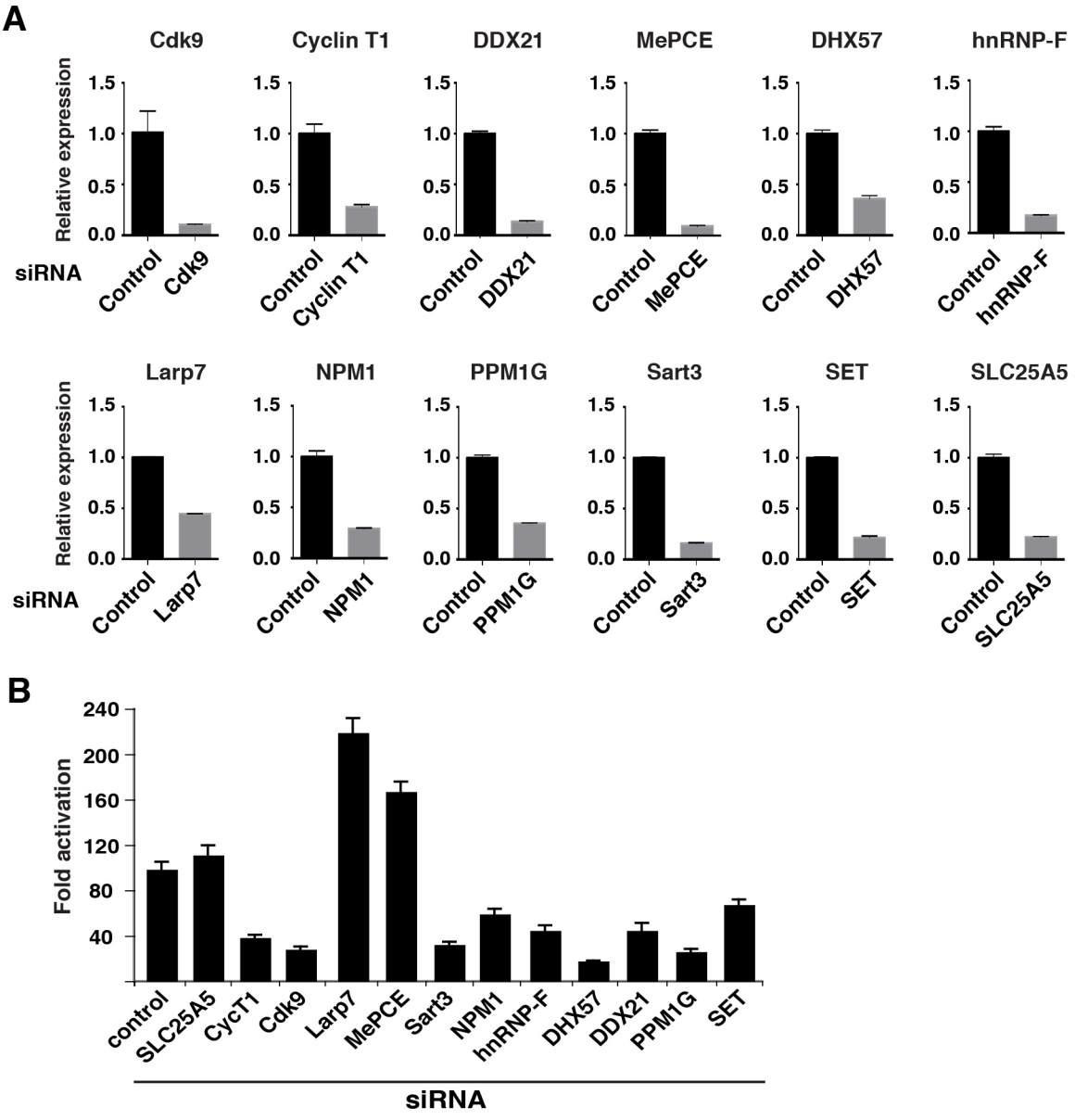

Figure 9. RNAi KD of the Tat-7SK RNP interactors decreases Tat activity from an HIV LTR reporter.

(A) HeLa cells were transfected with the indicated siRNAs. Levels of the indicated mRNAs were quantified by qRT-PCR using gene-specific primers. Relative expression denotes relative gene expression levels normalized to the Rpl19 gene (Mean ± SEM, n=3).

(B) HeLa cells transfected with siRNAs as in panel (A) were subsequently transfected with a FLAG-tagged Tat and an HIV LTR-luciferase plasmids (24 hours post-siRNA transfection) to measure transcription reporter activity. The fold activation (+/– Tat) normalized to a CMV-RL activity is shown (Mean ± SEM, n=3).

Figure 10

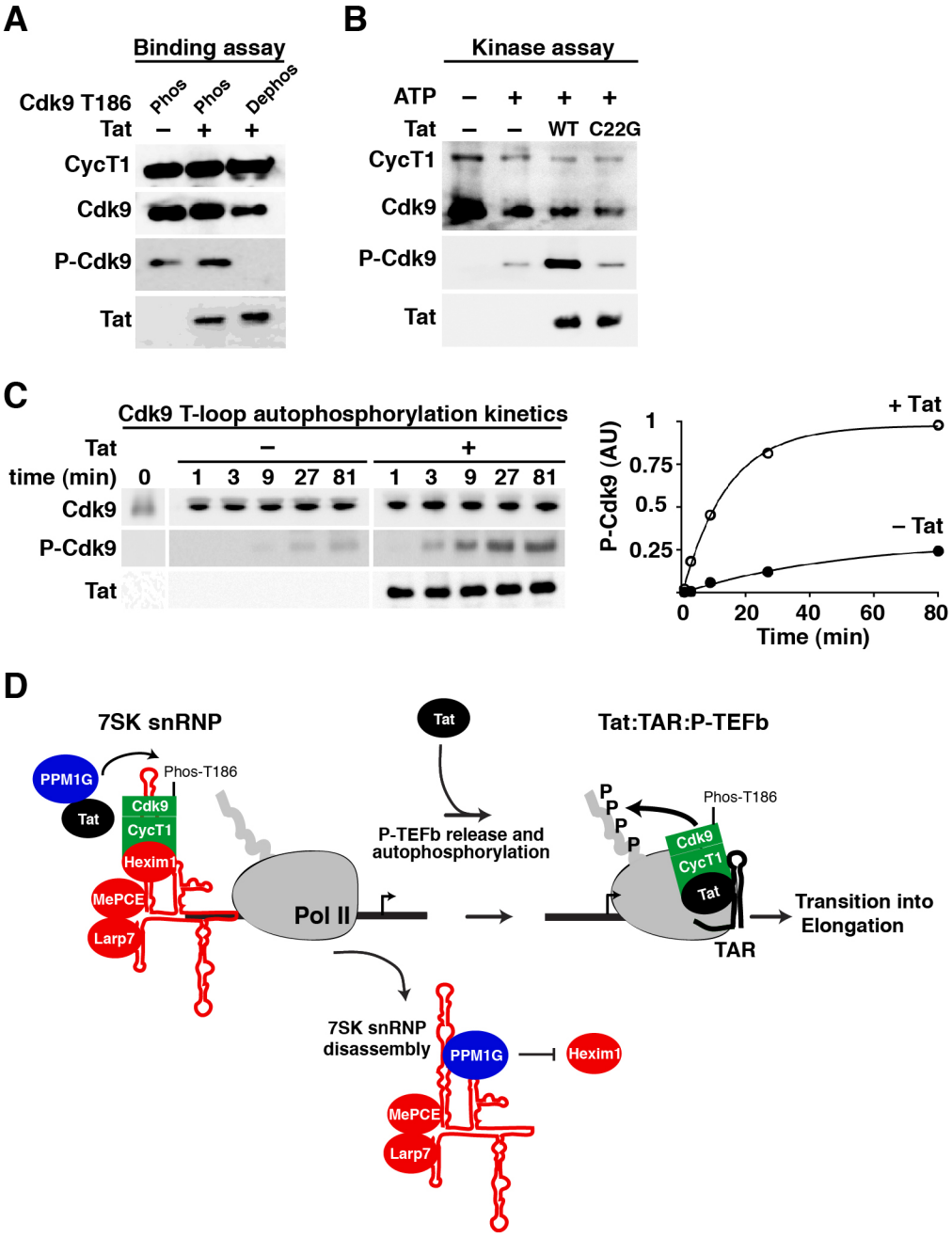

Figure 10. Tat Binds Dephosphorylated P-TEFb and Stimulates CDK9 T Loop Phosphorylation to Activate the Transition into Elongation

(A) P-TEFb containing a phosphorylated (Phos) or dephosphorylated (Dephos) CDK9 T loop (T186) was incubated with Tat. CDK9 was subsequently purified using FLAG beads to monitor for Tat binding by Western blot.

(B) Dephosphorylated P-TEFb was incubated with or without ATP or with ATP along with wild-type (WT) Tat or a nonfunctional C22G mutant in kinase conditions and analyzed by Western blot.

(C) Kinetics of CDK9 T loop auto-phosphorylation. Left, dephosphorylated P-TEFb was incubated with or without Tat in kinase conditions. Fractions were removed and assayed for CDK9 T loop phosphorylation (P-CDK9). Right, levels of P-CDK9 standardized to total CDK9 were expressed in arbitrary units (a.u.).

(D) Model of Tat-mediated enzymatic release of P-TEFb from the 7SK snRNP. Tat recruits PPM1G to the HIV promoter-bound 7SK snRNP. PPM1G dephosphorylates the T loop of CDK9 (Phos-T186) to disassemble the 7SK snRNP and release P-TEFb. Released P-TEFb is captured by Tat, which stimulates CDK9 T loop auto-phosphorylation to form transcription elongation competent complexes on TAR.

Figure 11

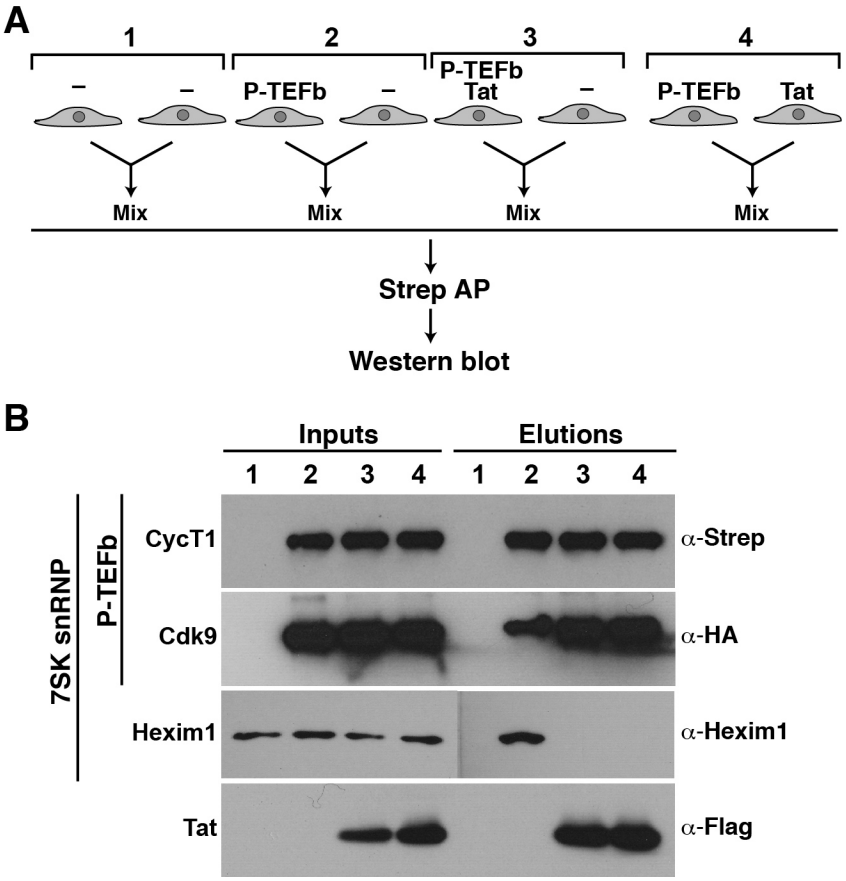

Figure 11. ***In vitro*** **association of P-TEFb and Tat leads to Hexim1 physical competition from the 7SK snRNP.**

(A) Schematic diagram of the population of HEK 293T cells examined in the mixing experiments. Cells were transfected with the indicated plasmids 48 hours before harvesting. P-TEFb is CycT1:S and CDK9:HA; Tat is FLAG-tagged; (–) denotes cells transfected with an empty pcDNA4 vector.

(B) The cells described in panel (A) were mixed as indicated before the cell lysis. Protein lysates were subjected to a Strep AP for 2 hours at 4°C and the eluted material was probed in Western blots with the indicated antibodies.

Figure 12

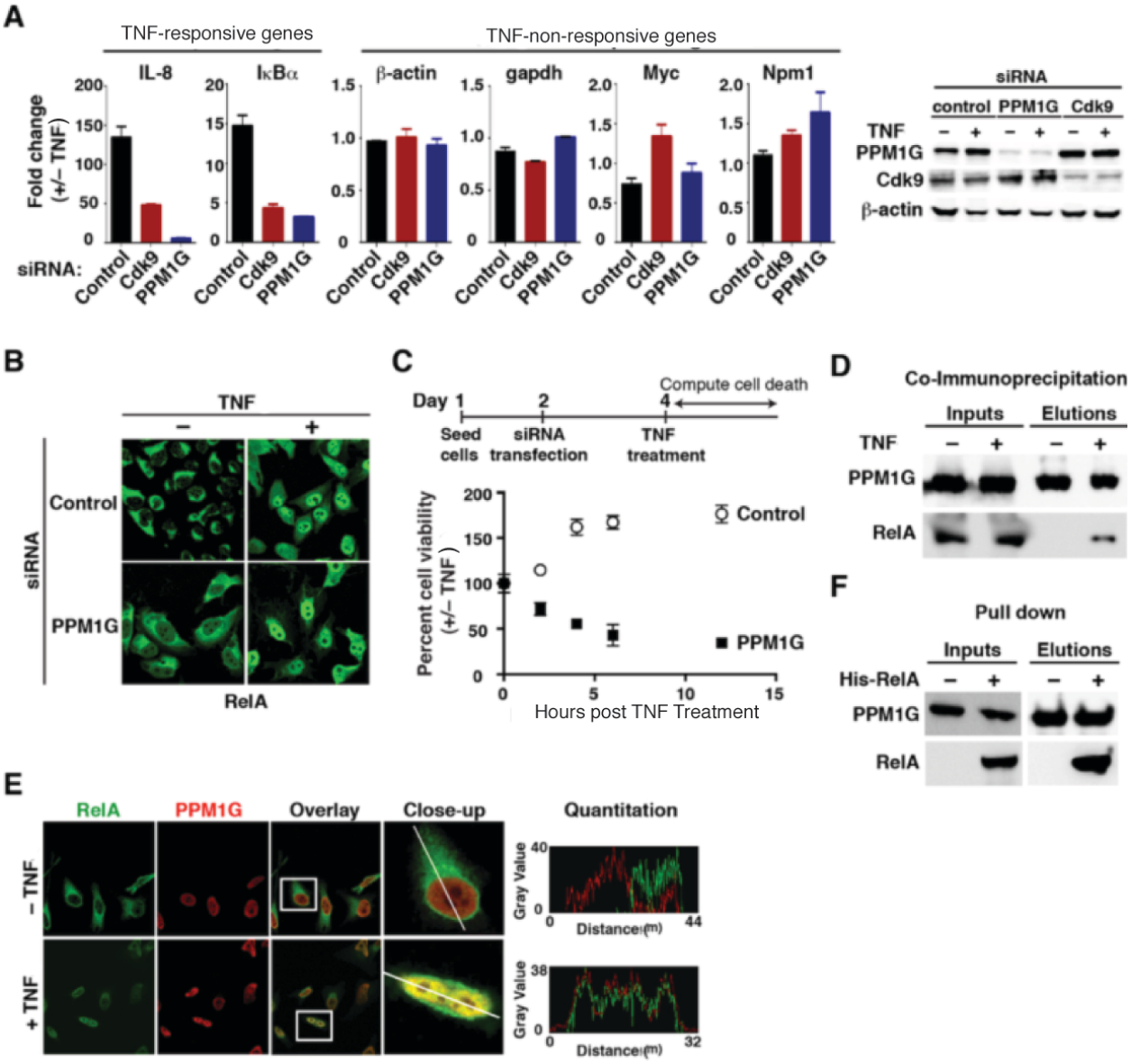

Figure 12. PPM1G Is a Nuclear Transcriptional Co-activator of NF-κB in the Inflammatory Pathway

(A) HeLa cells were transfected with the indicated siRNAs and treated with or without TNF. Left, the expression of TNF-responsive and -nonresponsive genes (normalized to Rpl19) was

measured by qRT-PCR and plotted as fold change ± TNF (mean ± SEM are shown; n = 3). Right, Western blots showing the validation of the KD.

(B) KD of PPM1G in HeLa cells does not alter NF-κB translocation to the nucleus upon TNF treatment.

(C) PPM1G KD sensitizes cells to TNF-induced apoptosis. HeLa cells were transfected with the indicated siRNAs. At 48 hours post-transfection, cells were treated with or without TNF and incubated for the indicated time points. The plot shows percentage of viable cells and is expressed as fold difference ± TNF (mean ± SEM are shown; n = 3).

(D) PPM1G and NF-κB interact in cells. HEK 293T cells were transfected with a Strep-tagged PPM1G plasmid and cells were treated with or without TNF, crosslinked with formaldehyde, and lysed under denaturing conditions. Strep AP was performed to analyze the interaction between PPM1G and NF-κB (RelA subunit).

(E) NF-κB and PPM1G co-localize in a stimulus-dependent manner in the cell nuclei. HeLa cells were treated with or without TNF, and confocal images are shown. A single cross-sectional image was acquired, and emission intensities were quantified to analyze the relative co-localization along the path indicated (histograms).

(F) NF-κB and PPM1G directly bind. Bacterially synthesized NF-κB was incubated with Strep-tagged PPM1G, and the final elution was analyzed by Western blot.

Figure 13

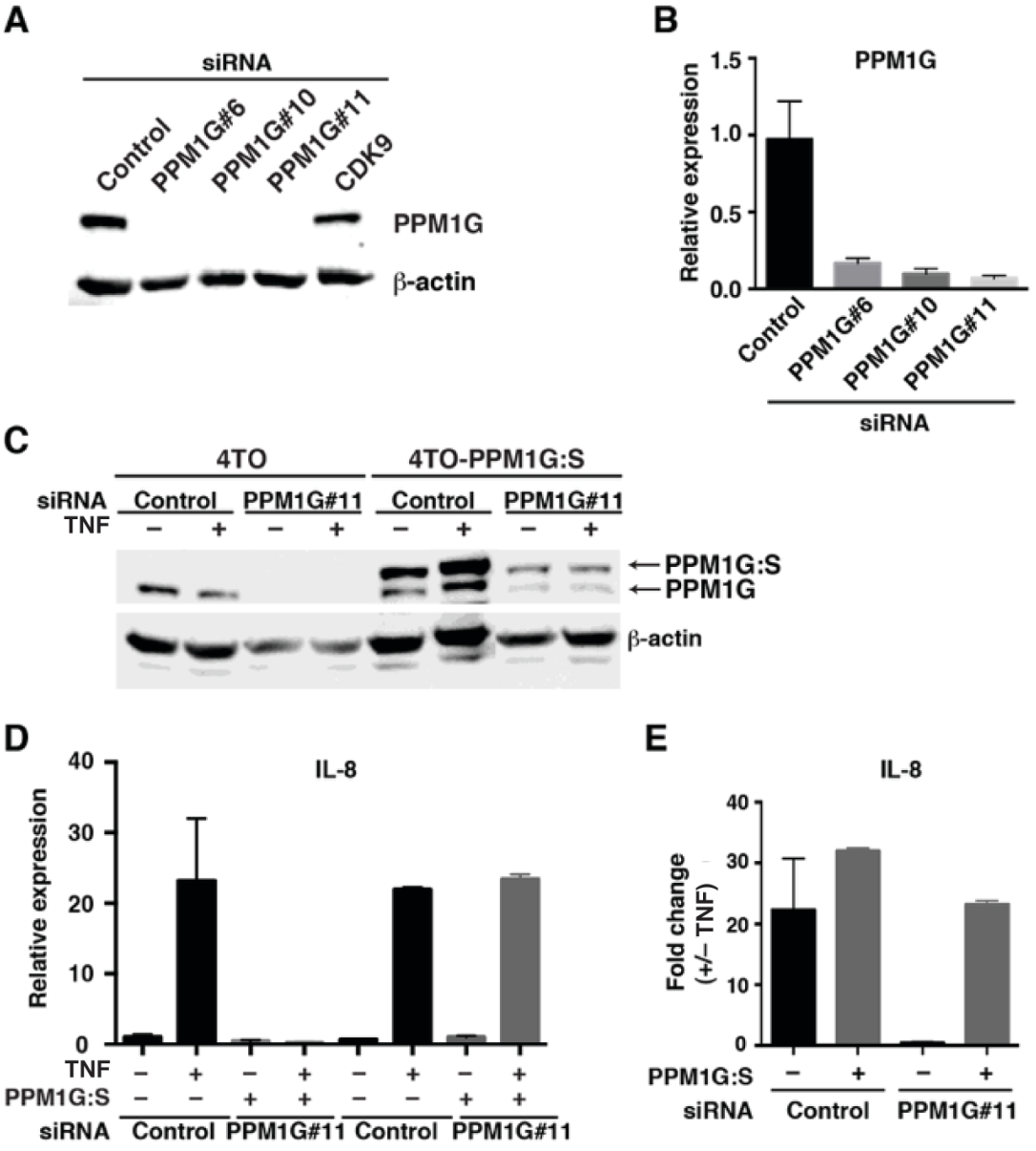

Figure 13. PPM1G RNAi with different siRNA duplexes and rescue experiment.

(A) HeLa cells were transfected with the indicated siRNAs and the expression of PPM1G and β-actin (loading control) was analyzed by Western blot.

(B) The relative expression of PPM1G (normalized to β-actin) in the samples from panel (A) was determined by qRT-PCR (Mean ± SEM, n=3).

(C) siRNA rescue experiment. HeLa cells were transfected with the indicated plasmids (4TO or 4TO-PPM1G:S) and one day later re-transfected with the indicated siRNAs. The expression of endogenous and transfected PPM1G and β-actin was analyzed by Western blot

(D) A fraction of the samples transfected as in panel (C) was used to extract RNA and prepare cDNA. The relative expression of IL-8 (normalized to β-actin) was determined by qRT-PCR (Mean ± SEM, n=3)

(E) IL-8 expression levels from panel (D) expressed as fold change (± TNF, Mean ± SEM, n=3)

Figure 14

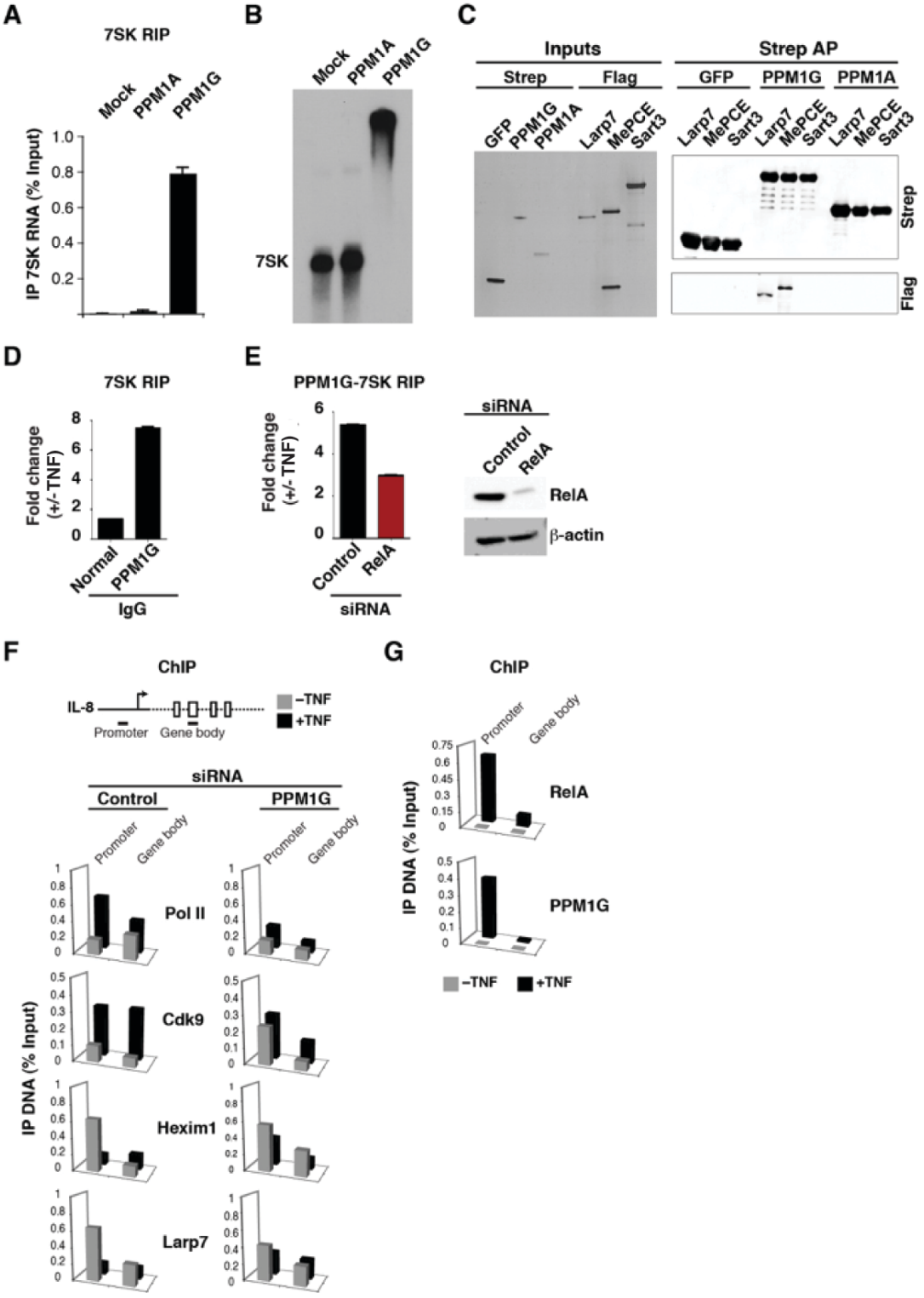
A
7SK RIP
Mock
PPM1A
PPM1G
IP 7SK RNA (% Input)
1.0
0.8
0.6
0.4
0.2
B
Mock
PPM1A
PPM1G
7SK
C
Inputs
Strep AP
Strep
Flag
GFP
PPM1G
PPM1A
Larp7
MePCE
Sart3
Strep
Flag
D
7SK RIP
Fold change (+/- TNF)
8
6
4
2
0
Normal
PPM1G
IgG
E
PPM1G-7SK RIP
Fold change (+/- TNF)
6
4
2
0
Control
RelA
siRNA
siRNA
Control
RelA
RelA
β-actin
F
ChIP
IL-8
Promoter
Gene body
–TNF
+TNF
siRNA
Control
PPM1G
Promoter
Gene body
Pol II
Cdk9
Hexim1
Larp7
IP DNA (% Input)
G
ChIP
Promoter
Gene body
RelA
PPM1G
IP DNA (% Input)
–TNF
+TNF

Figure 14. PPM1G Binds the 7SK snRNP and Participates in the NF-κB-Mediated Disassembly of the Promoter-Assembled snRNP at Inflammatory Responsive Genes

(A) HeLa cells were transfected with an empty vector (Mock) or Strep-tagged PPM1A or PPM1G plasmids, and levels of 7SK RNA co-immunoprecipitating were quantified by qRT-PCR (mean ± SEM are shown; n = 3).

(B) PPM1G binds 7SK RNA. Gel shift assay between *in vitro*-synthesized 7SK RNA and His-tagged, bacterial PPM1A or PPM1G. Mock denotes incubation of 7SK RNA in binding buffer with a purification from bacterial cells transformed with an empty vector.

(C) PPM1G binds 7SK snRNP core components. Left, a silver-stained gel of input proteins purified from HEK 293T in the presence of RNase A. The FLAG-tagged proteins were incubated with GFP, PPM1G, or PPM1A bound to Strep beads. Right, Western blots of the Strep AP.

(D) Stimulus-dependent recruitment of PPM1G to the 7SK RNA. HeLa cells were treated with or without TNF, endogenous PPM1G was immunoprecipitated using a specific serum, and levels of PPM1G-bound 7SK RNA were quantified by qRT-PCR. Normal immunoglobulin G (IgG) was used as negative control (mean ± SEM are shown; n = 4).

(E) PPM1G-7SK RIP. Left, HeLa cells were transfected with control or NF-κB siRNAs and treated with or without TNF, and a RIP assay was performed to quantify the amount of endogenous PPM1G-bound 7SK RNA by qRT-PCR (mean ± SEM are shown; n = 3). Right, Western blots to validate the KD.

(F) ChIP assay to analyze the distribution of Pol II and P-TEFb-7SK snRNP subunits at the IL-8 locus. HeLa cells were transfected with a control or PPM1G siRNA and treated with (black bars) or without (gray bars) TNF. Values represent the average of three independent experiments. The SEM is less than 5% and not shown for simplicity.

(G) ChIP assay to analyze the distribution of NF-κB and PPM1G at the IL-8 locus in HeLa cells treated with (black bars) or without (gray bars) TNF. Values represent the average of four independent experiments. The SEM is less than 5% and not shown for simplicity.

Figure 15

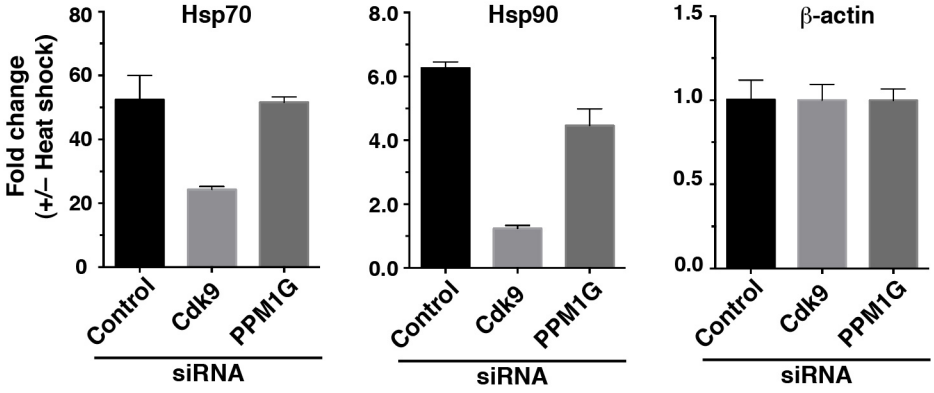

Figure 15. qRT-PCR analysis of gene expression in response to heat shock.

HeLa cells were exposed (or not) to heat shock stress (42°C) and harvested 2 hours post-treatment. Total RNA was isolated, and qRT-PCR was performed with gene-specific primers. Values were normalized to Rpl19 and are expressed as fold change over untreated cells (Mean ± SEM, n=3).

Table 1. List of high confidence interactors of the Tat – 7SK snRNP

Protein name	Accession number	Molecular Iight (kDa)
NPM1	P06748	32
SET	Q01105	33
PPM1G/PP2Cγ	O15355	59
hnRNP-F	P52597	46
MePCE/BCDIN3	Q7L2J0	74
DDX21	Q9NR30	87
Sart3/Tip110	Q15020	110
DHX57	Q6P158	155

CHAPTER IV – RESULTS

PART II

The 7SK snRNP Complex Interacts with the Transcriptional Regulator KAP1/TRIM28/TIF1β

Previous studies have shown that the inhibitory 7SK snRNP complex is recruited to gene promoters (**Figure 16A**) (Cherrier et al., 2013; D'Orso and Frankel, 2010a; D'Orso et al., 2012; Ji et al., 2013). However, the molecular basis of this recruitment mechanism and its precise role in the transcriptional cycle has yet to be understood. To identify factors that interact with the 7SK snRNP to potentially recruit it to gene promoters, we employed an affinity purification (AP) and mass spectrometry (Jadlowsky et al. 2014) approach, which has proven to be an efficient unbiased and comprehensive methodology to discover the nature and composition of protein complexes (Jager et al., 2012; Jager et al., 2011). Strep AP of the 7SK snRNP component Larp7 (Larp7:S) from nuclear fractions, but not GFP:S, revealed the existence of several specific interactors not present in control samples lacking Larp7 (**Figure 16B**). To reveal the identity of these interactors, we subjected the AP of Larp7 and control samples (GFP) for LC-MS/MS analysis. High-confidence interactors (based on their quantitative enrichment relative to control samples) are shown on **Table 2** along with their corresponding hypothetical bands in the silver stained gel of **Figure 16B**. In addition to the core 7SK snRNP components (Larp7, CycT1, CDK9, MePCE, Hexim1), reversibly-bound hnRNP proteins, PPM1G, and DDX21

(Barrandon et al., 2007; Calo et al., 2015; Hogg and Collins, 2007; Jeronimo et al., 2007; Krueger et al., 2008), we also identified one novel high-confidence interactor, the KAP1 protein (also known as TRIM28/TIF1β) (Iyengar and Farnham, 2011a; Iyengar et al., 2011b) (**Figures 16** and **C**). In contrast to this AP-MS dataset, KAP1 has not been identified in a previous Larp7 proteomics study (Krueger et al., 2008). The apparent discrepancy between these two datasets could be attributed to the higher affinity of the Strep AP (compared with regular antibody-mediated immunoprecipitation (IP) methods) thereby leading to more sensitivity.

To validate that KAP1 interacts with Larp7 as part of the 7SK snRNP complex, we generated HEK 293 T-REx cell lines inducibly expressing FLAG-tagged KAP1 (KAP1:FLAG), and GFP:FLAG (negative control) after DOX treatment. Indeed, using nuclear extracts we detected the primary components of the 7SK snRNP complex after FLAG IP of KAP1, but not GFP, including 7SK RNA but not another abundant snRNA such as U6 (**Figure 16D**). Because we detected the KAP1-Larp7 protein-protein interaction using epitope-tagged overexpressed components, we next performed an IP of endogenous Larp7, (which is constitutively bound to the 7SK snRNP (Krueger et al., 2008)), to demonstrate that the association occurs in a more biologically relevant system. Consistent with our previous results, we detected KAP1 in the IP of endogenous Larp7, as well as core 7SK snRNP components, but not when using a normal serum (IgG) (**Figure 16E**). Based on the material loaded on the Western blot we estimated that about 10-15% of the 7SK snRNP interacts with KAP1 in the conditions the cell lysate was generated. Importantly, while the KAP1-Larp7 interaction was largely insensitive to RNase, as expected for a direct protein-protein interaction, the KAP1-CDK9 interaction was RNA-dependent (**Figure 16F**), strongly indicating that KAP1 recruits the P-TEFb kinase through the Larp7 subunit of the 7SK snRNP complex.

To further validate the relevance of our findings (that KAP1 and the 7SK snRNP form part of a complex in cells) we performed high-resolution confocal microscopy of endogenous KAP1 and components of the 7SK snRNP (CDK9, Hexim1 and Larp7) and observed statistically significant co-localization indexes (p-value < 0.01) for all protein pair combinations (**Figure 16G**).

To further test whether KAP1 associates with the 7SK snRNP complex through direct interaction with the Larp7 subunit, we performed an *in vitro* binding assay using affinity purified proteins from HEK 293T cells under high-salt conditions, and observed that Larp7 directly binds to KAP1 *in vitro*, as well as MePCE, which is consistent with a previously characterized interaction between Larp7 and MePCE in the absence of 7SK RNA (Xue et al., 2010), but not to GFP (used as negative control) (**Figure 16H**). Although in the conditions used, the interaction between KAP1 and Larp7 appear not to be stoichiometric (as well as the control Larp7-MePCE interaction), we cannot rule out that Larp7 (in the absence of 7SK RNA) is biochemically poorly behaved or that the protein-protein interaction requires post-translational modifications to achieve higher binding affinity.

To validate that the cell line we had generated that inducibly expresses KAP1:FLAG does not substantially overexpress the protein, we performed a quantitative western blot after DOX treatment. While we would not separate KAP1:FLAG from endogenous KAP1, we observed that the total amount of KAP1 relative to –DOX treatment was ~2.7 fold higher (**Figure 17A**). We therefore concluded that the ratio of KAP1:FLAG to endogenous KAP1 were similar and use the KAP1:FLAG cell line to characterize how KAP1 interacts with the 7SK snRNP was appropriate. In further support of this, we took additional confocal images of KAP1 and components of the 7SK snRNP and calculated their co-localization indices. As expected,

KAP1 significantly co-localized with components of the 7SK snRNP (**Figures 17B-D**), but not with β-actin, which served as a negative control (**Figure 17E**).

Taken together, we have demonstrated through various approaches that KAP1 interacts with the 7SK snRNP. Given the relevance of P-TEFb in the transcriptional cycle, we next focused on studying the function of KAP1 in recruiting the 7SK snRNP to gene promoters to control the transition from transcription initiation into elongation.

KAP1 and the 7SK snRNP Complex Co-occupy the HIV Promoter-proximal Region

The 7SK snRNP complex occupies promoter-proximal regions of both HIV and cellular genes (Cherrier et al., 2013; D'Orso and Frankel, 2010a; Eilebrecht et al., 2014; Ji et al., 2013). Given that KAP1 physically associates with the 7SK snRNP and that KAP1 is also recruited to gene promoters (Hu et al, 2009; Bunch et al., 2014), we asked whether both are co-recruited to the promoter of HIV integrated into a Jurkat CD4+ T cell line known as clone E4 (Pearson et al., 2008). Using chromatin immunoprecipitation (ChIP) assays, we detected both KAP1 and the 7SK snRNP (CDK9, Larp7 and Hexim1) at the HIV promoter-proximal region surrounding the TSS (5'-LTR), but not inside the genome (at least with the amplicons used in qPCR assays) (**Figure 18A**).

The presence of duplicated LTRs makes it difficult to assign ChIP signal to individual LTRs. Therefore, to distinguish PCR amplifications of the 5'-LTR from the 3'-LTR, we used LTR-specific primer sets (+141 and +4230, respectively) (Jadlowsky et al., 2014). We noted that KAP1 displayed a preference for the 5'-LTR of HIV, consistent with previous reports of it occupying the LTR of endogenous retroelements in stem cells and cellular genes (Iyengar et al.,

2011; Rowe et al., 2010). A minor fraction of KAP1 was also detected at the 3'-LTR of the integrated provirus. However, none of the ChIP samples from components of the 7SK snRNP displayed significant enrichment at the 3'-LTR, demonstrating that the inhibitory snRNP, and the majority of KAP1, primarily co-occupy the 5'-LTR of HIV. In agreement with previous reports (D'Orso and Frankel, 2010a; D'Orso et al., 2012), the peak intensity of the 7SK snRNP and KAP1 closely mirrors the occupancy profile of Pol II and the PIC (as revealed by the occupancy of basal transcription factors Sp1 and TBP) at the promoter (**Figure 18A**).

Given that HIV integrates randomly and that integration site placement dictates basal transcriptional activity (Jordan et al., 2001; Schroder et al., 2002), we asked whether the occupancy of KAP1 and the 7SK snRNP at the HIV promoter was dependent on the integration site, chromatin context and promoter activity. To evaluate this, we performed ChIP-qPCR with the primer-pair that specifically amplifies the HIV promoter-proximal region (5'-LTR, +141) and measured transcription initiation levels across several cell-based models containing a single HIV integrant (Jordan et al., 2001; Pearson et al., 2008; Schroder et al., 2002). Notably, we observed that KAP1 and the 7SK snRNP (CDK9, Larp7, and Hexim1) are recruited only to promoters containing a PIC (as revealed by the presence of Sp1 and TBP) and transcriptionally engaged Pol II such as in the E4, 2B2D, 8.4 and 9.2 HIV clones (**Figures 18B–E**). Conversely, KAP1 and the 7SK snRNP do not appear to occupy the promoter of a provirus integrated in chromatin-dense regions without noticeable PIC and Pol II occupancy (6.3 clone, **Figure 18F**) or a promoter containing a mutation that impairs PIC assembly, Pol II recruitment, and transcription initiation (2B5 clone, **Figure 18G**).

Given the correlation between PIC assembly, Pol II pausing, and the presence of the KAP1-7SK snRNP complex at promoter-proximal regions, we asked if the HIV genomes in

these cell lines are transcriptionally active. To test this, we performed qRT-PCR assays to measure levels of promoter-proximal transcripts (indicative of transcription initiation), and observed that higher transcript levels (in the E4, 2B2D, 8.4 and 9.2 HIV clones) directly correlated with the presence of a PIC, transcriptionally engaged Pol II, and the KAP1-7SK snRNP complex at promoter-proximal regions (**Figure 18H**). On the contrary, 6.3 and 2B5 showed much less promoter-proximal transcripts (at least 7-10 fold less compared with active promoters like in the E4 clone). These results are consistent with previous evidence demonstrating that 6.3 exhibits very low transcript levels and 2B5 contains a non-functional promoter (Jadlowsky et al., 2014; Jordan et al., 2001; Pearson et al., 2008).

Taken together, these findings indicate that KAP1-7SK snRNP occupies the promoters of proviruses that show evidence of transcriptional initiation. The data is in agreement with previous evidence suggesting an intimate link between loading of the 7SK snRNP at gene promoters and transcriptionally engaged Pol II (D'Orso and Frankel, 2010a).

KAP1 Mediates Recruitment of the 7SK snRNP Complex to the HIV and Cellular Promoter-proximal Regions

Given that KAP1 associates with the 7SK snRNP and that both are recruited to chromatin, we hypothesized that KAP1 recruits the inhibitory snRNP to gene promoters to control the transition from transcription initiation into elongation. To test this model, we utilized a genetic interaction approach where the Jurkat HIV E4 clone was transduced with self-inactivating lentiviruses expressing non-target (shNT) or KAP1 (shKAP1) shRNAs. Importantly, this approach yielded efficient KD (KD) of KAP1 (~80-85%) without largely affecting normal levels

of the 7SK snRNP components, indicating that snRNP components only undergo subtle changes in response to KAP1 KD (**Figure 19A**). Given the minor fluctuations in protein levels observed, we asked whether the stability of the snRNP is disrupted when KAP1 expression is lost (Liu et al., 2014). To test this, we performed IP of endogenous Larp7 complexes in the shNT and shKAP1 cell lines and quantitated levels of co-immunoprecipitated 7SK snRNP components (Hexim1 and CDK9) as well as KAP1. The data demonstrated minimal alterations in 7SK snRNP integrity (due to reduced CDK9 levels) in response to KAP1 KD (shKAP1 cell line) (**Figure 19B**). Thus, although we observed minor fluctuations in the 7SK snRNP components, the overall stability of the snRNP appears to be KAP1-independent.

Using these cell lines, we then assessed whether KAP1 plays any role in recruiting the 7SK snRNP complex to the HIV and cellular gene promoters by ChIP. Importantly, since loss of KAP1 resulted in only minimal 7SK snRNP disruption, this strategy allowed us to quantify direct effects of KAP1 loss on the promoter-bound 7SK snRNP pool. Expectedly, KAP1 and the 7SK snRNP (CDK9, Larp7 and Hexim1) co-occupy the HIV promoter but not intragenic domains (gene body) in the control cell line (shNT). However, the occupancy of KAP1 and the 7SK snRNP at the HIV promoter was reduced by a factor of ~3–5-fold (depending on the subunit) upon KAP1 KD (shKAP1) (**Figure 19C**). Although Pol II occupancy at the HIV promoter increased ~2-fold in response to KAP1 KD, we did not detect any significant increase in Pol II levels at the gene body (**Figure 19C**), consistent with a previous report showing evidence of Pol II accumulation at some cellular promoters in response to KAP1 KD (Bunch et al., 2014). Notably, PIC assembly at the HIV promoter (as revealed by Sp1 and TBP occupancy) remained unaffected in response to KAP1 KD, indicating that loss of KAP1 directly affects 7SK

snRNP recruitment to the promoter without altering assembly of the basal transcriptional apparatus.

Given our previous findings of detecting the 7SK snRNP at NF-κB–regulated gene promoters, we wanted to test whether KAP1 mediates 7SK snRNP recruitment to cellular genes as observed with HIV. We chose two inflammatory-responsive PRGs: NF-κB inhibitor alpha (*NFKBIA*, encoding the protein IκBα) and Tumor Necrosis Factor Alpha-Induced Protein 3 (*TNFAIP3*, encoding the protein A20) due to the presence of a fully-assembled PIC at their promoters, P-TEFb–dependence, Pol II pausing, and a strong and rapid induction in the presence of inflammatory stimuli (TNF) through NF-κB (Diamant and Dikstein, 2013; Smale, 2010). Notably, both PRGs showed evidence of KAP1 and 7SK snRNP recruitment at their promoters but not in the gene bodies, as well as promoter-proximal paused Pol II prior to stimulation (**Figures 19D** and **E**). Similar to HIV, KAP1 KD produced a sharp reduction (~3-to-10 fold depending on the subunit) in the occupancy of the 7SK snRNP complex at the *NFKBIA* and *TNFAIP3* gene promoters (**Figures 19D** and **E,** respectively).

We also examined whether the KAP1-7SK snRNP complex is recruited to a constitutively and actively transcribed gene such as *ACTB* (encoding the protein β-actin), but we found no significant evidence of stable KAP1 enrichment at the promoter or an effect of KAP1 KD in Pol II levels at this gene (**Figure 19F**), consistent with a previous report (Iyengar et al., 2011a). This strongly indicates that KAP1 is deposited at promoter-proximal regions of inducible genes containing paused Pol II, and that KAP1 KD does not interfere with steady-state RNA levels of genes that are not direct targets. Similarly to the HIV promoter, PIC assembly (as revealed by the occupancy of the basal factors Sp1 and TBP) at the cellular *NFKBIA* and *TNFAIP3* gene promoters remains much unaffected in response to loss of KAP1 (**Figures 19D**

and **E**), indicating that KAP1 functions at a step downstream control transcription activation at inducible genes.

To further examine how the observed changes in KAP1 and 7SK snRNP occupancy at the HIV and cellular (*NFKBIA* and *TNFAIP3*) genes affect RNA levels from these promoters, we performed qRT-PCR assays of transcripts at steady state in KAP1 KD cells. There was no observable change in HIV transcripts and a small drop in *NFKBIA* and *TNFAIP3* expression in response to KAP1 KD (**Figures 19G–I**) without alterations in *β-ACTIN* levels (**Figure 19J**), demonstrating that at these inducible genes, KAP1 KD does not abruptly alter basal transcripts of the genes assayed.

As a control, we wanted to validate that our findings that KAP1 KD would result in the loss of the promoter assembled KAP1-7SK snRNP a separate HIV clonal cell line. We thus tested the effect of KAP1 KD on the occupancy of the 7SK snRNP in the cell line (2B2D). We had previously found that the 2B2D cell line gave a very similar phenotype to the WT E4 (**Figure 18**), and in agreement with that, we observed that KAP1 KD resulted in a decrease in KAP1-7SK snRNP occupancy at its promoter (**Figure 20A**). Moreover, we observed that KAP1 and the 7SK snRNP also occupied the promoter of the well-known Pol II paused gene *HSPA1B* (HSP70 protein) (**Figure 20B**), but not at the transcriptionally inactive gene regulated by estrogen in breast cancer (*GREB1*, encoding the protein GREB-1) (**Figure 20C**). *GREB1* served as a negative control because it does not contain an assembled PIC nor a transcriptionally engaged Pol II (Bunch et al., 2014; Lis et al., 2000), similar to the virtually inactive HIV promoter in the 6.3 and 2B5 clones (**Figures 18F** and **G,** respectively).

Pol II and PIC Assembly at Promoter-proximal Regions is Required for KAP1 and 7SK snRNP Recruitment

Our previous findings supported a model in which only transcriptionally active promoters allowed for placement of the KAP1-7SK snRNP complex at gene promoters. To test this further, we used two different chemicals, Triptolide (TPL) and 5,6-dichloro-1-β-D-ribofuranosyl-1H-benzimidazole (DRB), to block transcription initiation or elongation, respectively (**Figure 21A**). While TPL blocks the XPB helicase of the TFIIH complex and prevents formation of the open complex needed for transcription initiation, DRB targets the P-TEFb kinase thereby blocking transcription elongation but not initiation (Bensaude, 2011; Titov et al., 2011) (**Figure 21A**). Using these inhibitors for a short-term (1 hour) in the Jurkat HIV E4 clone we observed that DRB only interferes with the synthesis of promoter-distal transcripts, consistent with its role in blocking elongation; conversely, TPL blocks the synthesis of both promoter-proximal and -distal transcripts, in agreement with its role in blocking Pol II recruitment to gene promoters (**Figure 21B**).

To test whether TPL antagonizes the recruitment of the KAP1-7SK snRNP complex to PRG promoters, we performed ChIP assays in cells treated with DMSO (vehicle control), DRB, or TPL. We observed that TPL, but not DRB, significantly reduced the occupancy (by a factor of ~3–10-fold, depending on the subunit and target gene) of KAP1 and the 7SK snRNP (CDK9, Hexim1 and Larp7) at the HIV and cellular (*NFKBIA* and *TNFAIP3*) gene promoters (**Figures 21C–E**). Similarly, TPL blocked Pol II assembly at promoters, indicating that Pol II undergoes rapid abortive cycles of initiation and termination which are undetectable by ChIP due to its dynamic nature, consistent with data from a recent genome-wide study (Chen et al., 2015). As a

control, we observed that treatment with the elongation inhibitor DRB did not block promoter assembly of Pol II, KAP1, and the 7SK snRNP at these gene promoters. This finding is expected and is in agreement with the inhibitory role of DRB being downstream of assembly of the PIC and transcription initiation by Pol II. These results are consistent with previous studies showing reduction in Pol II assembly at promoters in response to TPL, even in shorter treatments (Jonkers et al., 2014) on a gene-to-gene basis (Henriques et al., 2013).

While TPL blocked the recruitment of Pol II, KAP1 and the 7SK snRNP to the three gene promoters, the active chromatin signature (H3K4me3) remains similar at the *NFKBIA* promoter or modestly increased (1.5-2–fold) at the HIV and *TNFAIP3* promoter-proximal regions (**Figures 21C–E**). Despite this unexpected result, which we attribute to the long-lived H3K4me3 mark compared with the dynamic nature of Pol II recruitment to promoters, and cycles of H3 methylation/demethylation (Guenther et al., 2007; Min et al., 2011), we conclude that PIC formation and Pol II pausing are required for assembly of the KAP1-7SK snRNP complex at these gene promoters. The data is consistent with previous observations indicating that KAP1 functions at a step downstream of PIC assembly to control transcription activation (**Figures 18, 19,** and **20**).

Given that TPL almost completely abolishes the ChIP signal of Pol II, KAP1, and the 7SK snRNP at the gene promoters evaluated, we examined if this short-term TPL treatment alters the stability of these proteins. Importantly, we found that total protein levels of Pol II, KAP1, and components of the 7SK snRNP (CDK9, Hexim1 and Larp7) were not altered in response to TPL (**Figure 21F**), thereby indicating that their reduced promoter density is not due to aberrant protein degradation.

Collectively, the data strongly supports the hypothesis that recruitment of KAP1 and the 7SK snRNP to gene promoters requires PIC formation, engagement of Pol II and transcription initiation, but not elongation. This data is consistent with our previous observations that KAP1-7SK snRNP complex is only detected at active HIV promoters containing an assembled PIC and transcriptionally engaged Pol II (**Figure 18, 19, and 20**).

KAP1-Mediated Recruitment of the 7SK snRNP Complex to Promoters Enables Rapid Transcriptional Kinetics in Response to Stimulation

Given that KAP1 recruits the 7SK snRNP complex to the promoters of HIV and cellular PRGs regulated by NF-κB, we reasoned that the loss of 7SK snRNP on promoters in response to KAP1 KD alters the kinetics of gene induction in response to stimuli. To examine this in detail, we monitored transcriptional kinetics of three NF-κB regulated PRGs (HIV, *NFKBIA*, and *TNFAIP3*) in response to stimulation (TNF) in the NT (shNT) or KAP1 (shKAP1) cell lines (**Figure 22**). Given that in response to TNF, HIV is sequentially activated by NF-κB and the viral activator Tat, we used the HIV *Tat null* cell line 2B2D (Pearson et al., 2008) (here referred to as HIV for simplicity) to solely monitor NF-κB–mediated transcriptional activation. While HIV RNA levels are similar in both cell lines prior to stimulation, we observed a sharp decrease in the magnitude of RNA synthesis in response to TNF in the shKAP1 cell line compared to the shNT control (**Figure 22A**). This decrease in magnitude was also observed, albeit with gene-specific differences, in the *NFKBIA* and *TNFAIP3* genes in response to TNF (**Figures 22B–C**). The reduced magnitude of RNA synthesis demonstrates that KAP1-mediated 7SK snRNP recruitment to PRG promoters is required for the rapid and robust induction upon stimulation.

To test if the delay in gene activation was due to the inability of the transcription factor NF-κB and Pol II to be properly recruited to the promoter in response to stimulation, we performed ChIP assays. For this, we treated the shNT and shKAP1 cell lines with TNF for different time points (both preceding and following the kinetics of RNA synthesis) and monitored occupancy of NF-κB and Pol II at the promoters of HIV, *NFKBIA*, and *TNFAIP3* (**Figures 22D–F**, respectively) and in downstream transcribed regions (**Figures 22G–I**, respectively). In response to TNF, NF-κB is recruited rapidly and similarly to the three gene promoters examined in both the shNT and shKAP1 cell lines, albeit with gene-specific differences in kinetics (**Figures 22D–F**), but not to the gene bodies (**Figures 22G–I**).

While the kinetics of NF-κB and Pol II recruitment to the HIV promoter in response to TNF was not impaired, CDK9 recruitment was drastically delayed and its magnitude (occupancy levels) was significantly antagonized (>3-fold reduced density in the shKAP1 cell line compared to the shNT control) (**Figure 22D**). We then asked if this delayed recruitment of CDK9 to the promoter was also observed in the *NFKBIA* and *TNFIAP3* gene promoters. Consistent with HIV, neither NF-κB nor Pol II was antagonized in their recruitment to the *NFKBIA* and *TNFIAP3* promoters in response to TNF (**Figures 22E** and **F**). However, CDK9 recruitment to these two gene promoters was strongly antagonized in the shKAP1 cell line compared to the shNT control (**Figures 22E** and **F**). This finding demonstrates that KAP1-mediated 7SK snRNP recruitment to promoters (before and after stimulation) facilitates the delivery of the P-TEFb kinase to promoters to promote gene activation.

To test if decreased RNA levels were a result of a non-processive Pol II as a consequence of delayed CDK9 recruitment, we again performed ChIP assays after a temporal TNF treatment to monitor levels of CDK9 and Pol II in the gene body of the three PRGs. Consistent with the

ChIP results at promoters, CDK9 and Pol II occupancy in the gene body of all three PRGs was decreased in the shKAP1 cell line in relation to the shNT control in response to TNF (**Figures 22G–I**). Notably, the reduced magnitude of CDK9 and Pol II in the gene bodies is consistent with decreased RNA levels in response to stimulation, indicative of a restriction in productive elongation.

In the case of HIV and *TNFAIP3*, more Pol II was recruited to the promoter in response to TNF (**Figures 22D** and **F**). However, this Pol II was not elongation-competent because gene body levels were lower in the shKAP1 cell line compared to the shNT control for both genes (**Figures 22G** and **I**), in agreement with the requirement of KAP1 for pause release, and the low levels of CDK9 observed in the gene body.

To address whether KAP1 KD simply delayed the transcriptional response to stimuli, we took our TNF time course out to 480 minutes. In agreement with our previous results, we observed that the shKAP1 cells demonstrated prolonged HIV transcriptional antagonization after treatment with TNF (**Figure 23A**). Importantly, this same trend was observed for the cellular genes *NFKBIA* and *TNFAIP3* (**Figures 23B** and **C,** respectively), albeit with gene to gene variability.

Together, we have established a new role for KAP1 in transcription elongation control through deposition of catalytically inactive P-TEFb at gene promoters. Furthermore, the data indicate that TF and P-TEFb recruitment to promoters is kinetically uncoupled, and that KAP1 facilitates the local delivery of P-TEFb (as part of the 7SK snRNP) to activate the kinase ‘on site’ and promote Pol II phosphorylation.

KAP1 is Dynamically Recruited to Promoters to Enable 7SK snRNP Loading and P-TEFb Delivery in Response to Stimulation

The finding that KAP1 mediates 7SK snRNP recruitment to promoters in basal conditions does not completely explain why more P-TEFb kinase is recruited in response to stimulation (to facilitate rapid NF-κB–mediated Pol II elongation). Given that the KAP1-7SK snRNP complex is present both on and off promoters, we hypothesized that additional P-TEFb molecules are recruited from the nucleoplasm to the promoter to act on newly incorporated Pol II to facilitate rapid elongation upon stimulation. Since KAP1 indirectly interacts with P-TEFb through the 7SK snRNP (**Figure 16**), we reasoned that in addition to recruitment of the 7SK snRNP to the promoter (which is required for the initial round of elongation), KAP1 facilitates recruitment of subsequent 7SK snRNP molecules during stimulation to continuously resupply the promoter with additional P-TEFb and facilitate subsequent rounds of elongation (see model in **Figure 24A**).

This hypothesis is supported by our previous results demonstrating that KAP1 KD resulted in both decreased basal 7SK snRNP recruitment levels (**Figures 19** and **20**) as well as antagonized P-TEFb recruitment to promoters upon stimulation (**Figures 22D–F**). We thus envisioned two potential mechanisms that take place once the promoter-bound 7SK snRNP is disassembled upon stimulation and P-TEFb released to facilitate the initial round of elongation: (1) KAP1 remains bound to the promoter to facilitate the recruitment of subsequent 7SK snRNP molecules (**Figure 24A,** top), or (2) the entire KAP1-7SK snRNP complex is continuously resupplied to promoters upon stimuli (**Figure 24A,** bottom). In either case, KAP1 facilitates the recruitment of inactive but primed kinase to promoters as part of the 7SK snRNP complex. To

discern between these two mechanisms, we took advantage of our inducible cell lines, where KAP1:FLAG or GFP:FLAG are only expressed in the presence of DOX (**Figure 17A**), and performed competition ChIP assays (van Werven et al., 2009). After a short time induction of KAP1:FLAG and GFP:FLAG, cells were treated with TNF to monitor if there is any KAP1 turnover at promoters through competition of endogenous KAP1 with induced KAP1:FLAG. In the event that new KAP1-7SK snRNP is recruited upon stimulation, the FLAG ChIP signal would increase temporally thereby showing competition with endogenous KAP1 for promoter binding. Conversely, if KAP1 remains stably bound at the promoter and more 7SK snRNP is recruited upon stimulation, then KAP1:FLAG would not replace promoter-bound endogenous KAP1. Importantly, the inducible KAP1:FLAG is expressed at similar levels to endogenous KAP1 thereby allowing us to perform the experiment in conditions where the transgene is not over-expressed (**Figure 17A**).

Upon stimulation with TNF, KAP1 occupancy at the *TNFAIP3* promoter remains virtually constant, and low-to-undetectable in the gene body (**Figure 24B**). However, a rapid recruitment of KAP1:FLAG, but not GFP:FLAG, to the promoter is observed in response to TNF, signifying that the endogenous KAP1 is being competed off with KAP1:FLAG (**Figures 24C** and **D**). This phenomena was exclusive to the *TNFAIP3* promoter as there was no difference between GFP:FLAG and KAP1:FLAG recruitment to the gene body (**Figures 24C** and **D**).

Importantly, as control, factor occupancy at the *HSPA1B* locus was examined, which shows evidence of KAP1 recruitment to the promoter but not to the gene body (**Figure 24E**), alongside with the 7SK snRNP and transcriptionally engaged Pol II (**Figure 20B**). Since *HSPA1B* is non-responsive to TNF, we reasoned that no turnover of KAP1 will occur at the promoter after stimulation, and therefore KAP1:FLAG will not compete off endogenous KAP1

assembled at the promoter. While lower levels of KAP1:FLAG, but not GFP:FLAG, were detected at the promoter before stimulation (probably because KAP1 has dynamic promoter interaction properties), no additional recruitment of KAP1:FLAG to the promoter nor gene body was observed upon stimulation (**Figures 24F** and **G)**, indicating that at non-responsive, highly-paused genes, KAP1 does not undergo any turnover at the promoter in response to stimulation. Notably, upon stimulation, CDK9 levels in the *TNFAIP3* promoter and gene body (but not at the *HSPA1B* locus) increase with similar kinetics to KAP1, potentially indicating that they are co-recruited (**Figures 24H** and **I**, respectively).

Taken together, the data demonstrate that KAP1-7SK snRNP remains stably bound to the promoter in basal conditions but it is dynamically recruited to gene promoters in response to stimulation. This indicated that KAP1 continuously recruits inactive, but primed P-TEFb to gene promoters and that the kinase can be activated 'on site' upon TF binding to hyper-phosphorylate Pol II and stimulate transcription elongation.

KAP1 and the 7SK snRNP Co-occupy Most Gene Promoters with Transcriptionally Engaged Pol II

Thus far, these results support a model in which KAP1 mediates placement of the 7SK snRNP complex to promoters of transcriptionally active but paused genes. To obtain genome-wide information, we performed ChIP-seq of KAP1 and 7SK snRNP components (CDK9, Hexim1 and Larp7) alongside with Pol II. Given our proposed link between activated genes and placement of the KAP1-7SK snRNP complex, the colonic cancer cell line HCT116 was used to compare levels of Pol II and KAP1-7SK snRNP at promoters with transcriptional activity by

using the nascent RNA sequencing data generated by the Espinosa lab using GRO-seq (Galbraith et al., 2013).

After ranking all Refseq genes in the human genome (30181 unique TSS) based on decreasing Pol II occupancy, we observed that KAP1 and components of the 7SK snRNP co-occupy promoter-proximal regions and that their distribution is similar, if not identical, to Pol II (**Figure 25A**). Consistent with our previous results, the distribution of KAP1-7SK snRNP at promoter-proximal regions also mirrors that of the active chromatin signature (H3K4me3), as well as H3K27Ac, but not the enhancer signature (H3K4me1) (**Figure 25A**), which is absent from promoter-proximal regions (Calo and Wysocka, 2013). Analysis of the GRO-seq dataset sorted based on ranked Pol II occupancy indicated that KAP1 and the 7SK snRNP complex co-occupy genes containing transcriptionally engaged Pol II (**Figure 25A**).

The location of paused Pol II and KAP1-7SK snRNP also matches the expected distribution of the active markers (H3K4me3 and H3K27Ac) at promoter-proximal regions, which are devoid of H3K4me1 (**Figure 25B**), in agreement with previous studies (Calo and Wysocka, 2013).

To determine the number of genes regulated by the KAP1-7SK snRNP complex, we first compiled a list of genes marked by the three 7SK snRNP components (Hexim1, CDK9, and Larp7) at promoter-proximal regions (using a 500-bp window from the summit of all peaks). Using this threshold, we identified that over 17,987 genes contain the 7SK snRNP, which corresponds to 59.6% of all annotated unique TSS's (**Figure 25C**). Although a large fraction of active genes are marked by both KAP1 and the 7SK snRNP, the number of genes identified is likely underestimated due to the absence of one of the three markers from some genes based on the high confidence threshold (FDR < 0.05) used in during peak calling. Using this class of

genes, we then determined the fraction of gene promoters in the human genome that are marked by co-occupancy of the 7SK snRNP, KAP1 and transcriptionally engaged Pol II (a Pol II ChIP-seq peak summit located with 500 bp of the TSS). In total, we observed that over 14,141 genes contain KAP1, the 7SK snRNP, and a poised Pol II, which corresponds to 46.9% of all annotated unique TSS's. Moreover, we observed that KAP1 and the 7SK snRNP were present at over 88% of TSS's marked with a poised Pol II (**Figure 25D**). We thus refer to these genes as KEC (KAP1-7SK snRNP early Elongation Complex) target genes. To visualize this, we performed an inspection of the ChIP-seq tracks for the gene *TNFAIP3,* which we had previously observed contained all components of the KEC as determined by ChIP-qPCR (**Figure 18**). Consistent with our findings, we were able to observe the KEC at the promoter of *TNFAIP3*, along with the markers H3K4me3 and H3K27ac as well as a nascent RNA, demonstrating that this gene is a transcriptionally active, Pol II poised gene (**Figure 25E**).

We wished to further expand on our genomics analysis of the KEC and thus. Clustering both Watson and Crick strands of the genome to Pol II showed that there was no strand specificity for Pol II and that our analysis was not biased to genes on either strand (**Figure 26A**). To gain further insight into the relative position of the KEC relative to Pol II, a ± 500 bp metagene analysis was done. We observed that KAP1 and all 7SK snRNP components (CDK9, Hexim1 and Larp7) peak at 200/250-bp downstream from the TSS (**Figure 26B**), slightly downstream of the Pol II pause peak located 100-bp downstream from the TSS, in agreement with previous reports (Danko et al., 2013; Muse et al., 2007). To further confirm out previous findings, we visualized ChIP-seq tracks for the genes *NFKBIA*, *HSPA1B* and *GREB1* for the KEC. As expected, only the transcriptionally active *NFKBIA* and *HSPA1B* contain a promoter assembled KEC, whereas the transcriptionally inactive *GREB1* did not (**Figures 26C–E**). At this

level of resolution, we cannot conclude that the KAP1-7SK snRNP resides ~100 bp downstream of the paused Pol II, but we did note that there were several locally distinct species of the snRNP in the genes *TNFAIP3* and *NFKBIA*. We therefore acknowledge that our metagene analysis can suffer a slight skewing effect through the deposition of multiple KAP1-7SK snRNP molecules at promoter proximal regions.

We have previously shown that KAP1 is strictly required for the induction of PRG in response to stimulation. However, the fact that KAP1 occupies most Pol II regulated genes in the human genome prompted us to test the effect of the loss of KAP1 in unstimulated gene expression. For this, HCT116 shNT and shKAP1 cell lines (which showed a KD of ~80%) were generated and used for qPCR assays to measure steady-state RNA levels. We observed that the majority of the transcripts examined have decreased RNA levels in the shKAP1 cell line (**Figure 26F**). Although GRO-seq experiments will be needed to examine this genome-wide, the data supports that KAP1 has indeed an effect in controlling basal transcriptional levels in addition to playing a critical role in the transcription of inducible genes.

A Model for KAP1-mediated Recruitment of the 7SK snRNP Complex to Promoters and Transcription Elongation

Collectively, KAP1–mediated 7SK snRNP recruitment to promoter-proximal regions plays a critical role in gene activation. We therefore propose a model in which transcription initiation primes a gene for rapid induction through the placement of the KEC at the promoter. In normal cells, TFs such as NF-κB locally capture P-TEFb for rapid gene activation in response to induction stimuli through KAP1-mediated placement of the 7SK snRNP. Importantly,

recruitment of KAP1–7SK snRNP to promoters does not affect the kinetics of TF-promoter DNA interaction in response to stimuli, rather it facilitates the rapid capture of P-TEFb by the TF to promote transcription elongation (initial round). From there, additional KEC is further recruited to promote subsequent rounds of elongation, and robust gene induction in response to activating stimuli (**Figure 27A**).

In KAP1 depleted cells, NF-κB and Pol II are recruited to target gene promoters in response to stimuli as well as in normal cells. Conversely, the loss of KAP1 drastically affects the normal kinetics of P-TEFb recruitment, delaying NF-κB-mediated Pol II phosphorylation and attenuating both the initial and subsequent rounds of transcription elongation (**Figure 27B**). Notably, the fact that TF recruitment to target gene promoters is not affected by loss of KAP1, it indicates that the recruitment of TF and the P-TEFb kinase is mechanistically and temporally uncoupled, providing novel insights into the mechanisms of transcriptional activation.

The genome-wide distribution of KAP1 and the 7SK snRNP at most transcriptionally active genes in the human genome underscores the importance of this complex in gene activation in basal conditions (no stimuli). However, due to the nature of transcriptional activation of constitutive and inducible genes, the last set of genes might be more sensitive to the loss of KAP1, probably due to the critical requirement of elongation factors for their rapid gene activation.

Figures and Figure Legends

Figure 16

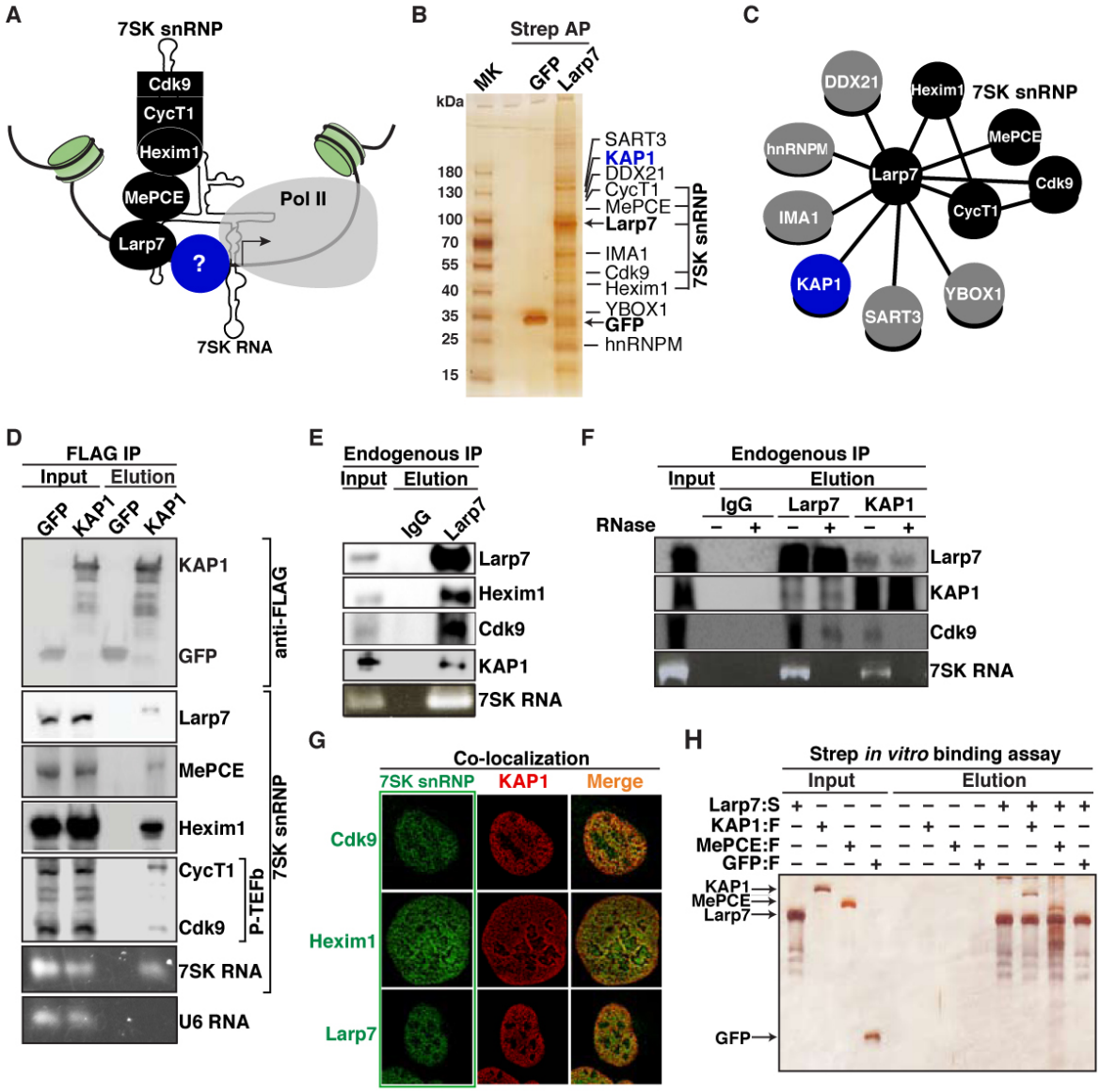

Figure 16. Identification of KAP1 as an Interactor of the 7SK snRNP Complex.

(A) Model for the tethering of the 7SK snRNP complex (black) to a gene promoter containing promoter-proximal paused Pol II through an as-yet unidentified factor(s) (blue).

(B) Strep affinity purification (AP) of GFP (GFP:S) and Larp7 (Larp7:S) from stable HEK 293T T-REx cell lines. An in-solution mixture was analyzed by LC-MS/MS analysis. High-confidence Larp7 protein Interactors were identified by mass spectrometry analysis (see also **Table 2**) and indicated by lines on the silver-stained gel.

(C) A network representation of protein-protein interactions between Larp7 and proteins identified in panel (B). Solid lines indicate direct protein interaction while dashed lines indicate 7SK RNA-mediated interactions based on data from this and other studies (Jeronimo et al., 2007; Krueger et al., 2008; McNamara et al., 2013; Calo et al., 2015).

(D) Validation of KAP1 and 7SK snRNP interaction. Stably-expressed FLAG-tagged GFP or KAP1 proteins were affinity purified from HEK 293 T-REx cell lines and interactions analyzed by Western blot with the indicated antibodies. The bottom panels are agarose gels showing qRT-PCR amplifications of 7SK and U6 RNAs. U6 was used as a negative IP control. Input and elution represent 0.5% and 5% of the initial material, respectively.

(E) Interaction between endogenous KAP1 and 7SK snRNP components. Lysates from HEK 293T cells were immunoprecipitated with anti-Larp7 (or normal IgG serum) and protein and 7SK RNA interactions were analyzed by Western blot or qRT-PCR, respectively. Input and elution represent 0.5% and 10% of the initial material, respectively.

(F) The interaction between KAP1 and P-TEFb, but not Larp7, is RNase sensitive. Endogenous IPs (Larp7 and KAP1) or a normal IgG serum was performed in the absence or presence of

RNaseA. Ethidium bromide stain of PCR amplified 7SK RNA demonstrates the efficiency of RNAse treatment. Input and elution represent 0.5% and 10% of the initial material, respectively. (G) Full 3D de-convoluted confocal microscopy images of KAP1 and 7SK snRNP components. U2OS cells were used to measure co-localization of KAP1 and 7SK snRNP components.

(H) Larp7 and KAP1 directly interact *in vitro*. KAP1:FLAG, MePCE:FLAG, and GFP:FLAG were purified using high-salt conditions and incubated with immobilized Strep-tagged Larp7 (Larp7:S) (+) or empty beads (–). Note that both KAP1 and Larp7 protein purifications are devoid of any co-purifying 7SK RNA (as judged based on qRT-PCR assays) thus indicating that the interaction is direct and not 7SK RNA-mediated.

Figure 17

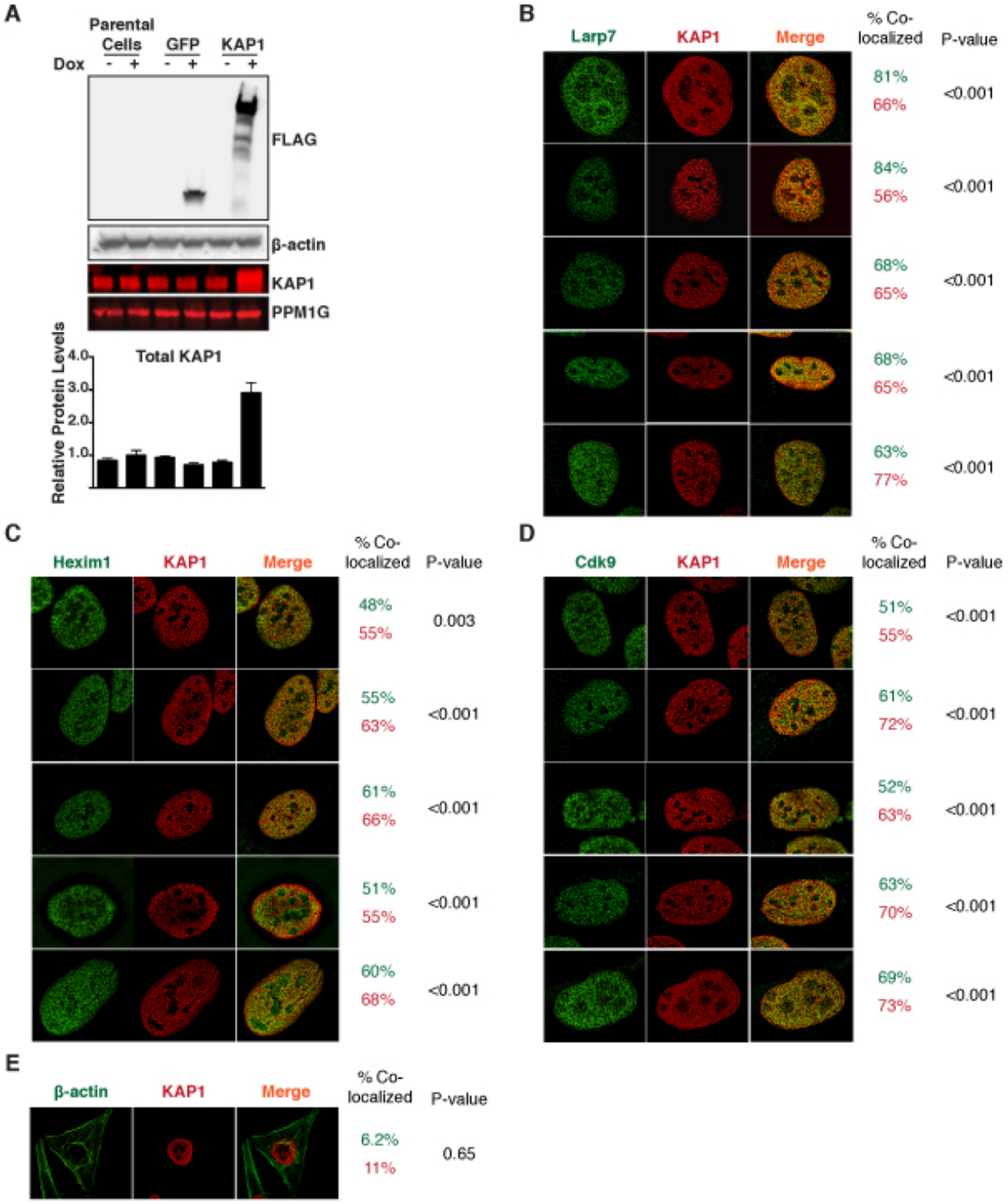
A
Parental Cells
GFP
KAP1
Dox - + - + - +
FLAG
β-actin
KAP1
PPM1G
Total KAP1
Relative Protein Levels
4.0
3.0
2.0
1.0
B
Larp7
KAP1
Merge
% Co-localized
P-value
81%
66%
<0.001
84%
56%
<0.001
68%
65%
<0.001
68%
65%
<0.001
63%
77%
<0.001
C
Hexim1
KAP1
Merge
% Co-localized
P-value
48%
55%
0.003
55%
63%
<0.001
61%
66%
<0.001
51%
55%
<0.001
60%
68%
<0.001
D
Cdk9
KAP1
Merge
% Co-localized
P-value
51%
55%
<0.001
61%
72%
<0.001
52%
63%
<0.001
63%
70%
<0.001
69%
73%
<0.001
E
β-actin
KAP1
Merge
% Co-localized
P-value
6.2%
11%
0.65

Figure 17. Generation of Inducible Cell Lines and Co-localization of KAP1 with Components of the 7SK snRNP Complex.

(A) HEK 293 T-REx cells stably expressing GFP:FLAG and KAP1:FLAG or no transgene (parental cells transfected with an empty pcDNA-4/TO plasmid) were treated with (+) or without (–) doxycycline (DOX). A FLAG Western blot shows the position of the KAP1:FLAG and GFP:FLAG proteins. β-actin is shown as loading control. Total protein levels of KAP1 (endogenous and KAP1:FLAG) in each cell line were quantified compared to PPM1G (large protein loading control) and plotted. A.U. denotes arbitrary units of protein levels.

(B) Full 3D de-convoluted confocal microscopy images of KAP1 and Larp7. The percentage KAP1 co-localization (or Pearson correlation coefficient, *r*) with Larp7 and vice versa (green) is shown. The probability (p-value) that the measured value of *r* from the two colors is significantly greater than values of *r* that would be calculated if there was only random overlap is shown.

(C) Full 3D de-convoluted confocal microscopy images of KAP1 and Hexim1. Percentage co-localization between KAP1 and Hexim1 (green) and vice versa, and a p-value was calculated based on random pixel distribution (calculated as in panel (B)).

(D) Full 3D de-convoluted confocal microscopy images of KAP1 and CDK9. Percentage co-localization between KAP1 co-localizing with CDK9 (green) and vice versa, and a p-value was calculated based on random pixel distribution (calculated as in panel (B)).

(E) Full 3D de-convoluted confocal microscopy image of KAP1 and β-actin, which served as a negative co-localization control. Percentage co-localization is shown of both red (KAP1) co-

localizing with green (β-actin) and vice versa and a p-value was calculated based on random pixel distribution (calculated as in panel (B)).

Figure 18

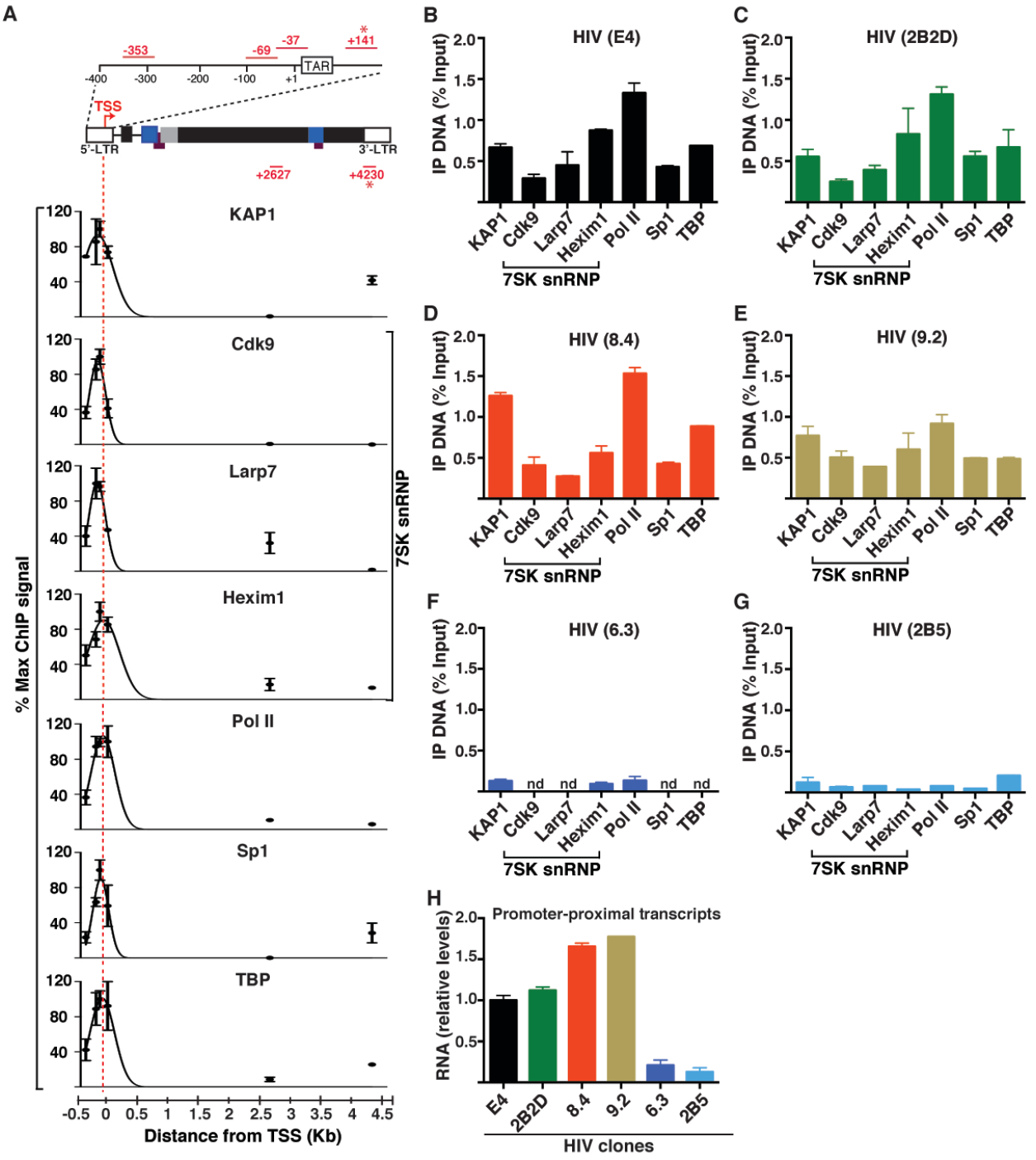

Figure 18. KAP1 and the 7SK snRNP Complex Co-occupy the HIV Promoter along with paused Pol II.

(A) (Top) Scheme of wild type HIV provirus in the E4 clone with the location of the amplicons used in ChIP-qPCR indicated. (Bottom) ChIP of KAP1, components of the 7SK snRNP (CDK9, Larp7 and Hexim1), total Pol II and PIC components (Sp1 and TBP) followed by qPCR using primer-pairs amplifying six different genomic regions (dots). For each antibody, maximal ChIP signal was arbitrarily set to 100% and all other amplifications were standardized to it (mean ± SEM are shown; n = 3). Amplicons marked with an asterisk (*) indicate 5'-LTR (+141) or 3'-LTR (+4230) specific. From each data set, best-fit non-linear regression curves were generated using Prism to approximate the ChIP signal throughout the HIV genome at a given location.

(B) KAP1 and the 7SK snRNP are detected at the HIV promoter in the E4 clone. ChIP assays using the antibodies indicated were done and followed by qPCR assays using the 5'-LTR-specific amplicon (+141). Percentage (%) of IP DNA respective to Input was calculated (mean ± SEM are shown, n = 3).

(C) KAP1 and the 7SK snRNP are detected at the HIV promoter in the 2B2D clone. Same as in panel (B).

(D) KAP1 and the 7SK snRNP are detected at the HIV promoter in the 8.4 clone. Same as in panel (B).

(E) KAP1 and the 7SK snRNP are detected at the HIV promoter in the 9.2 clone. Same as in panel (B).

(F) KAP1 and the 7SK snRNP are not detected at the HIV promoter in the 6.3 clone. Same as in panel (B).

(G) KAP1 and the 7SK snRNP are not detected at the HIV promoter in the 2B5 clone. Same as in panel (B).

(H) Proviruses containing transcriptionally engaged Pol II, KAP1, 7SK snRNP, and PIC components have higher levels of promoter-proximal transcripts. Total RNA was isolated from the different Jurkat HIV clones used in panels (B–G) and relative HIV promoter-proximal transcript levels (+141) were quantified and normalized to *ACTB* by qRT-PCR (mean ± SEM are shown, n = 3).

Figure 19

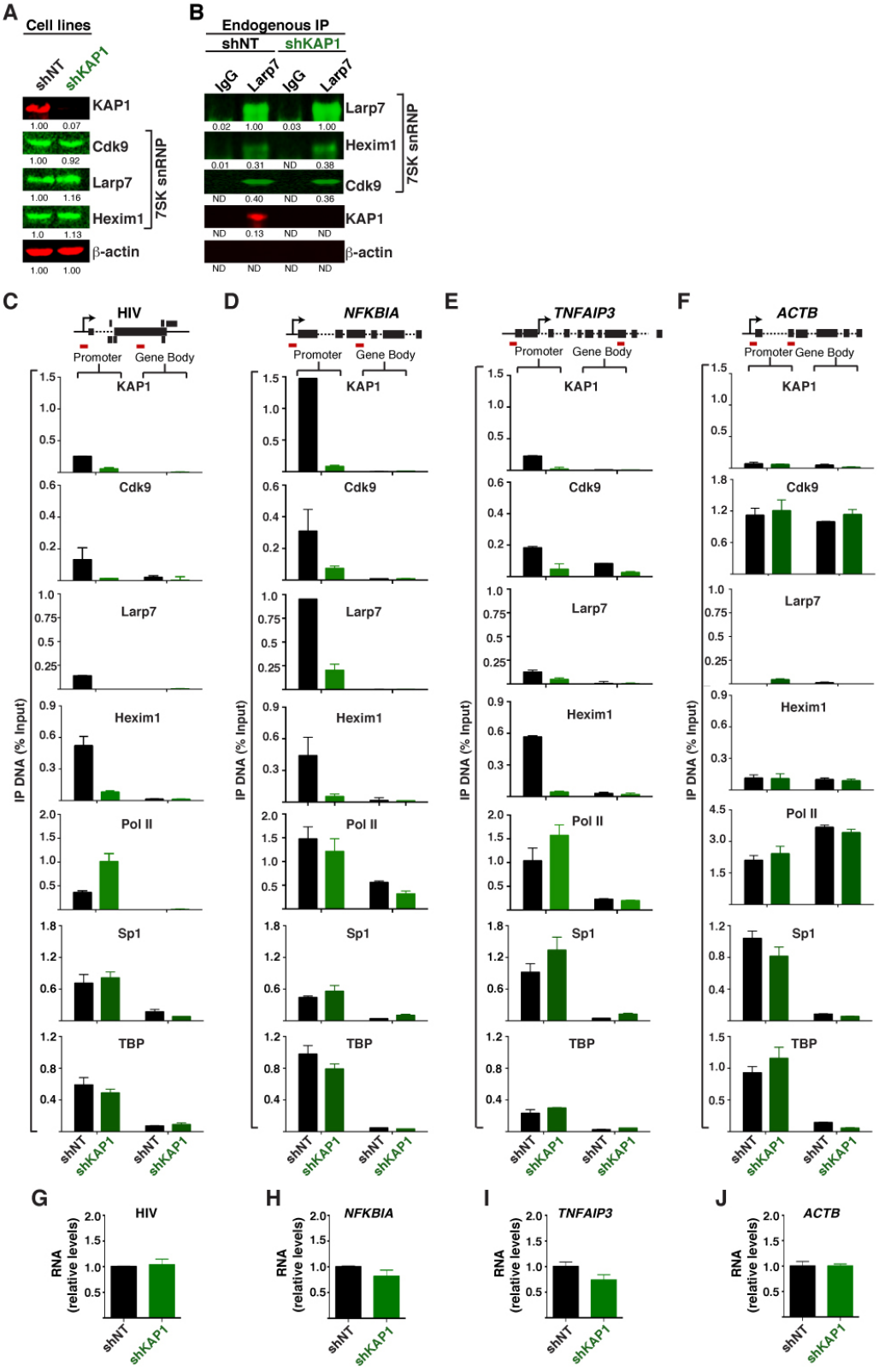
A
Cell lines
shNT
shKAP1
KAP1
1.00 0.07
Cdk9
1.00 0.92
Larp7
1.00 1.16
Hexim1
1.0 1.13
β-actin
1.00 1.00
7SK snRNP
B
Endogenous IP
shNT
shKAP1
IgG
Larp7
IgG
Larp7
Larp7
0.02 1.00 0.03 1.00
Hexim1
0.01 0.31 ND 0.38
Cdk9
ND 0.40 ND 0.36
KAP1
ND 0.13 ND ND
β-actin
ND ND ND ND
7SK snRNP
C
HIV
D
NFKBIA
E
TNFAIP3
F
ACTB
Promoter
Gene Body
KAP1
Cdk9
Larp7
Hexim1
Pol II
Sp1
TBP
IP DNA (% Input)
shNT
shKAP1
G
HIV
H
NFKBIA
I
TNFAIP3
J
ACTB
RNA (relative levels)

Figure 19. KAP1 Mediates Recruitment of the 7SK snRNP Complex to Promoter-proximal Regions.

(A) Generation of stable Jurkat HIV E4 cell lines transduced with non-target (shNT) or KAP1-directed (shKAP1) shRNAs. Jurkat E4 were transduced with lentiviruses encoding NT or KAP1 shRNAs and KD efficiency was monitored through Western blot with the indicated antibodies. Values underneath bands represent relative protein levels standardized to β-actin.

(B) KAP1 KD does not affect levels of Larp7-bound 7SK snRNP components. Lysates from panel (A) were immunoprecipitated using endogenous Larp7 or normal IgG serum as control to monitor if the 7SK snRNP was disrupted in response to KAP1 KD. Values underneath bands represent amount of IP proteins relative to input from panel (A).

(C) ChIP assay to analyze the distribution of KAP1, 7SK snRNP components (CDK9, Larp7 and Hexim1), Pol II and PIC (Sp1 and TBP) at the HIV locus in the shNT and shKAP1 cell lines from panel (A). The position of the amplicons used in ChIP-qPCR (Promoter and Gene Body) is shown with the schematic of the HIV genome and are the same used in **Figure 18A** (+141 and +2627). Values represent the average of three independent biological replicates minus a non-specific IgG control (mean ± SEM are shown, n = 3).

(D) ChIP assay to analyze the distribution of KAP1, 7SK snRNP components (CDK9, Larp7 and Hexim1), Pol II and PIC (Sp1 and TBP) at the *NFKBIA* locus in the shNT and shKAP1 cell lines from panel (A) (mean ± SEM are shown, n = 3). Same as in panel (C).

(E) ChIP assay to analyze the distribution of KAP1, 7SK snRNP components (CDK9, Larp7 and Hexim1), Pol II and PIC (Sp1 and TBP) at the *TNFAIP3* locus in the shNT and shKAP1 cell lines from panel (A) (mean ± SEM are shown, n = 3). Same as in panel (C).

(F) ChIP assay to analyze the distribution of KAP1, 7SK snRNP components (CDK9, Larp7 and Hexim1), Pol II and PIC (Sp1 and TBP) at the *ACTB* locus in the shNT and shKAP1 cell lines from panel (A) (mean ± SEM are shown, n = 3). Same as in panel (C).

(G) Relative HIV transcripts (+2627) levels quantified by qRT-PCR (mean ± SEM are shown, n = 3).

(H) Same as in panel (H) for *NFKBIA*.

(I) Same as in panel (H) for *TNFAIP3*.

(J) *ACTB* transcript levels were standardized to 1 for each cell line.

For panels G–I, all transcript levels were standardized to *ACTB* levels.

Figure 20

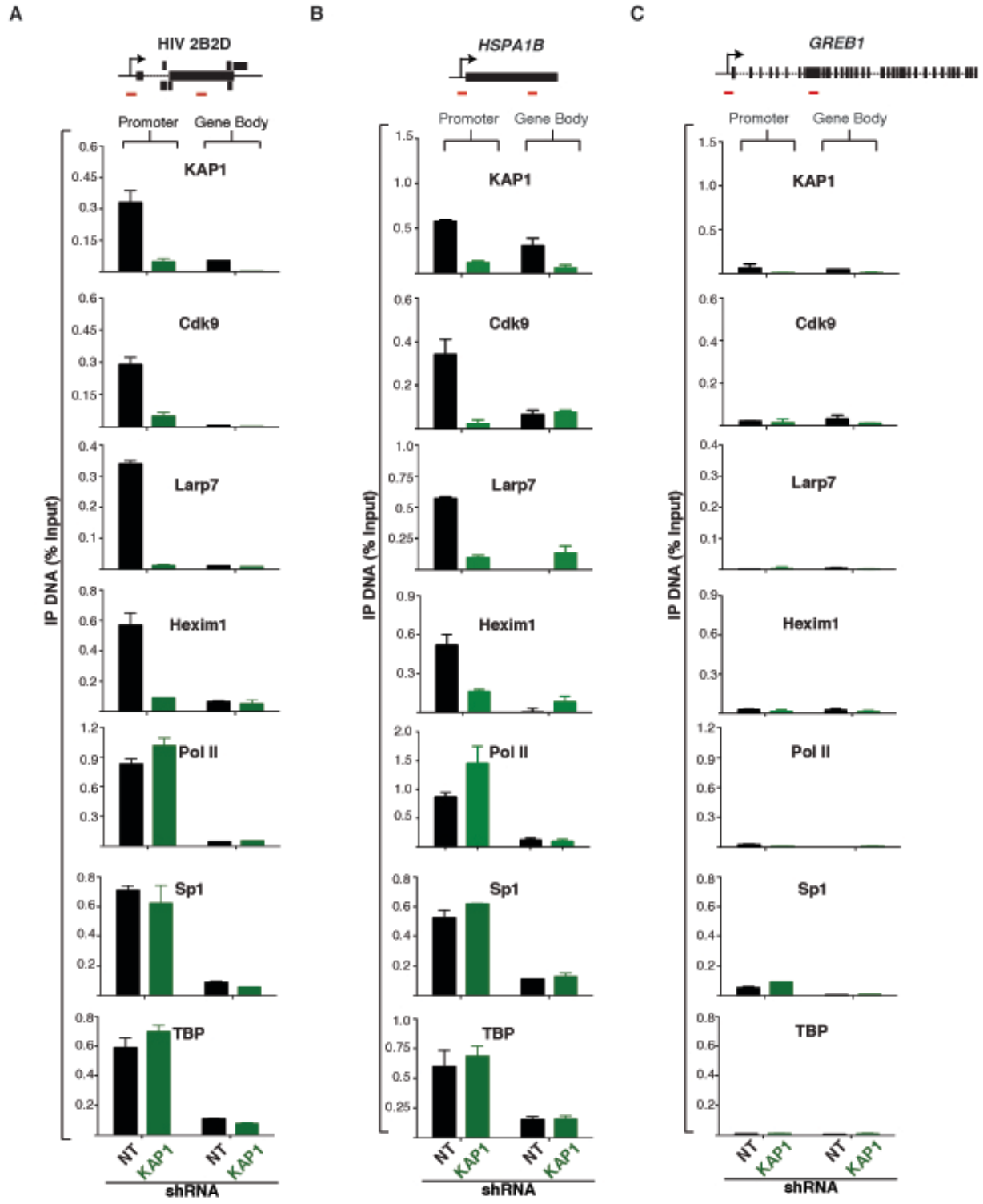
A
HIV 2B2D
Promoter
Gene Body
KAP1
Cdk9
Larp7
Hexim1
Pol II
Sp1
TBP
IP DNA (% Input)
NT
KAP1
NT
KAP1
shRNA
B
HSPA1B
Promoter
Gene Body
KAP1
Cdk9
Larp7
Hexim1
Pol II
Sp1
TBP
IP DNA (% Input)
NT
KAP1
NT
KAP1
shRNA
C
GREB1
Promoter
Gene Body
KAP1
Cdk9
Larp7
Hexim1
Pol II
Sp1
TBP
IP DNA (% Input)
NT
KAP1
NT
KAP1
shRNA

Figure 20. ChIP Assays on the HIV Clone 2B2D and HSPA1B, and the Negative Control GREB1.

(A) Jurkat HIV (*Tat null* 2B2D clone) NT (shNT) and KAP1 (shKAP1) cell lines were used for ChIP assays in unstimulated conditions. HIV promoter and gene body amplification locations are shown highlighted in red below to the gene schemes.

(B) NT and KAP1 shRNA cell lines were used for ChIP in unstimulated conditions for *HSPA1B*. Same as in panel (A).

(C) NT and KAP1 shRNA cell lines were used for ChIP in unstimulated conditions for *GREB1*. Same as in panel (A).

Figure 21

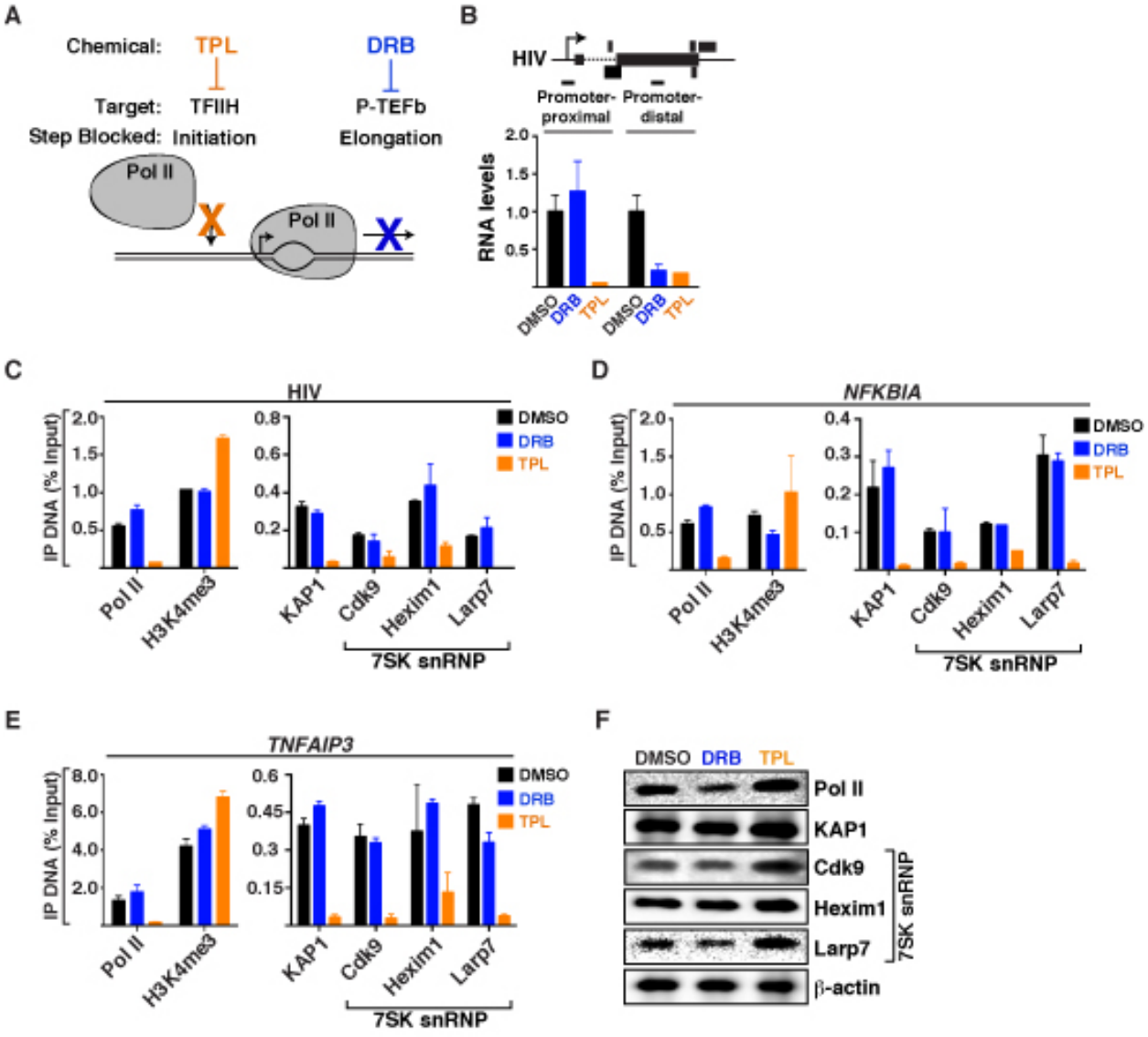

Figure 21. Blockage of Pol II and PIC Assembly at Gene Promoters Impairs KAP1 and 7SK snRNP Recruitment.

(A) Schematic representation of the mechanisms of transcriptional arrest by TPL and DRB. TPL halts transcription at the initiation stage by blocking the TFIIH helicases while DRB blocks transcription elongation by blocking P-TEFb.

(B) TPL inhibits the synthesis of promoter-proximal HIV transcripts. The Jurkat HIV E4 clone was treated with the indicated drugs (1 μM) for 1 hour and HIV promoter-proximal (+141) and -distal (+2627) transcripts were quantified by qRT-PCR and standardized to *ACTB* (mean ± SEM are shown; n = 3).

(C) ChIP assay to analyze the effect of DRB and TPL in the placement of KAP1 and the 7SK snRNP complex at the HIV promoter in the E4 clone. Cells were treated with DMSO (vehicle control), DRB, or TPL in the same conditions as in panel (B). ChIP assays were performed with the indicated antibodies and qPCR with the 5'-LTR amplicon (+141). Values represent the average of three independent experiments (mean ± SEM are shown; n = 3).

(D) ChIP-qPCR analysis of the *NFKBIA* promoter-proximal region as in panel (C).

(E) ChIP-qPCR analysis of the *TNFAIP3* promoter-proximal region as in panel (C).

(F) Treatment with DRB and TPL does not alter the stability of Pol II, KAP1 and 7SK snRNP components. Protein lysates from panel (C) were used for Western blot analysis with the indicated antibodies to test whether the drug treatments resulted in protein degradation.

Figure 22

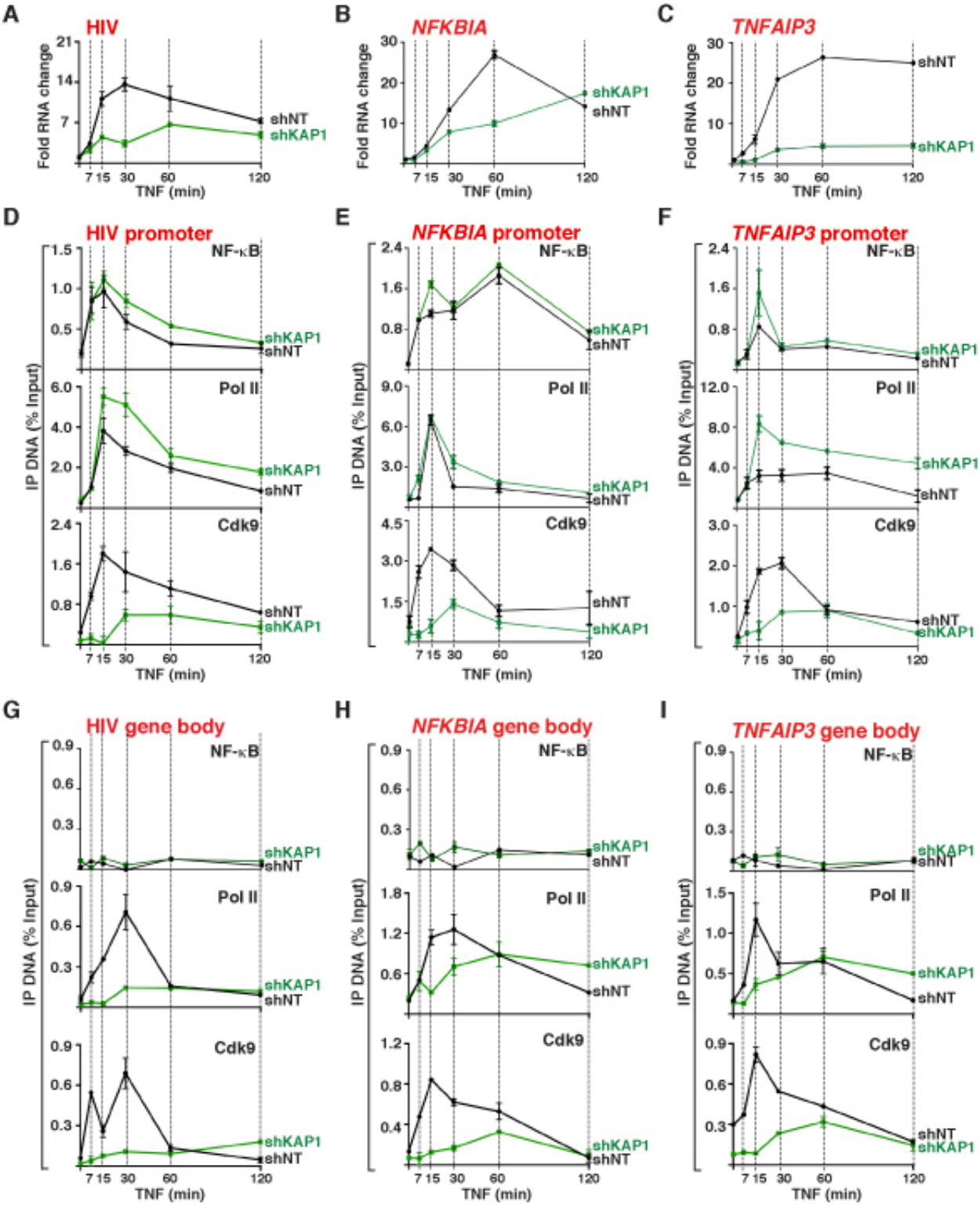
A
HIV
Fold RNA change
21
14
7
shNT
shKAP1
7 15 30 60 120
TNF (min)
B
NFKBIA
Fold RNA change
30
20
10
shKAP1
shNT
7 15 30 60 120
TNF (min)
C
TNFAIP3
Fold RNA change
30
20
10
shNT
shKAP1
7 15 30 60 120
TNF (min)
D
HIV promoter
IP DNA (% Input)
NF-κB
1.5
1.0
0.5
shKAP1
shNT
Pol II
6.0
4.0
2.0
shKAP1
shNT
Cdk9
2.4
1.6
0.8
shNT
shKAP1
7 15 30 60 120
TNF (min)
E
NFKBIA promoter
IP DNA (% Input)
NF-κB
2.4
1.6
0.8
shKAP1
shNT
Pol II
9.0
6.0
3.0
shKAP1
shNT
Cdk9
4.5
3.0
1.5
shNT
shKAP1
7 15 30 60 120
TNF (min)
F
TNFAIP3 promoter
IP DNA (% Input)
NF-κB
2.4
1.6
0.8
shKAP1
shNT
Pol II
12.0
8.0
4.0
shKAP1
shNT
Cdk9
3.0
2.0
1.0
shNT
shKAP1
7 15 30 60 120
TNF (min)
G
HIV gene body
IP DNA (% Input)
NF-κB
0.9
0.6
0.3
shKAP1
shNT
Pol II
0.9
0.6
0.3
shKAP1
shNT
Cdk9
0.9
0.6
0.3
shKAP1
shNT
7 15 30 60 120
TNF (min)
H
NFKBIA gene body
IP DNA (% Input)
NF-κB
0.9
0.6
0.3
shKAP1
shNT
Pol II
1.8
1.2
0.6
shKAP1
shNT
Cdk9
1.2
0.8
0.4
shKAP1
shNT
7 15 30 60 120
TNF (min)
I
TNFAIP3 gene body
IP DNA (% Input)
NF-κB
0.9
0.6
0.3
shKAP1
shNT
Pol II
1.5
1.0
0.5
shKAP1
shNT
Cdk9
0.9
0.6
0.3
shNT
shKAP1
7 15 30 60 120
TNF (min)

Figure 22. Loss of KAP1 Antagonizes P-TEFb Recruitment to Gene Promoters thereby Reducing Transcriptional Pause Release and Gene Induction in Response to Stimulation.

(A) KAP1 KD antagonizes transcriptional induction of HIV (2B2D *Tat null* clone, see text). shNT (black) and shKAP1 (green) cell lines were used to monitor transcriptional induction of HIV at several time points after TNF treatment using qRT-PCR. Graphs show fold RNA change (+/- TNF) of HIV (+2627 amplicon) standardized to *ACTB* (mean ± SEM are shown, n = 3).

(B) KAP1 KD antagonizes transcriptional induction of *NFKBIA* by TNF. Same as in panel (A).

(C) KAP1 KD antagonizes transcriptional induction of *TNFAIP3* by TNF. Same as in panel (A).

(D) Promoter recruitment of CDK9 is antagonized in the shKAP1 cell line at the HIV promoter. shNT (black) and shKAP1 (green) cell lines were treated with TNF and ChIP assays were performed (NF-kB, Pol II, and CDK9) at the indicated time points using the HIV 5'-LTR (+141) promoter-specific amplicon (mean ± SEM, n = 3).

(E) CDK9 recruitment to the *NFKBIA* promoter is antagonized in shKAP1 cells. Same as in panel (D).

(F) CDK9 recruitment to the *TNFAIP3* gene promoter is antagonized in shKAP1 cells. Same as in panel (D).

(G) KAP1 KD antagonizes accumulation of CDK9 and Pol II in the *HIV* gene body. NT and KAP1 shRNA cell lines were treated with TNF and ChIP (NF-κB, Pol II and CDK9) was performed using HIV gene body specific primers (+2627 amplicon) (mean ± SEM, n = 3).

(H) KAP1 KD antagonizes accumulation of CDK9 and Pol II in the *NFKBIA* gene body. Same as in panel (G).

(I) KAP1 KD antagonizes accumulation of CDK9 and Pol II in the *TNFAIP3* gene body. Same as in panel (G).

Figure 23

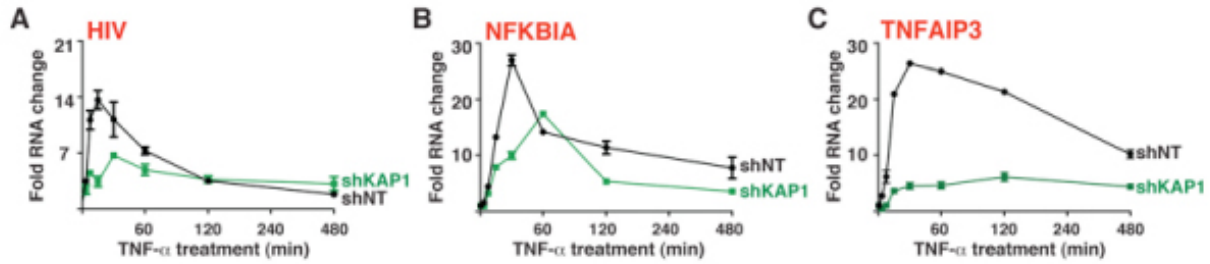

Figure 23. KAP1 KD Reduces Transcriptional Kinetics in Response to TNF.

(A) KAP1 KD (shKAP1 cell line) antagonizes transcriptional induction of HIV by TNF. Same as in **Figure 22A**, but the time course is extended to 480 min (mean ± SEM, n = 3).

(B) KAP1 KD (shKAP1 cell line) antagonizes transcriptional induction of *NFKBIA* by TNF. Same as in **Figure 22B**, but the time course is extended to 480 min.

(C) KAP1 KD (shKAP1 cell line) antagonizes transcriptional induction of *TNFAIP3* by TNF. Same as in **Figure 22C**, but the time course is extended to 480 min.

Figure 24

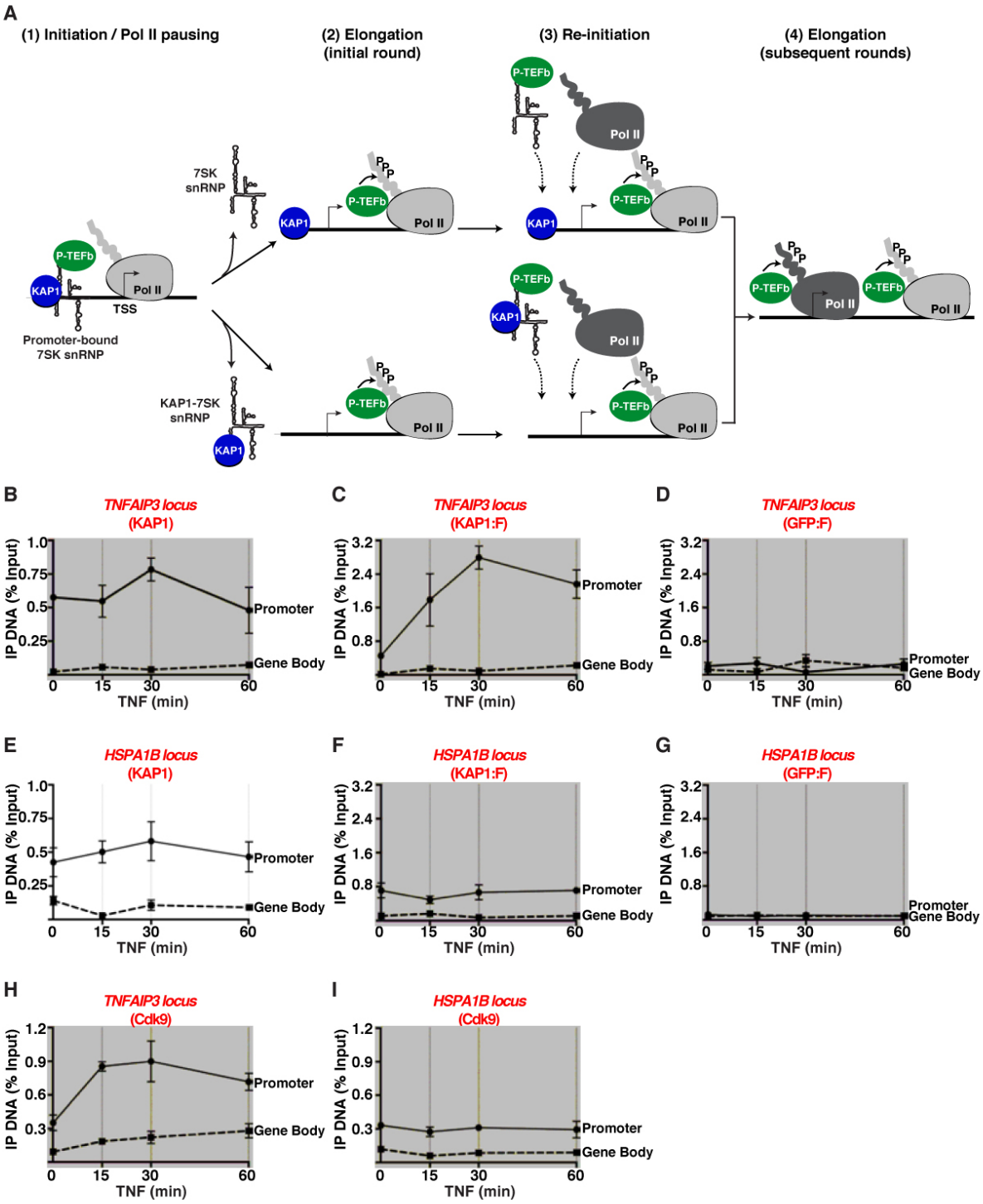
A
(1) Initiation / Pol II pausing
(2) Elongation (initial round)
(3) Re-initiation
(4) Elongation (subsequent rounds)
7SK snRNP
KAP1
P-TEFb
Pol II
TSS
Promoter-bound 7SK snRNP
KAP1-7SK snRNP
B
TNFAIP3 locus (KAP1)
C
TNFAIP3 locus (KAP1:F)
D
TNFAIP3 locus (GFP:F)
E
HSPA1B locus (KAP1)
F
HSPA1B locus (KAP1:F)
G
HSPA1B locus (GFP:F)
H
TNFAIP3 locus (Cdk9)
I
HSPA1B locus (Cdk9)
IP DNA (% Input)
TNF (min)
Promoter
Gene Body

Figure 24. KAP1 and P-TEFb are Continuously Recruited to the HIV Promoter in Response to Stimulation.

(A) Scheme of the two potential mechanisms of KAP1 and P-TEFb/7SK snRNP recruitment to promoters upon stimulation. See text for full description.

(B) HEK 293 T-REx cells were treated with TNF for the indicated times and ChIP (KAP1) was performed to analyze occupancy levels at the *TNFAIP3* locus.

(C) HEK 293 T-REx cells stably expressing KAP1:FLAG were treated with TNF for the indicated time points and ChIP (FLAG) was performed to analyze occupancy levels at the *TNFAIP3* locus.

(D) HEK 293 T-REx cells stably expressing GFP:FLAG were treated with TNF for the indicated time points and ChIP (FLAG) was performed to analyze occupancy levels at the *TNFAIP3* locus.

(E) HEK 293 T-REx cells treated with TNF for the indicated times and ChIP (KAP1) was performed to analyze occupancy levels at the *HSPA1B* locus.

(F) HEK 293 T-REx cells stably expressing KAP1:FLAG were treated with TNF for the indicated time points and ChIP (FLAG) was performed to analyze occupancy levels at the *HSPA1B* locus.

(G) HEK 293 T-REx cells stably expressing GFP:FLAG were treated with TNF for the indicated time points and ChIP (FLAG) was performed to analyze occupancy levels at the *HSPA1B* locus.

(H) HEK 293 T-REx cells treated with TNF for the indicated time points and ChIP (CDK9) was performed to analyze occupancy levels at the *TNFAIP3* locus.

(I) HEK 293 T-REx cells treated with TNF for the indicated time points and ChIP (CDK9) was performed to analyze occupancy levels at the *HSPA1B* locus.

Figure 25

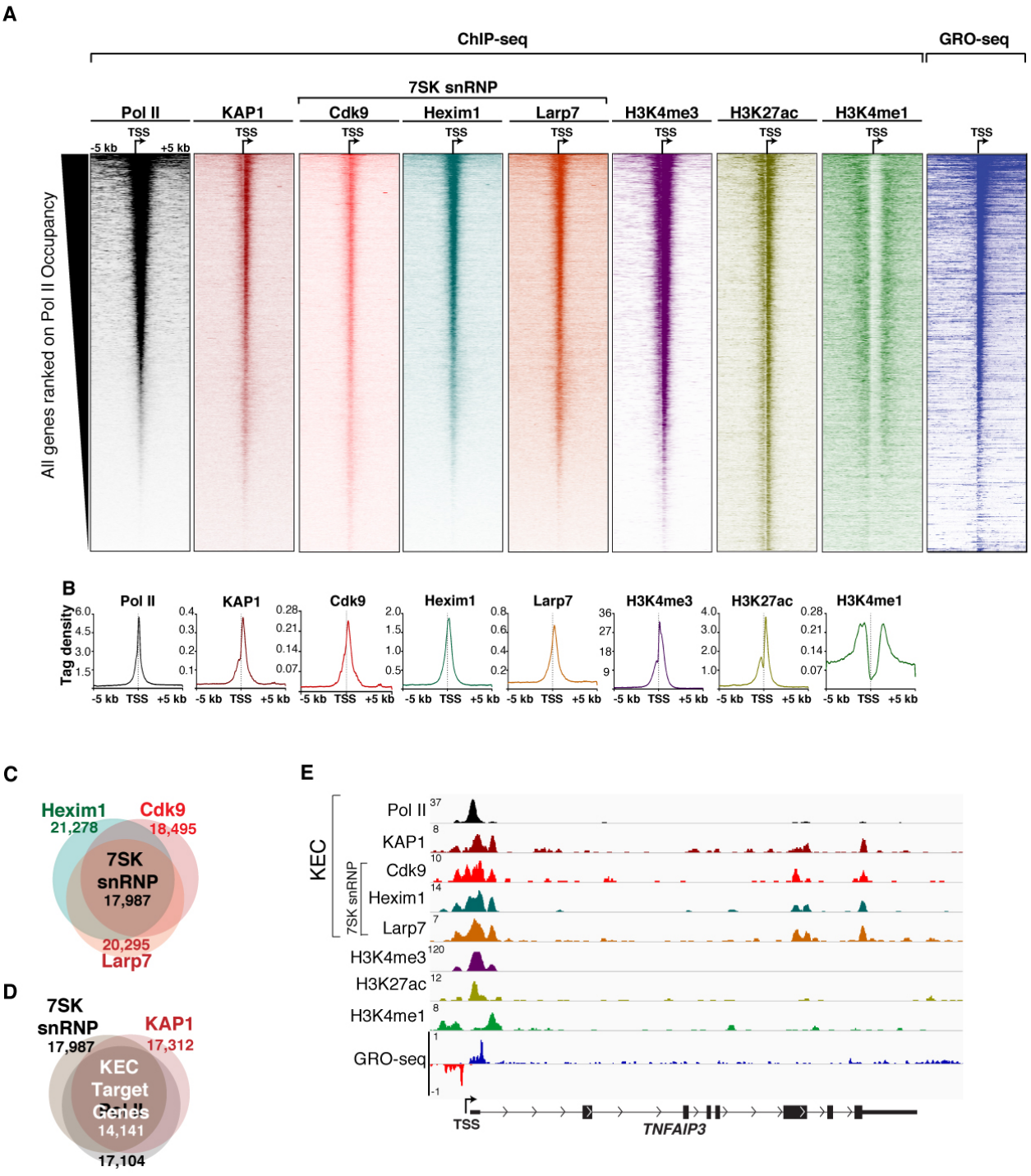

Figure 25. KAP1 and the 7SK snRNP Co-occupy Most Genes Containing Transcriptionally Engaged Pol II.

(A) Heat map representation of factor distribution relative at all Refseq genes (30,181) relative to the TSS in the HCT116 cell line. Genes are ranked based on decreased Pol II occupancy. GRO-seq dataset aligned to the Pol II ChIP-seq dataset. All heat maps are clustered to Pol II.

(B) Metagene analysis showing the distribution of each marker relative to the TSS.

(C) Identification of genes marked by the 7SK snRNP in the human genome. The number of genes containing Hexim1, Larp7, and CDK9 (7SK snRNP) at promoter-proximal regions (±250-bp threshold from the center of each peak) was calculated. Only genes containing all three proteins within this threshold were referred to as 7SK snRNP target genes.

(D) Identification of genes marked by the 7SK snRNP, KAP1 and Pol II. The number of genes containing the 7SK snRNP (from panel (C)), KAP1, and paused Pol II at promoter-proximal regions (±250-bp threshold from the center of each peak) was calculated. Genes containing all three markers were referred to as KEC (KAP1-7SK snRNP early Elongation Complex) target genes.

(E) Genome browser views of ChIP-seq at the *TNFAIP3* locus indicate presence of the KEC, paused Pol II and active chromatin signatures (H3K4me3 and H3K27Ac) at the promoter-proximal region. The GRO-seq track shows the position of the sense (blue) and antisense transcripts.

Figure 26

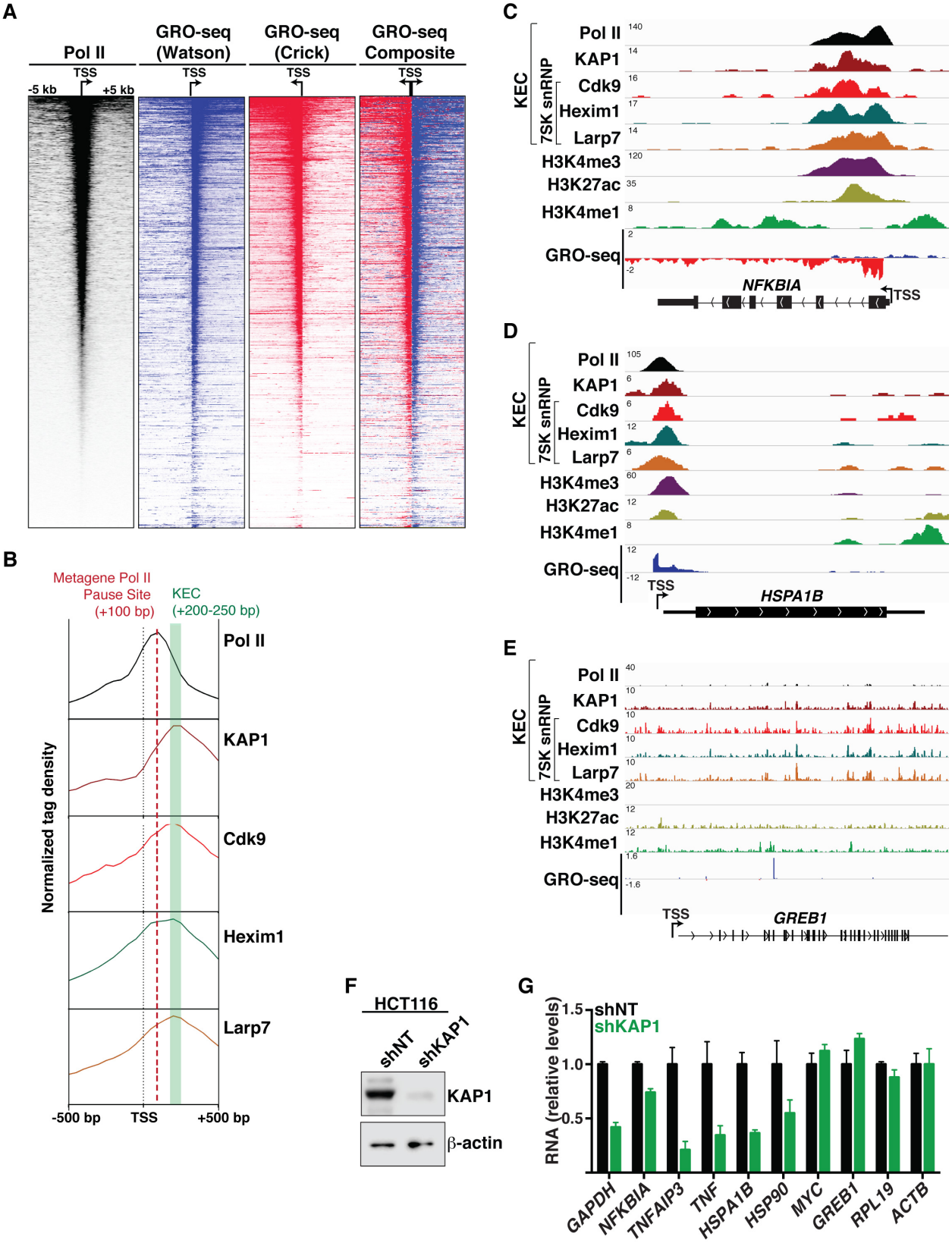
A
Pol II
GRO-seq (Watson)
GRO-seq (Crick)
GRO-seq Composite
TSS
-5 kb
+5 kb
B
Metagene Pol II Pause Site (+100 bp)
KEC (+200-250 bp)
Normalized tag density
Pol II
KAP1
Cdk9
Hexim1
Larp7
-500 bp
TSS
+500 bp
C
KEC
7SK snRNP
Pol II
KAP1
Cdk9
Hexim1
Larp7
H3K4me3
H3K27ac
H3K4me1
GRO-seq
NFKBIA
TSS
D
HSPA1B
E
GREB1
F
HCT116
shNT
shKAP1
KAP1
β-actin
G
RNA (relative levels)
shNT
shKAP1
GAPDH
NFKBIA
TNFAIP3
TNF
HSPA1B
HSP90
MYC
GREB1
RPL19
ACTB

Figure 26. The KEC Occupies Promoter-proximal Regions of Transcriptionally Active Genes.

(A) Heatmap representation of Pol II distribution at all Refseq genes (30181) relative to the TSS in the HCT116 cell line and GRO-seq data clustered to decreased Pol II occupancy. The GRO-seq sense strand (Watson, blue), antisense strand (Crick, red) and composite (Watson and Crick) information are shown.

(B) Metagene plots (±500-bp respective to the TSS) for Pol II, KAP1, Cdk9, Hexim1, and Larp7. The dashed lines indicate the summit of the metagenome peak for Pol II (+100-bp from the TSS) and for KAP1 and the 7SK snRNP components (+200-250-bp from the TSS).

(C) Genome browser view of ChIP-seq for the KEC target gene *NFKBIA* with histogram profiles of the indicated proteins along with GRO-seq RNA profile. *NFKBIA* is a transcriptionally active gene that was identified as containing an assembled KEC at the promoter-proximal region by ChIP q-PCR.

(D) Genome browser view of ChIP-seq for the KEC target gene *HSPA1B* with histogram profiles of the indicated proteins along with GRO-seq RNA profile. *HSPA1B* is a transcriptionally active gene that was identified as containing an assembled KEC at the promoter-proximal region by ChIP q-PCR.

(E) Genome browser view of ChIP-seq for the non-KEC target gene *GREB1* with histogram profiles of the indicated proteins along with GRO-seq RNA profile. *GREB1* is a transcriptionally inactive gene lacking detectable KEC assembly at the promoter-proximal region by ChIP qPCR in.

(F) Western blot of the HCT116 shNT and shKAP1 cell lines with the indicated antibodies shows efficient KAP1 KD.

Figure 27

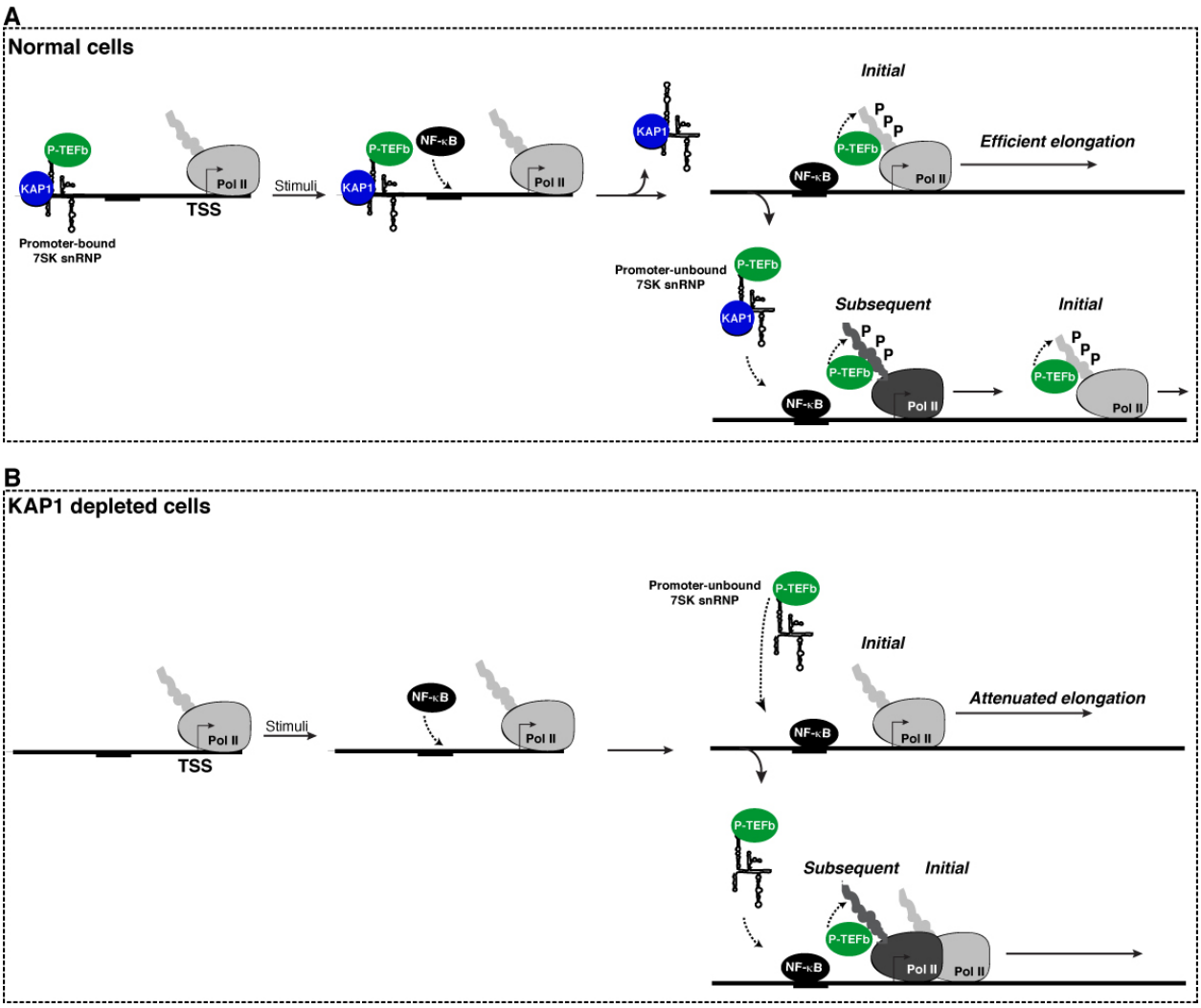

Figure 27. Model for the Role of the KAP1-7SK snRNP Complex in Transcriptional Pause Release.

(A) The KEC is recruited to transcriptionally paused genes. In response to stimulation, inducible TFs such as NF-κB recognize their target promoters and capture P-TEFb released from the KEC. KAP1 and the inhibitory 7SK snRNP components are evicted from the promoter. These molecular events allow for the initial round of elongation through P-TEFb–mediated

phosphorylation (P) of Pol II (light gray Pol II). The promoter-bound KEC facilitates rapid elongation (initial round). Additional incoming pools of nucleoplasmic KEC provide a continuous source of P-TEFb that phosphorylates newly incorporated Pol II (dark gray Pol II) to promote productive elongation (subsequent rounds).

(B) Loss of KAP1 results in the elimination of the KEC from promoters, antagonizing both the initial and subsequent recruitment of P-TEFb to the promoter, thereby attenuating Pol II elongation, thus affecting the magnitude of RNA synthesis in response to stimulation.

Table 2. List of high confidence interactors of the Larp7:S AP

Description	Protein	Peptide sequences	Coverage (%)	Molecular weight (kDa)
Larp7	AAH66945	87	81.80	67
hnRNPM	Q6P2D7	59	77.30	30
MePCE	Q7L2J0	47	67.80	74
DDX21	Q9NR30	54	60.90	87
SART3	Q15020	70	62.20	110
IMA1	P52292	27	55.40	58
Cyclin T1	O60563	36	54.80	81
Cdk9	P50750	21	47.80	43
Hexim1	O94992	27	74.90	41
YBOX1	P67809	20	73.10	36
KAP1	Q13263	46	68.70	89

CHAPTER V

CONCLUSIONS FROM CHAPTER III

The compartmentalization of factors enables rapid coordination of molecular reactions (Francastel et al., 2000; Hager et al., 2009). Transcription is one example of molecular crowding where a paused Pol II, negative elongation factors, and co-activators are collectively assembled at gene promoters to facilitate rapid gene expression. P-TEFb is a pleiotropic co-activator of gene expression initially thought to be recruited to promoters upon gene activation. Recent studies, however, revealed a close association between P-TEFb as part of the inhibitory 7SK snRNP at gene promoters, and this association is positively correlated with the degree to which Pol II pauses to prevent spurious transcription elongation (D'Orso and Frankel, 2010a; Ji et al., 2013; Lenasi and Barboric, 2010). Thus, the 7SK snRNP complex is not simply a reservoir of inactive P-TEFb that roams the nucleus but rather keeps a primed P-TEFb kinase at promoters of inducible genes.

Our findings elucidate a functional link between the recruitment of PPM1G by transcription factors and the release of P-TEFb from the promoter-assembled 7SK snRNP during gene activation. PPM1G dephosphorylates the CDK9 T loop to enzymatically disassemble the snRNP during activation of the HIV and cellular transcriptional programs by Tat and NF-κB, respectively. Other cellular phosphatases have been previously described to dephosphorylate P-TEFb *in vitro* and to dissociate Hexim1 from P-TEFb in response to signal-dependent activation (Ammosova et al., 2005; Chen et al., 2008; Wang et al., 2008); however, the mechanisms of transcription activation *in vivo* are poorly understood. Here, we demonstrate that Tat and NF-κB

recruit PPM1G to the promoter-assembled 7SK snRNP to release P-TEFb from the snRNP providing a functional link between CDK9 dephosphorylation and activation of specific elongation programs.

At the HIV promoter, the Tat-TAR interaction plays a key role in the disassembly of the 7SK snRNP *in vivo*, and thus, the nascent RNA likely coordinates the recruitment of the Tat-PPM1G complex to the HIV promoter to release P-TEFb from the 7SK snRNP. This model is supported through the observation that TAR deletion impairs the eviction of the 7SK snRNP even in the presence of Tat (D'Orso and Frankel, 2010a). Our results are consistent with a model whereby, upon disassembly of the 7SK snRNP, Tat captures released P-TEFb in an unphosphorylated state and enhances the rate of CDK9 T loop auto-phosphorylation to induce a permissive kinase state for TAR binding. Several previous observations are consistent with this model of activation: (1) Tat contacts the CDK9 T loop in the Tat-P-TEFb structure (Tahirov et al., 2010) and (2) phosphorylation of the T loop is critical for the folding and auto-phosphorylation of its CTD, a prerequisite for the assembly of Tat:TAR:P-TEFb (Baumli et al., 2008; Garber et al., 2000). Thus, these results suggest that the local transfer of P-TEFb from 7SK to TAR at the viral promoter follows a stepwise CDK9 dephosphorylation and rephosphorylation process.

Similar to Tat, we show that NF-κB utilizes P-TEFb (Barboric et al., 2001a; Hargreaves et al., 2009; Nissen and Yamamoto, 2000) and PPM1G during activation of the inflammatory transcriptional program. The inducible recruitment of PPM1G by NF-κB to promoters of inflammatory-responsive genes disassembles the 7SK snRNP and releases P-TEFb for transcription elongation. We provide evidence that KD of PPM1G impairs disassembly of the snRNP, thereby decreasing the transcription of inflammatory-responsive genes and sensitizing

cells to cytokine-induced apoptosis. The inducible disassembly of the 7SK snRNP at the IL-8 gene mimics the Tat-mediated snRNP eviction at the HIV promoter during the switch to elongation (D'Orso and Frankel, 2010a). Although NF-κB operates in a TAR-independent manner, we cannot rule out the scenario in which binding of NF-κB to its responsive elements and cooperation with RNA-processing factors that bind the nascent RNA (Ji et al., 2013) plays a similar role to Tat. At difference to the IL-8 gene, which requires P-TEFb to be induced during TNF signaling, the *NFKBIA* gene appears to be activated in a P-TEFb-independent manner in lung- and kidney-derived cell lines (Amir-Zilberstein et al., 2007; Luecke and Yamamoto, 2005). Because this assay was performed in HeLa, it is unclear whether the discrepancy in the requirement of P-TEFb is due to cell-type-specific differences.

PPM1G assembles into the 7SK snRNP in a reversible- and stimulus-dependent manner through direct contacts with both the RNA and core snRNP components. Because PPM1G binds the snRNP tightly, in the absence of P-TEFb and Hexim1, the enzyme can also block the reassociation of P-TEFb-Hexim1 onto the 7SK RNA, thereby maintaining a local pool of free P-TEFb (Barboric et al., 2007) to sustain multiple rounds of elongation. When Tat levels or the inflammatory stimuli are decreased, PPM1G can dissociate from the 7SK snRNP, thereby allowing for the recruitment of P-TEFb and Hexim1 into the snRNP to reset the system to the basal, repressed state. Thus, we propose that the reversible nature of the PPM1G-7SK snRNP interaction facilitates both the initiation and termination of the elongation programs.

CONCLUSIONS FROM CHAPTER IV

Several studies have documented that the 7SK snRNP complex is recruited to gene promoters (Cherrier et al., 2013; D'Orso and Frankel, 2010a; D'Orso et al., 2012; Ji et al., 2013). However, the molecular basis of this recruitment mechanism and its precise role in the transcriptional cycle has yet to be reconciled. In this study, we have established that the KAP1 protein physically recruits the 7SK snRNP (as part of the KEC) to the integrated HIV 5'-LTR and most cellular gene promoters containing transcriptionally engaged Pol II. Mechanistically, the KEC is assembled on chromatin where KAP1 recruits the 7SK snRNP by directly contacting the Larp7 subunit. Recruitment of the 7SK snRNP facilitates delivery of P-TEFb to promoters to induce Pol II phosphorylation, pause release and rapid activation of PRGs in response to environmental stimulation.

Besides controlling the rapid induction of PRGs, we have discovered that the KEC is placed at most Pol II regulated genes in the human genome. We propose that KAP1-mediated localized placement of inactive but primed P-TEFb kinase in the vicinity of promoter-proximal paused Pol II allows TFs to locally capture P-TEFb to couple transcription initiation with elongation. TFs like NF-κB bind P-TEFb to stimulate transcriptional elongation by Pol II (Barboric et al., 2001b). Although TNF stimulation triggers the recruitment of P-TEFb to NF-κB regulated genes, it remains poorly understood whether P-TEFb is tethered to DNA through interaction with the TF or through another mechanism. Our data suggest that there are two temporally distinct post-initiation events: 1. the transcription initiation-dependent recruitment of the KEC and 2. the further recruitment of KEC upon TF binding to its DNA responsive element. This model is consistent with the fact that, in the absence of stimulation, P-TEFb is already assembled at promoters as part of the KEC and without detectable TF loading. In summary, the primary function of the promoter-assembled 7SK snRNP is to keep a primed P-TEFb kinase.

Collectively, we propose that the 7SK snRNP, along with paused Pol II, be viewed as a mechanism for rapid and inducible gene expression through transcriptional pause release.

The KEC is recruited to genes promoters containing transcriptionally engaged Pol II, but not to inactive genes. In agreement with this model, chemical inhibition of initiation, which interferes with formation of the open complex, antagonizes KEC recruitment to gene promoters, as well as Pol II and the entire PIC. This is further exemplified through the use of multiple clonal cell lines containing a single HIV integrant in which only viral genomes with transcriptionally engaged Pol II (active paused genomes), but not transcriptionally inactive HIV genomes, show evidence of KEC occupancy. Noteworthy, KAP1 recruitment requires a nucleosome free region in the promoter-proximal region and ongoing transcription initiation because proviruses integrated in silent genomic regions (heterochromatin) or transcriptionally inactive genes in the human genome do not show evidence of transcriptionally engaged Pol II and KEC deposition at promoter-proximal regions.

Despite these functional insights supporting the role for the KEC in transcriptional pause release, one previous study reported that the 7SK snRNP complex occupies distal enhancers and plays an important role in transcriptional pause release through a specialized set of chromatin-modifying enzymes (Liu et al., 2013). We thus propose that the KEC functions from multiple genomic domains (promoters and enhancers) to properly control gene activation. Therefore, it would be interesting to examine the role of the KEC in controlling the transcription of non-coding RNAs from enhancers and other promoter-distal units as well as gene looping. Notwithstanding, these results provide compelling evidence of a novel transcriptional regulatory network that controls the rapid transition from transcription initiation into elongation by assembling at gene promoters.

Historically, KAP1 has been functionally linked to epigenetic silencing in embryonic stem cells through the deposition of repressive chromatin marks (H3K9me3) and formation of a heterochromatic environment in developmental genes and retroelements (Rowe et al., 2010; Wolf and Goff, 2007). However, our findings indicate a novel function of KAP1 in controlling the rapid induction of PRGs through the 7SK snRNP complex and probably transcription elongation control genome-wide. In the case of the transcriptional silencing function, KAP1 is recruited through interaction with the tetrapod-restricted family of KRAB zinc finger proteins (KRAB-ZFPs) (Iyengar and Farnham, 2011a; Iyengar et al., 2011b). Therefore, it would be informative for the field to further define the mode of KAP1 recruitment to promoter-proximal regions of all KEC target genes. Recently, Bunch *et al*. reported that KAP1 appears to have DNA-binding activity (Bunch et al., 2014). Additionally, KAP1 contains a bromodomain that can contact acetylated lysine residues on histones (Iyengar and Farnham, 2011b), indicating that there are multiple mechanisms for KAP1's recruitment onto chromatin. Thus, further examination of these properties is needed to precisely ascertain the recruitment mechanism.

Recent work has indicated that the loss of KAP1 triggered a subtle increase in RNA levels for a subset of genes in basal conditions (Bunch et al., 2014). Based on this data the authors hypothesized a role for KAP1 in the control of Pol II pausing. Our data (both in basal and stimulated conditions) indicate that KAP1 is needed for transcription elongation at both most Pol II regulated genes. However, the most dramatic effect due to loss of KAP1 in gene expression was observed in the attenuation of PRGs in response to stimulation. While these findings focused on the stimulus-dependent transcriptional regulation by KAP1, the KEC is assembled at most genes containing transcriptionally engaged Pol II. Therefore, further work to define how maintains transcriptional homeostasis in the absence of stimulation is needed.

Consistent with our findings, another zinc-finger B-cell specific transcriptional regulator (CTIP2/Bcl11b) also has been implicated in anchoring the 7SK snRNP to gene promoters regions (Cherrier et al., 2013), but the functional consequences were not investigated. Hence, this supports a model in which multiple cell-type specific transcriptional regulators recruit the 7SK snRNP to a diverse set of target genes to regulate cell fate choice, differentiation, and/or development. KAP1 belongs to a family of transcriptional regulators known as transcription intermediary factor 1 (TIF1) composed by three members: TIF1α/TRIM24, TIF1β/TRIM28 and TIF1γ/TRIM33. The TIF1γ member controls erythroid cell fate by regulating transcription elongation of a selective class of genes through P-TEFb (Bai et al., 2010; Rowe et al., 2010). Although it is unknown whether TIF1γ interacts with the 7SK snRNP to indirectly recruit P-TEFb at gene promoters, TIF1γ's strikingly high homology to KAP1 would support a similar mechanism of interaction (Iyengar and Farnham, 2011b). Therefore, further work on this arena will enhance our understanding of the TIF1/TRIM family in the activation of various transcriptional programs through the 7SK snRNP complex.

In conclusion, we have discovered and characterized a novel complex, the KEC, which facilitates transcriptional pause release by recruiting the 7SK snRNP nearby promoter-proximal paused Pol II. This localized placement of transcriptional regulators allows the cell to rapidly respond to activating stimuli, a mechanism that HIV has hijacked for rapid viral activation and spread.

CONLUSIONS – OVERALL AND FUTURE DIRECTIONS

In this dissertation, I have described a novel transcriptional signaling program in which the PPM1G phosphatase is inducibly recruited to a promoter to enzymatically disassembly the 7SK snRNP. The 7SK snRNP, which is anchored onto chromatin through KAP1, is continuously recycled to the promoter upon stimulation. This ensures that P-TEFb is not only primed for local extraction for the initial round of transcription elongation, but that it is further recruited to activate subsequent rounds of elongation. While this phenomenon was initially characterized through use of HIV and a few other TNF-responsive genes (D'Orso and Frankel, 2010a; Eilebrecht et al., 2014; McNamara et al., 2013), we observed that the KEC was present at the majority of Pol II poised genes. Therefore, while my work has primarily focused on the stimulus-dependent activation of inducible genes, clearly the KEC plays a substantial role in basal transcription and cell homeostasis. For example, the "house keeping" gene GAPDH is classified as KEC target gene using our stringent thresholds. KAP1 KD resulted in small, but repeatable decrease in basal transcripts of GAPDH. Therefore, GRO-seq or NET-seq approaches need to be employed in the future to examine how the KEC regulates the high proportion (88%) of all Pol II poised genes during normal growth conditions (Churchill and Weissman, 2011; Core et al., 2008).

While we noted that the majority of the KEC was present near the Pol II pause site, we did also note that Pol II is also deposited at intragenic regions, such as intron/exon junctions. It is well established that Pol II's traveling speed through intronic/exonic junctions is changed to allow for co-transcriptional RNA processing (Bittencourt et al., 2012; Bentley et al., 2005). Given the KEC's presence at exons, it is very tempting to speculate that, in addition to regulating

Pol II's promoter pause release, the KEC is also regulating Pol II's intragenic transcribing speed to allow for co-transcriptional RNA processing. Supporting this model are the recent findings from Ji et al. detailing an intimate relationship between splicing factors and the 7SK RNA (Ji et al., 2013). Monitoring Pol II speed (ChIP-seq) along with RNA abundance (GRO-seq or Next-seq) will further address how the KEC functions in intragenic regions and its role in co-transcriptional RNA processing.

The rate limiting step of recruiting the elongation factor P-TEFb to the paused Pol II has long been known to be a rate-limiting step in gene expression of HIV (Mancebo et al., 1997; Zhu et al., 1997), however it was only after this characterization that the scientific community appreciated that the majority of human genes are regulated in a similar manner (Peterlin and Price, 2006). Therefore, the use of HIV as a model system to study transcription has yielded major advances in both our understanding of the pathogenesis of HIV as well as basic cellular biology. Throughout this dissertation, I have outlined how the employment of HIV as a model system has provided crucial insight into mechanisms of gene expression by both HIV and the eukaryotic cell. The viral transcription factor Tat, like many cellular transcription factors, acts as a scaffold to recruit co-activators such as PPM1G to their targeted promoters and release Pol II from its paused state (Barboric et al., 2001b; Barboric et al., 2000; Galbraith et al., 2013; Gomes et al., 2006; Hargreaves et al., 2009; Kayama et al., 2008; Lis et al., 2000; Mancebo et al., 1997; Rahl et al., 2010; Zhu et al., 1997). Therefore, the findings outlined throughout this dissertation provide further insight into how HIV transcriptionally regulates its genome as well as how the cell maintains transcriptional homeostasis. Future work using HIV as a model system will continue to progress our knowledge of cellular biology as well as highlight novel targets amenable for drug therapy. However we cannot discount that HIV, given its small genome, has

adapted to bypass certain restriction barriers to transcription elongation such as promoter placement of the 7SK snRNP by KAP1. While cellular transcriptional pathways such as activation through NF-κB are subject to promoter placement of the KEC, ongoing work by Morton et al. has thus far demonstrated that Tat can still potently activate the HIV genome in the absence of a fully assembled KEC at the promoter (Morton et al, unpublished data). Therefore further characterization of how Tat and cellular transcription factors contrast in activation of their target genes is critical for our view of how the KEC regulates gene expression.

In conclusion, the characterization of a novel transcriptional signaling program has opened the door to a number of subsequent hypotheses. I feel that there is much more to learn about the KEC, and how it regulates gene expression, both HIV and cellular. It is particularly interesting that the KEC was present at both inducible genes (such as *NFKBIA*, *TNFAIP3*, *HSPA1B*, etc.) as well as seemingly constitutively active genes (such as *GAPDH*, *RPL19*, etc.). The finding that it is at 88% of all Pol II poised genes strongly supports a model in which the KEC plays a crucial role in maintaining transcriptional homeostasis as well as the activation of inducible pathways. I greatly look forward to how the D'Orso lab, and other labs in the field of gene expression, further examines the role of the KEC in the human genome and how HIV has hijacked this complex for its own rapid activation. Furthermore, how the KEC influences Pol II pause release and transcription elongation throughout intragenic barriers (nucleosomes, intragenically bound factors, and intron-exon units) retention of introns, and coordination of splicing factors, should be addressed in the coming years.

CHAPTER VI

ACKNOWLEDGEMENTS

I would like the personally thank my advisor Prof. Ivan D'Orso for his thankless efforts into training me, an admittedly occasionally stubborn student. Through his guidance, I have grown into a far more mature scientist, and feel confident that I can design thoughtful experiments to test an important hypothesis. Throughout my time here as a student, Prof. D'Orso has continuously challenged me to pay critical attention to detail, all while keeping me focused on the larger impact of my research. While I am very proud of the things that we have accomplished together, I am certain that Prof. D'Orso's greatest scientific achievements are still yet to come.

I would like to thank my thesis committee, Prof. Nicholas Conrad, Prof. Melanie Cobb, and Dr. Jeffrey Kahn for their constructive guidance throughout my dissertation. They have continuously challenged me to broaden the scope of my work and to carefully consider how my results fit into multiple fields of study, from cellular signaling pathways to HIV pathogenesis.

I would like to thank the past and current members of the D'Orso lab for their contributions to both the work that I have done. I would like to thank Jennifer McCann who was a tremendous help, particularly at the beginning of my dissertation research. She optimized a number of procedures that I utilized throughout my work, and her efforts have been instrumental not only in my dissertation, but in the continuing work of the D'Orso lab. I would like to thank Dr. Swapna Aravind-Gudipaty for her guidance on biochemical and biophysical approaches to my work and their respective interpretations. I would like to thank Jonathan Reeder for his help with bioinformatics and program writing. Without his help, much of the ChIP-seq data that was

generated would have taken much longer to analyze. Also, his critical thinking greatly facilitated how we chose to present the data in a manner that would make sense to those outside of the field. I would like to thank Emily Morton for her help in a number of assays, as well as her insights into presentations. Emily was invaluable to me, and she will undoubtedly continue to be an asset to the lab. I would like to thank Curtis Bacon and Nora Guadalupe Ramirez for their insight into my dissertation and their input on to how to best convey my results to those outside of the field.

I would like to thank Dr. Paul Luethy and Dr. Lauren Luethy for their guidance and friendship over the past 6 years. In addition to being marvelous scientists, I may never find another couple that are a wonderful humanitarians as them. I would like to thank Dr. Ryan Downey for his support and friendship throughout these 6 years, particularly for his insight in presentations as well as constructive feedback into the writing of this dissertation. I would also like to thank Dr. Stefan Bresson for his constructive feedback on this project throughout his graduate school career.

I would like to thank Prof. Mark Lehrman, Prof. Pingui Feng, and Prof. Nicholas Conrad for allowing me to rotate in their laboratories. The techniques that I learned during my rotations were instrumental to my growth as a scientist. I would also like to thank all of the members of their respective laboratories for training me

I am very grateful to Dr. E. Burstein for the NF-κB plasmid and the AIDS Research and Reference Reagent Program for reagents. I would also like to thank the UTSW Proteomic and Live Cell Imaging Cores (particularly Abhjiit Bruedge) for their help in protein identifications and high resolution microscropy, respectively. I am also very grateful to Dr. Jonathan Karn and Dr. Eric Verdin for generously sharing the Jurkat T cell lines used in the second half of this dissertation. Research reported in this dissertation was supported by the National Institute Of

Allergy And Infectious Diseases of the NIH under Award Numbers R56AI106514 and R01AI114362, and The Welch Foundation grant I-1782 to Iván D'Orso; and Cell and Molecular Biology Training grant T32 GM008203 and the NIH training grant T32 2T32AI007520-16.

Lastly, and most of all, I would like to thank my father Prof. Bernie McNamara who stirred up my interest in science when I was a child. Although he being an astronomer characterized the exponentially large, while I being a microbiologist characterize the exponentially small, we are both genuinely interested in the love of science. He did not push me into one specific field of science, but rather challenged me to push myself into a field that interested me.

Table 3. Primers used in this dissertation

Gene/Amplicon	Primer Number/Sequence (5'-3')	Purpose/Assay
RPL19	**354**/ATCGATCGCCACATGTATCA **355**/GCGTGCTTCCTTGGTCTTAG	qRT-PCR
ACTB Promoter	**251**/TGCACTGTGCGGCGAAGC **252**/TCGAGCCATAAAAGGCAA	ChIP
ACTB Gene Body	**1256**/GATGATGATATCGCCGCGCT **1257**/CTTCTCGCGGTTGGCCTTGG	qRT-PCR/ChIP
TNFAPI3 Promoter	**1142**/CAGCCCGACCCAGAGAGTCAC **1143**/CTTGGCCCGCCACGAA	ChIP
TNFAIP3 Gene Body	**1144**/GCTGCTGCCTCAGGGAAAGTC **1145**/CTCTTCTGTCCTTTTGGCCTC	qRT-PCR/ChIP
NFKBIA Promoter	**1146**/AAGGCTCACTTGCAGAGGG **1147**/GGACTGCTGTGGGCTCTG	ChIP
NFKBIA Gene Body	**1148**/TCCTGAGCTCCGAGACTTTC **1149**/GTAGTTGGTAGCCTTCAGG	qRT-PCR/ChIP
HSPA1B Promoter	**1380**/GGAAAACGGCCAGCCTGAGGAGCTG **1381**/TGGAAACGGCGGACGGGATCCGC	ChIP
HSPA1B Gene Body	**1384**/ATCCTGATGGGGGACAAGTCCGAGAAC **1385**/CGGGTTGGTTGTCGGAGTAGGTGG	qRT-PCR/ChIP
GREB1 Promoter	**1390**/AAATAAAGTTTTTTCAATGGAAGGCT TGCAGCTCTTG	ChIP

	1391/CGCTCAGCAGAGACGAAGAAAGGG	
GREB1 Gene Body	**1394**/GCCATGGGAACTTCCCTTACCTCTG **1395**/CCAGCTGGACCAGGTAGTAGACGGT	qRT-PCR/ChIP
HIV (-353)	**1093**/AAGGCTACTTCCCTGAT **1094**/TAGCACCATCCAAAGGTC	ChIP
HIV (-69)	**1360**/CTTGCTACAAGGGACTT **1361**/AGGGCTCGCCACTCC	ChIP
HIV (-37)	**1364**/CTTTCTACAAGGGACTTTCCGCTG **1365**/CTCCCAGGCTCAGATCTGGTC	ChIP
HIV (+141) (5'LTR specific)	**1111**/GCTTAAGCCTCAATAAAGCTTGCCTT GAG **1112**/GTCCTGCGTCGAGAGATCTCCTCTG	qRT-PCR/ChIP
HIV (+2627)	**1358**/TGAGGGACAATTGGAGAAGTGA **1359**/TCTGCACCACTCTTCTCTTTGC	qRT-PCR/ChIP
HIV (+4230) (3'LTR specific)	**1368**/ACAAGAGGAGGAAGAGGTGGGT **1369**/GCCCTGGTGTGTAGTTCTGCCA	ChIP
Myc	**860**/ACAGCTACGGAACTCTTGTGCGTA **861**/CAGCCAAGGTTGTGAGGTTGCATT	qRT-PCR
HSP90	**838**/CCTTCTATTTGTCCCACG **839**/ATCCTCCGAGTCTACCAC	qRT-PCR
TNF	**950**/CCCAGGGACCTCTCTCTAATC **951**/ATGGGCTACAGGCTTGTCACT	qRT-PCR
GAPDH	**239**/CCCTGTGCTCAACCAGT	qRT-PCR

	240/CTCACCTTGACACAAGCC	
MePCE	**507**/TGAAGCCAGAGCAGTTCAGTTCCTA **508**/TACACAGGACGCTGGAAGCCTTTA	RNAi validation
Larp7	**511**/TCCGGGATACTTTGGCAGCAATCT **512**/AGGATCTCGAGTTTCCAGCAGTGT	RNAi validation
CDK9	**517**/GTGTTCGACTTCTGCGAGCATGAC **518**/CTATGCAGGATCTTGTTTCTGTGG	RNAi validation
SLC25A5	**519**/GCAGCTGATGTGGGTAAAGCTGG **520**/CCGGAAGCATTCCCTTTGCAGT	RNAi validation
NPM1	**525**/GTTGTGAACTAAAGGCCGACAAAG **526**/TGTGCAACTCATCCTTTGCACCAG	RNAi validation
hnRNP-F	**527**/TGTATTGGTCTCCTGCTCCTAGA **528**/AGGACTGGTTTCTGTTGCTACC	RNAi validation
DDX21	**531**/CTGGGTGTTTGCTTTGATGTACC **532**/AGTTCTGGTTGCTCTGTGG	RNAi validation
DHX57	**537**/CCGCTGTTTGGATCCTGCTCT	RNAi validation

	538/GCTGCCATCCCTTATACGCTTG	
SET	**539**/AATATAACAAACTCCGCCAACC **540**/CTGGTCAAATAATGCAGTGCCTC	RNAi validation
PPM1G	**543**/AAAATGGCAACAGCGACAAG **544**/CACCTCATACCCACTGCTA	RNAi validation
Sart3	**549**/CTTTACTCGTGCCTTGGAGTAT **550**/CCGAGCTTTCTGCATGTTATTG	RNAi validation
CycT1	**551**/GCAGACTTTAGGCTTTGAACTAAC **552**/TACTGCAGGCTAAATGTGGT	RNAi validation
7SK RNA	**15**/GGATGTGAGGGCGATCTG **16**/GGAGCGGTGAGGGAGGAAG	RIP assay, Northern blot
U6 snRNA	**9**/CTCGCTTCGGCAGCACATATAC **10**/GGAACGCTTCACGAATTTGCGTG	RIP assay
IL-8	**245**/GGGCCATCAGTTGCAAATC **246**/TTCCTTCCGGTGGTTTCTTC	Inflammatory response ChIP (Promoter)

IL-8	**282**/GCCATAAAGTCAAATTTAGCTGGAA **283**/GTGCTTCCACATGTCCTCACA	Inflammatory response ChIP (Gene body)

Table 4. Plasmids used in this dissertation

Insert	Plasmid Name	Restriction Site	Purpose
Larp7	pcDNA-4/TO-Strep	HindIII/XhoI	Transient Expression
GFP	pcDNA-4/TO-Strep	BamHI/XhoI	Transient Expression
KAP1	pcDNA-4/TO-Strep	HindIII/XhoI	Transient Expression
GFP	pcDNA-4/TO-FLAG	BamHI/XhoI	Transient/Stable Expression
KAP1	pcDNA-4/TO-FLAG	HindIII/XhoI	Transient/Stable Expression
NT shRNA	pLKO.1	AgeI/EcoRI	shRNA-mediated RNAi
KAP1 shRNA	pLKO.1	AgeI/EcoRI	shRNA-mediated RNAi
Tat	pcDNA-4TO / Strep	BamHI - XhoI	Transient Expression
Tat	pcDNA-4TO / FLAG	BamHI - XhoI	Transient Expression
Larp7	pcDNA-4TO / Strep	BamHI - XhoI	Transient Expression
Larp7	pcDNA-4TO / FLAG	BamHI - XhoI	Transient

			Expression
Sart3	pcDNA-4TO / Strep	HindIII - XhoI	Transient Expression
Sart3	pcDNA-4TO / FLAG	HindIII - XhoI	Transient Expression
CycT1	pcDNA-4TO / Strep	HindIII - EcoRI	Transient Expression
CycT1	pcDNA-4TO / FLAG	HindIII - EcoRI	Transient Expression
CDK9	pcDNA-4TO / Strep	HindIII - XhoI	Transient Expression
CDK9	pcDNA-4TO / FLAG	HindIII - XhoI	Transient Expression
MePCE	pcDNA-4TO / Strep	HindIII - XhoI	Transient Expression
MePCE	pcDNA-4TO / FLAG	HindIII - XhoI	Transient Expression
DDX21	pcDNA-4TO / Strep	BamHI - XhoI	Transient Expression
DDX21	pcDNA-4TO / FLAG	BamHI - XhoI	Transient Expression
NPM1	pcDNA-4TO / Strep	BamHI - XhoI	Transient Expression

NPM1	pcDNA-4TO / FLAG	BamHI - XhoI	Transient Expression
hnRNP-F	pcDNA-4TO / Strep	BamHI – XhoI	Transient Expression
hnRNP-F	pcDNA-4TO / FLAG	BamHI - XhoI	Transient Expression
SET	pcDNA-4TO / FLAG	NotI - XhoI	Transient Expression
GFP	pcDNA-4TO / Strep	BamHI - XhoI	Transient Expression
GFP	pcDNA-4TO / FLAG	XhoI - ApaI	Transient Expression
PPM1G	pcDNA-4TO / Strep	BamHI - NotI	Transient Expression
PPM1G	pcDNA-4TO / FLAG	BamHI – NotI	Transient Expression

Table 5. Antibodies used in this dissertation

Target	**Company**	**Catalogue Number**	**Assay**
FLAG	Sigma	F3165	Western blot
Strep-Tag II	Novagen	71591	Western blot
b-actin	Santa Cruz Biotechnologies	SC-47778	Western blot
Cyclin T1	Santa Cruz Biotechnologies	SC-8127X	Western blot
MePCE	Santa Cruz Biotechnologies	SC-82542	Western blot
RNA Pol II CTD	Active Motif	61081	Western Blot, ChIP
CDK9	Santa Cruz Biotechnologies	SC-8338X	Western Blot, ChIP, ChIP-seq
Hexim1	Abcam	AB25388	Western blot, Immunofluorescence, ChIP, ChIP-seq
Larp7	Sigma	AV40847	Western blot, Immunofluorescence, ChIP, ChIP-seq
CDK9 (C-20)	Santa Cruz Biotechnologies	SC-484	ChIP-seq
RNA Pol II (N-20)	Santa Cruz	SC-899X	ChIP-seq

	Biotechnologies		
RNA Pol II Phospho-Serine 5 CTD	Abcam	AB5408	ChIP-seq
H3K4me1	Abcam	AB8895	ChIP-seq
H3K27ac	Abcam	AB4729	ChIP-seq
NF-κB NF-κB (p65)	Santa Cruz Biotechnologies	SC-372X	ChIP
TATA Binding Protein (TBP)	Abcam	AB818	ChIP
Sp1	Santa Cruz Biotechnologies	SC-59X	ChIP
Normal Rabbit IgG	Santa Cruz Biotechnologies	SC-2027	ChIP
Normal Mouse IgG	Santa Cruz Biotechnologies	SC-2025	ChIP
Phospho Thr186 CDK9	Cell Signaling	2549	Immunofluorescence
Alexa Fluor 546 Anti-Mouse IgG	Life Technologies	A11030	Immunofluorescence
Alexa Fluor 488 Anti-Rabbit IgG	Life Technologies	A11034	Immunofluorescence
Anti-Mouse-HRP	Santa Cruz Biotechnologies	SC-2005	Western blot
Anti-Rabbit-HRP	Santa Cruz	SC-2313	Western blot

	Biotechnologies		
Anti-Goat-HRP	Santa Cruz Biotechnologies	SC-2020	Western blot
Anti-Mouse 680	Li-Cor	P/N 926-68022	Western blot
Anti-Rabbit 800	Li-Cor	P/N 926-32213	Western blot
CDK9 (C12F7)	Cell Signaling	2316	ChIP-seq
CycT1	Santa Cruz Biotechnologies	sc-10750	Western Blot, ChIP
RNA Polymerase II S2P-CTD	Abcam	ab5095	ChIP
PPM1G/PP2Cγ	Santa Cruz Biotechnologies	sc-136320	Western Blot, IF
PPM1G/PP2Cγ	Abcam	ab70794	Western Blot, IF, RIP
SART3 (C3)	GeneTex	GTX107684	Western Blot
HIV-1 p24 (AG3.0)	AIDS Research and Reference Reagent Program	4121	Western Blot, ELISA

GST	Santa Cruz Biotechnologies	sc-33613	Western Blot

Table 6. siRNA's used in this dissertation

Gene	**Flexitube siRNA product name**	**Catalog number**
CDK9	Hs_CDK9_5	SI00605066
MePCE	Hs _BCDIN3_1	SI02778265
hnRNP-F	Hs_HNRPF_1	SI00300461
DHX57	Hs_DHX57_5	SI02757839
NPM1	Hs_NPM1_1	SI00300979
PPM1G	Hs_PPM1G_6	SI02658684
Negative control	AllStars Negative Control	SI1027280
PPM1G*	Hs_PPM1G_10	SI05016627
PPM1G*	Hs_PPM1G_11	SI05016634
PPM1A	HS_PPM1A_5	SI02659258

SLC25A5	Hs_SLC25A5_5	SI02654358
CycT1	Hs_CCNT1_1	SI00024073
Larp7	Hs_HDCMA18P_2	SI00434980
SET	Hs_SET_5	SI03021291
DDX21	Hs_DDX21_7	SI04311412
Sart3	Hs_SART3_7	SI04357304
RelA H	s_RELA_5	SI00301672

*** Denotes an siRNA that targets the 3'UTR of the gene RNA**

CHAPTER VII

REFERENCES

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nature Reviews Genetics *13*, 720-731.

Allemand, E., Hastings, M.L., Murray, M.V., Myers, M.P., and Krainer, A.R. (2007). Alternative splicing regulation by interaction of phosphatase PP2C gamma with nucleic acid-binding protein YB-1. Nature Structural & Molecular Biology *14*, 630-638.

Amir-Zilberstein, L., Ainbinder, E., Toube, L., Yamaguchi, Y., Handa, H., and Dikstein, R. (2007). Differential regulation of NF-kappa B by elongation factors is determined by core promoter type. Molecular and Cellular Biology *27*, 5246-5259.

Ammosova, T., Washington, K., Debebe, Z., Brady, J., and Nekhai, S. (2005). Dephosphorylation of CDK9 by protein phosphatase 2A and protein phosphatase-1 in Tat-activated HIV-1 transcription. Retrovirology *2*.

Ammensova, T., Yedavalli, V. R., Niu, X., Jerebtsova, M., Van Eynde, A., Beullens, M., Bollen, M., Jeang, K. T., and Nekhai. S. (2011). Expression of a protein phosphatase 1 inhibitor, cdNIPP1, increases CDK9 threonine 186 phosphorylation and inhibits HIV-1 transcription. J Biol Chem *5* 3798-3804.

Bai, X., Kim, J., Yang, Z., Jurynec, M.J., Akie, T.E., Lee, J., LeBlanc, J., Sessa, A., Jiang, H., DiBiase, A.*, et al.* (2010). TIF1gamma controls erythroid cell fate by regulating transcription elongation. Cell *142*, 133-143.

Barboric, M., Kohoutek, J., Price, J.P., Blazek, D., Price, D.H., and Peterlin, B.M. (2005). Interplay between 7SK snRNA and oppositely charged regions in HEXIM1 direct the inhibition of P-TEFb. Embo Journal *24*, 4291-4303.

Barboric, M., Lenasi, T., Chen, H., Johansen, E.B., Guo, S., and Peterlin, B.M. (2009). 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. Proceedings of the National Academy of Sciences of the United States of America *106*, 7798-7803.

Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B.M. (2001a). NF-kappa B binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. Molecular Cell *8*, 327-337.

Barboric, M., Taube, R., Nekrep, N., Fujinaga, K., and Peterlin, B.M. (2000). Binding of Tat to TAR and recruitment of positive transcription elongation factor b occur independently in bovine immunodeficiency virus. J Virol *74*, 6039-6044.

Barboric, M., Yik, J.H.N., Czudnochowski, N., Yang, Z., Chen, R., Contreras, X., Geyer, M., Peterlin, B.M., and Zhou, Q. (2007). Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. Nucleic Acids Research *35*, 2003-2012.

Barkett, M., and Gilmore, T.D. (1999). Control of apoptosis by Rel/NF-kappa B transcription factors. Oncogene *18*, 6910-6924.

Barrandon, C., Bonnet, F., Nguyen, V.T., Labas, V., and Bensaude, O. (2007a). The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. Mol Cell Biol *27*, 6996-7006.

Bartkowiak, B., Liu, P., Phatnani, H.P., Fuda, N.J., Cooper, J.J., Price, D.H., Adelman, K., Lis, J.T. and Greenleaf A.L. (2010). CDK12 is a transcription elongation associated

CTD kinase, the metazoan ortholog of yeast Ctk1. Genes and Development *24* 2303-2316.

Baumli, S., Lolli, G., Lowe, E.D., Troiani, S., Rusconi, L., Bullock, A.N., Debreczeni, J.E., Knapp, S., and Johnson, L.N. (2008). The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. Embo Journal *27*, 1907-1918.

Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? Transcription *2*, 103-108.

Bentley D. L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol *3* 251-256.

Bittencourt, D. and Auboeuf, D. (2012). Analysis of co-transcriptional RNA processing by RNA-ChIP Assay. Methods Mol Biol *809* 563-577.

Blazek, D., Kohoutek, J., Bartholomeeusen, K., Johansen, E., Hulinkova, P., Luo, Z., Cimermancic, P., Ule, J., and Peterlin B.M. (2011). The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. Genes and Development *25* 2158-2172.

Boesken, C.A., Farnung, L., Hintermair, C., Schachter, M.M., Vogel-Bachmayr, K., Blazek, D., Anand, K., Fisher, R.P., Eick, D., and Geyer, M. (2014). The structure and substrate specificity of human Cdk12/Cyclin K. Nature Communications *5*.

Bowman, E.A., and Kelly, W.G. (2014). RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation A tail of two kinases. Nucleus *5*, 224-236.

Bunch, H., Zheng, X., Burkholder, A., Dillon, S.T., Motola, S., Birrane, G., Ebmeier, C.C., Levine, S., Fargo, D., Hu, G.*, et al.* (2014). TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release. Nat Struct Mol Biol *21*, 876-883.

Calo, E., Flynn, R.A., Martin, L., Spitale, R.C., Chang, H.Y., and Wysocka, J. (2015). RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. Nature *518*, 249-253.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol Cell *49*, 825-837.

Chang, P.-C., Fitzgerald, L.D., Van Geelen, A., Izumiya, Y., Ellison, T.J., Wang, D.-H., Ann, D.K., Luciw, P.A., and Kung, H.-J. (2009). Kruppel-Associated Box Domain-Associated Protein-1 as a Latency Regulator for Kaposi's Sarcoma-Associated Herpesvirus and Its Modulation by the Viral Protein Kinase. Cancer Research *69*, 5681-5689.

Chen, F., Gao, X., and Shilatifard, A. (2015). Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide. Genes Dev *29*, 39-47.

Chen, R., Liu, M., Li, H., Xue, Y., Ramey, W.N., He, N., Ai, N., Luo, H., Zhu, Y., Zhou, N.*, et al.* (2008). PP2B and PP1 alpha cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca2+ signaling. Genes & Development *22*, 1356-1368.

Chen, R.C., Yang, Z.Y., and Zhou, Q. (2004). Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. Journal of Biological Chemistry *279*, 4153-4160.

Cherrier, T., Le Douce, V., Eilebrecht, S., Riclet, R., Marban, C., Dequiedt, F., Goumon, Y., Paillart, J.-C., Mericskay, M., Parlakian, A.*, et al.* (2013). CTIP2 is a negative regulator of P-TEFb. Proceedings of the National Academy of Sciences of the United States of America *110*, 12655-12660.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368-373.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.

Czudnochowski, N., Bosken, C.A., and Geyer, M. (2012). Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. Nature communications *3*, 842.

D'Orso, I., and Frankel, A.D. (2010a). HIV-1 Tat: Its Dependence on Host Factors is Crystal Clear. Viruses-Basel *2*, 2226-2234.

D'Orso, I., and Frankel, A.D. (2010b). RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation. Nat Struct Mol Biol *17*, 815-821.

D'Orso, I., Jang, G.M., Pastuszak, A.W., Faust, T.B., Quezada, E., Booth, D.S., and Frankel, A.D. (2012). Transition Step during Assembly of HIV Tat:P-TEFb Transcription Complexes and Transfer to TAR RNA. Molecular and Cellular Biology *32*, 4780-4793.

Danko, C.G., Hah, N., Luo, X., Martins, A.L., Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol Cell *50*, 212-222.

Das, A.K., Helps, N.R., Cohen, P.T.W., and Barford, D. (1996). Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 angstrom resolution. Embo Journal *15*, 6798-6809.

Debaisieux, S., Rayne, F., Yezid, H., and Beaumelle, B. (2012). The Ins and Outs of HIV-1 Tat. Traffic *13*, 355-363.

Diamant, G., and Dikstein, R. (2013). Transcriptional Control by NF-kappa B: Elongation in Focus. Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms *1829*, 937-945.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Diribarne, G., and Bensaude, O. (2009). 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. Rna Biology *6*, 122-128.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., and Davies, D.R., (1994). Crystal-structure of the catalytic domain of HIV-1 integrase - similarity to other polynucleotidyl transferases. EMBO Journal *9* 1551-1560.

Egly, J.-M., and Coin, F. (2011). A history of TFIIH: Two decades of molecular biology on a pivotal transcription/repair factor. DNA Repair *10*, 714-721.

Eilebrecht, S., Le Douce, V., Riclet, R., Targat, B., Hallay, H., Van Driessche, B., Schwartz, C., Robette, G., Van Lint, C., Rohr, O.*, et al.* (2014). HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. Nucleic Acids Res *42*, 4962-4971.

Faust, T., Frankel, A., and D'Orso, I. (2012). Transcription control by long non-coding RNAs. Transcription *3*, 78-86.

Fisher, R.P. (2005). Secrets of a double agent: CDK7 in cell-cycle control and transcription. Journal of Cell Science *118*, 5171-5180.

Francastel, C., Schubeler, D., Martin, D.I.K., and Groudine, M. (2000). Nuclear compartmentalization and gene activity. Nature Reviews Molecular Cell Biology *1*, 137-143.

Frankel, A. D., and Young, J. A. (1998). HIV-1: fiftenn proteins and an RNA. Annual Reviews in Biochemistry *67*, 1-25.

Galbraith, M.D., Allen, M.A., Bensard, C.L., Wang, X., Schwinn, M.K., Qin, B., Long, H.W., Daniels, D.L., Hahn, W.C., Dowell, R.D.*, et al.* (2013). HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. Cell *153*, 1327-1339.

Garber, M.E., Mayall, T.P., Suess, E.M., Meisenhelder, J., Thompson, N.E., and Jones, K.A. (2000). CDK9 autophosphorylation regulates high-affinity binding of the human immunodeficiency virus type 1 Tat-P-TEFb complex to TAR RNA. Molecular and Cellular Biology *20*, 6958-6969.

Ghosh, S., and Karin, M. (2002). Missing pieces in the NF-kappa B puzzle. Cell *109*, S81-S96.

Glover-Cutter, K., Larochelle, S., Erickson, B., Zhang, C., Shokat, K., Fisher, R.P., and Bentley, D.L. (2009). TFIIH-Associated Cdk7 Kinase Functions in Phosphorylation of C-Terminal Domain Ser7 Residues, Promoter-Proximal Pausing, and Termination by RNA Polymerase II. Molecular and Cellular Biology *29*, 5455-5464.

Gomes, N.P., Bjerke, G., Llorente, B., Szostek, S.A., Emerson, B.M., and Espinosa, J.M. (2006). Gene-specific requirement for P-TEFb activity and RNA polymerase II

phosphorylation within the p53 transcriptional program. Genes & Development *20*, 601-612.

Grewe, B., and Uberla, K. (2010). The human immunodeficiency virus type 1 Rev protein: menage a trois during the early phase of the lentiviral replication cycle. Journal of General Virology *91*, 1893-1897.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Guettler, T., and Goerlich, D. (2011). Ran-dependent nuclear export mediators: a structural perspective. Embo Journal *30*, 3457-3474.

Guo, J., and Price, D.H. (2013). RNA Polymerase II Transcription Elongation Control. Chemical Reviews *113*, 8583-8603.

Hager, G.L., McNally, J.G., and Misteli, T. (2009). Transcription Dynamics. Molecular Cell *35*, 741-753.

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of Inducible Gene Expression by Signal-Dependent Transcriptional Elongation. Cell *138*, 129-145.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Henriques, T., Gilchrist, D.A., Nechaev, S., Bern, M., Muse, G.W., Burkholder, A., Fargo, D.C., and Adelman, K. (2013). Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals. Mol Cell *52*, 517-528.

Hogg, J.R., and Collins, K. (2007). RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. RNA *13*, 868-880.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes and Development *23*, 837-848.

Iyengar, S., and Farnham, P.J. (2011a). KAP1 protein: an enigmatic master regulator of the genome. J Biol Chem *286*, 26267-26276.

Iyengar, S., Ivanov, A.V., Jin, V.X., Rauscher, F.J., 3rd, and Farnham, P.J. (2011b). Functional analysis of KAP1 genomic recruitment. Mol Cell Biol *31*, 1833-1847.

Jadlowsky, J.K., Wong, J.Y., Graham, A.C., Dobrowolski, C., Devor, R.L., Adams, M.D., Fujinaga, K., and Karn, J. (2014). Negative Elongation Factor Is Required for the Maintenance of Proviral Latency but Does Not Induce Promoter-Proximal Pausing of RNA Polymerase II on the HIV Long Terminal Repeat. Molecular and Cellular Biology *34*, 1911-1928.

Jager, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C., Shales, M., Mercenne, G., Pache, L., Li, K.*, et al.* (2012). Global landscape of HIV-human protein complexes. Nature *481*, 365-370.

Jager, S., Gulbahce, N., Cimermancic, P., Kane, J., He, N., Chou, S., D'Orso, I., Fernandes, J., Jang, G., Frankel, A.D.*, et al.* (2011). Purification and characterization of HIV-human protein complexes. Methods *53*, 13-19.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J.*, et al.* (2007). Systematic analysis of the

protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Molecular Cell *27*, 262-274.

Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., and Fu, X.-D. (2013). SR Proteins Collaborate with 7SK and Promoter-Associated Nascent RNA to Release Paused Polymerase. Cell *153*, 855-868.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nature Reviews Molecular Cell Biology *16*, 167-177.

Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. Embo Journal *22*, 1868-1877.

Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. Embo Journal *20*, 1726-1738.

Karn, J. (1999). Tackling Tat. Journal of Molecular Biology *293*, 235-254.

Kayama, H., Ramirez-Carrozzi, V.R., Yamamoto, M., Mizutani, T., Kuwata, H., Iba, H., Matsumoto, M., Honda, K., Smale, S.T., and Takeda, K. (2008). Class-specific regulation of pro-inflammatory genes by MyD88 pathways and IkappaBzeta. J Biol Chem *283*, 12468-12477.

Kim, Y.K., Bourgeois, C.F., Isel, C., Churcher, M.J., and Karn, J. (2002). Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type I Tat-activated transcriptional elongation. Molecular and Cellular Biology *22*, 4622-4637.

Kim, Y.K., Bourgeois, C.F., Pearson, R., Tyagi, M., West, M.J., Wong, J., Wu, S.-Y., Chiang, C.-M., and Karn, J. (2006). Recruitment of TFIIH to the HIV LTR is a rate-limiting step in the emergence of HIV from latency. Embo Journal *25*, 3596-3604.

Klasse, P.J. (2012). The molecular basis of HIV entry. Cellular Microbiology *14*, 1183-1192.

Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C.E., Cooper, J.J., Bensaude, O., Cohen, E.A.*, et al.* (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. Nucleic Acids Research *36*, 2219-2229.

Krueger, B.J., Varzavand, K., Cooper, J.J., and Price, D.H. (2010). The Mechanism of Release of P-TEFb and HEXIM1 from the 7SK snRNP by Viral and Cellular Activators Includes a Conformational Change in 7SK. Plos One *5*.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lenasi, T., and Barboric, M. (2010). P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. Rna Biology *7*, 145-150.

Li, Q.T., Price, J.P., Byers, S.A., Cheng, D.M., Peng, J.M., and Price, D.H. (2005). Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. Journal of Biological Chemistry *280*, 28819-28826.

Li, H., Handsaker, B., Wysoker A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R; 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell *151*, 56-67.

Lis, J.T., Mason, P., Peng, J., Price, D.H., and Werner, J. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes & Development *14*, 792-803.

Liu, P., Xiang, Y., Fujinaga, K., Bartholomeeusen, K., Nilson, K.A., Price, D.H., and Peterlin, B.M. (2014). Release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein (snRNP) activates hexamethylene bisacetamide-inducible protein (HEXIM1) transcription. J Biol Chem *289*, 9918-9925.

Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A.K., and Rosenfeld, M.G. (2013). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. Cell *155*, 1581-1595.

Liu, Y., Li, J.L., Kim, B.O., Pace, B.S., and He, J.J. (2002). HIV-1 tat protein-mediated transactivation of the HIV-1 long terminal repeat promoter is potentiated by a novel nuclear tat-interacting protein of 110 kDa, Tip110. Journal of Biological Chemistry *277*, 23854-23863.

Lorincz, M.C., and Schubeler, D. (2007). RNA polymerase II: Just stopping by. Cell *130*, 16-18.

Luecke, H.F., and Yamamoto, K.R. (2005). The glucocorticoid receptor blocks P-TEFb recruitment by NF kappa B to effect promoter-specific transcriptional repression. Genes & Development *19*, 1116-1127.

Luo, Z., Lin, C., and Shilatifard A. (2012). The super elongation complex (SEC) family in transcriptional control. Nature Reviews in Mol Cell Biol *13*, 543-547.

Malumbres, M. (2014). Cyclin-dependent kinases. Genome Biology *15*.

Mancebo, H.S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D.*, et al.* (1997a). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev *11*, 2633-2644.

Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhas, O., Sickmann, A., and Fischer, U. (2008). The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. Embo Reports *9*, 569-575.

Marshall, N.F., and Price, D.H. (1992). CONTROL OF FORMATION OF 2 DISTINCT CLASSES OF RNA POLYMERASE-II ELONGATION COMPLEXES. Molecular and Cellular Biology *12*, 2078-2090.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J Biol Chem. 270:12335–12338.

Marz, M., Donath, A., Verstraete, N., Nguyen, V.T., Stadler, P.F., and Bensaude, O. (2009). Evolution of 7SK RNA and Its Protein Partners in Metazoa. Molecular Biology and Evolution *26*, 2821-2830.

Mbonye, U., and Karn, J. (2011). Control of HIV Latency by Epigenetic and Non-Epigenetic Mechanisms. Current Hiv Research *9*, 554-567.

Mbonye, U., and Karn, J. (2014). Transcriptional control of HIV latency: Cellular signaling pathways, epigenetics, happenstance and the hope for a cure. Virology *454*, 328-339.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley D. L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385* 357-361.

McNamara, R.P., McCann, J. L. Aravind, S.A., and D'Orso, I. (2013). Transcription factors mediate enzymatic disassebly of promoter bound 7SK snRNP to locally recruit P-TEFb for elongation. Cell Reports *5* 1256-1288.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev *25*, 742-754.

Muniz, L., Egloff, S., and Kiss, T. (2013). RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. Nucleic Acids Research *41*, 4686-4698.

Muniz, L., Egloff, S., Ughy, B., Jady, B.E., and Kiss, T. (2010). Controlling Cellular P-TEFb Activity by the HIV-1 Transcriptional Transactivator Tat. Plos Pathogens *6*, 12.

Murphy, S., Diliegro, C., and Melli, M. (1987). THE INVITRO TRANSCRIPTION OF THE 7SK RNA GENE BY RNA POLYMERASE-III IS DEPENDENT ONLY ON THE PRESENCE OF AN UPSTREAM PROMOTER. Cell *51*, 81-87.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007a). RNA polymerase is poised for activation across

Napolitano, G., Lania, L., and Majello, B. (2014). RNA Polymerase II CTD Modifications: How Many Tales From a Single Tail. Journal of Cellular Physiology *229*, 538-544.

Nguyen, V.T., Kiss, T.S., Michels, A.A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature *414*, 322-325.

Nissen, R.M., and Yamamoto, K.R. (2000). The glucocorticoid receptor inhibits NF kappa B by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes & Development *14*, 2314-2329.

Osborn, L., Kunkel, S., and Nabel, G.J. (1989). TUMOR NECROSIS FACTOR-ALPHA AND INTERLEUKIN-1 STIMULATE THE HUMAN IMMUNODEFICIENCY VIRUS ENHANCER BY ACTIVATION OF THE NUCLEAR FACTOR KAPPA-B. Proceedings of the National Academy of Sciences of the United States of America *86*, 2336-2340.

Ott, M., Geyer, M., and Zhou, Q. (2011). The Control of HIV Transcription: Keeping RNA Polymerase II on Track. Cell Host & Microbe *10*, 426-435.

Palmieri, C., Trimboli, F., Scala, G., Quinto, I., and Bressler, P.B. (2003) Regulation of the HIV promoter/enhancer. Current Protoc Immunology. Chapter 12, Unit 12.

Pearson, R., Kim, Y.K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M., and Karn, J. (2008). Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. J Virol *82*, 12291-12303.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol Cell *23*, 297-305.

Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Molecular and Cellular Biology *20*, 2629-2634.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

Rayne, F., Debaisieux, S., Yezid, H., Lin, Y.-L., Mettling, C., Konate, K., Chazal, N., Arold, S.T., Pugniere, M., Sanchez, F.*, et al.* (2010a). Phosphatidylinositol-(4,5)-bisphosphate enables efficient secretion of HIV-1 Tat by infected T-cells. Embo Journal *29*, 1348-1362.

Rayne, F., Debaisieux, S., Bonhoure, A., and Beaumelle B. (2010b) HIV-1 Tat is uconventionally secreted through the plasma membrane. Cell Biol Int. *4* 409-413.

Robinson, J.T., Thorvalshdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotechnology *29*, 24-26.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J.*, et al.* (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237-240.

Ruelas, D.S., and Greene, W.C. (2013). An Integrated Overview of HIV-1 Latency. Cell *155*, 519-529.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc *3*, 1101-1108.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529.

Sedore, S.C., Byers, S.A., Biglione, S., Price, J.P., Maury, W.J., and Price, D.H. (2007). Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR. Nucleic Acids Research *35*, 4347-4358.

Sidoti-de Fraisse, C., Rincheval, V., Risler, Y., Mignotte, B., and Vayssiere, J.L. (1998). TNF-alpha activates at least two apoptotic signaling cascades. Oncogene *17*, 1639-1651.

Smale, S.T. (2010). Selective transcription in response to an inflammatory stimulus. Cell *140*, 833-844.

Smith, E., Lin, C., and Shilatifard A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes and Development *25*, 661-672.

Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., and Benkirane, M. (2010). HIV-1 Tat Assembles a Multifunctional Transcription Elongation Complex and Stably Associates with the 7SK snRNP. Molecular Cell *38*, 439-451.

Sundquist, W.I., and Kraeusslich, H.-G. (2012). HIV-1 Assembly, Budding, and Maturation. Cold Spring Harbor Perspectives in Medicine *2*.

Tahirov, T.H., Babayeva, N.D., Varzavand, K., Cooper, J.J., Sedore, S.C., and Price, D.H. (2010). Crystal structure of HIV-1 Tat complexed with human P-TEFb. Nature *465*, 747-U742.

Tee, W. W., Shen, S. S., Oksuz, O., Narendra, V., and Reinberg D. (2014). Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. Cell *156* 678-690.

Trudgain, D.C., Ridlova, G., Fischer R., Mackeen, M.M., Ternette, N., Acuto, O., Kessler B.M., and Thomas, B. (2011). Comparative evaluation of label-free SINQ normalized

spectral index quantitiation in the central proteomics facilities pipeline. Proteomics *11*, 2790-2797.

Titov, D.V., Gilman, B., He, Q.L., Bhat, S., Low, W.K., Dang, Y., Smeaton, M., Demain, A.L., Miller, P.S., Kugel, J.F.*, et al.* (2011). XPB, a subunit of TFIIH, is a target of the natural product triptolide. Nat Chem Biol *7*, 182-188.

Tyagi, M., Pearson, R.J., and Karn, J. (2010). Establishment of HIV Latency in Primary CD4(+) Cells Is due to Epigenetic Transcriptional Silencing and P-TEFb Restriction. Journal of Virology *84*, 6425-6437.

Van Lint, C., Bouchat, S., and Marcello, A. (2013). HIV-1 transcription and latency: an update. Retrovirology *10*.

van Werven, F.J., van Teeffelen, H.A., Holstege, F.C., and Timmers, H.T. (2009). Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. Nat Struct Mol Biol *16*, 1043-1048.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. Embo Journal *17*, 7395-7403.

Wang, Y., Dow, E.C., Liang, Y.-Y., Ramakrishnan, R., Liu, H., Sung, T.-L., Lin, X., and Rice, A.P. (2008). Phosphatase PPM1A Regulates Phosphorylation of Thr-186 in the Cdk9 T-loop. Journal of Biological Chemistry *283*, 33578-33584.

Wassarman, D.A., and Steitz, J.A. (1991). STRUCTURAL-ANALYSES OF THE 7SK RIBONUCLEOPROTEIN (RNP), THE MOST ABUNDANT HUMAN SMALL RNP OF UNKNOWN FUNCTION. Molecular and Cellular Biology *11*, 3432-3445.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46-57.

Xue, Y., Yang, Z., Chen, R., and Zhou, Q. (2010). A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. Nucleic Acids Res *38*, 360-369.

Yang, Z.Y., Zhu, Q.W., Luo, K.X., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature *414*, 317-322.

Yik, J.H.N., Chen, R.C., Nishimura, R., Jennings, J.L., Link, A.J., and Zhou, Q. (2003). Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. Molecular Cell *12*, 971-982.

Zhang, J. and Corden, J. L. (1991). Identification of phospohrylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. J Biol Chem *266* 2290-2296.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhou, M.S., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J.M., Price, D.H., and Brady, J.N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Molecular and Cellular Biology *20*, 5077-5086.

Zhou, Q., Li, T., and Price, D.H. (2012). RNA Polymerase II Elongation Control. Annual Review of Biochemistry, Vol 81 *81*, 119-143.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B., and Price, D.H. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev *11*, 2622-2632.