Title
Mechanistic insights into the roles of P-TEFb and its novel cofactors in tumorigenesis and HIV transcription

Author
He, Nanhai

Publication Date
2012

Peer reviewed|Thesis/dissertation

Mechanistic insights into the roles of P-TEFb and its novel cofactors in tumorigenesis and HIV transcription

By

Nanhai He

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Qiang Zhou, Chair
Professor Mike Botchan
Professor Tom Abler
Professor Fenyong Liu

Spring 2012

Abstract

Mechanistic insights into the roles of P-TEFb and its novel cofactors in tumorigenesis and HIV transcription

by

Nanhai He

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Qiang Zhou, Chair

Ongoing research in the field of transcription has given rise to the unappreciated role of elongation control as a rate limiting step for transcription, and the general transcription elongation factor P-TEFb therefore has taken the central stages. P-TEFb is composed of cyclin T1 and the cyclin-dependent kinase Cdk9. It stimulates transcription elongation by releasing the paused RNA Polymerase II (Pol II) through phosphorylating Pol II at Ser2 and antagonizing the effects of negative elongation factors. P-TEFb is not only essential for the transcription of the vast majority of cellular genes, but also critical for the expression of HIV genome. Elucidating how P-TEFb activity is controlled therefore plays a key role in advancing our understanding of cellular and disease-related transcription elongation.

The first part of this thesis presents my early Ph.D work, which focusing on the identification and characterization of a La-like protein PIP7S (also named as LARP7). I have shown that PIP7S binds and stabilizes nearly all the nuclear 7SK, which leading to sequestration and inactivation of P-TEFb. Consistent with the fact that PIP7S is frequently mutated in human tumors and the Drosophila homolog of PIP7S is a tumor suppressor, loss of PIP7S function disrupts epithelial differentiation and causes P-TEFb-dependent transformation.

The second part of this thesis introduces the identification of the Super Elongation Complex (SEC) and subsequent characterization of two of its subunits, AFF4 and ELL2. I have shown that through the bridging functions of AFF4 and Tat, P-TEFb and ELL2 combine to form a bifunctional elongation complex that greatly activates basal and HIV-1 transcription, respectively.

The third part of this thesis continues to dissect the functions of the other two SEC subunits ENL and AF9. I have shown that the homologous ENL and AF9 exist in separate SECs with similar but non-identical functions. ENL/AF9 also exists outside SEC when bound to Dot1L, which is found to inhibit SEC function. The

YEATS domain of ENL/AF9 targets SEC to Pol II on chromatin through contacting the PAFc complex.

To beloved Dandan, Lawrence and my parents

Table of Contents

List of Figures

Chapter 2

Chapter 3

List of Symbols and Abbreviations

AF9	ALL1-Fused gene from chromosome 9
AFF4	AF4/FMR2 family, member 4
CDK9	Cyclin-dependent kinase 9
ChIP	Chromatin immunoprecipitation
CTD	C-terminal domain
CYCT1	CyclinT1
DRB	5,6-dichloro-l-β-D-ribofuranosylbenzimidazole
DSIF	DRB-sensitivity inducing factor
ELL2	eleven-Nineteen Lysine-rich Leukemia Gene 2
EMSA	electrophoretic mobility shift assay
ENL	eleven Nineteen Leukemia
GST	Glutathione-S -Transferase
HIV-1	Human immunodeficiency virus type-1
HMBA	hexamethylene bisacetamide
IP	Immunoprecipitation
LARP7	La ribonucleoprotein domain family, member 7
LTR	long terminal repeat
mAb	monoclonal antibody
MELC	murine erythroleukemia cells
MLL	mixed lineage leukemia
NB	northern blotting
NE	nuclear extracts
NELF	negative elongation factor
ORF	open reading frame
PAFc	Polymerase Associated Factor complex
PIP7S	protein interacting with P-TEFb and 7SK
Pol II	RNA polymerase II
P-TEFb	Positive transcription elongation factor b

RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein
RT	room temperature
RRM	RNA Recognition Motif
SEC	Super Elongation Complex
shRNA	short hairpin RNA
snRNA	small nuclear ribonucleic acid
TAR	trans-acting response element
TCEP	tris(2-carboxyethyl)phosphine
TRM	Tat-TAR Recognition Motif
ug	microgram
ul	microliter
μM	micromolar
UTR	untranslated region
WB	Western blotting
WCE	whole cell extracts
WT	wildtype

Acknowledgements

I would like first to thank my advisor, Dr. Qiang Zhou, for giving me the opportunity to do my Ph.D work in his lab. His tremendous passion and enthusiasm for research have been motivating and inspiring me all the times, and his persistent supports and guidance have made everything in my thesis possible. I would also like to acknowledge all the members, both past and present, of the Zhou lab for making the lab an enjoyable place to work. I have learned an incredible amount from each person. I am also grateful to my committee members Drs. Mike Botchan, Tom Alber and Fenyong Liu for their time and guidance. I also want to thank the California HIV/AIDS Research Program for financial support.

Last and most importantly, I have to thank my wife Dandan and my parents, and their believing in my ability to succeed is always a comforting thought.

Chapter 1

General Introduction

Transcriptional elongation as a major gene expression control point

Transcription by RNA polymerase II (RNAPII) can be subdivided into multiple stages, of which the initiation and elongation stages are considered the two primary targets for controlling eukaryotic gene expression (1-3). However, during most of the past three decades, attention of the transcription field had been largely focused on the initiation stage, as the recruitment of RNAPII to a few model gene promoters was found to be the major rate-limiting step for their transcription (2, 4). In comparison, transcriptional control at the elongation stage had been generally neglected and viewed as a mechanism that is used by only a few selected genes and under highly specialized conditions.

A major paradigm shift in the transcription field occurred in 2007 when global analyses conducted in both Drosophila and human stem cells revealed that a large number of genes that often play important roles in controlling cell growth, renewal and differentiation have paused RNAPII at their promoter-proximal regions even under resting, un-stimulated conditions (5, 6). For these genes, transcriptional activation does not involve the recruitment of RNAPII and setting up a pre-initiation complex (PIC) at the promoters, which can be very time-consuming. Rather, the transition of RNAPII from its promoter-proximal paused state into highly productive elongation mode is the defining moment of their activation. Because these genes are already in a state of suspended transcription prior to activation, the subsequent induction of RNAPII elongation can proceed very rapidly, thus allowing highly sensitive and synchronous response that is essential for cell growth and developmental control. The widespread existence of paused RNAPII in metazoan genomes suggests that transcriptional elongation plays a much more prominent and general role in regulating gene expression than previously appreciated.

Tat activation of HIV-1 transcriptional elongation

The detection of promoter-proximal pausing of RNAPII at many gene promoters was indeed a major breakthrough in the transcription field at large. However, to those who had been studying HIV-1 gene expression control, the discovery came as no great surprise. Prior to the recognition of the general importance of elongation control, HIV-1 transcription had long been known to be regulated exquisitely at the elongation stage (7). In fact, HIV-1 had been used as a favored model system to study this phenomenon. As such, our understanding of elongation control has benefited greatly from studies of HIV-1 and its activation by a combination of viral and host transcription factors.

Unlike simpler retroviruses that rely almost exclusively on the host cellular machinery for replication, lentiviruses, of which HIV-1 is a prime example, encode additional accessory proteins that further control the viral life cycle. The transcriptional transactivator (Tat) is one such key accessory protein encoded by HIV-1. During active infection, Tat is expressed early after the proviral DNA integrates into the host genome. Without Tat, RNAPII has been found to clear the HIV-1 LTR successfully but soon pause, producing only short viral transcripts (7). Ample evidence indicates that Tat does not act alone and must cooperate with host cellular cofactors to activate HIV-1 transcription. After many years of effort aimed at elucidating the mechanism of Tat-transactivation, the late 1990s finally saw the identification of the human positive

transcription elongation factor b (P-TEFb) as a specific and essential human co-factor for Tat function (8, 9).

Consisting of the cyclin-dependent kinase 9 (CDK9) and its regulatory partner cyclin T1 (CycT1; other minor cyclin forms also exist but do not interact with Tat), P-TEFb is recruited to the HIV-1 LTR through interacting with Tat and the transactivation response (TAR) element, an RNA stem-loop structure formed by the nascent 5' end of viral transcripts that are synthesized before RNAPII pauses (for reviews, see (10, 11). Once positioned next to the paused polymerase, the CDK9 kinase phosphorylates its primary substrates, the C-terminal domain (CTD) of the largest subunit of RNAPII and a pair of negative transcription elongation factors, DSIF and NELF. These phosphorylation events antagonize the actions of the negative elongation factors and release RNAPII from promoter-proximal pausing, leading to the production of full-length viral transcripts (10, 11).

Cellular control of P-TEFb activity

P-TEFb is not a transcription elongation factor made just for HIV-1. Rather, its activity is also important for the expression of a vast majority of cellular genes in uninfected cells (12). Recent whole genome analyses have shown that inhibiting P-TEFb activity prevents the release of paused RNAPII at most active gene loci in embryonic stem cells (13), again underlining P-TEFb's general role in transcription.

Given the importance of transcriptional elongation in controlling both HIV-1 and cellular gene expression and a key role for P-TEFb in this process, one can expect that the activity of P-TEFb is tightly regulated in the cell in order to optimally address the transcriptional needs of both the virus and its host. Indeed, P-TEFb has been shown to interact with a variety of protein and RNA regulators, and these interactions dynamically modulate the level of active P-TEFb available for HIV-1 and cellular gene expression (11). For example, under normal growth conditions, more than half of the P-TEFb population in the HeLa nucleus are sequestered in a catalytically inactive complex termed the 7SK snRNP that also contains the 7SK snRNA and nuclear proteins HEXIM1 (or the homologous HEXIM2), LARP7 and MePCE (14-18). Within this complex, 7SK, an evolutionally conserved snRNA transcribed by RNA polymerase III, is protected against exonuclease cleavage by the actions of MePCE, the 7SK methylphosphate capping enzyme, as well as LARP7, a La-related protein associated with the 3'-poly(U) track of 7SK (14, 15, 19). In return, 7SK functions as a molecular scaffold to maintain the integrity of 7SK snRNP, which sequesters P-TEFb and allows the CDK9 kinase activity to be inhibited by HEXIM1/2 in a 7SK-dependent manner (18, 20).

The nuclear level of 7SK snRNP undergoes dynamic changes under a variety of conditions that globally affect cell growth and differentiation. For example, the exposure of cardiac myocytes to hypertrophic signals triggers the release of P-TEFb from 7SK snRNP, leading to an overall increase in cellular protein and RNA contents and hypertrophic growth (21, 22). Moreover, co-stimulating Jurkat T cells with anti-CD3/anti-CD28 antibodies to activate the T-cell receptor (TCR) pathway also results in the disruption of 7SK snRNP and liberation of P-TEFb (23). When cells are treated with stress-inducing agents such as DNA-damaging agents actinomycin D and

UV irradiation, and the kinase inhibitors DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole), flavopiridol, staurosporine, and H7 (1-(5-isoquinolinesulfonyl)-2-methylpiperazine), P-TEFb has also been found to dissociate from 7SK snRNP to mediate the stress-induced HIV-1 and cellular gene expression (16, 17, 24, 25). Finally, the 7SK snRNP level in murine erythroleukemia cells (MELC) shows a biphasic response upon the exposure to hexamethylene bisacetamide (HMBA; (26), an inducer of MELC differentiation. During the initial 1-2 hrs, HMBA induces a transient disruption of 7SK snRNP, which is then followed by a permanent increase in the levels of HEXIM1 and 7SK snRNP after a prolonged treatment (26). Taken together, all these observations are consistent with the notion that the 7SK snRNP serves as a reservoir, from which active P-TEFb can be withdrawn in response to increased demands for elevated gene expression during active cell growth and response to environmental changes/stress (11).

Brd4 recruitment of P-TEFb for general transcription but not Tat-transactivation

Once P-TEFb is released from the 7SK snRNP, it often finds its way into another complex that is characterized by the presence of the bromodomain protein Brd4 (Fig. 1; (27, 28). Brd4 is a ubiquitously expressed nuclear protein belonging to the BET (bromodomain and ET domain) family of proteins that contain two N-terminal tandem bromodomains and an extraterminal domain (29, 30). While the motif located near the C-terminus of Brd4 has been shown to be responsible for binding to P-TEFb (31), the two bromodomains residing in the N-terminal region of Brd4 are involved in the interaction with the acetylated tails of histone H3 and H4 (32). Notably, the association of Brd4 with acetylated chromatin can persist through mitosis (29, 32, 33). These properties enable Brd4 to recruit P-TEFb to a chromatin template, beginning around mid- to late anaphase and before nuclear envelope/lamina formation and nuclear import of other general transcription factors (33). This leads to activation of the expression of P-TEFb-dependent genes in the early G1 phase of the cell cycle (33, 34). As such, Brd4 has been proposed to play a key role in transmitting epigenetic memory across cell division (32, 35).

The recruitment of P-TEFb by Brd4 is likely to be important for general transcriptional elongation of both cellular and viral genes including basal, Tat-independent HIV-1 transcription (27, 28). However, a number of sequence-specific transcriptional activators (e.g. HIV-1 Tat, NF-κB, Myc, CIITA, and etc.) have been shown to interact with P-TEFb (36-39) and they could potentially bypass the requirement for Brd4 and deliver P-TEFb directly to their respective target genes (11). An excellent example to illustrate this point is provided by HIV-1 Tat. Evidence has been presented to show that the recruitments of P-TEFb by Tat and Brd4 are two mutually exclusive events that cannot occur at the same time (28, 31). In fact, Brd4 interferes with Tat-transactivation as it competes with Tat for binding to P-TEFb (28). Furthermore, it has been shown that overexpression of the P-TEFb-interacting domain (PID) located at the C-terminus of Brd4 disrupts the Tat-P-TEFb interaction and inhibits Tat transactivation and TNF-α-induced reactivation of latent HIV-1 (31).

Where does Tat get its P-TEFb?

The demonstrations that a major source of nuclear P-TEFb exists in the 7SK

snRNP raise the questions of where exactly Tat obtains its P-TEFb. The answer to this question has come from several independent studies all showing that Tat triggers the release of P-TEFb from 7SK snRNP in vitro and in vivo (40-42). Consistently, primary blood lymphocytes display a reduced amount of nuclear 7SK snRNP upon HIV-1 infection (40). The existence of multiple structural and sequence similarities between the two P-TEFb-containing ribonucleoprotein (RNP) complexes, with one containing the HIV-1 TAR RNA and the other the cellular 7SK snRNA, likely contributed to this phenomenon (11).

However, the exact mechanism used by Tat to extract P-TEFb from 7SK snRNP remains controversial. Several published studies show that this ability of Tat depends on the integrity of its N-terminal activation domain and stems from the high affinity interaction between Tat and CycT1, which allows Tat to directly displace HEXIM1 from CycT1 (40-42). Once P-TEFb leaves the complex, a conformational change in 7SK blocks re-association of HEXIM1 (43).

Revealing a different mechanism used by Tat to capture P-TEFb from the 7SK snRNP, a recent study implicates the Arginine-rich TAR-binding domain of Tat as critical in this process (44). Tat is shown to use this domain to interact with a portion of the 7SK snRNA, which is normally contacted by HEXIM1 but structurally similar to the Tat-binding site of HIV-1 TAR, and cause the release of P-TEFb (44). What remains to be seen from this study is how the observed Tat-7SK interaction will eventually be turned into the Tat/TAR/P-TEFb-containing complex that is necessary for Tat-transactivation.

Another possible mechanism by which Tat extracts P-TEFb from 7SK snRNP involves the use of protein phosphatase 1 (PP1). This enzyme has been demonstrated to play a key role in stress-induced disruption of 7SK snRNP through dephosphorylating Threonine186 located at the tip of the CDK9 T-loop (45). Interestingly, Tat has been shown to bind PP1 directly (46), which can presumably deliver the enzyme to the 7SK snRNP to induce the latter's disruption. Consistent with this notion, inhibition of PP1 in cultured cells is reported to block Tat activation of HIV-1 transcription and replication, and at the same time, increase the nuclear 7SK snRNP level (47). Since the engagement of PP1 by Tat will likely lead to the release of P-TEFb with the dephosphorylated CDK9 T-loop and thus catalytically inactive, it is postulated that there must be a subsequent, yet-to-be defined re-phosphorylation step to return P-TEFb to its active state prior to its contribution to Tat activation of HIV-1 transcription (45).

In addition to investigating the mechanisms of Tat disruption of 7SK snRNP, recent efforts have also been focused on determining the subnuclear location where this event takes place. In binding studies conducted in vitro, HEXIM1 has been shown to bind to the HIV-1 TAR RNA and inhibit P-TEFb kinase activity (40-42), implicating a possible association of the 7SK snRNP with the LTR through TAR. Using the chromatin immunoprecipitation (ChIP) assay, another study also detects the association of 7SK snRNP with the pre-initiation complex formed on the HIV-1 LTR (48). However, this association is shown to proceed in the absence of TAR RNA, and the synthesis of TAR actually triggers the release of P-TEFb for activated HIV-1 transcription (48). Although it is quite clear that 7SK snRNP can indeed be found on the

LTR during basal transcription (as indicated on the ΔTAR template), its level appears to be very low and cannot fully account for the high P-TEFb level detected on the LTR upon Tat activation (48), suggesting that the bulk of P-TEFb required for activated transcription may come from a different route.

Obviously, more studies are needed to fully understand the physiological significance of the association of 7SK snRNP with the HIV-1 LTR and determine exactly how and where the complex is targeted by Tat. Nevertheless, it is abundantly clear from the published data that Tat has the ability to not only recruit P-TEFb to the LTR but also increase the pool of functional P-TEFb for HIV-1 transcription through actively extracting P-TEFb from the 7SK snRNP.

Chapter 2

A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis

(This work was originally published in *Mol Cell* as: He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, Luo K, Zhou Q. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 2008, 29:588-99)

Summary

The general transcription factor P-TEFb stimulates RNA polymerase II elongation and co-transcriptional processing of pre-mRNA. Contributing to a functional equilibrium important for growth control, a reservoir of P-TEFb is maintained in an inactive snRNP where 7SK snRNA is a central scaffold. Here, we identify PIP7S as a La-related protein stably associated with and required for 7SK snRNP integrity. PIP7S binds and stabilizes nearly all the nuclear 7SK via 3' UUU-OH, leading to the sequestration and inactivation of P-TEFb. This function requires its La domain and intact C-terminus. The latter is frequently deleted in human tumors due to microsatellite instability-associated mutations. Consistent with the tumor suppressor role of a Drosophila homolog of PIP7S, loss of PIP7S function shifts the P-TEFb equilibrium toward the active state, disrupts epithelial differentiation and causes P-TEFb-dependent malignant transformation. Through PIP7S modulation of P-TEFb, our data thus link a general elongation factor to growth control and tumorigenesis.

Introduction

For many genes from flies to humans, transcriptional elongation is a key step to control their expression. During the transcription of these genes, RNA polymerase (Pol) II is paused early after initiation by negative transcription elongation factors (N-TEF). Reversing this block requires the positive transcription elongation factor b (P-TEFb), a kinase that phosphorylates the C-terminal domain (CTD) of the largest subunit of Pol II and N-TEF, and allows Pol II to produce full-length transcripts (10). This function of P-TEFb also facilitates the coupling of elongation with pre-mRNA processing (11, 26). P-TEFb is a heterodimer composed of Cdk9 and cyclin T1 (CycT1) (or the minor forms T2 or K) (10). Experiments employing RNAi or specific Cdk9 inhibitors suggest that P-TEFb is a general transcription factor important for the expression of a large number of genes (12, 49).

Recent studies indicate that most of P-TEFb in the nucleus exists in two distinct functional states (11, 26). In HeLa cells, about half of P-TEFb is in a catalytically inactive complex (referred herein as 7SK snRNP) that contains the 7SK snRNA and the HEXIM1 (or the minor isoform HEXIM2) protein (16-18, 20). Within this complex, 7SK mediates the interaction of P-TEFb with HEXIM1, which in turn inhibits P-TEFb kinase activity. Transcribed by RNA Pol III, 7SK is an abundant non-coding RNA of 331 nucleotides that is highly conserved in higher eukaryotes (50). In HeLa cells, approximately the other half of P-TEFb exists in an active complex containing the bromodomain protein Brd4 (27, 28). Brd4 recruits P-TEFb to chromatin templates through interacting with acetylated histones and the mediator complex, and this function is important for general elongation. Notably, P-TEFb can also be recruited by a number of gene-specific transcription factors (e.g. the HIV-1 Tat protein) to activate their target genes (26, 51).

Through alternately interacting with its positive (Brd4) and negative (HEXIM1/7SK) regulators, P-TEFb is maintained in a functional equilibrium (26, 51). Recent studies suggest that shifts in this equilibrium may underlie alternative pathways toward unrestricted growth or terminal division/differentiation (21, 22, 50, 52). According to this model, the 7SK snRNP represents a reservoir of activity that can respond to demand for P-TEFb-dependent transcription and cell proliferation by rapidly releasing active P-TEFb (26, 51).

Toward the goal of identifying other factors involved in P-TEFb regulation, we report here the identification and characterization of a new protein, termed PIP7S, which is intimately associated with all the nuclear 7SK and required for 7SK stability and 7SK snRNP integrity. PIP7S is homologous to human La protein, a UUU-OH sequence-specific RNA-binding protein that associates with nascent Pol III transcripts and protects them from degradation by 3' exonucleases. We show that PIP7S indeed binds 7SK RNA in a UUU-OH-dependent manner and confers RNA 3' end protection activity in an established in vivo assay. This function of PIP7S is required for 7SK stability, 7SK snRNP assembly and inhibition of P-TEFb-dependent transcription. In addition to the La domain (the La and RRM motifs) in the PIP7S N-terminal region, sequestration and inactivation of P-TEFb also require the C-terminus, which is frequently deleted in human tumors with microsatellite instability. Consistent with the

demonstration that the Drosophila homolog of PIP7S is a tumor suppressor, PIP7S knockdown shifts the P-TEFb equilibrium toward the active Brd4-bound state and causes P-TEFb-dependent malignant transformation and activation of key tumor-related genes. Together, these observations link the PIP7S-dependent modulation of P-TEFb activity to the global control of cell growth and tumorigenesis.

Experimental Procedures

Immunological reagents

The rabbit polyclonal anti-PIP7S antibodies were raised against the PIP7S C-terminal peptide (TQQASKHIRFSEYD; aa 569-582) and affinity-purified. Antibodies against α6 integrin (rat), E-cadherin (mouse) and Ki-67 (rabbit) were from Chemicon, BD and Zymed, respectively. The anti-HEXIM1 antibodies have been described previously (18). All other antibodies were from Santa Cruz Biotechnology.

Immunodepletion of PIP7S or Cdk9 from HeLa NE

Depletion was performed by incubating 100 μl of HeLa NE (~7 mg/ml) containing 0.2% NP-40 and 0.35 M NaCl with 7 μg of anti-PIP7S or anti-Cdk9 antibodies at 4°C for 30 min, followed by the incubation with 30 μl of protein A-Sepharose beads (Amersham Biosciences) for 30 min. Upon the removal of the beads, 7 μg of fresh antibodies were added to the NE for 30 min at 4°C followed by three successive rounds of incubation with 30 μl each of fresh protein A-Sepharose beads.

7SK RNA gel shift assay (EMSA)

^{32}P-labeled wild-type 7SK and 7SK(Δ4U's) were synthesized by T7 RNA polymerase from PCR-amplified DNA templates. For EMSA, 20 μl reactions were carried out in buffer D (20 mM HEPES-KOH [pH 7.9], 15% glycerol, 0.2 mM EDTA, 0.1% NP-40, 1 mM dithiothreitol, and 1 mM phenylmethylsulfonyl fluoride) supplemented with 100 mM KCl, 5 mM MgCl2, 1 μg BSA, 150 ng poly(rG), 2,000 cpm of radiolabeled 7SK, and 0.2 pM of affinity-purified wild-type or mutant f-PIP7S proteins as indicated. For antibody supershift of the PIP7S-7SK complex, 0.3 μg of anti-PIP7S or anti-HEXIM1 antibodies were also included. After incubating at RT for 20 min, the reactions were resolved on a 4% polyacrylamide (19:1, acrylamide:bisacrylamide) gel in 0.5x Tris-glycine at 4°C for 3 h at 250 V.

To purify wild type and mutant f-PIP7S proteins free of any associated factors for EMSA, anti-Flag immunoprecipitation was performed in micrococcal nuclease (MN)-treated NEs from transiently transfected HeLa cells. The immunoprecipitates were washed extensively with buffer D containing 1.0 M KCl (D1.0) and then buffer D0.1 (0.1 M KCl) before elution with the Flag peptide as described previously (25, 53). To treat NE with MN (Roche Applied Science), 75 units of MN were incubated with 1 ml of NE and 1mM CaCl2 for 15 min at RT. The reaction was stopped with 10 mM EDTA. Purity of the proteins was confirmed by silver staining and the concentrations estimated by comparing with the BSA standards.

Generation of PIP7S knockdown cell lines

The Cre-induced RNA interference (RNAi) was used to generate the HeLa-based PIP7S knockdown cell lines. The lentiviral vector pSico (54) was a gift

from Dr. Tyler Jacks at MIT. Either of the following two shRNA sequences: 5'-AAGTTAATCACCAAAGCTGAATTCAAGAGATTCAGCTTTGGTGATTAACTT TTT-3' and 5'-AATCACAGCTGGATTGAAAGATTCAAGAGATCTTTCAATCCAGCTGTGATT TTT-3' was cloned into pSico for knocking down PIP7S. The procedures for the generation of recombinant lentiviruses, infection of F1C2 cells, and selection of knockdown clones were as described (54). To induce the expression of shPIP7S, two rounds of infection 24 hr apart were performed by using 5-10 PFU/cell of the adenovirus expressing Cre recombinase (Gene Transfer Vector Core facility of University of Iowa).

To establish MCF10A-based PIP7S knockdown cells, retroviruses, generated with the pSUPER vector (Oligoengine, WA) containing the shPIP7S-expressing cassette, were produced in the GP2-293 packaging cell line (Clontech, CA). Infected cells were selected with 0.4 μg/ml puromycin for two weeks to obtain individual clones.

To rescue PIP7S expression in MCF10A knockdown cells, nucleotides A, C and C at positions 1671, 1674 and 1677 were changed to G, T and G, respectively, to obtain the shPIP7S-resistant f-PIP7S cDNA encoding wild-type PIP7S, which was stably transduced via lentiviral infection.

Quantitative real-time PCR (qRT-PCR)

qRT-PCR was performed with Applied Biosystem 7300 Real-Time PCR System and Finnyzme F-410L SYBR Green RT-PCR reagents following the manufacturers' instructions. PCR primers were designed using Integrated DNA Technologies' Primer Quest. PCR conditions include an initial denaturing step at 94°C for 2 min and 40 cycles of 94°C for 30 sec and 60°C for 90 sec. Threshold values (Ct) were calculated to obtain the relative folds of induction. All reactions were run in triplicates.

Chromatin immunoprecipitation (ChIP) asay

The assay conditions were as described (Yang et al., 2005). After DNA purification, PCRs containing α-[32P]dCTP (800 Ci/mmol) were performed for 24 cycles. Immunoprecipitated chromatin was analyzed first in pilot experiments to ensure that PCRs occurred in the linear range of amplification. The primers for amplifying the TGM-2 gene are: TGM-2-1: 5'-ACCTGAACAAACTGGCCGAG and TGM-2-2: 5'-CAGAGAAAGGCTCCAGGTTG; for PTHLH: PTHLH-1: 5'-TACAAAGAGCAGCCGCTCA and PTHLH-2: 5'-TTACCGTGAATCGAGCTCCAG; and for GAPDH: GAPDH-1: 5'-ACTGCCAACGTGTCAGTGGT and GAPDH-2: 5'-CATACCAGGAAATGAGCTTGAC.

3D culture and immunofluorescence staining of MCF10A cells in Matrigel (BD Biosciences)

Morphogenesis analysis of stable MCF10A vector control or PIP7S knockdown clones was performed as described (55). 3D structures were harvested from the wells where they were grown after 12-16 days. Microscopy was performed on Zeiss LSM 510 Meta at the Berkeley Biological Imaging Facility. The localizations of the markers were viewed in serial confocal cross sections (x–y axis) through each colony.

For the DRB treatment, the drug (from 100 mM stock solution in 40% DMSO and 60% DMEM) was added at the final concentration of 18 uM to the 3D cultures at day 4 and then replaced with fresh DRB on day 8.

S. pombe red-white colony assay for detecting La activity

S. pombe ySH9 cells containing the La-dependent suppressor tRNA allele, tRNASerUGA-C37:10 (21), were transformed to the ura4+ phenotype with pRep4X containing no or the indicated inserts. Transformants were streaked onto plates containing Edinburgh minimal media (EMM) supplemented with adenine (10 mg/L) and amino acids lacking leucine (56). Immunoblotting with anti-His6 antibody was performed as described (56).

Results

Affinity-purification and identification of PIP7S as a Cdk9-associated protein

To identify new factors that can regulate P-TEFb, Cdk9 and its associated factors were affinity-purified from nuclear extract (NE) of F1C2 cells, a HeLa-based cell line stably expressing Cdk9-Flag (Cdk9-f) (17). Inspection by silver staining reveals two new bands (one above and one below CycT1) besides the known Cdk9-associated factors (CycT1, 7SK and HEXIM1) and the non-specific bands that were also in the negative control (Fig. 2-1A).

Analyses by mass spectrometry revealed that the upper band contained a recently described protein called BCDIN3, which resides in 7SK snRNP and functions as a methylphosphate capping enzyme for 7SK RNA (17). The lower band together with minor ones in its immediate vicinity contained a 582-amino acid protein formerly known as HDCMA18p or DKFZp564K112 and now renamed as PIP7S (P-TEFb-interaction protein for 7SK stability). Notably, the same protein was also identified independently as an associated protein of HEXIM1 and 7SK (data not shown).

Sequence analysis revealed that the N-terminal region of PIP7S is homologous to human La (hLa) protein via the La and RRM motifs (called the La domain, see Fig. 2-5C). In addition, the Drosophila ortholog of PIP7S, termed multi-sex-combs (mxc), encodes a known tumor suppressor (57). Finally, in an unbiased and comprehensive screen of gastric tumors with microsatellite instability, the PIP7S gene was identified as having the second highest frequency (41.2%) of frame shift mutations (58). The mutations occur at an oligoadenylate stretch in the sequence encoding the C-terminal region of PIP7S, resulting in truncated proteins.

PIP7S is a component of 7SK snRNP

To determine the relationship between PIP7S and the other known Cdk9-associated factors, we performed immunoprecipitations in HeLa NE with anti-Cdk9, anti-PIP7S, or anti-HEXIM1 antibodies and analyzed the immune complexes by western and northern blotting (WB and NB, Fig. 2-1B). All three antibodies precipitated PIP7S together with 7SK, HEXIM1, Cdk9 and CycT1, all known components of 7SK snRNP. Furthermore, anti-Flag immunoprecipitations in NEs of two HeLa-based cell lines expressing either Brd4-f (MCAP) (27, 28) or f-HEXIM (HH8) (18) reveal that Brd4 and PIP7S/HEXIM1/7SK exist in two

mutually exclusive P-TEFb-containing complexes (Fig. 2-1C). Together, these data indicate PIP7S as a new subunit of 7SK snRNP.

The PIP7S-7SK interaction is stable under conditions of stress and high salt treatments

How does PIP7S interact with the rest of 7SK snRNP? HEXIM1 and 7SK have been shown to dissociate from P-TEFb in cells treated with certain stress-inducing agents (e.g. actinomycin D and DRB) that globally block transcription (16-18). Notably, both compounds also dissociated HEXIM1, Cdk9 and CycT1 but not 7SK from immunoprecipitated PIP7S (Fig. 2-1D). In a separate experiment, immunoprecipitated f-PIP7S remained tightly bound to 7SK in high salt (e.g. 0.8M KCl), whereas Cdk9, CycT1 and HEXIM1 were washed away (Fig. 2-1E). These data suggest that the PIP7S-7SK binding is independent of P-TEFb/HEXIM1 and stable under conditions of stress and high salt treatments.

PIP7S interacts with P-TEFb and HEXIM1 in a 7SK-dependent manner

7SK mediates the HEXIM1-P-TEFb interaction in 7SK snRNP (18, 20, 53, 59). To extend the analysis to PIP7S, we degraded 7SK in the immobilized, anti-Cdk9 and anti-f-PIP7S immune complexes by RNase A before washing and elution. This procedure disrupted the interactions of PIP7S with both P-TEFb and HEXIM1 (Fig. 2-1F). Thus, 7SK is required for stable interactions among all the protein components (i.e. PIP7S, HEXIM1 and P-TEFb dimer) within 7SK snRNP. Consistently, the Cdk9 T-loop and the CycT1 cyclin-box, both of which are essential for the 7SK-P-TEFb binding (18, 25, 53), are also required for the PIP7S-P-TEFb interaction.

PIP7S interacts with about half of P-TEFb but all of 7SK molecules in vivo

To determine the percentages of 7SK and P-TEFb associated with PIP7S in HeLa cells, PIP7S was immunodepleted from HeLa NE. The specificity and efficiency of the depletion were illustrated by the amounts of -tubulin in the depleted NEs and the effect of non-specific IgG in a mock-depletion (Fig. 2-2A). Remarkably, this procedure caused nearly complete co-depletion of 7SK (Fig. 2-2A & 2B), suggesting that PIP7S interacts with almost all of this RNA in vivo. Besides 7SK, 48% of CycT1 and 20% of HEXIM1 were also removed. These numbers are consistent with the previous reports that about half of nuclear P-TEFb and 15-20% of HEXIM1 are sequestered in 7SK snRNP (16, 17, 28). In a separate experiment to immunodeplete Cdk9 from HeLa NE, about 20% of PIP7S was also co-depleted, suggesting that like 7SK, the majority of PIP7S exist outside the 7SK-P-TEFb snRNP.

PIP7S knockdown markedly decreases nuclear levels of 7SK RNA and 7SK snRNP

Consistent with the observation that nearly all the nuclear 7SK were associated with PIP7S, stable knockdown of PIP7S by a short hairpin (sh)RNA, shPIP7S, reduced the levels of not only PIP7S but also 7SK by more than 90% in F1C2 cells as determined by serial dilution and quantification (Fig. 2-2C, lanes 1 & 2; also data not shown). Thus, nearly all 7SK molecules relied on PIP7S for stability. Moreover, underscoring the structural role of 7SK RNA in 7SK snRNP, the globally reduced 7SK level caused by shPIP7S also prevented the association of HEXIM1 with immunoprecipitated Cdk9-f/CycT1 and thus the formation of 7SK snRNP (Fig. 2-2C,

lane 5). Significantly, however, the amount of Brd4 bound to P-TEFb increased about 2-fold at the same time (compare lanes 4 & 5). This is most likely due to a shift of the P-TEFb equilibrium from about half to nearly all in the Brd4-bound state (28) in knockdown cells. Correlating with this change, P-TEFb affinity-purified from knockdown cells also displayed ~2-fold increase in kinase activity toward GST-CTD.

To illustrate the specificity of shPIP7S action, 7SK was shown to be the only species among total nuclear RNA that was drastically reduced in PIP7S knockdown cells (Fig. 2-2D). Moreover, the elimination of 7SK RNA/snRNP could be achieved by more than one shRNAs that target different regions of PIP7S (data not shown), effectively ruling it out as an off-target effect caused by a single shRNA sequence.

PIP7S knockdown enhances the P-TEFb-dependent HIV-1 transcription

Because PIP7S knockdown liberated P-TEFb from 7SK snRNP and shifted the P-TEFb equilibrium to the active Brd4-bound state, it is expected to increase HIV-1 transcription, which is highly responsive to P-TEFb's activity (17, 18). Indeed, shPIP7S significantly increased (6.6-fold) the HIV-1 LTR-driven luciferase expression, and shPIP7S plus shHEXIM1 further enhanced (18-fold) this effect (Fig. 2-2E). It is worth noting that besides HIV-1 transcription in HeLa cells, PIP7S knockdown also stimulated P-TEFb occupancy on and transcription from two endogenous genes in a different cell type (see below). Thus, despite the fact that PIP7S is not a kinase inhibitor (data not shown), it contributes to the sequestration and inactivation of P-TEFb by maintaining the integrity of both 7SK RNA and 7SK snRNP in vivo.

The 3' oligouridylate tail of 7SK as well as the La and RRM motifs of PIP7S are all required for the PIP7S-7SK binding

How does PIP7S control the stability of 7SK snRNA? hLa binds nascent Pol III transcripts via the 3' -UUU-OH sequence and sequester them away from exonucleases (60-62). The La motif confers the binding specificity for -UUU-OH. Coincidentally, 7SK is a Pol III transcript with the signature 3' oligouridylate tail (-UUUU-OH). Moreover, PIP7S is homologous to hLa and its La domain has all of the invariant and highly conserved amino acids involved in recognition of 3' UUU-OH (63). Based on these facts as well as the observation that nearly all the nuclear 7SK were associated with PIP7S, we postulated that PIP7S displays intrinsic La activity, which is responsible for 7SK's stability in vivo.

Three lines of evidence exist to support this notion. First, the 3' -UUUU-OH of 7SK and the La domain of PIP7S are both required for the PIP7S-7SK binding. In an electrophoretic mobility shift assay (EMSA), homogeneously purified PIP7S protein (Materials and Methods; also Fig. 2-3E) readily formed a complex with 32P-labeled 7SK, which could be partially super-shifted by anti-PIP7S but not anti-HEXIM1 antibody (Fig. 2-3A). More importantly, deletion of the 7SK 3' –UUUU-OH (Δ4U's, Fig. 2-3B), removal of the PIP7S La motif (ΔLM), or substitution of a highly conserved residue (Y127D) in the PIP7S RRM (Fig. 2-3C) all dramatically reduced the PIP7S-7SK binding. Moreover, the ΔLM and Y127D mutations also significantly blocked the 7SK snRNP formation in vivo (Fig. 2-3D and Fig. 2-5). These data agree completely with the description of interactions between

authentic La proteins and Pol III transcripts with 3’ UUU-OH.

PIP7S enhances the binding of HEXIM1 to 7SK

Further evidence implicating a La-like function in PIP7S is the demonstration by EMSA that PIP7S markedly enhanced the binding of HEXIM1 to 7SK to form a robust, slow-migrating complex (Fig. 2-3E, lanes 6-8), which could be super-shifted or disrupted by anti-PIP7S or anti-HEXIM1 antibodies, respectively (data not shown). Thus, besides its requirement for 7SK stability, the strong and independent PIP7S-7SK binding (Fig. 2-1D & 1E) may help recruit HEXIM1 to 7SK snRNP. This is reminiscent of classic La proteins that can facilitate the assembly of U snRNPs (64, 65).

The PIP7S N-terminal region partially compensates for the loss of La function to provide RNA 3' end protection

The final evidence in support of PIP7S's La-like activity is the demonstration that PIP7S partially compensated for the loss of La function in an established S. pombe-based assay that depends on ectopic La protein for RNA 3’ end protection against exonuclease digestion (21, 56). In this system, La activity, monitored by a red-white colony assay, is required for the 3' end protection-dependent maturation of a suppressor tRNASerUGA that suppresses a premature UGA stop codon in ade6-704. Under limiting adenine conditions, unsuppressed cells accumulated red pigment (e.g. cells with vector pRep4X, Fig. 2-4A). Expression of hLa, however, suppressed the stop codon and turned colonies white. Although wild-type PIP7S displayed little activity, PIP7S 1-217 containing just the La and RRM motifs (Fig. 2-4C) significantly reduced red pigment (Fig. 2-4A). Thus, when separated from the C-terminal region, the PIP7S La domain could indeed display RNA 3' protection activity that is characteristic of authentic La proteins. However, the PIP7S C-terminal region may determine functional specificity, allowing PIP7S to focus on 7SK rather than pre-tRNAs as its major or perhaps sole target in vivo (see Discussion below).

hLa cannot substitute for PIP7S in 7SK snRNP

If the PIP7S N-terminal region can function like hLa to protect RNA 3’ ends, we asked whether hLa could also replace PIP7S to enter 7SK snRNP. Data in Fig. 2-4D indicate that in sharp contrast to f-PIP7S, f-hLa was only able to co-precipitate less than 5% of total 7SK RNA but not any known protein components of 7SK snRNP. Consistently, endogenous 7SK snRNP obtained by anti-Cdk9 IP contained PIP7S but not hLa, despite the fact that the latter is highly abundant and readily detectable in NE (Fig. 2-4E). Thus, although hLa is homologous to PIP7S, it cannot substitute for PIP7S to enter 7SK snRNP.

The PIP7S C-terminus deleted in human gastric tumors and absent in hLa is required for binding to 7SK snRNP

Given that PIP7S is required for sequestering P-TEFb in inactive 7SK snRNP, we tested how the PIP7S mutations detected in gastric tumors (58) would affect this ability. To mimic these mutations, Δ1A, Δ2A and Δ3A, with a deletion of 1, 2 and 3 adenosines, respectively, from a microsatellite repeat of 8 A's (nucleotides 1206-1213) in the PIP7S C-terminal region, were created (Fig. 2-5B). While Δ3A, which has a restored open reading frame after skipping one amino acid, interacted with the 7SK

snRNP components normally, Δ1A and Δ2A, both of which acquire premature stop codons, almost completely lost this ability (Fig. 2-5A). Moreover, deletion of just the last 21 amino acids also largely abolished the interaction (Fig. 2-5). Thus, the PIP7S C-terminus deleted in tumors and absent in hLa is essential for the PIP7S-dependent sequestration of P-TEFb. This observation, together with the demonstration that replacing the PIP7S RRM motif, and to a lesser extent the La motif, with those of hLa prevented the interactions with 7SK snRNP by PIP7S mutants pLhR and hLpR, respectively (Fig. 2-5), explain why hLa was not detected in 7SK snRNP (Fig. 2-4D & E).

PIP7S knockdown in MCF10A cells disrupts the formation of highly organized, multi-cellular acini with a well-defined border

The observation that the tumor-derived PIP7S mutants failed to sequester and inactivate P-TEFb, which would otherwise promote cell growth (11, 21, 22, 26, 52), implicates the PIP7S deficiency as a potential contributor to cancer. Further support of this idea comes from PIP7S's homology to the Drosophila tumor suppressor MXC. Overall, the two proteins are 25% identical and 43% similar (6e-29). An even stronger homology (40% identity and 62% similarity; 3e-24) exists in their N-terminal regions that contain the La domain. Based on these observations, we postulated that like Drosophila MXC, PIP7S is a human tumor suppressor.

To test this hypothesis, we stably expressed shPIP7S in human normal mammary epithelial cell line MCF10A, which is an accepted model system for studying transformation (55, 66). PIP7S knockdown significantly decreased the levels of both 7SK RNA and 7SK snRNP (Fig. 2-6A&B) as in HeLa cells. When cultured in a three-dimensional (3D) reconstituted basement membrane (Matrigel) (66), MCF10A with an empty vector underwent morphological differentiation to form a multi-cellular structure with an organized spherical arrangement reminiscent of breast acinus in vivo (Fig. 2-6C). In contrast, colonies formed by knockdown cells were disorganized and showed irregular borders similar to those of the breast cancer cell line MCF7 (Fig. 2-6C).

PIP7S knockdown disrupts cell polarity, blocks mammary epithelial cell differentiation and causes transformation

To test whether PIP7S knockdown would induce morphological changes often associated with malignant transformation, the localizations of several key protein markers within the 3D colonies were examined by immunofluorescence. First, α6 integrin, a marker for epithelial polarity, was found predominantly at the basal and to a lesser extent the lateral side of a control colony, which also displayed a well-formed lumen (Fig. 2-6D). In contrast, it was completely mislocalized within the colony of knockdown cells and detected at the apical side and among cells that filled the central cavity (Fig. 2-6D), indicating a severe disruption of apicobasal polarity. Furthermore, E-cadherin, normally restricted to cell-cell junction, was also dramatically disorganized in the knockdown colony (Fig. 2-6D). Finally, Ki67, a proliferation marker, was detected much more frequently in the knockdown colony than in the control (Fig. 2-6D), indicating a lack of cell cycle withdrawal and apoptosis for many knockdown cells at this stage.

Importantly, all these shPIP7S-mediated changes could be completely reversed by the introduction of a siRNA-resistant PIP7S cDNA that expresses wild-type f-PIP7S (Fig. 2-6E), which was properly incorporated into 7SK snRNP as expected (Fig. 2-6F). Thus, the observed changes are unlikely to be off-target effects caused by a particular shPIP7S sequence.

Elevated P-TEFb activity is essential for shPIP7S-induced MCF10A transformation

To determine whether increased P-TEFb activity is required for shPIP7S to induce MCF10A transformation, we incubated DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole), a well-documented inhibitor of P-TEFb kinase (67, 68), with the vector control and knockdown cells grown in Matrigel (Fig. 2-7A). Although P-TEFb and Pol II elongation are efficiently blocked by greater than 50 μM DRB (67, 68), the ability of control cells to form acini-like structure was minimally affected by 15-20 μM of the drug (18 μM in Fig. 2-7A). In contrast, the same amount of DRB completely blocked transformation of PIP7S knockdown cells, which formed colonies with well-organized structures like control cells (Fig. 2-7A). Thus, the DRB-mediated partial reduction in P-TEFb activity effectively reversed the shPIP7S-induced transformation, consistent with the notion that P-TEFb is the target of PIP7S suppression of tumorigenesis.

PIP7S knockdown enhances P-TEFb's occupancy on and transcription from two key tumor-related genes

Finally, we tested whether the shPIP7S-mediated release of P-TEFb from 7SK snRNP would enhance P-TEFb's occupancy on and transcription from endogenous genes in PIP7S knockdown MCF10A cells. To this end, the PTHLH (parathyroid hormone-like hormone) and TGM-2 (transglutaminase 2) genes were selected because of their demonstrated roles in breast cancer development and the availability of reagents in the lab. Both genes display elevated expression in breast cancer tissues and cell lines and are strongly correlated with tumor invasion and metastasis (69, 70).

Consistent with the transformed phenotypes of PIP7S knockdown cells, a significant increase in transcription from both PTHLH and TGM-2 but not the P-TEFb-independent housekeeping gene GAPDH (glyseraldehyde-3-phosphate dehydrogenase) (21, 22, 50, 52) was detected by quantitative real-time PCR (qRT-PCR) in these cells (Fig. 2-7B). Importantly, the elevated PTHLH and TGM-2 expression was inhibited by flavopiridol (Fig. 2-7C), a specific Cdk9 inhibitor (12, 49), indicating these two genes as P-TEFb-dependent. Consistently, an enhanced occupancy of P-TEFb at the PTHLH and TGM-2 loci near the 3' end of the ORFs was also revealed by chromatin immunoprecipitation (ChIP, Fig. 2-7D). In contrast, no more than the background level of P-TEFb (obtained via control IgG) was detected at the GAPDH locus and this situation was unchanged by PIP7S knockdown. Together, these data are consistent with the model that PIP7S knockdown disrupts mammary epithelial differentiation and causes transformation through activating P-TEFb and increasing P-TEFb-dependent expression of key tumor-promoting genes.

Discussion

The 7SK snRNA has been likened to a molecular scaffold that holds all the protein components together within 7SK snRNP (20, 53, 59). Without this RNA, HEXIM1 is unable to interact stably with P-TEFb and inhibit Cdk9's kinase activity. For such an important function ascribed to 7SK, one would imagine that there must exist cellular mechanisms to ensure that its stability, and in turn, the integrity of 7SK snRNP are not compromised. Indeed, nature finds an effective, and in hindsight, rather obvious solution in PIP7S, a La-related protein, which binds 7SK through its 3' poly(U) tail and is requited for the stable accumulation of this RNA and formation of 7SK snRNP. Further underscoring the importance of 7SK, a recent study (17) demonstrates that the 5' end of this RNA is also protected in 7SK snRNP by a specific methylphosphate capping enzyme called BCDIN3.

Within 7SK snRNP, the PIP7S-7SK binding is direct, strong and independent of other proteins (Fig. 2- 1D, 1E & 3). This observation raises the possibility that the PIP7S-7SK sub-complex functions as a preexisting unit to nucleate the formation of 7SK snRNP in vivo. It has been suggested that while there are probably two copies each of Cdk9, CycT1 and HEXIM1 (or HEXIM2) proteins in 7SK snRNP, there is only one copy of 7SK RNA in this complex (27, 28, 71, 72). Given our inability to detect any PIP7S dimers (data not shown), we suspect that the PIP7S-7SK sub-complex exists as a monomeric unit within 7SK snRNP.

In contrast to PIP7S, the prototypical hLa, which is highly abundant and homologous to PIP7S, was not detected in 7SK snRNP (Fig. 2-4D & E). Moreover, only less than 5% of 7SK RNA were associated with hLA in HeLa NE (Fig. 2-4D), whereas nearly all were bound by PIP7S (Fig. 2-2A & 2B). Given the strong homology between PIP7S and hLa (their La motifs are 40% identical and 65% similar), why would 7SK snRNA/snRNP preferentially employ PIP7S but not hLa to maintain their stability in vivo? We believe that part of the answer may lie in the unique C-terminal region of PIP7S that is frequently deleted in human cancers and absent in hLa. In support of this idea, PIP7S mutants lacking the C-terminal region were unable to bind 7SK and enter 7SK snRNP (Fig. 2-5). Besides the C-terminal region, certain amino acids that distinguish the La domains of PIP7S and hLa also contribute to the functional difference between these two proteins. This is illustrated by the demonstration that replacing the PIP7S La and RRM motifs with those of hLa almost completely disrupted the interactions of PIP7S with 7SK snRNA/snRNP (Fig. 2-5). Together, these data provide a likely explanation for the specific involvement of PIP7S but not hLa in controlling 7SK stability in vivo.

Interestingly, the functional difference between PIP7S and hLa is also evident when it comes to their roles in tRNA processing. While hLa was fully capable of promoting tRNA maturation thereby compensating for the loss of La function in the fission yeast strain ySH9, only the C-terminally truncated PIP7S but not the full-length protein displayed a partial activity (Fig. 2-4A). It is possible that the C-terminal region and the unique residues in the N-terminal half of PIP7S render the protection of pre-tRNA 3' ends not a top priority of PIP7S under normal conditions. However, in gastric tumors with PIP7S frame shift mutations (58), the situation could be quite different. Here, the C-terminally truncated PIP7S may increase tRNA levels

as has been implicated in fission yeast (Fig. 2-4A), which could potentially contribute to malignant transformation according to the current model that increased tRNAs and other pol III transcripts promote tumorigenesis (73, 74). If this hypothesis is proven, it will indicate a dual role for the PIP7S frame shift mutations in cancer. On one hand, the truncated proteins are unable to sequester P-TEFb into inactive 7SK snRNP (Fig. 2-5) and thus would lead to a loss of tumor suppressor activity. On the other hand, the mutant PIP7S may also acquire a gain-of-function in promoting tumorigenesis in a manner reminiscent of that of activated oncogenes.

A key observation of the current study is that PIP7S are associated with nearly all the 7SK RNA and required to maintain their stability in vivo (Fig. 2-2). Considering that 7SK is a very abundant RNA species (~2 x 10^5 per cell) (50), PIP7S is expected to be as abundant as 7SK. Importantly, stable and simultaneous knockdown of both molecules to less than 10% of their normal levels in HeLa and MCF10A cells not only failed to decrease cell viability, it even caused transformation of MCF10A cells (Fig. 2-6 & 2-7). These observations suggest that PIP7S and 7SK are not required for cell survival. Rather, they contribute to the suppression of growth and transformation. We noticed that this conclusion apparently contradicts a previous claim that 7SK depletion by transfected siRNA leads to apoptosis (75). Since only one siRNA sequence was tested and there was no attempt to reverse the phenotype with functional, siRNA-resistant 7SK, the possibility of an off-target effect caused by the siRNA cannot be ruled out. Alternatively, the difference could be due to the fact that both PIP7S and 7SK were co-depleted in our system, whereas the previous study involves the depletion of 7SK only.

Notably, the anti-growth/anti-tumor function of PIP7S agrees well with the previous demonstration that MXC, the Drosophila homolog of PIP7S, is a confirmed tumor suppressor for preventing the overproliferation of lymph glands and circulating hemocytes in larvae (57). As for 7SK, previous studies have also assigned a growth inhibitory role for at least a portion of this RNA that are associated with P-TEFb. For example, induction of hypertrophic growth of cardiac myocytes has been shown to involve the dissociation of 7SK from P-TEFb and activation of Cdk9 kinase, which is limiting for cell growth (22). Conversely, in murine erythroleukemia cells (MELC) that are induced to undergo terminal division and differentiation by HMBA, the P-TEFb equilibrium is overwhelmingly shifted toward the inactive 7SK snRNP, where 7SK mediates the binding and inhibition of P-TEFb by HEXIM1 (21). These findings, coupled with the demonstrations that both Brd4 and HEXIM1 affect cell growth albeit in opposite manners, all point to P-TEFb as a common target for the global control of cell growth and differentiation by its associated regulators (26, 51). Since a key characteristic of cancer cells is unchecked growth, it is conceivable that the loss of PIP7S/7SK function in human cells results in a significant increase in the active pool of P-TEFb, leading to the elevated expression of key tumor-promoting genes such as PTHLH and TGM-2, unsuppressed cell growth and ultimately transformation. Our results thus implicate an important and direct role for P-TEFb and its control of transcriptional elongation in tumorigenesis.

Fig. 2-1 PIP7S is a component of 7SK snRNP. A. Cdk9-f and its associated factors (αFlag IP) were affinity-purified from F1C2 NE and analyzed on a silver-stained SDS-gel, with their identities indicated on the right. HeLa NE was used in a parallel procedure for control. The band marked with "Δ" was not reproducibly seen and contained no identifiable protein. B. Immunoprecipitates (IP) obtained with the indicated antibodies from HeLa NE were analyzed by western (WB) and northern blotting (NB). C. The Cdk9-CycT1 heterodimers and their associated factors were isolated by anti-Flag immunoprecipitation (IP, left panel) from NEs (right panel) of two HeLa-based cell lines expressing Brd4-f (MCAP) and f-HEXIM1 (HH8), respectively, and analyzed as in B. D. HeLa cells were treated with the indicated compounds. IP with the indicated antibodies were analyzed as in B. E. NEs (right panels) from HeLa cells, which were transfected with a vector expressing no protein (v) or f-PIP7S, were subjected to anti-Flag IP. The immobilized immune-complexes were washed with buffer D containing the indicated amounts of KCl. f-PIP7S and its associated factors were eluted and analyzed (left panels) as in B. F. The same NEs in E were subjected to IP with the indicated antibodies. The immobilized immune complexes were incubated with (+) or without (-) RNase A before washing. The eluted complexes were analyzed as in B.

Fig. 2-1

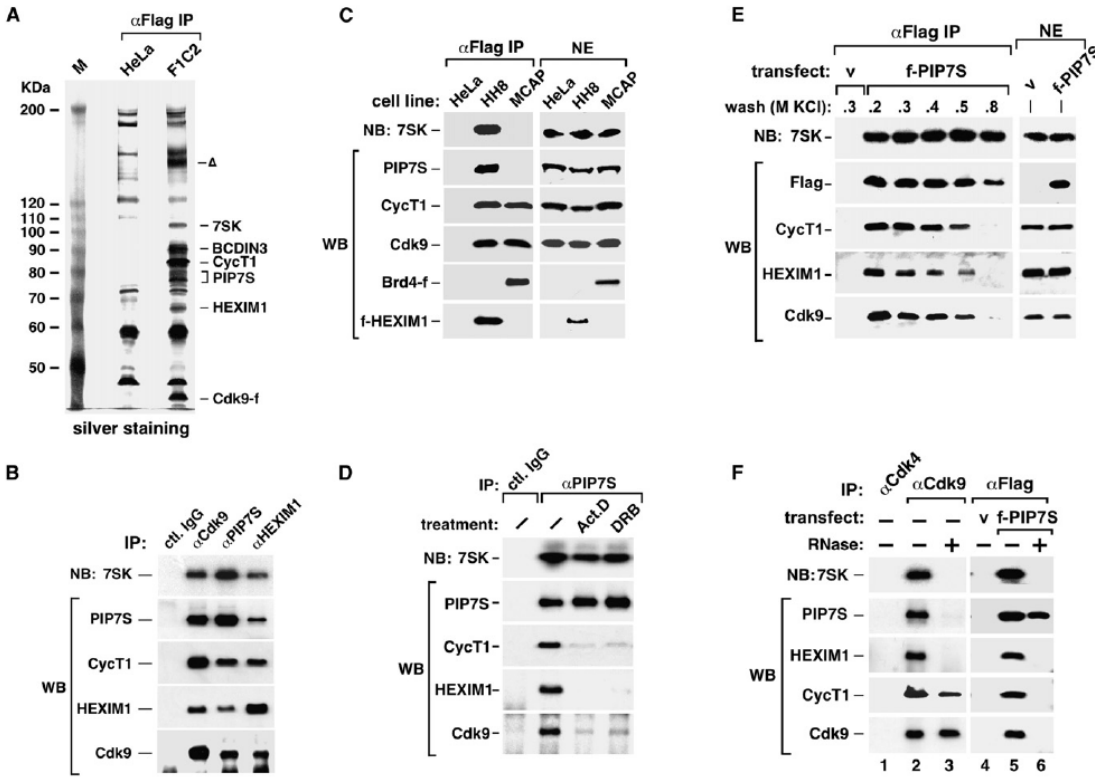

A
αFlag IP
M
HeLa
F1C2
KDa
200
120
110
100
90
80
70
60
50
Δ
7SK
BCDIN3
CycT1
PIP7S
HEXIM1
Cdk9-f
silver staining
B
IP:
ctl. IgG
αCdk9
αPIP7S
αHEXIM1
NB: 7SK
PIP7S
CycT1
HEXIM1
Cdk9
WB
C
αFlag IP
NE
cell line:
HeLa
HH8
MCAP
NB: 7SK
PIP7S
CycT1
Cdk9
Brd4-f
f-HEXIM1
WB
D
IP:
ctl. IgG
αPIP7S
treatment:
Act.D
DRB
NB: 7SK
PIP7S
CycT1
HEXIM1
Cdk9
WB
E
αFlag IP
NE
transfect:
v
f-PIP7S
wash (M KCl):
.3
.2
.4
.5
.8
NB: 7SK
Flag
CycT1
HEXIM1
Cdk9
WB
F
IP:
αCdk4
αCdk9
αFlag
transfect:
v
f-PIP7S
RNase:
NB:7SK
PIP7S
HEXIM1
CycT1
Cdk9
WB
1
2
3
4
5
6

Fig. 2-2 PIP7S knockdown decreases nuclear levels of 7SK RNA and 7SK snRNP but increases the Brd4-bound P-TEFb and P-TEFb-dependent transcription. A. NEs were subjected to immunodepletion with the indicated antibodies and then analyzed by northern (NB) and western blotting (WB). B. The levels of the indicated components in the depleted NEs were normalized to those of α-tubulin and quantified based on serial dilutions, with those in mock-depleted NE artificially set to 100%. The error bars represent mean +/- SD. C. Stable expression of shPIP7S from a Cre-controlled cassette was induced (+) or uninduced (-) in HeLa or F1C2 cells expressing Cdk9-f. The indicated factors in NEs (lanes 1 & 2) and anti-Flag IP (lanes 3-5) were analyzed as in A. D. Total RNA was extracted from NEs of HeLa cells expressing or not expressing shPIP7S, resolved on a denaturing gel, and stained with GelRed. E. HeLa cells were co-transfected with the HIV-1 LTR-luciferase reporter construct and a vector expressing no RNA, shPIP7S, shHEXIM1 or both. Luciferase activities were measured and the error bars represent mean +/- SD.

Fig. 2-2

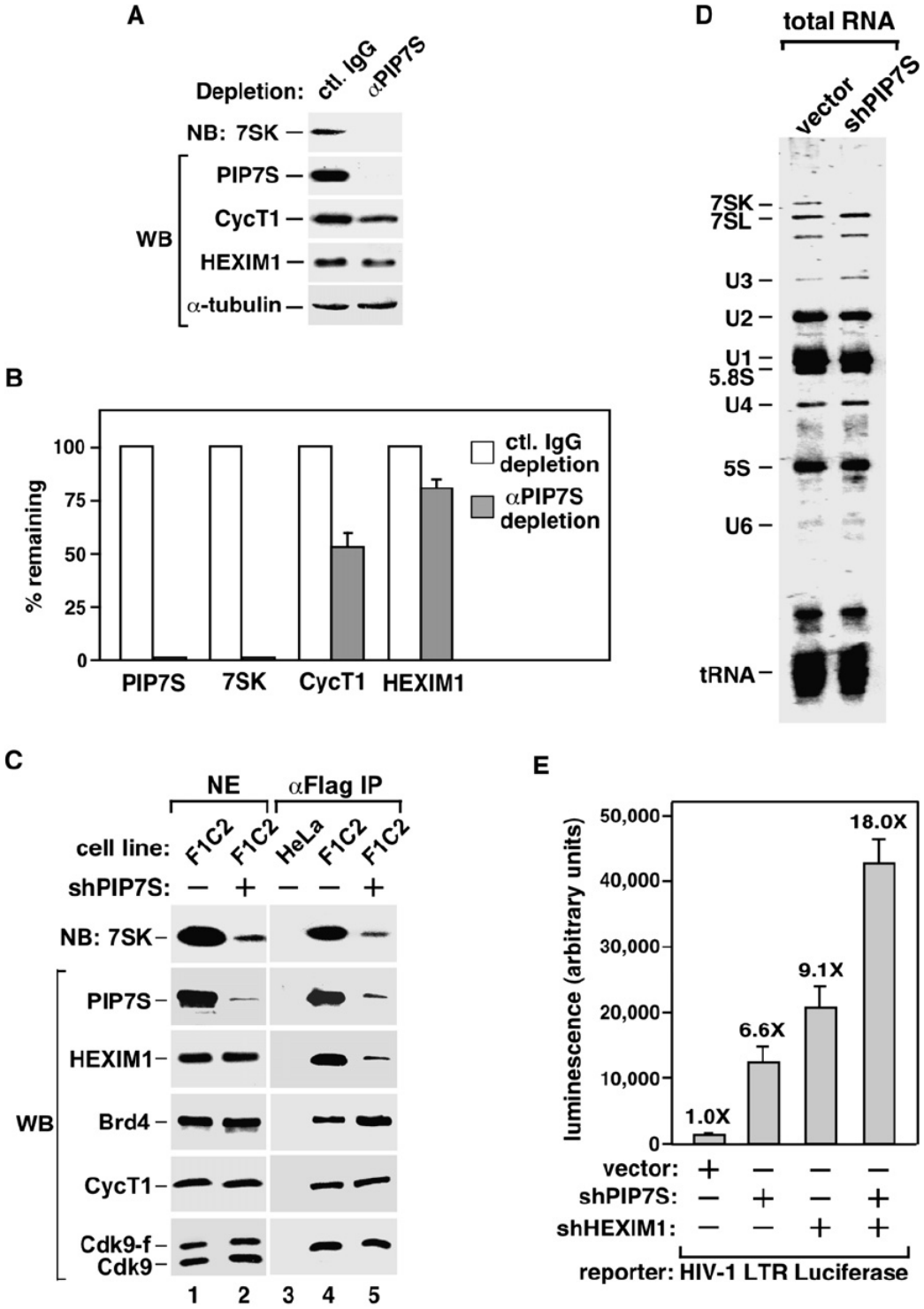

A
Depletion: ctl. IgG αPIP7S
NB: 7SK
PIP7S
CycT1
HEXIM1
α-tubulin
WB
B
% remaining
100
75
50
25
0
ctl. IgG depletion
αPIP7S depletion
PIP7S
7SK
CycT1
HEXIM1
C
NE
αFlag IP
cell line: F1C2 F1C2 HeLa F1C2 F1C2
shPIP7S: – + – – +
NB: 7SK
PIP7S
HEXIM1
Brd4
CycT1
Cdk9-f
Cdk9
WB
1 2 3 4 5
D
total RNA
vector
shPIP7S
7SK
7SL
U3
U2
U1
5.8S
U4
5S
U6
tRNA
E
luminescence (arbitrary units)
50,000
40,000
30,000
20,000
10,000
0
1.0X
6.6X
9.1X
18.0X
vector: + – – –
shPIP7S: – + – +
shHEXIM1: – – + +
reporter: HIV-1 LTR Luciferase

Fig. 2-3 The 7SK-PIP7S binding requires both the poly(U) tail of 7SK and the La domain of PIP7S and recruits HEXIM1 to 7SK snRNP. A. Affinity-purified PIP7S, anti-PIP7S (αPIP7S) and anti-HEXIM1 (αHEXIM1) antibodies were incubated as indicated with ^{32}P-labeled 7SK and analyzed by EMSA. B. Wild-type 7SK or its mutant Δ4U's was added to EMSA reactions containing increasing amounts of PIP7S (in 2-fold increments). C. Flag-tagged wild-type PIP7S and its mutants ΔLM and Y127D were affinity-purified, adjusted to similar concentrations by anti-Flag western blotting (right), and analyzed for their interactions with 7SK by EMSA (left). D. Flag-tagged wild-type PIP7S or Y127D was expressed in HeLa cells. Anti-Flag IP derived from NEs was analyzed by western (WB) and northern blotting (NB) as indicated. E. Left: Flag-tagged HEXIM1 and PIP7S were affinity-purified from transfected HeLa cells and examined on a silver-stained SDS-gel. Right: The two proteins were added at the indicated concentrations to EMSA reactions containing ^{32}P-labeled 7SK.

Fig. 2-3

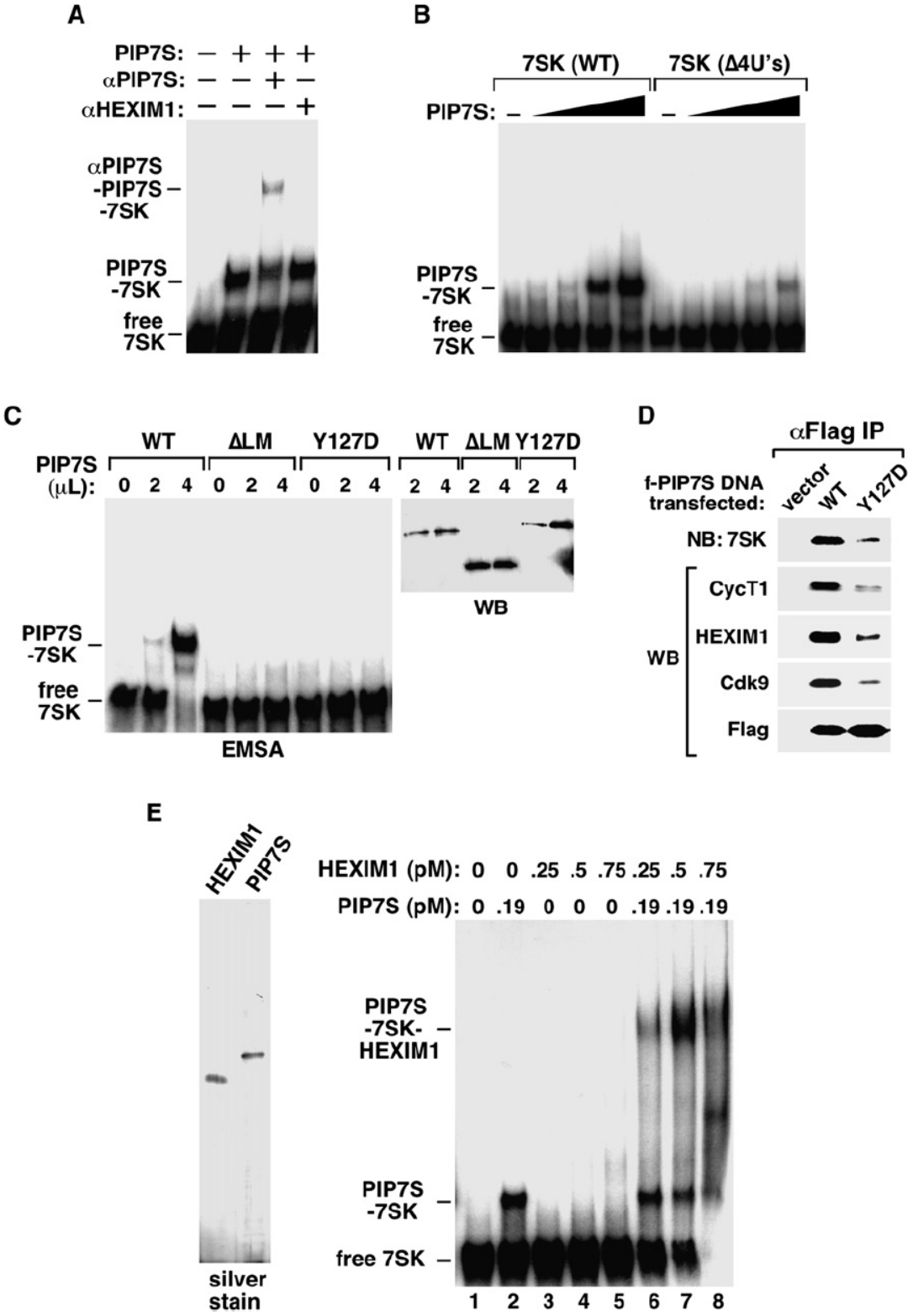

A
PIP7S: – + + +
αPIP7S: – – + –
αHEXIM1: – – – +
αPIP7S -PIP7S -7SK
PIP7S -7SK
free 7SK
B
7SK (WT)
7SK (Δ4U's)
PIP7S: – –
PIP7S -7SK
free 7SK
C
WT ΔLM Y127D WT ΔLM Y127D
PIP7S (μL): 0 2 4 0 2 4 0 2 4 2 4 2 4 2 4
PIP7S -7SK
free 7SK
EMSA
WB
D
αFlag IP
f-PIP7S DNA transfected: vector WT Y127D
NB: 7SK
WB
CycT1
HEXIM1
Cdk9
Flag
E
HEXIM1
PIP7S
silver stain
HEXIM1 (pM): 0 0 .25 .5 .75 .25 .5 .75
PIP7S (pM): 0 .19 0 0 0 .19 .19 .19
PIP7S -7SK- HEXIM1
PIP7S -7SK
free 7SK
1 2 3 4 5 6 7 8

Fig. 2-4 While PIP7S displays authentic La activity to protect RNA 3' end from exonuclease digestion, hLa cannot substitute for PIP7S in 7SK snRNP. A. PIP7S N-terminal region partially rescues defective tRNA processing in S. pombe lacking endogenous La function. ySH9 cells containing a La-dependent suppressor tRNA as well as a premature stop codon in ade6-704 were transformed with pRep4X containing no or the indicated inserts. Transformants were streaked onto plates supplemented with adenine. B. Expression of the His-tagged PIP7S was confirmed by anti-His6 western blotting. C. A diagram showing the domain structures of hLa, wild-type PIP7S and PIP7S 1-217. D. HeLa cells were transfected with a vector expressing nothing, f-PIP7S or f-hLa. Anti-Flag immunoprecipitates (IP) from NEs (bottom panel) were analyzed by western (WB) and northern blotting (NB) as indicated. E. IP obtained with the indicated antibodies from HeLa NE were analyzed as in D.

Fig. 2-4

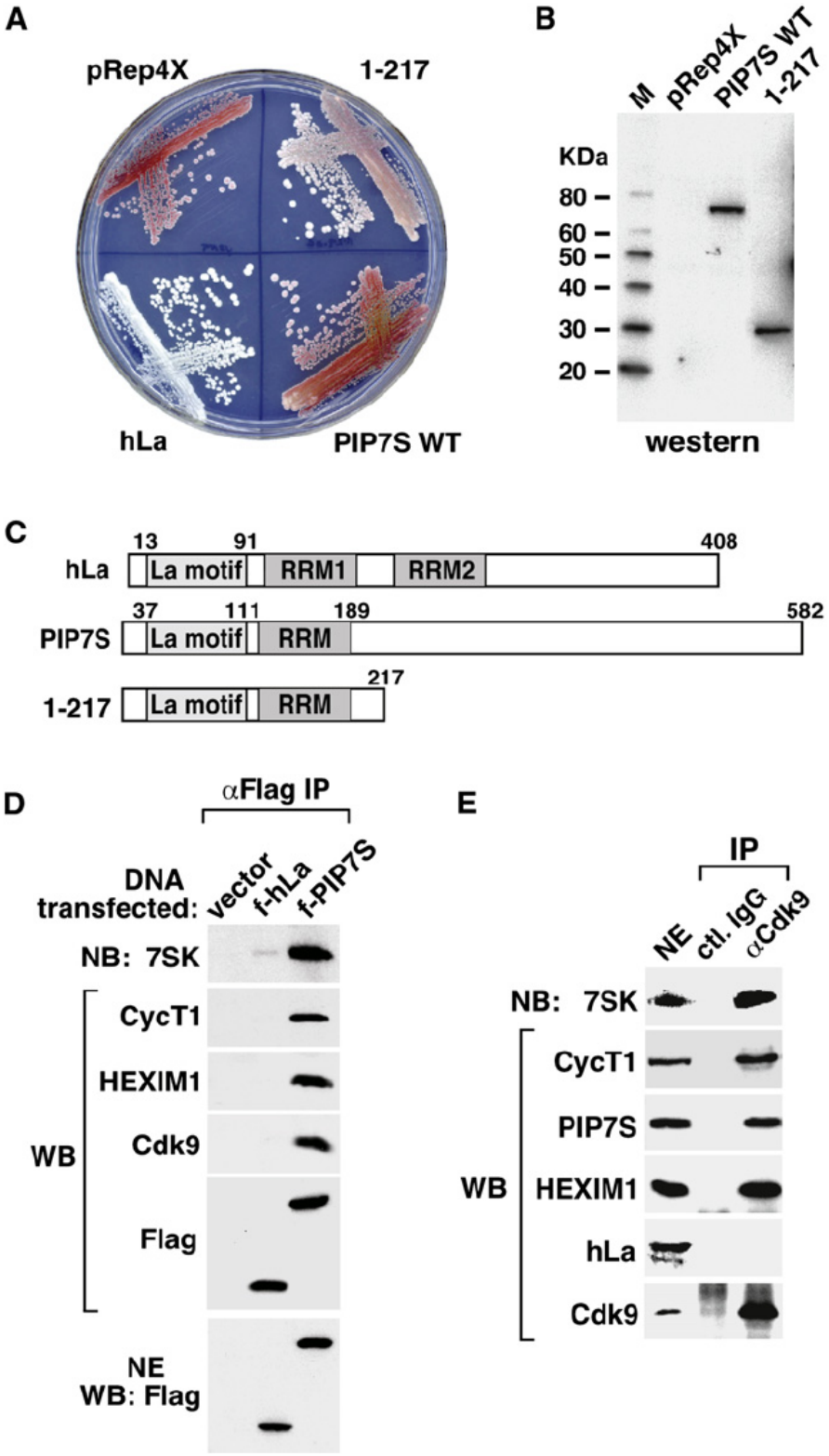

A
pRep4X
1-217
hLa
PIP7S WT
B
M
pRep4X
PIP7S WT
1-217
KDa
80
60
50
40
30
20
western
C
hLa
13
91
408
La motif
RRM1
RRM2
PIP7S
37
111
189
582
La motif
RRM
1-217
217
La motif
RRM
D
αFlag IP
DNA transfected:
vector
f-hLa
f-PIP7S
NB: 7SK
CycT1
HEXIM1
Cdk9
Flag
WB
NE WB: Flag
E
IP
NE
ctl. IgG
αCdk9
NB: 7SK
CycT1
PIP7S
HEXIM1
hLa
Cdk9
WB

Fig. 2-5 The N-terminal La domain and the C-terminus missing in cancer cells are both important for PIP7S to interact with 7SK snRNP. A series of Flag-tagged PIP7S mutants as diagrammed in B were expressed in HeLa cells. Anti-Flag IP derived from NEs were analyzed by western (WB) and northern blotting (NB) as indicated in A, with quantification of the binding data summarized in the last two columns in B. hLa: human La protein. RRM: RNA recognition motif.

Fig. 2-5

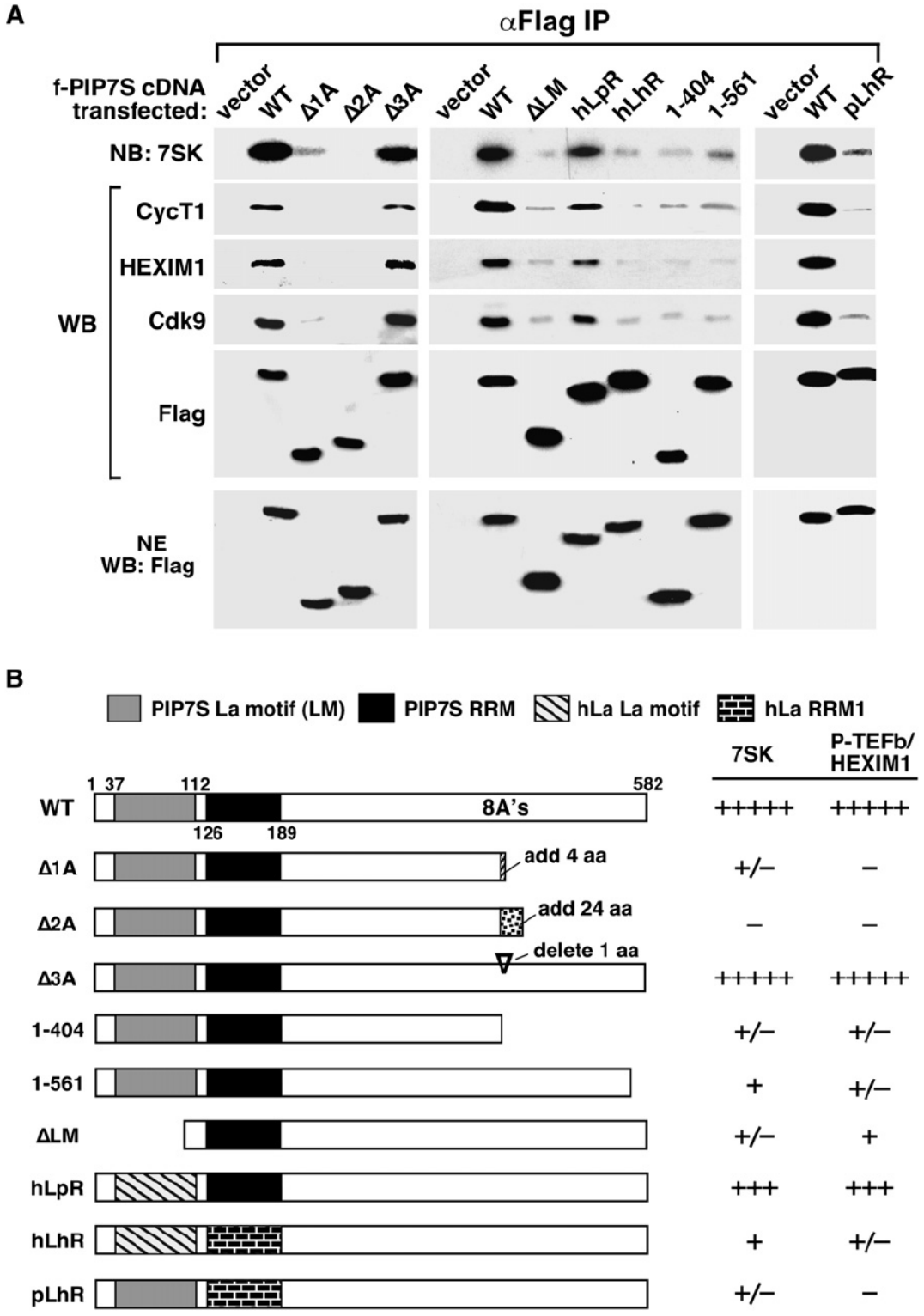

A
αFlag IP
f-PIP7S cDNA transfected:
vector
WT
Δ1A
Δ2A
Δ3A
vector
WT
ΔLM
hLpR
hLhR
1-404
1-561
vector
WT
pLhR
NB: 7SK
WB
CycT1
HEXIM1
Cdk9
Flag
NE
WB: Flag
B
PIP7S La motif (LM)
PIP7S RRM
hLa La motif
hLa RRM1
7SK
P-TEFb/ HEXIM1
1 37 112 582
WT
8A's
126 189
Δ1A
add 4 aa
Δ2A
add 24 aa
Δ3A
delete 1 aa
1-404
1-561
ΔLM
hLpR
hLhR
pLhR
+++++ +++++
+/– –
– –
+++++ +++++
+/– +/–
+ +/–
+/– +
+++ +++
+ +/–
+/– –

Fig. 2-6 PIP7S knockdown in MCF10A cells disrupts 7SK snRNP and mammary epithelial cell differentiation and causes transformation. A. NEs from stable MCF10A clones containing the vector expressing nothing or shPIP7S were examined by western (WB) and northern blotting (NB) for the indicated factors. B. Immunoprecipitates (IP) obtained with the indicated antibodies were analyzed as in A. C. Microscopic examination of the morphology of colonies formed in Matrigel from the MCF10A vector control or PIP7S knockdown cells, or the MCF7 cells. D. Colonies formed in Matrigel from the MCF10A vector control (vector) or PIP7S knockdown (shPIP7S) cells were harvested at day 12 and stained with the indicated antibodies or the DNA dye. E. PIP7S knockdown cells (shPIP7S) and their derivatives expressing shRNA-resistant f-PIP7S (rescue) were cultured in Matrigel and analyzed as in D. F. NEs from the indicated MCF10A clones were examined by WB (upper panel). f-PIP7S and its associated factors in αFlag IP were analyzed as in A.

Fig. 2-6

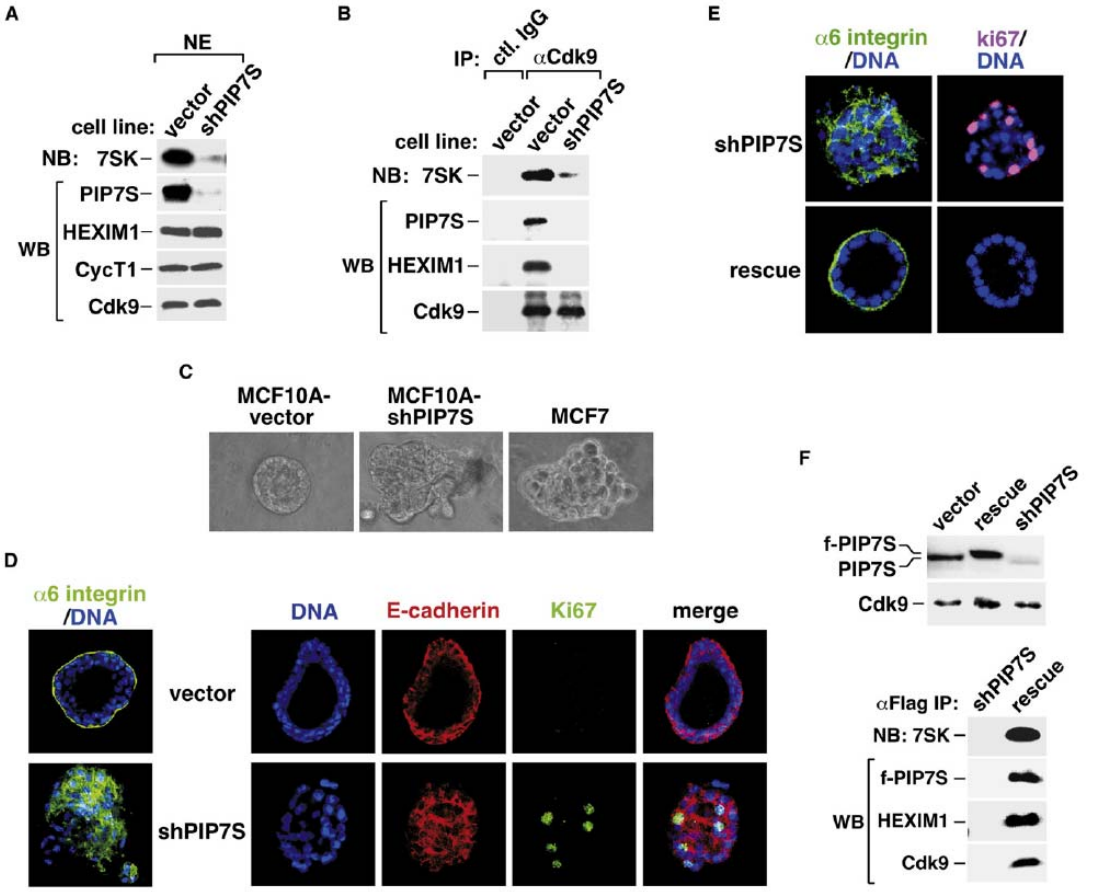
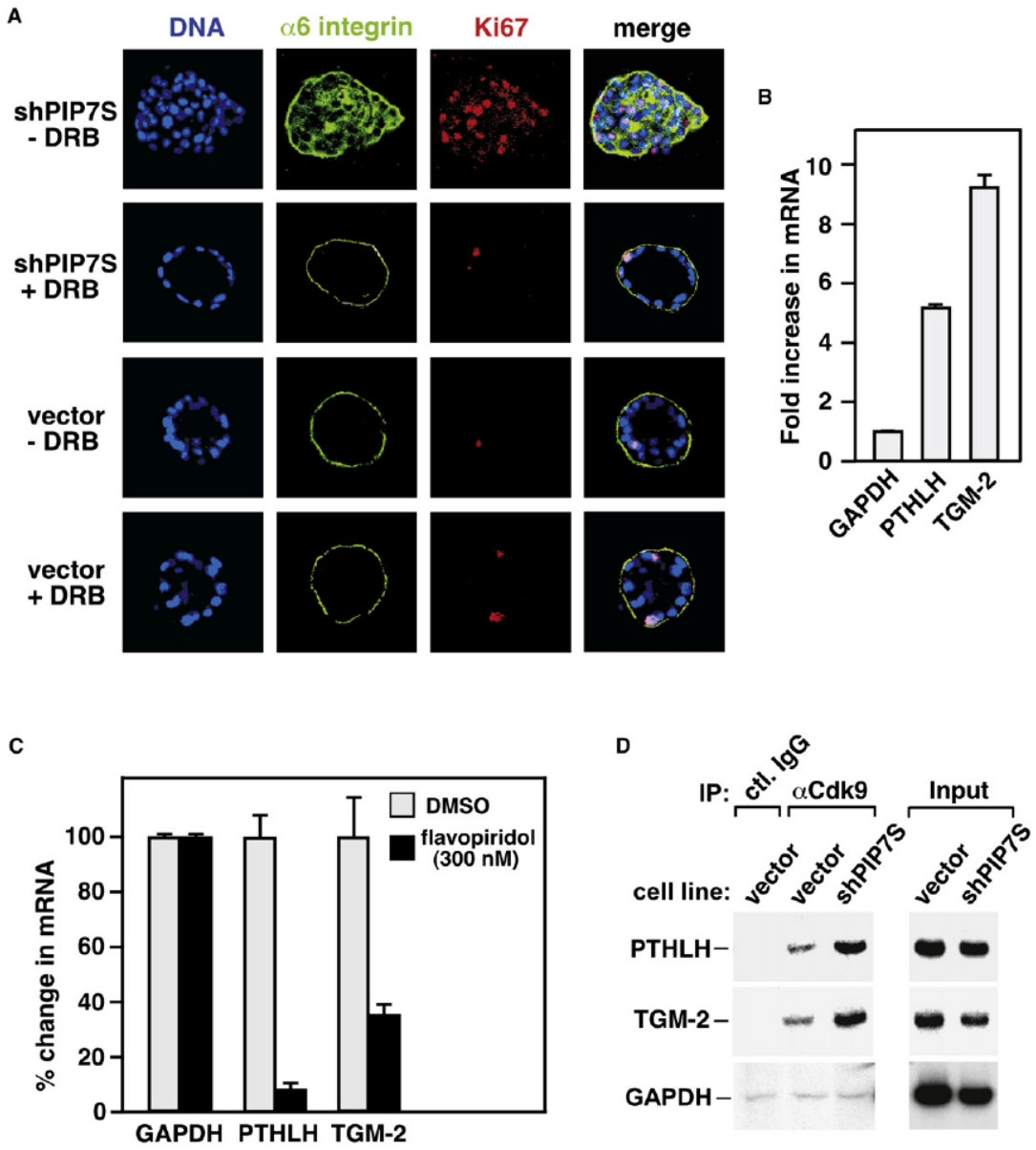

A
NE
cell line: vector shPIP7S
NB: 7SK
PIP7S
HEXIM1
CycT1
Cdk9
WB
B
IP: ctl. IgG αCdk9
cell line: vector vector shPIP7S
NB: 7SK
PIP7S
HEXIM1
Cdk9
WB
E
α6 integrin /DNA
ki67/ DNA
shPIP7S
rescue
C
MCF10A-vector
MCF10A-shPIP7S
MCF7
D
α6 integrin /DNA
DNA
E-cadherin
Ki67
merge
vector
shPIP7S
F
vector rescue shPIP7S
f-PIP7S
PIP7S
Cdk9
αFlag IP: shPIP7S rescue
NB: 7SK
f-PIP7S
HEXIM1
Cdk9
WB

Fig. 2-7 Elevated P-TEFb activity in PIP7S knockdown cells enhances P-TEFb's occupancy on and transcription from two key tumor-related genes and is essential for transformation. A. MCF10A vector control (vector) and PIP7S knockdown (shPIP7S) cells were allowed to form colonies in Matrigel in the presence of DRB and then stained with the indicated antibodies or the DNA dye. DMSO was used for the DRB (-) samples. B. qRT-PCR was used to examine the fold induction in mRNA levels from the indicated genes. The error bars represent mean +/- SD. C. PIP7S knockdown cells were treated with either DMSO or 300 nM flavopiridol for 5 hr. mRNA levels of the indicated genes were analyzed by qRT-PCR, with those in DMSO-treated cells artificially set at 100%. The error bars represent mean +/- SD. D. Chromatin immunoprecipitation with the indicated antibodies was performed. The regions close to the 3' end of the indicated genes were PCR-amplified from the precipitated DNA. Amplified signals from 10% of the input chromatin were also shown. The GAPDH signals were exposed 50% longer to reveal weak bands.

Fig. 2-7


A
DNA
α6 integrin
Ki67
merge
shPIP7S - DRB
shPIP7S + DRB
vector - DRB
vector + DRB
B
Fold increase in mRNA
0
2
4
6
8
10
GAPDH
PTHLH
TGM-2
C
% change in mRNA
0
20
40
60
80
100
DMSO
flavopiridol (300 nM)
GAPDH
PTHLH
TGM-2
D
IP:
ctl. IgG
αCdk9
Input
cell line:
vector
vector
shPIP7S
vector
shPIP7S
PTHLH
TGM-2
GAPDH

Chapter 3

HIV-1 Tat and host cellular AFF4 recruit two distinct transcription elongation factors into a bifunctional elongation complex for coordinated activation of HIV-1 transcription

(This work was originally published in *Mol Cell* as: He N, Liu M, Hsu J, Xue Y, Chou S, Burlingame A, Krogan NJ, Alber T, Zhou Q. HIV-1 Tat and host cellular AFF4 recruit two distinct transcription elongation factors into a bifunctional elongation complex for coordinated activation of HIV-1 transcription. *Mol Cell* 2010, 38:428-38)

Summary

Recruitment of the P-TEFb kinase by HIV-1 Tat to the viral promoter triggers the phosphorylation of RNA polymerase II C-terminal domain and escape of the polymerase from promoter-proximal pausing. It is unclear, however, if Tat recruits additional host factors that further stimulate HIV-1 transcription. Using a novel affinity-purification scheme, we have identified human transcription factor/co-activator AFF4 and elongation factor ELL2 as components of the Tat-P-TEFb complex. Through the bridging functions of Tat and AFF4, P-TEFb and ELL2 combine to form a bifunctional elongation complex that greatly activates HIV-1 transcription. Without Tat, AFF4 can mediate the ELL2-P-TEFb interaction, albeit inefficiently. Tat overcomes this limitation by bringing more ELL2 to P-TEFb, leading to the stabilization of ELL2 in a process that requires active P-TEFb. The ability of Tat to enable two different classes of elongation factors to cooperate and coordinate their actions on the same polymerase enzyme explains why Tat is such a powerful activator of HIV-1 transcription.

Introduction

Transcriptional elongation by RNA polymerase (Pol) II has recently been recognized as a major rate-limiting step for controlling the expression of many metazoan genes (1). In the absence of physiological stimuli, negative transcription elongation factors (N-TEF) pause Pol II transcribing these genes shortly after initiation. Signal-induced reversion of this block requires the positive transcription elongation factor b (P-TEFb), a heterodimer composed of CDK9 and cyclin T1 (CycT1) (or the minor forms T2 or K) (51). P-TEFb phosphorylates the C-terminal domain (CTD) of Pol II and N-TEF, leading to the production of full-length RNA transcripts (51). Besides P-TEFb, transcriptional elongation can also be stimulated by several other factors, including ELL1/2, TFIIS, TFIIF, and the Elongins, all of which enhance the processivity of Pol II through mechanisms that are quite different from that of P-TEFb (76). It is yet to be demonstrated if any of these factors work together with P-TEFb to coordinate their stimulatory effects on Pol II elongation.

Our understanding of the control of elongation, particularly by P-TEFb, has benefited tremendously from the investigations of the HIV-1 virus. Pausing of Pol II close to the transcription start site on the integrated proviral DNA template is a major rate-limiting step for HIV-1 gene expression. To antagonize this restriction, a small HIV-1 regulatory protein called Tat recruits the active form of host P-TEFb to the TAR RNA element located at the 5' end of all viral transcripts (51). This localized P-TEFb can phosphorylate the Pol II CTD and N-TEF, resulting in the production of the full-length HIV transcripts that are not only important for viral gene expression and but also serve as the genomic template for the next generation of virions.

In uninfected cells, P-TEFb functions as a general transcription factor required for the expression of a vast array of cellular genes (12, 49). As such, P-TEFb activity is subjected to stringent cellular control through the interactions with multiple factors (26). For example, more than half of nuclear P-TEFb exist in a catalytically inactive complex called 7SK snRNP that also contains the 7SK snRNA, the CDK9 kinase inhibitor HEXIM1, the LARP7/PIP7S and MePCE/BCDIN3 proteins (14-18, 59). In addition, a major fraction of P-TEFb also exists in a complex with the bromodomain protein Brd4, which recruits P-TEFb to cellular promoters through interactions with acetylated histones and the mediator complex (27, 28). Recent studies indicate that through alternately interacting with these positive and negative regulators, P-TEFb is kept in a functional equilibrium, which shifts in accordance to the cellular transcriptional demands and decisions between growth and differentiation (26, 51).

In light of the observations that P-TEFb can form distinct complexes under different conditions, we performed sequential affinity-purifications to identify factors that may interact with P-TEFb in the presence of Tat. Our experiments have identified ELL2, a previously described transcription elongation factor that stimulates Pol II elongation through suppressing transient pausing (77), and AFF4, a member of the AF4 family of transcription factors/co-activators (78), as two new proteins that exist in a single complex with Tat and P-TEFb. We found that under normal conditions, ELL2 is a short-lived protein, whose stability can be maintained through the interaction with P-TEFb in a process that is mediated by AFF4 and requires active

P-TEFb. However, the AFF4 activity is inadequate for activated HIV-1 transcription, and Tat greatly overcomes this limitation by recruiting more ELL2 to P-TEFb, resulting in the stabilization of ELL2 and synergistic activation of HIV-1 transcription by Tat and ELL2. Our data are consistent with the model that Tat and AFF4 bridge distinct elongation factors to form a bifunctional transcription elongation complex, allowing ELL2 and P-TEFb to cooperate and coordinate their actions and greatly stimulate the processivity of Pol II.

Experimental Procedures

Antibodies

The rabbit polyclonal anti-ELL2 antibodies were raised against the last 19 amino acids of ELL2 (AHIKRLIGEFDQQQAESWS; aa 622–640) and affinity-purified. The anti-AFF4 and -CycT1 antibodies were purchased from Abcam (ab57077) and Santa Cruz Biotechnology (sc-10750), respectively. The antibodies against CDK9, HEXIM1 and Brd4 have been described previously (18, 28).

Generation of the inducible Tat-HA-expressing lentiviral construct

The pcDNA4/TO-Tat-HA construct was made by inserting the Tat-HA sequence encoding the 86-aa Tat isoform derived from strain HXB2 into BamHI and EcoRI sites of pcDNA4/TO vector (invitrogen). A fragment of pcDNA4/TO-Tat-HA construct was cut out by Nru I and EcoRI and subsequently cloned into HpaI and EcoRI sites of a modified pSicoR-GFP lentiviral vector (54) containing neomycin expression cassette in order to generate the inducible Tat-HA-expressing lentiviral construct.

Generation of TTAC-8, a HEK293-based cell line stably expressing CDK9-F and inducibly expressing Tat-HA

The procedure for generating a T-RExTM-293 (Invitrogen)-based cell line that stably expresses CDK9-F and confers puromycin-resistance is essentially as described before (17). This cell line also stably expresses the tetracycline repressor protein. Next, to create a Tat-HA-inducible cell line in this background, the Tat-HA-expressing lentiviral construct described above was stably introduced through retroviral infection. The procedures for the production of recombinant lentivirus, infection of cells, and generation of puromycin- and neomycin-resistant colonies have been described before (45). Single colonies were picked and screened for the inducible expression of Tat-HA upon doxycycline treatment (1 μg/ml) for two days.

Purification of ELL2 and AFF4 as proteins associated with Tat-P-TEFb

Nuclear extracts (NEs) prepared from TTAC-8 cells stably expressing CDK9-F and inducibly expressing Tat-HA were incubated overnight with anti-Flag M2 agarose beads (Sigma). The beads were washed extensively with buffer D (20mM HEPES-KOH [Ph7.9], 15% glycerol, 0.2 mM EDTA, 0.2% NP-40, 1 mM dithiothreitol, and 1 mM phenylmethylsulfonyl fluride) containing 0.3M KCl (D0.3M), and the immobilized proteins were eluted with 0.5 μg/ml Flag peptide dissolved in D0.3M. The eluate was then subjected to a second precipitation with anti-HA agarose beads (Sigma) for 2 hours. The beads were washed extensively with D0.3M and then eluted by a low pH solution (200 mM glycine, pH 2.5). The

neutralized eluate (with 1/20 volume of 2M Tris-HCl, pH. 8.8) was subjected to analyses by SDS-PAGE followed by silver staining, and the bands were excised for identification by mass spectrometry.

In vitro binding assay

Full-length (FL) and N-terminally truncated (Δ1-300) F-AFF4 proteins as well as wild-type ELL2-F were affinity-purified from nuclear extracts (NE) of transfected HEK293T cells by anti-Flag immunoprecipitation. Prior to peptide elution, the beads were washed extensively with buffer D containing 1.0 M KCl and 0.5% NP-40. Similarly, the CycT1-HA/CDK9 complex immobilized on anti-HA agarose beads was isolated under highly stringent conditions from transfected cells. Prior to the binding assay, these proteins were checked by Western blotting to ensure that they were free of their normal binding partners found in the ELL2-AFF4-P-TEFb complex. For binding reactions, ~50 ng of ELL2-F and/or ~65 ng of either FL or Δ1-300 F-AFF4 proteins were incubated at 4oC with rotation with immobilized CycT1-HA/CDK9 purified from 50 μL NE for 90 min in a total volume of 250 μL in buffer D0.5M plus 0.2% NP-40. Upon washing extensively with D0.5M, CycT1-HA/CDK9 and its associated proteins were eluted with HA peptide (0.5 μg/ml) prepared in D0.1M. The eluates were subjected to western analysis.

Generation of AFF4 and ELL2 knockdown (KD) cells

The procedures for generating stable knockdown cells have been described previously (45). The shRNA sequences used for this purpose are:

shELL2 #8: 5’ GATCAACGCCAGAATTATAAGGATGTTCAAGAGACATCCTTATAATTCTGGC GTTTTTA 3’ and

shAFF4: 5’ GATCAAGCATCATGACAGATCTAGTTTCAAGAGAACTAGATCTGTCATGATG CTTTTTA 3’.

In addition, constructs expressing shELL2 #8 or the following two shRNA sequences are used in the luciferase reporter assay:

shELL2 #1: 5’ GATCAAATGATCCCCTCAATGAAGTTTCAAGAGAACTTCATTGAGGGGATC ATTTTTTA 3’

shELL2 #10: 5’ GATCAATAGGTGAATTTGACCAACATTCAAGAGATGTTGGTCAAATTCACCT ATTTTTA 3’

Chromatin immunoprecipitation (ChIP) assay

The ChIP assay was performed in two isogenic cell lines, one expressing Tat and the other containing an empty vector. Both cell lines also contain an integrated HIV-1 LTR-luciferase reporter gene and stably expressing ELL2-F. The procedure for performing the ChIP assay is essentially as described (28) with minor modifications. After DNA purification, PCRs containing α-[32P]dCTP (800 Ci/mmol) were performed for 24 cycles. Immunoprecipitated chromatin was analyzed first in pilot experiments to ensure that PCRs occurred in the linear range of amplification. Three regions corresponding to the promoter (-168-+80), interior (+1035-+1180), and

3'UTR (+2414-+2612) of the integrated HIV-1 LTR-luciferase gene were PCR amplified. The primers for amplifying the GAPDH gene are: GAPDH-1: 5'-ACTGCCAACGTGTCAGTGGT and GAPDH-2: 5'-CATACCAGGAAATGAGCTTGAC.

Pulse-chase analysis

Prior to the pulse-chase analysis, HEK293 cells expressing ELL2-F alone (~2x106 cells for each time point) or together with Tat-HA (~$1x10^6$ cells for each time point) were first starved in the conditioned medium lacking methionine and cysteine (Invitrogen) and supplemented with 10% dialyzed fetal calf serum (FBS). After 30 min at 37°C, the pulse medium containing EXPRES35S Protein Labeling Mix (73% methionine and 22% L-cysteine, Perkin Elmer) was added to each well at the final concentration of 0.5 mCi/mL, followed by incubation for 30 min at 37°C. Immediately thereafter, the plates were washed with the complete medium plus 10% regular FBS and chased for 1-8 hrs. Cells were directly lysed in 1X SDS loading buffer and centrifuged. The supernatants were diluted 5-fold with buffer D0.3M, pre-cleared, and subjected to anti-Flag immunoprecipitation. After extensive washing, the precipitates were eluted with 1X SDS loading buffer and subjected to analysis by SDS-PAGE followed by autoradiography.

Results

ELL2 and AFF4 associate with the Tat-P-TEFb complex

To identify novel cellular factors that control the Tat/P-TEFb-mediated activation of HIV-1 transcription, we performed affinity-purifications to isolate factors associated with the Tat-P-TEFb complex. Nuclear extracts (NEs) from an engineered human HEK293-based cell line stably expressing Flag-tagged CDK9 (CDK9-F) and inducibly expressing HA-tagged Tat (Tat-HA) were subjected to sequential co-immunoprecipitations (co-IPs) with anti-Flag and then anti-HA beads. CDK9-F was expressed at approximately the same level as the endogenous CDK9, while Tat-HA was produced under the control of a doxycycline-inducible promoter to a level similar to that of CDK9.

Inspection of the purified materials by SDS-PAGE followed by staining with silver reveals no detectable band prior to the induction of Tat-HA expression by doxycycline (Fig. 3-1A, lane 2). However, upon the induction of Tat-HA, the procedure resulted in the detection of new bands in addition to the known CycT1, Cdk9-F and Tat-HA proteins (Fig. 3-1A, lane 1), and the pattern was unchanged by pre-treating NEs with RNase A and DNase I prior to the purification (data not shown). Analyses by mass spectrometry revealed that the new bands contained proteins ELL2, AFF4, ENL and AF9.

ELL2 is related to ELL1 (49% identical and 66% similar), which was originally identified as a fusion partner of MLL (mixed lineage leukemia) in acute myeloid leukemia (79). It was subsequently re-discovered as a transcription factor that increases the catalytic rate and suppresses transient pausing of Pol II during elongation (80). ELL2 was reported to possess similar transcriptional activity as ELL1, and both are expressed in many of the same tissues (77).

AFF4 (also known as AF5q31 or MCEF) is a member of the AF4 family of transcription factors/co-activators. Like ELL1, it was also initially identified as a protein translocated to MLL in infant acute lymphoblastic leukemia (81). Using epitope-tagged CDK9 as bait, AFF4 was later found to associate with P-TEFb (78). However, ectopically expressed AFF4 failed to activate HIV-1 replication or transcription (78), for the reason that will become clear below.

ENL and AF9 are also fusion partners of MLL and involved in MLL-associated leukemia (for a review, see (82)). Although both proteins have previously been reported to interact with P-TEFb, other MLL fusion partners (e.g. AFF4 and AFF1), and the H3K79 methyltransferase DOT1L (83-85), this is the first time that they are specifically linked to Tat. Consistently, Sobhian et al. (submitted) have also identified ENL and AF9 as Tat-associated proteins (see their manuscript submitted back-to-back with ours). Despite the implication of these two proteins as critical for Tat/P-TEFb to stimulate HIV-1 transcription ((86) and our own unpublished data), exactly how they accomplish this task remains mostly unknown. Given the demonstrated interactions of ENL and AF9 with DOT1L (83-85), a thorough, full-scale investigation is needed in the future to study the mechanism by which ENL and AF9 contribute to Tat-transactivation and the possible involvement of DOT1L and H3K79 methylation in this process. For the current study, however, we decide to concentrate on the roles of ELL2 and AFF4 in modulating the Tat/P-TEFb-dependent HIV transcription.

ELL2, AFF4, Tat and P-TEFb exist in a single complex

First, the interactions of ELL2 and AFF4 with the Tat-P-TEFb complex purified through sequential co-IPs were confirmed by Western blotting using protein-specific antibodies (Fig. 3-1B). To determine whether ELL2 and AFF4 exist in the same or different complexes containing Tat-P-TEFb, we co-expressed Flag-tagged ELL2 (ELL2-F) and Tat-HA in HEK293 cells and performed sequential co-IPs with anti-Flag and then anti-HA affinity beads. Western analysis detected the associations of endogenous AFF4 and CDK9 with the affinity-purified ELL2-F-Tat-HA complex (Fig. 3-1C). The same complex was also obtained in HeLa cells expressing the tagged ELL2, CDK9 or/and Tat (data not shown). Taken together, these results indicate that ELL2, AFF4, Tat and P-TEFb likely exist in a single multi-subunit complex.

Tat-independent interactions of AFF4 and ELL2 with P-TEFb

It is important to point out that even in the absence of Tat, the anti-CDK9 immunoprecipitates derived from HEK293 NE also contained AFF4 and ELL2 as indicated by Western blotting (Fig. 3-1D). Likewise, ectopically expressed ELL2-F co-precipitated AFF4, CycT1 and CDK9 even when Tat was not present (Fig. 3-1E). Thus, the interactions of ELL2 and AFF4 with P-TEFb could occur independently of Tat. However, as will be demonstrated below, the presence of Tat can significantly enhance the interaction of ELL2 with P-TEFb.

To determine the fractions of cellular ELL2 and AFF4 that are associated with P-TEFb, we performed anti-CDK9 immunodepletion and found that the procedure removed virtually all of CDK9, ~60% of AFF4, and ~40% of ELL2 from NEs, indicating the sequestration of a major fraction of cellular ELL2 and AFF4 in the P-TEFb-containing complex.

The ELL2-AFF4-Tat-P-TEFb complex does not contain Brd4 or HEXIM1

The identification of a new P-TEFb-containing complex raised the question of whether it might be related to the two major P-TEFb complexes described in the past. Using the total CDK9-F immunoprecipitates harboring all the CDK9-associated factors as a reference, we performed Western analysis to determine whether the affinity-purified ELL2-AFF4-Tat-P-TEFb complex might contain Brd4, the P-TEFb recruitment factor for general transcription (27, 28), or HEXIM1, a signature subunit and inhibitor of CDK9 within the 7SK snRNP (18, 20). When the levels of CycT1, a common subunit shared among all the known P-TEFb-containing complexes, were normalized to about the same level, neither Brd4 nor HEXIM1 was detected in the anti-ELL2-F immune-complex even after a prolonged exposure (Fig. 3-1F). In contrast, abundant Brd4 and HEXIM1 were readily detected in the total CDK9-F immunoprecipitates. Thus, the ELL2-AFF4-Tat-P-TEFb complex (and also the Tat-free ELL2-AFF4-P-TEFb complex, data not shown) is physically distinct from 7SK snRNP or the Brd4-P-TEFb complex.

ELL2 depletion reduces both basal and Tat-activated HIV-1 transcription

To understand the functional significance of the associations of ELL2 and AFF4 with the Tat-P-TEFb complex in HIV-1 transcription, we employed short hairpin (sh)RNAs to reduce the expression of ELL2 and examined the impact on basal as well as Tat-dependent HIV-1 transcription. Compared to an ineffective control shRNA (shELL2 #10), the expression of either of the two effective shRNAs (shELL2 #1 and #8; with the latter causing ~75% ELL2 depletion) caused a significant reduction in both basal and Tat-dependent HIV-1 LTR-driven luciferase activity from the transfected reporter construct (Fig. 3-2A). Notably, the reduction in viral LTR activity mediated by shELL2 #1 and #8 correlated well with the reduction in the level of ELL2 expression (compare Fig. 3-2A and 2B). This general dependence on ELL2 for HIV-1 transcription, which was also observed with a stably integrated HIV-1 LTR-luciferase construct (data not shown), is highly reminiscent of the situation involving P-TEFb, whose loss of function has also been shown to reduce both basal and Tat-activated HIV-1 transcription (8, 68, 87).

Ectopically expressed ELL2, but not ELL1, synergizes with Tat to activate HIV-1 transcription in a TAR-dependent manner

Consistent with the observation that ELL2 is important for HIV-1 transcription, ectopically expressed ELL2 increased basal HIV-1 LTR activity from a reporter construct that contains either wild-type (wt) or a mutant (Δ) TAR element by 1.3- and 2.0-fold, respectively (Fig. 3-2C). When expressed alone, Tat increased HIV-1 transcription 82-fold in a strictly TAR-dependent manner. Remarkably, the combination of ELL2 and Tat increased HIV-1 transcription 284-fold from the wild-type TAR construct (Fig. 3-2C), demonstrating a strong synergistic effect of the two on the HIV-1 LTR. However, when the ΔTAR reporter construct was used, the synergism was lost and only a small positive effect similar to that caused by ELL2 alone remained (Fig. 3-2C). As shown in Supplemental Figure, ectopically expressed ELL2-F activated both basal and Tat-dependent HIV-1 transcription in a dose-dependent manner. The Tat/ELL2 synergism was also observed under conditions

of elevated ELL2-F expression and in cells containing a stably integrated HIV-1 LTR-luciferase reporter gene (data not shown).

Because the prototypical ELL1 transcription elongation factor is 66% similar and 49% identical to ELL2, we compared their abilities to synergize with Tat in activating HIV-1 transcription. Data in Fig. 3-2D indicate that ELL2, but not the homologous ELL1, displayed a significant synergism with Tat.

Tat increases the amount of ELL2 but not ELL1 associated with P-TEFb

To define the mechanism of this selectivity of Tat for ELL2, we examined the interactions of ELL1 and ELL2 with P-TEFb either in the absence or presence of Tat. Western analysis indicates that the expression of even a small amount of Tat significantly elevated the level of the co-expressed ELL2-F but not ELL1-F protein in NE (Fig. 3-2E). Correlating with more ELL2-F in the nucleus, there was a drastic increase in the amount of CDK9 bound to the immunoprecipitated ELL2-F but not ELL1-F (Fig. 3-2F), indicating a Tat-induced elevation of the level of the ELL2/P-TEFb-conaining complex. Importantly, Tat stimulated the association with P-TEFb by both the ectopically expressed ELL2-F and endogenous ELL2 protein (see below).

AFF4 synergizes with ELL2 to activate transcription from viral and cellular promoters

Because the Tat-P-TEFb complex also contains AFF4, we investigated the contribution of this factor to transcription from the HIV-1 LTR as well as several other viral and cellular promoters. Remarkably, just like the combination of Tat and ELL2, the co-expression of AFF4 and ELL2 also synergistically activated HIV-1 transcription (Fig. 3-3A). However, in contrast to the Tat/ELL2-induced synergistic activation of the HIV-1 LTR, which depended on TAR (Fig. 3-2C), the synergism displayed by the AFF4/ELL2 pair was not restricted to the HIV-1 promoter. Rather, the functional interaction was also detected using several other viral and cellular promoter constructs, although the effects were not as pronounced as that of the HIV-1 LTR (Fig. 3-3A). Unlike Tat, which is a sequence-specific transcription factor, neither AFF4 nor ELL2 is known to display any sequence-specific DNA or RNA-binding activity. This difference can explain the observations that the transcriptional synergism displayed by the Tat/ELL2 pair was HIV-1 TAR-dependent, whereas the synergism by AFF4/ELL2 was not.

AFF4 synergizes with ELL2 to activate basal but not Tat-dependent HIV-1 transcription

Our subsequent analyses employing the HIV-1 LTR-luciferase reporter assay indicate that although AFF4 and ELL2 synergistically activated basal HIV-1 transcription, this synergism was completely lost when Tat was present (Fig. 3-3B). Similarly, while ectopically expressed AFF4 was able to stimulate basal HIV-1 transcription by up to 17-fold, it enhanced the Tat-dependent transcription by only 2-fold (Fig. 3-3C).

Two questions are raised by the data in Fig. 3A-C: (1) what is the molecular basis of the synergism displayed by AFF4 and ELL2? and (2) why did the ectopically expressed AFF4 have very different effects on basal and Tat-activated HIV

transcription? To address the first question, we examined the levels of AFF4 and ELL2 that were expressed either alone or together in both nuclear (NE) and whole cell extracts (WCE). Just like the stable accumulation of ELL2 induced by Tat (Fig. 3-2E), the co-expression of F-AFF4 and ELL2-HA also significantly increased the level of ELL2-HA in both NE and WCE (Fig. 3-3D, top panel). As will be demonstrated below, this elevated ELL2 level was likely caused by the AFF4-induced stabilization of ELL2. Correlating with the higher ELL2 concentration, there was a marked increase in the amount of CDK9 bound to the immunoprecipitated ELL2-HA (Fig. 3-3D, bottom panel).

In addition to the significantly increase in the levels of ELL2 and ELL2/P-TEFb-containing complex, the co-expression of F-AFF4 and ELL2-HA also modestly increased the F-AFF4 concentrations (~2-fold) in both NE and WCE (Fig. 3-3D, top panel). Given that these two proteins synergistically activated transcription from a variety of cellular and viral promoters (Fig. 3-3A), it is possible that the elevated AFF4 level was due to this general increase in transcription from also the AFF4-expressing construct.

Tat and ectopically expressed AFF4 increase the levels of ELL2 and ELL2/P-TEFb-containing complex by a common mechanism

Why didn't the ectopically expressed AFF4 further promote Tat-transactivation? Since both proteins elevated the ELL2 level, we tested whether they exerted additive effects. While the expression of ELL2-HA with F-AFF4 or Tat alone increased the nuclear level of ELL2-HA as well as the amounts of CDK9 and CycT1 bound to ELL2-HA (Fig. 3-3E, compare lane 2 with lanes 3 & 4), simultaneous co-expression of ELL2-HA, F-AFF4 and Tat did not further enhance these effects (Fig. 3-3E, lane 5). This absence of additive effects by Tat and AFF4 suggests that these two proteins used a common mechanism to promote ELL2 accumulation and interaction with P-TEFb and that their effects were likely saturated under the current conditions. The fact that overexpressed AFF4 failed to further enhance Tat-transactivation also agrees with the previous observation that AFF4 overexpression cannot activate HIV-1 replication (78), which is Tat-dependent.

AFF4 synergizes more efficiently with ELL2 than with ELL1 to activate transcriptional elongation in vitro and in vivo

In addition to increasing the level of the ELL2/P-TEFb-containing complex in the cell, AFF4 and Tat also displayed a similar ability to distinguish between ELL1 and ELL2. Like Tat, AFF4 showed much stronger transcriptional synergism in combination with ELL2 than with ELL1 (Fig. 3-3F). This functional interaction is likely due to the ability of AFF4 to markedly increase the level of ELL2 but not ELL1 (Fig. 3-3G), which has been shown above to correlate with more ELL2 associated with P-TEFb (Fig. 3-3D & 3E).

It is worth noting that the stimulatory effect of AFF4/ELL2 has so far been demonstrated in only the luciferase reporter assay. To confirm that the stimulation indeed occurs at the transcriptional level, we first performed an in vitro transcription assay that employs the well-characterized HIV-1 LTR-G400 template for detecting the LTR-directed transcriptional elongation of a 400-bp G-less cassette (G400) inserted at

~1-kb downstream of the start site (88). The reactions also contained HeLa NE, in which P-TEFb and its associated factors were immuno-depleted with anti-Cdk9 antibodies. While the addition of isolated P-TEFb heterodimer or an equal molar mixture of AFF4/ELL2 into the depleted NE only weakly rescued transcription, the combination of P-TEFb plus AFF4/ELL2 dramatically enhanced the promoter-distal transcription of the G-less cassette (Fig. 3-4H). This result confirms that the positive effect of AFF4/ELL2 is indeed exerted at the level transcriptional elongation.

Finally, the AFF4/ELL2 stimulation of promoter-distal (i.e. elongation) but not -proximal transcription (i.e. initiation) was also confirmed by RT-PCR analysis of mRNA transcribed from the HIV LTR-luciferase reporter gene at two different locations. While the AFF4/ELL2 co-expression did not affect the abundance of transcripts mapping to the 5' end of the mRNA (HIV sequence +1 to +80), it did increase the level of transcripts corresponding to the 3' end (+1697 to +1804) of the luciferase gene (Fig. 3-4I).

AFF4 knockdown causes stronger inhibition of basal than Tat-dependent HIV-1 transcription and decreases the levels of ELL2 and ELL2/P-TEFb-containing complex in vivo

Further support for the idea that AFF4 and Tat act similarly to increase the level of the ELL2/P-TEFb-containing complex in vivo came from the effects of reducing AFF4 expression using an AFF4-specific shRNA (shAFF4). Although the reduction had a relatively mild effect (2-fold) on Tat-activated HIV-1 transcription, it affected Tat-independent transcription more severely (5-fold; Fig. 3-4A). Because of their similar function and mechanism of action, it is likely that the reduction in AFF4 in the knockdown (KD) cells was largely compensated by the continued presence of Tat. In contrast, the reduced level of the ELL2/P-TEFb-containing complex in the absence of both Tat and AFF4 resulted in a significant decrease in HIV-1 transcription.

Consistent with our finding that AFF4 promotes the stable accumulation of ELL2 in vivo, shRNA depletion of AFF4 was found to cause the co-depletion of ELL2 in HEK293 cells (Fig. 3-4B, left panel). Importantly, there was a concomitant decrease in the amounts of ELL2 bound to the immunoprecipitated CDK9 (Fig. 3-4B, right panel), indicating a reduction in the amount of the ELL2/P-TEFb-containing complex in KD cells. Similar results were also obtained in HeLa cells expressing shAFF4 (data not shown).

AFF4 mediates the ELL2-P-TEFb interaction

The reduced level of the ELL2/P-TEFb-containing complex upon AFF4 depletion could be a consequence of the reduced ELL2 protein level in AFF4 KD cells. Alternatively, it could be caused directly by the loss of AFF4 as a mediator of the ELL2-P-TEFb interaction, and the reduced ELL2 level could be due to a decrease in ELL2 stability when the protein is no longer in complex with P-TEFb and/or AFF4. To determine whether AFF4 could mediate the ELL2-P-TEFb interaction, we performed in vitro binding assays to test the interaction of ELL2-F with the immobilized CycT1-HA/Cdk9 heterodimer in the presence of either the full-length or a N-terminally truncated (Δ1-300) F-AFF4 protein. All four proteins, highly purified from transfected cells under stringent conditions involving high salt (1.0 M KCl) and

detergent (0.5% NP-40; see Materials & Methods), were confirmed by Western blotting to be free of any associated factors that are present in the ELL2/AFF4/P-TEFb-containing complex (data not shown). The Δ1-300 F-AFF4 was chosen as a control because it completely failed to interact with P-TEFb, but still maintained a nearly wild-type association with ELL2 in vivo (Fig. 3-4C), suggesting the existence of separable ELL2- and P-TEFb-binding domains in AFF4.

In the in vitro binding reactions (Fig. 3-4D), purified ELL2-F bound to the immobilized CycT1-HA/CDK9 in the presence of the full-length but not Δ1-300 F-AFF4 (left panel). In the absence of ELL2-F, the full-length but not the truncated F-AFF4 interacted with CycT1-HA/CDK9 directly (Fig. 3-4D, middle panel). These data suggest that AFF4 can bind to P-TEFb directly and that AFF4 bridges ELL2 and P-TEFb together. This bridging function explains the significantly decreased ELL2-P-TEFb interaction observed in AFF4 KD cells (Fig. 3-4B).

Tat further increases the amount of ELL2 bound to P-TEFb without affecting the AFF4-P-TEFb binding

Given that AFF4 mediated the ELL2-P-TEFb interaction in the absence of Tat, we examined whether Tat could affect the interaction of ELL2 with P-TEFb. When anti-CDK9 immunoprecipitates containing all endogenous proteins were isolated from cells either expressing or not expressing Tat-HA from a retroviral vector and analyzed by Western blotting, the interaction of Tat with P-TEFb was found to markedly increase the association of ELL2 with CDK9/CycT1 (Fig. 3-5A, top left panel). In contrast, although a small amount of ELL1 was found to interact with CDK9/CycT1 under Tat(-) conditions, the presence of Tat did not further enhance this interaction. Finally, just like the situation involving ectopically expressed ELL2-F and ELL1-F (Fig. 3-2), the Tat-enhanced formation of the ELL2/P-TEFb-containing complex also correlated well with the accumulation of more endogenous ELL2 but not ELL1 in the cell (Fig. 3-5A, bottom panels).

While Tat brought more ELL2 to P-TEFb, it did not significantly affect the interaction of AFF4 with P-TEFb (Fig. 3-5A, top left panel). Similarly, the AFF4-P-TEFb interaction also remained unchanged upon shRNA depletion of ELL2 (Fig. 3-5B). These data, together with our observation of a direct binding between AFF4 and P-TEFb in vitro (Fig. 3-4D, middle panel), reveal the autonomous nature of the AFF4-P-TEFb interaction, which occurs independently of Tat and ELL2 within the two P-TEFb-containing complexes.

The fact that Tat was required to recruit more ELL2 to P-TEFb explains why the previous attempts by others and us (15-17, 45, 78) to isolate P-TEFb-associated factors under Tat(-) conditions failed to identify ELL2. More importantly, it indicates that although AFF4 can function as a mediator between ELL2 and P-TEFb, it is insufficient for activated HIV-1 transcription, and Tat overcomes this limitation. One possible reason that keeps AFF4 from achieving its full capacity may be its relatively low concentration or/and efficiency in vivo. This view is supported by our demonstration that overexpression of AFF4 significantly increased the amount of ELL2 bound to P-TEFb (Fig. 3-3).

Further evidence indicating a functional consequence of the Tat-induced

increase in the ELL2/AFF4/P-TEFb-containing complex comes from the chromatin immunoprecipitation experiments employing a stably integrated HIV-1 LTR-driven luciferase reporter gene. The expression of Tat significantly increased the occupancy of ELL2-F, CDK9 and AFF4 at the HIV-1 promoter, the interior of the luciferase gene, and the 3' UTR region (Fig. 3-5C). In contrast, there was no Tat-mediated increase in bindings of these proteins to the control GAPDH locus (Fig. 3-5D), whose expression is Tat-independent. These results are consistent with the roles of ELL2 and P-TEFb as two critical elongation factors contributing simultaneously to Tat activation of HIV-1 transcription.

Physical association with active P-TEFb is required for Tat to promote ELL2 accumulation and formation of the ELL2/P-TEFb-containing complex

The data presented so far have revealed an important role for Tat and AFF4 in mediating the ELL2-P-TEFb interaction and elevating nuclear ELL2 level, which in turn allow Tat and AFF4 to synergize with ELL2 to activate HIV-1 transcription. Because P-TEFb is an integral component of the ELL2/AFF4-containing complexes and the enhanced ELL2-P-TEFb interaction correlates with increased ELL2 nuclear concentration, we determined whether the CDK9 kinase activity is required for complex formation and stable accumulation of ELL2.

When we treated cells with the CDK9 inhibitors DRB or flavopiridol (FVP), the Tat-induced nuclear accumulation of ELL2 was almost completely blocked (Fig. 3-6A). Because wild-type CDK9 kinase activity was apparently critical in this process, we asked whether the physical association with P-TEFb is also required. To address this question, the Tat mutant C22G, which is unable to bind P-TEFb due to the disruption of a key cysteine-zinc bridge (9), was co-expressed with ELL2-F. Compared to wild-type Tat, C22G could neither promote the accumulation of ELL2-F nor enhance the interactions of CDK9 and CycT1 with the immunoprecipitated ELL2-F (Fig. 3-6B). Together, these data suggest that not only the CDK9 kinase activity but also the physical association with active P-TEFb is required for Tat to promote the accumulation of ELL2 in the nucleus. These results suggest that the elevated ELL2 level is a direct result of ELL2 becoming sequestered in the Tat-P-TEFb complex. In other words, the enhanced ELL2-P-TEFb interaction is the cause rather than the consequence of the elevated ELL2 level in NE (see Discussion below).

The kinase-inactive CDK9 mutant blocks the Tat-induced nuclear accumulation of ELL2 and is defective in interacting with ELL2

To further characterize the dependence on active P-TEFb for Tat-induced nuclear accumulation of ELL2 and the ELL2-P-TEFb interaction, we employed in binding studies a CDK9 mutant, D167N, which is kinase-inactive and dominant-negative over wild-type CDK9 (89, 90). The expression of D167N CDK9 drastically reduced the ability of Tat to increase the nuclear concentration of ELL2-HA (Fig. 3-6C). Furthermore, D167N CDK9 also interacted with a significantly decreased amount of ELL2 under both Tat(+) and Tat(-) conditions (compare lanes 4 & 5 with lanes 2 & 3 in Fig. 3- 6C).

Not only was the formation of the Tat-containing ELL2-Tat-AFF4-P-TEFb

complex promoted by active P-TEFb, the stability of the Tat-free ELL2-AFF4-P-TEFb complex also depended on wild-type CDK9. Upon the inhibition of CDK9 kinase by DRB or FVP, the association of ELL2 with the immunoprecipitated CDK9-F/CycT1 heterodimer of P-TEFb was markedly reduced (Fig. 3-6D). In contrast, the AFF4-P-TEFb binding was only slightly affected, providing another indication of the stable and autonomous nature of this interaction within the ELL2-AFF4-P-TEFb complex. Taken together, these data illustrate a strong dependence on active P-TEFb for the ELL2-P-TEFb interaction, irrespective of whether the interaction is mediated by AFF4 alone or further promoted by HIV-1 Tat.

ELL2 is a short-lived protein whose stability can be significantly enhanced by the inhibition of the proteosome or the expression of Tat or AFF4

Our data so far reveal a critical role for Tat as well as both the ectopically expressed and endogenous AFF4 in promoting/maintaining the stable accumulation of ELL2 in vivo (Fig. 3-3D & 4B). These activities of Tat and AFF4 are likely responsible for their synergistic activation of HIV-1 transcription together with ELL2. To determine the mechanism of ELL2 accumulation, we treated cells with the proteosome inhibitor MG132 and found that both the ectopically expressed ELL2-F and its endogenous counterpart are short-lived proteins, whose half-lives were significantly prolonged by the drug (Fig. 3-7A & 7B).

A pulse-chase experiment employing 35S-labeled methionine and L-cysteine (Perkin Elmer) was performed to determine whether Tat could affect the stability of the co-expressed ELL2-F. In the absence of Tat, most of ELL2-F had a half-life of less than 1 hr, although a minor species with slightly slower mobility was considerably more stable (Fig. 3-7C). Interestingly, the presence of Tat significantly increased the abundance of this stable slow-migrating ELL2-F protein, and as a result, extended the half-live of the overall ELL2-F population to more than 4 hrs (Fig. 3-7C). This Tat-induced mobility shift of ELL2 was also evident in other experiments involving the co-expression of the two proteins (e.g. see 3-Fig. 6C, lanes 2 & 3). Treatment of ELL2-F purified from Tat-expressing cells with calf intestine phosphatase (CIP) caused the upper ELL2-F band to collapse into a single faster-migrating species, indicating that the upper band contains the phosphorylated form of ELL2 (Fig. 3-7E). Based on the demonstrations that active P-TEFb is essential for Tat-induced enhancement of the ELL2-P-TEFb interaction and accumulation of ELL2 (Fig. 3-6A-C), it is possible that CDK9 plays a role in phosphorylating ELL2 in the ELL2/Tat/AFF4/P-TEFb-containing complex.

Given that ectopically expressed AFF4 also promoted the accumulation of ELL2, we performed another pulse-chase experiment to determine whether AFF4 could stabilize ELL2. Indeed, expression of F-AFF4 increased the half-life of the overall ELL2-F population from less than 1 hr to more than 4 hrs (Fig. 3-7D). Like Tat, F-AFF4 also caused a small reduction in the mobility of ELL2-F in the SDS-gel, although the effect was less prominent than that caused by Tat. Taken together, these results illustrate that Tat and AFF4 similarly regulate the stability and transcriptional and P-TEFb-binding activities of ELL2.

Discussion

P-TEFb was first identified more than a decade ago as a specific host cellular cofactor for Tat activation of HIV-1 transcription (8, 9). Since then, this landmark discovery has provided the basic framework for our understanding of Tat function in the HIV-1 life cycle, and P-TEFb has gained universal acceptance as a functional Tat partner. Here, we expand the conventional view of the control of HIV transactivation by showing that the Tat-P-TEFb complex in fact contains additional components, ELL2, AFF4 and possibly also ENL and/or AF9 ((86) and our own unpublished data), although the precise roles of the latter two in this complex remains mostly unknown. The identification of these factors in complex with Tat-P-TEFb relied on a new method of double epitope-tagging followed by sequential immunoprecipitations. This method allows the selective isolation of factors that are directly involved in the Tat/P-TEFb-mediated HIV-1 transactivation, and can be adapted readily for future investigations of protein complexes containing at least two known subunits. We demonstrate that AFF4 and ELL2 cooperate with Tat and P-TEFb to stimulate HIV-1 transcription. ELL2 and AFF4 join P-TEFb as Tat partners essential for Tat-transactivation.

Like P-TEFb, ELL2 is also a potent transcription elongation factor. How can it be that the contribution to Tat-transactivation by this key factor, which is an integral part of the Tat-P-TEFb complex, has gone undetected in the past? We believe the answer lies in the fact that both the ability to associate with the Tat-P-TEFb complex and stability of ELL2 depend on wild-type CDK9 kinase activity (Fig. 3-6). Therefore, the contributions to Tat-transactivation by these two elongation factors are intertwined, and any molecular or pharmacological manipulations that reduce P-TEFb expression or activity also negatively affect ELL2, thus erasing their contributions simultaneously. In light of this phenomenon, our traditional view of the role of P-TEFb in Tat-transactivation is in need of a major revision. Namely, any stimulatory effects on basal and Tat-activated HIV-1 transcription that have previously been assigned to P-TEFb alone should in fact be attributed to both P-TEFb and ELL2.

P-TEFb stimulates the processivity of Pol II by phosphorylating the Pol II CTD and negative elongation factors DSIF and NELF (10, 76). These modifications convert Pol II into an elongation-competent form and antagonize the inhibitory effects of NELF and DSIF (10, 76). Employing a mechanism different from that of P-TEFb, ELL2 promotes elongation by keeping the 3' OH of nascent mRNA in alignment with the catalytic site, thus preventing Pol II backtracking (77, 80). Although both factors are considered general transcription factors, it is unclear whether they work independently or cooperatively during Pol II elongation of any DNA template. The identification of a Tat/AFF4-mediated interaction between ELL2 and P-TEFb provides for the first time strong evidence in support of a coordinated promotion of Pol II elongation on HIV-1 proviral DNA by different classes of elongation factors. By forming a bifunctional transcription elongation complex, ELL2 and P-TEFb can act on the same polymerase enzyme and at the same time. This coordination and cooperation between the two factors enhances the efficiency of Pol II elongation. The ability of Tat to promote AFF4 recruitment of two different elongation factors that

work by distinct mechanisms into a single complex explains why Tat is such a powerful activator of HIV-1 transcription.

It is interesting to note that a functional connection between ELL2 and P-TEFb has been implicated in previous studies. For example, ELL2 is required for the production of the secretory-specific form of IgH mRNA in plasma cells through promoting exon skipping and the use of a proximal poly(A) site (91). Factors that enhance the Pol II elongation rate can often facilitate exon skipping (92), presumably because the accelerated rate of elongation reduces the time in which splice sites are offered to the splicing apparatus. Interestingly, the ELL2-promoted processing of IgH mRNA correlates with a significant increase in Ser2 phosphorylation of the Pol II CTD as well as the occupancy of ELL2 close to the IgH promoter in plasma cells (93). Prior to our current study, it would have been difficult to explain how ELL2, an elongation factor that acts by directly modulating the catalytic rate of Pol II, can induce CTD phosphorylation on Ser2, which is thought to be performed by P-TEFb (76). The ELL2/AFF4/P-TEFb-containing complex described here provides a potential mechanism to link the phosphorylation of Ser2 to the ELL2-associated P-TEFb. An additional functional coupling between ELL2 and P-TEFb has been detected in Drosophila melanogaster. Upon RNAi-mediated silencing of CDK9 expression in larvae, a significantly reduced amount of Drosophila ELL, which is most similar to human ELL2 (90), was found on the chromosomes (94).

One common theme of the experiments presented in this study is that an increase in the ELL2-P-TEFb interaction, irrespective of whether it was mediated by AFF4 or further enhanced by Tat, is associated with enhanced ELL2 stability. The accumulation of stable ELL2 could be a cause or consequence of an elevated level of the ELL2/P-TEFb-containing complex. Alternatively, the two events could have no causal relationship at all. The observations that the enhanced ELL2 stability depended not only on the kinase activity of CDK9 but also on the physical association with P-TEFb indicate that the accumulation of stable ELL2 in vivo is very likely a direct consequence of more ELL2 sequestered in the P-TEFb-containing complexes. ELL2 is protected against proteolysis, because it is sequestered within these complexes. Given the observation that the stabilization correlated with a mobility shift of ELL2 (Fig. 3-6C & 7C&D), it will be interesting to determine whether the phosphorylation of ELL2 by P-TEFb may be involved.

In addition to the Tat-containing and Tat-free ELL2/AFF4/P-TEFb complexes described here, nuclear P-TEFb also occurs in the 7SK snRNP and the Brd4-containing complex (26). The 7SK snRNP, which harbors inactive P-TEFb, and the Brd4-P-TEFb complex have been shown to contain P-TEFb molecules that are in storage or being transported to a chromatin template, respectively (26). The notion that Brd4 only recruits P-TEFb to a promoter region but not directly participate in P-TEFb-mediated transcriptional elongation is based on the observation that the distribution patterns of Brd4 and P-TEFb are quite different throughout the transcription unit especially outside the promoter-proximal region (27). Furthermore, the Brd4-bound P-TEFb is also incompatible with Tat-activated HIV-1 transcription since Tat directly competes with Brd4 for binding to P-TEFb and the overexpression

of Brd4 interferes with Tat-transactivation (28, 31). These observations, together with the results obtained in this study, make the ELL2/AFF4/P-TEFb- and ELL2/Tat/AFF4/P-TEFb-containing complexes likely candidates for the forms of P-TEFb involved in general and Tat-dependent HIV-1 transcription, respectively.

Fig. 3-1 P-TEFb exists in two novel complexes containing ELL2/AFF4/Tat/P-TEFb and ELL2/AFF4/P-TEFb, respectively. A. CDK9-F, Tat-HA and their associated factors (lane 1) were isolated through sequential immunoprecipitations (IP; anti-Flag and then anti-HA) from NE of TTAC-8 cells upon the induction of Tat-HA expression and analyzed on a silver-stained SDS-gel, with their identities indicated on the left. NE derived from TTAC-8 cells prior to the induction of Tat-HA expression was used in a parallel procedure for control (lane 2). The molecular weight (MW) markers were in lane 3. B. The Flag & HA immunoprecipitates analyzed in A were examined by Western blotting for the indicated proteins. C. The αFlag & αHA sequential immunoprecipitates derived from NE of HEK293 cells expressing the indicated proteins were analyzed by Western blotting as in B. D. Immunoprecipitates obtained with the indicated antibodies were examined as in B. E. The αFlag immunoprecipitate derived from NE of ELL2-F-expressing cells were analyzed as in B. F. The parental HEK293 cells and the HEK293-based cell lines expressing the indicated proteins were subjected to anti-Flag immunoprecipitation. The immunoprecipitates were analyzed by Western blotting for the presence of the indicated proteins.

Fig. 3-1

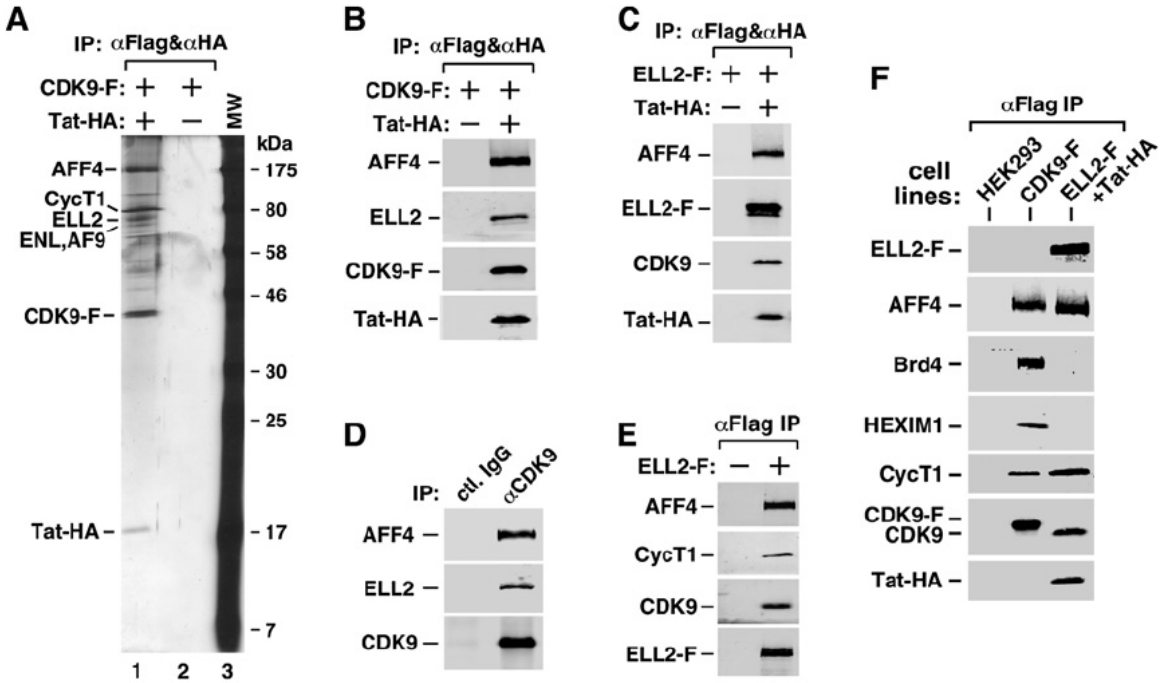

A
IP: αFlag&αHA
CDK9-F: + +
Tat-HA: + –
MW
kDa
AFF4
CycT1
ELL2
ENL,AF9
CDK9-F
Tat-HA
175
80
58
46
30
25
17
7
1 2 3
B
IP: αFlag&αHA
CDK9-F: + +
Tat-HA: – +
AFF4
ELL2
CDK9-F
Tat-HA
C
IP: αFlag&αHA
ELL2-F: + +
Tat-HA: – +
AFF4
ELL2-F
CDK9
Tat-HA
D
IP: ctl. IgG αCDK9
AFF4
ELL2
CDK9
E
αFlag IP
ELL2-F: – +
AFF4
CycT1
CDK9
ELL2-F
F
αFlag IP
cell lines: HEK293 CDK9-F ELL2-F +Tat-HA
ELL2-F
AFF4
Brd4
HEXIM1
CycT1
CDK9-F
CDK9
Tat-HA

Fig. 3-2 Tat increases the levels of ELL2 and the ELL2/P-TEFb-containing complex and synergizes with ELL2 but not ELL1 to activate HIV-1 transcription in a TAR-dependent manner. A. Luciferase activities were measured in extracts of cells co-transfected with the indicated shELL2-expressing constructs, the HIV-1 LTR-luciferase reporter gene, and a vector expressing Tat-HA (+) or nothing (-). The activity in cells expressing shELL2 #10 but no Tat was artificially set to 1. The error bars represent mean +/- SD. B. Western analyses of the levels of ELL2-F and α-Tubulin in cells transfected with the indicated shELL2-expressing constructs. C. Luciferase activities were measured and analyzed as in A in extracts of cells transfected with either the wild-type HIV-1 LTR-luciferase construct or a ΔTAR mutant construct, together with the indicated plasmids expressing ELL2-F (0.5 ug/well) and/or Tat-HA (0.01 ug/well). D. Luciferase activities were measured and analyzed as in A in cells transfected with the HIV-1 LTR-luciferase construct and the indicated ELL1-F, ELL2-F and/or Tat-HA-expressing plasmids. E. Western analysis of the indicated proteins in nuclear extracts (NE) of cells co-transfected with the indicated cDNA constructs or an empty vector (vec. or '-'). The Tat-HA construct was transfected in 2-fold increments. F. ELL1-F, ELL2-F and their associated CDK9 were isolated by anti-Flag immunoprecipitation from NE analyzed in E and examined by Western blotting.

Fig. 3-2

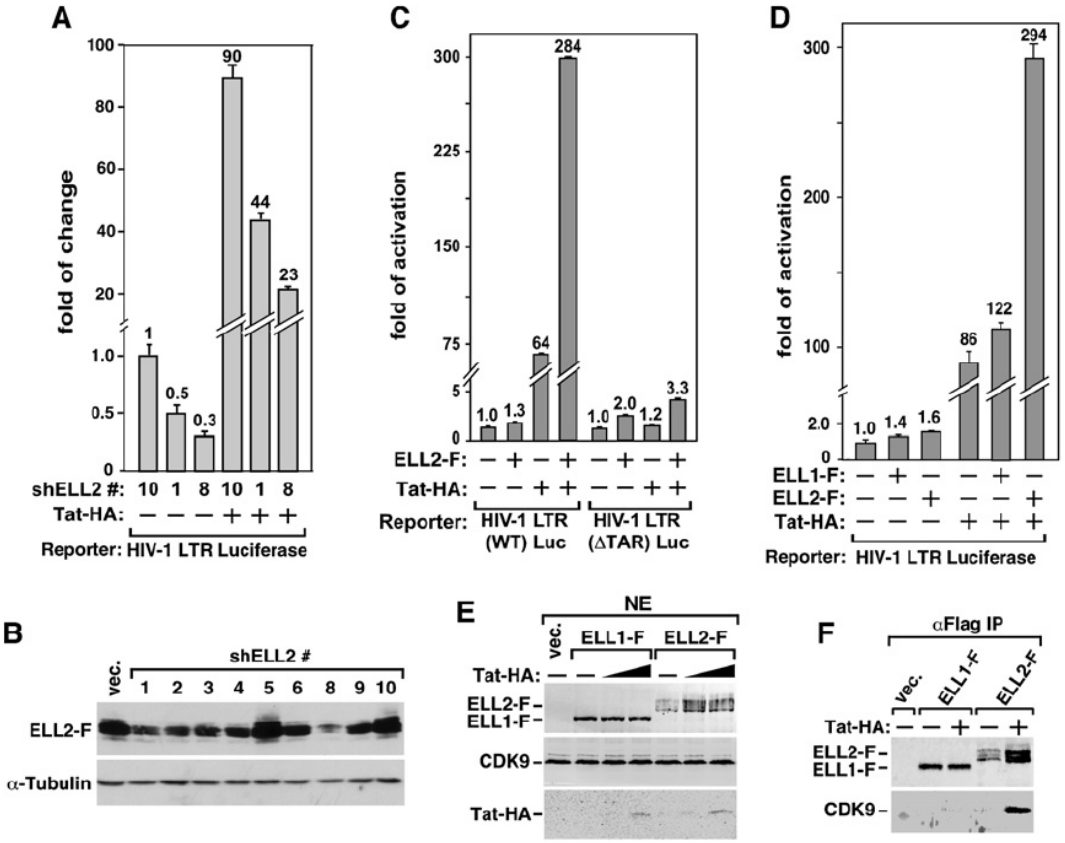

A
fold of change
100
80
60
40
20
1.0
0.5
0
1
0.5
0.3
90
44
23
shELL2 #: 10 1 8 10 1 8
Tat-HA: – – – + + +
Reporter: HIV-1 LTR Luciferase
C
fold of activation
300
225
150
75
5
0
284
64
1.0 1.3
1.0 2.0 1.2 3.3
ELL2-F: – + – + – + – +
Tat-HA: – – + + – – + +
Reporter: HIV-1 LTR (WT) Luc
HIV-1 LTR (ΔTAR) Luc
D
fold of activation
300
200
100
2.0
0
294
122
86
1.0 1.4 1.6
ELL1-F: – + – – + –
ELL2-F: – – + – – +
Tat-HA: – – – + + +
Reporter: HIV-1 LTR Luciferase
B
shELL2 #
vec. 1 2 3 4 5 6 8 9 10
ELL2-F
α-Tubulin
E
NE
vec. ELL1-F ELL2-F
Tat-HA: – – – –
ELL2-F
ELL1-F
CDK9
Tat-HA
F
αFlag IP
vec. ELL1-F ELL2-F
Tat-HA: – – + – +
ELL2-F
ELL1-F
CDK9

Fig.3-3 Ectopically expressed AFF4 behaves like Tat to synergize with ELL2 to stimulate transcriptional elongation through promoting ELL2 accumulation and association with P-TEFb. A. Luciferase activities were measured in extracts of cells transfected with the indicated promoter-luciferase reporter constructs together with the ELL2-F- or/and F-AFF4-expressing plasmids as indicated. For each promoter construct, the level of activity detected in the absence of any ELL2-F or F-AFF4 was set to 1 for easy comparison. The error bars represent mean +/- SD. B. Luciferase activities were measured and analyzed as in A in cells transfected with the indicated reporter and cDNA expression constructs. The activity detected in the absence of any transfected protein was set to 1. C. The luciferase reporter assay was performed as in A in cells transfected with the indicated reporter and cDNA expression constructs. D. & E. Nuclear extracts (NE) from cells transfected with the indicated cDNA constructs (top panel) and anti-HA immunoprecipitates (IP) derived from NE (bottom panel) were analyzed by Western blotting (WB) for the presence of the indicated proteins. F. The luciferase reporter assay was performed as in A in cells transfected with the indicated reporter and cDNA expression constructs. G. Whole cell extracts from the same cells as analyzed in F were examined by Western blotting for the expression of the indicated proteins. H. Transcription reactions containing the CDK9-depleted NE, DNA template HIV-1 LTR-G400, and the indicated purified proteins were performed. The 400-bp RNA fragment transcribed from a G-less cassette located at ~1-kb downstream of the HIV-1 promoter is indicated. I. mRNAs transcribed from the HIV-1 LTR-luciferase reporter gene and purified from cells transfected with either an empty vector (-) or the F-AFF4/ELL2-F-expressing constructs were subjected to RT-PCR analysis with primers that amplify the two indicated regions. RT: reverse transcriptase.

Fig. 3-3

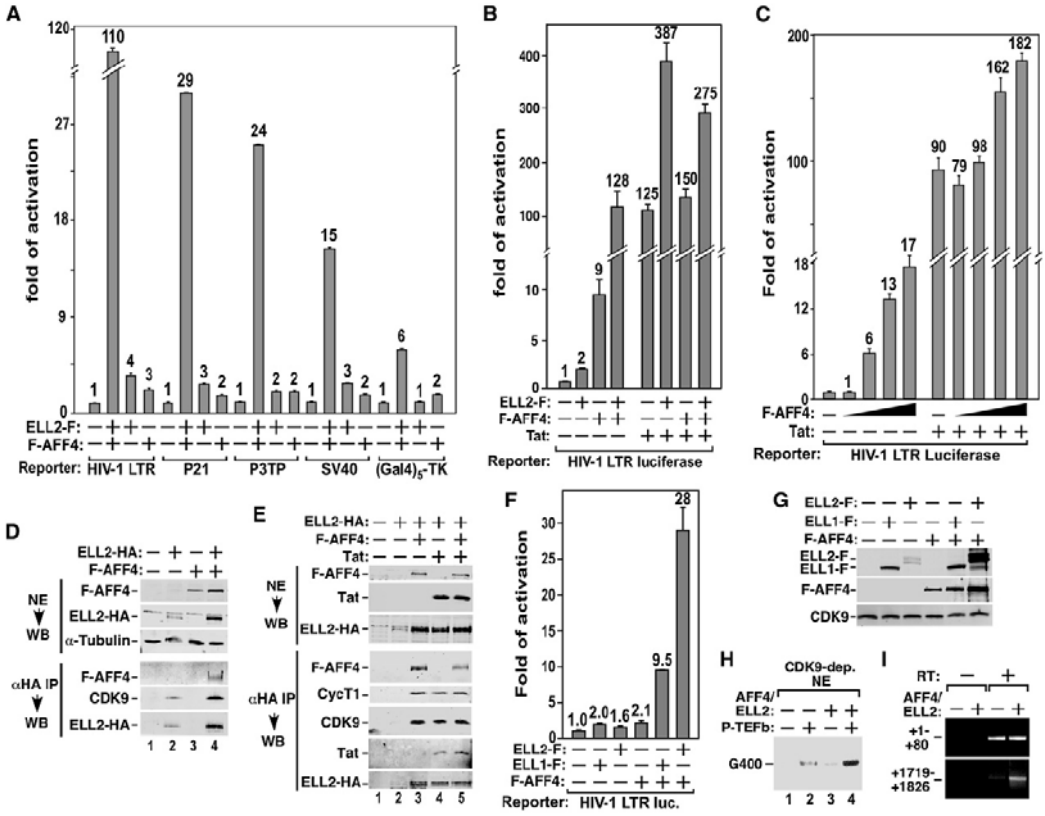

A
fold of activation
ELL2-F:
F-AFF4:
Reporter: HIV-1 LTR P21 P3TP SV40 (Gal4)5-TK
B
fold of activation
ELL2-F:
F-AFF4:
Tat:
Reporter: HIV-1 LTR luciferase
C
Fold of activation
F-AFF4:
Tat:
Reporter: HIV-1 LTR Luciferase
D
ELL2-HA:
F-AFF4:
NE
WB
F-AFF4
ELL2-HA
α-Tubulin
αHA IP
WB
F-AFF4
CDK9
ELL2-HA
1 2 3 4
E
ELL2-HA:
F-AFF4:
Tat:
NE
WB
F-AFF4
Tat
ELL2-HA
αHA IP
WB
F-AFF4
CycT1
CDK9
Tat
ELL2-HA
1 2 3 4 5
F
Fold of activation
ELL2-F:
ELL1-F:
F-AFF4:
Reporter: HIV-1 LTR luc.
G
ELL2-F:
ELL1-F:
F-AFF4:
ELL2-F
ELL1-F
F-AFF4
CDK9
H
CDK9-dep. NE
AFF4/ ELL2:
P-TEFb:
G400
1 2 3 4
I
RT:
AFF4/ ELL2:
+1- +80
+1719- +1826

Fig. 3-4 AFF4 bridges the ELL2-P-TEFb interaction and is required for the stable accumulation of ELL2 in the cell. A. The HIV-1 LTR-driven luciferase expression was measured in cells transfected with the HIV-1 LTR-luciferase reporter gene and constructs expressing shAFF4 or/and Tat-HA. The level of activity detected in the absence of shAFF4 or Tat-HA was set to 1, with the error bars representing mean +/- SD. B. Nuclear extracts (NE) of HEK293 cells either harboring an empty vector (-) or stably expressing shAFF4 (left panel) as well as immunoprecipitates (IP) obtained with a control IgG or the anti-CDK9 antibody from these extracts (right panel) were examined by Western blotting for the presence of the indicated proteins. C. NE and anti-Flag immunoprecipitates derived from cells transfected with an empty vector (-) or the construct expressing either the full-length (FL) or a truncated (Δ1-300) Flag-tagged AFF4 (F-AFF4) were analyzed by Western blotting for the indicated proteins. D. The indicated proteins were incubated with the immobilized CycT1-HA/CDK9 in binding reactions and the bound proteins were eluted and analyzed by Western blotting (left & middle panels). Ten percent of the input FL and Δ1-300 F-AFF4 proteins were also examined by anti-Flag Western blotting (right panel).

Fig. 3-4

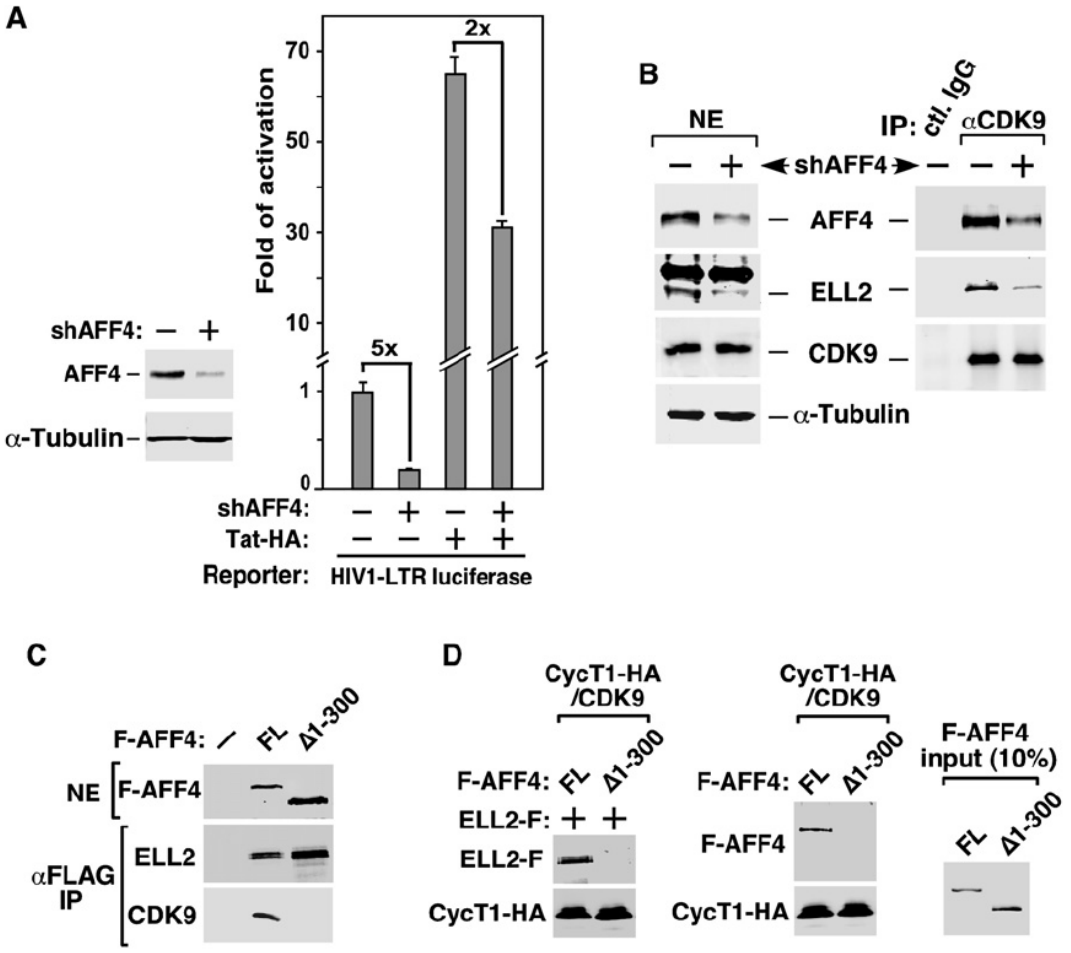

A
shAFF4: – +
AFF4 –
α-Tubulin –
Fold of activation
70
50
30
10
1
0
2x
5x
shAFF4: – + – +
Tat-HA: – – + +
Reporter: HIV1-LTR luciferase
B
NE
IP: ctl. IgG
αCDK9
– + shAFF4 – – +
– AFF4 –
– ELL2 –
– CDK9 –
– α-Tubulin
C
F-AFF4: FL Δ1-300
NE F-AFF4
αFLAG IP
ELL2
CDK9
D
CycT1-HA /CDK9
F-AFF4: FL Δ1-300
ELL2-F: + +
ELL2-F
CycT1-HA
CycT1-HA /CDK9
F-AFF4: FL Δ1-300
F-AFF4
CycT1-HA
F-AFF4 input (10%)
FL Δ1-300

Fig. 3-5 Tat further increases the amount of ELL2 bound to P-TEFb, leading to an increased association of the ELL2/AFF4/P-TEFb-containing complex with the HIV-1 chromatin template. A. Nuclear extracts (NE) derived from HEL293 cells infected with retroviruses either expressing (+) or not expressing (-) Tat-HA were subjected to Western blotting with the indicated antibodies (top right). The NEs were then immunoprecipitated with anti-CDK9 (top left), anti-ELL2 (bottom left), anti-ELL1 antibodies (bottom right) or a non-specific rabbit IgG. The immunoprecipitates were analyzed by Western blotting for the presence of the indicated proteins. B. Anti-Flag immunoprecipitates derived from cells stably expressing CDK9-F and harboring either an empty vector (-) or the shELL2 #8-expressing plasmid were analyzed by Western blotting for the indicated proteins. C. Chromatin immunoprecipitation (ChIP) with anti-Flag, anti-CDK9 and anti-AFF4 antibodies was performed in cells containing the integrated HIV-1 LTR-luciferase reporter gene and stably expressing ELL2-F. Three regions corresponding to the promoter, interior, and 3' UTR of the integrated HIV-1 LTR-luciferase gene (bottom panel) were PCR-amplified from the precipitated and purified DNA. Amplified signals from 5 and 10% of the input chromatin were also shown. D. ChIP assay was performed as in C at the GAPDH locus with the indicated antibodies. The region close to the 3' end of the gene was PCR-amplified from the precipitated DNA.

Fig. 3-5

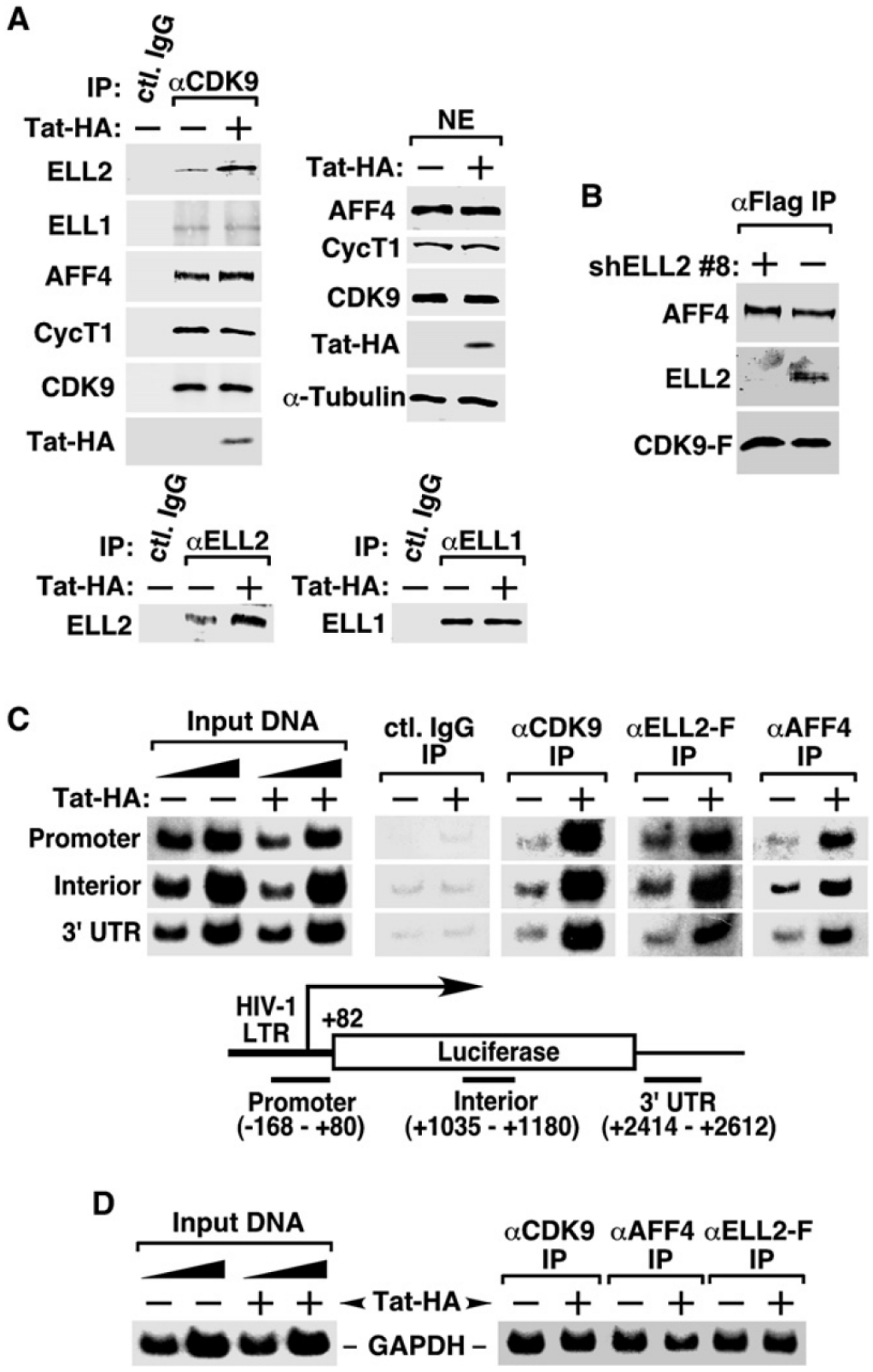

A
IP: ctl. IgG αCDK9
Tat-HA: – – +
ELL2
ELL1
AFF4
CycT1
CDK9
Tat-HA
NE
Tat-HA: – +
AFF4
CycT1
CDK9
Tat-HA
α-Tubulin
B
αFlag IP
shELL2 #8: + –
AFF4
ELL2
CDK9-F
IP: ctl. IgG αELL2
Tat-HA: – – +
ELL2
IP: ctl. IgG αELL1
Tat-HA: – – +
ELL1
C
Input DNA
ctl. IgG IP
αCDK9 IP
αELL2-F IP
αAFF4 IP
Tat-HA: – – + + – + – + – + – +
Promoter
Interior
3' UTR
HIV-1 LTR
+82
Luciferase
Promoter (-168 - +80)
Interior (+1035 - +1180)
3' UTR (+2414 - +2612)
D
Input DNA
αCDK9 IP
αAFF4 IP
αELL2-F IP
– – + + Tat-HA – + – + – +
– GAPDH –

Fig. 3-6 Dependence on active P-TEFb for ELL2 accumulation and interaction with P-TEFb, irrespective of whether the interaction is mediated by AFF4 alone or further promoted by HIV-1 Tat. A. NE derived from HEK293 cells transfected with the indicated cDNA constructs and treated with the indicated drugs were analyzed by Western blotting for the levels of ELL2-F and α-Tubulin. B & C. NE (left panel in B and top panel in C) and anti-Flag immunoprecipitates (right in B and bottom in C) derived from cells transfected with the indicated cDNA constructs were analyzed by Western blotting for the presence of the indicated proteins. D. F1C2 cells stably expressing CDK9-F were either untreated or treated with the indicated drugs. FVP: flavopiridol. NE (left panel) and anti-Flag immunoprecipitates (right panel) derived from NE were analyzed by Western blotting.

Fig. 3-6

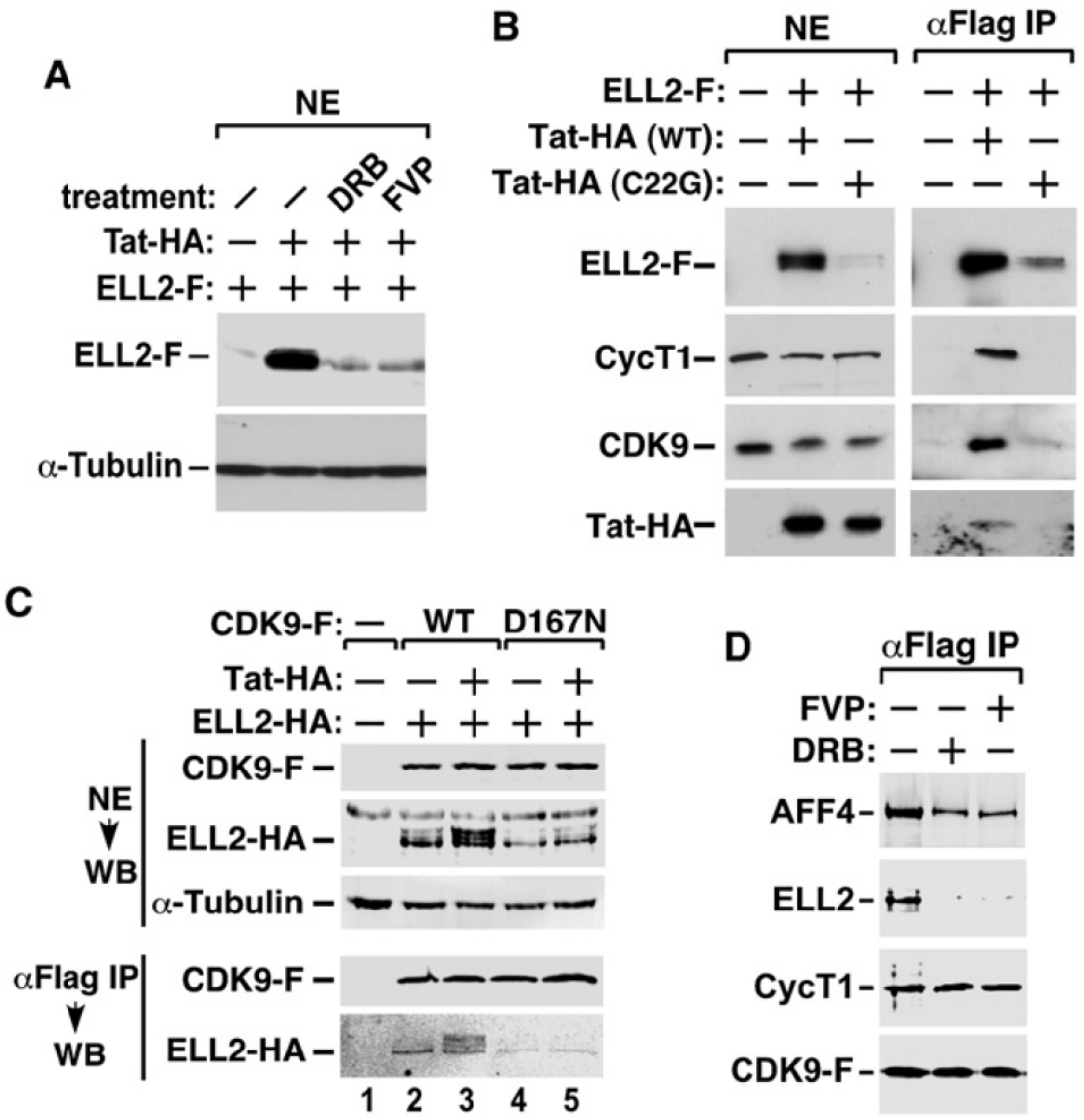
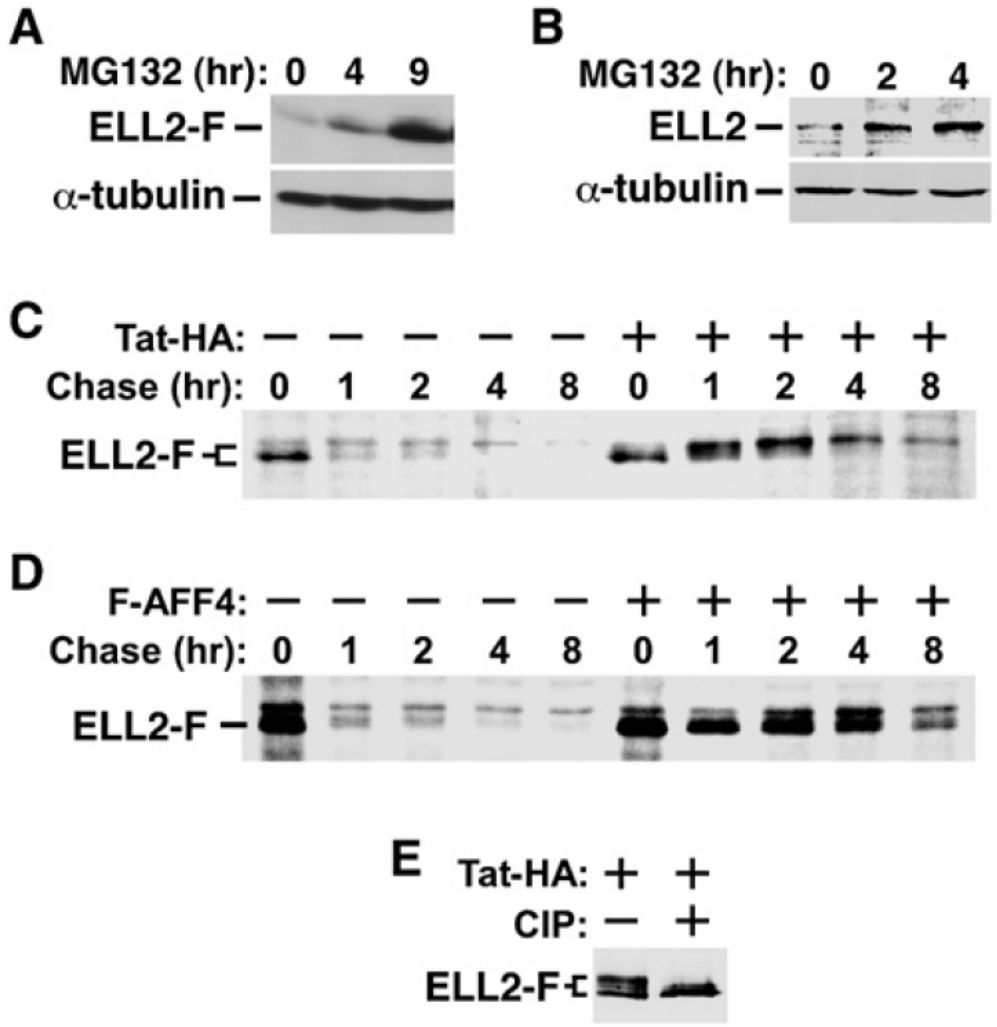

A
NE
treatment: / / DRB FVP
Tat-HA: – + + +
ELL2-F: + + + +
ELL2-F –
α-Tubulin –
B
NE
αFlag IP
ELL2-F: – + + – + +
Tat-HA (WT): – + – – + –
Tat-HA (C22G): – – + – – +
ELL2-F–
CycT1–
CDK9–
Tat-HA–
C
CDK9-F: – WT D167N
Tat-HA: – – + – +
ELL2-HA: – + + + +
NE
WB
CDK9-F –
ELL2-HA –
α-Tubulin –
αFlag IP
WB
CDK9-F –
ELL2-HA –
1 2 3 4 5
D
αFlag IP
FVP: – – +
DRB: – + –
AFF4–
ELL2–
CycT1–
CDK9-F–

Fig. 3-7 ELL2 is a short-lived protein whose stability can be significantly enhanced by Tat or AFF4. A. & B. HEK293 cells containing an ELL2-F-expressing vector (A) or nothing (B) were treated with MG132 for the indicated lengths of time. ELL2-F (A) and its endogenous counterpart (B) were detected by anti-Flag (A) and anti-ELL2 (B) Western blotting, with α-Tubulin serving as a loading control. C. & D. The ELL2-F-producing cells transfected with either an empty vector (-) or a construct expressing Tat-HA (B) or F-AFF4 (C) were pulsed-labeled with ^{35}S-labeled methionine and L-cysteine and then chased for the indicated periods of time. ELL2-F was then immunoprecipitated and analyzed by SDS-PAGE followed by autoradiography. E. ELL2-F affinity-purified from cells co-expressing Tat-HA were incubated with calf intestine phosphatase (CIP) and analyzed by anti-Flag Western blotting.

Fig. 3-7


A
MG132 (hr): 0 4 9
ELL2-F
α-tubulin
B
MG132 (hr): 0 2 4
ELL2
α-tubulin
C
Tat-HA: – – – – – + + + + +
Chase (hr): 0 1 2 4 8 0 1 2 4 8
ELL2-F
D
F-AFF4: – – – – – + + + + +
Chase (hr): 0 1 2 4 8 0 1 2 4 8
ELL2-F
E
Tat-HA: + +
CIP: – +
ELL2-F

Chapter 4:

ENL, AF9 and PAFc connect the Super Elongation Complex (SEC) to RNA polymerase II on chromatin

(This work was originally published in *PNAS* as: He N, Chan CK, Sobhian B, Chou S, Xue Y, Liu M, Alber T, Benkirane M, Zhou Q. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *PNAS* 2011, 108: E636-45)

Summary

The Super Elongation Complex (SEC), containing transcription elongation activators/co-activators P-TEFb, ELL2, AFF4/1, ENL and AF9, is recruited by HIV-1 Tat and mixed lineage leukemia (MLL) proteins to activate the expression of HIV-1 and MLL-target genes, respectively. In the absence of Tat and MLL, however, it is unclear how SEC is targeted to RNA polymerase (Pol) II to stimulate elongation in general. Furthermore, although ENL and AF9 can bind the H3K79 methyltransferase Dot1L, it is unclear whether these bindings are required for SEC-mediated transcription. Here, we show that the homologous ENL and AF9 exist in separate SECs with similar but non-identical functions. ENL/AF9 contacts the scaffolding protein AFF4 that uses separate domains to recruit different subunits into SEC. ENL/AF9 also exists outside SEC when bound to Dot1L, which is found to inhibit SEC function. The YEATS domain of ENL/AF9 targets SEC to Pol II on chromatin through contacting the PAFc complex. This explains the YEATS domain's dispensability for leukemogenesis when ENL/AF9 is translocated to MLL, whose interactions with PAFc and DNA likely substitute for the PAFc/chromatin-targeting function of the YEATS domain.

Introduction

Accumulating evidence has implicated the elongation stage of RNA polymerase II (Pol II) transcription as a major rate-limiting step for the expression of a large number of metazoan genes, especially those that control cell growth, renewal and differentiation (1, 5, 6). During elongation, the processivity of Pol II is regulated by a set of transcription factors, which had been thought to exist as separate entities and impact on the Pol II elongation complex independently of one another. However, recent data from us and others indicate that at least two well-defined transcription elongation factors of different classes reside in a single multisubunit complex termed SEC (super elongation complex, (95)) to cooperatively activate transcription (86, 95).

The first elongation factor found in SEC is human positive transcription elongation factor b (P-TEFb). Consisting of CDK9 and cyclin T1 (CycT1), P-TEFb functions by phosphorylating the C-terminal domain (CTD) of the largest subunit of Pol II and negative elongation factors DSIF and NELF. These events antagonize the actions of the negative factors, release Pol II from promoter-proximal pausing, and trigger the production of full-length mRNA transcripts (26, 51). The second elongation factor in SEC is ELL2, which promotes elongation by keeping the 3' OH of nascent mRNA in alignment with the catalytic site to prevent Pol II backtracking (77). Besides P-TEFb and ELL2, SEC also contains transcription factors/co-activators ENL, AF9, AFF4 (AF5q31), AFF1 (AF4) and probably others (86, 95). Among these, AFF4 is known to mediate the interaction between ELL2 and P-TEFb and maintain the integrity of SEC (95). While AFF1 has been shown to interact with AFF4 (96), it remains to be seen whether the interaction between these two homologous proteins can exist in a single SEC complex.

The biological importance of SEC has been highlighted by the recent discoveries that it is targeted by at least two sequence-specific transcriptional activators that play important roles in human diseases. First, the HIV-1 Tat protein is shown to interact with and recruit SEC to the HIV-1 LTR to stimulate viral transcription in a SEC-dependent manner (86, 95). In addition, Tat also promotes SEC formation, which in turn stabilizes ELL2, an otherwise short-lived protein rapidly degraded by the proteasome (95).

Besides HIV/AIDS, SEC is also found to be important for generating certain types of acute leukemias that involve chromosomal translocations of the mixed lineage leukemia *(MLL)* gene (85, 95, 96). At least three components of SEC, namely AFF4, ENL and AF9, are known as fusion partners of MLL (85). When fused to the DNA-binding domain of MLL, these proteins deliver SEC and its powerful elongation stimulatory activity to the MLL target genes to promote leukemic transformation (85, 95, 96).

Despite the demonstrations that SEC exists in cells under conditions that are free of Tat and MLL-translocations and is essential for metazoan transcriptional elongation in general (86, 95, 96), little is known about how this complex is targeted to Pol II on a chromatin template in the absence of sequence-specific recruitment factors like Tat and MLL. Furthermore, compared to the other subunits of SEC, the contributions of ENL and AF9 to SEC function have not been well characterized. So far, the only clue to how these two proteins may be involved in gene expression control arises from their

reported interactions with Dot1L (84, 97), a methyltransferase responsible for histone H3 lysine79 (H3K79) methylation (98, 99). As this modification is often found in actively transcribed regions (58, 100), the associations of ENL/AF9 with P-TEFb and DotlL are believed to recruit the latter two into a large complex for efficient activation of transcription (85, 101), an intriguing proposition that is yet to be proven. Besides the uncertainty about the Dot1L-ENL/AF9 interaction, it is also unclear whether the evolutionarily conserved YEATS domain located in the N-terminal regions of ENL and AF9, which are often missing in the MLL-ENL/AF9 fusions and unimportant for leukemogenesis (102), is required for SEC function.

Here, presenting answers to these questions, we show that the highly homologous ENL and AF9 exist in separate SEC complexes that display similar but non-identical functions. Within a SEC, ENL/AF9 is connected to the rest of the complex through the scaffolding protein AFF4, which uses separate domains to bind to different SEC subunits. Our data further show that Dot1L does not reside in or associate with SEC. Rather, it competes with AFF4 for binding to ENL/AF9 and inhibits SEC function. Finally, in the absence of sequence-specific recruitment factors, the YEATS domain of ENL/AF9 targets SEC to chromatin through contacting the PAFc transcription elongation complex (103, 104), and through PAFc, the elongating Pol II. This observation explains why the N-terminal regions of ENL and AF9 are dispensable for leukemogenesis when fused to MLL, whose abilities to bind PAFc (105) and DNA likely bypass the requirement for the PAFc/chromatin-targeting function of the YEATS domain.

Experimental Procedures

Antibodies

The anti-ENL (A302-267A), –AF9 (A300-595A) and –ELL2 (A302-505A-1) antibodies were purchased from Bethyl Laboratories, Inc. (Montgomery, TX). The anti-Dot1L (ab72454) and anti-H3 di-methyl K79 (ab3594) antibodies were purchased from Abcam. The antibodies against CDK9 and AFF4 have been described previously (95).

Co-immunoprecipitation (co-IP)

All co-IPs were performed in nuclear extracts (NEs) prepared from HeLa cells transfected with either specific cDNAs, siRNAs or shRNA-expressing constructs as indicated. The anti-Flag and anti-HA agarose beads were purchased from Sigma-Aldrich. After incubation at 4°C for 2 hr, the immunoprecipitates were washed with buffer D0.3M (20 mM HEPES, pH 7.9, 10% glycerol, 0.3 M KCl, 0.2 mM EDTA, 0.2% NP-40, 1 mM DTT, and 0.5 mM PMSF) unless indicated otherwise. The purified materials were eluted off the beads with buffers containing synthetic Flag or HA peptides as described (106) and analyzed by Western blotting with the indicated antibodies.

In vitro binding assay

Proteins used for in vitro binding assay were affinity-purified under highly stringent conditions (1.0 M KCl plus 0.5% NP-40) to strip away their binding partners as described (95). HA-AFF4 or the CycT1-HA/CDK9 complex immobilized on anti-HA-agarose beads was isolated under similar conditions. Prior to the binding

assay, all the proteins were checked by Western blotting to ensure that they were free of their normal binding partners. For the binding reactions, approximately 100 ng of each individual protein was incubated with immobilized HA-AFF4 or CycT1-HA/CDK9 isolated from 50 μL of NE. The washing and eluting conditions were as described (95).

The expression and purification of PAF1, GST-ENL-N and GST-ENL-C from recombinant E. coli were described in detail in Supplemental Materials. Pull-down assays were performed using 10 ug of each protein in a final volume 1ml of Buffer400 (400 mM NaCl, 20 mM HEPES, pH 8.0, 0.5 mM TCEP, 10% glycerol, and 0.1% NP-40). Proteins were incubated on a rotator for 30 min at 4oC followed by the addition of 10 uL GST beads and further incubation for 30 min. Beads were then washed in Buffer400 and eluted with SDS-PAGE sample buffer.

Chromatin immunoprecipitation (ChIP) assay

The ChIP assay was carried out essentially as described (107) with minor modifications. After formaldehyde fixation, HeLa cells (2x107) or a HeLa-based cell line containing a stably integrated HIV-1 LTR Luciferase reporter gene and transfected with the indicated ENL/AF9-expressing constructs were incubated in lysis buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl2, 10 mM KCl and 1% NP-40) on ice for 10 minutes and the nuclei were collected and re-suspended in sonication buffer (15 mM Tris-Hcl, pH 8.0, 0.1% SDS, 1% Triton X-100, 1 mM EDTA, 150 mM NaCl). For RNAi-coupled ChIP analyses, HeLa cells were treated for 48 hr with specific siRNA duplexes (Dharmacon) using INTERFERin^{TM} (Polyplus) and then subjected to ChIP analyses as mentioned above. Primers used to amplify the HEXIM1 gene are: forward 5'-TAACACCACGCAGTTCCTCATGGA-3' and reverse 5'-TGAGCTCCTGCTTGCTCATGTTCT-3'. For the c-Myc gene, the forward primer is: 5'-ACTCGGTGCAGCCGTATTTCTACT-3' and the reverse primer is: 5'-GCAGCAGCTCGAATTTCTTCCAGA-3'. The nucleotide sequences of the PCR primers used to examine the occupancy of ENL/AF9 at the HIV-1 promoter, the interior of the luciferase coding sequence and the 3' untranslated region (3' UTR) were essentially as described (95).

Generation of inducible PAF1 knockdown cells

The DNA oligonucleotide encoding the PAF1-specific shRNA (shPAF1; 5'-GATCAAGGTGGCAGTGACAATGATTTTCAAGAGAAATCATTGTCACTGCCA CCTTTTTTA-3') was cloned into the pSuperior.retro.neo+GFP construct (OligoEngine), which was introduced by retroviral infection into the T-RExTM-293 (Invitrogen)-based cell line stably expressing CDK9-F as described previously (95). The procedures for the production of recombinant retroviruses, infection of cells, and generation of neomycin-resistant colonies have been described previously (95). Single colonies were picked and screened for inducible knockdown of endogenous PAF1 upon the treatment with doxycycline (1 μg/ml) for two days.

Generation of ENL and AF9 knockdown cells

The procedure for generating a HeLa-based cell line in which the expressions of ENL and AF9 were silenced simultaneously has been described previously (95). The shRNA sequences used in the current procedure are:

ENL-ctrl.:
5’GATCAATCAGTGCACCGTCCAGGTGTTCAAGAGACACCTGGACGGTGCACTGATTTTTA 3’

ENL-sh3:

5’GATCAAGGTCTGCTTCACCTACGACTTCAAGAGAGTCGTAGGTGAAGCAGACCTTTTTA 3’

AF9-ctrl.:

5’GATCGCGGTCCGGAGCACAGTAACAGCTTCCTGTCACTGTTACTGTGCTCCGGACCGCTTTTTA 3’

AF9-sh10:

5’GATCGAGTTACCTGGAAACATCTGGGCTTCCTGTCACCCAGATGTTTCCAGGTAACTCTTTTTA 3’

siRNA-mediated knockdown of Dot1L expression

HeLa cells containing an integrated LTR-luciferase reporter construct were transfected twice with Dot1L-specific or control scrambled siRNA (ctl.). Six days post transfection cells were transduced with a retrovirus encoding for Tat or empty vector. siRNA sequences used: Ctl.: 5’-auguauuggccuguauuagtt -3'; siDot1L: 5’-uguauuggccuguauuagtt -3’.

Results

ENL and AF9 interact with P-TEFb through the scaffolding protein AFF4s

In light of our previous demonstration that AFF4 bridges the ELL2-P-TEFb interaction in SEC (95), we asked whether AFF4 plays a similar role in mediating the interactions of ENL and AF9 with P-TEFb. To answer this question, short hairpin (sh)RNA-mediated depletion of AFF4 was performed in HeLa cells to assess its impact on the associations of ENL and AF9 with P-TEFb. Although the depletion did not affect the total levels of ENL and AF9 in nuclear extracts (NE; Fig. 4-1A, left panel), it significantly reduced the levels of ENL and AF9 bound to the immunoprecipitated CDK9 (right panel).

Next, in vitro binding reactions were performed to test whether AFF4 could directly bridge the interaction between purified P-TEFb and ENL/AF9 in the absence of other SEC components. All the proteins added to the reactions were highly purified from transfected HeLa cells under stringent conditions (1.0 M KCl plus 0.5% NP-40) and confirmed by SDS-PAGE followed by silver-staining to be free of any contaminating proteins (95). As indicated in Fig. 4-1B, purified ENL and AF9 did not interact with immobilized CycT1-HA/CDK9 unless AFF4 was also present in the same reactions. Together, these in vitro and in vivo binding data highlight the important role for AFF4 to serve as a molecular scaffold to mediate the interactions of P-TEFb with not only ELL2 but also ENL and AF9.

ENL and AF9 directly interact with AFF4 through their C-terminal regions

Since AFF4 was shown to interconnect P-TEFb and ENL/AF9 in the absence of other SEC components, we reasoned that it must be able to make direct and simultaneous contacts with P-TEFb and ENL/AF9. In fact, a direct binding between

AFF4 and P-TEFb has been demonstrated previously (95). To show that ENL and AF9 can also directly contact AFF4, in vitro binding reactions employing highly purified proteins were conducted. While wild-type (WT) ENL and AF9 were able to bind to immobilized HA-AFF4 directly, the C-terminally truncated ENL (1-431) and AF9 (1-480) were not (Fig. 4-1C). The requirement for the C-terminal regions of ENL and AF9 for binding to AFF4, which in turn allowed the formation of a complete SEC complex, was also confirmed in vivo in transfected cells (Fig. 4-1D & 1E).

Separate regions of AFF4 are used to interact with different subunits of SEC

Given the demonstrations that the scaffolding protein AFF4 can directly and simultaneously contact each and every subunit of SEC (Fig. 4-2 and (95)), we would like to map its regions that are involved in these interactions. A series of Flag-tagged AFF4 deletion mutants were tested for their ability to co-precipitate the other components of SEC. When the levels of WT and mutant AFF4 in the anti-Flag immunoprecipitates (IP) were normalized to a similar level, the mutant missing the first 300 residues, Δ1-300, was found to interact with ELL2, ENL and AF9 normally but not CDK9 (Fig. 4-2, lane 3). In contrast, a short AFF4 truncation mutant containing just the first 300 amino acids (1-300) was able to efficiently pull down CDK9 but not any other components of SEC (lane 6). Thus, the first 300 residues of AFF4 contain an independent domain that is necessary and sufficient for P-TEFb-binding.

Similarly, the region between aa 301 and 600 in AFF4 appears to contain an ELL2-binding domain. This is indicated by the demonstration that the deletion mutant Δ1-600 failed to interact with ELL2 and P-TEFb, but still retained wild-type ability to bind to ENL and AF9 (Fig. 4-2, lane 4). On the other hand, a 300 residue-long segment of AFF4 encompassing positions 301 to 600 (301-600) displayed reduced but clearly above-the-background level of binding to ELL2 but not any other components of SEC (lane 7). Finally, the region between amino acids 601 and 900 of AFF4 likely contains an independent binding domain for ENL and AF9, as the fragment encompassing this region interacted efficiently with these two proteins but not any other subunits of SEC (lane 8) and the deletion mutant lacking the entire N-terminal 900 amino acids (Δ1-900) failed to associate with any component of SEC (lane 5). Together, these data indicate that AFF4 uses separate domains to interact with different subunits of SEC and serves as a platform to nucleate the assembly of SEC.

ENL and AF9 exist in separate SEC complexes

ENL and AF9 are highly homologous proteins, with their N-terminal YEATS domains and C-terminal coiled-coil domains showing a particularly high degree of identity (82%). Given this homology and the above demonstration that both proteins bound to the same region (aa 601-900) of AFF4 (Fig. 4-2), we asked whether they exist simultaneously in a single SEC complex. To address this question, anti-Flag immunoprecipitations were performed in extracts of cells co-expressing HA-tagged AF9 and Flag-tagged ENL or AFF4. While AF9-HA co-precipitated with F-AFF4 and CDK9 as expected, it did not co-precipitate with ENL-F (Fig. 4-3A). The reciprocal co-IP experiment employing anti-HA beads also failed to detect the interaction

between AF9-HA and ENL-F. These data suggest that AF9 and ENL do not exist in the same complex, although each can establish independent interactions with other SEC subunits.

To obtain further evidence in support of this notion, we performed immunodepletion to remove all ENL proteins from HeLa NE (Fig. 3B). This procedure caused the co-depletion of more than 75% of AFF4, ~30% of CDK9, but very few AF9 and the internal control protein, α-Tubulin, from NE. This result is consistent with the idea that ENL and AF9 do not exist in the same SEC complex. Since a major fraction of nuclear CDK9 is normally sequestered in the 7SK snRNP (26), it is not surprising to see that more than half of CDK9 were still present in the ENL-depleted NE. However, the co-depletion of more than 70% of AFF4 suggests that the AF9-containing SEC contains at most 30% of total AFF4 in NE, assuming that no AFF4 exists outside of the SEC complexes. Thus, the ENL-containing SEC appears to be the predominant form between the two SEC complexes.

ENL and AF9 compete for binding to AFF4

A likely reason for the failure of ENL and AF9 to coexist in the same SEC is revealed subsequently in an in vitro binding assay employing highly purified proteins. The addition of increasing levels of WT AF9-F, but not the C-terminally truncated AF9-F 1-480 that is defective for AFF4-binding (Fig. 4-1C), into binding reactions progressively reduced the amounts of F-AFF4 bound to the immobilized ENL-HA (Fig. 4-3C). Thus, the bindings of AF9 and ENL to AFF4, which occur through the same region (aa 601-900) of AFF4 (Fig. 4-2), were mutually exclusive, which explains their failure to exist in the same SEC complex.

Loss of ENL is compensated by increased AF9 expression, but not vice versa

The existence of two versions of SEC with one containing AF9 and the other ENL makes it important to examine whether these two homologous proteins have similar or different functions. Toward this goal, stable shRNA-mediated depletion of either ENL or AF9 was performed and the effect on SEC formation was assessed by anti-CDK9 immunoprecipitation followed by Western blotting. Notably, the reduction of total and CDK9-bound ENL in HeLa cells by shENL triggered a marked increase in the amounts of AF9, but not AFF4 or ELL2, in both NE and anti-CDK9 IP (Fig. 4-3D). In contrast, shRNA depletion of AF9 did not produce any significant change in ENL expression or sequestration into SEC (Fig. 4-3E). Thus, the loss of the ENL-containing SEC was compensated by the increased formation of the AF9-containing SEC, but not vice versa. This is likely caused by the fact that the former complex is the predominant form between the two SECs, and the compensation serves to prevent a major reduction in the overall SEC level in the cell.

AF9 and ENL have similar but non-identical functions in mediating SEC-dependent transcription

In agreement with the above demonstration that AF9 increased expression to compensate for the loss of ENL but not vice versa, we consistently observed that shRNA depletion of ENL had a smaller effect on HIV-1 transcription, which is shown to depend on the SEC function (86, 95), than did the depletion of AF9. For example, using the HIV-1 LTR-driven luciferase expression as readout, shENL reduced basal

and Tat-activated HIV LTR activity by 29% and 24%, respectively (Fig. 4-3F & 3G), whereas shAF9 decreased the two by 40% and 37%, respectively, when the knockdown efficiency was normalized between the two shRNAs. Despite this difference, the overall impact on the HIV-1 LTR by the individual depletion of AF9 or ENL was relatively minor. Only when both proteins were co-depleted at the same time, a more drastic reduction in HIV-1 transcriptions (72% reduction for basal and 69% for Tat-activated transcription) was observed (Fig. 4-3F & 3G). The fact that the depletion of ENL or AF9 alone only partially suppressed SEC-dependent HIV-1 transcription whereas the depletion of both had a much more significant effect indicates that the two proteins have similar but not completely identical functions.

Dot1L is not a component of SEC

The identification of ENL/AF9 as a subunit of SEC raises the issue of whether the methyltransferase Dot1L, a well-known partner of ENL and AF9 (84, 97), also exists and plays a key role in this complex. Dot1L is responsible for the methylation of H3K79 (98, 99), a modification that is often associated with actively transcribed genes (58, 100). Because of their overall positive influence on transcription and common connection through ENL/AF9, Dot1L and the key SEC subunits P-TEFb and ELL have been depicted in recent models as components of a larger complex for coordinated activation of transcription (85, 101).

To determine whether Dot1L is associated with SEC, anti-Flag immunoprecipitations were performed in NE of HeLa cells expressing F-Dot1L or F-AFF4. While F-AFF4 readily co-precipitated with the SEC components ENL, AF9 and CDK9 as expected, F-Dot1L was only able to pull down about the same amounts of ENL and AF9 but not CDK9 under identical conditions (Fig. 4-4A). This suggests that the interactions of Dot1L with ENL/AF9 likely occurred outside of SEC, a notion also proposed by Lin et al. (85, 95, 96) without showing the data.

It has previously been shown that HIV-1 Tat can bind to and promote the formation of SEC (86, 95). Consistent with this observation, transiently expressed Tat-HA precipitated SEC subunits CDK9 and AFF4 but not Dot1L, although a separate ENL-Dot1L interaction was easily detected in the same experiment (Fig. 4-4B).

ENL/AF9 cannot simultaneously interact with AFF4 and Dot1L

The data in Fig. 4-1 show that ENL and AF9 use their C-terminal regions to contact AFF4. These same regions, however, are also responsible for direct interactions of ENL and AF9 with Dot1L (84, 97). To explain why Dot1L was not detected in SEC, we asked whether ENL/AF9 could bind to Dot1L and AFF4 at the same time. To answer this question, an in-vitro binding/competition assay employing highly purified proteins was performed. In agreement with the data in Fig. 4-1, only WT ENL-F, but not the C-terminally truncated ENL-F 1-430, bound to the immobilized HA-AFF4 (Fig. 4-4C). However, when increasing amounts of Dot1L were added to the reactions, decreasing levels of WT ENL were found to associate with HA-AFF4 (Fig. 4-4C). A similar observation was also obtained with AF9 (Fig. 4-4D). Thus, Dot1L and AFF4 directly compete for binding to ENL and AF9, which explains why Dot1L was not detected in SEC. These results are consistent with the

observations by Yokoyama et al. (96) that ENL co-precipitated with Dot1L and AFF4 when co-expressed in HEK293 cells and that Dot1L and AFF4 failed to co-precipitate under the same conditions.

Dot1L inhibits SEC-dependent transcription

Even though Dot1L is not an integral component of SEC, we asked whether it might still be required for basal and Tat-activated HIV-1 transcription, which is shown to proceed in a SEC-dependent manner (86, 95). To this end, the siRNA-mediated silencing of Dot1L expression was performed, which was found to also significantly reduce the nuclear level of H3K79me2 but not AF9 as expected (Fig. 4-4F). Surprisingly, the loss of Dot1L and H3K79me2 not only failed to inhibit the HIV-1 LTR; it actually enhanced the luciferase production from a stably integrated HIV-1 LTR-driven lucfierase reporter gene under both Tat(+) and (-) conditions (Fig. 4-4E). It is interesting to note not all genes depend on Dot1L for expression. The data above place the HIV-1 LTR among a group of actively transcribed genes that are known to be inhibited by Dot1L (108, 109).

The YEATS domains of ENL and AF9 are unimportant for SEC formation but essential for SEC-dependent transcription

ENL and AF9 are two of the most frequent MLL fusion partners for inducing acute leukemias. When fused to MLL, their C-terminal regions are necessary and sufficient for oncogenic activities of the fusion products (96, 102). In contrast, their N-terminal regions that contain the highly conserved YEATS domain (110) are largely dispensable for leukemic transformation.

To determine whether the YEATS domains of AF9 and ENL play any role in SEC function, we compared WT AF9 and ENL and their mutants lacking the N-terminal YEATS domain for their abilities to promote luciferase gene expression driven by the HIV-1 LTR, whose activity requires a functional SEC (86, 95). Whereas WT AF9 and ENL activated the LTR by 9.3 and 5.4 folds, respectively, the deletion mutants produced only 1.3- and 2.1-fold increase, despite their stable accumulation in transfected cells (Fig. 4-5A). These results reveal a critical role for the YEATS domains in mediating the SEC-dependent HIV-1 transcription. However, our subsequent co-IP experiments indicate that like their WT counterparts, the N-terminal deletion mutants of AF9 and ENL were fully capable of interacting with all the other components of SEC (Fig. 4-5B & 5C). Thus, the YEATS domains contribute to the SEC function not through maintaining the integrity of the complex.

The YEATS domains of ENL and AF9 display chromatin-targeting function

Given the observations that the YEATS domains is crucial for AF9/ENL to support the SEC-dependent HIV-1 transcription but dispensable for MLL-AF9/ENL to cause leukemia (96, 102), we wanted to know exactly how this domain contributes to SEC function. Since MLL confers the chromatin-targeting function to the MLL-AF9/ENL fusions, we asked whether the YEATS domain plays a similar role in promoting the interaction of SEC with chromatin in the absence of a sequence-specific recruitment factor like Tat or MLL. To this end, the chromatin immunoprecipitation (ChIP) assay was performed to test the associations of WT ENL and AF9 as well as their YEATS domain deletion mutants with a stably integrated

HIV-1 LTR-lucfierase reporter gene. While WT ENL and AF9 were detected readily on the chromatin template at the promoter, interior of the luciferase ORF and the 3' UTR regions, the deletion mutants displayed significantly reduced occupancy at these locations (Fig. 4-6A and 6B), indicating that the YEATS domain is crucial for ENL/AF9 to interact with the HIV-1 chromatin template.

The YEATS domain directly binds to the PAF1 subunit of PAFc to target SEC to chromatin

What could be the functional target on chromatin that is used by the YEATS domain to deliver SEC to Pol II? Recently, an interaction of SEC with the human Polymerase-Associated Factor complex (PAFc), which is known to associate with Pol II during productive elongation (103, 106), has been discovered in cells expressing HIV-1 Tat (86). It is yet to be shown whether this interaction also occurs under Tat(-) conditions and what role it may play during SEC-dependent transcription. As the first step toward answering these questions, we asked whether the YEATS domain might be required for the SEC-PAFc interaction. Indeed, compared to WT ENL, the mutant missing the N-terminal YEATS domain (aa 113-559) interacted normally with CDK9 but not PAF1, the scaffolding subunit of the multi-component PAFc (Fig. 4-6C). In comparison, the ENL mutant lacking the C-terminal AFF4/P-TEFb-binding domain (aa 1-430) showed wild-type interaction with PAF1 but not CDK9.

Just like ENL, AF9 also depended on the YEATS domain to interact with PAF1 (Fig. 4-6D). The subsequent in vitro GST pull-down assay further reveals a direct physical interaction between recombinant PAF1 and the ENL YEATS domain that is fused to GST (GST-ENL-N; Fig. 4-6E). In contrast, GST-ENL-C, which retains the C-terminal AFF4/P-TEFb-binding region, failed to bind to PAF1.

PAFc connects SEC to Pol II on chromatin templates

Given that the YEATS domain mediated the interaction of SEC with PAFc, which is known to associate with Pol II during elongation (103), we postulated that this domain must also play a key role in allowing SEC to ultimately reach its functional target Pol II. Indeed, co-IP experiments in Fig. 4-7A reveal that compared to WT ENL, the YEATS-deleted mutant (aa113-559) consistently showed decreased binding to RPB1, the largest subunit of Pol II. To ensure that the observed interactions with endogenous PAF1 and RPB1 by transfected WT ENL-F and AF9-F are not a result of transient overexpression of the latter two proteins, we expressed specific shRNAs to enable simultaneous knockdown of both ENL and AF9, which were shown to exist in two separate SECs and display similar functions (Fig. 4-3). Confirming the dependence on the YEATS-containing ENL/AF9 for the interactions of endogenous SEC with PAFs and Pol II, the combination of shENL and shAF9 markedly decreased the amounts of PAF1 and RPB1 associated with the immunoprecipitated CDK9 (Fig. 4-7B).

Theoretically speaking, the above-described YEATS domain-dependent interaction between SEC and Pol II could proceed independently of PAFc. To prove that PAFc indeed acts as a bridge to interconnect SEC and Pol II, we performed shRNA-mediated depletion of PAF1 under inducible conditions. Upon the induction of PAF1 depletion by doxycycline, which activated shPAF1 expression, a significant

reduction in the amount of RPB1 bound to the immunoprecipitated CDK9-F, a key SEC subunit, was observed (Fig. 4-7C).

Given the important role for PAFc in mediating the interaction of SEC with elongating Pol II, we predicted that PAFc must also be required to target SEC to chromatin templates that encompass endogenous genes. Indeed, siRNA-mediated PAF1 knockdown in HeLa cells was found to markedly reduce the associations of the SEC subunit CDK9 with two endogenous gene loci c-Myc and HEXIM1 (Fig. 4-7D), which have been shown to depend on P-TEFb for expression (39, 51). Very similar observations were also obtained in 293T cells. Taken together, the data above support a model (Fig. 4-7E) that the YEATS domain of ENL/AF9 functions as a bridge to connect SEC to PAFc, and through PAFc, the elongating Pol II on a chromatin template. Once positioned next to Pol II, SEC applies its two functional modules, P-TEFb and ELL2, to synergistically stimulate the processivity of Pol II and facilitate elongation-coupled mRNA 3' processing, leading to the production of full-length, polyadenylated mRNA transcripts in a Dot1L-independent manner.

Discussion

Of all the SEC subunits, ENL and AF9 had been the least characterized up until this moment, which had prevented a thorough understanding of the SEC functions in normal and disease-associated transcriptional elongation. The current study represents a major effort aimed at correcting this deficiency. Our data indicate that ENL and AF9, which display extensive sequence homology, compete for binding to the same region of the scaffolding protein AFF4 and thus cannot reside in the same complex. The AF9-SEC and ENL-SEC complexes, with the latter appearing to be more predominant in HeLa cells, show similar but not completely identical functions. It is possible that the existence of two homologous SEC complexes with complementary functions can better address the needs of diverse cellular and viral genes to enable efficient transcriptional elongation under different conditions. Our data further show that within each SEC complex, ENL/AF9 is connected to the other subunits through AFF4, which acts like an assembly platform and uses separate regions to contact different subunits to nucleate the formation of SEC.

Notably, the notion of separate AF9-SEC and ENL-SEC complexes is also supported by recent glycerol gradient analysis of isolated SEC complexes, which reveals incompletely overlapping distributions of ENL and AF9 despite their very similar sizes and close identity (86, 95). Thus, although the two complexes share the common subunits AFF4, ELL2, CDK9, and Cyclin T1 (as revealed in the current study), they may have somewhat different compositions or structures. Future in-depth analyses are necessary to identify this difference, which could be responsible for the non-identical but complementary functions of the two complexes.

Another key finding of the current study concerns the methyltransferase Dot1L, which is a major binding partner of ENL and AF9 and has received much attention because of its modification of H3K79 and possible involvement in leukemogenesis (111). Contrary to the models proposed in several recent articles (85, 101), our data

indicate that Dot1L does not associate with SEC. Rather, it competes with AFF4 for binding to ENL/AF9, suggesting that the Dot1L-ENL/AF9 interactions occur outside of the SEC complex, a conclusion that was also reached under different experimental conditions (96). Consistent with this conclusion, it was recently shown that the MLL-ENL fusion delivers Dot1L and SEC as two separate entities to the MLL target loci. However, contradicting with the general presumption of the field, it is the recruitment of SEC that contributes predominantly to MLL-dependent leukemogenesis (96).

The methylation of H3K79, including mono-, di- and tri-methylation, is likely carried out exclusively by Dot1L (99). So far, these modification marks have displayed a complex relationship with gene transcription. For example, genome-wide analyses in Drosophila indicate that hypermethylated H3K79 is frequently enriched within actively transcribed genes, whereas hypomethylated H3K79 is generally associated with inactive genes (100). A subsequent high-resolution profiling of histone methylations in the human genome has painted a more detailed picture by showing that while mono-methylation of H3K79 is linked to gene activation, tri-methylation is associated with repression (108). In an attempt to examine the relationship between H3K79 methylation and transcriptional control at a single gene locus, it was found that H3K79 di-methylation marks developmental activation of the -globin gene but is reduced upon LCR-mediated high-level transcription (109). Taken together, these studies send a clear message that H3K79 methylation and its responsible enzyme Dot1L play complicated and sometimes conflicting roles in controlling gene expression. Depending on the specific genes and conditions involved, they can exert either a positive or negative influence on transcription. The inhibitory effect of Do1L on HIV-1 transcription observed in the current study provides yet another example supporting the notion that Do1L and H3K79 methylation are not always associated with transcriptional activation.

The final important finding of the current study is about the highly conserved YEATS domain in ENL and AF9. Many YEATS domain-containing proteins are components of histone-modifying and transcription complexes (110). However, the function of this domain itself remains poorly understood thus far. Prior to the current study, the only clue suggesting how it might contribute to the activity of SEC comes from the observation that the YEATS domain of ENL interacts with histones H3 and H1 in vitro (112). These interactions, which are yet to be confirmed in vivo, could in principle play a role in attracting SEC to a chromatin template. However, since H3 and H1 are not specifically associated with active transcription, their interactions with ENL/AF9 are not expected to recruit SEC to only the actively transcribed genes, let alone to keep SEC continuously engaged in co-migrating with its functional target, Pol II, during elongation.

These tasks, while challenging for histones H3 and H1, would be fairly straightforward and natural for the multi-subunit and multi-functional PAFc to accomplish. This is because PAFc is well-known for its association with the elongating Pol II (103, 104), contribution to transcriptional elongation on chromatin templates (103) and participation in transcription-coupled mRNA 3' processing (113).

Furthermore, it also interacts directly with SEC, although the physiological significance of this interaction was unknown at the time (86). These functions make PAFc an ideal candidate for the YEATS domain of ENL/AF9 to bind and deliver SEC to Pol II on a chromatin template. Indeed, data presented here completely agree with this notion and indicate that the YEATS domain directly binds to the PAF1 subunit, which is a scaffolding molecule to reinforce the binary interactions between other subunits of PAFc (103). This interaction is shown to target SEC to Pol II and chromatin and allow SEC to stimulate productive elongation and likely also transcription-coupled mRNA polyadenylation. This latter role of SEC is supported by the demonstrations that the SEC component P-TEFb plays a key role in 3' processing and that the distribution of PAFc is concentrated toward the 3' end of genes (3, 13, 114).

It is interesting to note that SII/TFIIS, another well-known transcription elongation factor that acts by inducing transcript cleavage in arrested elongation complexes and permitting paused Pol II to proceed downstream, was recently shown to cooperate with PAFc to bind to Pol II and stimulate elongation (103). It will be interesting to test whether the interactions of PAFc with SII and SEC occur simultaneously or in an exclusive manner, which will determine whether different elongation activities as represented by SEC and SII can work on the same polymerase enzyme to achieve synergistic activation.

The discovery of the chromatin/PAFc-targeting function of the YEATS domain has also helped clarify a long-standing confusion stemming from the observations that this domain confers autonomous transactivation of the SV40 minimal promoter (112) and contributes to the SEC function (Fig. 4-5A) but is nevertheless dispensable for oncogenic transformation in the context of the MLL-ENL/AF9 translocations (96, 102). It is interesting to note that the MLL portion of the fusion proteins contains multiple DNA binding structures that enable both sequence-specific and -nonspecific bindings to the target loci (102) and also the CxxC-RD2 domain for interacting with PAFc (105). It is highly likely that these activities of MLL, which are known to be essential for transformation, can effectively substitute for the PAFc/chromatin-targeting function of the YEATS domain. Besides ENL and AF9, 132 additional proteins in 59 different eukaryotes are also known to possess the YEATS domain (110). Future studies will shed light on whether the ability of this domain to target chromatin, PAFc and Pol II is evolutionarily conserved and how it may contribute to the biological functions of diverse YEATS domain-containing proteins.

Fig. 4-1 The scaffolding protein AFF4 directly binds to the C-terminal regions of ENL and AF9 to mediate their interactions with P-TEFb. A. Nuclear extracts (NE) were prepared from HeLa cells either containing an empty vector or expressing the AFF4-specifc shRNA (shAFF4) and subjected to immunoprecipitation (IP) with anti-CDK9 or an irrelevant rabbit IgG as a control. The isolated NE (left panel) and immunoprecipitates (right panel) were analyzed by Western blotting with the indicated antibodies. B. The indicated highly purified proteins were incubated with immobilized CycT1-HA/CDK9 in vitro and the bound proteins were eluted and analyzed by Western blotting (right). Five percent of the input proteins were also examined by anti-Flag Western blotting (left). C. In vitro binding reactions were performed by incubating highly purified wild-type (WT) or C-terminally truncated ENL-F or AF9-F (schematic diagram on the right) with immobilized HA-AFF4. The bound proteins and 10% of the soluble input proteins were analyzed by Western blotting with the indicated antibodies. D. WT or mutant ENL and AF9, all Flag-tagged, were expressed in transfected HeLa cells. Anti-Flag immunoprecipitates were examined by Western blotting for the indicated ENL/AF9-associated factors.

Fig. 4-1

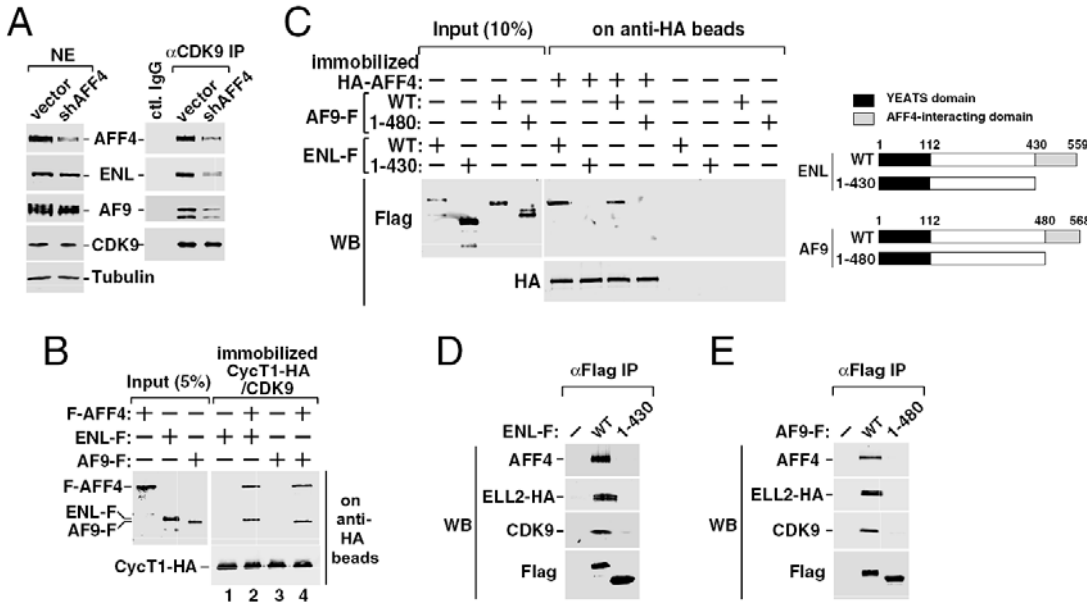

A
NE
αCDK9 IP
vector
shAFF4
ctl. IgG
vector
shAFF4
AFF4
ENL
AF9
CDK9
Tubulin
C
Input (10%)
on anti-HA beads
Immobilized
HA-AFF4:
AF9-F
WT:
1-480:
ENL-F
WT:
1-430:
WB
Flag
HA
YEATS domain
AFF4-interacting domain
ENL
WT
1-430
AF9
WT
1-480
B
Input (5%)
immobilized
CycT1-HA
/CDK9
F-AFF4:
ENL-F:
AF9-F:
F-AFF4
ENL-F
AF9-F
CycT1-HA
on
anti-
HA
beads
1 2 3 4
D
αFlag IP
ENL-F:
WT
1-430
AFF4
ELL2-HA
CDK9
Flag
WB
E
αFlag IP
AF9-F:
WT
1-480
AFF4
ELL2-HA
CDK9
Flag
WB

Fig. 4-2 Separate regions of AFF4 are used to interact with different subunits of SEC. NEs derived from HeLa cells, which were transfected with cDNA constructs expressing either WT Flag-tagged AFF or the various deletion mutants as indicated, were subjected to anti-Flag IP. The immunoprecipitates were analyzed by Western blotting with the indicated antibodies. The diagram at the bottom summarizes the findings of the binding study.

Fig. 4-2

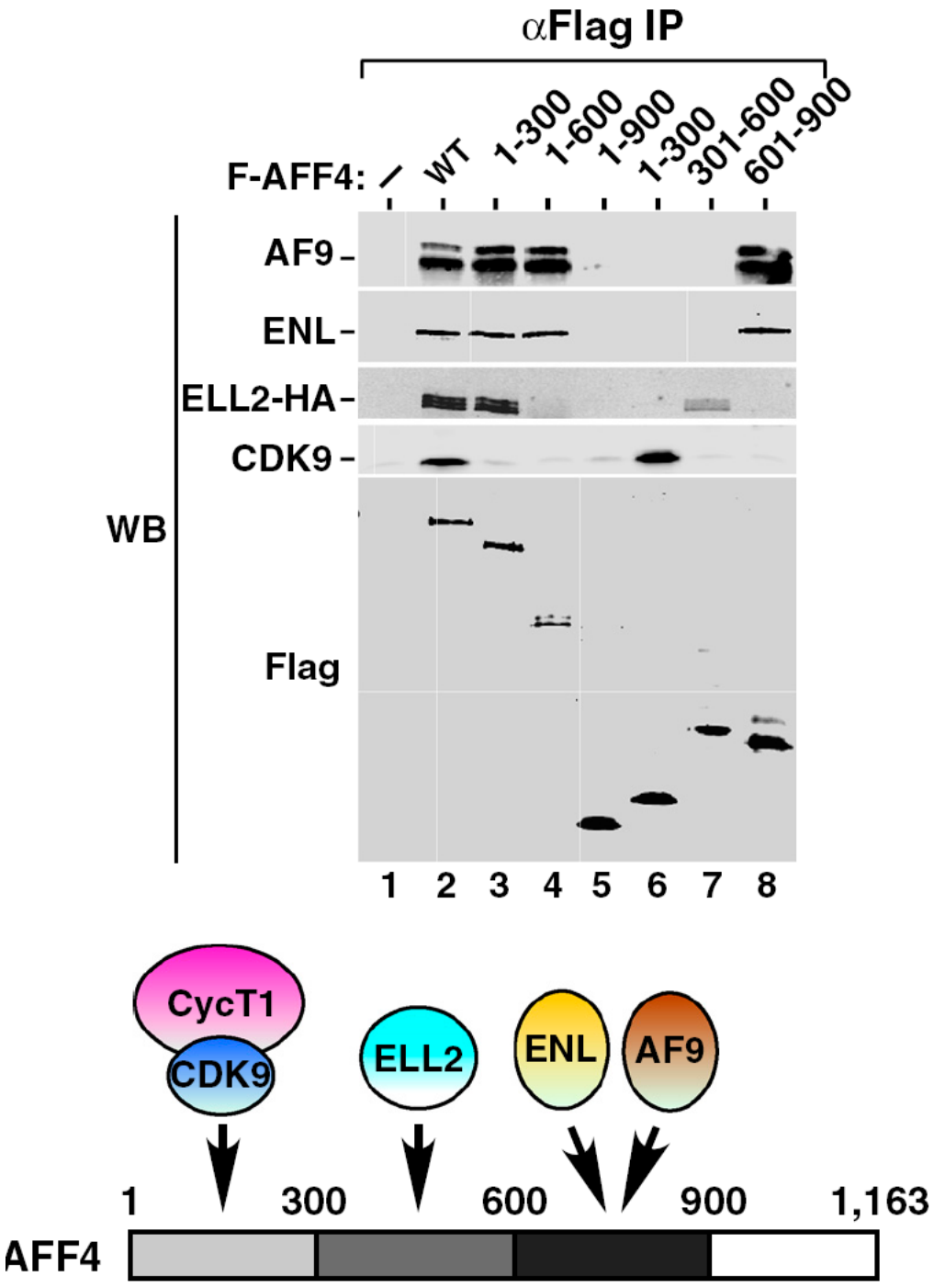

αFlag IP
F-AFF4: / WT 1-300 1-600 1-900 1-300 301-600 601-900
AF9
ENL
ELL2-HA
CDK9
WB
Flag
1 2 3 4 5 6 7 8
CycT1
CDK9
ELL2
ENL
AF9
1 300 600 900 1,163
AFF4

Fig. 4-3 ENL and AF9 do not exist in the same SEC complex and exert similar but non-identical functions in supporting SEC-dependent HIV-1 transcription. A. Anti-Flag immunoprecipitates were obtained from NE of cells transfected with the various cDNA constructs and analyzed by Western blotting with the indicated antibodies. B. HeLa NE were subjected to immunodepletion to remove ENL and the depleted NE were analyzed for the presence of the indicated proteins by Western blotting. C. In vitro binding reactions contained constant amounts of ENL-HA immobilized on anti-HA beads and F-AFF4 in solution. WT or C-terminally deleted AF9-F were either not added (-) or added (+) in 3-fold increments into the binding reactions. The bound and input proteins were examined by Western blotting as indicated. D. & E. NE from HeLa cells expressing either shENL (D) or shAF9 (E) were subjected to immunoprecipitation (IP) with anti-CDK9 or an irrelevant rabbit IgG. The isolated NE (left) and immunoprecipitates (right) were analyzed by Western blotting with the indicated antibodies. F. & G. HeLa cells containing an integrated HIV-1 LTR-luciferase reporter gene were transfected with the Tat cDNA and/or plasmids expressing shENL, shAF9, or an irrelevant control sequence (ctl.). Luciferase activities were measured in cell extracts. The basal LTR activity in F and Tat-activated LTR activity in G were artificially set to “1” and “100”, respectively, for easy comparison. The error bars represent mean +/- SD.

Fig. 4-3

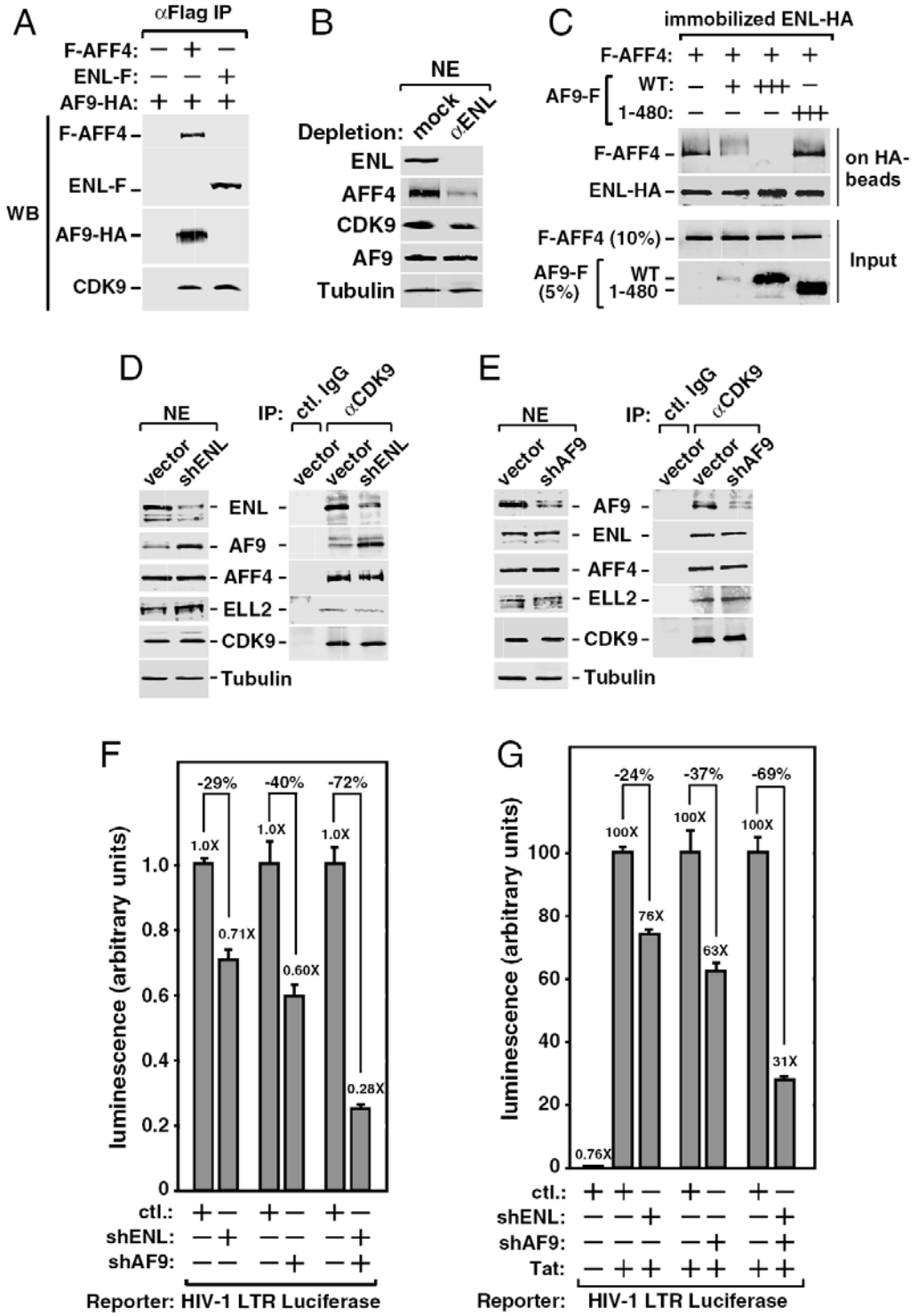

A
αFlag IP
F-AFF4: − + −
ENL-F: − − +
AF9-HA: + + +
WB
F-AFF4
ENL-F
AF9-HA
CDK9
B
NE
Depletion: mock αENL
ENL
AFF4
CDK9
AF9
Tubulin
C
immobilized ENL-HA
F-AFF4: + + + +
AF9-F WT: − + +++ −
1-480: − − − +++
F-AFF4
ENL-HA
on HA-beads
F-AFF4 (10%)
AF9-F (5%) WT
1-480
Input
D
NE
IP: ctl. IgG αCDK9
vector shENL vector vector shENL
ENL
AF9
AFF4
ELL2
CDK9
Tubulin
E
NE
IP: ctl. IgG αCDK9
vector shAF9 vector vector shAF9
AF9
ENL
AFF4
ELL2
CDK9
Tubulin
F
luminescence (arbitrary units)
-29%
-40%
-72%
1.0X
0.71X
0.60X
0.28X
ctl.: + − + − + −
shENL: − + − − − +
shAF9: − − − + − +
Reporter: HIV-1 LTR Luciferase
G
luminescence (arbitrary units)
-24%
-37%
-69%
100X
76X
63X
31X
0.76X
ctl.: + + − + − + −
shENL: − − + − − − +
shAF9: − − − − + − +
Tat: − + + + + + +
Reporter: HIV-1 LTR Luciferase

Fig. 4-4 Dot1L competes with AFF4 for binding to ENL/AF9 and does not exist in SEC. A. & B. HeLa cells were transfected with the indicated cDNA constructs. NE (left panels) and immunoprecipitates isolated from NE (right panels) with either anti-Flag (A) or anti-HA beads (B) were examined by Western blotting for the presence of the indicated proteins. C. & D. In vitro binding assay was performed in reactions containing HA-AFF4 immobilized on anti-HA beads, WT or the C-terminally truncated AF9 or ENL, and different amounts of Dot1L. The bound and input proteins were examined by Western blotting as indicated. E. HeLa cells containing a copy of integrated HIV-1 LTR-luciferase reporter construct were transfected with Dot1L-specific or control scrambled siRNA (ctl.). Six days post transfection, cells were transduced with a retrovirus encoding for Tat or empty vector. Luciferase was measured 24 hrs post transduction. Fold activation is the ratio between luciferase values obtained in the presence or absence of Tat for each siRNA. The graph represents mean and standard error obtained from three independent experiments. F. Knock down efficiency of siRNA treatment was measured by Western blot with specific antibodies as indicated.

Fig. 4-4

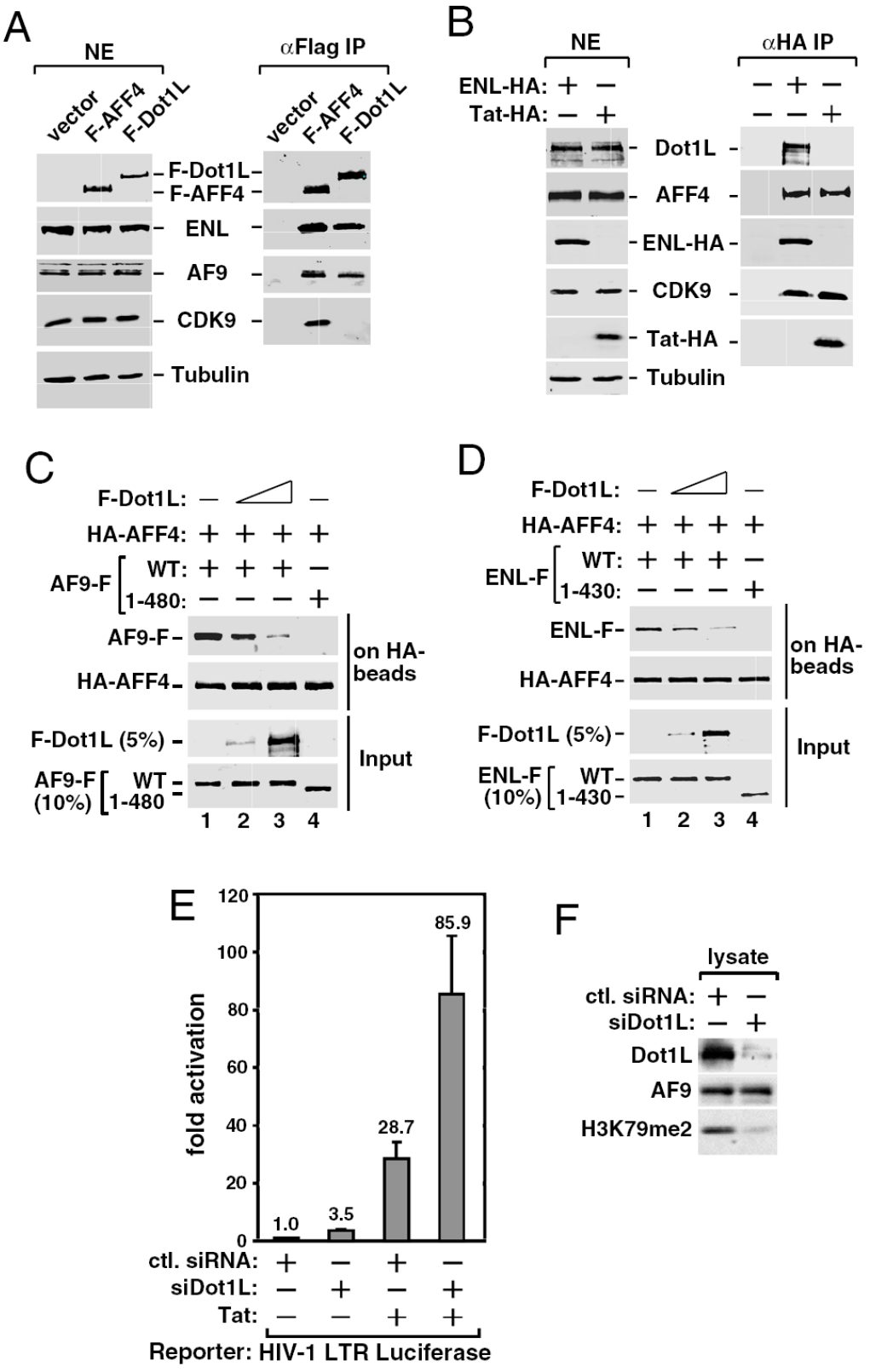

A
NE
αFlag IP
vector
F-AFF4
F-Dot1L
F-Dot1L
F-AFF4
ENL
AF9
CDK9
Tubulin
B
NE
αHA IP
ENL-HA:
Tat-HA:
Dot1L
AFF4
ENL-HA
CDK9
Tat-HA
Tubulin
C
F-Dot1L:
HA-AFF4:
AF9-F
WT:
1-480:
AF9-F
HA-AFF4
on HA-beads
F-Dot1L (5%)
AF9-F (10%)
WT
1-480
Input
1 2 3 4
D
F-Dot1L:
HA-AFF4:
ENL-F
WT:
1-430:
ENL-F
HA-AFF4
on HA-beads
F-Dot1L (5%)
ENL-F (10%)
WT
1-430
Input
1 2 3 4
E
fold activation
120
100
80
60
40
20
0
85.9
28.7
3.5
1.0
ctl. siRNA:
siDot1L:
Tat:
Reporter: HIV-1 LTR Luciferase
F
lysate
ctl. siRNA:
siDot1L:
Dot1L
AF9
H3K79me2

Fig. 4-5 The YEATS domains of ENL and AF9 are not required for SEC formation but essential for SEC-dependent HIV-1 transcription. A. HeLa cells containing an integrated HIV-1 LTR-luciferase reporter gene were transfected with the indicated expression constructs. Left panel: luciferase activities were measured in cell extracts, with the activity in cells transfected with an empty vector artificially set to "1". The error bars represent mean +/- SD. Right panels: Western analysis of the levels of WT and N-terminally deleted AF9-F and ENL-F in transfected cells. B. & C. HeLa cells were transfected with the indicated cDNA constructs. NE (left panels) and anti-Flag immunoprecipitates isolated from NE (right panels) were examined by Western blotting for the presence of the indicated proteins. A non-specific band is indicated by asterisks (*).

Fig. 4-5

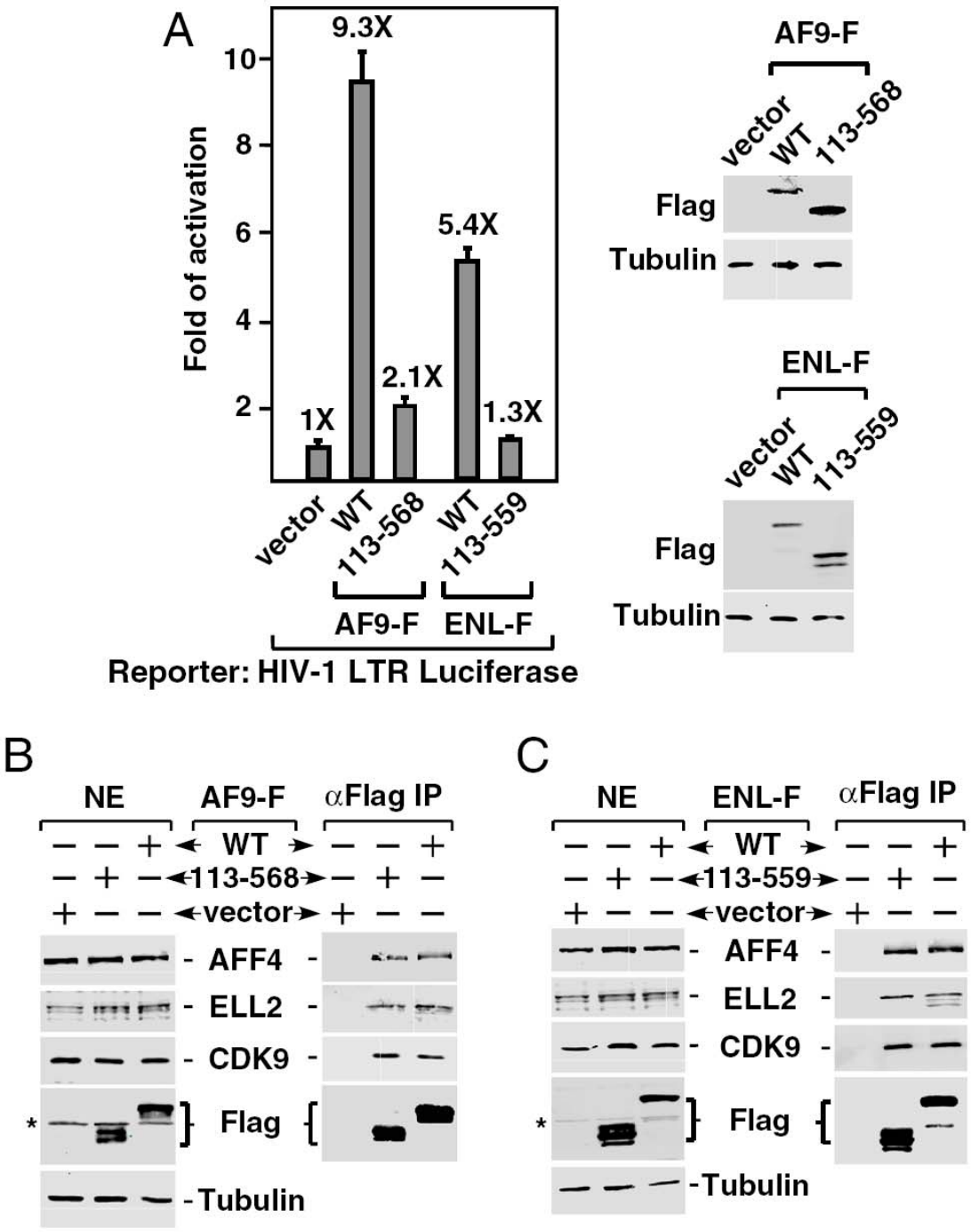

A
9.3X
5.4X
2.1X
1.3X
1X
Fold of activation
10
8
6
4
2
vector
WT
113-568
WT
113-559
AF9-F
ENL-F
Reporter: HIV-1 LTR Luciferase
AF9-F
vector
WT
113-568
Flag
Tubulin
ENL-F
vector
WT
113-559
Flag
Tubulin
B
NE
AF9-F
αFlag IP
WT
113-568
vector
AFF4
ELL2
CDK9
Flag
Tubulin
*
C
NE
ENL-F
αFlag IP
WT
113-559
vector
AFF4
ELL2
CDK9
Flag
Tubulin
*

Fig. 4-6 The ENL/AF9 YEATS domain interacts directly with PAF1 to target SEC to a chromatin template. A. & B. Chromatin immunoprecipitation (ChIP) with the anti-Flag antibody was performed in cells containing an integrated HIV-1 LTR-luciferase reporter gene and transiently expressing WT or N-terminally deleted AF9-F (A) or ENL-F (B). Three regions corresponding to the promoter, interior, and 3' UTR of the integrated reporter gene were qPCR-amplified from the precipitated and purified DNA and shown as percentages of the input chromatin. The error bars represent mean +/- SD. The levels of WT and mutant AF9-F or ENL-F in NE were examined by anti-Flag Western blotting in the right panels. C. & D. HeLa cells were transfected with the indicated ENL-F (C) or AF9-F-expressing constructs (D). NE (left panels) and anti-Flag immunoprecipitates isolated from NE (right panels) were examined by Western blotting for the presence of the indicated proteins. A non-specific band in D is indicated by an asterisk (*). E. The GST pull-down assay was performed with the indicated proteins present in the reactions. After extensive washing, the proteins bound to the GST beads were detected by silver staining. GST-ENL-N and GST-ENL-C contain amino acids 1-154 and 433-559 of ENL, respectively.

Fig. 4-6

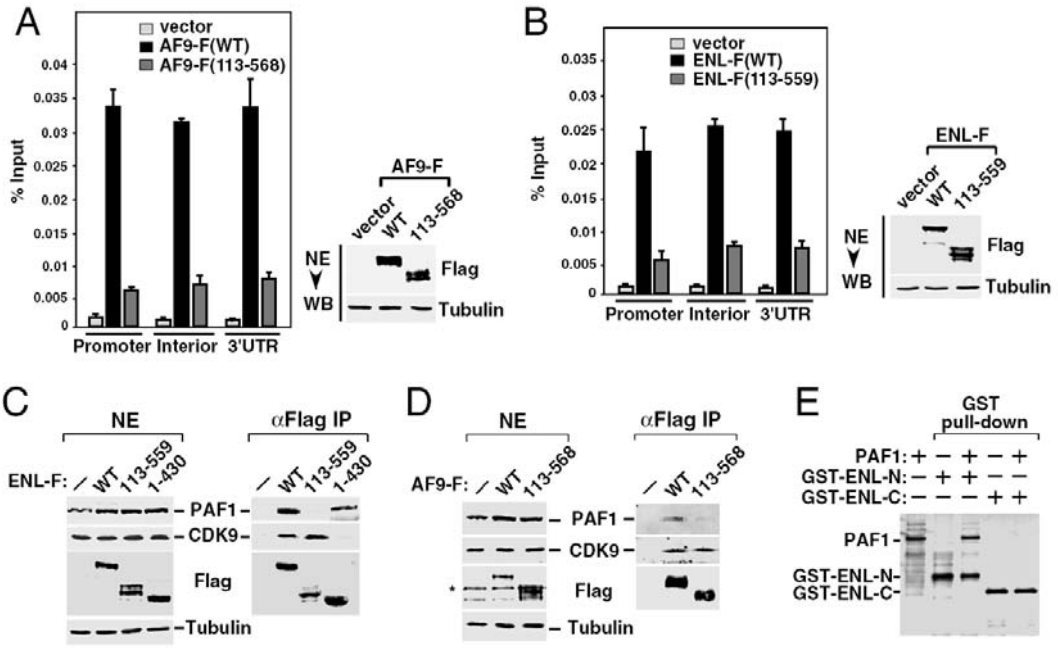
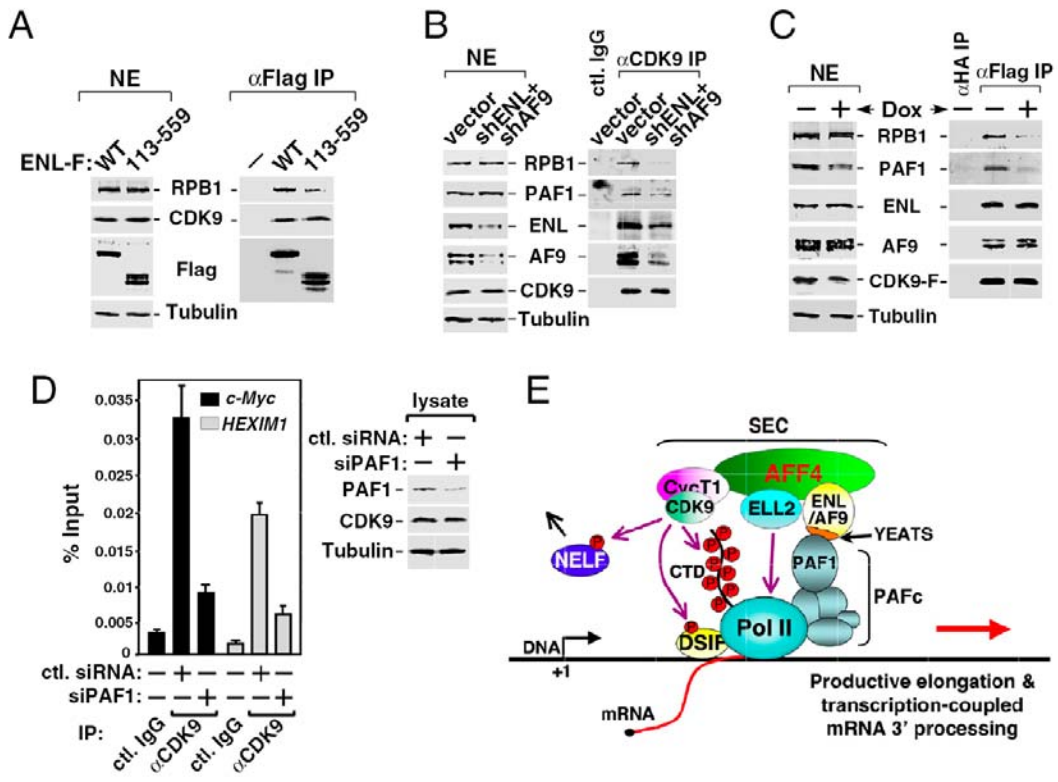

A
vector
AF9-F(WT)
AF9-F(113-568)
% Input
Promoter Interior 3'UTR
AF9-F
vector
WT
113-568
NE
WB
Flag
Tubulin
B
vector
ENL-F(WT)
ENL-F(113-559)
% Input
Promoter Interior 3'UTR
ENL-F
vector
WT
113-559
NE
WB
Flag
Tubulin
C
NE
αFlag IP
ENL-F:
WT
113-559
1-430
PAF1
CDK9
Flag
Tubulin
D
NE
αFlag IP
AF9-F:
WT
113-568
PAF1
CDK9
Flag
Tubulin
E
GST pull-down
PAF1:
GST-ENL-N:
GST-ENL-C:
PAF1
GST-ENL-N-
GST-ENL-C-

Fig. 4-7 PAFc connects SEC to Pol II. A. HeLa cells were transfected with the indicated ENL-F-expressing constructs. NE (left panels) and anti-Flag immunoprecipitates isolated from NE (right panels) were examined by Western blotting for the presence of the indicated proteins. B. & C. NE from HeLa cells either containing an empty vector or expressing the indicated shRNAs (B) or from the inducible shPAF1-expressing cells treated with (+) or without (-) doxycycline (Dox) to induce shPAF1 expression (C) were subjected to immunoprecipitation (IP) with the indicated antibodies. The isolated NE (left) and immunoprecipitates (right) were analyzed by Western blotting with the indicated antibodies. D. The ChIP assay was performed in HeLa cells with either the anti-CDK9 or an irrelevant control antibody. Interior regions of the c-Myc and HEXIM1 gene were amplified by qPCR from the precipitated and purified DNA and shown as percentages of the input chromatin. The error bars represent mean +/- SD. E. A model showing the recruitment of the SEC complex, which contains either ENL or AF9 and is assembled around the scaffolding protein AFF4, to the elongating Pol II through the interaction of the ENL/AF9 YEATS domain with the PAF1 subunit of PAFc. This configuration allows SEC to use its P-TEFb and ELL2 functional modules to exert a multitude of effects that include the phosphorylation of the Pol II CTD and elongation factors DSIF and NELF (the latter is released upon phosphorylation) by CDK9 and the suppression of Pol II pausing by ELL2. These events synergistically activate productive elongation and likely also transcription-coupled mRNA 3’ processing.

Fig. 4-7


A
NE
αFlag IP
ENL-F: WT 113-559 – WT 113-559
RPB1
CDK9
Flag
Tubulin
B
NE
ctl. IgG
αCDK9 IP
vector shENL+ shAF9 vector shENL+ shAF9
RPB1
PAF1
ENL
AF9
CDK9
Tubulin
C
NE
αHA IP
αFlag IP
– + Dox – – +
RPB1
PAF1
ENL
AF9
CDK9-F
Tubulin
D
c-Myc
HEXIM1
% Input
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
ctl. siRNA: – + – – + –
siPAF1: – – + – – +
IP: ctl. IgG αCDK9 ctl. IgG αCDK9
lysate
ctl. siRNA: + –
siPAF1: – +
PAF1
CDK9
Tubulin
E
SEC
AFF4
CycT1
CDK9
ELL2
ENL /AF9
YEATS
NELF
CTD
PAF1
PAFc
DSIF
Pol II
DNA
+1
mRNA
Productive elongation & transcription-coupled mRNA 3' processing

Chapter 5

Conclusions and Perspectives

Conclusions

P-TEFb, positive transcription elongation factor b, is composed of CDK9 and its regulatory partner cyclin T1 (CycT1; or the minor form T2 or K) (10). Through phosphorylating its primary substrates, the C-terminal domain (CTD) of the largest subunit of RNA polymerase II and the negative transcription elongation factor, N-TEF, P-TEFb releases the paused Pol II at promoter-proximal region and triggers productive elongation to generate full-length mRNA transcripts (10, 11). Studies suggest P-TEFb activity is tightly controlled in the cells through its association with different co-factors. A large cellular pool of P-TEFb exists in a P-TEFb/7SK snRNP complex, where its enzymatic activity is suppressed by the concerted effect of the other components in the complex, including 7SK snRNA, HEXIM1 (16-18). Since none of the previous known components in the 7SK snRNP is a RNA binding protein, the question here is how the 7SK snRNA is protected, which is normally mediated by RNA binding protein(s). Identification of PIP7S here provides insights into how the protection of 7SK is achieved. I have shown that the PIP7S is a La-related protein and intimately associated with all the nuclear 7SK and required for 7SK stability and 7SK snRNP integrity. This requires its La domain and C-terminus. The latter is frequently deleted in human tumors due to microsatellite instability-associated mutations. Consistent with the tumor suppressor role of a Drosophila homolog of PIP7S, loss of PIP7S function shifts the P-TEFb equilibrium toward the active state, disrupts epithelial differentiation and causes P-TEFb-dependent malignant transformation. Through PIP7S modulation of P-TEFb, our data link a general elongation factor to growth control and tumorigenesis.

On the other hand, unlike the 7SK snRNP, the bromo domain-containing protein Brd4 associated P-TEFb complex has been considered as a positive P-TEFb population (27, 28). However, emerging evidence has been presented to show that the recruitments of P-TEFb by Tat and Brd4 are two mutually exclusive events that cannot occur at the same time (28, 31), which necessitating investigation of what P-TEFb complex is associated with Tat for HIV transcription. To address this question, a tandem affinity-purification approach employing anti-Flag and then anti-HA beads was used to purify the complex that contains both HA-tagged Tat and Flag-tagged CDK9 (95). Although numerous cellular factors that can bind to either Tat or P-TEFb separately had been identified in the past, this was the first attempt to specifically isolate factors that are integral components of a complex(es) that contains both proteins. Analyses of the purified materials by mass spectrometry reveal that in addition to Tat, CDK9 and CycT1, the Tat-P-TEFb complex also contains ELL2, AFF4, ENL and AF9 (95). Importantly, the binding of these factors to Tat and P-TEFb has also been independently confirmed by Sobhian et al. (86). Through directly isolating the Tat-associated proteins, the latter study has also identified several additional proteins (86). However, CDK9, CycT1, ELL2, AFF4, ENL and AF9 are likely the core subunits of a single multi-component complex now called the Super Elongation Complex (86, 95).

The association of Tat with SEC has been shown to serve two complementary purposes. First, Tat can recruit SEC that contains at least two well-established

elongation factors of different classes to the viral LTR. This allows P-TEFb and ELL2 to act simultaneously on the same polymerase enzyme in a cooperative manner, which is most likely responsible for the powerful elongation activity attributed to Tat. Secondly, Tat has been shown to markedly promote SEC formation. ELL2 turns out to be a short-lived protein that is rapidly degraded by the proteasome (95). However, its stability can be greatly enhanced by the presence of Tat in a process that likely requires the CDK9 kinase activity (95). Tat-mediated ELL2 stabilization and accumulation in the cell allows more ELL2 to be sequestered into SEC, thus promoting SEC formation and SEC-dependent HIV-1 transcription.

ENL and AF9, the two highly homologous core subunits of SEC, in fact exist in separate SECs that display similar but non-identical functions. In the absence of sequence-specific recruitment factors such as Tat and MLL, the evolutionarily conserved YEATS domain of ENL/AF9 targets SEC to chromatin by contacting the Polymerase-Associated Factor complex (PAFc), and through PAFc, the paused Pol II. This explains why this domain is dispensable for leukemogenesis when ENL/AF9 is translocated to MLL (96, 102), whose DNA-binding activity likely substitutes for the chromatin-targeting function of the YEATS domain. Finally, contrary to popular belief, the histone lysine79 (H3K79) methyltransferase Dot1L, which is a well-known binding partner of ENL and AF9, competes with AFF4 for binding to ENL/AF9 and thus does not reside in SEC and also is unnecessary for SEC function.

Perspectives

Human P-TEFb was first identified in 1997 as a specific host cellular cofactor for Tat activation of HIV-1 transcription (8, 9). Since then, this landmark discovery has provided the basic framework for our understanding of Tat function in the HIV-1 life cycle. Recently, the conventional view of HIV-1 gene expression control has been significantly expanded by the demonstrations that Tat captures P-TEFb from the 7SK snRNP, the major P-TEFb reservoir in the nucleus, promotes the formation of a novel P-TEFb complex termed SEC that also contains elongation factor ELL2 and several other transcription factors/co-factors, and delivers SEC to the paused RNAPII on the viral LTR (86, 95). The ability to allow P-TEFb and ELL2, representatives of two different classes of elongation factors, to act on the same polymerase enzyme explains why Tat is such a powerful transcriptional activator. These findings have not only provided fresh mechanistic insights into the control of HIV-1 gene expression but also revealed new targets for the development of improved anti-viral treatments as well as more specific and efficient therapeutic strategies to eradicate the latent HIV-1 reservoirs.

Despite these progresses, there are still a number of outstanding questions that remain to be answered. First, the structures and functions of both 7SK snRNP and SEC await further characterization, which will enable us to better understand how and where Tat captures P-TEFb from 7SK snRNP and also the precise mechanism by which SEC stimulates HIV-1 and cellular transcriptional elongation. Furthermore, the relationship among several known P-TEFb-containing complexes requires further clarification. For example, it is known that the recruitment of P-TEFb by Brd4 occurs

predominantly at the promoter region and is important for general transcriptional elongation (27, 28). Once recruited by Brd4, it is unclear how P-TEFb is eventually converted to SEC that is believed to be the form of P-TEFb actually engaged in elongation. In addition, to determine how SEC contributes to transcriptional elongation in general, it is imperative to perform genome-wide analyses of the function and distribution of SEC under both normal and disease conditions. Finally, the investigation into the control of SEC formation and stability may hold the key to the development of effective strategies to reactivate latent HIV-1 and suppress the progression of aggressive acute leukemias.

We provided evidence to show that ENL and AF9 could recruit SEC to RNA Pol II through PAFc (115), however, unlike our model, Conway and her colleagues have recently found that the Mediator subunit MED26 can also bridge the interaction between SEC and RNA Pol II (116). This actually raises an interesting question, which is how SEC is exactly recruited to RNA Pol II. These two models seem to differ from each other at first glance, but they are not conflicting with each other at all. MED26 was shown to directly interact with EAF1/2, the ELL1/2 associated factors, therefore bringing the SEC close to RNA Pol II and hence facilitating transition of Pol II into elongation stage of transcription (116). Since SEC is a big protein complex containing many subunits, both of the ENL/AF9-PAFc and ELL1/2-EAF1/2-MED26 axes are likely needed to ensure the proper and firm hookup between SEC and RNA Pol II during transcription. Indeed, depletion of either PAF1, a critical subunit of the PAFc, or MED26, greatly decreased the occupancy of SEC on *c-myc* gene (115, 116), consistent with the idea that both PAFc and MED26 play important roles in recruiting SEC to RNA Pol II. It is still unclear whether this is true for other SEC regulated genes and further genome-wide analysis should provide insights into this issue.

Work Cited

1. Core LJ & Lis JT (2008) Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319(5871):1791-1792.
2. Kuras L & Struhl K (1999) Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* 399(6736):609-613.
3. Saunders A, Core LJ, & Lis JT (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7(8):557-567.
4. Ptashne M (2005) Regulation of transcription: from lambda to eukaryotes. *Trends Biochem Sci* 30(6):275-279.
5. Guenther MG, Levine SS, Boyer LA, Jaenisch R, & Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130(1):77-88.
6. Muse GW, *et al.* (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39(12):1507-1511.
7. Kao SY, Calman AF, Luciw PA, & Peterlin BM (1987) Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330(6147):489-493.
8. Mancebo HS, *et al.* (1997) P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev* 11(20):2633-2644.
9. Wei P, Garber ME, Fang SM, Fischer WH, & Jones KA (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(4):451-462.
10. Peterlin BM & Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23(3):297-305.
11. Zhou Q & Yik JH (2006) The Yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation. *Microbiol Mol Biol Rev* 70(3):646-659.
12. Chao SH & Price DH (2001) Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J Biol Chem* 276(34):31793-31799.
13. Rahl PB, *et al.* (2010) c-Myc regulates transcriptional pause release. *Cell* 141(3):432-445.
14. He N, *et al.* (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 29(5):588-599.
15. Jeronimo C, *et al.* (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27(2):262-274.
16. Nguyen VT, Kiss T, Michels AA, & Bensaude O (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* 414(6861):322-325.
17. Yang Z, Zhu Q, Luo K, & Zhou Q (2001) The 7SK small nuclear RNA inhibits

the CDK9/cyclin T1 kinase to control transcription. *Nature* 414(6861):317-322.

18. Yik JH*, et al.* (2003) Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* 12(4):971-982.
19. Xue Y, Yang Z, Chen R, & Zhou Q (2010) A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res* 38(2):360-369.
20. Michels AA*, et al.* (2004) Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *Embo J* 23(13):2608-2619.
21. Huang F, Wagner M, & Siddiqui MA (2004) Ablation of the CLP-1 gene leads to down-regulation of the HAND1 gene and abnormality of the left ventricle of the heart and fetal death. *Mech Dev* 121(6):559-572.
22. Sano M*, et al.* (2002) Activation and function of cyclin T-Cdk9 (positive transcription elongation factor-b) in cardiac muscle-cell hypertrophy. *Nat Med* 8(11):1310-1317.
23. Natarajan M, August A, & Henderson AJ (2010) Combinatorial signals from CD28 differentially regulate human immunodeficiency virus transcription in T cells. *J Biol Chem* 285(23):17338-17347.
24. Biglione S*, et al.* (2007) Inhibition of HIV-1 replication by P-TEFb inhibitors DRB, seliciclib and flavopiridol correlates with release of free P-TEFb from the large, inactive form of the complex. *Retrovirology* 4:47.
25. Chen R, Yang Z, & Zhou Q (2004) Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. *J Biol Chem* 279(6):4153-4160.
26. He N, Pezda AC, & Zhou Q (2006) Modulation of a P-TEFb functional equilibrium for the global control of cell growth and differentiation. *Mol Cell Biol* 26(19):7068-7076.
27. Jang MK*, et al.* (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19(4):523-534.
28. Yang Z*, et al.* (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* 19(4):535-545.
29. Dey A*, et al.* (2000) A bromodomain protein, MCAP, associates with mitotic chromosomes and affects G(2)-to-M transition. *Mol Cell Biol* 20(17):6537-6549.
30. Jeanmougin F, Wurtz JM, Le Douarin B, Chambon P, & Losson R (1997) The bromodomain revisited. *Trends Biochem Sci* 22(5):151-153.
31. Bisgrove DA, Mahmoudi T, Henklein P, & Verdin E (2007) Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. *Proc Natl Acad Sci U S A* 104(34):13690-13695.
32. Dey A, Chitsaz F, Abbasi A, Misteli T, & Ozato K (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A* 100(15):8758-8763.

33. Yang Z, He N, & Zhou Q (2008) Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression. *Mol Cell Biol* 28(3):967-976.
34. Mochizuki K*, et al.* (2008) The bromodomain protein Brd4 stimulates G1 gene transcription and promotes progression to S phase. *J Biol Chem* 283(14):9040-9048.
35. Loyola A & Almouzni G (2004) Bromodomains in living cells participate in deciphering the histone code. *Trends Cell Biol* 14(6):279-281.
36. Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, & Peterlin BM (2001) NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell* 8(2):327-337.
37. Eberhardy SR & Farnham PJ (2001) c-Myc mediates activation of the cad promoter via a post-RNA polymerase II recruitment mechanism. *J Biol Chem* 276(51):48562-48571.
38. Kanazawa S, Okamoto T, & Peterlin BM (2000) Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. *Immunity* 12(1):61-70.
39. Kanazawa S, Soucek L, Evan G, Okamoto T, & Peterlin BM (2003) c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. *Oncogene* 22(36):5707-5711.
40. Barboric M*, et al.* (2007) Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res* 35(6):2003-2012.
41. Schulte A*, et al.* (2005) Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. *J Biol Chem* 280(26):24968-24977.
42. Sedore SC*, et al.* (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 Res* 35(13):4347-4358.
43. Krueger BJ, Varzavand K, Cooper JJ, & Price DH (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(8):e12335.
44. Muniz L, Egloff S, Ughy B, Jady BE, & Kiss T (2010) Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. *PLoS Pathog* 6(10):e1001152.
45. Chen R*, et al.* (2008) PP2B and PP1alpha cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca2+ signaling. *Genes Dev* 22(10):1356-1368.
46. Ammosova T*, et al.* (2005) Nuclear targeting of protein phosphatase-1 by HIV-1 Tat protein. *J Biol Chem* 280(43):36364-36371.
47. Ammosova T*, et al.* (2011) Expression of a protein phosphatase 1 inhibitor, cdNIPP1, increases CDK9 threonine 186 phosphorylation and inhibits HIV-1 transcription. *J Biol Chem* 286(5):3798-3804.
48. D'Orso I & Frankel AD (2010) RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation. *Nat Struct Mol Biol*

17(7):815-821.

49. Shim EY, Walker AK, Shi Y, & Blackwell TK (2002) CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the C. elegans embryo. *Genes Dev* 16(16):2135-2146.
50. Wassarman DA & Steitz JA (1991) Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol Cell Biol* 11(7):3432-3445.
51. Contreras X, Lenasi T, & Peterlin BM (2006) HIV latency: present knowledge, future directions. *Future Virol* 1(6):733-745.
52. Turano M, *et al.* (2006) Increased HEXIM1 expression during erythroleukemia and neuroblastoma cell differentiation. *J Cell Physiol* 206(3):603-610.
53. Yik JH, Chen R, Pezda AC, Samford CS, & Zhou Q (2004) A human immunodeficiency virus type 1 Tat-like arginine-rich RNA-binding domain is essential for HEXIM1 to inhibit RNA polymerase II transcription through 7SK snRNA-mediated inactivation of P-TEFb. *Mol Cell Biol* 24(12):5094-5105.
54. Ventura A, *et al.* (2004) Cre-lox-regulated conditional RNA interference from transgenes. *Proc Natl Acad Sci U S A* 101(28):10380-10385.
55. Debnath J, Muthuswamy SK, & Brugge JS (2003) Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* 30(3):256-268.
56. Intine RV, *et al.* (2000) Control of transfer RNA maturation by phosphorylation of the human La antigen on serine 366. *Mol Cell* 6(2):339-348.
57. Remillieux-Leschelle N, Santamaria P, & Randsholt NB (2002) Regulation of larval hematopoiesis in Drosophila melanogaster: a role for the multi sex combs gene. *Genetics* 162(3):1259-1274.
58. Morillon A, Karabetsou N, Nair A, & Mellor J (2005) Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Mol Cell* 18(6):723-734.
59. Michels AA, *et al.* (2003) MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. *Mol Cell Biol* 23(14):4859-4869.
60. Keene JD (1996) RNA recognition by autoantigens and autoantibodies. *Mol Biol Rep* 23(3-4):173-181.
61. Maraia RJ & Bayfield MA (2006) The La protein-RNA complex surfaces. *Mol Cell* 21(2):149-152.
62. Wolin SL & Cedervall T (2002) The La protein. *Annu Rev Biochem* 71:375-403.
63. Teplova M, *et al.* (2006) Structural basis for recognition and sequestration of UUU(OH) 3' temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* 21(1):75-85.
64. Pannone BK, Xue D, & Wolin SL (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for

RNA polymerase III transcripts. *Embo J* 17(24):7442-7453.

65. Xue D, Rubinson DA, Pannone BK, Yoo CJ, & Wolin SL (2000) U snRNP assembly in yeast involves the La protein. *Embo J* 19(7):1650-1660.
66. Shaw KR, Wrobel CN, & Brugge JS (2004) Use of three-dimensional basement membrane cultures to model oncogene-induced changes in mammary epithelial morphogenesis. *J Mammary Gland Biol Neoplasia* 9(4):297-310.
67. Marshall NF, Peng J, Xie Z, & Price DH (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* 271(43):27176-27183.
68. Zhu Y, *et al.* (1997) Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev* 11(20):2622-2632.
69. Mehta K, Fok J, Miller FR, Koul D, & Sahin AA (2004) Prognostic significance of tissue transglutaminase in drug resistant and metastatic breast cancer. *Clin Cancer Res* 10(23):8068-8076.
70. Torricelli C, *et al.* (2006) Role of PTHrp and PTHrp-engaged pathways in MCF-7 cells migration/invasion. *Matrix Biol* 25(2):104-111.
71. Li Q, *et al.* (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. *J Biol Chem* 280(31):28819-28826.
72. Yik JH, Chen R, Pezda AC, & Zhou Q (2005) Compensatory contributions of HEXIM1 and HEXIM2 in maintaining the balance of active and inactive positive transcription elongation factor b complexes for control of transcription. *J Biol Chem* 280(16):16368-16376.
73. Reina JH, Azzouz TN, & Hernandez N (2006) Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS One* 1:e134.
74. White RJ (2004) RNA polymerase III transcription and cancer. *Oncogene* 23(18):3208-3216.
75. Haaland RE, Herrmann CH, & Rice AP (2005) siRNA depletion of 7SK snRNA induces apoptosis but does not affect expression of the HIV-1 LTR or P-TEFb-dependent cellular genes. *J Cell Physiol* 205(3):463-470.
76. Sims RJ, 3rd, Belotserkovskaya R, & Reinberg D (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev* 18(20):2437-2468.
77. Shilatifard A, *et al.* (1997) ELL2, a new member of an ELL family of RNA polymerase II elongation factors. *Proc Natl Acad Sci U S A* 94(8):3639-3643.
78. Estable MC, *et al.* (2002) MCEF, the newest member of the AF4 family of transcription factors involved in leukemia, is a positive transcription elongation factor-b-associated protein. *J Biomed Sci* 9(3):234-245.
79. Thirman MJ, Levitan DA, Kobayashi H, Simon MC, & Rowley JD (1994) Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia. *Proc Natl Acad Sci U S A* 91(25):12110-12114.
80. Shilatifard A, Lane WS, Jackson KW, Conaway RC, & Conaway JW (1996) An RNA polymerase II elongation factor encoded by the human ELL gene.

Science 271(5257):1873-1876.

81. Taki T, *et al.* (1999) AF17q25, a putative septin family gene, fuses the MLL gene in acute myeloid leukemia with t(11;17)(q23;q25). *Cancer Res* 59(17):4261-4265.
82. Harper DP & Aplan PD (2008) Chromosomal rearrangements leading to MLL gene fusions: clinical and biological aspects. *Cancer Res* 68(24):10024-10027.
83. Bitoun E, Oliver PL, & Davies KE (2007) The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum Mol Genet* 16(1):92-106.
84. Mueller D, *et al.* (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* 110(13):4445-4454.
85. Mueller D, *et al.* (2009) Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS Biol* 7(11):e1000249.
86. Sobhian B, *et al.* (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell* 38(3):439-451.
87. Zhou Q, Chen D, Pierstorff E, & Luo K (1998) Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *Embo J* 17(13):3681-3691.
88. Zhou Q & Sharp PA (1995) Novel mechanism and factor for regulation by HIV-1 Tat. *Embo J* 14(2):321-328.
89. De Falco G, *et al.* (2000) Physical interaction between CDK9 and B-Myb results in suppression of B-Myb gene autoregulation. *Oncogene* 19(3):373-379.
90. Garber ME, *et al.* (1998) The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev* 12(22):3512-3527.
91. Martincic K, Alkan SA, Cheatle A, Borghesi L, & Milcarek C (2009) Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nat Immunol* 10(10):1102-1109.
92. Kornblihtt AR, de la Mata M, Fededa JP, Munoz MJ, & Nogues G (2004) Multiple links between transcription and splicing. *Rna* 10(10):1489-1498.
93. Shell SA, Martincic K, Tran J, & Milcarek C (2007) Increased phosphorylation of the carboxyl-terminal domain of RNA polymerase II and loading of polyadenylation and cotranscriptional factors contribute to regulation of the ig heavy chain mRNA in plasma cells. *J Immunol* 179(11):7663-7673.
94. Eissenberg JC, Shilatifard A, Dorokhov N, & Michener DE (2007) Cdk9 is an essential kinase in Drosophila that is required for heat shock gene expression, histone methylation and elongation factor recruitment. *Mol Genet Genomics* 277(2):101-114.
95. Chen R, Liu M, Zhang K, & Zhou Q (2010) Isolation and functional

characterization of P-TEFb-associated factors that control general and HIV-1 transcriptional elongation. *Methods*.

96. Yokoyama A, Lin M, Naresh A, Kitabayashi I, & Cleary ML (2010) A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell* 17(2):198-212.
97. Zhang W, Xia X, Reisenauer MR, Hemenway CS, & Kone BC (2006) Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCalpha in an aldosterone-sensitive manner. *J Biol Chem* 281(26):18059-18068.
98. Feng Q, *et al.* (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* 12(12):1052-1058.
99. Jones B, *et al.* (2008) The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet* 4(9):e1000190.
100. Schubeler D, *et al.* (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18(11):1263-1271.
101. Fromm G & Adelman K (2010) Taking MLL through the MudPIT: identification of novel complexes that bring together MLL-fusion proteins and transcription elongation factors. *Mol Cell* 37(4):449-450.
102. Slany RK, Lavau C, & Cleary ML (1998) The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol Cell Biol* 18(1):122-129.
103. Kim J, Guermah M, & Roeder RG (2010) The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell* 140(4):491-503.
104. Shi X, *et al.* (1996) Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. *Mol Cell Biol* 16(2):669-676.
105. Muntean AG, *et al.* (2010) The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* 17(6):609-621.
106. Chen Y, *et al.* (2009) DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev* 23(23):2765-2777.
107. Boyer LA, *et al.* (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122(6):947-956.
108. Barski A, *et al.* (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129(4):823-837.
109. Sawado T, *et al.* (2008) H3 K79 dimethylation marks developmental activation of the beta-globin gene but is reduced upon LCR-mediated high-level transcription. *Blood* 112(2):406-414.
110. Schulze JM, Wang AY, & Kobor MS (2009) YEATS domain proteins: a diverse family with many links to chromatin modification and transcription.

Biochem Cell Biol 87(1):65-75.

111. Okada Y, *et al.* (2005) hDOT1L links histone methylation to leukemogenesis. *Cell* 121(2):167-178.
112. Zeisig DT, *et al.* (2005) The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene* 24(35):5525-5532.
113. Nagaike T, *et al.* (2011) Transcriptional activators enhance polyadenylation of mRNA precursors. *Mol Cell* 41(4):409-418.
114. Ahn SH, Kim M, & Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* 13(1):67-76.
115. He N, *et al.* (2011) Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proc Natl Acad Sci U S A* 108(36):E636-645.
116. Takahashi H, *et al.* (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* 146(1):92-104.