Schizosaccharomyces pombe Pol II transcription elongation factor ELL functions as part of a rudimentary super elongation complex

Sneha Gopalan[1,2], Dana M. Gibbon[1], Charles A.S. Banks[1], Ying Zhang[1], Laurence A. Florens[1], Michael P. Washburn[1,3], Preeti Dabas[4], Nimisha Sharma[4], Christopher W. Seidel[1], Ronald C. Conaway[1,5] and Joan W. Conaway[1,5,*]

[1]Stowers Institute for Medical Research, Kansas City, MO 64110, USA, [2]The Open University, Milton Keynes, UK, [3]Department of Pathology and Laboratory Med icine, University of Kansas Medical Center, Kansas City, KS 66160, USA, [4]University School of Biotechnology, G.G.S.Indraprastha University, New Delhi 110078, India and [5]Department of Biochemistry and Molecular Biology, University of Kansas Medical Center, Kansas City, KS 66160, USA

ABSTRACT

ELL family transcription factors activate the overall rate of RNA polymerase II (Pol II) transcription elongation by binding directly to Pol II and suppressing its tendency to pause. In metazoa, ELL regulates Pol II transcription elongation as part of a large multisubunit complex referred to as the Super Elongation Complex (SEC), which includes P-TEFb and EAF, AF9 or ENL, and an AFF family protein. Although orthologs of ELL and EAF have been identified in lower eukaryotes including *Schizosaccharomyces pombe,* it has been unclear whether SEC-like complexes function in lower eukaryotes. In this report, we describe isolation from *S. pombe* of an ELL-containing complex with features of a rudimentary SEC. This complex includes *S. pombe* Ell1, Eaf1, and a previously uncharacterized protein we designate Ell1 binding protein 1 (Ebp1), which is distantly related to metazoan AFF family members. Like the metazoan SEC, this *S. pombe* ELL complex appears to function broadly in Pol II transcription. Interestingly, it appears to have a particularly important role in regulating genes involved in cell separation.

INTRODUCTION

Pol II elongation factor ELL was initially identified in rat liver nuclear extracts and purified by its ability to stimulate the overall rate of elongation by suppressing transient pausing by polymerase at many sites along the DNA (1). Two additional ELL family members, ELL2 (2) and ELL3 (3), were subsequently identified in mammals. Although ELL family members can bind to Pol II and stimulate elongation *in vitro* in the absence of any other proteins, their activities are markedly increased when they are bound by an activator protein referred to as an ELL-associated factor (in mammals, EAF1 or EAF2) (4–6).

More recently, ELL was found to be present in mammalian cells in several multiprotein complexes with roles in regulating Pol II transcription. One type of ELL-containing complex, referred to as the Super Elongation Complexes or SECs, includes an ELL protein, an EAF protein, the cyclin-dependent kinase P-TEFb (Cdk9/Cyclin T), either of the YEATS domain proteins AF9 or ENL, and an AF4/FMR (AFF) family protein (AFF1, AFF2, AFF3, or AFF4) (7–11), which acts as a scaffold upon which the other SEC components assemble (12). SECs regulate Pol II promoter-proximal pausing and transcription elongation of a large number of messenger and other RNAs (13). Notably, the genes encoding multiple SEC subunits, including ELL, AF9, ENL and AFF family members, have been found as translocation partners of the MLL gene in leukemia, and it has been proposed that mis-regulation of Pol II transcription elongation by the resulting mutant SECs underlies oncogenesis (1,7,14).

Although orthologs of SEC subunits are found in all metazoa, it has not been clear whether SEC-like complexes are present in lower eukaryotes, where Pol II transcription programs are generally less complex. In a previous study, we exploited bioinformatics to identify proteins that share weak sequence similarity with ELL and EAF in *Schizosaccharomyces pombe* and many other fungi, but not in *Saccharomyces cerevisiae*. In biochemical experiments, we showed that the potential *S. pombe* ELL and EAF orthologs (Ell1 and Eaf1, respectively) act similarly to their metazoan counterparts to stimulate elongation *in vitro* by *S. pombe* Pol II

---

*To whom correspondence should be addressed. Tel: +1 816 926 4091; Email: jlc@stowers.org
Present address: Dana M. Gibbon, Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR 97331, USA.

(15). P-TEFb and AF9/ENL family proteins also have apparent orthologs in *S. pombe*: Cdk9 and Pch1 are *S. pombe* orthologs of the P-TEFb subunits Cdk9 and Cyclin T, while Tfg3, found as a component of the transcription initiation and elongation factor TFIIF, the initiation factor TFIID, and the chromatin remodeler SWI/SNF, is a YEATS domain protein similar to AF9 and ENL (16–19). However, to date, no AFF family members have been identified in fungi.

In light of these findings, we sought to determine whether *S. pombe* Ell1 and Eaf1 might be present in cells as components of an SEC-like complex. To do so, we purified Ell1 and Eaf1 and exploited MudPIT mass spectrometry to identify associated proteins. These efforts resulted in discovery of an ELL-containing complex composed of Ell1, Eaf1 and a previously uncharacterized sequence orphan we refer to as Ell1 binding protein 1 (Ebp1), which is distantly related in sequence to the metazoan AFF family subunits of the SEC complex. Like the metazoan SEC, this ELL-containing complex appears to function broadly in Pol II transcription; however, it appears to have a particularly important role in modulating expression of genes involved in cell separation. Taken together, our findings identify and provide functional insight into a likely evolutionary precursor of the metazoan SEC complex.

MATERIALS AND METHODS

Strains

The *S. pombe* strains used are listed in Table 1. Cells were grown at 32°C in rich medium (YES) supplemented with adenine, histidine, leucine, and uracil (225 μg/ml). The *ell1Δ*, *eaf1Δ* and *ebp1Δ* strains were generated by replacing the coding region of the gene of interest with the kanMX6 or natMX6 marker. To generate the 972 h- strains used for mass spectrometry analyses, Ell1 or Eaf1 were C-terminally FLAG-tagged by replacing their stop codons with sequences encoding 3× FLAG followed by a stop codon and the NatMX6 marker (20).

Antibodies

Rabbit polyclonal antibodies were raised against purified, recombinant Ell1, Eaf1 and Ebp1 (Bio-Synthesis, Inc.) and purified using Melon™ Gel (ThermoFisher Scientific) according to the manufacturer's instructions. Anti-Pol II (4H8) was obtained from Abcam (ab5408). Mouse monoclonal anti-Myc (M4439) and anti-HA (H9658), and rabbit anti-FLAG (F7425) antibodies, as well as anti-FLAG agarose (A2220) and anti-Myc agarose (E6654) beads, were from SIGMA. Alexa Fluor 680-goat anti-mouse and Rabbit IgG DyLight™ 800 antibodies were from Invitrogen (A21058) and Rockland Antibodies (611-745-127), respectively.

Analysis of proteins associated with FLAG-tagged Ell1 and Eaf1 in *S. pombe*

Cells were grown at 32°C in rich medium (YES) supplemented with adenine, histidine, leucine and uracil (225 μg/ml) and harvested at mid-log phase. Preparation of whole cell lysates and anti-FLAG immunopurification of

Ell1- or Eaf1-associated proteins, as well as no-Tag negative control, were performed in biological duplicates as described (21). FLAG-immunopurified proteins were treated with benzonase and TCA precipitated before analysis by multidimensional protein identification technology (MudPIT) as described (22,23). Tandem mass (MS/MS) spectra were interpreted using SEQUEST against a database of *S. pombe* proteins (downloaded from NCBI on March 7, 2017), and complemented with 177 sequences from usual contaminants (human keratins, IgGs, proteolytic enzymes). To estimate false positive discovery rates (FDRs), each sequence was randomized keeping amino acid composition and length the same, and the resulting 'shuffled' sequences were added to the 'normal' database (doubling its size) and searched at the same time. Peptide/spectrum matches were sorted and selected using DTASelect with the following criteria set: spectra/peptide matches were only retained if they had a DeltaCn of at least 0.08, and minimum XCorr of 1.8 for singly, 2.0 for doubly, and 3.0 for triply charged spectra. To be considered, peptides had to be fully tryptic and at least 7 amino acids long. Peptide hits from multiple runs were compared using CONTRAST (24). Combining all runs, proteins had to be detected by at least 2 such peptides or by 1 peptide with 2 independent spectra. Under these criteria, FDRs were on average 0.1, 0.2 and 0.76% at the spectra, peptide, and protein levels, respectively.

Expression and purification of recombinant proteins in insect cells

cDNAs encoding wild-type Ell1, Eaf1 and Ebp1 containing N-terminal FLAG, HA or c-Myc epitope tags were subcloned into pBacPAK8. Recombinant baculoviruses were generated with the BacPAK expression system (Clontech) using pBacPAK6 viral DNA prepared as described (25). Sf9 insect cells were cultured at 27°C in Sf-900 II SFM (Invitrogen). Flasks containing 1×10^8 Sf9 cells were infected with the recombinant baculoviruses. Forty-eight hours after infection, cells were collected and lysed in 15 ml of ice-cold buffer containing 50 mM HEPES–NaOH, pH 7.9, 300 mM NaCl, 5 mM $MgCl_2$, 0.2% Triton-X-100 and 20% (v/v) glycerol with 1% (v/v) protease inhibitor cocktail (Sigma P8849). Lysates were centrifuged at 40 000 rpm for 30 min at 4°C. Lysates from 1×10^8 cells were incubated with 0.5 ml anti-FLAG (M2) or anti-Myc agarose beads overnight at 4°C. The beads were washed three times with TRIS-buffered saline (TBS, 25 mM Tris–HCl, pH 7.4, 137 mM NaCl, 2.7 mM KCl). To prepare proteins for transcription assays or antibody generation, bound proteins were eluted at 4°C with TBS containing 10% (v/v) glycerol and 0.3 mg/ml 3× FLAG peptide. For IP-western assays, proteins were eluted at room temperature with 0.3 mg/ml 3× FLAG or Myc peptides, analyzed by western blotting, and visualized using a LI-COR Biosciences Odyssey Infrared Imager.

ChIP-seq

To prepare chromatin sufficient for ∼4 ChIP-seq assays, *S. pombe* strains were grown in 200 ml YES to mid log phase. Cells were crosslinked by adding formaldehyde directly to cultures to a final concentration of 1% and incubating for

Table 1. *S. pombe* strains used in this study

Strain	Genotype		Used in Fig
972	*h-*	This study	Table 2
Ell1-FLAG (972)	*h- ell1+::3Xflag natMX6*	This study	Table 2
Eaf1-FLAG (972)	*h- eaf1+:: 3Xflag natMX6*	This study	Table 2
PEM-2	*h- ade6-M210 leu1-32 ura4-D18 mat1_m-cyhS, smt0 rpl42::cyhR(sP56Q)*	*Roguev et al., 2007b)*	Figures 3, 4A, B, C and 5B, C
ell1∆ (PEM-2)	*PEM-2; ∆ell1::natMX6*	This study	Figures 3, 4A, B, C and 5B, C
eaf1∆ (PEM-2)	*PEM-2; ∆eaf1::natMX6*	This study	Figures 3, 4A and 5B, C
ebp1∆ (PEM-2)	*PEM-2; ∆ebp1::natMX6*	This study	Figures 3, 4A and 5B, C
ED666	*h+ ade6-M210 ura4-D18 leu1-32*	*Bioneer Inc.*	Figure 4C
ell1∆ (ED666)	*h+ ∆ell1::kanMX6 ade6-M210 ura4-D18 leu1-32*	*Bioneer Inc.*	Figure 4C
eaf1∆ (ED666)	*h+ ∆eaf1::kanMX6 ade6-M210 ura4-D18 leu1-32*	*Bioneer Inc.*	Figure 4C
ebp1∆ (Ed666)	*h+ ∆ebp1::kanMX6 ade6-M210 ura4-D18 leu1-32*	*Bioneer Inc.*	Figure 4C
PP138	*h- ade6-M216 leu1-32 ura4-D18 his3-D1*	This study	Figure 4D
ell1∆ (PP138)	*h- ell1::kanMX6 ade6-M216 leu1-32 ura4-D18 his3-D1*	This study	Figure 4D
Ell1 rescue (*PP138*)	*h- ell1::kanMX6 ade6-M216 leu1-32::ell1/leu1ura4-D18 his3-D1*	This study	Figure 4D

15 min at room temperature. Crosslinking reactions were quenched by addition of 10 ml of 2.5 M glycine for 15 min. Cells were harvested by centrifugation, washed twice with ice-cold phosphate-buffered saline (PBS), and resuspended in 2 ml of Buffer A (40 mM HEPES–KOH (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate) supplemented with 0.2% SDS and 1% (v/v) protease inhibitor cocktail. After addition of an equal volume of acid washed glass beads (Sigma G8772), cells were disrupted by vortexing for 60 min at 4°C using a Vortex-Genie with a Turbomix attachment. Lysates were recovered from the beads, the total volume was increased to ∼2 ml by addition of lysis buffer, and lysates were sonicated in a 15 ml conical centrifuge (Falcon) at 4°C using a Fisher Scientific 550 Sonic Dismembrator with a 1/8' probe (11 cycles 10 s on, 30 s off) to generate chromatin fragments in the range 200–800 bp. The sonicate was spun at 4°C, 14 000 rpm for 20 min, and the soluble chromatin was retained.

400 µl chromatin was mixed with 400 µl Buffer A with 1 mM AEBSF and 80 µl 10% sarkosyl. 50 µl of Dynabeads Protein G or Dynabeads M-280 sheep anti-mouse IgG were incubated for 2 h at 4°C with 10 µg of rabbit or mouse antibodies, respectively, and washed with Buffer A. Dynabead-bound antibodies were incubated overnight at 4°C with chromatin, and then washed sequentially once with Buffer A, twice with Buffer A with 500 mM NaCl, once with buffer containing 0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris–HCl, pH 8.0, and twice with 10 mM Tris–HCl, pH 8.0, 1 mM EDTA. Bound complexes were eluted from the beads by incubating with 400 µl of 10 mM Tris–HCl pH 8.0, 1 mM EDTA, 250 mM NaCl, 1% SDS at 65°C for 30 min. To prepare input samples, 60 µl of chromatin extract was RNase treated and combined with 140 µl TE, pH 8.0 (10 mM Tris–HCl, 1 mM EDTA), 200 µl proteinase K buffer (50 mM Tris–HCl, pH 8.0, 12.5 mM EDTA, 300 mM NaCl, 1% SDS). 2 µl 10 mg/ml Proteinase K (Roche) was added to ChIP and input samples and left overnight at 65°C. DNA was purified by phenol/chloroform extraction and ethanol precipitation.

Sequencing libraries were prepared using the KAPA HTP library preparation kit (Kapa Biosystems) and size selected with 2% Pippin Prep gels (Sage Science). Pooled libraries were sequenced using an Illumina HiSeq 2500 platform. 50 bp single reads were aligned to *S. pombe* ASM294v2 with

bowtie2 (2.2.9) with default settings (26). Aligned reads were coordinate sorted and indexed with samtools (v1.3.1). Reads with MAPQ < 10 were removed to filter out low quality and multi-mapped reads (27). Replicates of ChIPs performed with deletion strains *ell1∆*, *eaf1∆* and *ebp1∆* were merged using samtools merge. Broad peaks were called with macs2 callpeaks (2.1.1) with the following parameters: -nomodel -broad -f BAM -gsize 12100000 -B with a q-value cutoff of 0.1; merged deletion strain datasets were used as inputs for Ell1, Eaf1 and Ebp1 ChIPs, respectively (28). For Pol II (4H8) peak calling, a *q*-value cutoff of 0.01 was used. Peaks were assigned to the nearest gene using bedtools intersect with the following parameters: -a -b -wa- wb (29). Genes with associated peaks in only one replica and antisense, mitochondrial, rRNA, and tRNA genes were excluded from downstream analyses. ChIP-seq results were visualized using Integrative Genomics Viewer (IGV) v2.3 (30,31).

RNA-seq

200 ml cultures were grown to mid log phase (OD600 = 0.8–1.0), harvested by centrifugation, washed in ice-cold DEPC-treated water, pulverized in liquid nitrogen using a mortar and pestle, and re-suspended in 10 ml 50 mM sodium acetate (pH 5.2) with 1% SDS. After addition of 10 ml phenol: chloroform 5:1 (equilibrated with sodium acetate, pH 5.2), samples were vortexed briefly and incubated at 65°C in a water bath for 2 min. Samples were incubated on ice for 1 minute and centrifuged at 7500g for 5 min. After an additional four extractions with phenol:chloroform, RNA was ethanol precipitated and purified using RNeasy mini spin columns (Qiagen). RNA quality was assessed using a Bioanalyzer. PolyA-selected mRNA libraries were prepared using the TruSeq Standard mRNA LT Sample Prep Kits, size-selected with 2% Pippin Prep gels, and sequenced using an Illumina HiSeq 2500 platform with single read lengths of 50 bp. Reads were aligned to *S. pombe* ASM294v2 with TopHat (v2.1.1) with default settings (32). Aligned reads were coordinate sorted and indexed with samtools (v1.3.1) (27). A counts table for unambiguous uniquely mapped reads with MAPQ score >10 was generated. Genes with an absolute $\log_2$ fold change of 0.58 or more and a Benjamini-Hochberg false discovery rate <0.05 were identified as differentially expressed.

Ell1 rescue

The Ell1 ORF was cloned into the pDUAL vector and inserted at the leu1 locus in *ell1Δ* (*PP138*) to generate the Ell1 rescue strain. 1 μg of total RNA from 3 independent biological replicates of wild type (*PP138*), *ell1Δ* (*PP138*), and the rescue strain was reverse transcribed with an oligo(dT) primer, and the cDNA was analyzed by qPCR. RNA levels were calculated relative to the levels of actin (*act1+*) using the $\Delta\Delta Ct$ function in the iQ5 analysis software (Bio-RAD). Details of RNA preparation and qPCR are provided in Supplemental Material.

Assay of cell separation defects

Cell separation phenotypes were assessed essentially as described (33), except that cells were stained with 8 μg/ml propidium iodide and 10 nM calcofluor white, both from Sigma. Prior to staining, cells were grown in EMM supplemented with 225 mg/L each of adenine, histidine, leucine, lysine, and uracil and grown to an OD600 of 0.4–0.5. Imaging was performed on a Nikon Eclipse Ti with a CSU-W1 Yokogawa disc. A 405/488/561/640 dichroic was used. 405 nm laser light was used to excite Calcofluor White, and 561 nm laser light was used to excite propidium iodide. Emission for Calcofluor white and propidium iodide was collected through 435–485 nm and 590–650 nm BP filters, respectively, using a Hamamatsu ORCA Flash 4.0 sCMOS camera. A Nikon 60× ApoTIRF 1.49 NA oil objective was used. Z-stacks were acquired with a spacing of 0.7 microns, and a max projection was applied prior to image analysis. At least 500 cells per genotype were scored manually for the number of septa per cell. Samples were blinded before imaging and only unblinded after determining the number of cells with no, one or two or more septa.

RESULTS AND DISCUSSION

Ell1 and Eaf1 copurify with a previously uncharacterized gene product distantly related to AFF family proteins

To identify Ell1- and Eaf1-associated proteins, we introduced sequences encoding C-terminal 3X FLAG tags into the *ell1+* and *eaf1+* genes by homologous recombination. Ell1-FLAG and Eaf1-FLAG interacting proteins were purified from the resulting *S. pombe* strains by anti-FLAG immunopurification and analyzed by mass spectrometry using multi-dimensional protein identification technology (MudPIT). Both Ell1 and Eaf1 copurified with each other and with a previously uncharacterized sequence orphan encoded by the gene *SPAC6G9.15c* (Table 2). The predicted SPAC6G9.15c ORF encodes a 498-amino acid protein; sequencing of a cDNA generated by 5-RACE confirmed the

predicted SPAC6G9.15c ORF. Because of its association with Ell1 and Eaf1, we have designated the SPAC6G9.15c gene product 'Ell1 binding protein 1' or 'Ebp1'.

We sought to identify potential Ebp1 orthologs in other species. In initial BLAST searches, we identified proteins orthologous to the complete Ebp1 sequence only in *S. octosporus* and *S. cryophilus*; however, proteins with sequences similar to the carboxy-terminal half of Ebp1 were found in S. *japonicus* and numerous other fungal species, but not in *S. cerevisiae*. To identify potential orthologs in more distantly related species, we used PSI-BLAST to search NCBI reference proteins, using the most highly conserved, C-terminal half of Ebp1 as query. By the third PSI-BLAST iteration, AFF family members from *Aedes aegypti* and *Aedes albopictus* appeared with a borderline E-value of 0.7, and by the fourth iteration, numerous AFF family members, including insect and vertebrate AFF4s, were retrieved with E-values ranging from 0.057 (*Tetranychus urticae*) to 9.7 (*Kryptolebias marmoratus*).

AFF family proteins from higher eukaryotes share an approximately 250 amino acid, C-terminal signature sequence (Figure 1A). Of note, the region of similarity between Ebp1 and AFF4 proteins was limited to the C-terminal signature sequence (Figure 1B, alignment). Although the molecular function of the C-terminal signature sequence remains poorly defined, it has been reported to contribute to transformation of myeloid progenitors by MLL-AFF4 chimeras (14).

Outside of the conserved C-terminal region, AFF proteins are predicted to be largely disordered (34). Sequence based secondary structure predictions revealed that like AFF proteins, the amino terminal portion of Ebp1 is largely unstructured (Figure 1C). Interestingly, the predicted disorder profiles of Ebp1 and AFF4 were strikingly similar within their C-termini (Supplementary Figure S1).

An Ebp1-containing ELL complex reconstituted with recombinant proteins

To characterize interactions between Ell1, Eaf1, and Ebp1, we used baculovirus vectors to express in insect cells various combinations of Ell1, Eaf1 and Ebp1 epitope tagged with FLAG, Myc, or HA, respectively, and performed anti-FLAG or anti-Myc agarose chromatography. Ell1 copurified with both Eaf1 and Ebp1 when all three proteins were expressed together or when Ell1 was expressed with either Eaf1 or Ebp1 (Figure 2A), indicating that Ell1 can interact independently with Eaf1 and Ebp1. In contrast, an interaction between Eaf1 and Ebp1 was detected only in cells that also expressed Ell1 (Figure 2B), arguing that Eaf1 and Ebp1 are linked through Ell1 (Figure 2C).

Table 2. Ell1 and Eaf1 copurify with a previously uncharacterized protein, Ebp1. *S. pombe* cell lysates were subjected to anti-FLAG agarose immunochromatography, and the resulting proteins were analysed by MudPIT mass spectrometry. The table shows the number of peptides (P) and spectra (S) for Ell1, Eaf1, and Ebp1 detected in samples purified from wild type cells (No-Tag Control) or cells expressing either Ell1-FLAG or Eaf1-FLAG

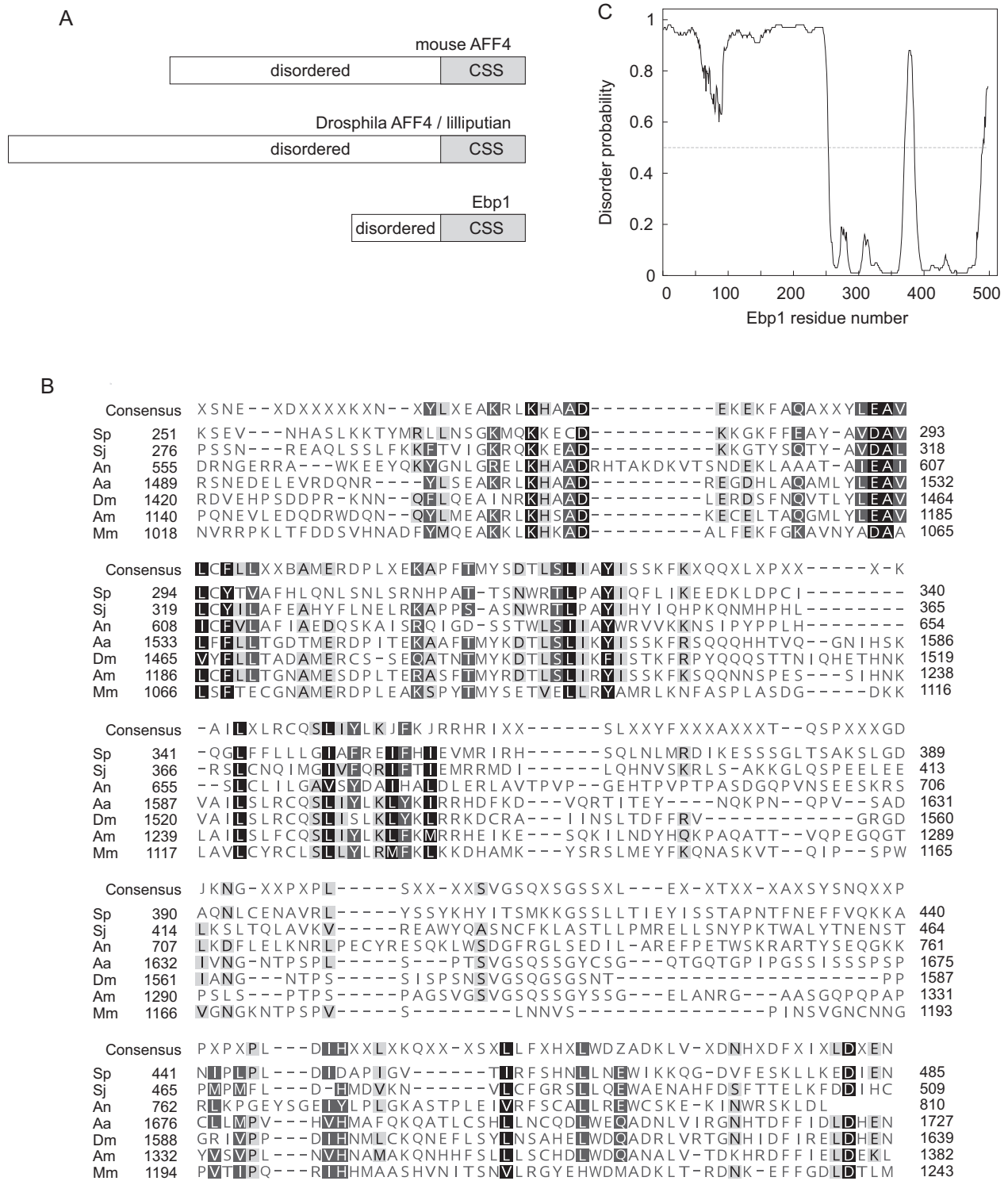|	Ell1- Flag IP				Eaf1-Flag IP				No-Tag control			
	Replicate 1		Replicate 2		Replicate 1		Replicate 2		Replicate 1		Replicate 2	
	P	S	P	S	P	S	P	S	P	S	P	S
Ell1	38	419	46	900	34	392	34	759	0	0	0	0
Eaf1	21	261	17	449	19	393	25	563	0	0	0	0
Ebp1	32	163	28	308	21	394	14	121	0	0	0	0

Figure 1. Similarities between Ebp1 and AFF4 proteins from higher eukaryotes. (**A**) Schematic representation of AFF4 proteins from mouse and *Drosophila melanogaster* and Ebp1. CSS, conserved signature sequence. (**B**) Multiple sequence alignment of the C-terminus of Ebp1 from *S. pombe* (Sp, NP_594124.1), *Schizosaccharomyces japonicus* (Sj, XP_002173449.1), and *Aspergillus nidulans* (An, XP_662712.1) with the C-terminal conserved signature sequences from AF4/FMR family members from *Aedes aegypti* (Aa, XP_021701574.1), *Drosophila melanogaster* (Dm, NP_001137775.1), *Apis mellifera* (Am, XP_016773562.1), and mouse (Mm, NP_032058.2). An initial alignment was generated using the Constraint-based Multiple Alignment Tool (COBALT; https://www.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi) with default parameters. The alignment was then manually adjusted using the Genious protein sequence alignment editing tool (Biomatters Inc.). (**C**) Predicted disorder profile of Ebp1, calculated using DISOPRED3 (43) at the PSIPRED Protein Sequence Analysis Workbench (44).
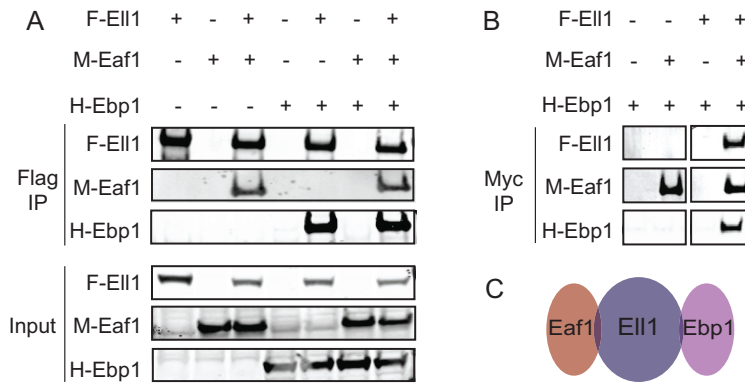
Figure 2. Ell1 bridges Ebp1 and Eaf1. Lysates from Sf9 cells, coinfected with baculoviruses encoding FLAG-Ell1, Myc-Eaf1 and HA-Ebp1 in the combinations indicated in the figure, were subjected to immunoprecipitations with (**A**) anti-FLAG and (**B**) anti-Myc antibodies. Bound proteins were eluted and analyzed by Western blotting using the indicated antibodies. Inputs (4%) and IPs shown in panel A are from the same exposures of the same Western blots. The IP samples shown in panel **B** were analyzed on a different immunoblot from those shown in panel A. (**C**) Schematic representation of *S. pombe* ELL complex.

The Ell1-Eaf1 complex has been shown previously to stimulate Pol II elongation *in vitro* (15). To ask whether Ebp1 affects Ell1-Eaf1's ability to stimulate elongation, we performed *in vitro* transcription assays using recombinant proteins. Under the conditions used in our assays, Ebp1 did not stimulate elongation by Pol II, either in the presence or absence of Ell1-Eaf1 (Supplementary Figure S2).

Ell1, Eaf1, and Ebp1 are co-recruited to genes in vivo

If Ell1, Eaf1 and Ebp1 function in cells as part of the same multi-protein complex one would expect them to co-occupy the same genomic loci. We therefore generated polyclonal antibodies specific for *S. pombe* Ell1, Eaf1, and Ebp1 and used them in ChIP-seq assays to measure Ell1, Eaf1 and Ebp1 occupancy genome-wide in wild type *S. pombe*. To control for the specificities of these antibodies, we also performed ChIP-seq assays in *ell1Δ*, *eaf1Δ*, and *ebp1Δ* strains. In addition, to assess the degree of overlap between ELL complex and Pol II occupancies, we also performed Pol II ChIP-seq.

We observed many regions of overlapping Ell1, Eaf1 and Ebp1 enrichment, most of which also overlapped with peaks of Pol II. Examples are shown in Figure 3A. Validating the specificities of the antibodies used for these experiments, Ell1, Eaf1, and Ebp1 signals were dramatically reduced in ChIPs performed using the *ell1Δ*, *eaf1Δ* and *ebp1Δ* strains, respectively (data from *ell1Δ*, *eaf1Δ* or *ebp1Δ* strains, respectively, is overlaid in lighter shades on the darker ChIP signals from wild type cells in Figure 3A). To assess more rigorously the overlap of Ell1, Eaf1 and Ebp1 occupancies, we used the peak finding algorithm, MACS2 (Model-based Analysis of ChIP-seq) to identify genes with peaks that exhibit significant enrichment in wild type *S. pombe* compared to the deletion strains. As shown in Figure 3B, there was a very high degree of overlap between Ell1-, Eaf1- and Ebp1-occupied genes. Of note, 752 of 762 Ebp1-occupied genes were also occupied by both Ell1 and Eaf1. In addition, nearly all Ell1- and Eaf1-occupied genes overlapped with genes occupied by Pol II (Figure 3C and Supplementary Figure S3).

As described earlier, in higher eukaryotes, ELL and EAF are found as part of the multiprotein SEC, which also includes P-TEFb (Cdk9/Cyclin T). Although neither Cdk9 nor Cyclin T (Pch1) copurified with the ELL complex from *S. pombe* under the conditions we used, we wished to determine whether they occupy common genomic loci. We therefore compared our ChIP-seq datasets to a published *S. pombe* Cdk9 ChIP-chip dataset (35). As shown in Figure 3A, D, and Supplementary Figure S3 we find that occupancies of Ell1, Eaf1 and the AFF family protein-like Ebp1 protein correlate remarkably well with that of Cdk9, raising the possibility that the *S. pombe* ELL complex could function with P-TEFb as part of a rudimentary SEC.

A common set of genes targeted by Ell1, Eaf1 and Ebp1

By visual inspection, we noted that some genes, such as *sme2*, *ace2* and *adg1*, appeared to exhibit a higher than average enrichment of Ell1, Eaf1 and Ebp1 relative to Pol II (Figure 4A). Local regression analysis revealed a positive correlation between Ell1 and Pol II occupancies ($r^2 = 0.6$; Figure 4B); however, it also identified a set of genes that were more highly enriched for Ell1 (Figure 4B, red and Supplementary Table S1). Similar patterns were observed for Eaf1 and Ebp1 (Supplementary Table S1 and Supplementary Figure S4).

To assess the relationship between ELL complex occupancy and Ell1-, Eaf1-, and Ebp1-dependent effects on gene regulation, we performed RNA-seq and qPCR-based assays. Only a small number of genes were differentially expressed in *ell1Δ*, *eaf1Δ* or *ebp1Δ* cells compared to wild type. However, we observed that genes exhibiting high Ell1 occupancy relative to Pol II were modestly, but significantly, more down-regulated in *ell1Δ* cells in multiple strain backgrounds (Figure 4C and D). Importantly, the same group of genes also tended to be down-regulated in both *eaf1Δ* and *ebp1Δ* cells, identifying a common set of genes targeted by components of the ELL complex (Figure 4C).

To explore cellular pathways impacted by the ELL complex, we queried the STRING (Search Tool for the Retrieval of Interacting Genes) database (36) to determine
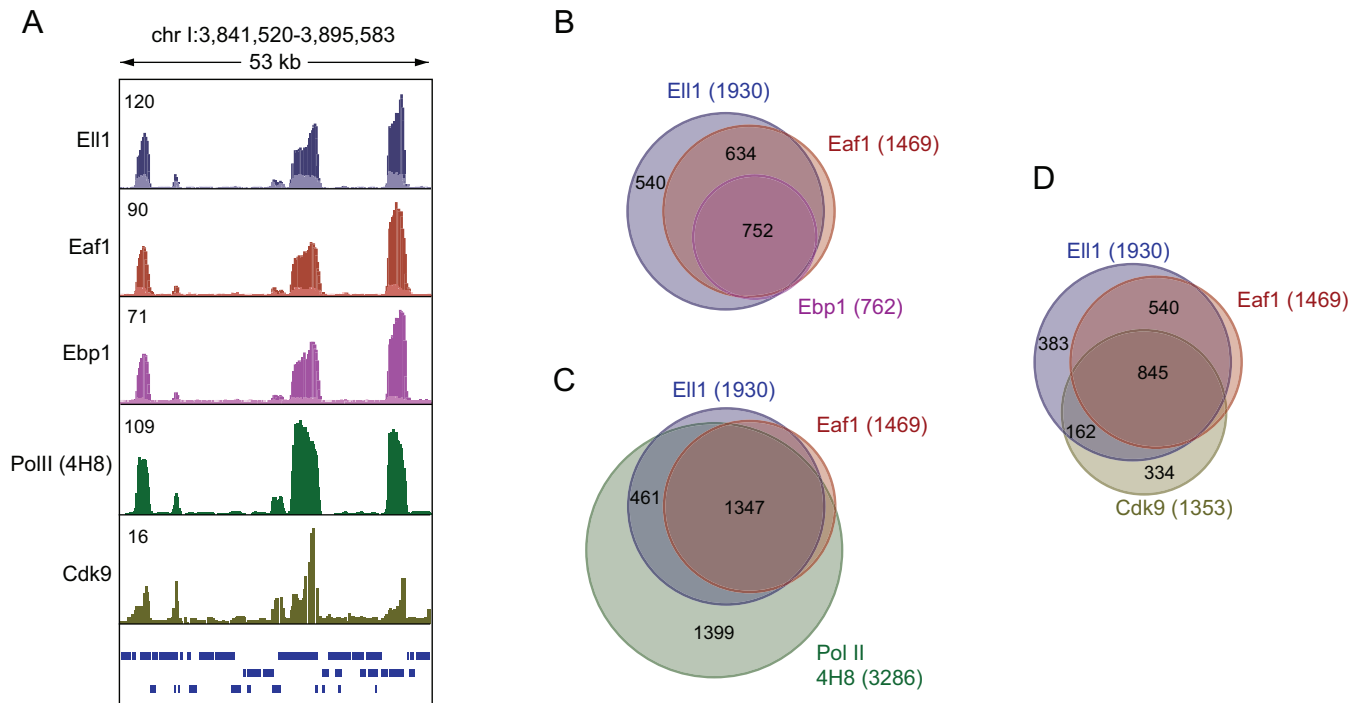
Figure 3. A rudimentary SEC-like complex in *S. pombe*. (**A**) Representative IGV browser shots of Ell1, Eaf1, Ebp1, Pol II (4H8), and Cdk9 ChIP signals. In the Ell1, Eaf1, and Ebp1 tracks, data from *ell1Δ*, *eaf1Δ*, or *ebp1Δ* strains, respectively, is overlaid in lighter shades on the darker ChIP signals from wild type cells. Cdk9 ChIP-chip data (GSE16498) is from (35). The blue bars at the bottom represent genes. Venn diagram showing overlap between genes occupied by Ell1, Eaf1, and (**B**) Ebp1, (**C**) Pol II and (**D**) Cdk9 in wild type *S. pombe*.

whether the proteins encoded by this group of genes are members of shared physical and/or functional interaction networks. This analysis identified several clusters of inter-acting nodes (Figure 5A). The most striking of these is made up of a collection of co-expressed genes that are referred to as the cell separation or eng1 gene cluster (37) and in-cludes many genes of the Sep1-Ace2 pathway (38). Tran-scription of genes encoding the cell cycle transcription fac-tor Ace2, the β-glucosidase Exg1, and the sequence orphan SPCC306.11 is activated by the forkhead family transcrip-tion factor Sep1 (39,40). Ace2 in turn activates expression of genes encoding the glycosyl hydrolases Agn1 and Eng1, the cell cycle glycoproteins Adg1 and Adg2, β-glucosidase Adg3, chitin synthase regulatory protein Chf4, and the me-dial ring protein Mid2 ((33) and references therein). These cell cycle regulated genes are expressed in late G2 and M phases and play key roles as regulators of cell separation (33,37,39).

That expression of transcripts of Ace2-Sep1 pathway genes is decreased in *ell1Δ*, *eaf1Δ*, and *ebp1Δ* cells pre-dicted that these mutants might exhibit cell separation de-fects. *ace2Δ* cells exhibit a strong cell separation phenotype in which nearly all cells have 1, 2 or more septa, while other mutants of the pathway often have substantially milder phe-notypes (33). Indeed, microscopic examination of these mu-tants revealed they had mild defects in cell separation (Fig-ure 5B and C). In particular, the percentage of septated cells was increased by ∼2–3-fold in all three mutants, similar to what is seen in *adg2* or *adg3* deletions, and consistent with a slight delay in digestion of septa during cell separation.

How might the ELL complex regulate the cell separation cluster of genes? Although we cannot rule out the possi-bility that ELL complex-dependent regulation of at least some genes in the cell separation cluster occurs indirectly *via* Ace2, our ChIP-seq data indicates that the ELL com-plex is highly enriched not only at *ace2* but also at the other genes of the cluster as might be expected if it directly regu-lates their transcription. In this regard, it is noteworthy that the promoters of Ace2-regulated genes include binding site motifs for other transcription factors, including members of the forkhead family (37,41). Indeed, recent studies suggest that the promoters of a number of Ace2-dependent genes are also bound and regulated by Sep1 (41). In light of these findings, we believe it is reasonable to propose that the ELL complex may directly regulate transcription of cell separa-tion genes by a mechanism that is at least in part indepen-dent of Ace2 itself.

CONCLUSION

In summary, in this report, we describe isolation from *S. pombe* and characterization of a Pol II elongation fac-tor ELL-containing complex that shares features with the metazoan SEC. In mammalian cells, the multisubunit SEC is assembled combinatorially from one of three ELL fam-ily members, one of two EAF family members, the cyclin-dependent kinase T-TEFb (Cdk9/Cyclin C), one of two AF9 family members, and one of four AFF family mem-bers, which provides the scaffold upon which the SEC as-sembles. The *S. pombe* SEC-like complex that we have iso-lated is markedly smaller than the metazoan SEC, with just
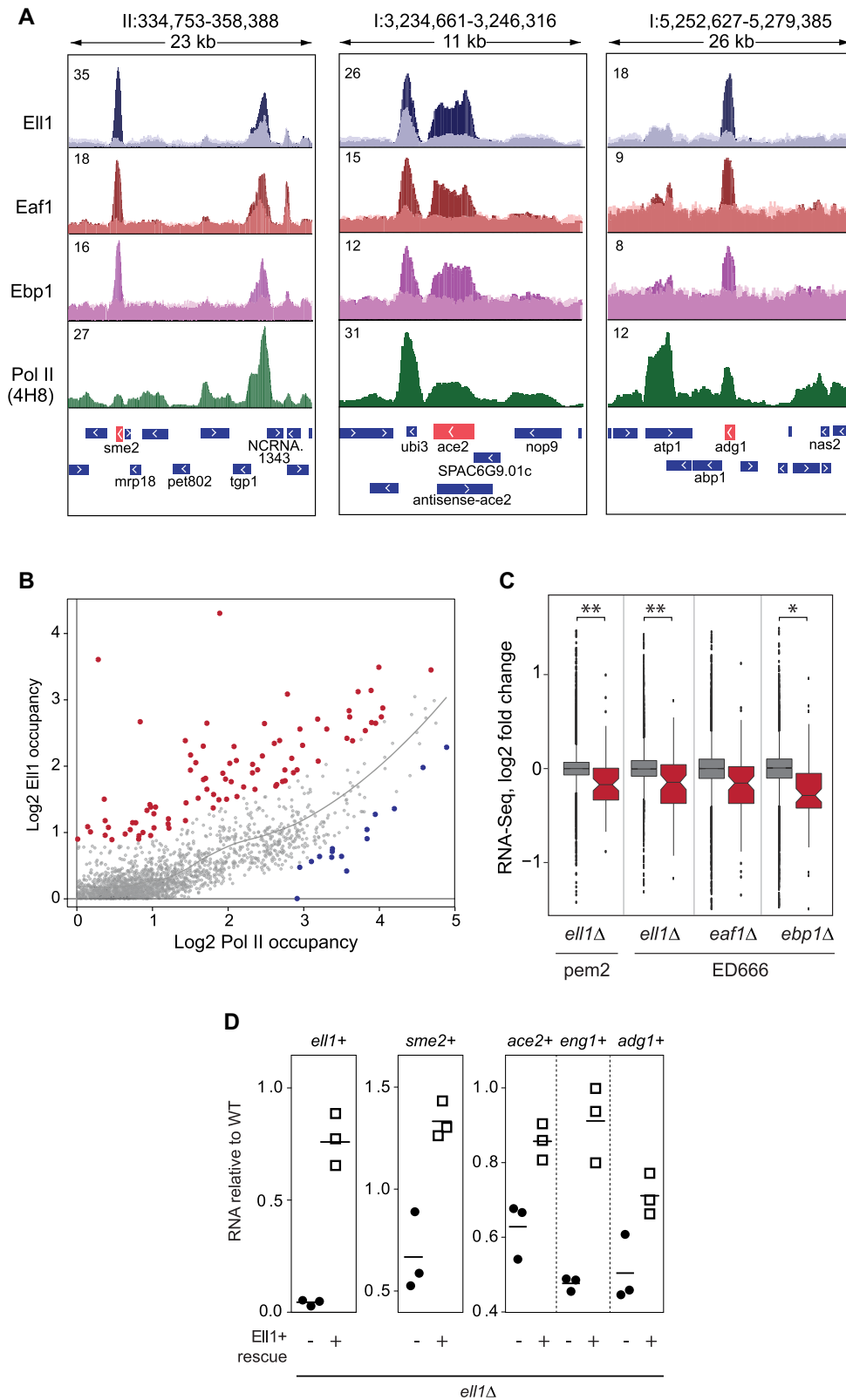
Figure 4. A common set of genes targeted by Ell1, Eaf1, and Ebp1. (**A**) IGV browser shots showing genes (highlighted in red) with high enrichment of Ell1, Eaf1, and Ebp1 relative to Pol II. (**B**) LOESS correlation between Ell1 and Pol II occupancies. Ell1 occupancy is defined as ChIP/Input (wild type) divided by ChIP/Input (*ell1Δ*); Pol II occupancy is ChIP/input. Outliers are indicated in red (Studentized residual > 1.96) and blue (Studentized residual ≤ −1.96). (**C**) Genes with high Ell1 occupancy relative to Pol II (red; Studentized residual >1.96) tend to exhibit greater reductions in transcript abundance upon deletion of *ell1+*, *eaf1+*, or *ebp1+* than other genes (gray, Studentized residual ≤ 1.96). *$P < 10^{-2}$; **$P < 10^{-4}$. (**D**) Reintroduction of wild type Ell1 into *ell1Δ* cells rescues gene expression. Levels of the indicated transcripts relative to the *act1+* transcript were determined by RT-qPCR.
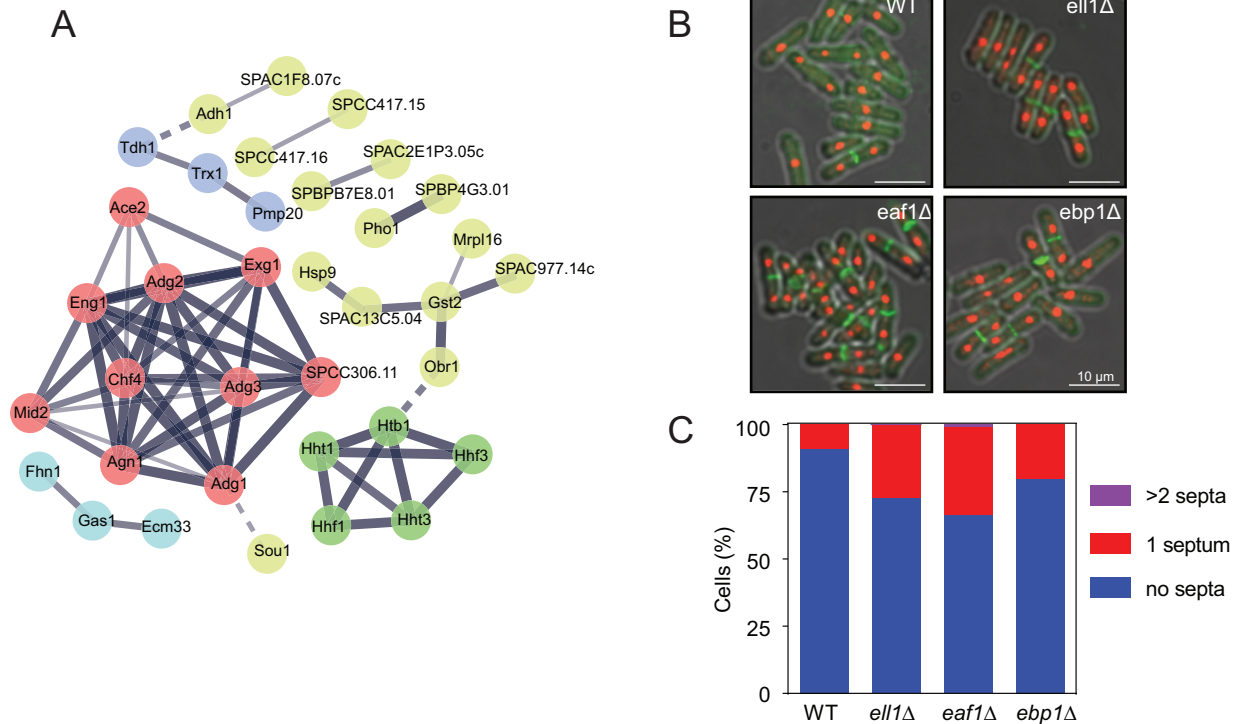
Figure 5. The ELL complex regulates cell separation genes. (**A**) Genes from the Sep1-Ace2 pathway are enriched among genes with high Ell1 occupancy relative to Pol II. Genes exhibiting Studentized residual >1.96 were submitted for comparison to the STRING database of known and predicted protein-protein interactions. Interactions were limited to neighborhood, experiments, co-occurrence, and co-expression. The width of the network edges reflects confidence scores. Colors represent kmeans clustering. (**B**) Representative microscopic images of wild type, *ell1Δ*, *eaf1Δ*, and *ebp1Δ* strains, stained with propidium iodide (red) to visualize nuclei and calcofluor white (green) to visualize septa. Red and green fluorescent images are overlaid on transmitted light images. (**C**) Phenotypes of mutants lacking the *ell1+*, *eaf1+* or *ebp1+* genes. The fraction of cells with no septa (blue), 1 septum (red), or 2 or more septa was determined by counting more than 500 cells for each strain.

three subunits: *S. pombe* orthologs of the ELL and EAF proteins, and a previously uncharacterized sequence orphan we designate Ebp1, which is distantly related to metazoan AFF proteins. We find that the Ell1, Eaf1 and Ebp1 proteins closely co-localize throughout the genome in ChIP-seq experiments and regulate transcription of an overlapping set of genes. Thus, although Ebp1 does not affect Ell1-Eaf1 activity in *in vitro* assays of transcription elongation, and its precise function remains to be determined, our data support the notion that the three subunits of the *S. pombe* SEC-like complex indeed function together in cells to regulate Pol II transcription.

Even though *S. pombe* expresses apparent orthologs of P-TEBb and AF9 (referred to as Cdk9/Pch1 and Tfg3), they did not copurify with the *S. pombe* ELL complex. We did, however, find that *S. pombe* P-TEFb co-localizes with the Ell1, Eaf1, and Ebp1 proteins across the genome, consistent with the possibility that these proteins could function together to regulate Pol II transcription much like their metazoan orthologs do as integral subunits of the SEC. It is perhaps not surprising that the *S. pombe* Af9 ortholog Tfg3 is not a component of the Ell1-containing, SEC-like complex, because, unlike metazoan AF9, *S. pombe* Tfg3 is brought to the Pol II transcription complex *via* its stable interaction with TFIIF, of which it is an integral subunit.

Finally, we have identified a collection of Sep1 and Ace2-regulated cell separation genes as common targets for reg-

ulation by the *S. pombe* SEC-like complex. These genes are all transcribed at the end of the cell cycle, during septation, under circumstances where the precise timing of their transcription is critical and likely relies on maintaining optimal Pol II function for this brief period. In light of our finding that the Ell1, Eaf1 and Ebp1 proteins are not essential for expression of these cell separation genes, but, instead appear to augment their expression, it is possible that the true role of the *S. pombe* SEC-like complex is to 'fine-tune' Pol II transcription. In this regard, it is noteworthy that mammalian ELL is believed to perform such a fine-tuning function by simultaneously increasing the overall rate of Pol II elongation (1) and reducing the frequency of Pol II 'accidents,' such as backtracking followed by transcriptional arrest (42).

DATA AVAILABILITY

All high throughput sequencing data are available at (GEO Accession Number GSE115636). Other original data underlying this paper can be accessed from the Stowers Original Data Repository at http//www.stowers.org/research/publications/ LIBPB-1312.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Shilatifard,A., Lane,W.S., Jackson,K.W., Conaway,R.C. and Conaway,J.W. (1996) An RNA polymerase II elongation factor encoded by the human ELL gene. *Science*, **271**, 1873–1876.
2. Shilatifard,A., Duan,D.R., Haque,D., Florence,C., Schubach,W.H., Conaway,J.W. and Conaway,R.C. (1997) ELL2, a new member of an ELL family of RNA polymerase II elongation factors. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 3639–3643.
3. Miller,T., Williams,K., Johnstone,R.W. and Shilatifard,A. (2000) Identification, cloning, expression, and biochemical characterization of the testis-specific RNA polymerase II elongation factor ELL3. *J. Biol. Chem.*, **275**, 32052–32056.
4. Kong,S.E., Banks,C.A.S., Shilatifard,A., Conaway,J.W. and Conaway,R. (2005) ELL-associated factors 1 and 2 are positive regulators of RNA polymerase II elongation factor ELL. *PNAS*, **102**, 10094–10098.
5. Simone,F., Polak,P.E., Kaberlein,J.J., Luo,R.T., Levitan,D.A. and Thirman,M.J. (2001) EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein. *Blood*, **98**, 201–209.
6. Simone,F., Luo,R.T., Polak,P.E., Kaberlein,J.J. and Thirman,M.J. (2003) ELL-associated factor 2 (EAF2), a functional homolog of EAF1 with alternative ELL binding properties. *Blood*, **101**, 2355–2362.
7. Lin,C., Smith,E.R., Takahashi,H., Lai,K.C., Martin-Brown,S., Florens,L., Washburn,M.P., Conaway,J.W., Conaway,R.C. and Shilatifard,A. (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**, 429–437.
8. Sobhian,B., Laguette,N., Yatim,A., Nakamura,M., Levy,Y., Kiernan,R. and Benkirane,M. (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell*, **38**, 439–451.
9. He,N., Liu,M., Hsu,J., Xue,Y., Chou,S., Burlingame,A., Krogan,N.J., Alber,T. and Zhou,Q. (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**, 428–438.
10. Takahashi,H., Parmely,T.J., Sato,S., Tomomori-Sato,C., Banks,C.A.S., Kong,S.E., Szutorisz,H., Swanson,S.K., Martin-Brown,S., Washburn,M.P. *et al.* (2011) Human mediator subunit Med26 functions as a docking site for transcription elongation factors. *Cell*, **146**, 92–104.
11. Luo,Z., Lin,C., Guest,E., Garrett,A.S., Mohaghegh,N., Swanson,S., Marshall,S., Florens,L., Washburn,M.P. and Shilatifard,A. (2012) The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol. Cell Biol.*, **32**, 2608–2617.
12. Chou,S., Upton,H., Bao,K., Schulze-Gahmen,U., Samelson,A.J., He,N., Nowak,A., Lu,H., Krogan,N.J., Zhou,Q. *et al.* (2013) HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E123–E131.
13. Luo,Z., Lin,C. and Shilatifard,A. (2012) The super elongation complex (SEC) family in transcriptional control. *Nat. Rev. Mol. Cell Biol.*, **13**, 543–547.
14. Yokoyama,A., Lin,M., Naresh,A., Kitabayashi,I. and Cleary,M.L. (2010) A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell*, **17**, 198–212.
15. Banks,C.A., Kong,S.E., Spahr,H., Florens,L., Martin-Brown,S., Washburn,M.P., Conaway,J.W., Mushegian,A. and Conaway,R.C. (2007) Identification and characterization of a schizosaccharomyces pombe RNA polymerase II elongation factor with similarity to the metazoan transcription factor ELL. *J. Biol. Chem.*, **282**, 5761–5769.
16. Monahan,B.J., Villen,J., Marguerat,S., Bahler,J., Gygi,S.P. and Winston,F. (2008) Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nat. Struct. Mol. Biol.*, **15**, 873–880.
17. Kimura,M., Suzuki,H. and Ishihama,A. (2002) Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in Schizosaccharomyces pombe involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Mol. Cell. Biol.*, **22**, 1577–1588.
18. Kimura,M. and Ishihama,A. (2004) Tfg3, a subunit of the general transcription factor TFIIF in Schizosaccharomyces pombe, functions under stress conditions. *Nucleic Acids Res.*, **32**, 6706–6715.
19. Cairns,B.R., Lynn Henry,N. and Kornberg,R.D. (1996) TFG3/TAF30/ANC1, a component of the yeast SWI/SNF complex that is similar to the leukemogenic proteins ENL and AF-9. *Mol. Cell. Biol.*, **16**, 3308–3316.
20. Bahler,J., Wu,J.Q., Longtine,M.S., Shah,N.G., McKenzie,A. III, Steever,A.B., Wach,A., Philippsen,P. and Pringle,J.R. (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast*, **14**, 943–951.
21. Tsai,K.L., Yu,X., Gopalan,S., Chao,T.C., Zhang,Y., Florens,L., Washburn,M.P., Murakami,K., Conaway,R.C., Conaway,J.W. *et al.* (2017) Mediator structure and rearrangements required for holoenzyme formation. *Nature*, **544**, 196–201.
22. Florens,L. and Washburn,M.P. (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol. Biol.*, **328**, 159–175.
23. Washburn,M.P., Wolters,D. and Yates,J.R. III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.*, **19**, 242–247.
24. Tabb,D.L., McDonald,W.H. and Yates,J.R. III (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.*, **1**, 21–26.
25. Kitts,P.A. and Possee,R.D. (1993) A method for producing recombinant baculovirus expression vectors at high frequency. *BioTechniques*, **14**, 810–817.
26. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
27. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
28. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
29. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
30. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.*, **14**, 178–192.
31. Robinson,J.T., Thorvaldsdottir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
32. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
33. Alonso-Nunez,M.L., An,H., Martin-Cuadrado,A.B., Mehta,S., Petit,C., Sipiczki,M., del Rey,F., Gould,K.L. and de Aldana,C.R.

(2005) Ace2p controls the expression of genes required for cell separation in Schizosaccharomyces pombe. *Mol. Biol. Cell*, **16**, 2003–2017.

34. He,N., Chan,C.K., Sobhian,B., Chou,S., Xue,Y., Liu,M., Alber,T., Benkirane,M. and Zhou,Q. (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**, E636–E645.

35. Coudreuse,D., van Bakel,H., Dewez,M., Soutourina,J., Parnell,T., Vandenhaute,J., Cairns,B., Werner,M. and Hermand,D. (2010) A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in S. pombe. *Curr. Biol.*, **20**, 1053–1064.

36. Szklarczyk,D., Morris,J.H., Cook,H., Kuhn,M., Wyder,S., Simonovic,M., Santos,A., Doncheva,N.T., Roth,A., Bork,P. *et al.* (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.*, **45**, D362–D368.

37. Oliva,A., Rosebrock,A., Ferrezuelo,F., Pyne,S., Chen,H., Skiena,S., Futcher,B. and Leatherwood,J. (2005) The cell cycle-regulated genes of Schizosaccharomyces pombe. *PLoS Biol.*, **3**, e225.

38. Bahler,J. (2005) A transcriptional pathway for cell separation in fission yeast. *Cell Cycle*, **4**, 39–41.

39. Rustici,G., Mata,J., Kivinen,K., Lio,P., Penkett,C.J., Burns,G., Hayles,J., Brazma,A., Nurse,P. and Bahler,J. (2004) Periodic gene expression program of the fission yeast cell cycle. *Nat. Genet.*, **36**, 809–817.

40. Garg,A., Futcher,B. and Leatherwood,J. (2015) A new transcription factor for mitosis: in Schizosaccharomyces pombe, the RFX transcription factor Sak1 works with forkhead factors to regulate mitotic expression. *Nucleic Acids Res.*, **43**, 6874–6888.

41. Suarez,M.B., Alonso-Nunez,M.L., del Rey,F., McInerny,C.J. and Vazquez de Aldana,C.R. (2015) Regulation of Ace2-dependent genes requires components of the PBF complex in Schizosaccharomyces pombe. *Cell Cycle*, **14**, 3124–3137.

42. Elmendorf,B.J., Shilatifard,A., Yan,Q., Conaway,J.W. and Conaway,R.C. (2001) Transcription factors TFIIF, ELL, and elongin negatively regulate SII- induced nascent transcript cleavage by Non-arrested RNA polymerase II elongation intermediates. *J. Biol. Chem.*, **276**, 23109–23114.

43. Ward,J.J., Sodhi,J.S., McGuffin,L.J., Buxton,B.F. and Jones,D.T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.*, **337**, 635–645.

44. Buchan,D.W., Minneci,F., Nugent,T.C., Bryson,K. and Jones,D.T. (2013) Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res.*, **41**, W349–W357.