THE PROTEIN SOCIETY

The yeast transcription elongation factor Spt4/5 is a sequence-specific RNA binding protein

Amanda J. Blythe,[1] Berra Yazar-Klosinski,[2] Michael W. Webster,[3] Eefei Chen,[4] Marylène Vandevenne,[3] Katerina Bendak,[3] Joel P. Mackay,[3] Grant A. Hartzog,[2] and Alice Vrielink[1]*

[1]School of Chemistry and Biochemistry, University of Western Australia, Crawley, Western Australia, 6009, Australia
[2]Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, California 95064
[3]School of Molecular Bioscience, University of Sydney, Sydney, New South Wales 2006, Australia
[4]Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064

Received 22 May 2016; Accepted 29 June 2016
DOI: 10.1002/pro.2976
Published online 4 July 2016 proteinscience.org

Abstract: The heterodimeric transcription elongation factor Spt4/Spt5 (Spt4/5) tightly associates with RNAPII to regulate both transcriptional elongation and co-transcriptional pre-mRNA processing; however, the mechanisms by which Spt4/5 acts are poorly understood. Recent studies of the human and Drosophila Spt4/5 complexes indicate that they can bind nucleic acids *in vitro*. We demonstrate here that yeast Spt4/5 can bind in a sequence-specific manner to single stranded RNA containing AAN repeats. Furthermore, we show that the major protein determinants for RNA-binding are Spt4 together with the NGN domain of Spt5 and that the KOW domains are not required for RNA recognition. These findings attribute a new function to a domain of Spt4/5 that associates directly with RNAPII, making significant steps towards elucidating the mechanism behind transcriptional control by Spt4/5.

Keywords: Spt4/5; transcription elongation; RNA binding; SELEX; RNA polymerase; transcription elongation factor

Introduction

Transcription is a highly dynamic and regulated process, which in eukaryotes is carried out by three multi-subunit RNA polymerases (RNAPI, II, and III) to produce distinct classes of RNA. These RNAPs are related through their common evolutionary histories, their structures and the mechanisms by which they transcribe RNA in a DNA-template-dependent manner. The catalytic cores of multi-subunit polymerases found in all living organisms display deep conservation.[1] In contrast, the general regulatory proteins that assist and direct the activities of these RNAPs generally exhibit significant variation in structure and function across the three kingdoms of life and between functional classes of polymerase. However, a single family of transcriptional regulators displays the same degree of conservation as seen in RNAP, namely the Spt5/NusG family.[1] These proteins are essential for life and are known to regulate transcription elongation in eukaryotes, archaea and bacteria. Their strict conservation across all domains of life suggests that they carry out an ancient, core function in transcription. The details of that function are, however, largely unknown.

Eukaryotic Spt5 is a large multi-domain protein consisting of an N-terminal acidic domain, a NusG

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Alice Vrielink, School of Chemistry and Biochemistry, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia. E-mail: alice.vrielink@uwa.edu.au

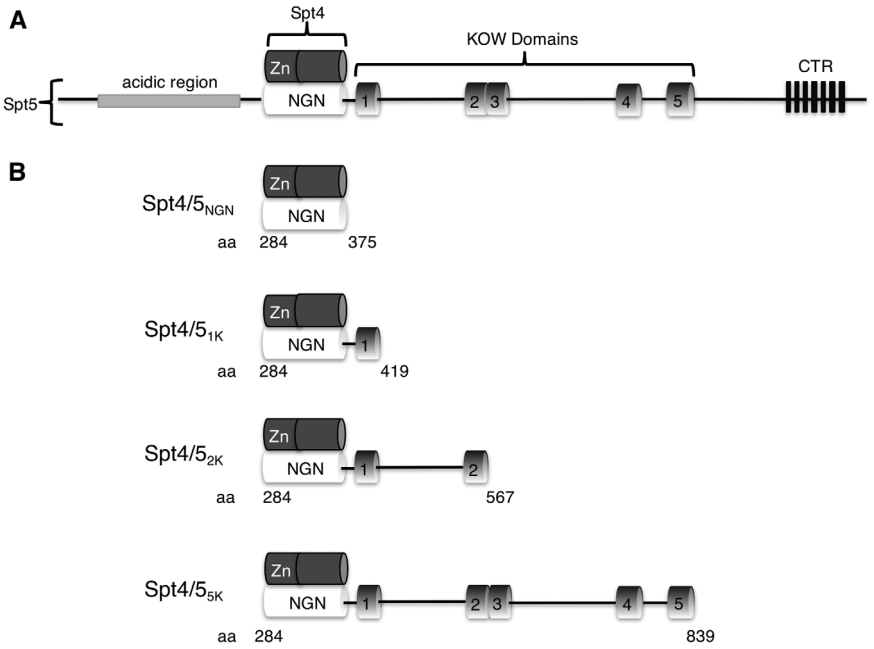


Figure 1. The organization of the Spt4/5 heterodimer and design of expression constructs. An updated Spt5 domain delineation has since been published.[42] A: The domains of Spt4/5, including the zinc finger (Zn) of Spt4 and the acidic N-terminus, the five Kyrpides, Ouzounis, and Woese (KOW) domains and C-terminal repeat (CTR) region of Spt5. B: The names and domain truncations of the Spt4/5 constructs utilized in this study. The amino acid (aa) boundaries of each construct are shown.

N-terminal (NGN) domain, several Kyprides, Ouzounis, Woese (KOW) domains and a set of C-terminal repeats (CTRs) whose sequence varies across species[2,3] [Fig. 1(A)]. The conserved core of Spt5, comprising the NGN domain and a single KOW domain, is found in archaeal Spt5 and the bacterial homolog NusG. In eukaryotes and archaea (but not bacteria), Spt5 forms a noncovalent complex with a small zinc finger protein, Spt4 (RpoE" in archaea), via its NGN domain.[4–6]

The mechanism by which the Spt4/5 heterodimer regulates elongation is not well understood; however, recent structural studies have suggested that the NGN domains of NusG and Spt4/5 bind directly to RNAP, bridging its central cleft.[7–10] This arrangement effectively seals the DNA into the elongation complex and may prevent the disengagement of the DNA template from the transcribing RNAP, enhancing the processivity of the elongating RNAP. This model also places Spt4/5/NusG in a location suitable for allosterically modulating the RNAP active site and interacting with nucleic acids in the transcription elongation complex (TEC). Supporting this model, archaeal Spt4/5 has been shown to interact with both ds- and ssDNA;[11] NusG has recently been shown to interact with T-rich nontemplate DNA in the transcription bubble inducing pausing of the elongation complex;[12] and, Spt5 has been shown to interact with nontemplate DNA in the transcription bubble and upstream of the elongating RNAP II, which appears to critical for the ability of Spt5 to modulate transcription pausing or arrest.[13]

Several observations indicate that, in addition to regulating elongation, Spt4/5 may couple the activities of the TEC to pre-mRNA processing. The CTR region of Spt5 is required for normal elongation control[14,15] and genetically interacts with the CTD of RNAPII.[16,17] Through its regulated phosphorylation, the CTR of Spt5 may serve as a scaffold for the cooperative assembly of transcription and RNA processing factors.[15,16,18–21] Biochemical and proteomic studies show that Spt5 associates with a wide variety of 5' and 3' RNA processing factors including RNA capping enzymes, polyadenylation factors and RNA cleavage factors.[22–27] Additionally, *spt5* mutations affect splicing, polyadenylation and nuclear export of mRNA;[26,28–31] Spt5 has also been shown to facilitate splicing independently of transcription.[32]

Recent evidence suggests that Spt4/5 may exert its effects on elongation or processing through direct contacts with the nascent transcript. Both NusG and Spt4/5 can interact with nucleic acids.[11–13,33–36] Furthermore, *in vitro* transcription studies suggest that Spt4/5 associates with nascent transcripts soon after they emerge from the elongating polymerase, leading to the model that efficient association of Spt4/5 with RNAPII may depend on transcript binding.[13,37,38] Several observations have led to the suggestion that the KOW domains of Spt5 may mediate RNA binding. KOW domains are found in RNA helicases and ribosomal proteins,[39] and in rRNA processing factor Mtr4 they have been observed to directly contact RNA.[39–41] In yeast Spt5, the first KOW domain and sequences immediately

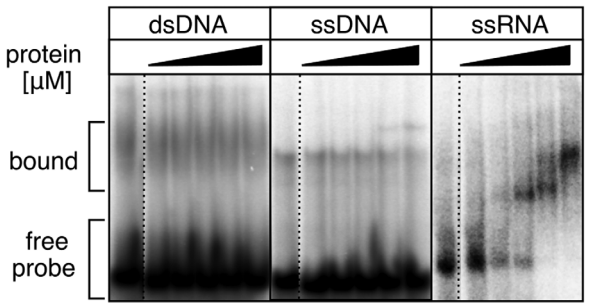


Figure 2. Gel shifts showing that Spt4/5 binds ssRNA with higher affinity than ssDNA or dsDNA. Increasing concentrations of protein from 0-10 μ*M* were electrophoresed with approximately 3 n*M* ^{32}P labeled dsDNA, ssDNA, and RNA pentaprobes. The dotted line distinguishes lanes cut between the probe alone and other samples for each gel.

downstream of it bind RNA and DNA with micromolar affinities.[42] In contrast, the 2nd through 5th KOW domains of Spt5 do not appear to bind nucleic acids, and may instead help mediate binding to polymerase.[42–44] As yet, it is not clear that KOW domains 2-5 of Spt5 contribute to the nucleic acid binding activity of Spt4/5.

We show here that Spt4/5 has sequence-specific RNA-binding activity, and that the Spt4/5$_{NGN}$ heterodimer—but not the KOW domains—is responsible for this functionality. We further show that Spt4/5 displays specificity for RNA bearing multiple AA repeat elements, and discuss ways in which Spt4/5 may use this motif to exert its regulatory effects on transcription elongation.

Results

Spt4/5 binds to ssRNA in vitro

To address the RNA-binding properties of Spt4/5, we first tested the ability of a recombinant yeast Spt4/5 complex containing the Spt5 NGN and all five KOW domains (Spt4/5$_{5K}$) [Fig. 1(B)] to bind double stranded DNA (dsDNA), single stranded DNA (ssDNA) and single stranded RNA (ssRNA) pentaprobes using electrophoretic mobility shift assays (EMSAs). Pentaprobes are overlapping 100-nt oligonucleotides encompassing all possible 5-nt sequences.[45,46] Spt4/5$_{5K}$ gave rise to a clear concentration dependent shift in the ssRNA probe (Fig. 2), whereas no binding to the dsDNA probe was observed and only small amounts of ssDNA shifting could be seen at the highest protein concentrations. These data show that the Spt4/5 complex binds preferentially to ssRNA.

Spt4/5 is a sequence-specific ssRNA-binding protein

To determine whether Spt4/5 recognizes ssRNA with any sequence or structural specificity, we carried out systematic evolution of ligands by exponential enrichment (SELEX) experiment[47,48] using a library containing a random 24-nt region. Selection of the library was performed against an Spt4/5 heterodimer in which Spt5 comprised the NGN domain and the first KOW domain (Spt4/5$_{1K}$), a version containing the NGN domain and the first 2 KOW domains (Spt4/5$_{2K}$) and Spt4/5$_{5K}$, in order to identify any differences in binding specificity between the constructs. Ubiquitin was used as a negative control. Enrichment of the SELEX pool through the rounds of selection was monitored following every second round using RNA EMSA (REMSA) [Fig. 3(A)] with the enrichment defined by the complete protein dependent shift of the library for any given round. In comparison to the negative control, the enrichment of the Spt4/5 constructs was evident after two to seven rounds of selection. The rate of library enrichment was related to the number of KOW domains in the Spt4/5 complexes; a significant proportion of the library was enriched after round two for Spt4/5$_{5K}$, round four for Spt4/5$_{2K}$ and round seven for Spt4/5$_{1K}$ [Fig 3(A)].

For each multi-KOW construct, we found that the most significantly enriched motif (5_AA) is 14 nt long and consists of the sequence AANAANAANAANAA, where N denotes any nucleotide [Fig. 3(B)]. Structure predictions carried out for sequences containing this motif revealed a distinct lack of secondary structure (data not shown). To determine whether the repetitive AA motif was selected for its ability to bind Spt4/5 or as a result of PCR bias (or some other bias), REMSAs and microscale thermophoresis (MST) were carried out using ssRNA probes lacking the AA repeats [Fig. 3(C), 3(D)]. Strong selectivity for the AA repeat sequence was observed; 5_AA bound to Spt4 and a truncated form of Spt5 containing just the NGN domain (Spt4/5$_{NGN}$; see section below for details regarding this construct) with a dissociation constant of 0.65 μ*M*, whereas there was no measurable binding to sequences lacking the AA repeats. Thus, the specificity of Spt4/5 for ssRNA is sequence based.

The Spt4/5 NGN domain is sufficient for RNA binding

Since there was no difference in the enriched SELEX motif between the three Spt4/5 constructs, we considered that some of the KOW domains may be dispensable for RNA binding. To determine the essential RNA-binding core of Spt4/5, we carried out REMSAs using a 5_AA containing sequence obtained from the SELEX experiment (AA$_{rich}$, Supporting Information Table S1) as the target probe. We tested the binding of a series of Spt4/5 complexes comprising: (i) Spt4 with the just NGN domain of Spt5 (Spt4/5$_{NGN}$) or (ii) Spt4/5$_{1K}$, Spt4/5$_{2K}$ and Spt4/5$_{5K}$. Surprisingly, the complex composed only of Spt4 and the isolated NGN domain (Spt4/5$_{NGN}$) bound RNA

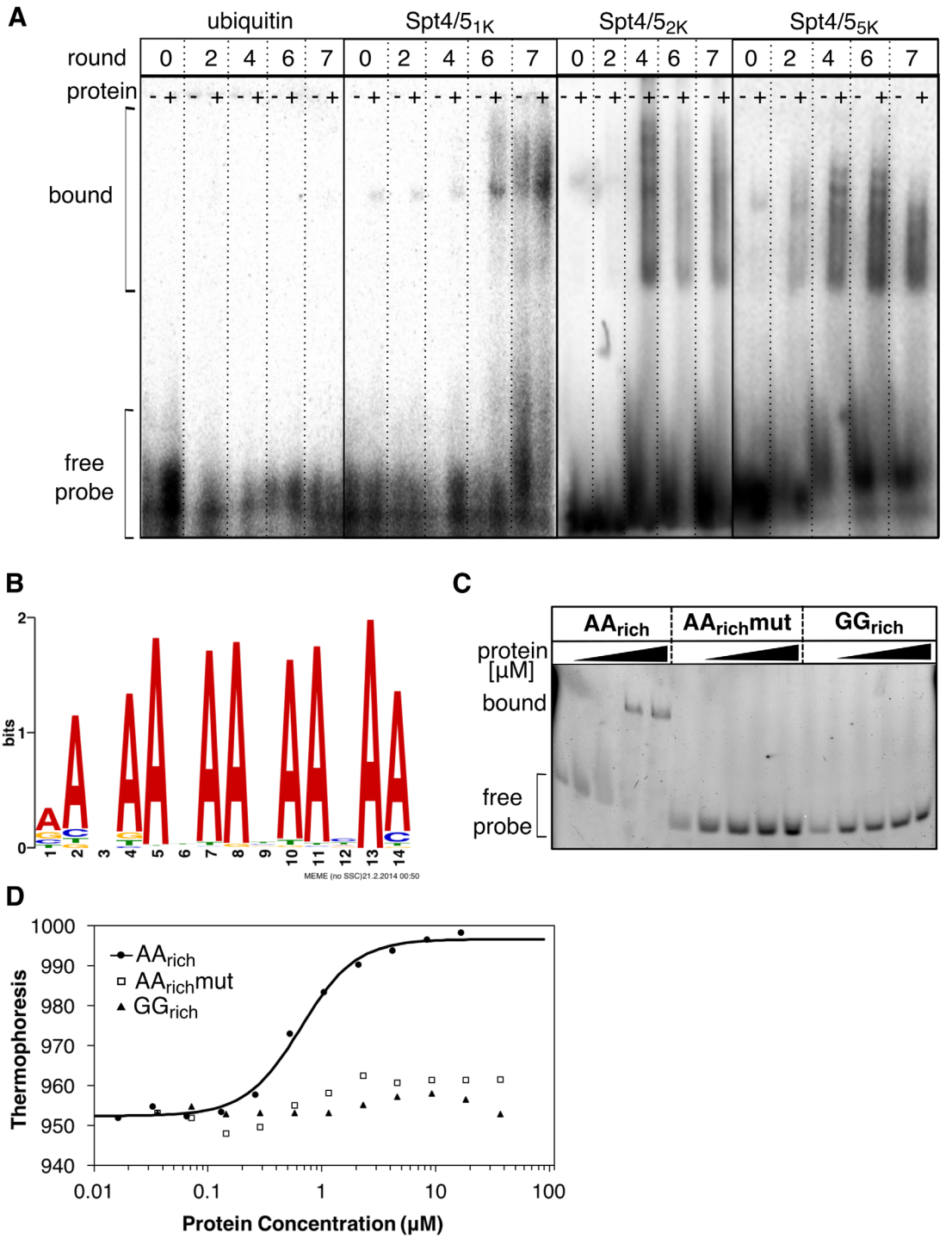


Figure 3. Spt4/5 SELEX shows enrichment of RNA containing AA repeats. A: REMSAs of the initial library (0) and round 2, 4, 6, and 7 for each multi-KOW domain Spt4/5 complex and for the ubiquitin control. ^{32}P labeled probes were incubated with (+) or without (-) 10 μ*M* protein and electrophoresed on a polyacrylamide gel. The dashed lines delineate each round. B: The sequence logo derived from MEME analysis of 2000 unique sequences obtained from a SELEX experiment of Spt4/5$_{5K}$. The motif encompasses 86% of those sequences. C: REMSAs. Fluorescently labeled RNA probes 24nt in length were electrophoresed in the presence of increasing concentrations of Spt4/5$_{5K}$. AA$_{rich}$ is a representative 24nt sequence obtained from SELEX; it harbors 5 AA repeats each separated by a single nucleotide. AA$_{rich}$mut is similar to AA$_{rich}$ with the AA repeats mutated to non-A nucleotides. GG$_{rich}$ comprises the same sequence as AA$_{rich}$ with the AA repeats replaced by GG. D: Microscale thermophoresis (MST) data measuring the change in thermophoresis of fluorescently labeled RNA sequences in the presence of increasing concentrations of Spt4/5$_{NGN}$. A K_d of 0.65 ± 0.2 μ*M* was derived for the binding of Spt4/5 to AA$_{rich}$. No significant change in thermophoresis was observed when Spt4/5 was incubated with either AA$_{rich}$mut or GG$_{rich}$. Symbols show actual data points while the curves represent a fit of the data to a simple 1:1 binding isotherm, yielding K_d estimates and the error associated with each fit. Each data point shown is the average of three independent measurements.

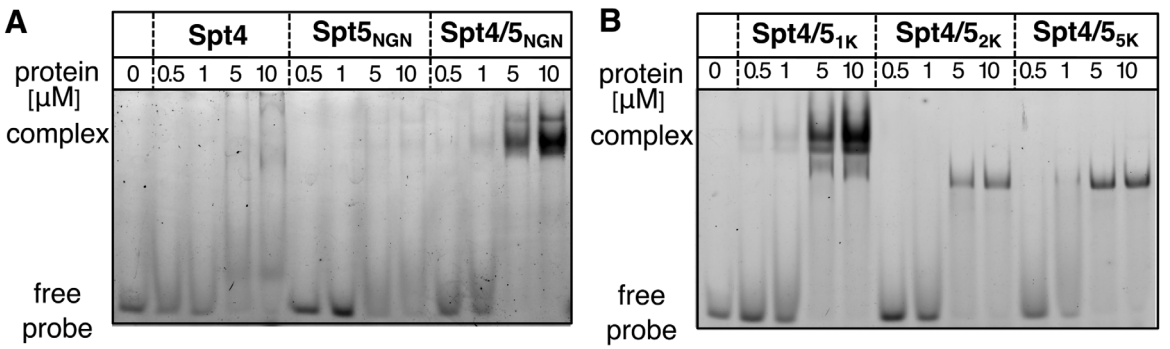


Figure 4. The Spt4/5_{NGN} domain is sufficient for RNA binding. A: 5 n*M* AA_{rich} RNA was electrophoresed in the presence of increasing concentrations of Spt4, $Spt5_{NGN}$ the Spt4/5_{NGN} complex. B: 5 n*M* AA_{rich} RNA was electrophoresed in the presence of increasing concentrations of Spt4/5 complexes with one (Spt4/5_{1K}), two (Spt4/5_{2K}) or five (Spt4/5_{5K}) KOW domains. All probes were fluorescently labeled.

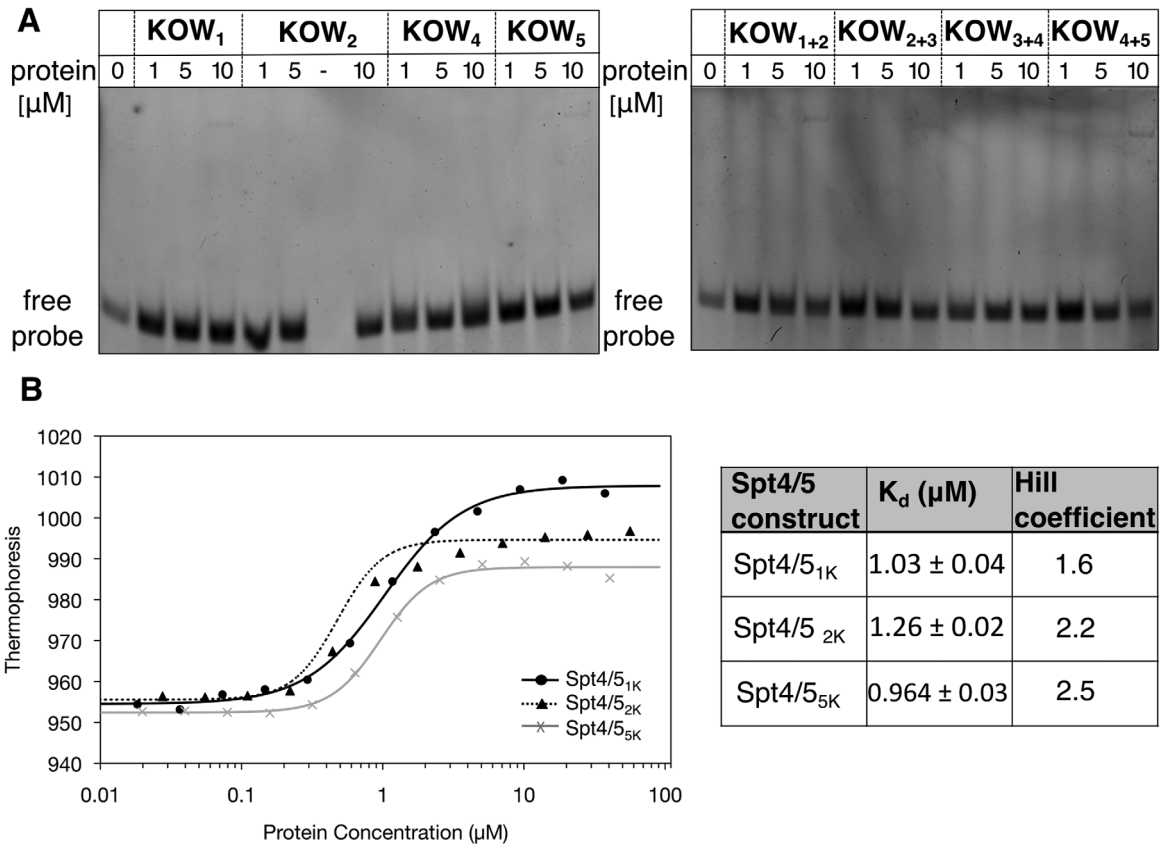
with an affinity that was indistinguishable from the affinities measured for the KOW-containing complexes (Fig. 4). These data demonstrate that the KOW domains are not required for sequence-specific RNA binding. Multiple shifted species were observed for Spt4/5_{NGN} and Spt4/5_{1K}, suggesting that more than one Spt4/5 monomer is able to bind a single RNA molecule. In contrast, binding of the Spt4/5_{2K} and Spt4/5_{5K} complexes resulted in a single stable species [Fig. 4(B)], suggesting that the presence of multiple KOW domains perhaps inhibits the binding of multiple Spt4/5 complexes to the RNA probe. Intriguingly, the shifted Spt4/5_{2K} and Spt4/5_{5K} complexes showed greater mobility through the gel than the smaller Spt4/5_{NGN} and Spt4/5_{1K} complexes; however, the multiple species observed for these smaller constructs suggests a mixture of protein:RNA complexes with different stoichiometries. This situation would result in an increase in the overall molecular weight of the complexes and could thus account for the greater retardation of migration through the gel. Addition of Spt4 or $Spt5_{NGN}$ alone causes the disappearance of the unbound probe, indicating that each possesses a degree of RNA-binding activity [Fig. 4(A)]. However, there is no corresponding appearance of a stable shifted species, suggesting that both have a reduced affinity and/or a greater rate of dissociation than the Spt4/5 heterodimer. Thus, both Spt4 and Spt5 are required to make high-affinity and sequence-specific interactions with RNA.

The role of KOW domains in the RNA binding of Spt4/5

That Spt4/5_{NGN} is the minimal RNA-binding region of the complex contradicts the prevailing idea in the literature that the KOW domains are responsible for the nucleic acid binding capabilities of Spt4/5.[8,10,33,42] There appeared to be very weak interactions between RNA and the tandem KOW_{1+2}, and KOW_{4+5} domains [Fig. 5(A)]. These low affinity interaction (most likely of the order of 100 m*M*) are unlikely to be specific; however, the KOW_{1+2} interaction may be related to the recently described RNA binding activity of the Spt5 Kow1 + linker region, which was reported to be more than 10 fold weaker than the RNA binding affinity of Spt4/5_{NGN} reported here.[42] These weak interactions and the observation that no RNA interactions were discernable for the purified KOW domains alone or other tandem constructs [Fig. 5(A)] suggests that the KOW domains do not make a significant contribution to the high-affinity and sequence-specific binding observed for the Spt4/5_{NGN} complex. However, it is notable that Spt4/5 complexes with more than one KOW domain enriched AA-bearing RNA sequences in the SELEX experiment in fewer rounds of selection. Furthermore, MST experiments on the different Spt4/5 heterodimers binding to 5_AA RNA [Fig. 5(B)] show the same microscopic Kd ($\sim$1 μ*M*) along with a clear increase in the calculated Hill coefficient as more KOW domains are added. These measurements suggest the existence of positive cooperativity; that is, that a single full-length Spt4/5 heterodimer binds AA-rich RNA with a micromolar affinity and that the presence of this protein on the RNA increases the affinity of a subsequent heterodimer for the same RNA molecule. In this scenario, the single REMSA band observed for Spt4/5_{5K} binding to RNA would represent a single multimeric complex (i.e. 2:1 or more) that is formed cooperatively, whereas the multiple bands observed for Spt4/5_{NGN} represent a mixture of multiple protein-RNA assemblies (i.e. 1:1 and 2:1 complexes) binding without cooperativity.

Spt4/5 heterodimer formation is associated with conformational change

Since neither Spt4 nor Spt5 bound RNA strongly in isolation, we considered the possibility that a conformational change of one or both subunits is required for high-affinity binding. We tested this hypothesis by comparing far-UV circular dichroism (CD) spectra



Spt4/5 construct	K_d (μM)	Hill coefficient
Spt4/5_{1K}	1.03 ± 0.04	1.6
Spt4/5_{2K}	1.26 ± 0.02	2.2
Spt4/5_{5K}	0.964 ± 0.03	2.5

Figure 5. KOW domains are not involved in RNA binding *in vitro* A. 5 n*M* of fluorescently labeled AA_{rich} RNA was electrophoresed in the presence of increasing concentrations of single or tandem KOW domains. No significant RNA binding by either the single or tandem KOW domains alone was observed. B. The interaction of AA_{rich} RNA with increasing concentrations of Spt4/5_{1K}, Spt4/5_{2K} or Spt4/5_{5K}, as determined by MST. The presence of KOW domains does not affect the RNA-binding affinity of Spt4/5. Symbols show actual data points while the curves represent a fit of the binding data giving K_d estimations and the error associated with the fit. Each data point is the average of 3 independent measurements.

for Spt4, $Spt5_{5K}$ and Spt4/5_{5K} [Fig. 6(A)]. Spectra characteristic of a structured protein (containing a mixture of α-helices and β-strands) were obtained for the individual Spt4 and $Spt5_{5K}$ proteins, as expected from the available domain/subcomplex structures of the heterodimer.[5] Interestingly the spectrum of the Spt4/5_{5K} complex was distinct from a concentration weighted average of the spectra of Spt4 and $Spt5_{5K}$. Thus, the formation of the Spt4/5 heterodimer is associated with a conformational change in one or both subunits and this change is likely to promote RNA-binding activity.

The NGN domain of Spt4/5 contains a novel heterodimeric RRM

In light of the RNA-binding activity seen for Spt4/5_{NGN}, the topology of the NGN domain of Spt5 was compared to other RNA binding modules, revealing a remarkable similarity to the well-established RNA recognition motif (RRM) [Fig. 6(B)]. This similarity was noted previously after the determination of the crystal structure of the bacterial homologue NusG,[33] but has eluded the descriptions of the eukaryotic and archaeal complexes.[4,5,10,44] The NGN domain of yeast Spt5 contains the typical RRM arrangement of four antiparallel β-strands packed against two α-helices although its α1-helix is extended and it has an additional C-terminal α-helix (α3) that partially obscures what would be the RNA-binding surface of a canonical RRM [Fig. 6(B)]. Spt4 interacts with the NGN domain of Spt5 through the alignment of their respective β-sheets.[5] In the context of an RRM, this effectively extends the β-sheet surface through an additional two β-strands and brings in another α-helix that further obscures the canonical RNA-binding surface [Fig. 6(C)]. Most canonical RRMs interact with ssRNA through three highly conserved aromatic rings located within two RNP motifs on the β-sheet surface. Spt5 NGN lacks these canonical RNP sequences and conserved surface aromatics suggesting that Spt4/5 binds RNA in a noncanonical manner.

Discussion

Our *in vitro* data show that yeast Spt4/5 is a sequence-specific RNA-binding protein that recognizes multiple AA repeat sequences with an affinity that is consistent with those observed for many other sequence-specific RNA-binding proteins.[49] We

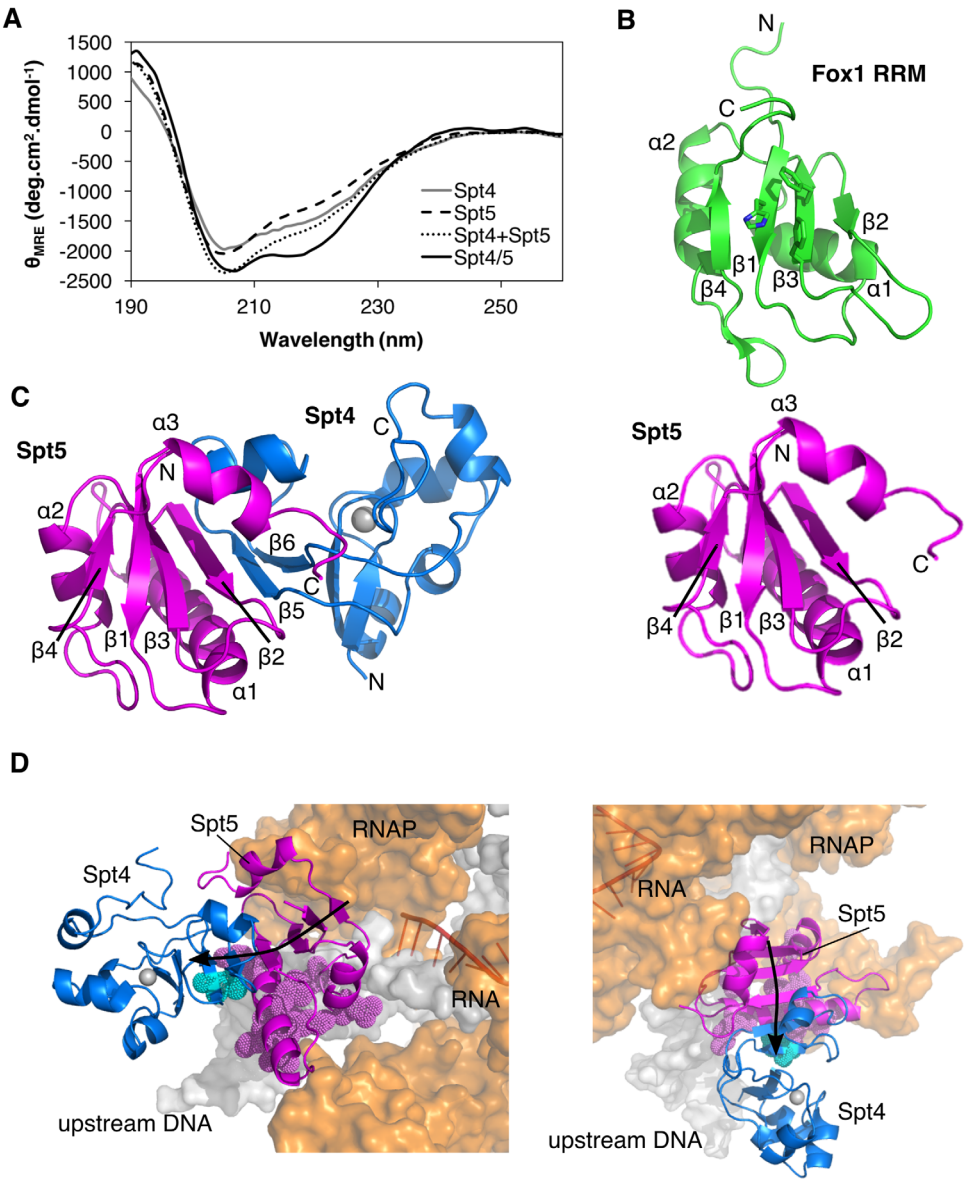


Figure 6. Structural insight into the mechanism of RNA binding by Spt4/5. A: Far-UV CD spectra of Spt4, $Spt5_{5K,}$ and $Spt4/5_{5K}$ compared to the theoretical sum of Spt4 + $Spt5_{5K}$ show that Spt4 modulates the structure of Spt5. B: Cartoon representation of Human Fox 1 (PDB entry 2ERR; green) showing a canonical RRM topology, and of the NGN domain of Spt5 (PDB entry 2EXU;[5] magenta) showing a similar RRM topolgy with an extended α1 helix and an additional C-terminal helix (α3). C: The structure of the yeast $Spt4/5_{NGN}$ heterodimer (PDB entry 2EXU). Spt4 (blue) contributes two additional β-strands to the RRM of Spt5 (magenta). D: A model of yeast $Spt4/5_{NGN}$ (PDB entry 2EXU) bound to yeast RNAPII (PDB entry 5C4X[69]) showing the RRM surface (black arrow) is distinct from the surface/residues proposed to interact with both RNAPII and the upstream DNA (cyan/violet dot representation). Spt4/5 is shown in cartoon representation. RNAPII (orange) and DNA (white) is shown in surface representation and the nascent RNA (red) in cartoon representation. The model was made by superimposing the yeast Spt4/5 and RNAPII with the crystal structure of the archaeal Spt4/5:RNAPII complex.[8].

were surprised to discover that Spt4/5 binds RNA in a sequence-specific manner, given that the distribution of Spt4/5 on genes closely mirrors that of RNAPII.[50] Intriguingly, a recent study demonstrated that Spt4 is required for normal transcription of genes containing trinucleotide repeats and that genes with $(AAA)_N$ repeats showed the strongest dependence on Spt4.[51] Thus, the RNA binding activity of Spt4/5 may facilitate expression of genes with A-rich sequence motifs *in vivo*.

The $NusG_{NGN}$ domain alone has recently been shown to be sufficient for DNA binding;[12] however, we show that both Spt4 and Spt5 are required to make significant RNA contacts. This raises the question of whether the specific RNA-binding activity we observed for yeast Spt4/5 is conserved in other Spt5 homologs, especially NusG, which lacks an Spt4 homolog. The recruitment of *Drosophila melanogaster* and *S. cerevisiae* Spt5 to the RNAP elongation complex is dependent on the presence of a

nascent transcript longer than 18nt.[13,37,38] This is in direct contrast to NusG, which does not require RNA interactions to associate with RNAP.[12] Therefore, we propose that RNA-binding by Spt4/5 may be specific to the eukaryotic elongation complexes.

RNA binding activity has been suggested for NusG[33] and the yeast Sp4/5[34] although the interactions were not biochemically characterized and the studies did not delineate the specific domains responsible for the interaction. The KOW domains of NusG/Spt5 have previously been proposed to mediate nucleic acid binding;[8,10,33,42] however, we show here that the NGN domain alone is sufficient to bind RNA. Our binding data suggest that the KOW domains promote the cooperative binding of multiple Spt4/5 heterodimers to RNA but the biological relevance of this cooperative binding is yet to be established. It is also possible that the KOW domains mediate interactions with other components of the TEC. Consistent with this hypothesis, fragments of Spt5 containing KOW domains but lacking the NGN or CTR domains are capable of binding RNAPII in pulldown assays,[34,43,44] and the KOW 4-5 domains have been shown to make extensive contacts with the dissociable subunit of RNAP, Rpb4/7.[13,52]

The NGN domain of Spt5 can be considered a variant RRM, one of the most common RNA-binding domains. Interestingly, in archaea, humans and yeast Spt5, the β-sheet of the Spt5 RRM is extended by its heterodimerization interface with Spt4.[5] Extension of the β-sheet has been previously reported in proteins containing tandem RRMs, including the poly(A)-binding protein[53] and the polypyrimidine tract binding protein;[54] this extension is thought to help accommodate longer RNA motifs and achieve higher affinity for the RNA target. However, extension of the β-sheet by heterodimerization as seen here with Spt4/5 has never before been reported. Thus, the conformational change in either Spt4 and/or Spt5 upon formation of the heterodimer is likely mediated by this novel inter-protein RRM and is consistent with our observations that both Spt4 and Spt5 are required for the interaction with RNA.

Sequence-specific DNA binding activity has been observed for NusG[12] and its homolog RfaH,[55] and archaeal Spt4/5 has been shown to bind both ssDNA and dsDNA.[11] Although there is growing evidence to suggest that Spt4/5 interacts with DNA in the context of transcription elongation complexes, we did not observe significant DNA-binding activity for yeast Spt4/5. This suggests that in the absence of RNAP, the primary nucleic acid binding activity of Spt4/5 is directed toward RNA. The NGN domain of RfaH is recruited to elongation complexes via its recognition of the *ops* sequence element.[56] Mutational analysis of RfaH localized the DNA recognition to a polar patch on the surface of the NGN domain[56] that is analogous to the Spt4/5$_{NGN}$ surface proposed to interact with the nontemplate DNA in the TEC[7,11,13] but distinct from the RRM binding surface [Fig. 6(D)]. Thus while the modes are likely different, the recruitment of Spt4/5 to the TEC via recognition of A-rich sequences presents an attractive model for the sequence-specific RNA-binding activity of Spt4/5$_{NGN}$. The specific RNA-interactions we observed do not rule out the possibility that yeast Spt4/5 can also interact nonspecifically with other nucleic acids in the context of the elongation complex; rather, it suggests that the complex regulation of the eukaryotic transcription elongation machinery by Spt4/5 is subject to multiple modes of nucleic acid modulation.

Consistent with this idea, recent studies suggest that the proposed DNA binding activities of Spt4/5 lie on distinct faces of the protein and are separate to the Spt4/5 RRM surface [Fig. 6(D)]. The proposed interaction site for DNA-binding by the archaeal, yeast and human Spt4/5 complex is a highly conserved basic patch on the] α1 helix of Spt5$_{NGN}$ [Fig. 6(D)].[7,11,13] Homologous regions on the NGN domains of NusG and its paralog RfaH have also been proposed to contact the nontemplate strand in the transcription bubble.[12,56–59] Furthermore, Spt4 together with the first KOW domain of Spt5 and the NGN-KOW1 linker are proposed to contact nucleic acids upstream of the transcription bubble [Fig. 6(D)].[42] Other published data show that additional regions on Spt5, other than the ones already characterized for the RNAP-Spt4/5 clamp interface, interact with the surface of RNAPII.[13,42,52,60,61] These observations suggest multiple modes of interaction with RNAPII, and it is becoming a common train of thought that Spt4/5 is able to associate with multiple regions of RNAPII and nucleic acids in the transcription elongation complex.[13,42] This may provide a mechanism whereby nucleic acid binding by Spt4/5 triggers allosteric modulation of RNAPII activity.

We suggest three possible nonexclusive models for functions mediated by the sequence-specific RNA-binding activity of Spt4/5. First, the binding of A-rich RNA sequences by Spt4/5 might promote the recruitment of the heterodimer to RNAP to stabilize the transcription elongation complex [Fig. 7(A)]. Consistent with this idea, several prior *in vitro* studies have demonstrated increased association of Spt4/5 with elongation complexes that contain RNA transcripts long enough to protrude beyond the polymerase[37,38] and Spt5 has been shown to cross-link to RNA near the exit channel.[13] Second, because Spt4/5 also associates with pre-mRNA processing and regulatory factors,[22–26] sequence-specific RNA-binding by Spt4/5 might facilitate pre-mRNA processing by enhancing the recruitment of processing machinery to their pre-mRNAs targets [Fig. 7(B)]. Such a function would explain observations of RNA processing

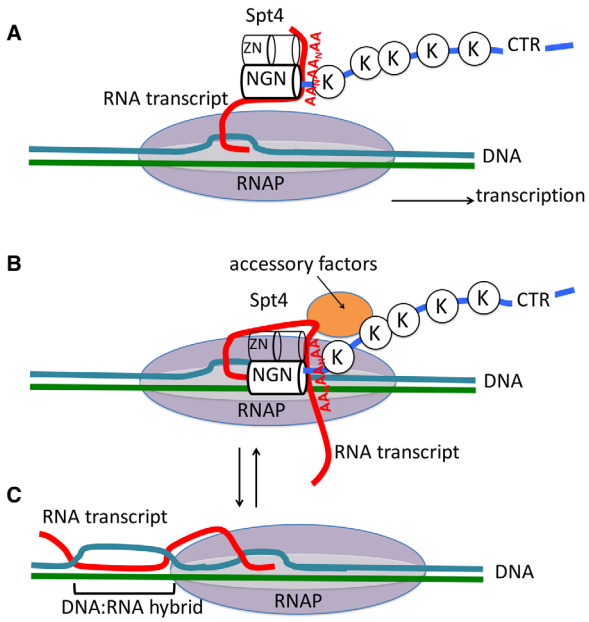


Figure 7. Possible functions mediated by Spt4/5's RNA-binding activity. A: RNA binding by Spt4/5 may promote the recruitment of the heterodimer to RNA polymerase II (RNAP) to stabilize the transcription elongation complex. B: By binding over the central cleft of elongating RNAP, Spt4/5 may effectively seal the DNA template into the elongation complex, ensuring polymerase processivity. In addition, the KOW domains and the Spt5 CTR may recruit additional accessory factors that can interact with the elongating polymerase, RNA transcript or DNA template as it enters or exits the polymerase. **C.** In the absence of Spt4/5 and factors that associate with it, R-loops—persistent DNA:RNA hybrids—may form behind the elongating polymerase, leading to elongation defects and DNA damage.

defects in *spt4* and *spt5* mutants (reviewed in Ref. 3). Finally, by binding A-rich tracts in nascent transcripts Spt4/5 may directly or indirectly influence formation or stability of extended RNA:DNA hybrids or R-loops [Fig.7(C)]. Consistent with this idea, *spt4* mutations show strong genetic interactions with a number of mutations associated with R-loop formation[62,63] and, recent evidence suggests that polyA tracts are a major contributor to the formation of hybrid regions in *S. cerevisiae*.[64] Further experiments to elucidate the cross-talk between RNA-bound Sp4/5 and the rest of the TEC will allow these possible mechanisms to be distinguished.

Materials and Methods

Cloning

The cloning of *SPT4* and constructs of *SPT5* encompassing the NGN domain plus one KOW domain (residues 284-419), two KOW domains (residues 284-567) and five KOW domains (residues 284-839) has been described previously.[65] All other constructs were cloned using standard restriction based techniques. Primer sequences are listed in Supporting Information Table S1.

Expression and purification

Spt4 and all Spt4/5 constructs were expressed and purified as described previously.[65] The KOW domain constructs were transformed into *E. coli* Rosetta 2 (DE3) (Novagen) for expression. Transformants containing the recombinant plasmids were incubated in Luria-Bertani (LB) growth medium supplemented with 50 μg mL^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol at 37°C. When the optical density at 600 nm reached 0.6–0.8, protein expression was induced with 0.2 m*M* IPTG. The induced culture was incubated at 25°C overnight, harvested by centrifugation (4000*g*, 4°C, 15 min), the bacterial pellets resuspended in 50 m*M* Tris pH 8, 0.2 M NaCl, 20 m*M* imidazole, 1% Triton-X and lysed using an Emulsiflex C5 high pressure homogeniser (Avestin). The lysate was clarified by centrifugation (20,000 x g, 4°C, 45 min) before loading onto a 1 ml GSTrap column (GE Healthcare) and washing with 50 m*M* Tris pH 8, 300 m*M* NaCl, 10% glycerol. Affinity chromatography was performed on an ÄKTA purifier (GE Healthcare) and recombinant protein eluted with a glutathione gradient (50 m*M* Tris pH 8, 150 m*M* NaCl, 0.005% Triton-X, 20 m*M* glutathione).

Radiolabeled pentaprobe preparation

Pentaprobes were prepared as described.[45] DNA pentaprobes were 5' end-labeled with [γ-^{32}P] ATP (PerkinElmer) using T4 polynucleotide kinase. [α-^{32}P] UTP was incorporated into RNA pentaprobes during *in-vitro* transcription. Unincorporated [^{32}P] nucleotides were removed from DNA pentaprobe labeling reactions using Sephadex® G-25 Quick Spin™ columns and gel purification for the RNA pentaprobes. The reactions were diluted to a consistent count-rate as measured by a Geiger counter and then diluted 50-fold to give an optimal signal for visualization on a Phosphor screen with a Typhoon PhosphoImager™ FLA 9000 (GE Healthcare). Pentaprobe concentrations used for gel shift reactions were estimated to be approximately 0.1 n*M*.

RNA electromobility shift assay

Radiolabeled nucleic acid pentaprobes (~0.1 n*M*) were incubated with protein samples in gel shift buffer (10 m*M* MOPS pH 7.0, 50 m*M* KCl, 5 m*M* $MgCl_2$, 10% glycerol and 0.03 mg mL^{-1} heparin) at 4°C for 30 min. The binding reactions were electrophoresed on a pre-equilibrated 6% native acrylamide/bisacrylamide gel (19:1) in 0.5X Tris-borate buffer at 200 V for 2 h at room temperature. Gels were visualized on a Phosphor screen with a Typhoon PhosphoImager™ FLA 9000 (GE Healthcare).

5' 56FAM (Integrated DNA Technologies) labeled RNA probes (5 n*M*) were incubated with protein samples in gel shift buffer at 4°C for 30 min. The binding reactions were electrophoresed on a pre-equilibrated 6% native acrylamide/bisacrylamide gel (19:1) in 0.5X Tris-Hepes buffer at 65 mA for 1 h at 4°C. Gels were visualized with a Typhoon Trio Variable Mode Imager (Amersham Biosciences) using 488 nm/520 nm excitation/emission wavelengths. The sequences of the RNA probes are listed in Supporting Information Table S1.

Selex

The initial SELEX library was prepared by treating a synthetic oligonucleotide pool containing a 25-nt random sequence flanked by two primer binding sites with DNA polymerase I large (Klenow) fragment (New England Biolabs) at 37°C for 30 min and purified using a PCR clean up kit (QIAGEN). The library was transcribed using the T7-RiboMAX™ Large Scale RNA production system (Promega). Unincorporated nucleotides were removed with Sephadex® G-25 Quick Spin™ columns (Roche) and RNA extracted by phenol/chloroform and ethanol-precipitation. Binding reactions were carried out in SELEX Buffer (40 m*M* MOPS pH 7.0, 20 m*M* KCl, 10 m*M* $MgCl_2$, 10% glycerol, 0.2% Triton X-100, 0.2 m*M* PMSF, 2 m*M* DTT). Each 150 μl binding reaction contained 5–80 pmol of protein i m*M* obilized on MagneHis™ Ni-Particles (Promega) containing 0.06–1.6 mg mL^{-1} heparin sulphate and 80 pmol RNA, and mixed for 60 min at 4°C. Unbound RNA was removed and the beads were washed 5 times with SELEX buffer (500 μL). Bound RNA was eluted from the i m*M* obilized protein in water by vigorous mixing at 95°C for 15 min. The selected RNA was ethanol precipitated and reverse-transcribed using a complementary primer, then amplified by 10 or 15 cycles of PCR with Pfu DNA polymerase. The PCR products were transcribed into RNA and applied to fresh protein coupled MagneHis™ Ni-Particles and the cycle was repeated. A total of seven rounds of SELEX were completed. The primers and SELEX oligonucleotide are listed in Supporting Information Table S1.

High-throughput sequencing and analyses of selected sequences

Purified PCR products from each round of SELEX were re-amplified with barcoded primers and sequenced on the HiSeq2000 (Illumina) sequencing platform using a single-end, 50 nucleotide sequencing protocol at the University of California, Los Angeles Bioengineering. The primers used for the addition of the barcodes are listed in Supporting Information Table S1. Each barcode contained between 2.4 and 5.5 million reads. A workable data set of 10,000 sequences per barcode was taken, and barcodes and duplicate sequences removed. 2000 of these unique sequences were randomly selected and subjected to analysis with the MEME suite. RNA secondary structure analysis was performed using the online RNA secondary structure prediction tool Context Fold.[66]

Circular dichroism spectroscopy and analysis

For Spt4, $Spt5_{5K}$ and $Spt4/5_{5K}$ analysis, far-UV spectra were measured on a CD spectrophotometer (AVIV 60DS, Lakewood, NJ) using a 200-μm path length quartz cuvette. Data were collected every 1 nm with a 1 nm bandwidth in the 180–320 nm wavelength region using an integration time of 8 s per step. The far-UV CD spectra represent the average of 15 scans for $Spt5_{5K}$ and Spt4/5 and 22 scans for Spt4. The data sets were collected for three different sample preparations (~20 μM) each for Spt4, Spt5, and Spt4/5. CD spectra were measured at room temperature.

Microscale thermophoresis

MST experiments were performed on a Monolith NT.115 system (NanoTemper Technologies) using 95% LED and 40% IR-laser power for wild-type Spt4/5 constructs or 60% LED and 40% IR-laser power for Spt4/5 mutants. Laser on and off times were set at 30 s and 5 s respectively. Dilution series were prepared for unlabeled protein in 50 m*M* Hepes pH 7.4, 0.15 M KCl, 0.01% Tween-20 and 0.005% RNase inhibitor with 5' 56-FAM labeled RNA oligonucleotides (IDT) at a final concentration of 50 n*M*. Measurements were performed in standard treated capillaries (NanoTemper Technologies) and the data from three replicate measurements were combined and analyzed using the implemented fitting software NT Analysis (NanoTemper Technologies). Binding isotherms were fitted using the Hill method.

Structure analysis

All structural figures were generated with PyMOL.[67]

Acknowledgments

The authors would like to thank the sequencing facility at UCLA run by Prof. Matteo Pellegrini for running the deep sequencing.

References

1. Werner F, Grohmann D (2011) Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol 9:85–98.
2. Ponting CP (2002) Novel domains and orthologues of eukaryotic transcription elongation factors. Nucleic Acids Res 30:3643–3652.

3. Hartzog GA, Fu J (2013) The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. Biochim Biophys Acta 1829:105–115.
4. Hirtreiter A, Damsma GE, Cheung A, Klose D, Grohmann D, Vojnic E, Martin ACR, Cramer P, Werner F (2010) Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic Acids Res 38:4040–4051.
5. Guo M, Xu F, Yamada J, Egelhofer T, Gao Y, Hartzog GA, Teng M, Niu L (2008) Core structure of the yeast Spt4-Spt5 complex: a conserved module for regulation of transcription elongation. Structure 16:1649–1658.
6. Wenzel S, Martins B, Rosch P, Wohrl B (2009) Crystal structure of the human transcription elongation factor DSIF hSpt4 subunit in complex with the hSpt5 dimerization interface. Biochem J 425:373–380.
7. Bernecky C, Herzog F, Baumeister W, Plitzko JM, Cramer P (2016) Structure of transcribing mammalian RNA polymerase II. Nature 529:551–554.
8. Martinez-Rucobo FW, Sainsbury S, Cheung ACM, Cramer P (2011) Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. Embo J 30:1302–1310.
9. Sevostyanova A, Belogurov GA, Mooney RA, Landick R, Artsimovitch I (2011) The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Mol Cell 43:253–262.
10. Klein BJ, Bose D, Baker KJ, Yusoff ZM, Zhang X, Murakami KS (2011) RNA polymerase and transcription elongation factor Spt4/5 complex structure. Proc Natl Acad Sci USA 108:546.
11. Guo G, Gao Y, Zhu Z, Zhao D, Liu Z, Zhou H, Niu L, Teng M (2015) Structural and biochemical insights into the DNA-binding mode of MjSpt4p:Spt5 complex at the exit tunnel of RNAPII. J Struct Biol 192:418–425.
12. Yakhnin AV, Murakami KS, Babitzke P (2016) NusG is a sequence-specific RNA polymerase pause factor that binds to the non-template DNA within the paused transcription bubble. J Biol Chem 291:5299–5308.
13. Crickard JB, Fu J, Reese JC (2016) Biochemical analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. J Biol Chem 291: 9853–9870.
14. Diamant G, Amir-Zilberstein L, Yamaguchi Y, Handa H, Dikstein R (2012) DSIF restricts NF-kappaB signaling by coordinating elongation with mRNA processing of negative feedback genes. Cell Rep 2:722–731.
15. Chen H, Contreras X, Yamaguchi Y, Handa H, Peterlin BM, Guo S (2009) Repression of RNA polymerase II elongation in vivo is critically dependent on the C-terminus of Spt5. PLoS One 4:e6918.
16. Schneider S, Pei Y, Shuman S, Schwer B (2010) Separable functions of the fission yeast Spt5 carboxyl-terminal domain (CTD) in capping enzyme binding and transcription elongation overlap with those of the RNA polymerase II CTD. Mol Cell Biol 30:2353–2364.
17. Lindstrom DL, Hartzog GA (2001) Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae. Genetics 159:487–497.
18. Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, Hahn S (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Mol Cell Biol 29:4852–4863.
19. Zhou K, Kuo WH, Fillingham J, Greenblatt JF (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. Proc Natl Acad Sci USA 106:6956–6961.
20. Wier AD, Mayekar MK, Heroux A, Arndt KM, Vandemark AP (2013) Structural basis for Spt5-mediated recruitment of the Paf1 complex to chromatin. Proc Natl Acad Sci USA 110:17290–17295.
21. Qiu H, Hu C, Gaur NA, Hinnebusch AG (2012) Pol II CTD kinases Bur1 and Kin28 promote Spt5 CTR-independent recruitment of Paf1 complex. Embo J 31:3494–3505.
22. Wen Y, Shatkin AJ (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. Genes Dev 13: 1774–1779.
23. Pei Y, Shuman S (2002) Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. J Biol Chem 277:19639–19648.
24. Lindstrom DL, Squazzo SL, Muster N, Burckin TA, Wachter KC, Emigh CA, McCleery JA, Yates JR III, Hartzog GA (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. Mol Cell Biol 23:1368–1378.
25. Shen Z, St-Denis A, Chartrand P (2010) Cotranscriptional recruitment of She2p by RNA pol II elongation factor Spt4-Spt5/DSIF promotes mRNA localization to the yeast bud. Genes Dev 24:1914–1926.
26. Burckin T, Nagel R, Mandel-Gutfreund Y, Shiue L, Clark TA, Chong JL, Chang TH, Squazzo S, Hartzog G, Ares M Jr (2005) Exploring functional relationships between components of the gene expression machinery. Nat Struct Mol Biol 12:175–182.
27. Mayer A, Schreieck A, Lidschreiber M, Leike K, Martin DE, Cramer P (2012) The Spt5 C-terminal region recruits yeast 3 ' RNA cleavage factor I. Mol Cell Biol 32:1321–1331.
28. Xiao Y, Yang YH, Burckin TA, Shiue L, Hartzog GA, Segal MR (2005) Analysis of a splice array experiment elucidates roles of chromatin elongation factor Spt4-5 in splicing. PLoS Comput Biol 1:e39.
29. Bucheli ME, Buratowski S (2005) Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. Embo J 24:2150–2160.
30. Kaplan CD, Holland MJ, Winston F (2005) Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. J Biol Chem 280:913–922.
31. Cui Y, Denis CL (2003) In vivo evidence that defects in the transcriptional elongation factors RPB2, TFIIS, and SPT5 enhance upstream poly(A) site utilization. Mol Cell Biol 23:7887–7901.
32. Diamant G, Eisenbaum T, Leshkowitz D, Dikstein R (2016) Analysis of subcellular RNA fractions revealed a transcription-independent effect of TNFalpha on splicing mediated by Spt5. Mol Cell Biol 36:1342–1353.
33. Steiner T, Kaiser JT, Marinkovic S, Huber R, Wahl MC (2002) Crystal Strucures of transcription factor NusG in light of its nucleic acid and protein-binding activities. Embo J 21:4641–4653.
34. Viktorovskaya OV, Appling FD, Schneider DA (2011) Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly. J Biol Chem 286:18825–18833.
35. Andrulis ED, Guzman E, Daring P, Werner J, Lis J (2000) High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev 14:2635–2649.
36. Hanna MM (1993) Photocrosslinking analysis of protein-RNA interactions in *E. coli* transcription complexes. Cell Mol Biol Res 39:393–399.

37. Cheng B, Price DH (2008) Analysis of factor interactions with RNA polymerase II elongation complexes using a new electrophoretic mobility shift assay. Nucleic Acid Res 36:e135.
38. Missra A, Gilmour DS (2010) Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. Proc Natl Acad Sci USA 107:11301–11306.
39. Weir JR, Bonneau F, Hentschel J, Conti E (2010) Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. Proc Natl Acad Sci USA 107:12139–12144.
40. Zhang W, Dunkle JA, Cate JH (2009) Structures of the ribosome in intermediate states of ratcheting. Science 325:1014–1017.
41. Selmer M, Dunham CM, Murphy FVT, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313:1935–1942.
42. Meyer PA, Li S, Zhang M, Yamada K, Takagi Y, Hartzog GA, Fu J (2015) Structures and functions of the multiple KOW domains of transcription elongation factor Spt5. Mol Cell Biol 35:3354–3369.
43. Ivanov D, Kwak YT, Guo J, Gaynor RB (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. Mol Cell Biol 20:2970–2983.
44. Yamaguchi Y, Wada T, Watanabe D, Takagi T, Hasegawa J, Handa H (1999) Structure and function of the human transcription elongation factor DSIF. J Biol Chem 274:8085–8092.
45. Bendak K, Loughlin FE, Cheung V, O'Connell MR, Crossley M, Mackay JP (2012) A rapid method for assessing the RNA-binding potential of a protein. Nucleic Acids Res 40:e105.
46. Kwan AH, Czolij R, Mackay JP, Crossley M (2003) Pentaprobe: a comprehensive sequence for the one-step detection of DNA-binding activities. Nucleic Acids Res 31:e124.
47. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510.
48. Djordjevic M (2007) SELEX experiments: new prospects, applications and data analysis in inferring regulatory pathways. Biomol Eng 24:179–189.
49. Yang X, Li H, Huang Y, Liu S (2013) The dataset for protein-RNA binding affinity. Protein Sci 22:1808–1811.
50. Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, Cramer P (2010) Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol 17:1272–1278.
51. Liu CR, Chang CR, Chern Y, Wang TH, Hsieh WC, Shen WC, Chang CY, Chu IC, Deng N, Cohen SN, Cheng T-H (2012) Spt4 is selectively required for transcription of extended trinucleotide repeats. Cell 148: 690–701.
52. Li W, Giles C, Li S (2014) Insights into how Spt5 functions in transcription elongation and repressing transcription coupled DNA repair. Nucleic Acids Res 42: 7069–7083.
53. Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. Cell 98:835–845.
54. Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL, Allain FH T (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science 309:2054–2057.
55. Belogurov GA, Mooney RA, Svetlov V, Landick R, Artsimovitch I (2009) Functional specialization of transcription elongation factors. Embo J 28:112–122.
56. Belogurov GA, Sevostyanova A, Svetlov V, Artsimovitch I (2010) Functional regions of the N-terminal domain of the antiterminator RfaH. Mol Microbiol 76:286–301.
57. Artsimovitch I, Landick R (2002) The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell 109:193–203.
58. Yakhnin AV, Yakhnin H, Babitzke P (2008) Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. Proc Natl Acad Sci USA 105:16131–16136.
59. Yakhnin AV, Babitzke P (2014) NusG/Spt5: are there common functions of this ubiquitous transcription elongation factor? Curr Opin Microbiol 18:68–71.
60. Ujvari A, Luse DS (2006) RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. Nat Struct Mol Biol 13:49–54.
61. Andrecka J, Lewis R, Bruckner F, Lehmann E, Cramer P, Michaelis J (2008) Single-molecule tracking of mRNA exiting from RNA polymerase II. Proc Natl Acad Sci USA 105:135–140.
62. Huertas P, Aguilera A (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell 12:711–721.
63. Rondon AG, Garcia-Rubio M, Gonzalez-Barrera S, Aguilera A (2003) Molecular evidence for a positive role of Spt4 in transcription elongation. Embo J 22: 612–620.
64. Wahba L, Costantino L, Tan FJ, Zimmer A, Koshland D (2016) S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes Dev 30:1327–1338.
65. Blythe A, Gunasekara S, Walshe J, Mackay JP, Hartzog GA, Vrielink A (2014) Ubiquitin fusion constructs allow the expression and purification of multi-KOW domain complexes of the Saccharomyces cerevisiae transcription elongation factor Spt4/5. Protein Expr Purif 100:54–60.
66. Zakov S, Goldberg Y, Elhadad M, Ziv-Ukelson M (2011) Rich parameterization improves RNA structure prediction. J Comput Biol 18:1525–1542.
67. Shroedinger, LLC. The PyMOL Molecular Graphics System, Version 1.5.0.4.
68. Barnes CO, Calero M, Malik I, Graham BW, Spahr H, Lin G, Cohen AE, Brown IS, Zhang Q, Pullara F, Trakselis MA, Kaplan CD, Calero G (2015) Crystal structure of a transcribing RNA polymerase II complex reveals a complete transcription bubble. Mol Cell 59: 258–269.