Title

HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold

Journal

Proceedings of the National Academy of Sciences of the United States of America, 110(2)

ISSN

0027-8424

Authors

Chou, Seemay

Upton, Heather

Bao, Katherine

et al.

Publication Date

2013-01-08

DOI

10.1073/pnas.1216971110

Peer reviewed

HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold

Seemay Chou[a,1], Heather Upton[a], Katherine Bao[a,2], Ursula Schulze-Gahmen[a], Avi J. Samelson[a], Nanhai He[a,3], Anna Nowak[a], Huasong Lu[a], Nevan J. Krogan[b,c,d], Qiang Zhou[a], and Tom Alber[a,e,4]

[a]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720; [b]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158; [c]QB3, California Institute for Quantitative Biosciences, San Francisco, CA 94158; [d]J. David Gladstone Institutes, San Francisco, CA 94158; and [e]QB3, California Institute for Quantitative Biosciences, Berkeley, CA 94720

Edited by Bryan R. Cullen, Duke University, Durham, NC, and accepted by the Editorial Board November 16, 2012 (received for review October 2, 2012)

The HIV-1 Tat protein stimulates viral gene expression by recruiting human transcription elongation complexes containing P-TEFb, AFF4, ELL2, and ENL or AF9 to the viral promoter, but the molecular organization of these complexes remains unknown. To establish the overall architecture of the HIV-1 Tat elongation complex, we mapped the binding sites that mediate complex assembly in vitro and in vivo. The AFF4 protein emerges as the central scaffold that recruits other factors through direct interactions with short hydrophobic regions along its structurally disordered axis. Direct binding partners CycT1, ELL2, and ENL or AF9 act as bridging components that link this complex to two major elongation factors, P-TEFb and the PAF complex. The unique scaffolding properties of AFF4 allow dynamic and flexible assembly of multiple elongation factors and connect the components not only to each other but also to a larger network of transcriptional regulators.

paused RNA polymerase II | intrinsically disordered proteins | super elongation complex | MLL-fusion complex

RNA polymerase II (Pol II) activity is tightly regulated throughout the steps of eukaryotic transcription. Each stage—initiation, clearance from the promoter, elongation, and termination—is licensed by specific factors that serve as checkpoints (1–4). Pol II transcription downstream of the promoter also involves intricate crosstalk between elongation and posttranscriptional events, such as splicing (5–7). Differential phosphorylation of the Pol II C-terminal domain (CTD) during transcription allows preferential binding of stage-specific regulators (8). Historically, focus has been placed on the control of transcription initiation, but mounting evidence suggests that elongation is the rate-limiting step for many highly expressed genes during cell growth and differentiation (2, 9–11).

A major regulator of transcription elongation is the positive transcription elongation factor b (P-TEFb). Comprising a heterodimer of the CDK9 kinase and cyclin T1 (CycT1), P-TEFb phosphorylates Ser2 of the CTD heptad repeat, YSPTSPS. Disruption of this activity inhibits elongation (8). By also phosphorylating negative elongation factor (NELF) and DRG sensitivity-inducing factor (DSIF) (12–14), two factors that block Pol II, P-TEFb acts as a gatekeeper for the escape of paused Pol II into elongation. Because promoter escape and efficient elongation are important for transcription of the HIV genome, HIV infection is hypersensitive to elongation defects. The HIV-1 Tat protein recruits active P-TEFb to the HIV promoter by binding both the CycT1 subunit and the transactivation response (TAR) element in the nascent HIV mRNA. With Zn^{2+} in the protein interface, Tat folds onto P-TEFb and anchors recognition of TAR. This bridging linkage by Tat highlights the central roles of P-TEFb in promoting not only transcriptional elongation but also HIV pathogenesis (15–17).

Several other classes of human Pol II-associated elongation factors have been identified, including transcription factor S-II (TFIIS), the eleven-nineteen lysine-rich leukemia protein (ELL), and the polymerase-associated factor complex (PAFc) (2, 4, 18–20). The ELL family (ELL1–3) interacts with Pol II and directly enhances its catalytic rate (21). PAFc, which is required for H2B ubiquitylation and H3K4 and H3K79 methylation, also interacts with Pol II and stimulates elongation on a chromatin template (20). Although some of these factors, such as ELL and PAFc, can enhance transcription elongation directly in vitro, the functions of other factors implicated in transcriptional elongation remain elusive. Genetic and biochemical studies indicate that many factors function in large complexes or in conjunction with RNA-processing activities to stimulate transcription in vivo (22). For example, PAFc interacts with SII/TFIIS (20) and ENL/AF9 both physically and functionally to stimulate elongation (23), highlighting the importance of coordination and cooperativity among different transcriptional regulators.

Recent proteomics studies revealed that HIV-1 Tat recruits P-TEFb not as an isolated heterodimer but as part of a large, stoichiometric complex containing additional transcription elongation factors. Tat-P-TEFb partners include ELL2 and the mixed-lineage leukemia (MLL) fusion partners, AFF4 and the homologs ENL and AF9 (24–28). This complex belongs to a family of assemblies that have been named "super elongation complexes" (SECs) (27) to reflect roles in normal transcription and "MLL-fusion complexes" (29) because of the activities of certain subunit chimeras in promoting myeloid leukemias. The SECs form a combinatorial family of related assemblies containing homologous subunits and also normally are recruited to a subset of human genes occupied by paused polymerases (30, 31). In embryonic stem cells, several SEC components reside at actively transcribed genes and are required for stimulating transcription during differentiation (31), pointing to a wider, essential role for these factors in regulating transcriptional elongation. mRNA-knockdown experiments suggest some redundancy among members of the AFF and ELL families, but these factors also may have specialized functions (18, 30). For example, HIV transcription is stimulated specifically by AFF4 and ELL2 (26, 28, 32).

Defining SEC structure is critical for understanding the roles of this family of complexes in the transcription of HIV and

Author contributions: S.C., H.U., K.B., U.S.-G., A.J.S., N.H., A.N., H.L., N.J.K., Q.Z., and T.A. designed research; S.C., H.U., K.B., U.S.-G., A.J.S., N.H., and A.N. performed research; S.C., H.U., K.B., U.S.-G., A.J.S., N.H., A.N., H.L., N.J.K., Q.Z., and T.A. analyzed data; and S.C., H.U., N.J.K., Q.Z., and T.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.R.C. is a guest editor invited by the Editorial Board.

[1]Present address: Department of Microbiology, University of Washington, Seattle, WA 98195.

[2]Present address: Department of Immunology, Duke University School of Medicine, Durham, NC 27710.

[3]Present address: Salk Institute for Biological Studies, San Diego, CA 92186.

[4]To whom correspondence should be addressed: E-mail: tom@ucxray.berkeley.edu.

See Author Summary on page 395 (volume 110, number 2).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216971110/-/DCSupplemental.

metazoan genes. Although functional and biochemical studies have revealed SEC components, the structural organization and molecular mechanisms of assembly have not been defined. Analysis of SECs in vivo suggests that AFF4 mediates complex formation through discrete binding segments (26–28), but no structural or functional domains have been identified in this protein. Moreover, the heterogeneity of SEC assemblies in vivo has hindered characterization of the direct interactions in single complexes. Here we use in vitro reconstitution, binding-site analyses, and cell-based assays of elongation-factor binding and transcriptional stimulation to characterize the overall architecture of the SEC recruited by HIV-1 Tat. AFF4 emerges as a flexible central scaffold with dispersed binding sites for CycT1, ELL2, and ENL or AF9. Unlike ordered scaffolds, AFF4 is largely unstructured, and assembly sites map to 20- to 70-aa segments distributed along the sequence. Binding of CycT1, ELL2, and ENL to AFF4 is neither dependent on HIV-1 Tat nor interdependent. The proteins that bind AFF4 are modular and bifunctional. For example, ELL2, ENL, and AF9 each have a C-terminal domain that contacts AFF4 and an N-terminal domain that interacts with PAF1, suggesting multiple AFF4c components physically link P-TEFb to PAF1c. These results support a model for AFF4 as a flexible tether at the core of the complex, with CycT1, ELL2, and ENL/AF9 bridging P-TEFb to a larger network of transcription factors that bind and regulate RNA Pol II.

Results

AFF4 Directly Assembles Components Through Distinct Binding Regions. To map direct interactions that mediate assembly, we expressed and purified segments of each SEC subunit and assayed for the formation of stable subcomplexes. AFF4 contains nearly 1,200 amino acids, but bioinformatic analysis revealed no recognizable structural domains. Accordingly, we expressed segments differing in length by 300 residues. In contrast, the other SEC subunits contain structurally defined domains that guided our expression strategy (Fig. 1*A*). From engineered *Escherichia coli* expression strains, we purified the cyclin domain of CycT1 (residues 1–268 and 1–303), the C-terminal occludin domain of ELL2 (residues 518–640), and the N-terminal YEATS domain of ENL. In addition, we identified a C-terminal domain conserved in the paralogs AF9 (residues 420–568) and ENL (residues 433–559) that has predicted structural similarity to the T1 domain of Brd4 (Fig. S1), a positive regulator of P-TEFb (33, 34). Because the AF9 and ENL C-terminal domains interact competitively with AFF4 (23, 29, 30), we analyzed one of the paralogs, ENL. To purify complexes of HIV-1 Tat and P-TEFb, we adapted baculovirus coexpression strategies (35).

In vivo, AFF4 recruits SEC components through defined regions in the first 900 amino acids (23, 29, 30). $AFF4_{1–300}$ associates in vivo with P-TEFb, $AFF4_{300–600}$ recruits the C-terminus of ELL2, and $AFF4_{600–900}$ binds competitively to the C-terminal domain of homologs ENL and AF9. Because direct contacts within the SEC cannot be mapped from these studies, we assayed interactions between purified domains of the components in vitro. Using analytical gel-exclusion chromatography, we found that $AFF4_{1–300}$ and $CycT1_{1–268}$ coeluted as a complex (Fig. 1*B*). Likewise, $AFF4_{300–600}$:$ELL2_{518–640}$ and $AFF4_{600–900}$:$ENL_{433–559}$ formed stable 1:1 complexes (Fig. 1 *C* and *D*). Thus, AFF4 recruited components by direct binding in a stoichiometric and modular fashion to regions distributed along the sequence. Consistent with interactions defined by coexpression of SEC subunits in Sf9 cells (29), the CDK9 subunit of P-TEFb was not required for AFF4 recognition.

AFF4 Is an Intrinsically Disordered Scaffold. The lack of identifiable structural motifs in AFF4 raised the question of how it supports complex formation. Strikingly, sequence-based secondary structure analysis (36) predicted that 94% of the AFF4-binding region (residues 1–900) is disordered (Fig. 2*A*). The sequence also is rich in hydrophilic residues (26% Ser + Thr, 16% Lys + Arg, and 13% Asp + Glu) and Gly (5%) that would favor a flexible, unfolded structure. To test the prediction that the AFF4 -regions are intrinsically unfolded, we assessed the susceptibility of the purified scaffold segments to proteolysis in vitro using trace amounts of proteinase K. $AFF4_{1–300}$, $AFF4_{300–600}$, and $AFF4_{600–900}$ were hypersensitive to proteinase K (Fig. 2*B*). Limited proteolysis failed to produce large, stable fragments of these AFF4 polypeptides, suggesting intrinsic disorder over the entire sequence.

To assess whether the SEC subunits mask proteolytic cleavage sites or promote AFF4 folding, we compared the limited proteolysis of AFF4 alone with that of the complexes with $CycT1_{1–268}$, $ELL2_{518–640}$, or $ENL_{433–559}$. In contrast to AFF4, the CycT1, ELL2, and ENL domains were resistant to proteolysis under these conditions, highlighting the extensive disorder of AFF4 even in the presence of the other SEC subunits (Fig. 2*B*). Although slight differences were observed in the fragmentation patterns, the overall similarity of AFF4 proteolysis fragments >10 kDa in the absence and presence of the ELL2- and ENL-binding domains (Fig. 2*B*) indicated that these partners do not protect large segments of the scaffold.

Flexible Linkers Connect Conserved, Hydrophobic Binding Modules on AFF4. To test further the idea that AFF4 presents short binding sequences, we finely mapped the binding sites of CycT1, ELL2, and ENL. Because protein–protein interactions often are driven by the burial of hydrophobic surfaces, we used a hydropathy plot to define candidate binding segments in AFF4. Unlike typical globular proteins, AFF4 contains only a few short hydrophobic clusters interspersed between hydrophilic stretches (Fig. 3*A*). The ELL2- and ENL-binding domains of AFF4 contain only one or two major hydrophobic clusters, respectively. This pattern of islands of hydrophobic segments separated by low-complexity linkers is preserved in the *Drosophila* AFF4 ortholog, Lilliputian (Fig. S2). This qualitative conservation suggests that these hydrophobic regions mediate assembly.

To test the hypothesis that AFF4 hydrophobic sites mediate associations with ELL2 and ENL, we used a native gel-shift assay to detect binding of ELL2 and ENL domains to ∼20-residue peptides encompassing the hydrophobic clusters. A purified synthetic peptide corresponding to $AFF4_{318–337}$, but not $AFF4_{303–322}$, changed the electrophoretic mobility of $ELL2_{518–640}$ (Fig. 3*B*), suggesting specific binding. $AFF4_{600–900}$ contains two hydrophobic clusters, but a deletion mutant $AFF4_{600–744}$ missing the second cluster preserved binding (Fig. S3*A*), suggesting that $AFF4_{745–900}$ is not required for interaction with ENL. Therefore, we analyzed a synthetic peptide corresponding to the first hydrophobic cluster, $AFF4_{710–729}$. This 20-residue peptide shifted $ENL_{433–559}$ on a native gel (Fig. 3*B*), indicating that this segment is sufficient to bind and recruit ENL. In conjunction with the extensive disorder of AFF4, these data suggest that AFF4 recruits ELL2 and ENL directly via discrete, short, hydrophobic binding modules connected by linker regions that remain flexible upon complex assembly.

In contrast to the ELL2- and ENL/AF9-binding segments, $AFF4_{1–300}$ contains multiple hydrophobic clusters (Fig. 3*A*) that might mediate binding to CycT1. To test whether these hydrophobic regions correspond to one or several CycT1-binding sites within $AFF4_{1–300}$, we subjected the $AFF4_{1–300}$ to limited proteolysis with proteinase K and identified peptide fragments that retained the ability to pull down with CycT1 in vitro (Fig. S3*C*). Mass spectrometry revealed that these bands contain AFF4 fragments 18–98 and 17–122, suggesting that the N-terminal end of $AFF4_{1–300}$ mediates the interaction with CycT1. To explore further the requirements for assembly, we analyzed binding of recombinant $CycT1_{1–268}$ to truncations of AFF4 using a native gel-shift assay. $CycT1_{1–268}$ electrophoretic mobility was changed

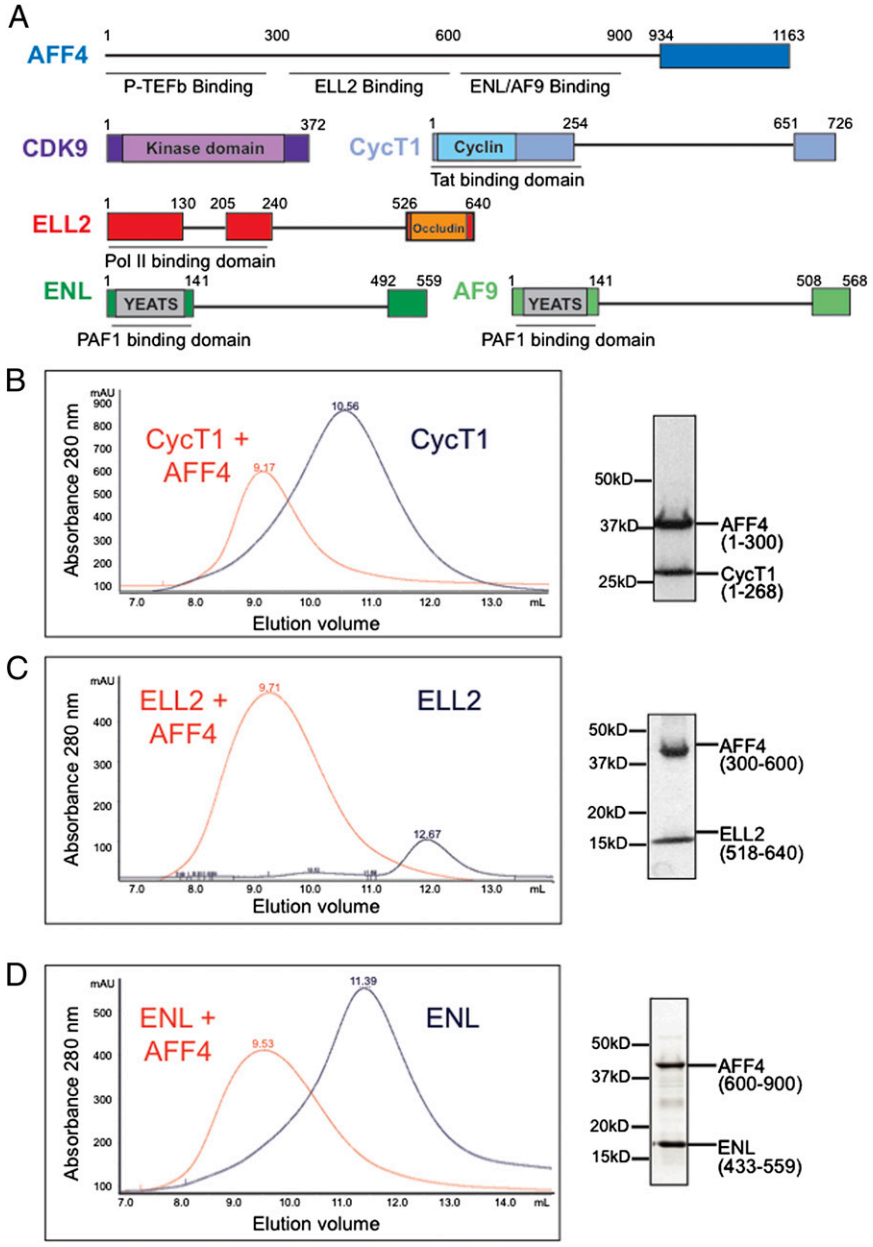

Fig. 1. AFF4 directly binds CycT1, ELL2, and ENL through defined segments. (*A*) Domain architecture of SEC subunits. (*B–D*) Pairs of purified protein constructs were incubated at equimolar concentrations and separated by gel-exclusion chromatography. (*Left*) For comparison, elution profiles for protein pairs are overlaid with profiles obtained in separate experiments for each AFF4 partner in isolation. (*Right*) SDS/PAGE of the peak fraction of each mixture shows both proteins in complex.

by $AFF4_{1–300}$, $AFF4_{1–230}$, $AFF4_{1–209}$, and $AFF4_{2–73}$. However, deletion of the N-terminal 79 residues ($AFF4_{80–300}$) eliminated the effect on $CycT1_{1–268}$ mobility (Fig. 3*B*), demonstrating that the N terminus of $AFF4_{1–300}$ is necessary and sufficient for binding CycT1 of P-TEFb.

To explore the possibility that larger complexes could protect bigger AFF4 fragments, we used limited proteolysis to probe a five-protein complex comprising HIV-1 Tat-P-TEFb purified from baculovirus-infected cells assembled with equimolar amounts of GST-$ELL2_{518–640}$ and $AFF4_{1–368}$ (Fig. 3*D*). This SEC subcomplex, which encompasses the scaffold-binding sites for CycT1 and ELL2, was stable to purification by gel-exclusion chromatography (Fig. 3*D*, lane 1). P-TEFb and ELL2 domains were protease resistant, as seen in the binary complexes, but Tat and AFF4 were protease hypersensitive. No stable AFF4 fragments >10 kDa were observed. Mass spectroscopy of the reaction purified using SDS PAGE revealed peptides including $Tat_{8–49}$, $AFF4_{18–98}$, $AFF4_{215–225}$, and $AFF4_{298–363}$, which include the regions of AFF4 that interact directly with binding partners.

Recruitment of Binding Partners Is Coupled to Folding of AFF4 Assembly Sites. The results described above suggest that the first 900 residues of AFF4 form a largely disordered scaffold, even in complex with one or more binding partners. To determine if the hydrophobic AFF4 segments fold in complexes with the partner domains, we determined whether the AFF4-binding sites form a secondary structure in the binary complexes. Circular dichroism (CD) spectra of purified peptides corresponding to AFF4-binding sites ($AFF4_{2–73}$, $AFF4_{318–337}$, and $AFF4_{710–729}$) revealed the distinct absence of secondary structure (Fig. 4 *A–C*, red). In contrast,

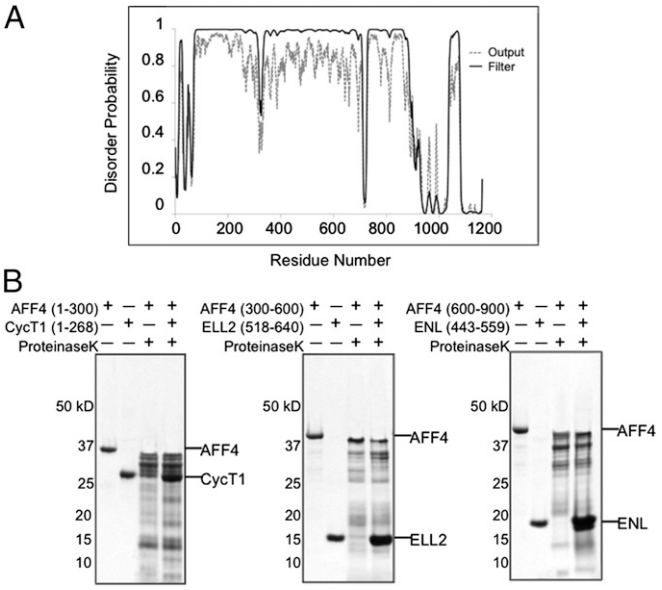

Fig. 2. AFF4 is largely disordered in the absence and presence of binding partners. (*A*) Predicted disorder profile of AFF4 reveals large regions are predicted to lack structure. Profile was calculated using DISOPRED (36) using a 15-residue window (Output) with a 2% false-positive cutoff (Filter). (*B*) High sensitivity to limited proteolysis suggests that AFF4 is natively unfolded. Recombinant $AFF4_{1–300}$, $AFF4_{300–600}$, or $AFF4_{600–900}$ was incubated with proteinase K (1:4,000) without and with the cognate partner for 10 min at 4 °C and analyzed by SDS/PAGE. AFF4 segments were digested rapidly in comparison with the binding partners. Complex formation with individual binding partners does not significantly alter the patterns of large fragments.

the binding partners ($CycT1_{1–268}$, $ELL2_{518–640}$, and $ENL_{433–559}$) gave spectra corresponding to folded proteins (Fig. 4 *A–C*, blue). Upon mixing stoichiometric amounts of the AFF4 sites and the respective cognate partners (Fig. 4 *A–C*, purple vs. green), all three AFF4 assembly sites showed a significant increase in the CD signal resulting from complex formation. Because the CycT1, ELL2, and ENL domains are well structured in isolation, the acquisition of a secondary structure promoted by binding likely occurs in the short, hydrophobic binding sites in AFF4 that fold locally upon assembly.

Subunits Associate Through AFF4. AFF4, ELL2, and ENL associate with HIV-1 Tat-P-TEFb in vivo, and Tat enhances recruitment of ELL2 to the complex (26, 28, 32). These data suggest that Tat may bind AFF4, ELL2, and ENL directly to mediate further assembly of the components with P-TEFb. To determine whether Tat has a direct role in recruiting ELL2 or ENL, we tested the ability of purified Tat-P-TEFb to bind these factors. The interactions were measured using an in vitro affinity chromatography assay with full-length Flag-AFF4 affinity-purified from HeLa cells under stringent conditions that removed the endogenous partners. Recombinant truncated Tat-P-TEFb purified from baculovirus-infected cells was incubated with equimolar amounts of GST-$ELL2_{518–640}$, GST-$ENL_{433–559}$, or both $ELL2_{518–640}$ and GST-$ENL_{433–559}$ in the absence and presence of full-length AFF4 (Fig. 5). GST pull-downs of binding reactions revealed that the truncated Tat-P-TEFb associates with the C-terminal domains of ELL2 and ENL in the presence, but not in the absence, of AFF4 (Fig. 5, lanes 6–9). Neither Tat-P-TEFb nor P-TEFb alone stably binds $ELL2_{518–640}$ or $ENL_{433–559}$. Thus, AFF4 mediates incorporation of these ELL2 and ENL domains into the complex in vitro (lane 11).

AFF4-Dependent Complex Assembly Is Important for Transcriptional Activation. To investigate the functional significance of the AFF4 scaffolding sites defined in vitro, we measured the recruitment of SEC subunits by AFF4 in HeLa cells. We tested the importance of AFF4 interaction sites by introducing alanine substitutions at various hydrophobic residues in the full-length scaffold and measuring the association of the other SEC subunits in vivo (Fig. 6). Wild-type or mutant AFF4-Flag constructs were transfected into HeLa cells, and binding of the SEC subunits was measured in anti-Flag immunoprecipitations of nuclear extracts (Fig. 6*B* and Fig. S4 *A–C*). Several tandem alanine substitutions in the CycT1-binding site (Pro33Ala/L34Ala, Val41Ala/Thr42Ala, Arg51Ala/Ile52Ala, Met55Ala/Leu56Ala) decreased CycT1 associated with AFF4, whereas levels of other SEC subunits (ELL2, ENL, and AF9) remained unperturbed. Similarly, alanine substitutions in the cognate AFF4-binding sites specifically reduced levels of associated ELL2 (Ile300Ala, Leu305Ala, Val313Ala, Val316Ala, Ile319Ala, Trp327Ala, Ile334Ala, Thr340-Ala) and ENL or AF9 (Leu705Ala, Leu714Ala, Leu715Ala, Val716Ala, Ile718Ala, Leu720Ala, Thr724Ala, Leu714Ala/Ile718Ala) (Fig. 6 *A* and *B* and Fig. S4 *A–C*). These results suggest that hydrophobic interactions with AFF4 mediate SEC assembly and that AFF4 binds independently to P-TEFb, ELL2, and ENL/AF9.

To define the boundaries of the functional sites, we tested the effects of AFF4 mutations on AFF4-dependent transcriptional stimulation in HeLa cells. The SEC is essential for both Tat-dependent and Tat-independent transcription from the HIV promoter (28). HIV-1 Tat efficiently recruits the endogenous SECs to TAR and stabilizes ELL2, rendering Tat-dependent transcription relatively insensitive to the ectopic expression of AFF4. In contrast, without Tat, the AFF4 concentration limits SEC activity on the viral LTR, and transcription depends on AFF4 expression in a dose-dependent manner (28). Accordingly, to quantify the effects of AFF4 mutations in the presence of the endogenous wild-type protein, we measured the stimulation by AFF4 variants of Tat-independent, basal transcription of a luciferase reporter gene under the control of the HIV promoter.

An alanine scan across the ELL2-binding site, for example, revealed that substitutions between AFF4 His294 and Pro348 could reduce transcriptional stimulation by at least twofold (Fig. 6*C* and Fig. S5*A*). This region, including the segment where mutations caused the largest reductions (Gln303–Thr340), is bigger than the $AFF4_{318–337}$ peptide that binds the ELL2 C-terminal domain in vitro. In contrast, minimal effects were observed at sites flanking $AFF4_{294–348}$, suggesting that the observed phenotypes are linked to ELL2 recruitment. Likewise, an alanine scan of the ENL/AF9-binding region revealed that single-alanine substitutions between AFF4 Lys699 and Tyr731 can reduce transcriptional stimulation more than fourfold (Fig. 6*C* and Fig. S5*B*). This functional region also is larger than the $AFF4_{710–729}$ peptide that binds ENL in vitro. We submitted elsewhere a similar analysis of tandem alanine mutations across the CycT1-binding site, $AFF4_{2–73}$. These AFF4 sites, important for transcriptional activity in vivo, encompass the hydrophobic peptides that are sufficient for binding SEC subunits in vitro.

ELL2 and ENL also Bind AFF1. The AFF4 homolog AFF1 shares several binding partners with AFF4 in vivo, including ENL and P-TEFb. Comparison of the AFF1 and AFF4 sequences reveals that the ENL-binding site of AFF4 (residues 699–731) shares only 51% sequence identity with AFF1 (Fig. S6*A*). This low sequence identity suggests that not all residues contribute to binding or that AFF1 and AFF4 bind ENL differently. On the other hand, the AFF4 residues 294–348 that constitute the ELL2-binding site are 73% conserved in AFF1, including 100% identity of the crucial 318–337 site (Fig. S6*A*). These patterns suggest that AFF1 also binds ELL2. To test this prediction, we coprecipitated Flag-tagged AFF1 and ELL2 in vivo. AFF1

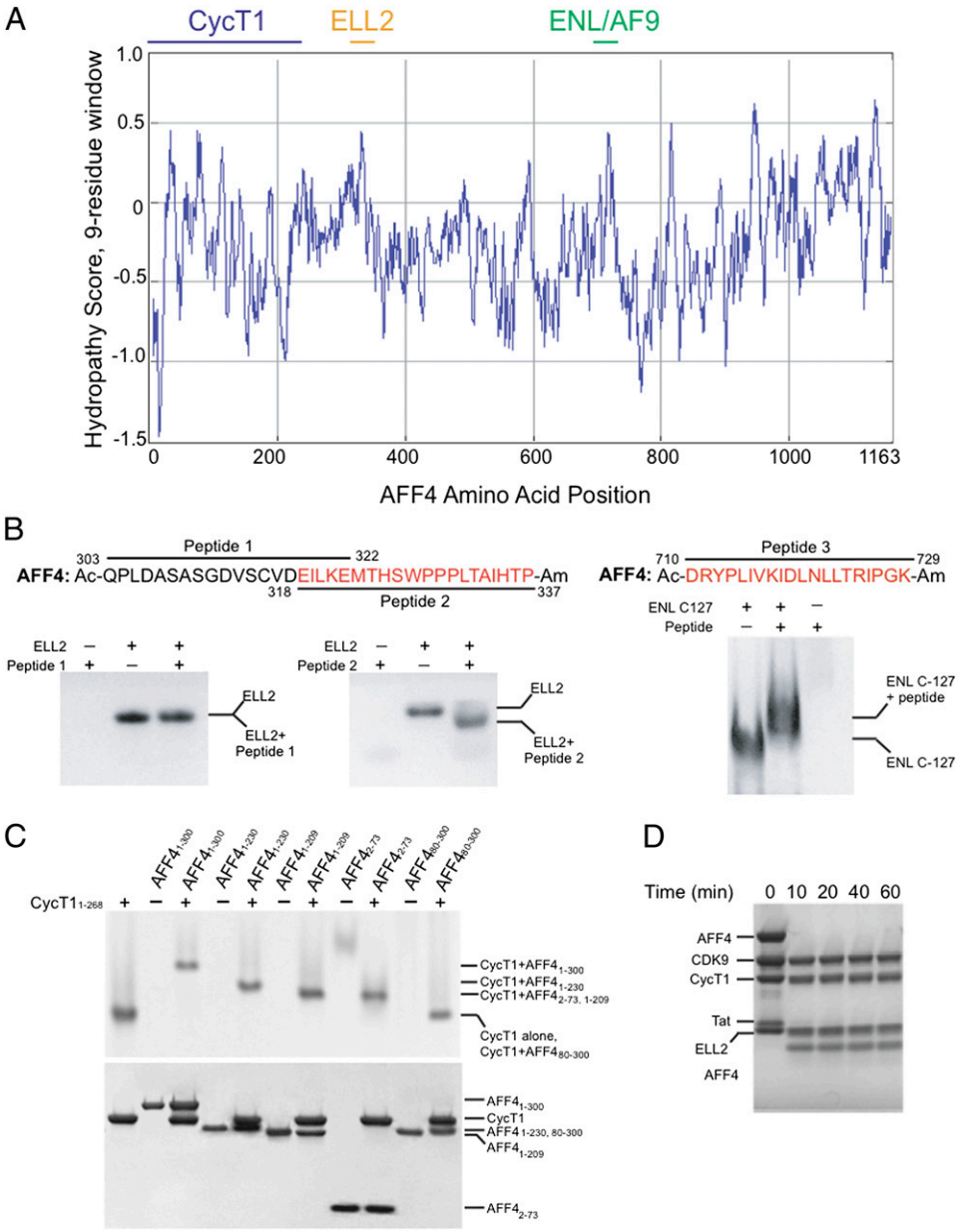

Fig. 3. AFF4 recruits partners via short, hydrophobic clusters. (*A*) Hydropathy plot of AFF4 calculated using a nine-residue window. Highest-scoring regions above 0.25 correspond to CycT1-, ELL2-, and ENL-binding segments. (*B*) Native gel electrophoresis with peptides corresponding to hydrophobic clusters identified in AFF4. $AFF4_{318–337}$ shifts the mobility of $ELL2_{518–640}$, and $AFF4_{710–729}$ shifts the mobility of $ENL_{433–559}$. (*C*) The N-terminal segment of $AFF4_{1–300}$ is required for CycT1 binding. (*Upper*) Native gel electrophoresis of $CycT1_{1–268}$ with (+) and without (–) AFF4 fragments. With the exception of the isolated $AFF4_{2–73}$, the AFF4 fragments in isolation do not enter the native gel. $AFF4_{1–300}$, $AFF4_{1–230}$, $AFF4_{1–209}$, and $AFF4_{2–73}$ shifted the mobility of $CycT1_{1–268}$, but $AFF4_{80–300}$ did not. (*Lower*) Control SDS/PAGE gel shows the composition of each sample. (*D*) Limited proteolysis of a five-protein subcomplex—Tat-P-TEFb-$AFF4_{1–368}$-$ELL2_{518–640}$—with trypsin shows rapid loss of the bands for AFF4 and Tat.

bound ELL2 as well as other SEC components, including ENL, AF9, and CDK9 (Fig. S6*B*).

ELL2 N-Terminal Domain Bridges the AFF4 Complex to PAFc. The PAF1 protein, the scaffold of PAFc, physically connects the SEC to Pol II by interacting directly with the N-terminal YEATS domain of ENL or AF9 (30). ENL/AF9 functions as a bridge—predicted flexible linker connects the ENL/AF9 C-terminal domain that interacts with AFF4 to the ENL/AF9 N-terminal domain that interacts with PAF1. Sequence-based secondary structure analysis of ENL/AF9 and ELL2 reveals similar organization of ordered domains (Fig. 7*A*). AF9, ENL, and ELL2 have small, ordered N- and C-terminal domains separated by hydrophilic linkers with low sequence complexity. This arrangement of binding modules raises the possibility that ELL2 also might bridge AFF4 to another transcriptional regulator.

Because the interaction between ENL/AF9 and PAF1 brings PAFc into close physical proximity with the SEC, we explored whether ELL2 also interacts with the PAF1 scaffold. We found that ELL2 coimmunoprecipitates with PAF1 in vivo (Fig. 7*B*). To determine whether this interaction is mediated indirectly through ENL/AF9, we assessed the coprecipitation of ELL2 deletions with PAF1. If ENL/AF9 mediate the ELL2/PAF1 interaction indirectly through AFF4c, a C-terminal ELL2 deletion that

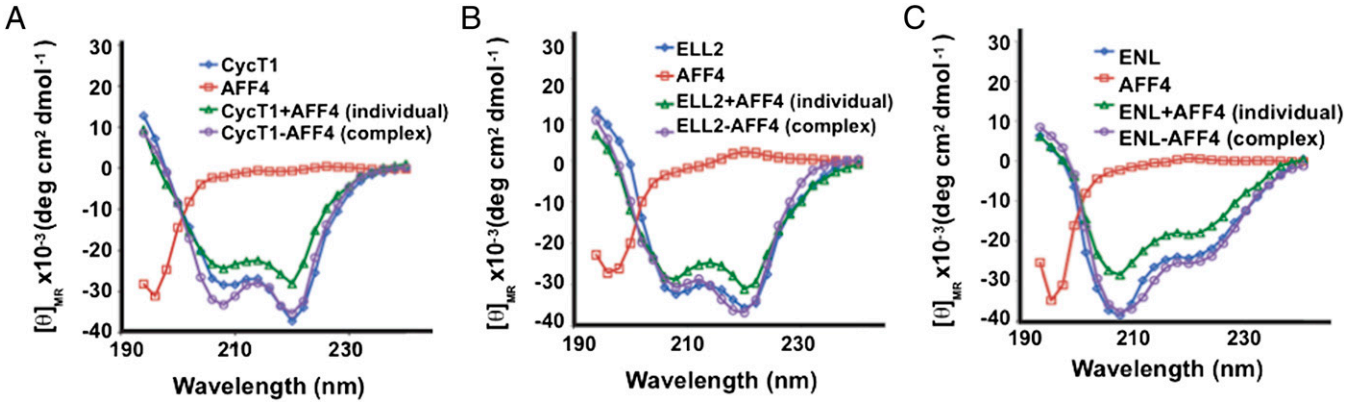

Fig. 4. Binding of AFF4 to partners changes the local structural landscape. (*A*–*C*) CD spectra of AFF4 assembly sites in the presence and absence of their binding partners. Mean residue ellipticity is shown for isolated peptides corresponding to recruitment sites along AFF4 (red), AFF4-binding partners (blue), unassembled protein pairs in a divided cuvettete (green), and the mixtures of AFF4 with protein partners (purple). (*A*) Spectra for $CycT1_{1–268}$ and $AFF4_{2–73}$ separately and in complex. (*B*) Spectra for $ELL2_{518–640}$ and $AFF_{318–337}$ separately and in complex. (*C*) Spectra for $ENL_{443–559}$ and $AFF_{710–729}$ separately and in complex.

prevents binding of ELL2 to AFF4 also would eliminate association with PAF1. However, the C-terminal deletion mutant ELL2 (Δ499–640), but not the N-terminal deletion mutant ELL2 (Δ50–194), coprecipitated with PAF1 (Fig. 7*C*).

To confirm that this interaction is direct, we coexpressed His-PAF1 with either GST-$ELL2_{50–194}$ or GST-$ELL2_{518–640}$ in *E. coli* and assessed complex formation using affinity chromatography. Only GST-$ELL2_{50–194}$ copurified with His-PAF1 (Fig. 7*C*). Therefore, the N-terminus of ELL2, which is not required for AFF4c assembly, directly binds PAF1 independently of ENL/AF9. These results indicate that ELL2, like ENL and AF9, connects AFF4c to PAFc.

Discussion

Large protein assemblies mediate transcription elongation. Although progress has been made toward identifying and characterizing the components of transcription elongation complexes, information is lacking about the structural organization of these assemblies crucial for understanding their specific functions. In addition to their large size, the intrinsic disorder of complex components creates challenges for structural analysis. Using a combination of interaction mapping and limited proteolysis, we defined the overall architecture of the HIV-1 Tat SEC and identified contacts between the complex and another transcriptional regulator, PAFc. The biochemical mapping of SEC-subunit binding sites in AFF4 enabled the identification of single- and double-residue substitutions that reduce the cellular activity of this nearly 1,200-aa scaffold.

Our in vitro reconstitutions using purified components reveal several fundamental principles of SEC organization. Scaffolds often are structurally defined platforms (e.g., Cullin-RING ligases) that control the spatial organization of partner proteins (37). In contrast, flexible tethering (exemplified by axin, BRCA1, p300 histone acetyltransferase, and the Ste5 kinase scaffold) also can increase the avidity of interactions among subunits and sometimes allosterically control signaling components (37–39).

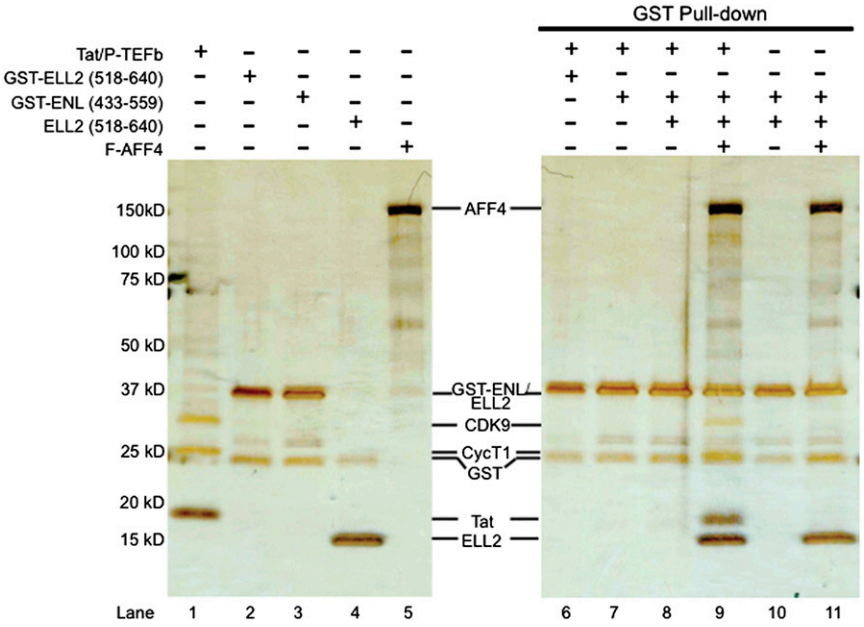

Fig. 5. HIV-1 Tat does not mediate complex assembly directly. AFF4 is required for assembly of Tat-P-TEFb with $ELL2_{518–640}$ and $ENL_{433–559}$ in vitro. (*Left*) Silver-stained gel of purified AFF4c components. (*Right*) GST pull-downs of Tat-P-TEFb with ELL2 and ENL ± full-length Flag-AFF4.

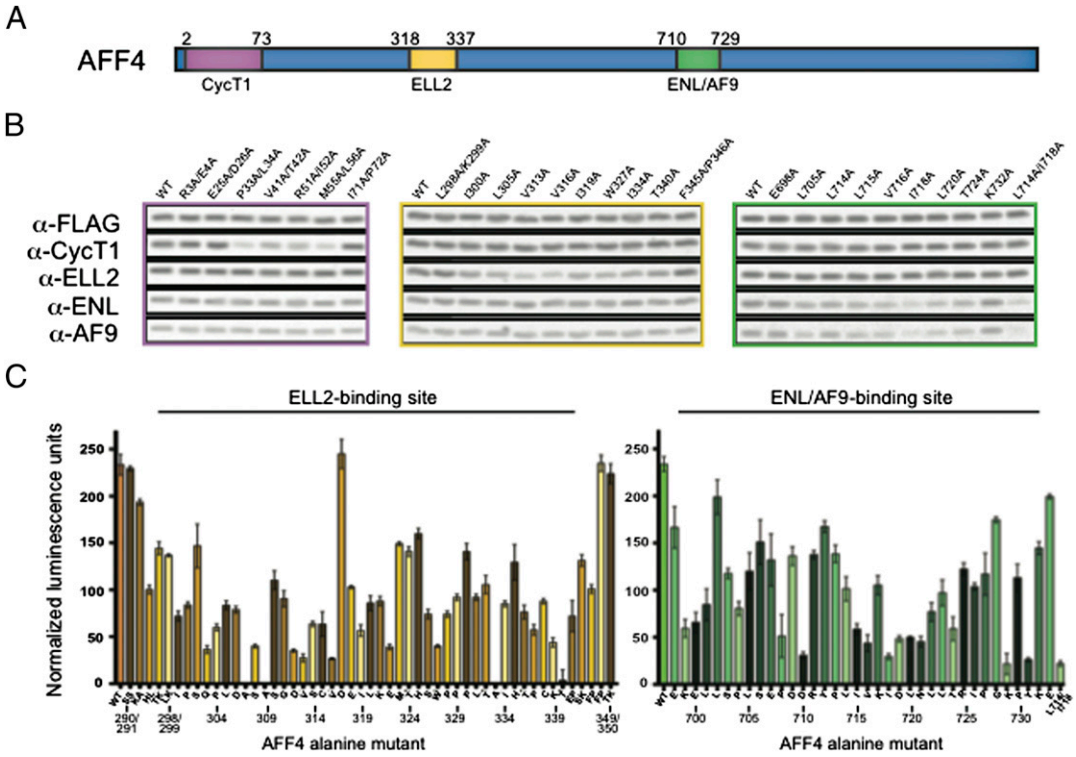

Fig. 6. Short, hydrophobic sites in AFF4 mediate SEC assembly and activity in HeLa cells. (*A*) Schematic representation of in vitro SEC assembly sites on AFF4. (*B*) AFF4-binding site residues influence complex assembly in vivo. Anti-Flag immunoprecipitations of nuclear extracts prepared from HeLa cells transfected with AFF4-Flag variants were analyzed by Western blotting. Coprecipitation of CycT1 was decreased significantly in several CycT1-binding site mutants. Substitutions in the ELL2 and ENL/AF9-binding sites caused similar reductions in binding of the cognate subunit. (*C*) Transcriptional effects of Ala mutations at ELL2- (yellow, *Left*) and ENL/AF9- (green, *Right*) binding sites of AFF4. Luciferase activity, as a surrogate for transcription from the HIV LTR, was measured in extracts of HeLa cells cotransfected with a luciferase reporter construct and an expression vector for the indicated AFF4 mutants. Activity was normalized to AFF4 expression levels. Values represent the mean of three independent assays.

The AFF4 scaffold is a strikingly disordered protein that coordinates binding partners through 20- to 70-aa sites interspersed with flexible linker regions. The complex components assemble on AFF4 like flags on a line. The long, unstructured nature of AFF4 implies that flexibility may be an important organizational principle required for SEC function. This flexibility, as in other intrinsically disordered scaffolds (39, 40), likely modulates binding affinity, allows the coordination of multiple components over long distances, and provides mechanisms for dynamic adaptation to new binding partners and spatial requirements.

The binding modes of AFF4 partners also provide insights into how flexibility promotes specific functions. ELL2 and ENL/AF9 have small, independently folded N- and C-terminal domains separated by linker regions with little predicted structure. These proteins bind AFF4 via their C-terminal domains and recruit PAF1 through their N-terminal domains. These properties allow ELL2 and ENL/AF9 to bridge the AFF4c and PAFc flexibly, which may be important for crosstalk between the complexes during transcription. This activity remains to be reconciled with other reported functions for ELL family members, including binding to Mediator through Med26 (41) and stimulating Pol II (21). With less flexibility, CycT1 links CDK9 and AFF4. These binary connections bridge CDK9 to RNA Pol II through SECs and PAFc.

A major challenge in defining the functions of natively unstructured proteins is the identification of interaction domains (39). Minimal AFF4 protein interaction sites mapped to the few short, hydrophobic segments in the scaffold sequence. These data suggest that we have identified most of the hydrophobic AFF4 protein-binding sites. One exception is the C-terminal segment, which is a candidate for mediating additional, uncharacterized interactions (25, 29). The correlation between hydrophobicity and protein–protein binding in AFF4 could reflect a general property of disordered proteins. Nonetheless, the short AFF4 peptides that are sufficient to bind and fold in vitro represent operationally minimized recognition sequences contained within larger functional sites. For example, the ELL2-binding peptide $AFF4_{318–337}$ is part of a larger segment, $AFF4_{294–348}$, that contains residues important for ELL2 binding and transcriptional stimulation in HeLa cells. Similarly, the ENL-binding peptide $AFF4_{710–729}$ is contained within a larger functional segment encompassing $AFF4_{698–731}$.

In contrast to the reported failure of complexes of ELL with ELL-associated factors (EAF1/2) to bind AFF1 or AFF4 (29), we observed a direct interaction between $AFF4_{300–600}$ and $ELL2_{518–640}$ (Fig. 1*C*). A 20-residue segment conserved in AFF4 paralogs was sufficient to mediate this interaction in vitro. In addition, our assembly of purified SEC subcomplexes did not recapitulate a reported direct interaction between coexpressed ELL and CycT1 (29). These differences in apparent interactions may arise from differences in assay formats, the activities of the EAFs, potential distinctions between ELL and ELL2, or the use of full-length versus truncated SEC subunits.

Our data show that SEC subunits can form independent binary complexes with the scaffold in vitro and that AFF4 mutations that individually reduce binding of P-TEFb, ELL2, or ENL/AF9 in HeLa cells do not affect the associations of the other subunits. These results suggest an inherent combinatorial multiplicity that may be modulated to allow greater functional diversity (42). Potential dynamic variations in the composition of these complexes have been suggested for the HIV promoter (17). AFF1

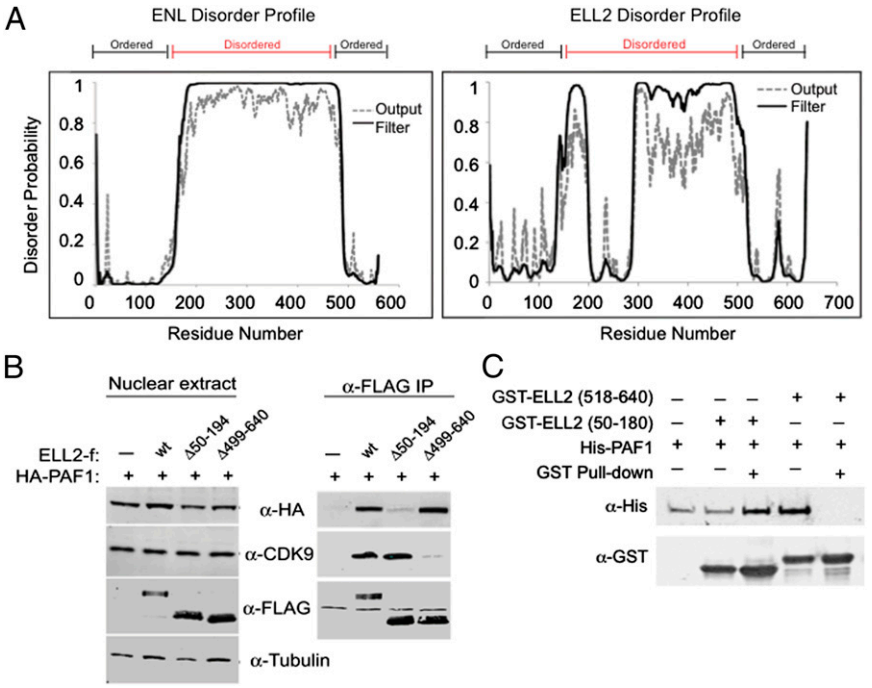

Fig. 7. ELL2 connects AFF4c to PAFc by binding PAF1. (*A*) ENL and ELL2 contain continuous regions of predicted disorder between the N- and C-terminal domains. DISOPRED disorder profiles of ENL and ELL2 are shown. (*B*) ELL2 in HeLa cells associates with PAF1. Nuclear extracts (*Left*) and anti-Flag immunoprecipitations (*Right*) prepared from HeLa cells cotransfected with HA-PAF1 and Flag-ELL2 wild-type, ELL2 (Δ499–640), or ELL2 (Δ50–194) were analyzed by Western blotting. HA-PAF1 coprecipitation was diminished significantly by the Flag-ELL2 (Δ50–194) deletion, suggesting that the ELL2 N terminus is required for mediating the interaction with PAF1. (*C*) The N terminus of ELL2 binds PAF1 directly. GST-ELL2Δ518–640 and GST-ELL2Δ50–194 were coexpressed with His-PAF1 in *E. coli*. Western analysis of GST affinity capture of cell lysate shows copurification of His-PAF1 with GST-ELL2Δ50–194 but not GST-ELL2Δ518–640.

also binds P-TEFb in vivo (25, 42), and ENL and AF9 compete for scaffold binding (30). With four, three, and two human homologs of AFF4, ELL, and ENL, respectively, as well as posttranslational modifications of the subunits, including P-TEFb (14, 43, 44), these data demonstrate the potential for considerable variability in SEC composition. Alternatively, specific complexes may assemble preferentially or cooperatively. Contacts that are not sufficient to stabilize binary complexes may mediate cooperative interactions. For example, the presence of HIV-1 Tat enhances ELL2 recruitment (26, 32). However, Tat is not required for AFF4-dependent complex assembly, and Tat does not recruit the C-terminal domains of ELL2 or ENL directly in vitro. To define the functions of these factors and to understand the specific requirement for the AFF4 complex in HIV transcription, it will be necessary to determine how SEC composition is regulated both spatially and temporally in cells. The model of AFF4 as a flexible scaffold with dispersed, short, hydrophobic binding sites that recruit bifunctional connecting proteins provides a road map to define and distinguish the activities of SEC assemblies.

Materials and Methods

Flag-AFF4 was purified from HeLa cells (23). P-TEFb and Tat-P-TEFb ($Tat_{1–86}$, $CDK9_{1–330}$, $CycT1_{1–264}$) were purified from Sf9 cells infected with baculovirus expression vectors. AFF4 fragments were cloned into the pET28b expression vector (Invitrogen), CycT1 cDNA fragments into the pETDuet-1 expression vector (Invitrogen), and ELL2 and ENL fragments into pGEX-6P-3 (GE Healthcare). Proteins were expressed in *E. coli* BL21 CodonPlus-(DE3)-RIL cells (Stratagene) by isopropylthiogalactoside induction for 16 h at 16 °C. For purification of His_6-tagged and GST-tagged proteins, cells were resuspended in buffer A [20 mM Hepes (pH 7.5), 0.5 M NaCl, 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP), 10% (vol/vol) glycerol, 0.2 mM 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride], and proteins were purified by gradient elution from a Ni-NTA affinity column (GE Healthcare) or a GST-FF column (GE Healthcare) with imidazole or glutathione, respectively. Tags were cleaved by incubation with 1:50 $(His)_6$-tobacco etch virus protease for 22 h at 4 °C. Proteins were purified further using a Superdex S75 gel-filtration column (GE Healthcare).

For in vitro precipitations, proteins were incubated at equimolar ratios at 4 °C for 30 min in 0.4 M NaCl, 20 mM Hepes (pH 8.0), 5% (vol/vol) glycerol, 0.5 mM TCEP. GST-Sepharose or Ni-NTA agarose was added to cognate reactions and incubated for 20 min at room temperature. Beads were washed with 0.4 M NaCl, 20 mM Hepes (pH 8.0), 10% (vol/vol) glycerol, 0.5% Nonidet P-40, 0.5 mM TCEP and boiled in SDS/PAGE buffer. Bound proteins were analyzed by gel electrophoresis.

Peptide-binding reactions were carried out in 100 mM NaCl, 20 mM Hepes (pH 8.0), 0.5 mM TCEP. ELL2 or ENL domains (2 μg) were incubated in the presence or absence of 2 μg purified synthetic peptide (obtained from the University of Utah School of Medicine DNA/Peptide Core Facility) for 15 min on ice. Reactions were separated in a 4–20% Tris-glycine native gel by electrophoresis at 4 °C for 3 h. Complex assembly also was examined by analytical gel-exclusion chromatography. Proteins were combined in 150 mM NaCl, 20 mM Hepes (pH 8.0), 0.5 mM TCEP at 4 °C for 20 min and separated on a Sephadex S75 column (GE Healthcare) using an AKTA Explorer FPLC system (GE Healthcare). Eluted complexes were compared with individual input proteins and molecular weight standards. For limited proteolysis, proteins were mixed at equimolar ratios for 30 min at room temperature, and pK (1:4,000) was added for 15 min at room temperature in 150 mM NaCl, 20 mM Hepes (pH 8.0), 0.5 mM TCEP. Limited proteolysis with trypsin was carried out on ice for various times. Reactions were stopped by the addition of SDS/PAGE buffer.

CD spectra were recorded on a JASCO J-815 Circular Dichroism Spectrometer using a 10-mm path length divided quartz cuvettete. Individual proteins or reactions containing equimolar concentrations of protein were equilibrated in 40 mM sodium fluoride, 50 mM potassium phosphate, pH 7.5 for data collection at 190–240 nm at 25 °C with a bandwidth of 1 nm.

For coimmunoprecipitations, nuclear extracts from HeLa cells transfected with specific cDNA constructs (23) were incubated with anti-Flag or anti-HA agarose beads (Sigma-Aldrich) for 2 h at 4 °C. The beads were washed with 0.3 M KCl, 20 mM Hepes (pH 7.9), 10% (vol/vol) glycerol, 0.2 mM EDTA, 0.2% N, 1 mM DTT, 0.5 mM PMSF. Proteins were eluted with buffers containing synthetic Flag or HA peptides and analyzed by Western blotting (28).

Transcriptional stimulation by AFF4 variants was measured as described (28) using a luciferase assay of extracts of HeLa cells cotransfected with an expression vector for AFF4-Flag and a reporter plasmid encoding the luciferase

gene transcribed from the HIV promoter. Briefly, HeLa cells were cotransfected with 100 ng of an HIV-LTR firefly luciferase reporter construct and 350 ng of pCDNA3.1 containing AFF4 using the 25-kDa linear polyethyleneimine reagent (Sigma-Aldrich). After 48 h, the cells were lysed in passive lysis buffer (Promega) containing 0.5 mM PMSF for 5 min at 25 °C. The lysates were incubated with firefly luciferase substrate, and luminescence was measured on a SpectraMax L microplate reader (Molecular Devices). The relative luminescence was normalized to the concentration of AFF4 in the cell determined by Western blotting using an anti-Flag primary antibody.

ACKNOWLEDGMENTS. We thank Ann Fischer for help with baculovirus production and tissue culture; David King, Tony Lavarone, and Lori Kohlstaedt for mass spectrometry; Scott Endicott and Bob Schackmann for peptide synthesis and purification; Alberto Stolfi for cloning assistance; and Christoph Grundner and James Fraser for support and advice. This work was supported by National Institutes of Health Grants R01AI41757-11 and R01AI095057) (to Q.Z.) and P50GM82250 (to N.J.K. and T.A.) and by an Innovative, Developmental Exploratory Award (IDEA) from the California HIV/AIDS Research Program (to T.A.). S.C. and N.H. held predoctoral fellowships from the California HIV/AIDS Research Program.

1. 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.
2. 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.
3. Ptashne M (2005) Regulation of transcription: From lambda to eukaryotes. *Trends Biochem Sci* 30(6):275–279.
4. Saunders A, Core LJ, Lis JT (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7(8):557–567.
5. Muñoz MJ, de la Mata M, Kornblihtt AR (2010) The carboxy terminal domain of RNA polymerase II and alternative splicing. *Trends Biochem Sci* 35(9):497–504.
6. Carrillo Oesterreich F, Bieberstein N, Neugebauer KM (2011) Pause locally, splice globally. *Trends Cell Biol* 21(6):328–335.
7. Lenasi T, Barboric M (2010) P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. *RNA Biol* 7(2):145–150.
8. Buratowski S (2009) Progression through the RNA polymerase II CTD cycle. *Mol Cell* 36(4):541–546.
9. Muse GW, et al. (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39(12):1507–1511.
10. 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.
11. Core LJ, Lis JT (2008) Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319(5871):1791–1792.
12. Wada T, et al. (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* 12(3):343–356.
13. Yamaguchi Y, et al. (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97(1):41–51.
14. Cho S, Schroeder S, Ott M (2010) CYCLINg through transcription: Posttranslational modifications of P-TEFb regulate transcription elongation. *Cell Cycle* 9(9):1697–1705.
15. Barboric M, Peterlin BM (2005) A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. *PLoS Biol* 3(2):e76.
16. Karn J (1999) Tackling Tat. *J Mol Biol* 293(2):235–254.
17. Ott M, Geyer M, Zhou Q (2011) The control of HIV transcription: Keeping RNA polymerase II on track. *Cell Host Microbe* 10(5):426–435.
18. 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.
19. Shilatifard A, et al. (1997) ELL2, a new member of an ELL family of RNA polymerase II elongation factors. *Proc Natl Acad Sci USA* 94(8):3639–3643.
20. 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.
21. Shilatifard A, Haque D, Conaway RC, Conaway JW (1997) Structure and function of RNA polymerase II elongation factor ELL. Identification of two overlapping ELL functional domains that govern its interaction with polymerase and the ternary elongation complex. *J Biol Chem* 272(35):22355–22363.
22. Arndt KM, Kane CM (2003) Running with RNA polymerase: Eukaryotic transcript elongation. *Trends Genet* 19(10):543–550.
23. 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 USA* 108(36):E636–E645.
24. Jäger S, et al. (2011) Purification and characterization of HIV-human protein complexes. *Methods* 53(1):13–19.
25. 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.
26. 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.
27. Lin C, et al. (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* 37(3):429–437.
28. He N, et al. (2010) HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol Cell* 38(3):428–438.
29. Biswas D, et al. (2011) Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc Natl Acad Sci USA* 108(38):15751–15756.
30. Luo Z, et al. (2012) The super elongation complex family of RNA polymerase II elongation factors: Gene target specificity and transcriptional output. *Mol Cell Biol* 32(13):2608–2617.
31. Lin C, et al. (2011) Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes Dev* 25(14):1486–1498.
32. He N, Zhou Q (2011) New insights into the control of HIV-1 transcription: When Tat meets the 7SK snRNP and super elongation complex (SEC). *J Neuroimmune Pharmacol* 6(2):260–268.
33. 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.
34. 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.
35. Tahirov TH, et al. (2010) Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* 465(7299):747–751.
36. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 337(3):635–645.
37. Good MC, Zalatan JG, Lim WA (2011) Scaffold proteins: Hubs for controlling the flow of cellular information. *Science* 332(6030):680–686.
38. Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447(7147):1021–1025.
39. Cortese MS, Uversky VN, Dunker AK (2008) Intrinsic disorder in scaffold proteins: Getting more from less. *Prog Biophys Mol Biol* 98(1):85–106.
40. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6(3):197–208.
41. Takahashi H, et al. (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* 146(1):92–104.
42. Smith E, Lin C, Shilatifard A (2011) The super elongation complex (SEC) and MLL in development and disease. *Genes Dev* 25(7):661–672.
43. Flatz L, et al. (2011) Single-cell gene-expression profiling reveals qualitatively distinct CD8 T cells elicited by different gene-based vaccines. *Proc Natl Acad Sci USA* 108(14): 5724–5729.
44. Sakane N, et al. (2011) Activation of HIV transcription by the viral Tat protein requires a demethylation step mediated by lysine-specific demethylase 1 (LSD1/KDM1). *PLoS Pathog* 7(8):e1002184.