# SR Proteins Collaborate with 7SK and Promoter-Associated Nascent RNA to Release Paused Polymerase

Xiong Ji,[1] Yu Zhou,[2] Shatakshi Pandit,[2] Jie Huang,[1] Hairi Li,[2] Charles Y. Lin,[4] Rui Xiao,[2] Christopher B. Burge,[4] and Xiang-Dong Fu[1,2,3,*]
[1]State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China
[2]Department of Cellular and Molecular Medicine
[3]Institute of Genomic Medicine
University of California, San Diego, La Jolla, CA 92093-0651, USA
[4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
*Correspondence: xdfu@ucsd.edu
http://dx.doi.org/10.1016/j.cell.2013.04.028

## SUMMARY

**RNAP II is frequently paused near gene promoters in mammals, and its transition to productive elongation requires active recruitment of P-TEFb, a cyclin-dependent kinase for RNAP II and other key transcription elongation factors. A fraction of P-TEFb is sequestered in an inhibitory complex containing the 7SK noncoding RNA, but it has been unclear how P-TEFb is switched from the 7SK complex to RNAP II during transcription activation. We report that SRSF2 (also known as SC35, an SR-splicing factor) is part of the 7SK complex assembled at gene promoters and plays a direct role in transcription pause release. We demonstrate RNA-dependent, coordinated release of SRSF2 and P-TEFb from the 7SK complex and transcription activation via SRSF2 binding to promoter-associated nascent RNA. These findings reveal an unanticipated SR protein function, a role for promoter-proximal nascent RNA in gene activation, and an analogous mechanism to HIV Tat/TAR for activating cellular genes.**

## INTRODUCTION

The expression of protein-coding genes in mammalian genomes begins with the assembly of the preinitiation complex (PIC) that brings RNA polymerase II (RNAP II) to gene promoters, which has been long considered a major step in regulated gene expression (Lee and Young, 2000). However, after transcript initiation and promoter clearance, RNAP II frequently pauses near the transcription start site (TSS) on numerous genes, and regulated RNAP II pause release has now been recognized as a critical step in gene activation (Adelman and Lis, 2012).

Promoter clearance has been linked to phosphorylation on Ser5 in the heptapeptide repeat of the C-terminal domain (CTD) of the large subunit of RNAP II. This event is catalyzed by TFIIH (consisting of CDK7 and cyclin H) and allows the recruitment of the capping enzymes to protect the 5′ end of nascent RNA (Bentley, 2005). RNAP II is frequently paused within 20–40 nt downstream from the TSS, and its release requires the recruitment of P-TEFb (consisting of CDK9 and cyclin T), a kinase that is responsible for phosphorylating the negative elongation factor (NELF) and DRB-sensitive-inducing factor (DSIF), as well as RNAP II CTD at Ser2 and perhaps Ser5 positions (Czudnochowski et al., 2012). This series of events is correlated with RNAP II entry into the elongation phase of transcription (Saunders et al., 2006; Zhou et al., 2012).

A large body of literature indicates that P-TEFb is distributed in two separate pools in the nucleus (Peterlin and Price, 2006). One pool contains active P-TEFb associated with paused RNAP II in the promoter-proximal region, where a series of rearrangements eventually links the kinase to the superelongation complex (SEC) to initiate productive elongation (He et al., 2010; Lin et al., 2010; Sobhian et al., 2010; Takahashi et al., 2011). The other P-TEFb pool appears to be reversibly sequestered in the 7SK complex, a multisubunit ribonucleoprotein particle composed of the 7SK noncoding RNA, P-TEFb, the specific P-TEFb inhibitor protein HEXIM1, the La-like protein LARP7, and MePCE (Peterlin and Price, 2006).

Our current view of P-TEFb recruitment arises from studies on Tat-activated transcription on the HIV-1 promoter (Ott et al., 2011; Peterlin and Price, 2006). The HIV genome encodes a transcriptional transactivator, Tat, which binds to the transactivation response (TAR) element at the 5′ end of nascent viral RNA to release paused RNAP II at the HIV-1 promoter. In this process, Tat binding to TAR enhances P-TEFb recruitment from the nucleoplasm or directly from the 7SK complex to transcriptionally engaged RNAP II (Krueger et al., 2010; Ott et al., 2011). Despite a refined understanding of these events at a viral promoter, it has been unclear how P-TEFb is recruited to cellular gene promoters to activate transcription.

SR proteins are a family of RNA-binding proteins involved in both constitutive and regulated splicing (Lin and Fu, 2007) as well as in integrating multiple steps in RNA metabolism in mammalian cells (Zhong et al., 2009). Here, we show that a

unique SR protein SRSF2 (originally known as SC35) is associated with gene promoters as part of the 7SK complex, mediates the release of P-TEFb from the 7SK complex in an RNA-dependent manner, facilitates the recruitment of P-TEFb and other key transcription elongation factors to gene promoters, and activates transcription via promoter-proximal nascent RNA. These data reveal that SRSF2 functions like HIV Tat as a transcription activator and also assign an active role of short, promoter-associated RNA (Esteller, 2011) in acting like HIV TAR to activate transcription.

## RESULTS

### SRSF2 Is Preferentially Recruited to Active Gene Promoters

Our previous work demonstrated that SRSF2 plays an active role in transcription elongation in addition to its traditional function in RNA splicing (Lin et al., 2008). SRSF2 is the only SR protein retained in the nucleus, likely due to its intimate association with genomic DNA, consistent with its role in transcription (Sapra et al., 2009). To understand its mechanism in transcription, we performed chromatin immunoprecipitation sequencing (ChIP-seq) analysis of SRSF2 in mouse embryonic fibroblasts (MEFs) derived from conditional knockout mice (Lin et al., 2005). In these cells, the endogenous gene is replaced by an HA-tagged SRSF2 gene expressed from a tet-off promoter, permitting efficient depletion of the protein with the tet analog Dox (Figure S1A available online). We previously showed that HA-SRSF2 is expressed at a comparable level to and provides all essential functions of the endogenous gene (Lin et al., 2005). As a control, we performed parallel analyses on another SR protein (SRSF1) using a similarly constructed MEF line. Both endogenous SRSF1 and SRSF2 could also be detected with specific antibodies (Figure S1B), allowing for validation of anti-HA antibody-generated data when necessary (see below).

The ChIP-seq analysis unexpectedly revealed an abundance of sequence tags near gene promoters for both SRSF1 and SRSF2 (Figures 1A and 1B), which is not observed with total input DNA from sonicated chromatin or by anti-HA ChIP-seq from MEFs not expressing any HA-tagged protein (data not shown). SRSF2 is more frequently associated with gene promoters than SRSF1 (Figure 1A), though the tag densities at their binding sites are globally concordant (Figure S1C). We validated the association of both SR proteins on a large panel of gene promoters by ChIP-qPCR (Figure S1D).

### SRSF1 and SRSF2 Associate with Distinct Sets of DNA and RNA Sequences

The association of SR proteins with gene promoters might reflect early function in interacting with nascent RNA to facilitate cotranscriptional RNA processing. To test this possibility, we mapped their interactions with RNA by UV-crosslinking immunoprecipitation in the same cells using the same anti-HA antibody, followed by deep RNA sequencing (CLIP-seq, see Pandit et al., 2013). Strikingly, ChIP-seq and CLIP-seq revealed completely distinct profiles on DNA and RNA, as illustrated on the *hnRNPH1* gene (Figure 1B). This distinction is also evident from the meta-

analysis, showing that SRSF2 primarily associates with DNA at gene promoters but with RNA on internal exons in gene bodies (Figure 1C), where its well-defined role in exon inclusion is executed. We further note that SRSF2 crosslinks to RNA near the TSS but with a lower efficiency compared to internal exons (Figure 1C, note the scale difference in the y axes), indicative of complex interplay with DNA and RNA in the promoter-proximal regions. SRSF1 exhibited essentially identical patterns in these analyses (data not shown).

The functional significance of the prevalent association of SRSF1 and SRSF2 with DNA is evidenced by the positive correlation with levels of gene expression determined by RNA-seq (Figure 1D), which is consistent with an active role of the SR proteins in transcription. Because of their similar profiles on DNA, we wondered whether SRSF1 and SRSF2 depend on one another for association with genomic DNA. Antibodies against endogenous SR proteins generated ChIP-seq profiles identical to those with the anti-HA antibody (Figures S1E and S1F), which permitted analysis of one SR protein after the other is depleted. We found that depleting either SR protein greatly diminished the association of the other SR protein with genomic DNA (Figure 1E). Using the normalized data, the SR ChIP-seq signals at the TSS are much reduced in the absence of the other SR protein (Figures S1G and S1H). These observations explain many similar functional requirements later observed for both SR proteins in vivo (see below).

### SR Proteins Are Involved in the Regulation of Transcription Pause Release

Association of SR proteins with both promoter DNA and RNA near the TSS suggests a role in promoter-proximal events that involve nascent RNA. We pursued this idea by monitoring the impact of SR proteins on RNAP II occupancy and nascent RNA production. We performed RNAP II ChIP-seq and global nuclear run-on coupled with deep sequencing (GRO-seq) (Core et al., 2008) and found that depletion of either SR protein induced the accumulation of RNAP II and nascent RNA at the TSS, as illustrated on several genes (Figures 2A, 2B, S2A), which is also evident from meta-analyses of the genome-wide data (Figures 2C and S2B). These data suggest that SR proteins facilitate the release of RNAP II paused near gene promoters, a critical regulatory step recently shown to require key transcription factors and regulators (Byun et al., 2012; Rahl et al., 2010; Sawarkar et al., 2012).

To quantify induced transcription pausing in response to SR protein depletion, we calculated the "traveling ratio" (TR) on individual genes, which is defined by the average RNAP II density near the promoter (−30 to +300 nt from TSS) divided by that in the gene body (+300 nt to the end of gene). Depletion of either SRSF1 or SRSF2 caused RNAP II accumulation at promoters relative to gene bodies (increased TR) based on both RNAP II ChIP-seq and GRO-seq signals, the latter of which reflects transcriptionally engaged RNAP II (Figures 2D and S2C). The TR changes induced by SR protein depletion are highly statistically significant and reminiscent of the effect of blocking P-TEFb with a small-molecule inhibitor or inhibiting c-Myc (Rahl et al., 2010). To rule out the possibility that depletion of any RNA-binding protein capable of interacting with nascent RNA may
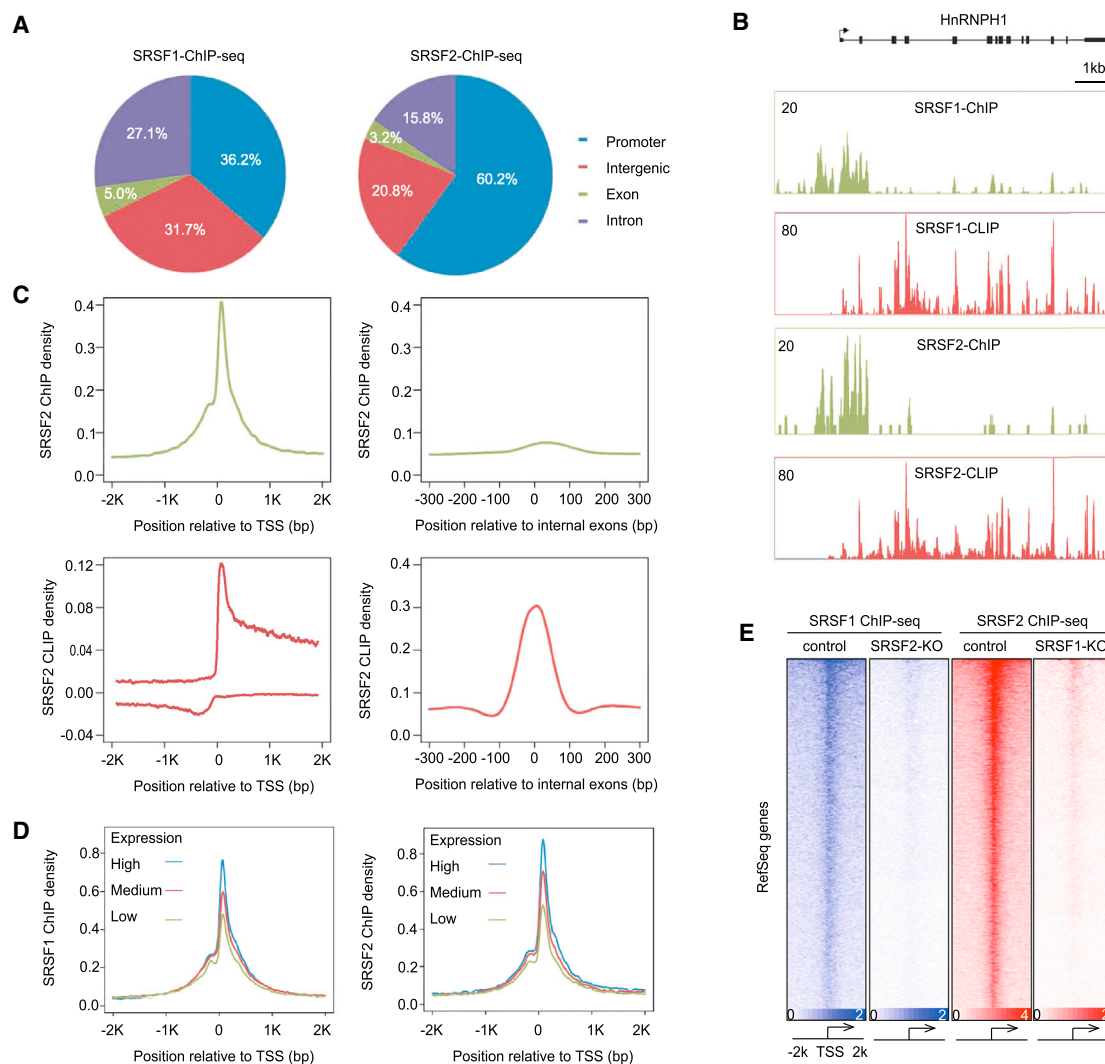
**Figure 1. SR Proteins SRSF1 and SRSF2 Interact with DNA on Gene Promoters and RNA on Exonic Regions**

(A) Genomic distribution of SR protein ChIP tags (SRSF1 total tags = 10,529,663; SRSF2 total tags = 5,199,318), showing that SRSF1 and SRSF2 have similar binding patterns with a significant fraction mapped to gene promoters in each case.

(B) SR protein ChIP-seq and CLIP-seq signals on the representative *hnRNPH1* gene. y axis indicates normalized tags per million, with the floor set to 0. The SR CLIP-seq data sets (SRSF1 total tags = 3,694,535; SRSF2 total tags = 4,874,935) on the same MEFs are from the published work (Pandit et al., 2013).

(C) Metagene analysis of SRSF2 ChIP-seq (green) and CLIP-seq (red) data at the TSS (based on 23,158 annotated TSS), compared to SRSF2 signals on internal exons (based on 149,352 annotated mouse exons). y axis indicates tags per million per gene.

(D) Correlation between SR ChIP-seq signals at the TSS and gene expression analyzed by using all genes with unique and nonoverlapping TSSs. Genes were divided into three groups based on RNA-seq: high (n = 2,829), medium (n = 2,829), and low (n = 2,828). p value is $< 2.2 \times 10^{-16}$ on all pairwise comparisons according to two-tailed Kolmogorov-Smirnov test. y axis indicates tag density per million per gene.

(E) Heatmaps of SR-DNA interactions near the TSS in cells depleted of a different SR protein. Raw tag counts from the same amounts of starting cells were used for comparisons (SRSF1 ChIP-seq tags in WT MEFs = 4,538,963; SRSF1 ChIP-seq tags in SRSF2-depleted MEFs = 551,590; SRSF2 ChIP-seq tags in WT MEFs = 9,489,245; SRSF1 ChIP-seq tags in SRSF2-depleted MEFs = 551,933).

See also Figure S1.

produce a similar effect, we examined hnRNP A or B and found no effect on RNAP II TR in response to knockdown of either protein (Figure 2E).

We next assessed the relationship between TR changes induced by SR protein depletion and corresponding changes in gene expression. Based on RNAP II ChIP density or GRO-seq signals, we divided genes into three bins based on the magnitude of TR changes in response to SRSF2 depletion and found that the largest increases in TR are correlated with reduced gene expression measured by RNA-seq (Figure 2F). Similar results were also obtained for SRSF1 (Figure S2D). These data suggest a strong influence (direct or indirect) of SR proteins on gene transcription in addition to their traditional functions in RNA processing.
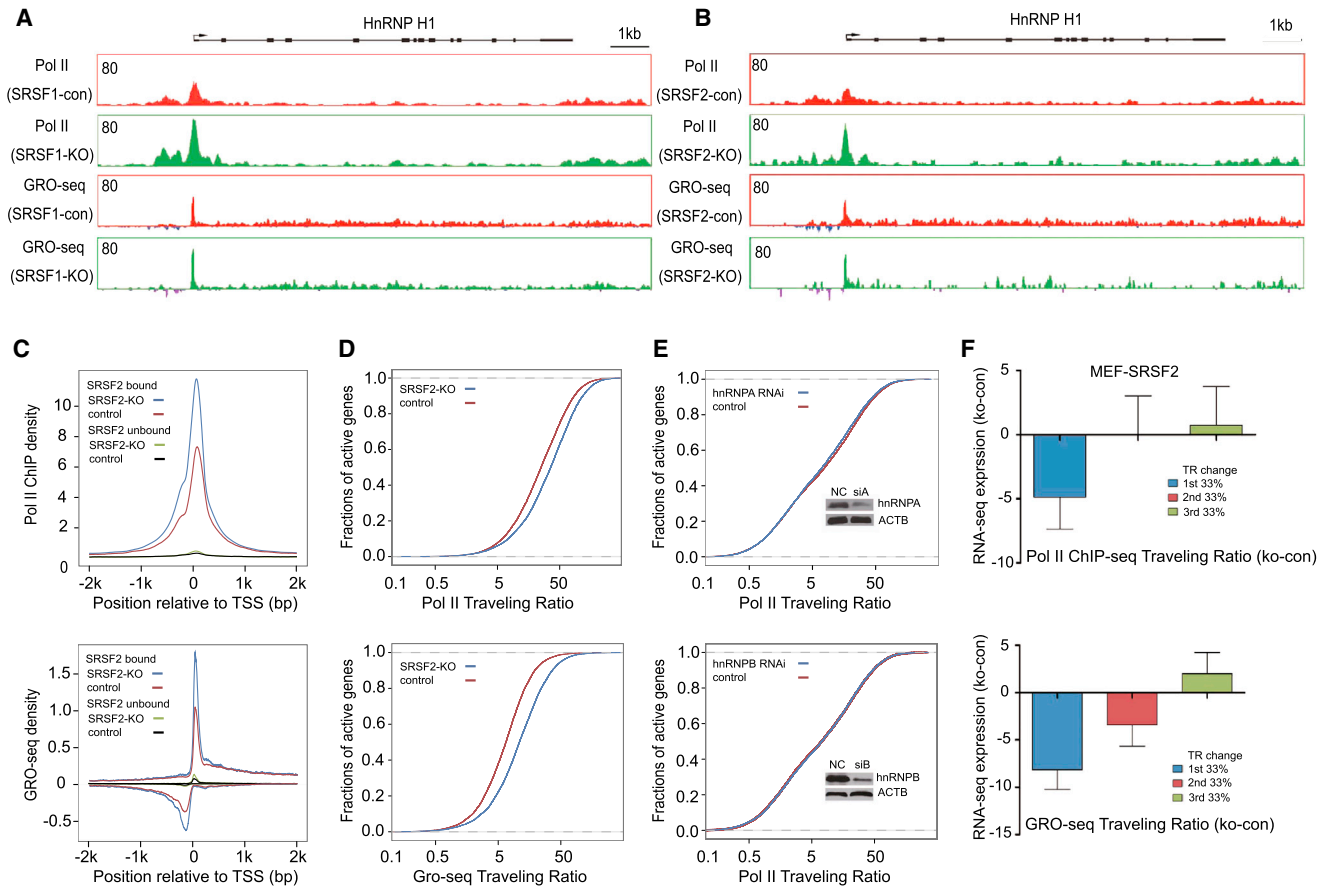
**Figure 2. SR Proteins Are Required for RNAP II Pause Release**

(A and B) UCSC genome browser views of RNAP II ChIP-seq (detected by N20) and GRO-seq signals on the representative *hnRNPH1* gene before and after Dox-induced depletion of SRSF1 (A) or SRSF2 (B) in MEFs. y axis indicates normalized tags per million, with the floor set to 0.

(C) Metagene analysis of RNAP II ChIP-seq (top) or GRO-seq (bottom) signals at the TSS (n = 23,037) in response to SRSF2 depletion. SRSF2-bound and unbound genes were separately compared. The differences are significant (p < 2.2 × 10$^{-16}$) based on two-tailed KS test. y axis indicates normalized tags per million per TSS.

(D) Shift of traveling ratio (TR) based on RNAP II ChIP-seq (top) or GRO-seq (bottom) data sets of active genes in response to SRSF2 depletion (n = 5,703, p < 2.2 × 10$^{-16}$) according to two-tailed KS test in both cases.

(E) TR differences based on RNAP II ChIP-seq signals in MEFs depleted of hnRNP A (top) or hnRNP B (bottom). The knockdown effects were verified by western blotting (insets).

(F) TR shifts based on RNAP II ChIP-seq (top) or GRO-seq (bottom) correlated with induced gene expression in response to SRSF2 depletion. Averaged changes in gene expression (FDR < 0.05) detected by RNA-seq were plotted against three groups of genes evenly divided according to their TR differences from large to small.

See also Figure S2.

## The SR-Promoter Interactions Depend on RNA, but Not Ongoing Transcription

Early studies indicated that SR proteins interact with the RNAP II complex (Misteli and Spector, 1999) but in an RNA-dependent manner (Sapra et al., 2009). To determine whether ongoing transcription is required for such interactions, we performed ChIP analysis on several gene promoters in response to α-amanitin treatment, which largely abolished ongoing transcription based on RT-qPCR analysis of nascent RNA but only modestly reduced RNAP II occupancy near the TSS (Figure 3A). Unexpectedly, we found little or no effect of α-amanitin treatment on SR ChIP signals on the *hnRNPH1* and *TMSB4X* promoters, indicating that

ongoing transcription may not be a prerequisite for SR proteins to associate with genomic DNA (Figure 3A).

To determine whether the association of SR proteins with DNA is dependent on RNA, we performed ChIP-qPCR analyses on multiple SR protein target genes in the presence of RNase T1 or RNase A and found that SR ChIP signals on both promoter and gene body are sensitive to the RNase treatment (Figures 3B and S3A). Some remaining ChIP signals might result from formaldehyde-mediated crosslinking of the SR proteins to other DNA-binding proteins. These results suggest that some RNA of unknown identity may be responsible for linking SR proteins to genomic DNA near the TSS.
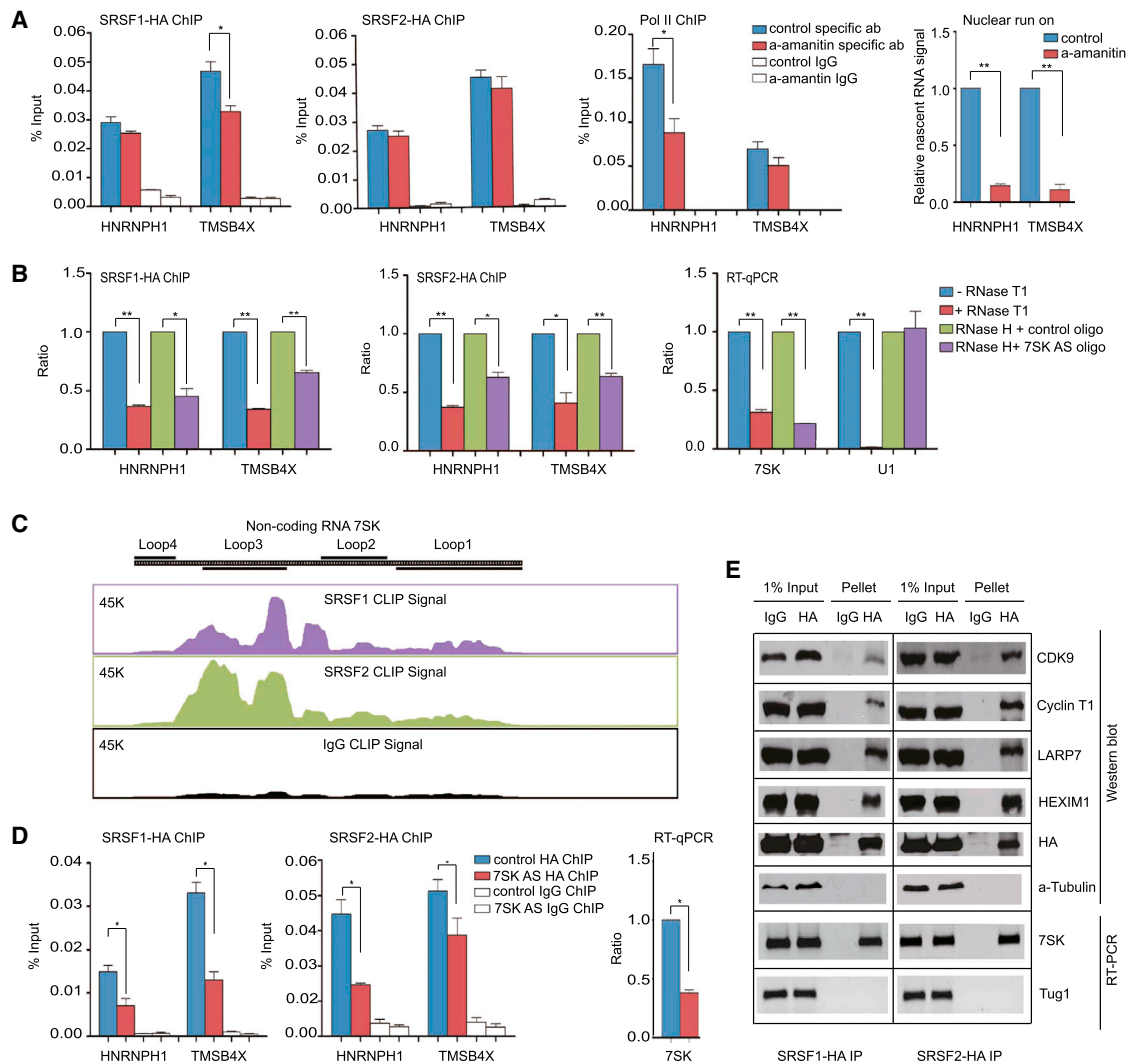
**Figure 3. Noncoding 7SK RNA Mediates SR Protein Binding to Gene Promoters**

(A) ChIP-qPCR analysis of SR protein interaction with two gene promoters (*HNRNPH1* and *TMSB4X*) in MEFs mock treated with DMSO or treated with α-amanitin (left two panels). The effects of α-amanitin on RNAP II occupancy and nascent RNA (produced during nuclear run-on) at the TSS regions of the two genes were determined by ChIP or RT-qPCR (right two panels).

(B) ChIP-qPCR analysis of SR protein interaction with gene promoters using cell lysate treated with RNase T1 or RNase H plus anti-7SK oligo (7SK AS) (left two panels). 7SK level was measured by RT-qPCR; U1 snRNA served as a negative control (right).

(C) SR protein CLIP-seq signals on the 7SK RNA. IgG CLIP served as a negative control. y axis indicates normalized tags per million, with the floor set to 0.

(D) ChIP-qPCR analysis of SR protein interaction with gene promoters in response to degradation of the 7SK RNA by an anti-7SK oligo in MEFs. A scrambled oligo served as a negative control. The far-right panel shows the level of the 7SK RNA measured by RT-qPCR under each treatment condition.

(E) Co-IP/western blotting analysis, showing SR proteins as part of the 7SK complex.

Data are shown in (A), (B), and (D) as mean ± SD. *p < 0.05 and **p < 0.005 based on Student's t test. See also Figure S3.

## The 7SK Noncoding RNA Mediates SR-Promoter Interactions

An increasing number of noncoding RNAs have been shown to mediate protein interactions with genomic DNA (Rinn and Chang, 2012). We searched for abundant cellular RNAs bound by SR proteins in our CLIP-seq data sets, paying particular attention to those previously implicated in the regulation of transcription. Interestingly, we noted that both SRSF1 and SRSF2 crosslink extensively to the 7SK noncoding RNA (note the scale of the y axis in Figure 3C), known to regulate transcription elongation by control-

ling the elongation factor P-TEFb (Ott et al., 2011; Peterlin and Price, 2006). Both SR proteins specifically bound the third stem loop, adjacent to the regions bound by other relatively stable components of the 7SK complex (Krueger et al., 2010).

To test whether the 7SK RNA mediates the association of SR proteins with genomic DNA, we used RNase H and a 7SK-antisense oligonucleotide to specifically degrade the RNA (Figure 3B). Similar to RNase T1 treatment, specific degradation of the 7SK RNA greatly reduced SR protein ChIP signals on gene promoters, whereas the scrambled control oligonucleotide had

no effect (Figure 3B). We also transfected a DNase-resistant 2′-O-methyl anti-7SK oligonucleotide into MEFs to degrade the 7SK RNA in vivo and again observed significant reduction of SR protein ChIP signals on gene promoters (Figure 3D). These data strongly suggest that the 7SK RNA plays a major role in mediating the SR protein-promoter DNA interactions.

The evidence for SR proteins crosslinking to the 7SK RNA, in conjunction with their colocalization in the cell (Prasanth et al., 2010), suggests that both SR proteins may be part of the 7SK complex. To test this hypothesis, we immunoprecipitated HA-tagged SRSF1 or SRSF2 followed by RT-PCR analysis of the 7SK RNA and western analysis of previously characterized components of the 7SK complex, including P-TEFb (CDK9/CyclinT1), LARP7, and HEXIM1. We found that the anti-HA immunoprecipitate contained all established components of the 7SK complex, but not the abundant polycomb-body-associated Tug1 noncoding RNA (Yang et al., 2011), which served as a negative control (Figure 3E). The associations were sensitive to RNase T1 treatment in vitro (Figure S3B) or degradation of the 7SK RNA in vivo (Figure S3C). These results demonstrate that both SR proteins are part of the 7SK complex under physiological conditions despite the fact that these SR proteins were not detected in highly purified 7SK complex in previous proteomics studies (Yang et al., 2005; Yik et al., 2003).

## The 7SK Complex Is Intimately Associated with Active Gene Promoters

The observations that SRSF1 and SRSF2 are part of the 7SK complex and that the 7SK RNA is critical for their association with gene promoters raise an intriguing possibility that these factors may represent a previously unknown molecular assembly near the TSS. By cellular fractionation, we found that ~50% of the 7SK complex could be readily extracted under mild detergent conditions, likely representing the soluble pool of the 7SK complex in the nucleoplasm, and the remaining half of the 7SK complex was associated with the chromatin fraction and releasable with DNase I treatment (Figure 4A).

The ability to detect a significant amount of the 7SK complex on chromatin is reminiscent of a recent observation that both CDK9 and HEXMI1 appear to interact with the HIV-1 promoter even before Tat induction, indicating that the 7SK complex may be more closely associated with genomic DNA than previously thought (D'Orso and Frankel, 2010). To extend this observation, we conducted ChIP-qPCR analysis and detected both CDK9 and HEXIM1 on multiple endogenous gene promoters (Figures 4B and S4A). In these experiments, we note that the standard formaldehyde-based crosslinking protocol is robust for ChIP-qPCR, but not for genome-wide analyses by ChIP-seq. Reasoning that this might reflect multiple protein-mediated associations between the 7SK complex and genomic DNA, which may not be efficiently preserved by formaldehyde, we employed a recently described glutaraldehyde-based crosslinking strategy, which appears to be more effective in mapping noncoding RNA-containing complexes to mammalian genomes, even though this method is anticipated to cause higher background due to extensive crosslinking induced by glutaraldehyde (Chu et al., 2011). Under these conditions, ChIP-seq with specific antibodies revealed the interactions of both CDK9 and

HEXIM1 with genomic DNA at the promoter-proximal regions, and the IgG control showed a modest degree of enrichment as predicted by the relatively high background of the glutaraldehyde-based method (Figure 4C).

We note that the ChIP-seq signals of the 7SK complex components are broadly associated with gene promoters (the peaks occupy ~1 kb on both sides of gene promoters), similar to the binding patterns seen on the HIV-1 gene (D'Orso and Frankel, 2010). These data suggest that the 7SK complex functions at endogenous gene promoters as well as at the HIV promoter. Consistently, we found that the association of the 7SK complex with a given endogenous gene promoter is positively correlated with the degree to which RNAP II pauses at the TSS of that promoter, a relationship that is also true for SRSF1 and SRSF2 (Figure S4B). Furthermore, P-TEFb (CDK9) and the SR proteins co-occupy a large set of gene promoters (Figure 4D, left for SRSF2; data not shown for SRSF1). This relationship likely reflects function because of extensive overlap between the set of genes that responded to DRB inhibition of P-TEFb with increased RNAP II pausing at their TSSs and the set that responded similarly to depletion of either SR protein (Figure 4D, right for SRSF2; data not shown for SRSF1). Together, these results strongly implicate functional cooperation between SR proteins and the 7SK complex at endogenous promoters.

## RNA Triggers SR Protein Release along with P-TEFb from the 7SK Complex

Having established the presence of at least two SR proteins in the 7SK complex, we next asked whether such interactions might be perturbed by the presence of RNA with high-affinity binding sites for specific SR proteins. We modified a P-TEFb release assay described previously (Krueger et al., 2010) by incubating the 7SK complex brought down with anti-HEXIM1 antibody with increasing amounts of RNA-containing SR protein-binding sites (schematic in Figure 4E). We selected the sequence GAAGGA, a high-affinity binding site for multiple SR proteins that has been characterized as an exonic-splicing enhancer (ESE) (Cavaloc et al., 1999). A pyrimidine-rich sequence (UUCUCU) incapable of interacting with SR proteins was tested as a negative control. We found that the added ESE RNA released both SRSF1 and SRSF2 from the 7SK complex and, strikingly, that SR protein release was accompanied by progressive release of CDK9, whereas HEXIM1 remained associated with beads (Figure 4E, blue boxed lanes 6–8).

We also tested a 2′-O-methyl oligonucleotide antisense to the SR protein-binding site in the 7SK RNA for its ability to bump off SR proteins and found that the specific antisense oligonucleotide, but not a nonspecific control, could compete off both SR proteins as well as CDK9 from the immunopurified 7SK complex (Figure 4E, red boxed lanes 12–14). The 7SK RNA remained intact under these conditions (data not shown). Collectively, these data demonstrate RNA-induced coordinated release of SR proteins and P-TEFb from the 7SK complex. We envision that nascent RNA might trigger this process during transcription activation.

## SRSF1 and SRSF2 Connect P-TEFb to RNAP II

Previous work showed that SR proteins are associated with RNAP II, which we further confirmed by reciprocal IP using
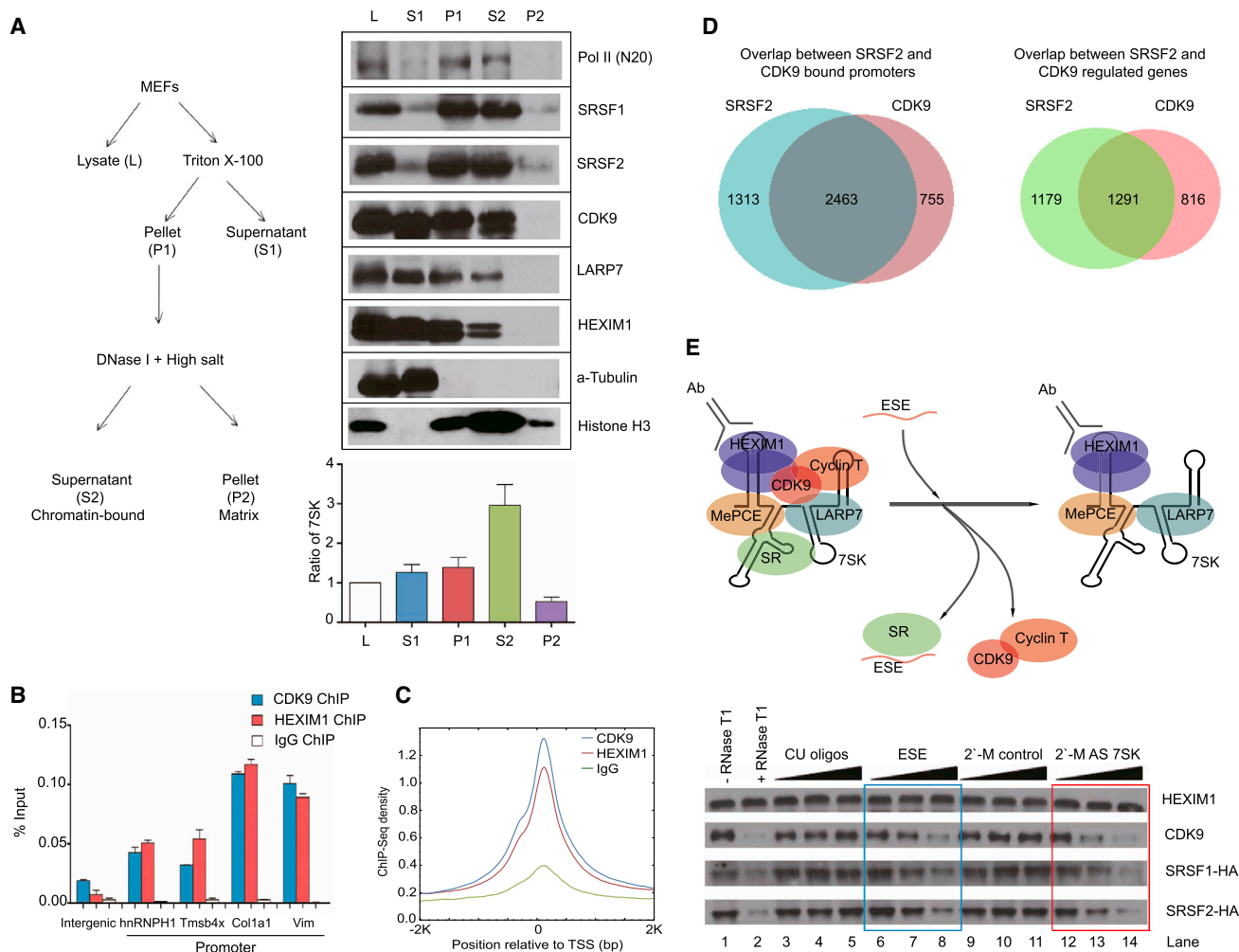
**Figure 4. SR Proteins Mediate P-TEFb Release from the 7SK Complex in an RNA-Dependent Manner**

(A) Experimental strategy used to fractionate MEFs. Both active and inhibitory components of the 7SK complex are equally distributed between the soluble (S1) and chromatin-bound fraction (P1 or S2). Histone H3 and α-tubulin served as chromatin-bound and unbound markers.

(B) ChIP-qPCR analysis of CKD9 and HEXIM1 interactions on four gene promoters in glutaraldehyde-crosslinked MEFs. "Intergenic" indicates a region ~5 kb upstream the *Vim* gene promoter.

(C) Genome-wide analysis of CDK9 (blue) and HEXIM1 (red) interactions near the TSS (n = 23,037) in glutaraldehyde-crosslinked MEFs. Note some background enrichment with IgG control (green) under this experimental condition. $p < 2.2 \times 10^{-16}$ is calculated based on two-tailed KS test. y axis indicates normalized tags per million per gene.

(D) Venn diagrams of genomic interactions between CDK9 and SRSF2 detected by ChIP-seq (left) and the induction of RNAP II pausing on P-TEFb-dependent versus SRSF2-dependent genes (right), indicating extensive physical and functional relationships ($p < 2.2 \times 10^{-16}$, hypergometric test) between these two factors.

(E) RNA-dependent release of SR proteins and CDK9 from anti-HEXIM1 IPed 7SK complex. (Top) The strategy for the RNA-mediated P-TEFb release assay. (Bottom) Western blotting analysis of SR proteins and CDK9 released from the 7SK complex with increasing amounts of RNA. Blue and red boxes, respectively, highlight dosage-dependent P-TEFb release induced by the purine-rich ESE and the 2′-O-methyl oligo complementary to the mapped SR-binding site in the 7SK RNA.

Data in (A) and (B) are shown as mean ± SD. See also Figure S4.

anti-HA and anti-RNAP II antibodies (Figures S5A and S5B). Because of similar effects with both SRSF1 and SRSF2 observed thus far, we focused on SRSF2 in the remaining in vivo studies until the experiments designed to define the direct role of specific SR proteins in transcription activation. RNase treatment or degradation of the 7SK RNA greatly reduced the association of RNAP II with SRSF2 (Figures 5A and 5B), whereas in vivo depletion of SRSF2 modestly increased the association of

P-TEFb subunit CDK9 with the 7SK RNA, likely due to the replacement of SRSF2 by other SR proteins, which might slightly reduce the dynamics of the 7SK complex in the cell (Figure 5C). Importantly, these data show that the 7SK complex connects SRSF2 to RNAP II.

Knockdown of SRSF2 caused a dramatic and selective reduction of RNAP II phosphorylation at the P-TEFb target site Ser2, but not at Ser5 positions (Figure 5D). Interestingly, degrading
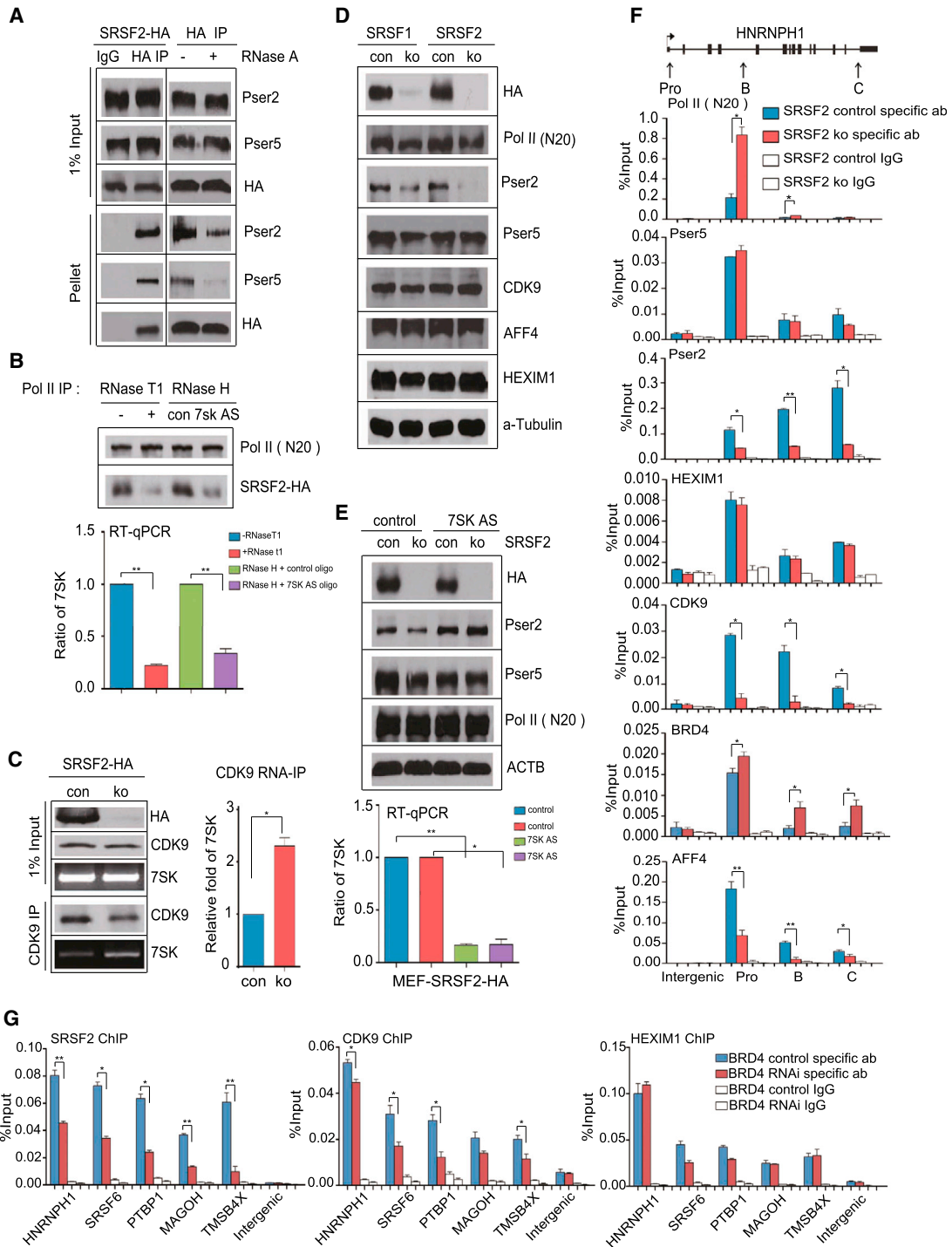
**Figure 5. 7SK RNA Connects SR Proteins to RNAP II, and SR Proteins Are Required for SEC Recruitment to Gene Promoters**

(A) Co-IP/western blotting analysis, showing RNA-dependent association of RNAP II with HA-tagged SRSF2.

(B) Reciprocal co-IP/western blotting analysis, demonstrating RNA-dependent association of HA-tagged SRSF2 with IPed RNAP II. Levels of the 7SK RNA were determined by RT-qPCR under different treatment conditions (bottom).

(C) Ribo-IP analysis, showing slightly increased association of CDK9 with 7SK in SRSF2-depleted MEFs. (Right) Levels of CDK9-associated 7SK quantified by RT-qPCR.

(D) Western blotting analysis, showing diminished Ser2-phosphorylated RNAP II (Pser2) in SR protein-depleted MEFs. Specific antibodies were used to detect total RNAP II (N20), Ser2- and Ser5-phosphorylated RNAP II (Pser2 and Pser5), CDK9, AFF4, and HEXIM1 before and after Dox-induced depletion of SRSF1 or SRSF2. α-tubulin served as loading control.

*(legend continued on next page)*

the 7SK RNA reversed the reduction of RNAP II Ser2 phosphorylation in MEFs (Figure 5E). This result could be explained if the increased soluble pool of released P-TEFb in 7SK RNA knockdown cells compensated for the defects in SRSF2-dependent delivery of P-TEFb to RNAP II at gene promoters, as suggested earlier (Yik et al., 2003; Young et al., 2007). As a result of this compensation, degradation of the 7SK RNA reduced the association between SRSF2 and P-TEFb but caused little change in RNAP II traveling ratio (Figure S5C).

## SEC Is Recruited to Gene Promoters in an SRSF2-Dependent Manner

To further explore the mechanism for SRSF2-dependent RNAP II pause release, we performed ChIP analyses of multiple key transcription factors implicated in transcription elongation before and after depletion of the SR protein. On the *hnRNPH1* gene, SRSF2 depletion caused the accumulation of total RNAP II (detected with N20) at the gene promoter but had little effect on levels of Ser5-phosphorylated RNAP II or HEXIM1 (Figure 5F). In contrast, the ChIP signals of Ser2-phosphorylated RNAP II and CDK9 were greatly diminished (Figure 5F). These results were also evident on multiple other genes that we examined (Figure S5D). We next asked whether the impairment might result from reduced recruitment of Brd4, a chromatin-associated factor known for its critical role in P-TEFb recruitment to RNAP II at gene promoters. We found that the Brd4 ChIP signals changed little or even increased to some extent in SRSF2-depleted cells (Figures 5F and S5D). Conversely, Brd4 RNAi significantly reduced the ChIP signals of both SRSF2 and CDK9, but not HEXIM1, on the majority of their target gene promoters that we examined (Figure 5G). These data are fully consistent with the proposed role of Brd4 as a key chromatin "receptor" for P-TEFb.

To understand how various defects that we detected in SRSF2-depleted cells might cause RNAP II to pause at gene promoters, we assessed the recruitment of the superelongation complex (SEC) to gene promoters. A previous study showed that various forms of the SEC all contain AFF4, which is essential for transcription elongation (Lin et al., 2010). By ChIP-qPCR, we found that SRSF2 depletion dramatically impaired the recruitment of AFF4 to multiple gene promoters with the exception of SRSF6 (Figures 5F and S5D), consistent with the selective effects of SEC on different genes (Luo et al., 2012). Collectively, the data suggest a chain of events that lead to SEC recruitment, and defects in this chain cause RNAP II pausing in the promoter-proximal regions.

## SRSF2 Mediates ESE-Dependent Transcriptional Activation

As a family, SR proteins have the capacity to bind diverse RNA sequences. In particular, SRSF2 seems to recognize highly degenerate ESEs (Daubner et al., 2012), suggesting its ability to act on diverse sequences in transcribed RNA in mammalian transcriptome. Because a high-affinity SR protein-binding sequence is able to release P-TEFb from the 7SK complex, we hypothesized that RNA elements in nascent promoter-associated transcripts may provide critical signals to induce a chain of events similar to that triggered by Tat through binding to TAR on the HIV-1 promoter. Using luciferase reporters driven by several commonly used promoters, each of which carries a different 5′ UTR, we found that SRSF2 actively promoted gene expression from the HIV-1 promoter in transfected HEK293T cells (Figure 6A), and this effect was independent of but synergistic with the activity of Tat on TAR (Figure S6A). Importantly, we found that SRSF2 could be coimmunoprecipitated with Brd4, and Brd4 RNAi abolished SRSF2-dependent transcriptional response on the HIV-1 promoter (Figures S6B and S6C). These data suggest that SRSF2 may act like HIV Tat but via different *cis*-acting element(s) in transcription activation.

In these functional analyses, we returned to compare between SRSF2 and SRSF1 to determine the specificity and mechanism in transcriptional activation by different SR proteins. Interestingly, we observed that SRSF1 activated reporters driven by both the HIV-1 and HSV promoters, but not by the CMV promoter, indicating a degree of specificity in these reporter-based assays (Figure 6A). In light of the previous finding that the shuttling SRSF1, but not nonshuttling SRSF2, was able to activate translation of ESE-containing luciferase reporters in the cytoplasm (Sanford et al., 2004), we asked whether each of the activation events might result from enhanced transcription (increase in mRNA) or elevated translation (assessed by increase in luciferase activity without corresponding increase in mRNA). The results suggested that the effect of SRSF1 on HIV-1 was transcriptional, whereas its impact on HSV was primarily translational. In contrast, the effect of SRSF2 on HIV-1 was largely transcriptional (Figure 6A).

We next chose an endogenous gene (*hnRNPH1*), which depended on both SRSF1 and SRSF2 for efficient expression in MEFs to link its 3 kb core promoter and 5′ UTR to a luciferase reporter. We found that the first 100 nt sequences in its 5′ UTR were necessary and sufficient for transcriptional activation (Figure 6B). Interestingly, this transcription unit responded positively to SRSF2 overexpression and negatively to SRSF2 downregulation by RNAi, but not to SRSF1 overexpression, although we detected some effect with SRSF1 RNAi. We further tested the response of the HSV-based reporter to both of the SR proteins using a tethering approach by engineering a tandem repeat of RNA elements that can be recognized by a specific Pumilio 1 (PUF) RNA-binding motif (RRM) and by scoring the reporter response to individual SR proteins fused to the PUF RRM (Wang et al., 2009). We found that tethered

(E) Restoration of RNAP II Ser2 phosphorylation in MEFs depleted of both SRSF2 and 7SK RNA. (Bottom) Levels of 7SK determined by RT-qPCR.

(F) Requirement of SRSF2 for efficient recruitment of the superelongation complex (SEC) to gene promoters. In response to SRSF2 depletion, Ser2-phosphorylated RNAP II and two common components (CDK9 and AFF4) of SEC were dramatically reduced on *HnRNPH1* gene, but the levels of HEXIM1 and chromatin-bound Brd4 were unaffected.

(G) Brd4 requirement for the recruitment of CDK9 and SRSF2 to gene promoters.

Data in (B), (C), (E), (F), and (G) are shown as mean ± SD. *p < 0.05 and **p < 0.005 based on Student's t test. See also Figure S5.
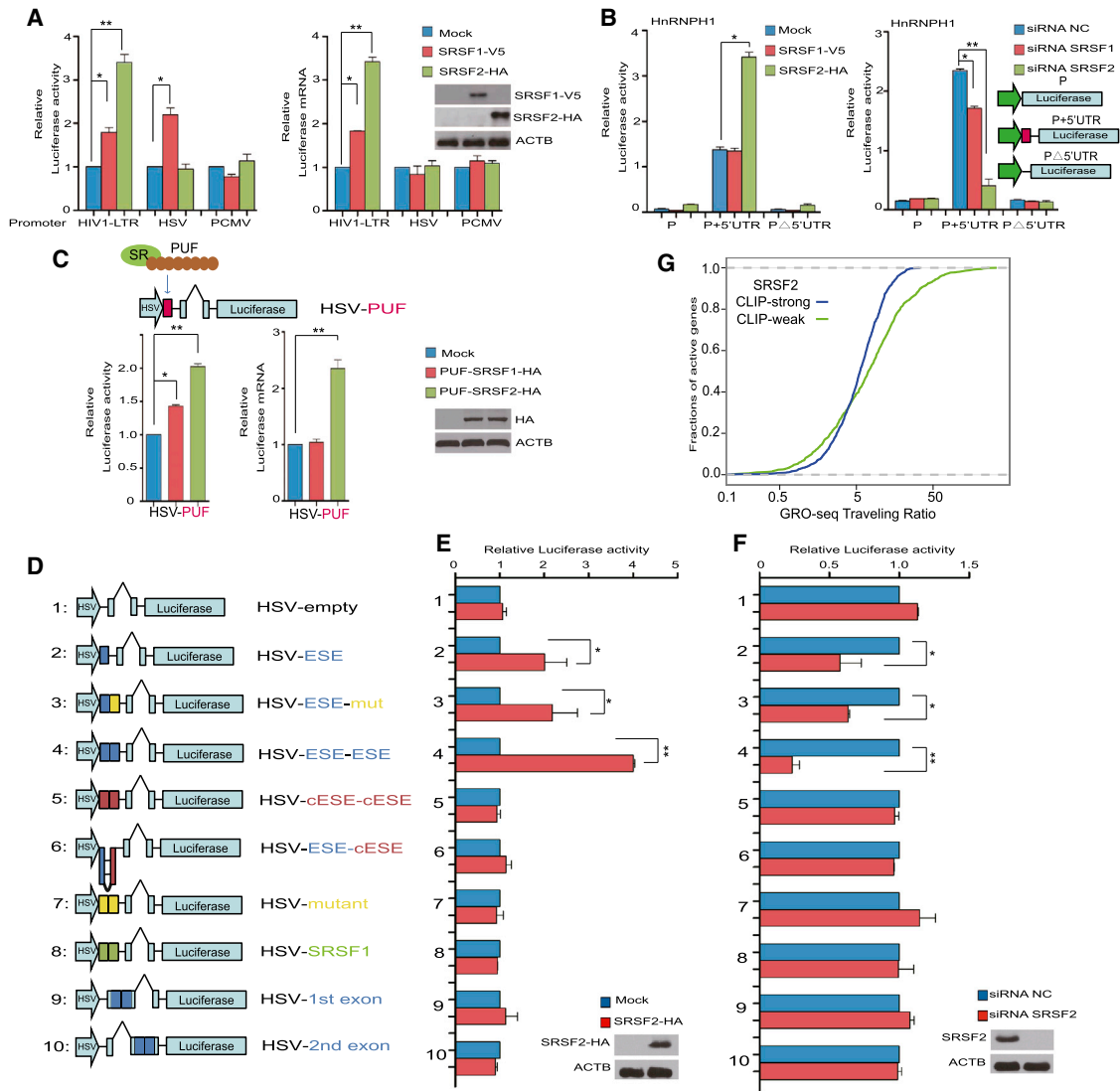
**Figure 6. High-Affinity RNA Elements Mediate Transcriptional Activation by SRSF2**

(A) Luciferase assay in transfected HEK293T cells, showing activation of the HIV-1 promoter by overexpressed V5-tagged SRSF1 or HA-tagged SRSF2 (inset). SRSF1 activated the HSV-driven reporter at the translational level (no induction of mRNA), but SRSF2 had no effect on HSV. None of the SR proteins activated the CMV promoter.

(B) Luciferase assay of reporters constructed from the *hnRNPH1* gene in response to SR protein overexpression (left) or RNAi (right). Specific constructs are illustrated on the right.

(C) Tethered SRSF2 activated transcription from an HSV-based reporter containing a specific PUF binding motif (red box). The SRSF2-PUF fusion protein activated transcription, whereas SRSF1-PUF fusion protein stimulated translation of the reporter. A plasmid not carrying any SR-coding sequences served as a negative control.

(D–F) Schematic presentation of HSV-based reporters (D). Dual luciferase assays based on PCMV (internal control) and HSV-ESE reporters in response to SRSF2 overexpression (E) or RNAi (F). (Insets) Protein levels monitored by western blotting.

(G) Comparison of traveling ratio (TR) of transcriptionally engaged RNAP II (based on GRO-seq signals) on two groups of genes with high (blue) or low/no (green) SRSF2 CLIP-seq signals on their 300 nt TSS-associated RNA. Genes with lower TR are more linked than those with higher TR to SRSF2 binding on RNA near the TSS. The differences are highly significant ($p < 2.2 \times 10^{-16}$) based on the two-tailed KS test.

Data in (A), (B), (C), (E), and (F) are shown as mean ± SD. *$p < 0.05$ and **$p < 0.005$ based on Student's t test. See also Figure S6.

SRSF2 activated the reporter at the transcriptional level, whereas tethered SRSF1 enhanced the translation of the reporter (Figure 6C). Using this approach, we performed preliminary survey for multiple other SR family members. The data (not shown) suggest that SRSF2 may be the only SR protein capable of functioning as a general transcription activator, whereas other SR proteins may largely affect transcription via indirect mechanisms (see Discussion).

## Single-Stranded ESE Near TSS Is Required for Transcriptional Activation

The ESEs cloned after the TSS may simply function as downstream enhancers to activate transcription, and increased reporter expression may also indirectly result from enhanced RNA stability or transport. To address these possibilities, we engineered the HSV-based reporter, which does not seem to contain any SRSF2-responsive element (see Figure 6A). This reporter also carries a splicing unit in front of the luciferase-coding sequences (construct 1, Figure 6D), which permitted us to determine positional requirements for engineered ESEs to activate transcription. We selected a well-characterized SRSF2-responsive ESE from the second exon of the β-globin gene (Schaal and Maniatis, 1999) and prepared a series of constructs to test the response to SRSF2 overexpression (Figure 6E, insert) or knockdown (Figure 6F, insert).

We first examined the reporters containing one or two copies of SRSF2 ESE inserted in the 5′ UTR (constructs 2 to 4), finding that SRSF2 overexpression activated (whereas SRSF2 RNAi diminished) transcription in an ESE-dependent manner and that the reporter containing two ESEs showed stronger responses to SRSF2 overexpression or knockdown than did the reporter containing a single ESE. In contrast, reporters containing two copies of a control antisense ESE (cESE) (construct 5), mutant ESE (construct 7), or SRSF1 ESE (construct 8) all failed to respond to SRSF2 overexpression or knockdown. Importantly, when one ESE and one cESE were cloned adjacent with one another in the reporter (which has a potential to form a hairpin, construct 6), we detected no response, suggesting that the ESE must be exposed as single-stranded RNA to function in SRSF2-mediated transcription activation. This experimental strategy, which has been used to demonstrate Tat binding to TAR in nascent viral transcripts to activate transcription (Berkhout et al., 1990), suggests that SRSF2 activates transcription via the promoter-proximal nascent RNA.

We next moved the SRSF2 ESE from the promoter-proximal region to 60 nt downstream from the TSS or to the second exon (constructs 9 and 10) in the reporter. We found that the ESE no longer mediated the response to SRSF2 overexpression or knockdown (Figures 6E and 6F). This position-sensitive effect argues against RNA stability, transport, or splicing-related mechanisms because the ESE in either exon 1 or 2 is expected to have similar effects on those pathways. Together, these data suggest that SRSF2-binding motifs near the 5′ end of nascent RNA may act as critical signals for transcription activation, a mechanism that is highly reminiscent of the Tat-TAR interaction during the activation of the HIV-1 promoter.

Finally, to relate the ESE effect from reporter-based studies to global activation of gene expression, we determined how SRSF2-RNA interactions near the TSS might be correlated to the traveling ratio of transcriptionally engaged RNAP II measured by nascent RNA production. Because CLIP-seq signals are generally lower near the TSS than on internal exons among ~8,700 genes that showed sufficient GRO-seq signals on both TSS and gene body, we selected the top 1,000 genes with significant CLIP-seq signals in the first 300 nt and compared them with the bottom 1,000 genes that had little or no detectable CLIP-seq signals in the same region. We found that transcriptionally engaged RNAP II (measured by associated GRO-seq signals) on genes with little SRSF2 CLIP-seq signals near their TSS tend to be much more paused than those with high SRSF2 CLIP-seq signals, especially among 60% of genes in both groups with high (>5) traveling ratios (Figure 6G). These findings suggest that transcriptionally engaged RNAP II near the TSS of genes with higher SRSF2 CLIP-seq signals is more efficient in entering the gene body, therefore providing global evidence that SRSF2 interacts with promoter-associated RNAs to enhance transcription elongation.

## DISCUSSION

### An Unexpected Role of SRSF2 in the Regulation of Transcription Pause Release

Transcriptional pause release has been increasingly realized as a major step in regulated gene expression (Adelman and Lis, 2012). As depicted in Figure 7 (upper-right), HIV Tat has the ability to extract P-TEFb from the 7SK complex (Krueger et al., 2010). This process is likely mediated by the protein-protein interaction between Tat and the cyclin T subunit of P-TEFb. The released P-TEFb is recruited to paused RNAP II via Tat binding to TAR on nascent RNA, which is assisted by chromatin-bound Brd4 (Peterlin and Price, 2006). These events eventually trigger the transition of transcriptionally engaged RNAP II on the HIV-1 promoter from the pausing to the elongating state. At this point, the field is still debating whether the released P-TEFb first joins its nucleoplasmic pool before being recruited to paused RNAP II or whether these two steps might be more locally coupled at gene promoters (indicated by the two dashed arrows in the model).

We see many parallels between the actions of Tat and SRSF2 in the regulation of transcription pause release (Figure 7, lower-right). We provide evidence that the SR protein is part of the 7SK complex (Figure 3E) and that high-affinity RNA for the SR protein can induce P-TEFb release from the complex (Figure 4E). SRSF2 can also be coimmunoprecipitated with both P-TEFb and Brd4 (Figures S3B and S6B). Importantly, like the HIV Tat/TAR system, SRSF2 appears to use chromatin-bound Brd4 to enhance the recruitment of P-TEFb to paused RNAP II (Figures 5F and 5G), and the SRSF2 ESE on nascent promoter-proximal RNA is sufficient for SRSF2-mediated activation of gene expression (Figure 6). However, two critical questions await future studies. One is whether an ESE on nascent promoter-associated RNA is directly responsible for inducing SRSF2 release from the 7SK complex, and the other is how the released P-TEFb might be recruited to paused RNAP II near the TSS (indicated by the two dashed arrows in the bottom panel of the model).

### The 7SK Complex as Part of Megadalton Promoter Complexes

It does not seem to be an efficient mechanism for released P-TEFb to first diffuse around in the nucleoplasm before being attracted to gene promoters for transcription activation. Frankel and colleagues first detected a close association of the 7SK complex with the HIV-1 promoter before Tat induction (D'Orso and Frankel, 2010). We now provide genome-wide evidence for this spatial arrangement by showing preferential association
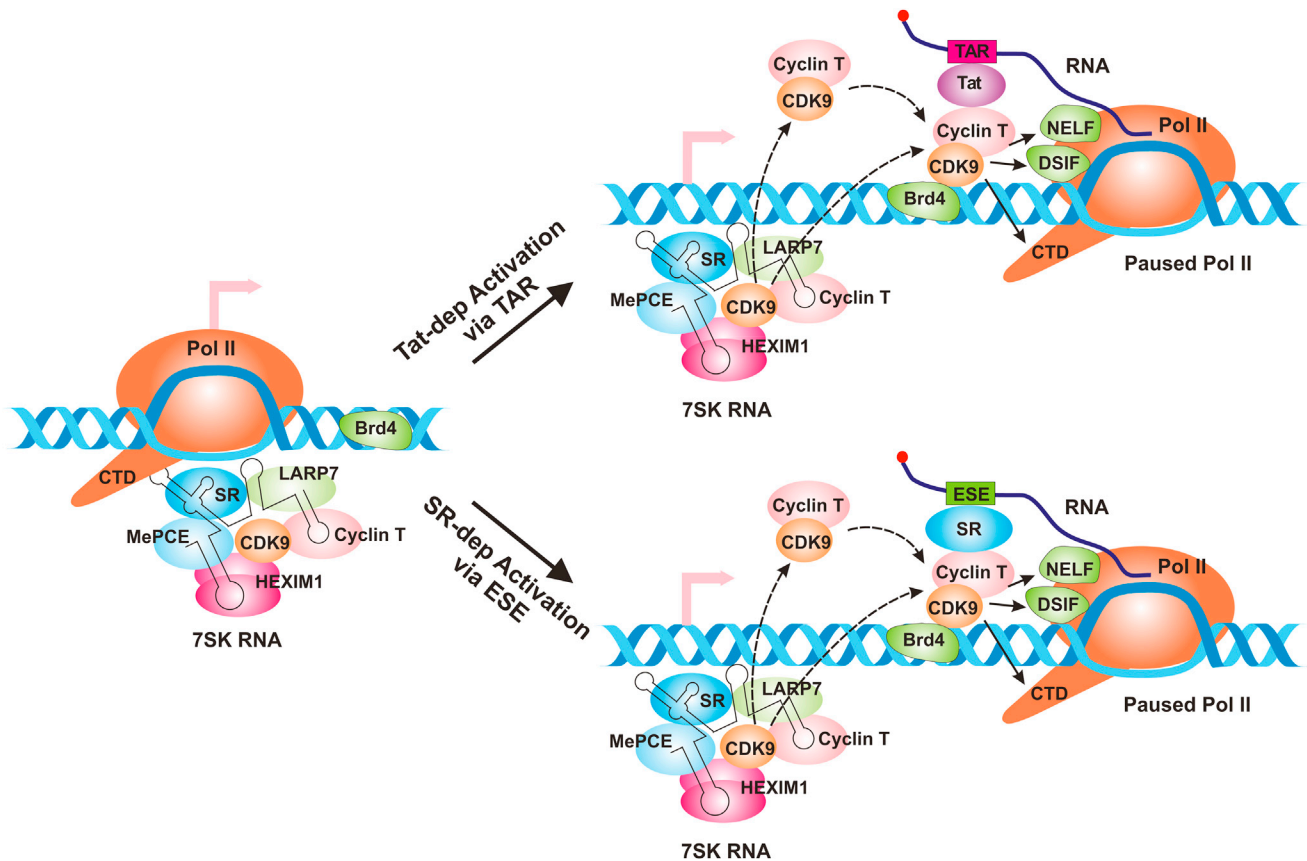
**Figure 7. A Working Model for SR Protein-Dependent Transcriptional Activation**

SR proteins and the 7SK RNA complex are intimately associated with genomic DNA near the promoter-proximal region (left). This proximity may allow more efficient local switches during gene activation. During Tat-dependent activation of the HIV-1 promoter (upper-right), Tat binding to TAR induces relocation of P-TEFb (CDK9:cyclin T) from the 7SK complex to paused RNAP II. This process may be facilitated by direct protein-protein interactions between Tat and cyclin T and between Brd4 and CDK9. It is currently unclear whether released P-TEFb is directly recruited to RNAP II or indirectly via the nucleoplasmic pool, as indicated by the dashed arrows. During transcription pause release on cellular genes (lower-right), SR proteins are also associated with gene promoters as part of the 7SK complex. We speculate that, by taking advantage of local assembly, an SRSF2-binding site (ESE) emerging from RNAP II may induce the SR protein to switch from the 7SK RNA to nascent RNA, thereby triggering the coordinated release of P-TEFb from the 7SK complex. Again, the released P-TEFb may go through two separate routes before being recruited to paused RNAP II at the TSS, as indicated by the dashed arrows. In both the Tat/TAR and SR/ESE systems, chromatin-bound Brd4 may enhance the association of released P-TEFb with RNAP II at the TSS. Recruited P-TEFb will phosphorylate RNAP II and some key factors, such as NELF and DSIF, resulting in transcription pause release.

of both the inhibitory (HEXIM1) and active (CDK9) components of the 7SK complex with active gene promoters. Such intimate association suggests an intriguing possibility that P-TEFb may undergo a local switch from the 7SK complex to gene promoters in both Tat- and SRSF2-dependent gene activation.

Most constitutively active genes may use such a local switch mechanism for basal level transcription. However, some highly induced genes may attract additional P-TEFb from the nucleoplasmic pool because a large amount of extra P-TEFb was clearly recruited to the HIV-1 promoter during Tat-mediated gene induction (D'Orso and Frankel, 2010) and to the Hsp70 gene promoter in response to heat shock (Zobeck et al., 2010). Interestingly, it appears that SR proteins are also recruited to some highly induced genes in both *Drosophila* (Champlin et al., 1991) and mammalian cells (Sapra et al., 2009). These observations suggest that SR proteins may be additionally recruited either independently or together with P-TEFb to some rapidly

induced genes for both transcriptional activation and cotranscriptional RNA processing.

### Promoter-Associated Transcripts as Signals for Transcriptional Elongation

Various genome-wide studies have revealed short transcripts associated with the TSS, but their functional significance has remained undefined (Esteller, 2011). The study of Tat-dependent gene expression demonstrates the importance of the promoter-proximal TAR element in transcription activation, but it has been unclear whether this is a widely used mechanism for activation of cellular genes in mammalian cells. We now provide evidence that SRSF2 is able to activate transcription via a promoter-proximal ESE. Because of its broad RNA-binding specificity, by recognizing a highly degenerate SSNG motif (S = G or C, N = any nucleotides) (Daubner et al., 2012), this SR protein appears to be particularly suitable for such a role in activating a large array

of cellular genes via binding to diverse sequences in the promoter-proximal nascent RNA.

## Implications on Dynamic Nuclear Structures Involved in Regulated Gene Expression

Our findings have interesting implications in the organization of mammalian genomes in the nucleus. The 7SK complex colocalizes with SR proteins in nuclear speckles, a specific nuclear domain enriched with the splicing machinery. As suggested earlier and emphasized more recently, nuclear speckles are likely the consequence of clustered gene expression events (Dundr and Misteli, 2010; Singer and Green, 1997), which may permit efficient coupling between transcription and cotranscriptional RNA processing (Han et al., 2011). Our current discovery has gone one step further by showing that SRSF2 can directly activate transcription, whereas SRSF1 and perhaps other SR family members may largely affect transcription indirectly via their contribution to the structural integrity of the speckled nuclear domain, which has been recently shown to be critical for efficient gene expression (Tripathi et al., 2012). However, it would be premature to rule out the possibility that other SR proteins may affect transcription via RNA elements away from TSS, as we clearly detected SRSF1-dependent activation of a HIV-based reporter. Importantly, our findings emphasize that transcription and cotranscriptional RNA splicing are not simply temporally linked; rather, their efficient coupling likely results from highly integrated machineries to permit coregulation of gene expression at both transcriptional and posttranscriptional levels in mammalian cells.

## EXPERIMENTAL PROCEDURES

### Cell Culture, Protein/RNA Analyses, and Reporter-Based Assays

We treated SRSF1-MEFs with Dox for 4 days and SRSF2-MEFs with Dox for 2 days to deplete respective SR protein to ∼90% because different SR proteins are depleted at different rates. Cell fractionation was performed as described (Cernilogar et al., 2011). RNA and protein analysis by RT-qPCR and western blotting were according to standard procedures, as detailed in the Extended Experimental Procedures, and transient transactivation assays were carried out in HEK293T cells after SR protein overexpression or RNAi.

### P-TEFb Release Assay from the IPed 7SK Complex

The P-TEFb release assay was modified from a published protocol (Krueger et al., 2010). Anti-HEXIM1 was used to IP the 7SK complex followed by incubation with increasing amounts of RNA oligonucleotide at 37°C for 15 min on a thermal mix. After washing, the content that remained on beads was analyzed by SDS-PAGE.

### Genome-wide Analyses

ChIP-qPCR and ChIP-seq were performed as described (Wang et al., 2011). RNA-seq was performed by using the MAPS protocol to profile gene expression based on tags detected at the 3′ end of individual transcripts (Fox-Walsh et al., 2011). The global nuclear run-on assay (GRO-seq) was according to a published protocol (Wang et al., 2011). Statistical analysis of data is detailed in the Extended Experimental Procedures.

### ACCESSION NUMBERS

The ChIP-seq, RNA-seq, and GRO-seq data on SRSF1- and SRSF2-MEFs are available at the Gene Expression Omnibus under the accession number GSE45517. The CLIP-seq data for SRSF1 and SRSF2 are available under GSE44583 and GSE44591.

## REFERENCES

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. 13, 720–731.

Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr. Opin. Cell Biol. 17, 251–256.

Berkhout, B., Gatignol, A., Rabson, A.B., and Jeang, K.T. (1990). TAR-independent activation of the HIV-1 LTR: evidence that tat requires specific regions of the promoter. Cell 62, 757–767.

Byun, J.S., Fufa, T.D., Wakano, C., Fernandez, A., Haggerty, C.M., Sung, M.H., and Gardner, K. (2012). ELL facilitates RNA polymerase II pause site entry and release. Nat. Commun. 3, 633.

Cavaloc, Y., Bourgeois, C.F., Kister, L., and Stévenin, J. (1999). The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. RNA 5, 468–483.

Cernilogar, F.M., Onorati, M.C., Kothe, G.O., Burroughs, A.M., Parsi, K.M., Breiling, A., Lo Sardo, F., Saxena, A., Miyoshi, K., Siomi, H., et al. (2011). Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature 480, 391–395.

Champlin, D.T., Frasch, M., Saumweber, H., and Lis, J.T. (1991). Characterization of a Drosophila protein associated with boundaries of transcriptionally active chromatin. Genes Dev. 5, 1611–1621.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell 44, 667–678.

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

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

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

Daubner, G.M., Cléry, A., Jayne, S., Stevenin, J., and Allain, F.H. (2012). A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. EMBO J. 31, 162–174.

Dundr, M., and Misteli, T. (2010). Biogenesis of nuclear bodies. Cold Spring Harb. Perspect. Biol. 2, a000711.

Esteller, M. (2011). Non-coding RNAs in human disease. Nat. Rev. Genet. 12, 861–874.

Fox-Walsh, K., Davis-Turak, J., Zhou, Y., Li, H., and Fu, X.D. (2011). A multiplex RNA-seq strategy to profile poly(A+) RNA: application to analysis of transcription response and 3′ end formation. Genomics 98, 266–271.

Han, J., Xiong, J., Wang, D., and Fu, X.D. (2011). Pre-mRNA splicing: where and when in the nucleus. Trends Cell Biol. 21, 336–343.

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.

Krueger, B.J., Varzavand, K., Cooper, J.J., and Price, D.H. (2010). The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK. PLoS ONE 5, e12335.

Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. Annu. Rev. Genet. 34, 77–137.

Lin, S., and Fu, X.D. (2007). SR proteins and related factors in alternative splicing. Adv. Exp. Med. Biol. 623, 107–122.

Lin, S., Xiao, R., Sun, P., Xu, X., and Fu, X.D. (2005). Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. Mol. Cell 20, 413–425.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. Nat. Struct. Mol. Biol. 15, 819–826.

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.

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.

Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol. Cell 3, 697–705.

Ott, M., Geyer, M., and Zhou, Q. (2011). The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe 10, 426–435.

Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares, M., Jr., and Fu, X.-D. (2013). Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. Mol. Cell 50, 223–235.

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

Prasanth, K.V., Camiolo, M., Chan, G., Tripathi, V., Denis, L., Nakamura, T., Hübner, M.R., and Spector, D.L. (2010). Nuclear organization and dynamics of 7SK RNA in regulating gene expression. Mol. Biol. Cell 21, 4184–4196.

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

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81, 145–166.

Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. Genes Dev. 18, 755–768.

Sapra, A.K., Ankö, M.L., Grishina, I., Lorenz, M., Pabis, M., Poser, I., Rollins, J., Weiland, E.M., and Neugebauer, K.M. (2009). SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. Mol. Cell 34, 179–190.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell Biol. 7, 557–567.

Sawarkar, R., Sievers, C., and Paro, R. (2012). Hsp90 globally targets paused RNA polymerase to regulate gene expression in response to environmental stimuli. Cell 149, 807–818.

Schaal, T.D., and Maniatis, T. (1999). Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. Mol. Cell. Biol. 19, 261–273.

Singer, R.H., and Green, M.R. (1997). Compartmentalization of eukaryotic gene expression: causes and effects. Cell 91, 291–294.

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.

Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A., 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.

Tripathi, V., Song, D.Y., Zong, X., Shevtsov, S.P., Hearn, S., Fu, X.D., Dundr, M., and Prasanth, K.V. (2012). SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles. Mol. Biol. Cell 23, 3694–3706.

Wang, Y., Cheong, C.G., Hall, T.M., and Wang, Z. (2009). Engineering splicing factors with designed specificities. Nat. Methods 6, 825–830.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474, 390–394.

Yang, Q., Lin, C., Liu, W., Zhang, J., Ohgi, K.A., Grinstein, J.D., Dorrestein, P.C., and Rosenfeld, M.G. (2011). ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell 147, 773–788.

Yang, Z., Yik, J.H., Chen, R., He, N., Jang, M.K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Mol. Cell 19, 535–545.

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

Young, T.M., Tsai, M., Tian, B., Mathews, M.B., and Pe'ery, T. (2007). Cellular mRNA activates transcription elongation by displacing 7SK RNA. PLoS ONE 2, e1010.

Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G., and Fu, X.D. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol. Cell 35, 1–10.

Zhou, Q., Li, T., and Price, D.H. (2012). RNA Polymerase II Elongation Control. Annu. Rev. Biochem. 81, 119–143.

Zobeck, K.L., Buckley, M.S., Zipfel, W.R., and Lis, J.T. (2010). Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. Mol. Cell 40, 965–975.