STRUCTURAL AND BIOCHEMICAL STUDIES OF THE TRANSCRIPTION FACTORS SPT6 AND SPN1

by

Seth M. McDonald

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

The University of Utah

May 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Seth M. McDonald**

has been approved by the following supervisory committee members:

Christopher P. Hill	, Chair	**11/19/2012** Date Approved
Tim Formosa	, Member	**11/19/2012** Date Approved
Bradley R. Cairns	, Member	**11/19/2012** Date Approved
Dana Carroll	, Member	**11/19/2012** Date Approved
Michael S. Kay	, Member	**11/19/2012** Date Approved

and by **Wesley I. Sundquist**, Chair of the Department of **Biochemistry**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

Eukaryotic transcription and mRNA processing depend upon the coordinated interactions of many proteins, including Spn1 and Spt6, which are conserved across eukaryotes, are essential for viability, and associate with each other in some of their biologically important contexts. Spt6 functions at several important regulatory steps in transcription, including nucleosome reassembly, transcription elongation, and mRNA processing and export. As a histone chaperone, Spt6 is important for reassembly of nucleosomes in the wake of elongating RNA polymerase II, a process that is required to regulate transcription initiation and prevent inappropriate transcription from repressed promoters as well as cryptic intragenic transcription start sites. In conjunction with Spt6, Spn1 coordinates the recruitment of mRNA processing and export factors, such as Yra1 and the exosome, thereby enabling biogenesis of mature and export competent mRNA molecules. The functional relevance of Spn1 and Spt6 in chromatin organization and mRNA maturation is well established, although mechanistic details of how these processes are performed are poorly understood. In order to enhance our understanding of the molecular details of Spn1 and Spt6 functions, this thesis has focused on structural, biochemical and functional studies of Spn1 and Spt6 from *Saccharomyces cerevisiae*.

The work presented in this dissertation establishes the structures of the entire ordered region of the Spt6 protein, the ordered core of the Spn1 protein, and the Spn1 core in complex with the binding determinant of Spt6. Additionally, we demonstrate the capacity of Spt6 to interact with factors that very likely influence Spt6 function, including histones and nucleosomes. The structures and the functional data described in this dissertation have enhanced our understanding of how Spt6 binds and chaperones histones, and describes a novel role for Spn1 in regulating the histone chaperone activity of Spt6. The Spt6 and Spn1 structures and the biochemical assays developed in this work will aid in future functional and mechanistic studies that will aim to develop a complete molecular and mechanistic model for each Spt6 and Spn1 function.

TABLE OF CONTENTS

ABSTRACT…….. iii

ACKNOWLEDGMENTS …… vii

Chapter

1. INTRODUCTION……... 1

Transcription and mRNA Biogenesis…… 1
Transcription in the Context of Chromatin…… 8
The Conserved Spn1 Protein…….. 20
Spn1 and Spt6 in mRNA Processing and Export…… 21
Spt6 is a Transcription Elongation Factor…… 22
Spt6 is a Histone Chaperone…….. 23
Goals of This Dissertation……... 25
Outline of Chapters…….. 28
References…… 31

2. CRYSTAL STRUCTURES OF THE *S. CEREVISIAE* SPT6 CORE AND TANDEM SH2 DOMAIN……. 41

Abstract…….. 42
Introduction…… 43
Results and Discussion……..…… 44
Materials and Methods……..……. 55
Acknowledgements……. 56
References…… 57
Supplementary Data…… 59

3. STRUCTURE AND BIOLOGICAL IMPORTANCE OF THE SPN1-SPT6 INTERACTION, AND ITS REGULATORY ROLE IN NUCLEOSOME BINDING..... 64

Summary..... 65
Introduction..... 65
Results..... 66
Discussion..... 70
Experimental Procedures..... 73
Acknowledgments..... 73
References..... 73
Supplmental Information..... 76

4. MECHANISM OF SPT6 HISTONE CHAPERONE ACTIVITY AND ITS REGULATION BY SPN1..... 92

Summary..... 92
Introduction..... 93
Results and Discussion..... 98
Experimental Procedures..... 124
References..... 128

5. CONCLUSIONS AND ONGOING RESEARCH..... 134

Summary..... 134
Functional Implications for Spt6 and Spn1..... 135
Biochemical Analysis of Spt6..... 139
Conclusions..... 144
References..... 145

ACKNOWLEDGMENTS

I would like to acknowledge all of my co-workers in the Hill lab for enabling a fun, relaxing, unassuming, and scientifically simulating work environment. Devin Close and Matt Sdano were critical to making the work presented in this dissertation possible. I would also like to thank Frank Whitby and Heidi Schubert for their patience, leadership and guidance in all practical aspects of crystallography. A big thank you to the members of my thesis committee: Tim Formosa, Brad Cairns, Dana Carroll, and Michael Kay, who have provided invaluable guidance throughout my graduate career. I would like extend a special thank you to the members of our 'nucleosome subroup', especially Dave Kemble, Matt Sdano, Devin Close, and Tim Formosa for always being open to discussion and providing important insights into my work. In particular, I would like to thank my advisor Chris Hill for being an exceptional mentor. Chris has a passion for life and science unrivaled by any person I have met. I would also like to thank my family and friends who have provided unending support for me throughout my life and especially during my graduate career. Most of all, I would like to thank my beautiful wife Ashley. Without her unconditional love and support, I would have crumbled under the stress of graduate school long ago. My incredible children, Myron and Adeline, are not only examples of Ashley's

tireless love and support, but are also beacons of happiness that serve to bring me back to reality after a hard day's work. I would also like to acknowledge financial support from the University of Utah Graduate Research Fellowship program, and to Elsevier Science for permission to use published articles and figures within this dissertation.

CHAPTER 1

INTRODUCTION

Transcription and Biogenesis of mRNA

The process of transcription is the synthesis of RNA molecules from a DNA template. RNA molecules play numerous roles in many cellular processes, including peptidyl-transferase activity during protein synthesis (Ramakrishnan, 2002), splicing of messenger RNA (mRNA) molecules (Semlow and Staley, 2012), and the inhibition of transcription and translation (Mello and Conte, 2004; Verdel et al., 2004). Fascinatingly, the emergence of enzymatically active RNA molecules, such as self-replicating ribozymes, may have been a critical early step in the development of life on earth (Johnston et al., 2001). RNA polymerase (RNAP) is the multi-subunit enzyme that catalyzes the polymerization of ribonucleotides into RNA chains that are complimentary to the DNA template (Hurwitz, 2005). RNAP enzymes are essential to life and are found in every organism, including some viruses (Ahlquist, 2002; Korkhin et al., 2009). The RNAP from bacteria is structurally and mechanistically related to each of the three types of RNAPs in eukaryotes, RNA polymerase I, II and III (Ebright, 2000). The most highly studied eukaryotic RNAP is RNA polymerase II (RNAPII), which

is responsible for transcribing protein coding genes, the precursors to mRNA, and most microRNAs (Lee et al., 2004). The transcriptional process in modern organisms, from bacteria to humans, employs strikingly similar topological features (Kornberg, 2007; Zakrzewska and Lavery, 2012), and consists of three major steps, initiation, elongation, and termination (Figure 1-1A and B). The following discussion will focus primarily on eukaryotic transcription.

From initiation to elongation

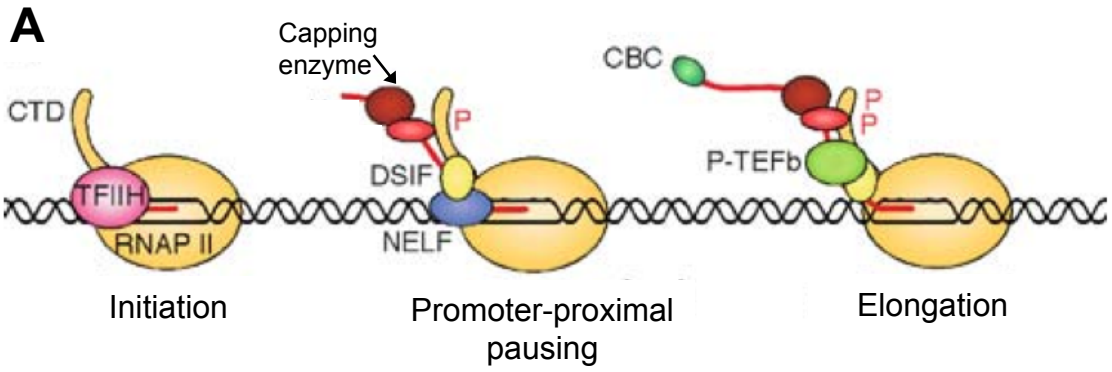
Initiation begins with assembly of the pre-initiation complex and recognition of core promoter elements within genomic DNA. In bacteria, RNAP is aided in promoter recognition by sigma factors (Feklistov and Darst, 2011), whereas eukaryotic RNAPII utilizes a set of general transcription factors unrelated to their bacterial counterparts (Figure 1-1A) (Liu et al., 2011). After promoter recognition, the DNA is partially unwound to form an open complex that allows the DNA to enter the active site of the enzyme (Goodrich and Tjian, 1994; Kornberg, 2007). Synthesis of the first 8-9 ribonucleotides produces a DNA-RNA heteroduplex, which stabilizes the open complex, permits promoter escape, and the transition to the elongation step (Kornberg, 2007). Although the precise determinants of the transition to an elongation competent complex are not known, most evidence suggests that the length of the nascent transcript is crucial for making contacts to RNAPII that allow collapse of the transcription bubble and escape from the promoter region (Luna et al., 2008).

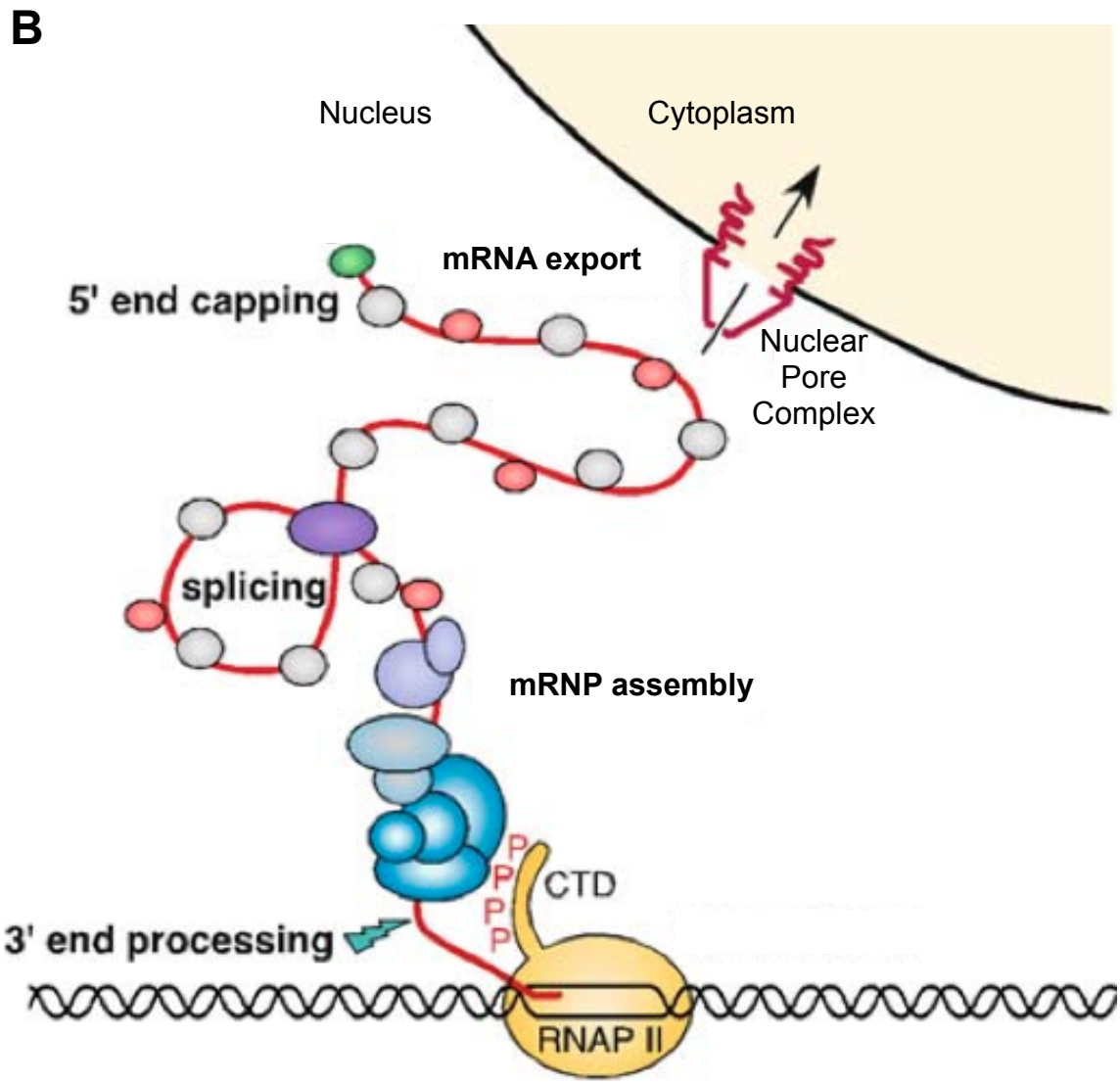
Figure 1-1. Overview of Transcription and mRNA Maturation

Adapted from (Luna et al., 2008)

A) Simplified diagram of the first steps of transcription from initiation to productive elongation. Initiation in eukaryotes is aided by general transcription factors such as TFIIH. Promoter-proximal pausing induced by the Spt4/Spt5 complex (DSIF) and NELF coincides with 5'-end capping of the nascent mRNA transcript. Release from pausing to productive elongation requires the kinase activity of PTEF-b

B) Schematic of the co-transcriptional processing of nascent transcripts required to produce mature, export competent mRNA molecules. Capping, splicing and 3'-end processing are regulated co-transcriptionally and export from the nucleus requires adaptor proteins that bind the mRNA and interact with factors at nuclear pore complex.


A
Capping enzyme
CBC
CTD
P
P
P
DSIF
P-TEFb
TFIIH
RNAP II
NELF
Initiation
Promoter-proximal pausing
Elongation


B
Nucleus
Cytoplasm
mRNA export
5' end capping
Nuclear Pore Complex
splicing
mRNP assembly
P
P
P
P
CTD
3' end processing
RNAP II

Promoter escape in eukaryotes coincides with the start of the phosphorylation cycle of the C-terminal domain of the largest subunit of RNAPII (RNAPII-CTD). The RNAPII-CTD consists of heptapeptide repeats of the sequence YSPTSPS, where each of the serines, as well as the tyrosine, can be phosphorylated (Mayer et al., 2012). This domain serves as a scaffold for recruiting a number of factors whose binding depends on the phosphorylation pattern of the RNAPII-CTD, which changes throughout the transcription cycle and coordinates different events in the maturation of the mRNA molecule (Buratowski, 2003; Luna et al., 2008). Prior to becoming a wholly productive elongation complex, continued adjustments are made to the complex that leads to transcriptional pausing near the promoter region. Factors such as the Spt4/Spt5 complex and negative elongation factor (NELF) bind the transcription complex and inhibit transcription (Saunders et al., 2006). Release from this inhibition into productive elongation requires phosphorylation events mediated by the positive transcription elongation factor complex-b (PTEF-b), which phosphorylates the Spt4/Spt5 complex, NELF, and the RNAPII-CTD (Figure 1-1A) (Bres et al., 2008). The Spt4/Spt5 complex remains associated to the elongation complex in accordance with its positive role in transcription elongation (Saunders et al., 2006).

Termination

Transcription termination occurs after passage of the polyadenylation (polyA) signal by the elongation complex. Nascent mRNA molecules are typically cleaved 20-30 nucleotides downstream of the polyA signal by the concerted effort of several 3'-end processing factors (Luna et al., 2008). After cleavage, polyA polymerase adds the polyA tail, and the tail is bound by polyA binding proteins that protect the transcript from degradation (Figure 1-1B) (Luna et al., 2008). Two mechanistic and non-mutually exclusive models have been proposed to explain termination (Buratowski, 2005; Rosonina et al., 2006). The first is an 'allosteric' model in which conformational changes within the elongation complex following the passage of the polyA signal cause a decrease in processivity and subsequent disassembly. In support of this model, a single cleavage factor has been shown to efficiently disassemble paused elongation complexes (Zhang et al., 2005; Zhang and Gilmour, 2006). The second model, the 'torpedo' termination model, proposes that the exonucleolytic degradation of the uncapped nascent transcript that remains after cleavage of the pre-mRNA at the polyA signal somehow signals the elongation complex to terminate. Support for this model comes from observations that the 5'-3' exonuclease XRN2 in humans (Rat1 in yeast) is required for efficient termination (Kim et al., 2004; Luo et al., 2006; West et al., 2004).

mRNA processing and quality control

It is important to note that the processing required for production of mature, stable, and export competent mRNA molecules occurs co-transcriptionally (Figure 1-1B) and begins once the nascent transcript emerges from RNAPII (Luna et al., 2008). Capping of the nascent transcript, the methyl guanylation of the 5'-end, happens coincidentally with promoter-proximal pausing (Moteki and Price, 2002; Rasmussen and Lis, 1993); likely serving as a checkpoint to ensure that only properly capped transcripts are extended. The cap-binding protein complex then binds the monomethyl cap and is required for further maturation events including splicing and export from the nucleus (Izaurralde et al., 1994). Some of the required 3'-end processing factors associate with promoter regions, implying that initiation and termination are connected and may impact each other's efficiency (Krishnamurthy et al., 2004; Mayer et al., 2012). The splicing, or removal of introns from the nascent transcript, can occur during or after transcription, and ensures that the correct protein code is contained within the transcript (Semlow and Staley, 2012). RNA-binding proteins, such as Yra1 (REF1/Aly in humans), are recruited to the transcription elongation complex and are required for mRNA export to the cytoplasm (Johnson et al., 2009; Strasser and Hurt, 2000; Yoh et al., 2007). Improperly processed transcripts are retained in the nucleus and degraded by the nuclear exosome complex (Saguez et al., 2005), while factors at nuclear pore

complexes also contribute to the transcriptional quality control mechanism (Galy et al., 2004).

Transcription in the Context of Chromatin

The cells of all organisms contain DNA molecules that constitute the genome and provide a blueprint or code for all aspects of cell survival. The fundamental difference between bacteria and eukaryotes is that bacterial cells are not partitioned into intracellular compartments or organelles, their biology is less complicated, and as such, their genomes are smaller and not sequestered in a membrane bound nucleus. The DNA molecules in human cells are extremely long, on the order of two meters if all molecules were stitched together, and eukaryotic organisms have therefore evolved an effective packaging system termed chromatin that enables the compaction and storage of the genome inside the nucleus of each cell (Li and Reinberg, 2011). It is not simply enough to package efficiently, however, because much of the DNA needs to be accessed on a regular basis, and chromatin therefore needs to be dynamic. For example, every time a cell divides the entire genome must be replicated; a process that requires access to the DNA. Similarly, the synthesis of RNA and protein molecules that drive cellular processes requires accessing, processing, and repackaging relevant segments of the genome. Thus, packaging, accessing, and repackaging of genomic DNA is intimately linked to biological function and is highly regulated. This intimate link can have profound consequences if there is

even a slight error in any of the regulatory steps of genomic access. Errors in these processes can vary in severity from simple replication or transcription mistakes, which can be corrected by inherent cellular proofreading and repair pathways, to catastrophic defects leading to cell death or uncontrolled cellular proliferation as in tumor formation and cancers (Cairns, 2001). The complexity of genomic regulation and our continuing advancement in understanding these processes remains an intriguing and essential task in biological and medical research.

Eukaryotic genome organization

Chromatin structure

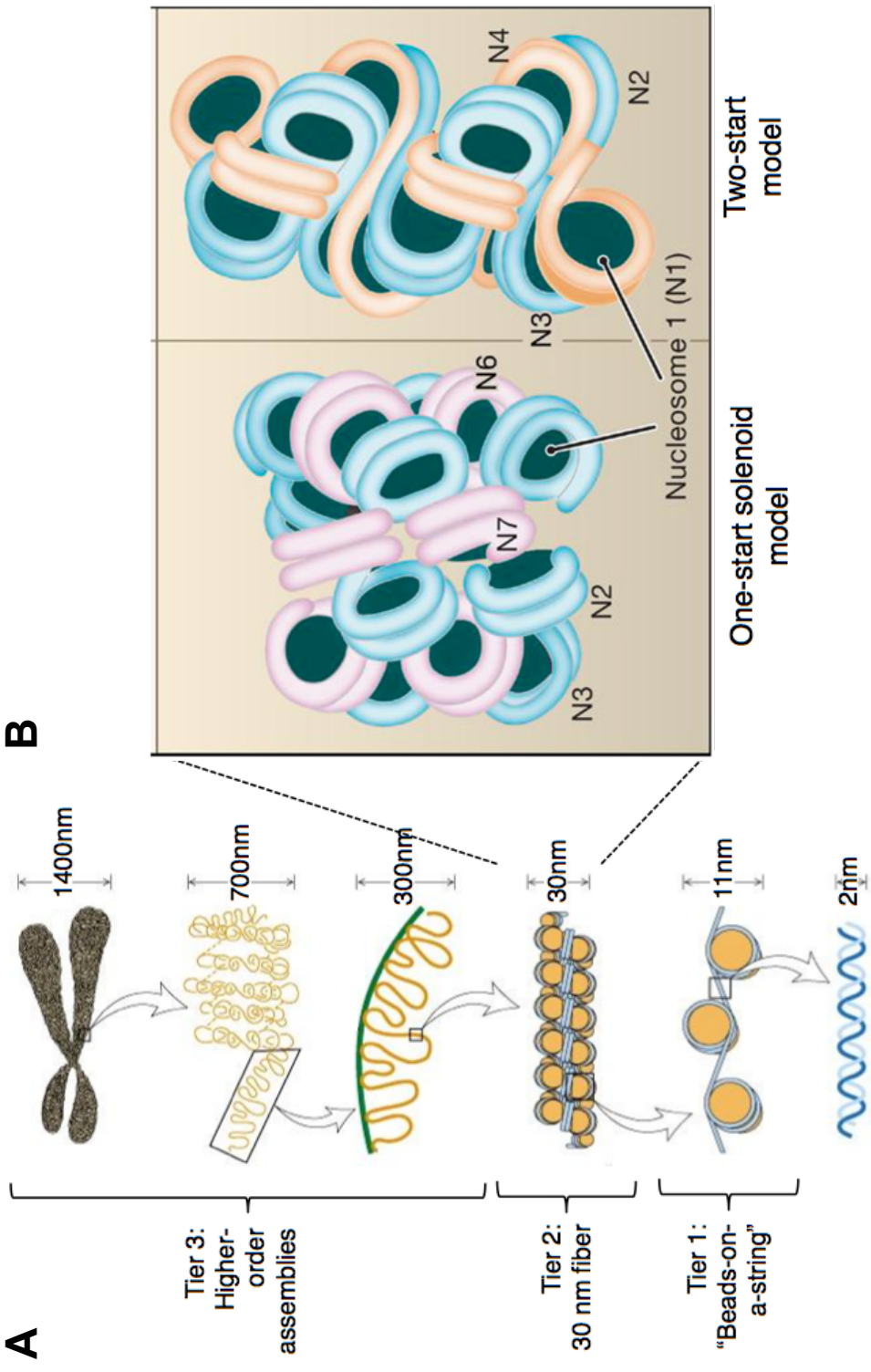
The overall structure and regulation of chromatin is nearly identical from single-celled yeasts to multicellular organisms and humans (Li and Reinberg, 2011). Chromatin structure ranges from highly condensed mitotic chromosomes to single nucleoprotein complexes called nucleosomes (Figure 1-2). The nucleosome is the fundamental repeating unit of chromatin and consists of approximately 150 bp of DNA wrapped around an octameric protein core made of two copies of each of four core histone proteins, H2A, H2B, H3, and H4 (Luger et al., 1997). Beyond the packaging into individual nucleosomes, higher order chromatin compaction is classified by a hierarchical scheme (Li and Reinberg, 2011). Repeating units of single nucleosomes connected by regions of free DNA is the most basic level of chromatin organization and resembles ‘beads on a

Figure 1-2. Overview of the Hierarchy of Chromatin Structure

Adapted from (Lodish et al., 2000; and Li and Reinberg, 2011)

A) Diagram of the three tiered hierarchical classification scheme for chromatin compaction states.

B) Competing models for the arrangement of nucleosomes within the 30 nm fiber highlight the varied degree of compaction, and the likelihood of multiple conformational states.


A
B
1400nm
700nm
300nm
30nm
11nm
2nm
Tier 3: Higher-order assemblies
Tier 2: 30 nm fiber
Tier 1: "Beads-on-a-string"
N4
N2
N3
N6
N7
N2
N3
Nucleosome 1 (N1)
Two-start model
One-start solenoid model

string'. Further compaction is loosely defined by internucleosomal contacts likely mediated through the N-terminal tails of the histone proteins resulting in a fiber approximately 30 nm in diameter. Precise arrangement of nucleosomes within the fiber remains unknown despite continuing efforts (Li and Reinberg, 2011), and is likely to be dynamic with a multitude of conformations. The nature of the interactions required to further compact the fibers are also poorly understood, and seem to vary with tissue type and cell cycle phase.

Chromatin can be separated into two distinct types in nondividing cells called heterochromatin and euchromatin (Kwon and Workman, 2011). Heterochromatin is condensed, tightly packed, and is generally associated with silenced portions of the genome such as centromeres, telomeres, and genes whose protein products are not required for specific tissue or cell types. Euchromatin is lightly packed and is where the majority of gene transcription takes place. The efficacy of chromatin organization extends beyond a means for genome compaction, and is intimately linked to nearly all nuclear processes including transcription, DNA replication and repair, and kinetochore and centromere formation. Although chromatin regulation of nuclear processes is ubiquitous, the role of nucleosomes in transcription has led to a greater understanding of the explicit contributions of chromatin to cellular function.

When the discovery was made that eukaryotic DNA was packaged into nucleosomes it was clear that higher-ordered chromatin assemblies would be a significant impediment to transcription (Kornberg, 1974; Kornberg and Lorch,

1992). Because RNAPII requires a relatively free DNA template to synthesize mRNA molecules, the presence of nucleosomes throughout the length of a given gene would severely encumber mRNA and hence protein synthesis (Saunders et al., 2006). An enormous amount of research in the past 15 years has shown that histones are not simply inert obstacles, but specifically contribute to the regulation of gene transcription. Histones can regulate the specificity of gene transcription through covalent modifications that function to recruit activating or repressive protein factors or prevent repressive higher-order chromatin assemblies (Li and Reinberg, 2011). Histones also serve as steric blocks to the underlying genetic information, with displacement of intimate histone-DNA contacts being achieved through ATP-dependent and independent processes (Tyler, 2002). Consequently, the assembly and architecture of the nucleosome is central to the regulation of gene expression, as virtually all regulatory processes involve chemical or conformational perturbations to nucleosomal components.

Nuclesome structure

Timothy Richmond and colleagues published the first high-resolution structure of the nucleosome core particle in 1997 (Luger et al., 1997). Prior to this, the disc-like shape and the general composition of the nucleosome were already well established (Kornberg, 1977; Thomas and Kornberg, 1975). The four core histone proteins each adopt a fold with a helix-turn-helix-turn-helix motif, termed the histone fold, as well as partially disordered N-terminal ‘tails’

(Figure 1-3). These 'tails' are highly enriched in lysine, arginine, and serine residues that are capable of accepting covalent modifications that aid in the regulation of gene expression (Henikoff, 2008; Jenuwein and Allis, 2001; Kouzarides, 2007). The histone fold is optimal for dimerization and protein-protein interactions required for formation of the core histone octamer, around which approximately 150 bp of DNA is wrapped (Sullivan and Landsman, 2003). Each histone protein is relatively small, ranging from 10-13 kDa, and highly basic on the accessible surface where direct contacts are made with the negatively charged phosphodiester backbone of the DNA. Within the octameric core the histones are packaged as a heterotetramer of H3 and H4, and two H2A-H2B dimers. In vitro, isolated histone H3-H4 tetramers and H2A-H2B dimers can be reconstituted (Luger et al., 1999), and therefore it is believed that free histones in the cell exist in these oligomeric states. Recent work has indicated that H3-H4 may also exist as a heterodimeric complex under physiological conditions (Donham et al., 2011), although this topic remains up for debate.

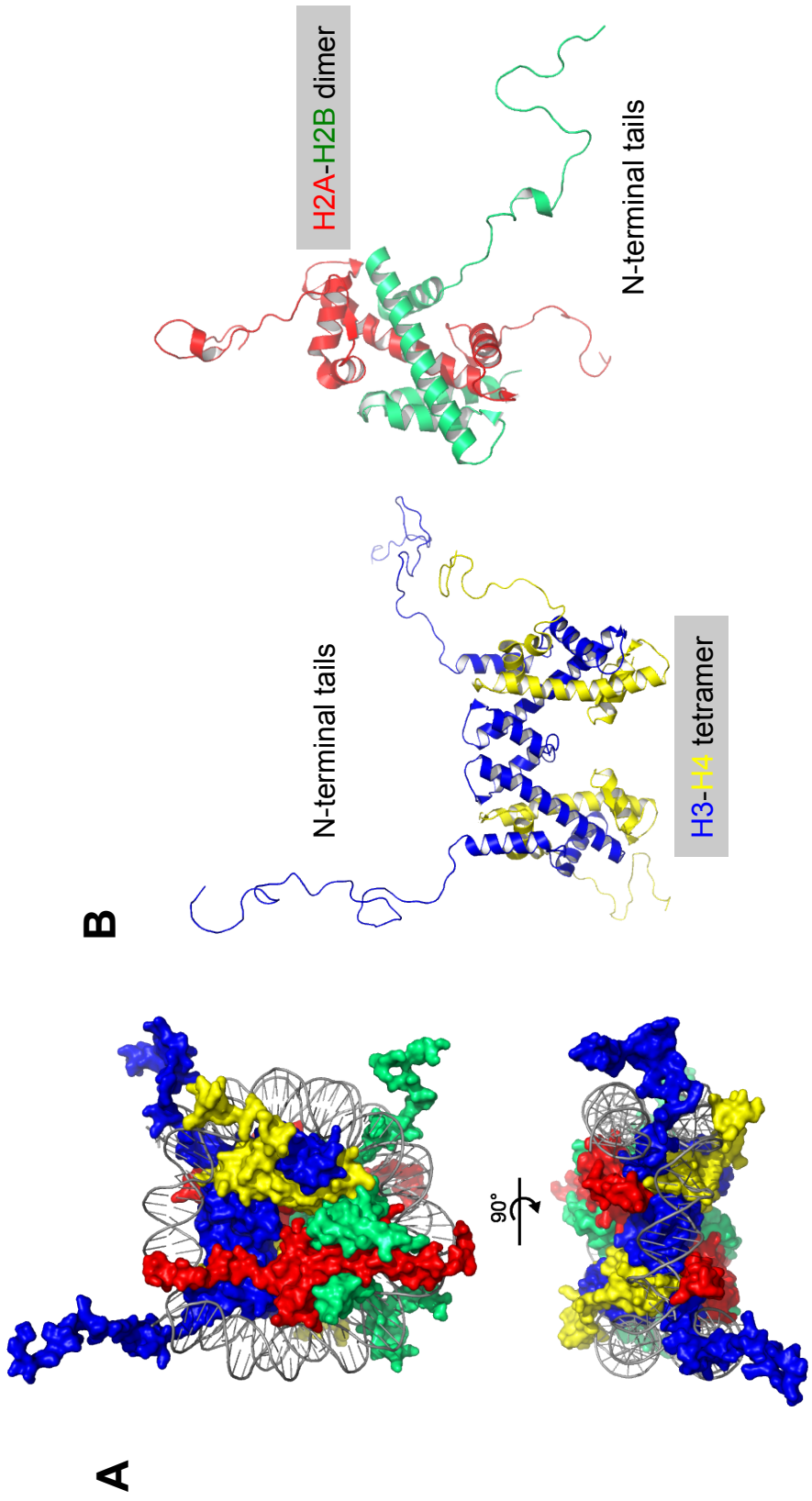
The N-terminal 'tails' of the histone proteins do not appear to make crucial architectural contributions to the nucleosome, but almost certainly contribute to higher-order chromatin assemblies (Li and Reinberg, 2011). This function of the 'tails' is highly dependent on the chemical modifications of the histones within a single nucleosome. Our understanding of the regulation and nature of chemical modifications of histone 'tails' is ever increasing, and the composition of

Figure 1-3. Nucleosome Composition and Structure

Figure generated with PyMol (Delano, 2002) using model PDB ID: 1kx5 (Davey et al., 2002)

A) Orthogonal views of the nucleosome core particle. Histones are depicted in a surface representation and colored as follows: H2A, red; H2B, green; H3, blue; H4, yellow. DNA is grey and shown as a cartoon representation

B) Models of the histone H3-H4 tetramers and H2A-H2B dimers, colored as in (A). N-terminal tails are highlighted.


A
90°
B
H2A-H2B dimer
N-terminal tails
N-terminal tails
H3-H4 tetramer

modifications on the nucleosomes throughout a gene yields a code that can significantly influence the transcriptional status.

The histone code

The residues within the N-terminal histone 'tails' can be covalently modified by chemical moieties such as phosphoryl, methyl, and acetyl groups, and can even be ubiquitylated (Jenuwein and Allis, 2001; Kouzarides, 2007; Strahl and Allis, 2000). Covalent 'tail' modifications can influence transcription and access to the underlying DNA in one of two ways, the first being the contributions made to higher-order chromatin assemblies (Li and Reinberg, 2011). The second way the chemical modification of histone tails influences transcription and DNA access is through the recruitment of various factors involved in transcription, chromatin modification, and chromatin remodeling (Taverna et al., 2007). These factors utilize domains that recognize specific modifications and thus can interpret the 'histone code' or the combinatorial manifestation of modifications present throughout the length of a given gene. In the simplest explanation, acetylation is typically a transcriptional activation signal, while the absence of acetyl marks is generally associated with repressed transcription. Overall, the common purpose of the 'histone code' is to establish the chemical signatures required to coordinate the recruitment of proteins and protein complexes that influence DNA accessibility. As a result, the 'histone code' directly contributes to chromatin-mediated transcriptional regulation.

Recent work highlights the complexity of the histone code by challenging the notion that modifications are established in an identical, or symmetric fashion on both copies of each of the four core histones (Voigt et al., 2012). This work established that asymmetric modifications exist on mononucleosomes within native chromatin, and further demonstrates that asymmetric modifications signal different biological outcomes than symmetric modifications.

Histone chaperones and chromatin remodelers

Nucleosomes throughout the genome exist in a state of structural flux between fully assembled and disassembled forms (Polo and Almouzni, 2006; Tyler, 2002). The balancing of this equilibrium can have dramatic consequences on transcription, and is mediated through the direct actions of two groups of proteins: Histone chaperones and chromatin remodelers.

Histone chaperones, such as Spt6, are thought to use inherent binding energy to displace histones from intimate contacts with the DNA and/or prevent nonproductive histone-DNA contacts to allow the proper reassembly of nucleosomes (Andrews et al., 2010; Winkler et al., 2011). Histone chaperones facilitate DNA access and maintain the integrity of chromatin throughout processes that perturb nucleosomal structure, such as transcription and DNA replication. The inherent similarities of these processes allow many histone chaperones to function during each process, where they can be active in both the disassembly of nucleosomes in front of the respective polymerase, and

reassembly after passage. The histone chaperones FACT and Asf1 have been shown to have physical and genetic interactions with components of the transcription and replication machinery (Biswas et al., 2005; Formosa, 2012; Formosa and Nittis, 1998; Mousson et al., 2007; Park and Luger, 2006; Rocha and Verreault, 2008; VanDemark et al., 2006), while other chaperones, such as Spt6, appear to be specific to the transcription process (Das et al., 2010; Hondele and Ladurner, 2011; Mello and Almouzni, 2001).

Chromatin remodelers utilize the energy of ATP binding and hydrolysis to drive a motor-like ratchet system that can translocate DNA with respect to nucleosomal histones, or even evict histones completely (Cairns, 2005). Remodelers are also responsible for appropriate spacing of nucleosomes throughout the genome, as well as facilitating the exchange of histone variants, and thus contribute to both the regulation of higher-order chromatin assemblies and to nucleosome composition and spacing, each of which inherently impacts transcriptional regulation (Clapier and Cairns, 2009).

Histone chaperones and chromatin remodelers utilize different mechanisms to exert their impact on chromatin integrity; however, they also cooperate in the precise regulatory pathways that ultimately result in transcriptional activation or repression (Tyler, 2002). Spt6 is a protein that has an integral role in gene expression and chromatin maintenance conferred by its functions as a histone chaperone, transcription elongation factor, and modulator

of mRNA processing. Spt6 binds the Spn1 protein (also called Iws1) and they function together in many of these processes.

The Conserved Spn1 Protein

The *SPN1* gene was originally identified as a key regulator of transcription from genes that are regulated postrecruitment of RNAPII, and is essential to yeast viability (Fischbeck et al., 2002). *SPN1* encodes a 410 residue, 46 kDa protein with a central ordered core region (residues 148-295) that is flanked on both sides by regions predicted to be disordered (Ward et al., 2004). Interestingly, expression of the ordered core region complements a deletion of *SPN1*, indicating that the core provides the major function(s) of Spn1 in vivo (Fischbeck et al., 2002). The ordered core of Spn1 is also highly conserved in sequence, and hence presumably structure, from yeast to human (Pujari et al., 2010). Spn1 was also identified as a protein that interacts with Spt6, and has been reported to bind with Spt6 in some but not all of Spt6's functional states (Krogan et al., 2002; Lindstrom et al., 2003; Yoh et al., 2007; Zhang et al., 2008). For example, Spt6 can be coimmunopurified with three distinct Spt4/5-RNAPII complexes, whereas Spn1 is found in only two of these complexes (Lindstrom et al., 2003). The *CYC1* gene of *S. cerevisiae* provides an example of how the Spn1-Spt6 interaction contributes to postrecruitment regulation (Zhang et al., 2008). RNAPII is constitutively bound to the *CYC1* promoter, but is kept from elongating because it interacts with Spn1, which in turn inhibits the Swi/Snf

nucleosome remodeling complex from promoting transcription. During activation, Spt6 binds to Spn1, and repression of Swi/Snf recruitment is relieved.

Spn1 is also needed to achieve normal recruitment of the histone methyltransferase HYPB/Setd2 (Yoh et al., 2008) and the elongation factor TFIIS (Ling et al., 2006; Zhang et al., 2008) to RNAPII complexes traversing active genes. HYPB/ Setd2 methylates histone H3K36, which in turn recruits Rpd3-type histone deacetylases to restore chromatin to the repressive hypoacetylated state and block inappropriate transcription (Yoh et al., 2008). In contrast to their antagonistic relationship in activating postrecruitment initiation, Spn1 and Spt6 each contribute toward restoration of repressive chromatin. Human Spn1/IWS1 also binds the protein arginine methyltransferase PRMT5, which methylates the elongation factor Spt5 and thereby regulates its interaction with RNAPII (Liu et al., 2007). Spn1 can additionally function through interactions with pathway-specific regulatory factors, such as the Arabidopsis steroid hormone responsive transcription factor BES1, which recruits Spn1 to the promoter and transcribed regions of activated genes (Li et al., 2010). Spn1 therefore contributes in several ways to the appropriate functioning of RNAPII.

Spn1 and Spt6 in mRNA Processing and Export

Spt6 and Spn1 collaborate to promote the appropriate processing and export of mRNA molecules. Spt6 is required for proper 3'-end formation by preventing premature 3' processing at upstream polyA signals (Bucheli and

Buratowski, 2005; Kaplan et al., 2005). Additionally, Spt6 binding to the Ser2-phosphorylated RNAPII-CTD (Diebold et al., 2010; Sun et al., 2010; Yoh et al., 2007; Yoh et al., 2008) enhances recruitment of RNA processing and export factors, and *Drosophila* Spt6 copurifies with the RNA processing exosome complex (Andrulis et al., 2002). Both *SPN1* and *SPT6* have also been implicated in mRNA splicing in *S. cerevisiae* (Burckin et al., 2005), and binding of human Spn1/IWS1 to the mRNA export factor REF1/Aly is important for recruitment of REF1/Aly to the body of the c-Myc gene during transcription (Yoh et al., 2007). Further, human Spn1/IWS1 and Spt6 are required for the proper splicing of HIV-1 transcripts, and depletion of either results in the global retention of polyadenylated mRNA molecules in the nucleus (Yoh et al., 2007). Therefore, Spn1 and Spt6 play a pivotal role in modulating the processing and export of mRNA transcripts.

Spt6 is a Transcription Elongation Factor

The essential yeast protein Spt6 was originally identified in a screen for factors that alter normal transcription initiation (Clark-Adams and Winston, 1987; Denis, 1984; Neigeborn et al., 1987; Simchen et al., 1984). The role of Spt6 in transcription initiation has been ascribed to the ability of Spt6 to chaperone histones to promote reassembly of nucleosomes in the wake of RNAPII, thereby reestablishing the default repressive chromatin state that prevents inappropriate initiation of transcription (Adkins and Tyler, 2006; Bortvin and Winston, 1996;

Cheung et al., 2008; Kaplan et al., 2003). While of great importance, maintaining repressive chromatin appears to be just one of Spt6's roles. For example, Spt6 also enhances the elongation rate of RNAPII (Hartzog et al., 1998; Kaplan et al., 2005; Kaplan et al., 2000; Lindstrom et al., 2003) on nucleosome-free DNA templates in vitro (Endoh et al., 2004; Hartzog et al., 1998; Keegan et al., 2002; Yoh et al., 2007), as well as on chromatin templates in vivo (Ardehali et al., 2009). Spt6 is also classified as a transcription elongation factor because it colocalizes with RNAPII in *Drosophila* at sites of active transcription (Andrulis et al., 2000; Kaplan et al., 2000; Zobeck et al., 2010), and can be coimmunopurified with RNAPII throughout the length of an actively transcribed gene, indicative of a factor that travels along with the transcription elongation complex (Mayer et al., 2010). Together, these data indicate that Spt6 plays a number of mechanistically distinct roles during transcription.

Spt6 is a Histone Chaperone

The histone chaperone activity of Spt6 is necessary to maintain the integrity of chromatin to prevent spurious transcription from repressed promoters and cryptic intragenic transcription start sites (Adkins and Tyler, 2006; Cheung et al., 2008; Kaplan et al., 2003). Spt6 appears to exert this function through direct interactions with histones, with an apparent preference for the histone H3-H4 tetramer, and it promotes the deposition of histones and presumably the assembly of nucleosomes on plasmid DNA (Bortvin and Winston, 1996). The

PHO5 promoter provides an example of Spt6's role in nucleosome reassembly upon repression of transcription (Adkins and Tyler, 2006). Under high intracellular phosphate conditions, the transcriptional activator Pho4 is retained in the cytoplasm and transcription from *PHO5* is repressed. In low phosphate conditions, Pho4 is translocated to the nucleus where it binds the *PHO5* promoter and activates transcription. Nucleosome disassembly from the *PHO5* promoter, which is essential for transcriptional activation, is mediated by the histone chaperone Asf1 (Adkins et al., 2004; Adkins et al., 2007). Nucleosome reassembly at the *PHO5* promoter, which is required to repress transcription, is mediated by Spt6 (Adkins and Tyler, 2006). Surprisingly, *SPT6* mutants that are defective in nucleosome reassembly not only fail to repress *PHO5* transcription, but also allow transcriptional reactivation in the absence of the activator Pho4. Spt6 appears to perform this same function at other inducible genes, such as *PHO8*, *ADH2*, *ADY2*, and *SUC2*. Spt6 is also required for nucleosome reassembly throughout the open reading frame (ORF) of *PHO5*, and prevents transcription initiation from a cryptic transcription start site in the constitutively active *FLO8* gene (Cheung et al., 2008; Kaplan et al., 2003). Overall, these data indicate an extensive role for Spt6 in promoting nucleosome reassembly at promoters and within ORFs of inducible and constitutively active genes to prevent inappropriate transcription.

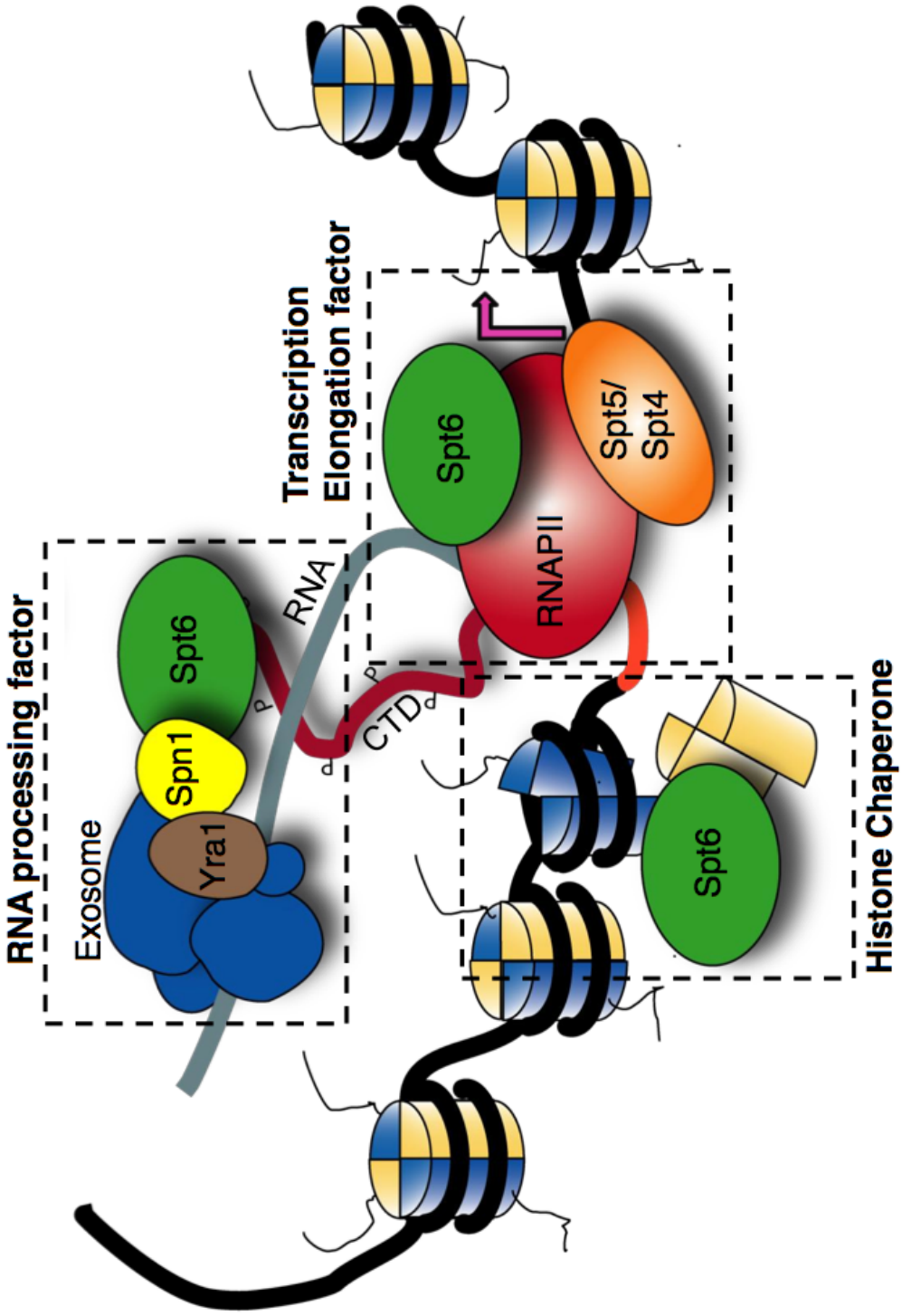
Goals of this Dissertation

Spt6 and Spn1 are transcription factors that bind one another and are each essential for yeast viability. Spt6 functions as a histone chaperone, a transcription elongation factor, and corroborates with Spn1 in mRNA processing (Figure 1-4). Although the importance of Spt6 and Spn1 in these processes is clear, little is known of the precise mechanistic details for any of their proposed functions. Therefore, the overall goal of this dissertation is to examine the molecular details of the functions of Spt6 and Spn1. The primary objectives were to structurally characterize Spt6, to describe the Spn1-Spt6 interaction structurally, biochemically, and genetically, and to perform a biochemical characterization of the Spt6 histone chaperone activity. The structures and functional data described in this dissertation have enhanced our understanding of how Spt6 binds and chaperones histones, and describes a novel role for Spn1 in regulating the histone chaperone activity of Spt6. The Spt6 and Spn1 structures and the biochemical assays developed in this work will aid in future functional and mechanistic studies aimed at developing a complete molecular and mechanistic model for each of the Spt6 and Spn1 functions.

Figure 1-4. Overview of the Functional Roles of Spt6 and Spn1

Figure made by Devin Close, used with permission.

The multifaceted Spt6 protein is a histone chaperone, transcription elongation factor, and collaborates with its binding partner Spn1 to recruit factors required for proper mRNA processing and export.


RNA processing factor
Exosome
Spt6
Spn1
Yra1
RNA
P
CTD
Transcription
Elongation factor
Spt6
Spt5/
Spt4
RNAPII
Histone Chaperone
Spt6

Outline of Chapters

Chapter 2: Crystal Structures of the *S. cerevisiae* Spt6 Core and C-Terminal Tandem SH2 Domain (D. Close, et al.)

The work described in Chapter 2 was originally published in the May 13, 2011 issue of Journal of Molecular Biology and is reprinted here in the published format. The focus of this work was to structurally characterize the yeast Spt6 protein, to investigate the Spt6 nucleic acid binding activity, and determine specificity for the C-terminal tandem SH2 domain in binding to phospho-isoforms of peptides mimicking the RNAPII-CTD.

The work in this chapter was initiated by Sean Johnson, a former postdoc in the lab, who solved the 3.3 Å resolution structure of the Spt6 core. Devin Close took that aspect of the project to completion and solved the higher resolution structures of the Spt6 core, as well as the structure of the C-terminal tSH2 domain. Devin was also critical in developing the nucleic acid and peptide binding assays used in the paper. I performed some of the cloning of the tSH2 domain mutants, and Matt Sdano and I helped to complete the peptide binding assays. One of my goals was to demonstrate nucleic acid binding specificity by mutational analysis of the $(HhH)_2$ domain, which ultimately proved unsuccessful. Overall, the structural and biochemical analysis of Spt6 presented in this chapter provide an invaluable foundation to further our understanding of Spt6 function, and helped to guide the structural and biochemical studies presented in Chapters 3, 4, and 5.

Chapter 3: Structure and Biological Importance of the Spn1-Spt6 Interaction, and its Regulatory Role in Nucleosome Binding (S.M. McDonald and D. Close, et al.)

The work described in Chapter 3 was originally published in the December 10, 2010 issue of Molecular Cell and is reprinted here in published format. The focus of this work was the structural, biochemical, and genetic characterization the Spn1-Spt6 interaction. We determined crystal structures of the ordered core region of Spn1 alone and in complex with the binding determinant of Spt6. Mutating interface residues greatly diminished binding in vitro and caused strong phenotypes in vivo, including a defect in maintaining repressive chromatin. Nucleosome binding by Spt6 was disrupted in the presence of Spn1, revealing a potential role for Spn1 in regulation of the histone chaperone activity of Spt6.

The work in this chapter was performed in close collaboration with Devin Close and Hua Xin. Devin and I worked together to solve the Spn1 structure and the structure of the Spn1-Spt6 complex, while I performed all of the biochemical experiments examining the Spn1-Spt6 interaction. Members in the lab of our collaborator Tim Formosa, especially Hua Xin, performed all of the genetic experiments and the nucleosome binding assays. Overall, our results point to a novel function for Spn1 in regulating the histone chaperone activity of Spt6, and provide a structural foundation for future work investigating the mechanism of Spn1 functions. This work also provided significant impetus for the work presented in Chapter 4.

Chapter 4: Mechanism of Histone Chaperone Activity of Spt6 and its Regulation by Spn1

Work in this chapter describes the biochemical analysis of the direct interactions between Spt6 and the two nucleosomal histone subcomplexes, H2A-H2B dimers and H3-H4 tetramers [$(H3\text{-}H4)_2$]. We further demonstrate that the Spn1-Spt6 and the Spt6-$(H3\text{-}H4)_2$ interactions are mutually exclusive, and that Spt6 competes with DNA for $(H3\text{-}H4)_2$ binding. These observations lead us to propose a direct-competition mechanism for the histone chaperone activity of Spt6, and strongly support a role for Spn1 as a regulatory switch in this process. We also determined that the mode of interaction between Spt6 and each of the two histone subcomplexes appears to be different, with only the interaction with H2A-H2B being highly sensitive to elevated salt concentrations. Further, Spt6 appears to bind the ordered core of the histone subcomplexes with the same 30 residue N-terminal segment that is required for Spn1 binding.

The work in this chapter was initiated by Devin Close, who aided in the development of the fluorescence anisotropy assay used to quantify binding to histones, while I performed all of the assays presented here. Overall, the data presented in this chapter strongly support our previous hypothesis that Spn1 plays a regulatory role in the histone chaperone activity of Spt6, and suggest a mechanism by which Spt6 chaperones histones by direct competition with DNA.

Chapter 5: Conclusions and Ongoing Research

Chapter 5 summarizes the structural and biochemical work presented in Chapters 2, 3, and 4, and discusses the functional implications for Spn1 and Spt6. In addition, this chapter describes preliminary structural studies on the Spt6-H2A-H2B complex, and the development of a small-scale proteomics approach to identifying new Spn1 and Spt6 binding partners. Preliminary structural studies on the Spt6-H2A-H2B complex, which are being undertaken in close collaboration with Matt Sdano, have yielded protein-containing crystals that diffract to low (~7 Å) resolution. Matt also developed the small-scale proteomics approach and initial results reveal interactions between the Spt6 C-terminal domain and several factors involved in transcription, including Yra1, an mRNA export adaptor protein, and Tom1, an E3-ubiquitin ligase of the HECT-class that is important for mRNA export and excess histone degradation. Long-term goals include structural, biochemical, and genetic characterization of confirmed direct binding partners. Overall, the work in this chapter, combined with the structural and functional data presented in the rest of the dissertation, provides a foundation for future work aimed at assessing molecular details of the proposed Spn1 and Spt6 functions.

References

Adkins, M.W., Howar, S.R., and Tyler, J.K. (2004). Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. Mol Cell *14*, 657-666.

Adkins, M.W., and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. Mol Cell *21*, 405-416.

Adkins, M.W., Williams, S.K., Linger, J., and Tyler, J.K. (2007). Chromatin disassembly from the PHO5 promoter is essential for the recruitment of the general transcription machinery and coactivators. Mol Cell Biol *27*, 6372-6382.

Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses, and RNA silencing. Science *296*, 1270-1273.

Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A., and Luger, K. (2010). The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. Mol Cell *37*, 834-842.

Andrulis, E.D., Guzman, E., Doring, P., Werner, J., and Lis, J.T. (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.

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. Nature *420*, 837-841.

Ardehali, M.B., Yao, J., Adelman, K., Fuda, N.J., Petesch, S.J., Webb, W.W., and Lis, J.T. (2009). Spt6 enhances the elongation rate of RNA polymerase II in vivo. EMBO J *28*, 1067-1077.

Biswas, D., Yu, Y., Prall, M., Formosa, T., and Stillman, D.J. (2005). The yeast FACT complex has a role in transcriptional initiation. Mol Cell Biol *25*, 5812-5822.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

Bres, V., Yoh, S.M., and Jones, K.A. (2008). The multi-tasking P-TEFb complex. Curr Opin Cell Biol *20*, 334-340.

Bucheli, M.E., and Buratowski, S. (2005). Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. Embo J *24*, 2150-2160.

Buratowski, S. (2003). The CTD code. Nat Struct Biol *10*, 679-680.

Buratowski, S. (2005). Connections between mRNA 3' end processing and transcription termination. Curr Opin Cell Biol *17*, 257-261.

Burckin, T., Nagel, R., Mandel-Gutfreund, Y., Shiue, L., Clark, T.A., Chong, J.L., Chang, T.H., Squazzo, S., Hartzog, G., and Ares, M., Jr. (2005). Exploring functional relationships between components of the gene expression machinery. Nat Struct Mol Biol *12*, 175-182.

Cairns, B.R. (2001). Emerging roles for chromatin remodeling in cancer biology. Trends Cell Biol *11*, S15-21.

Cairns, B.R. (2005). Chromatin remodeling complexes: strength in diversity, precision through specialization. Curr Opin Genet Dev *15*, 185-190.

Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol *6*, e277.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Clark-Adams, C.D., and Winston, F. (1987). The SPT6 gene is essential for growth and is required for delta-mediated transcription in Saccharomyces cerevisiae. Mol Cell Biol *7*, 679-686.

Das, C., Tyler, J.K., and Churchill, M.E. (2010). The histone shuffle: histone chaperones in an energetic dance. Trends Biochem Sci *35*, 476-489.

Denis, C.L. (1984). Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics *108*, 833-844.

Diebold, M.L., Loeliger, E., Koch, M., Winston, F., Cavarelli, J., and Romier, C. (2010). Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II. J Biol Chem *285*, 38389-38398.

Donham, D.C., 2nd, Scorgie, J.K., and Churchill, M.E. (2011). The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4-DNA complexes. Nucleic Acids Res *39*, 5449-5458.

Ebright, R.H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J Mol Biol *304*, 687-698.

Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D.K., Aida, M., Inukai, N., Narita, T., Yamada, T., Furuya, A.*, et al.* (2004). Human Spt6 stimulates

transcription elongation by RNA polymerase II in vitro. Mol Cell Biol *24*, 3324-3336.

Feklistov, A., and Darst, S.A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. Cell *147*, 1257-1269.

Fischbeck, J.A., Kraemer, S.M., and Stargell, L.A. (2002). SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant. Genetics *162*, 1605-1616.

Formosa, T. (2012). The role of FACT in making and breaking nucleosomes. Biochim Biophys Acta *1819*, 247-255.

Formosa, T., and Nittis, T. (1998). Suppressors of the temperature sensitivity of DNA polymerase alpha mutations in Saccharomyces cerevisiae. Mol Gen Genet *257*, 461-468.

Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell *116*, 63-73.

Goodrich, J.A., and Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell *77*, 145-156.

Hartzog, G.A., Wada, T., Handa, H., and Winston, F. (1998). Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. Genes Dev *12*, 357-369.

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet *9*, 15-26.

Hondele, M., and Ladurner, A.G. (2011). The chaperone-histone partnership: for the greater good of histone traffic and chromatin plasticity. Curr Opin Struct Biol *21*, 698-708.

Hurwitz, J. (2005). The discovery of RNA polymerase. J Biol Chem *280*, 42477-42485.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657-668.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Johnson, S.A., Cubberley, G., and Bentley, D.L. (2009). Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. Mol Cell *33*, 215-226.

Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E., and Bartel, D.P. (2001). RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science *292*, 1319-1325.

Kaplan, C.D., Holland, M.J., and 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.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

Kaplan, C.D., Morris, J.R., Wu, C., and Winston, F. (2000). Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. Genes Dev *14*, 2623-2634.

Keegan, B.R., Feldman, J.L., Lee, D.H., Koos, D.S., Ho, R.K., Stainier, D.Y., and Yelon, D. (2002). The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription in the zebrafish embryo. Development *129*, 1623-1632.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517-522.

Korkhin, Y., Unligil, U.M., Littlefield, O., Nelson, P.J., Stuart, D.I., Sigler, P.B., Bell, S.D., and Abrescia, N.G. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. PLoS Biol *7*, e1000102.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R.D. (1977). Structure of chromatin. Annu Rev Biochem *46*, 931-954.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. Proc Natl Acad Sci U S A *104*, 12955-12961.

Kornberg, R.D., and Lorch, Y. (1992). Chromatin structure and transcription. Annu Rev Cell Biol *8*, 563-587.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. (2004). Ssu72 Is an RNA polymerase II CTD phosphatase. Mol Cell *14*, 387-394.

Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J.F. (2002). RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol *22*, 6979-6992.

Kwon, S.H., and Workman, J.L. (2011). The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. Bioessays *33*, 280-289.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-4060.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Curr Opin Genet Dev *21*, 175-186.

Li, L., Ye, H., Guo, H., and Yin, Y. (2010). Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. Proc Natl Acad Sci U S A *107*, 3918-3923.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., 3rd, and Hartzog, G.A. (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.

Ling, Y., Smith, A.J., and Morgan, G.T. (2006). A sequence motif conserved in diverse nuclear proteins identifies a protein interaction domain utilised for nuclear targeting by human TFIIS. Nucleic Acids Res *34*, 2219-2229.

Liu, X., Bushnell, D.A., Silva, D.A., Huang, X., and Kornberg, R.D. (2011). Initiation complex structure and promoter proofreading. Science *333*, 633-637.

Liu, Z., Zhou, Z., Chen, G., and Bao, S. (2007). A putative transcriptional elongation factor hIws1 is essential for mammalian cell proliferation. Biochem Biophys Res Commun *353*, 47-53.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol Biol *119*, 1-16.

Luna, R., Gaillard, H., Gonzalez-Aguilera, C., and Aguilera, A. (2008). Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. Chromosoma *117*, 319-331.

Luo, W., Johnson, A.W., and Bentley, D.L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev *20*, 954-965.

Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science *336*, 1723-1725.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol *17*, 1272-1278.

Mello, C.C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. Nature *431*, 338-342.

Mello, J.A., and Almouzni, G. (2001). The ins and outs of nucleosome assembly. Curr Opin Genet Dev *11*, 136-141.

Moteki, S., and Price, D. (2002). Functional coupling of capping and transcription of mRNA. Mol Cell *10*, 599-609.

Mousson, F., Ochsenbein, F., and Mann, C. (2007). The histone chaperone Asf1 at the crossroads of chromatin and DNA checkpoint pathways. Chromosoma *116*, 79-93.

Neigeborn, L., Celenza, J.L., and Carlson, M. (1987). SSN20 is an essential gene with mutant alleles that suppress defects in SUC2 transcription in Saccharomyces cerevisiae. Mol Cell Biol *7*, 672-678.

Park, Y.J., and Luger, K. (2006). Structure and function of nucleosome assembly proteins. Biochem Cell Biol *84*, 549-558.

Polo, S.E., and Almouzni, G. (2006). Chromatin assembly: a basic recipe with various flavours. Curr Opin Genet Dev *16*, 104-111.

Pujari, V., Radebaugh, C.A., Chodaparambil, J.V., Muthurajan, U.M., Almeida, A.R., Fischbeck, J.A., Luger, K., and Stargell, L.A. (2010). The transcription factor Spn1 regulates gene expression via a highly conserved novel structural motif. J Mol Biol *404*, 1-15.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. Cell *108*, 557-572.

Rasmussen, E.B., and Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proc Natl Acad Sci U S A *90*, 7923-7927.

Rocha, W., and Verreault, A. (2008). Clothing up DNA for all seasons: Histone chaperones and nucleosome assembly pathways. FEBS Lett *582*, 1938-1949.

Rosonina, E., Kaneko, S., and Manley, J.L. (2006). Terminating the transcript: breaking up is hard to do. Genes Dev *20*, 1050-1056.

Saguez, C., Olesen, J.R., and Jensen, T.H. (2005). Formation of export-competent mRNP: escaping nuclear destruction. Curr Opin Cell Biol *17*, 287-293.

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

Semlow, D.R., and Staley, J.P. (2012). Staying on message: ensuring fidelity in pre-mRNA splicing. Trends Biochem Sci *37*, 263-273.

Simchen, G., Winston, F., Styles, C.A., and Fink, G.R. (1984). Ty-mediated gene expression of the LYS2 and HIS4 genes of Saccharomyces cerevisiae is controlled by the same SPT genes. Proc Natl Acad Sci U S A *81*, 2431-2434.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Strasser, K., and Hurt, E. (2000). Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. EMBO J *19*, 410-420.

Sullivan, S.A., and Landsman, D. (2003). Characterization of sequence variability in nucleosome core histone folds. Proteins *52*, 454-465.

Sun, M., Lariviere, L., Dengl, S., Mayer, A., and Cramer, P. (2010). A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). J Biol Chem *285*, 41597-41603.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Thomas, J.O., and Kornberg, R.D. (1975). An octamer of histones in chromatin and free in solution. Proc Natl Acad Sci U S A *72*, 2626-2630.

Tyler, J.K. (2002). Chromatin assembly. Cooperation between histone chaperones and ATP-dependent nucleosome remodeling machines. Eur J Biochem *269*, 2268-2274.

VanDemark, A.P., Blanksma, M., Ferris, E., Heroux, A., Hill, C.P., and Formosa, T. (2006). The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Mol Cell *22*, 363-374.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Voigt, P., Leroy, G., Drury, W.J., 3rd, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically modified nucleosomes. Cell *151*, 181-193.

Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F., and Jones, D.T. (2004). The DISOPRED server for the prediction of protein disorder. Bioinformatics *20*, 2138-2139.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522-525.

Winkler, D.D., Muthurajan, U.M., Hieb, A.R., and Luger, K. (2011). Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events. J Biol Chem *286*, 41883-41892.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev *21*, 160-174.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes & Development *22*, 3422-3434.

Zakrzewska, K., and Lavery, R. (2012). Towards a molecular view of transcriptional control. Curr Opin Struct Biol *22*, 160-167.

Zhang, L., Fletcher, A.G., Cheung, V., Winston, F., and Stargell, L.A. (2008). Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. Mol Cell Biol *28*, 1393-1403.

Zhang, Z., Fu, J., and Gilmour, D.S. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. Genes Dev *19*, 1572-1580.

Zhang, Z., and Gilmour, D.S. (2006). Pcf11 is a termination factor in Drosophila that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. Mol Cell *21*, 65-74.

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.

CHAPTER 2

CRYSTAL STRUCTURES OF THE *S. CEREVISIAE* SPT6 CORE AND C-TERMINAL TANDEM SH2 DOMAIN

Devin Close, Sean J. Johnson, Matthew A. Sdano,

Seth M. McDonald, Howard Robinson,

Tim Formosa, and

Christopher P. Hill

(2011)

Journal of Molecular Biology

408, 697-713

Note: See Chapter 1, Outline of Chapters, for a description of my contributions to this publication

doi:10.1016/j.jmb.2011.03.002 J. Mol. Biol. (2011) 408, 697–713

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Crystal Structures of the *S. cerevisiae* Spt6 Core and C-Terminal Tandem SH2 Domain

Devin Close[1], Sean J. Johnson[2], Matthew A. Sdano[1], Seth M. McDonald[1], Howard Robinson[3], Tim Formosa[1] and Christopher P. Hill[1]*

[1]*Department of Biochemistry, University of Utah, Salt Lake City, UT 84112-5650, USA*
[2]*Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322-0300, USA*
[3]*Department of Biology, Brookhaven National Laboratory, Upton, NY 11973, USA*

Received 2 December 2010; received in revised form 24 February 2011; accepted 1 March 2011
Available online 17 March 2011

Edited by K. Morikawa

Keywords:
protein structure;
protein function;
gene expression;
crystallography

The conserved and essential eukaryotic protein Spt6 functions in transcription elongation, chromatin maintenance, and RNA processing. Spt6 has three characterized functions. It is a histone chaperone capable of reassembling nucleosomes, a central component of transcription elongation complexes, and is required for recruitment of RNA processing factors to elongating RNA polymerase II (RNAPII). Here, we report multiple crystal structures of the 168-kDa Spt6 protein from *Saccharomyces cerevisiae* that together represent essentially all of the ordered sequence. Our two structures of the ~900-residue core region reveal a series of putative nucleic acid and protein–protein interaction domains that fold into an elongated form that resembles the bacterial protein Tex. The similarity to a bacterial transcription factor suggests that the core domain performs nucleosome-independent activities, and as with Tex, we find that Spt6 binds DNA. Unlike Tex, however, the Spt6 S1 domain does not contribute to this activity. Crystal structures of the Spt6 C-terminal region reveal a tandem SH2 domain structure composed of two closely associated SH2 folds. One of these SH2 folds is cryptic, while the other shares striking structural similarity with metazoan SH2 domains and possesses structural features associated with the ability to bind phosphorylated substrates including phosphotyrosine. Binding studies with phosphopeptides that mimic the RNAPII C-terminal domain revealed affinities typical of other RNAPII C-terminal domain-binding proteins but did not indicate a specific interaction. Overall, these findings provide a structural foundation for understanding how Spt6 encodes several distinct functions within a single polypeptide chain.

**Corresponding author.* E-mail address: chris@biochem.utah.edu.
Present address: D. Close, Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

Abbreviations used: CTD, C-terminal domain; DLD, death-like domain; dsDNA, double-stranded DNA; FA, fluorescence anisotropy; HhH, helix–hairpin–helix; HtH, helix–turn–helix; NSLS, National Synchrotron Light Source; OB, oligonucleotide–oligosaccharide binding; PDB, Protein Data Bank; pSer, phosphoserine; pTyr, phosphotyrosine; RNAPII, RNA polymerase II; SH2, Src homology 2; SSRL, Stanford Synchrotron Radiation Laboratory; WT, wild type; SeMet, selenomethionine substituted; NIH, National Institutes of Health.

Introduction

Gene expression in eukaryotes relies on a synergistic relationship between transcription, RNA processing, and chromatin structure.[1,2] The specific positioning, composition, and posttranslational modification of nucleosomes defines a code for chromatin-templated transcriptional regulation. Moreover, transcription is intimately tied to mRNA processing, surveillance, and export from the nucleus. This coordination relies on precise cooperation among many proteins, with Spt6 being remarkable for playing multifaceted roles in several distinct processes.

Spt6 (suppressor of Ty 6) was originally discovered in *Saccharomyces cerevisiae* as a gene that influences general transcription through manipulation of chromatin structure at upstream promoter elements.[3] Subsequently, Spt6 has been implicated in a variety of biological processes in organisms ranging from yeasts to human, including embryogenesis in zebrafish,[4] multiple stages of development in *Drosophila*,[5] gut morphogenesis in *Caenorhabditis elegans*,[6] signal transduction in mammals,[7,8] and pathogenesis of human immunodeficiency virus.[9,10] The broad utility of Spt6 stems from its ability to perform multiple functions as a histone chaperone, a transcription elongation factor, and a modulator of RNA transcript processing.

Spt6 is required for reassembly of nucleosomes in the wake of an elongating RNA polymerase II (RNAPII), a function that has profound regulatory effects at both intergenic and intragenic start sites.[11,12] Spt6 binds directly to histones and nucleosomes *in vitro*,[13,14] and these activities may contribute to the nucleosome reassembly function. In addition, Spt6 recruits the H3K36 methyltransferase Set2 to the transcription complex,[15] providing a link between the processes of transcription and histone modification. While its roles in modifying and reassembling nucleosomes indirectly influence the elongation rate, Spt6 also directly affects RNAPII, as it stimulates elongation on nucleosome-free DNA templates *in vitro*.[9,16] This role as an elongation factor independent of its effects on chromatin may also be significant *in vivo*, as knocking down Spt6 caused a decrease in the RNAPII elongation rate even in regions where the chromatin was considered to be permissive to transcription.[16] Yet another role as a modulator of transcript processing is indicated by the association of Spt6 with the Rrp6 subunit of the *Drosophila* exosome RNA processing complex[17] and by the requirement for Spt6 to prevent premature 3′ processing at cryptic polyadenylation signals upstream of the appropriate sites.[18] It has also been demonstrated that mammalian Spt6 can bind RNAPII C-terminal domain (CTD) phosphorylated at Ser2 by the P-TEFb kinase and that this interaction can subsequently promote recruitment of RNA processing/export factors such as REF1/Aly.[9,15] Binding to the phosphorylated RNAPII CTD is mediated by a Src homology 2 (SH2) domain that is located near the C-terminus of Spt6 and is conserved from yeast to human.[9] SH2 domains typically recognize phosphorylated tyrosine residues, are ubiquitous in metazoans, and are the primary recognition motif in phosphorylation-mediated signal transduction cascades.[19] Strikingly, the Spt6 SH2 domain is the only SH2 domain predicted to occur in the yeast

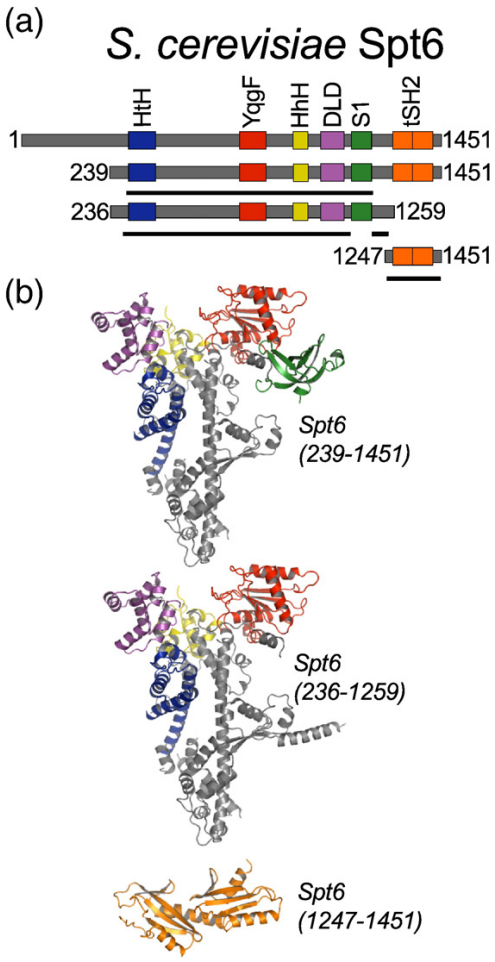


Fig. 1. Spt6 structures. (a) Schematic representation of full-length (top) Spt6 and constructs used for crystallization. The three crystallized constructs (239–1451, 236–1259, and 1247–1451) are shown below, with black continuous lines indicating regions of the protein constructs visible in each of the crystal structures. (b) Three independently determined Spt6 crystal structures colored by domain as in (a).

proteome.[20] Spt6 therefore participates in a wide range of functions affecting transcription, with each activity requiring different subsets of its multiple distinct functional domains.

We have determined multiple crystal structures of Spt6 from *S. cerevisiae* and find that, consistent with the range of functional domains inferred from previous studies, it comprises a series of structural domains whose homologs are known to function in nucleic acid binding and/or protein–protein interactions. The core of the structure comprises several recognizable structural motifs and, in composite, resembles the bacterial transcription factor Tex.[21] A C-terminal region that is tethered to the core by a flexible linker adopts a novel tandem SH2 domain comprising two closely associated SH2 folds, one of which corresponds to the previously predicted SH2 domain of Spt6 and contains many of the standard binding determinants characteristic of this family, while the other lacks these features but contributes to a putative specificity pocket of the more canonical SH2 domain. Our structure of the Spt6 tandem SH2 domain resembles two recently reported homologous Spt6 structures.[22,23] We also show that the Spt6 core domain has DNA-binding activity, and we examine the interaction between the Spt6 tandem SH2 domain and RNAPII-derived peptides for evidence of a phosphorylation-dependent interaction with the CTD.

Results and Discussion

Crystal structures of Spt6(236–1259), Spt6(239–1451), and Spt6(1247–1451)

We have determined three crystal structures that together comprise the entire ordered region of the 1451-residue *S. cerevisiae* Spt6 protein (Figs. 1–3). Based on these structures and sequence analysis, Spt6 residues 1–297, 456–464, 485–500, 562–566, 1003–1008, 1211–1217, and 1441–1451 are likely to

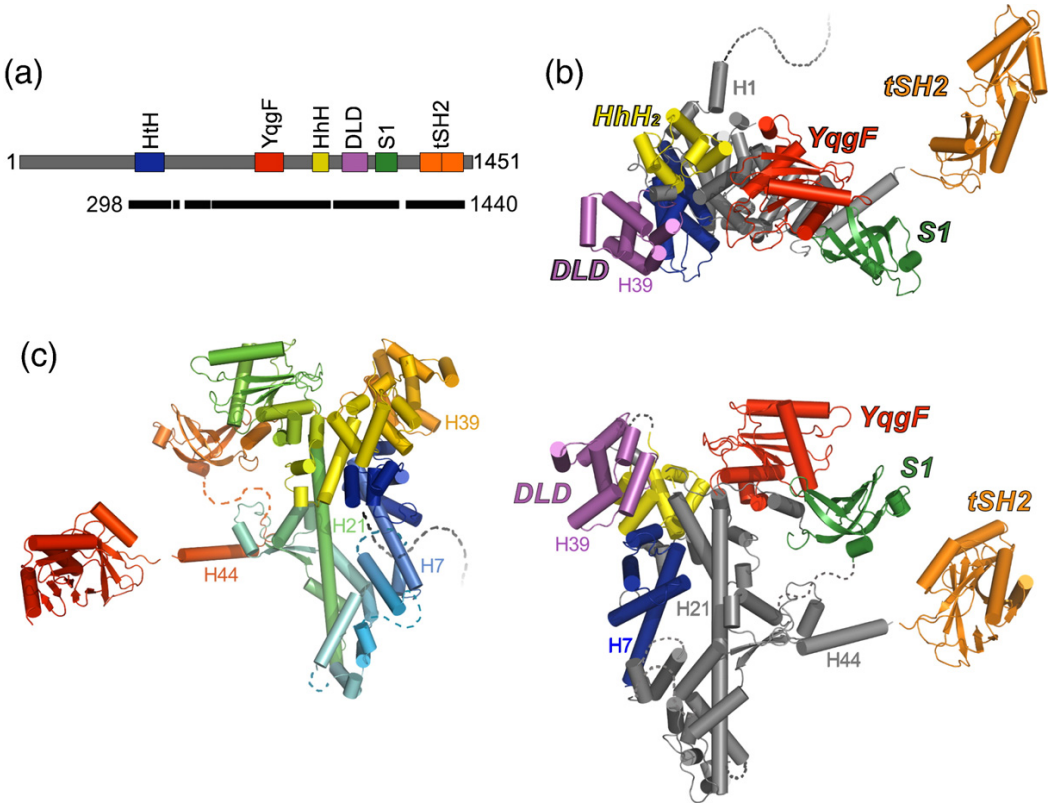


Fig. 2. Composite model of Spt6. (a) Schematic model of the Spt6 protein as in Fig. 1a. Black bar indicates segments of the protein represented by the composite model. (b) Two views of the composite model of Spt6 colored by domain. Approximately two hundred C-terminal residues (tSH2 domain) are expected to be highly mobile with respect to the core. Secondary structure elements mentioned in the text are labeled. Broken lines represent regions of the Spt6 protein that are not visible in our structures and are likely to be disordered, including the first 297 residues. (c) Composite model of Spt6 colored from the N- to C-terminus (blue to red). View orientation is rotated 180° from that of the lower image in (b).

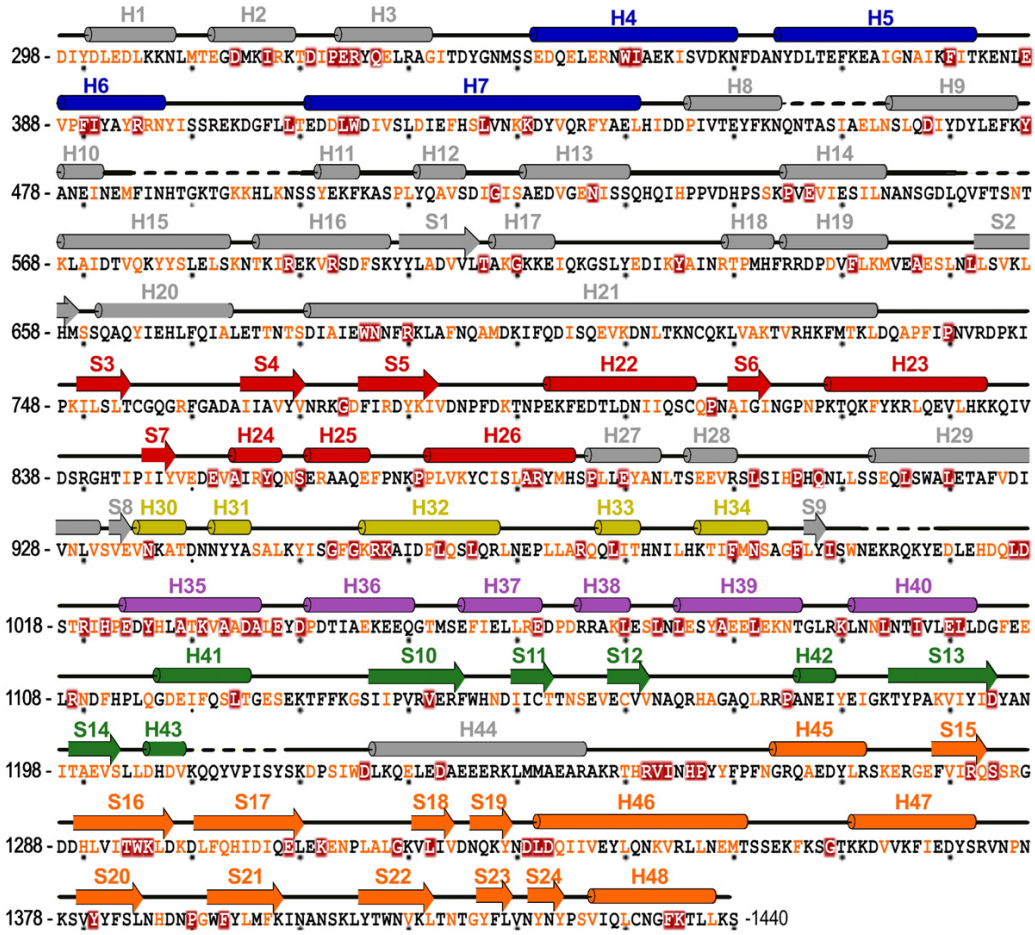


Fig. 3. Spt6 sequence. Spt6 sequence present in the composite model with corresponding secondary structure elements colored by domain as in Figs. 1 and 2. Coloring of sequence represents degree of conservation [dark-red background, invariant; orange font, conserved] in an alignment (see Supplemental Fig. S1) of proteins from *S. cerevisiae*, *Schizosaccharomyces pombe*, *C. elegans*, *Drosophila melanogaster*, *Danio rerio*, and *Homo sapiens*. Alignment was performed using T-coffee.[24] Broken lines indicate regions of disorder that are not included in the model(s).

be disordered in the full-length protein, at least in the absence of binding partners. Spt6 displays multiple recognizable structural domains whose homologs have been implicated in binding of nucleic acids or proteins, although conservation at the sequence level is low, and only three of these domains [$(HhH)_2$, YqgF, and S1] in the Spt6 core were predicted from the sequence.[25]

Full-length Spt6 expressed poorly in *Escherichia coli* and did not yield crystals. In contrast, two different Spt6 constructs that lacked the first ~235 residues and either lacked the C-terminal 192 residues [Spt6 (236–1259)] or extended to the C-terminus [Spt6 (239–1451)] expressed well, and the resulting proteins crystallized in different space groups (Table 1). The Spt6(236–1259) structure was determined by two-wavelength anomalous diffraction using data collected to 2.7 Å from crystals of selenomethionine-substituted (SeMet) protein and was refined against 2.6-Å native data to R/R_{free} values of 22.4%/26.5%. The Spt6(239–1451) structure was determined by molecular replacement using the Spt6(236–1259) structure as a search model and refined against 3.3-Å data to R/R_{free} values of 26.5%/30.8%.

Together, Spt6(236–1259) and Spt6(239–1451) display an ordered structure for the Spt6 core (residues 298–1248), which is primarily helical (54.6% helical, 8.3% strand, 37.1% coil), has overall dimensions of 110 Å × 77 Å × 36 Å, and is quite similar between the two structures (RMSD = 1.2 Å over 747 out of 763 pairs of C^{α} atoms that are ordered in both structures). The core is built around an ~80-Å-long central helix (H21, 680–733), with the rest of the

Table 1. Crystallization conditions

	Spt6(236–1259)	Spt6(236–1259)	Spt6(239–1451)	Spt6(1247–1451)	Spt6(1247–1451)
Crystal	SeMet	Native	Native	SeMet	Native
Protein solution[a]	15 mM Tris, 5% glycerol	15 mM Tris, 5% glycerol	50 mM Tris, 10% glycerol, 5 mM β-mercaptoethanol	15 mM Tris, 5% glycerol	15 mM Tris, 5% glycerol
Protein concentration (mg/ml)	9	8	10	26	26
Reservoir solution	50 mM 4-morpholineethanesulfonic acid (pH 6.0), 13% methyl-2,4-pentanediol, 0.1 M KCl, 5 mM $MgCl_2$	50 mM $Mg(HCO_2)_2$, 10% glycerol	0.1 M Na–4-morpholineethanesulfonic acid (pH 6.5), 13% PEG 4000, 0.2 M $MgCl_2$, 12% glycerol	22% PEG 4000, 0.3 M $(NH_4)_2SO_4$	0.1 M Tris (pH 5.5), 25% PEG 3350, 0.2 M $(NH_4)_2SO_4$
Drop ratio (μl:μl), protein:reservoir	2:2	3:1	3:1.5	2: 2	1:2
Additive	—	0.6 μl 1.0 M NDSB-256	—	—	—
Temperature (°C)	13	13	13	21	21
Cryo solution[b]	30% methyl-2,4-pentanediol	35% glycerol	30% glycerol	30% glycerol	20% glycerol

PEG, polyethylene glycol.
[a] All protein solutions were in 100 mM NaCl and Tris, pH 7.5 (defined at room temperature).
[b] Cryoprotectant solutions were made the same as for the reservoir solution, but including glycerol.

protein chain wrapping around this helix at both ends to give an overall V shape. Additionally, Spt6 (239–1451) displayed interpretable density for the S1 domain (residues 1128–1210), whereas the corresponding density in Spt6(236–1259) was poorly defined, and the S1 domain was not included in this refined structure. As discussed below, we favor a model in which the S1 domain is highly mobile in solution. Interestingly, the 240 C-terminal residues of Spt6(239–1451), encompassing all of the residues absent from the shorter Spt6(236–1259) construct, are not visible in the electron density. The Spt6(239–1451) crystals have an estimated solvent content of 60% with cavities that could accommodate the 240 C-terminal residues, which suggests that this region is tethered by a flexible linker that is mobile both in solution and in the Spt6(239–1451) crystals.

Guided by secondary structure predictions and limited proteolysis experiments (data not shown), we expressed and crystallized the Spt6 C-terminal region, Spt6(1247–1451). This structure was determined by the selenomethionine single-wavelength anomalous diffraction method using 2.7-Å-resolution data and refined to R/R_{free} values of 20.7%/25.4%. A second native crystal form yielded 2.1-Å-resolution data, and this model was refined to R/R_{free} values of 17.9%/21.2%. The SeMet and native proteins crystallized in different space groups with one and four molecules in the asymmetric unit, respectively (Table 2). Surprisingly, we found that the Spt6 C-terminal region comprises not one SH2 domain as anticipated,[20,28] but two SH2 folds that are packed tightly against each other to form a tandem SH2 domain composed of N-terminal (nSH2, residues 1250–1353) and C-terminal (cSH2, residues 1354–1440) folds. While this manuscript was in the final stages of preparation, equivalent tSH2 structures that overlap closely with our refined model were reported for Spt6 homologs from *Candida glabrata*[22] (RMSD = 1.0 Å over 184 C^{α} atoms with 87% sequence identity) and *Antonospora locustae*[23] (RMSD = 1.6 Å over 164 C^{α} atoms with 24% sequence identity).

N-terminal region

The first ~300 amino acids of Spt6 are extremely acidic, with an overall charge of −62 and a predicted p*I* of 4.3, and are also predicted to be disordered.[29] This is consistent with our Spt6(236–1259) and Spt6 (239–1451) structures, which lacked discernible density prior to residue 298. Despite the lack of inherent order, this region of Spt6 is functionally important. Residues 239–268 bind the essential Spn1/Iws1 protein and overlap with residues required for nucleosome binding.[14] Furthermore, the *spt6-1002* allele, a deletion of residues 2–122, displays synthetic lethality with deletion of the gene encoding the transcription factor Paf1.[18]

Helix–turn–helix domain

Residues 336–442 resemble a DNA-binding helix–turn–helix (HtH) motif, as seen in transcription factors such as c-Myb and the bacterial sigma factors.[30] Nevertheless, the structure is not consistent with binding DNA in a canonical manner; binding of Spt6 H7 in the major groove of DNA in the manner predicted for a canonical HtH domain would cause steric clashes between bound DNA and the rest of the structure. It is possible that Spt6 undergoes conformational changes upon binding DNA or that the Spt6 HtH domain serves as a protein–protein interaction motif, as occurs with members of the PWI subgroup of HtH domains.[31] The Spt6 HtH overlaps with the U4/U6 ribonucleoprotein Prp3 PWI domain [RMSD of 2.7 Å, 69 C^{α}, Protein Data Bank (PDB) code 1x4q] and the Nab2 PWI domain (RMSD of 2.9 Å, 73 C^{α}, PDB code 2v75), conserves the eponymous PWI motif as NWI (Asn349-Trp350-Ile351), and could utilize the equivalent protein-binding surface without invoking a conformational change in Spt6.

YqgF homologous domain

The Spt6 YqgF domain (residues 735–887) resembles members of the YqgFc superfamily, such as the *E. coli* protein YqgF and the RuvC class of Holliday junction resolvases.[25] The alignment is especially close with *E. coli* RuvC (RMSD of 2.9 Å, 117 C^{α}, PDB code 1hjr). Despite this similarity, the putative Spt6 YqgF catalytic site lacks carboxylate side chains that are critical for coordinating magnesium ions

Table 2. Data collection and refinement statistics

	Se-Spt6 (236–1259) peak	Se-Spt6 (236–1259) inflection	Spt6 (236–1259)	Spt6 (239–1451)	Se-Spt6 (1247–1451) peak	Spt6 (1247–1451)
Data collection						
Space group	$P2_12_12_1$		$P2_12_12_1$	$P3_121$	$I4_122$	$P2_12_12_1$
Unit cell dimensions (Å)	$a=114.0$		$a=115.1$	$a=118.7$	$a=97.5$	$a=78.8$
	$b=116.4$		$b=116.2$	$b=118.7$	$b=97.5$	$b=105.2$
	$c=122.7$		$c=117.4$	$c=214.4$	$c=132.4$	$c=119.5$
Molecules per asymmetric unit	1		1	1	1	4
Solvent content (%)	63.7		62.4	60.0	63.5	53.1
Beamline[a]	SSRL 11-1	SSRL 11-1	SSRL 7-1	NSLS X29	SSRL 9-1	Home
Wavelength (Å)	0.97886	0.97922	0.97773	1.10000	0.97908	1.54178
Resolution (Å)	50–2.7	50–2.7	32–2.6	46–3.3	35–2.7	30–2.1
High-resolution shell (Å)	(2.8–2.7)	(2.8–2.7)	(2.7–2.6)	(3.42–3.3)	(2.8–2.7)	(2.18–2.1)
No. of unique reflections	83,751	82,946	49,595	27,025	8481	58,489
No. of total reflections	261,886	257,334	434,557	182,336	649,419	258,498
Mean I/σ_I	22.1 (2.4)	22.1 (2.1)	29.8 (4.3)	39.8 (3.3)	25.1 (3.7)	16.8 (3.2)
Completeness (%)	95.8 (75.8)	94.7 (70.1)	99.8 (99.9)	99.5 (99.9)	92.2 (74.5)	99.5 (99.2)
R_{sym} (%)[b]	4.8 (35.7)	4.7 (38.3)	6.5 (43.9)	5.8 (56.9)	7.0 (29.6)	6.6 (47.8)
Refinement						
R_{cryst}[c]/R_{free}[d] (%)			22.4/26.5	26.5/30.8	20.7/25.4	17.9/21.2
No. of non-H atoms						
Protein			6788	6876	1575	6988
Solvent			48	0	19	491
$\langle B \rangle$ (Å^2)			103.7	167.6	74.7	44.2
RMSD bond lengths (Å)			0.006	0.013	0.008	0.004
RMSD bond angles (°)			0.895	1.59	1.12	0.757
Ramachandran outliers (%)			0.0	0.8	0.0	0.40
Ramachandran favored (%)			95.6	93.6	97.9	97.5
Rotamer outliers (%)			0.94	0.3	0.6	1.4

Values in parentheses correspond to the high-resolution shell.
Refinement statistics were determined by PHENIX[26] and MolProbity.[27]
[a] Data were collected at the SSRL, the NSLS, or on a Rigaku MicroMax-007HF rotating anode X-ray generator with a copper anode and VariMax confocal optics and a Rigaku R-AXIS IV image plate detector (home).
[b] $R_{sym}=(\sum|(I-\langle I\rangle)|)/(\sum I)$, where $\langle I\rangle$ is the average intensity of multiple measurements.
[c] $R_{cryst}=(\sum|F_{obs}-F_{calc}|)/(\sum|F_{obs}|)$.
[d] R_{free} is the R_{cryst} based on ~1000 (at least 10%) of the reflections that were excluded from refinement.

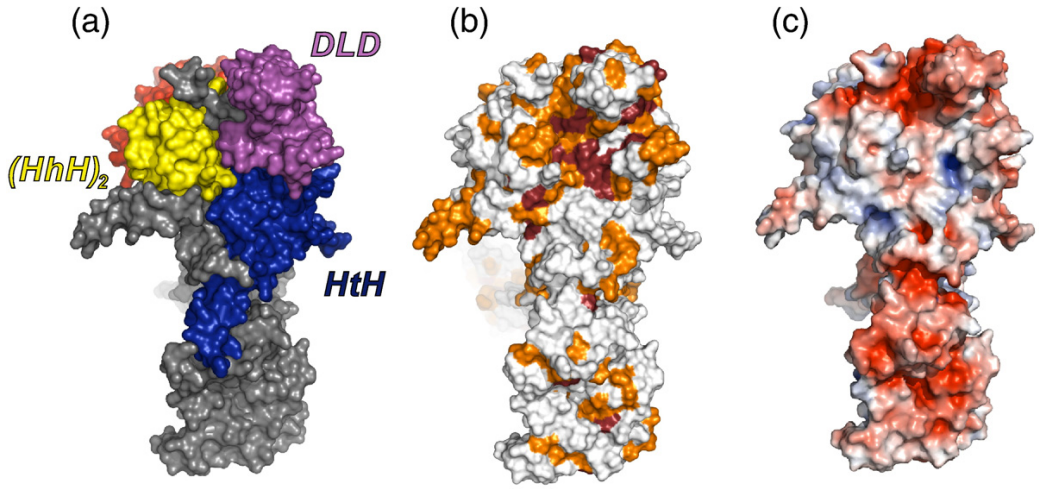


Fig. 4. The most conserved surface of the Spt6 core. (a) View of the Spt6 core showing the interface between HtH, DLD, and $(HhH)_2$ domains. (b) Same orientation as (a) but colored by conservation to illustrate the high level of surface conservation at the intersection of these domains, especially on the DLD. Coloring represents degree of conservation as described in Fig. 3 and Supplemental Fig. S1. (c) Same as (b) but colored by electrostatic potential (−5 to +5 kT/e).

that mediate phosphodiester bond hydrolytic cleavage.[32] Thus, it does not appear that the Spt6 YqgF fold is capable of nuclease activity using a catalytic mechanism similar to that of RuvC or related RNase H-fold nucleases.

Helix–hairpin–helix domain

Residues 933–1002 form two consecutive helix–hairpin–helix (HhH) motifs that pack together through highly conserved hydrophobic residues at an ~90° angle to form a $(HhH)_2$ domain that resembles known double-stranded DNA (dsDNA) binding domains.[33] These include proteins such as *E. coli* RNA polymerase α CTD (RMSD of 2.3 Å, 56 C^{α}, PDB code 1lb2) and the Holliday junction-binding protein RuvA (RMSD of 2.2 Å, 60 C^{α}, PDB code 1bvs). The first Spt6 HhH represents a characteristic HhH motif in the relative angle of the antiparallel helices and the presence of the Gly-hydrophobic-Gly motif within the hairpin loop.[33] The second HhH motif is more variant, as also observed in other $(HhH)_2$ domains, including DNA polymerase β and 5′ to 3′ exonucleases. Though

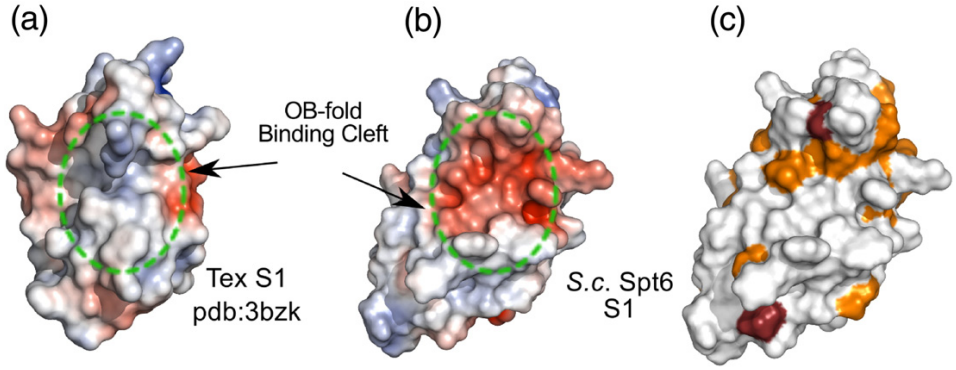


Fig. 5. Surface representations of the Spt6 and Tex S1 domains. (a) Electrostatic surface representation (−5 to +5 kT/e) of the S1 domain from *Pseudomonas aeruginosa* Tex (PDB code 3bzk) with the position of the nucleic acid binding OB-fold cleft approximated by the circle with the dotted green line. (b) Electrostatic surface representation (−5 to +5 kT/e) of the *S. cerevisiae* Spt6 S1 domain in the same orientation as in (a) showing a clustering of negative charge in the putative OB-fold binding cleft. (c) *S. cerevisiae* Spt6 S1 domain in the same orientation as (b) but colored by conservation in the same color scheme as in Fig. 3 to illustrate the low level of conservation within the region equivalent to the binding cleft of canonical S1 domains.

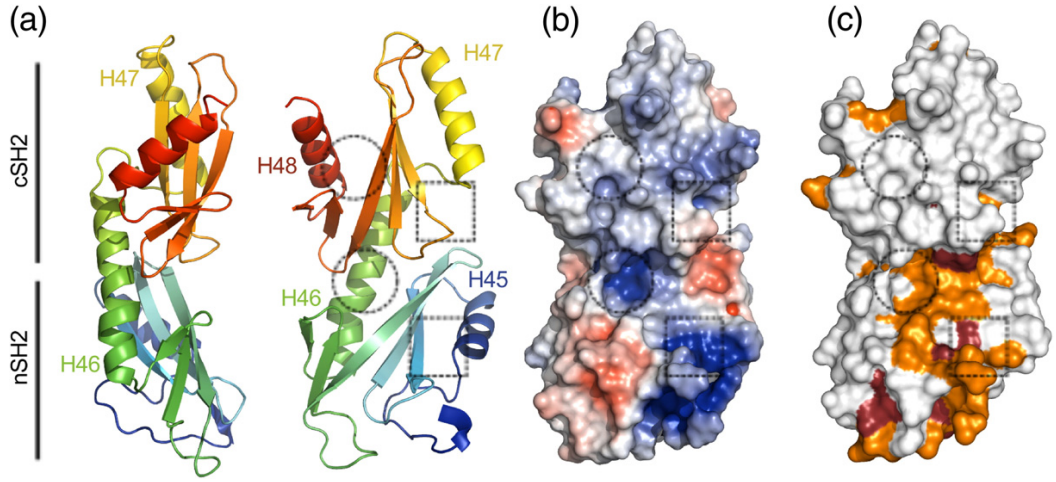


Fig. 6. The *S. cerevisiae* Spt6 tSH2 domain. (a) Two views of a cartoon representation of the tSH2 domain colored from the N- to C-terminus (blue to red; residues 1247–1440 are shown). Dotted squares and circles show the approximate positions of the canonical SH2 domain pTyr and specificity pockets, respectively. (b) Surface representation colored by electrostatic potential surface (−5 to +5 kT/e). (c) Same as (b) but colored by residue conservation in the same color scheme as in Figs. 3 and 4.

$(HhH)_2$ folds are primarily found in proteins that interact with DNA, they also occur in proteins that mediate protein–protein interactions, such as the sterile α motif proteins.[34] Notably, the first 28 ordered residues (~298–325) of Spt6 wrap around the $(HhH)_2$ domain in a fashion that would occlude binding of a canonical $(HhH)_2$ domain to a dsDNA ligand, although this interaction could be transient. The absence of corresponding density for H1 and part of H2 in the initial SeMet Spt6(236–1259) and Spt6(239–1451) maps suggests that these N-terminal residues might adopt a different conformation to allow binding of a physiological partner *in vivo*.

Death-like domain

Residues 1019–1104 form a prominent lobe of the structure that resembles members of the death domain superfamily. Death domains typically serve as recognition modules in proteins that assemble and activate inflammatory and apoptotic complexes.[35] The Spt6 death-like domain (DLD) maintains the characteristic overall topology of death domains, consisting of a six-helix bundle with three stacked antiparallel helices but with an additional helix inserted between the final two helices of the bundle (H39 in Figs. 2 and 3). Spt6 aligns reasonably well with several known death domain superfamily proteins, including the caspase-2-activating PIDDosome PIDD protein subunit component (RMSD of 3.0 Å, 60 C^{α}, PDB code 2of5). Although it is unlikely that the Spt6 DLD functions in an apoptotic process in yeast, its prominent location and the observation that it displays the most highly conserved region of the Spt6 surface suggest that it mediates important intermolecular interactions (Fig. 4).

S1 domain

A mostly unstructured linker of 15 residues leads to the S1 domain (residues 1129–1219), which adopts the canonical S1/oligonucleotide–oligosaccharide binding (OB)-fold of a β-barrel composed of two three-stranded β-sheets where strand 1 (S10) is shared by both sheets.[36] Despite the structural similarity, Spt6 lacks the typical S1 binding cleft residues that are important for binding nucleic acids.[36] In addition, the predicted electrostatic potential surface does not appear conducive to nucleic acid binding, shows a low level of conservation, and, as discussed below, is not required for dsDNA binding (Fig. 5). This is in contrast to the distantly related bacterial Tex protein, which loses its capacity to bind DNA or RNA in the absence of the S1 domain.[21] OB folds are used to bind partners other than nucleic acids, including oligosaccharides and proteins;[36,37] thus, it remains possible that the Spt6 S1 domain is used for an important interaction that does not involve nucleic acids.

Tandem SH2 domain

The S1 domain is followed by an unstructured ~10-residue segment and an ~30-Å helix (H44; 1227–1247) that buries ~440 $Å^2$ of accessible surface area against the core in the Spt6(236–1259) structure (Fig. 2). While some density is present for H44 in the

Spt6(239–1451) maps, this region is too disordered for reliable model building, which may indicate that this interface is not always formed in solution. H44 links the core to a tandem SH2 domain (tSH2; residues 1250–1440) that comprises N-terminal (nSH2; residues 1250–1353) and C-terminal (cSH2; 1353–1440) folds that associate through an ~800-Å^2 interface to form a single structural unit (Fig. 6). Both nSH2 and cSH2 conform to the standard SH2 domain fold (standard SH2 nomenclature in parenthesis) of an N-terminal helix (αA), a central three-stranded β-sheet (βB–βD), a small two-strand extension to the β-sheet (βE–βF), and a second helix (αB).[38] Intervening loops are labeled based on their relative position between these elements (e.g., the BC loop connects the βB and βC strands). The Spt6 nSH2 and cSH2 superimpose well with each other (RMSD of 2.1 Å, 69 C$^\alpha$; Fig. 7) and with the multitude of other characterized SH2 domains, such as those from v-Src kinase (nSH2: RMSD of 2.0 Å, 82 C$^\alpha$; cSH2: RMSD of 2.1 Å, 80 C$^\alpha$, PDB code 1sps; Fig. 7) and Nck2 (nSH2: RMSD of 2.1 Å, 83 C$^\alpha$; cSH2: RMSD of 1.9 Å, 73 C$^\alpha$, PDB code 2cia).

The relative orientation of the two SH2 folds that comprise the Spt6 tSH2 domain is unlike that of previously reported tandem SH2 domains from

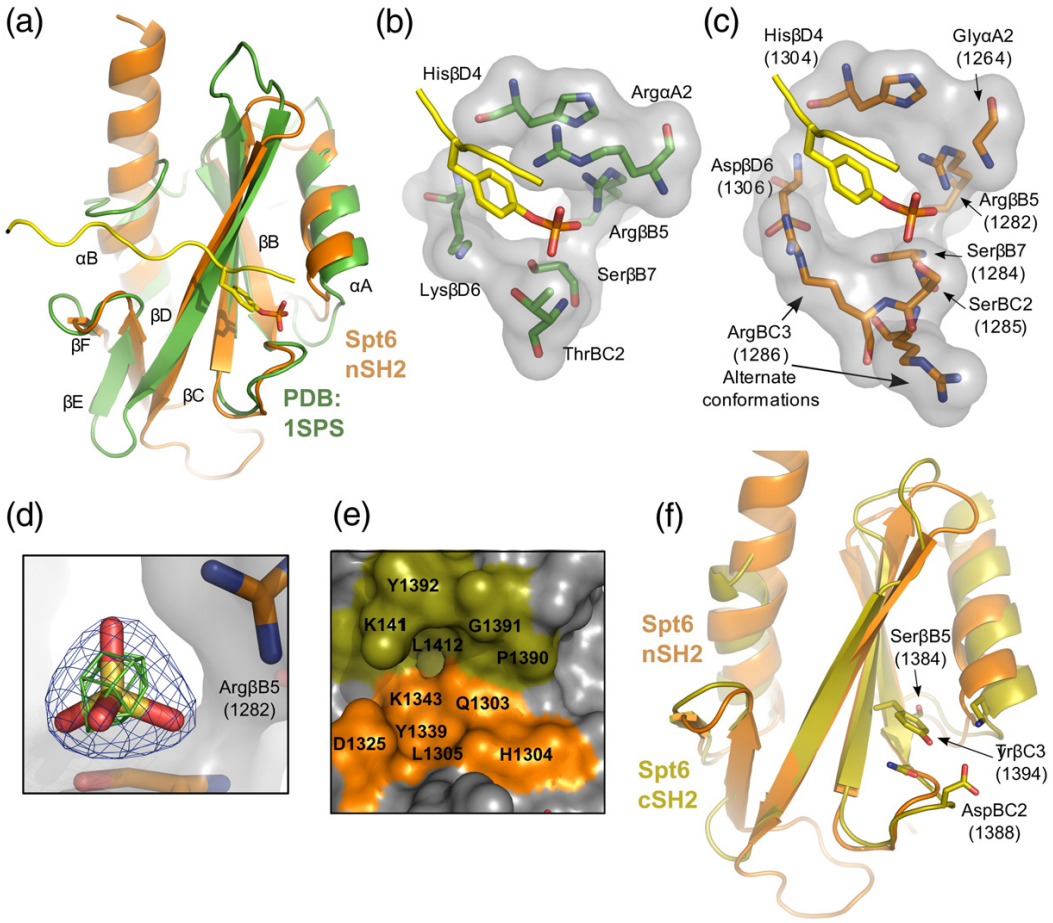


Fig. 7. tSH2 binding pockets. (a) Overlay of Spt6 nSH2 (orange) and the v-Src Kinase SH2 domain (green; PDB code 1sps) bound to a pTyr ligand (yellow). Secondary structure elements are labeled based on the standard SH2 nomenclature.[38] (b) Detailed view of the pTyr binding pocket of 1sps. Residues contributing to the coordination of the pTyr ligand are shown. (c) Same as (b) but for the Spt6 nSH2, with the pTyr peptide from 1sps positioned after the overlap on the SH2 protein domains. (d) Electron density for a sulfate bound in the tSH2 crystal structures. Blue density represents the $2mF_o - DF_c$ map contoured at 2.0 σ, and green density represents an anomalous difference Fourier map contoured at 3.0 σ. (e) The putative nSH2 specificity pocket. Residues from both nSH2 (orange) and cSH2 (olive) line the nSH2 specificity pocket. (f) Alignment of the Spt6 nSH2 (orange) and cSH2 (olive) folds. Residues that protrude into the typical location of the pTyr binding pocket of the cSH2 fold are shown.

other proteins. The αB helix of nSH2 undergoes an ~20° kink where the αB helix would end in a canonical SH2 domain and extends along the backside of the cSH2 fold to form an extensive hydrophobic packing interface with the cSH2 central β-sheet. The relative orientation of nSH2 and cSH2 folds therefore appears to be constrained, consistent with the observation that superposition of the five crystallographically independent tSH2 domains observed in our various crystal forms [one in SeMet Spt6(1247–1451) and four in native Spt6(1247–1451)] indicates a maximum relative rotation of ~8° between nSH2 and cSH2 folds.

Implications for ligand binding by the tSH2 domain

The extent to which Spt6 nSH2 corresponds to a prototypic SH2 domain is seen in a comparison with v-Src (PDB code 1sps) (Fig. 7a–c). The primary determinants of phosphate binding are preserved in nSH2, whereas the positively charged Arg/Lys side chains of classical SH2 domains that flank the aromatic ring of phosphotyrosine (pTyr) ligands are absent from their usual positions. Conserved residues include the consensus FLVRES (FVIRQS, 1279–1284 in Spt6) sequence motif that contributes the phosphate-coordinating Arg1282 and Ser1284 side chains (Spt6 numbering). Moreover, the following Ser1285 Spt6 side chain is also well positioned to hydrogen bond the phosphate and functionally substitute the ThrBC2 of the Src SH2 domain. Other positions within SH2 domains that are important for binding phosphate include HisβD4 (Spt6 H1304) and Ser/ThrβC4 (Spt6 Thr1294), whose side chains hydrogen bond the Arg1282 side chain in an optimum orientation for phosphate binding and, in cognate SH2–ligand complexes, also form a main chain to main chain hydrogen bond with the ligand residue following the pTyr.

Classical SH2 domains typically have a basic residue at the αA2 position that binds against one side of the tyrosine ligand aromatic ring where it forms an amino–aromatic interaction with the π ring and also hydrogen bonds with both the phosphate and the tyrosine main-chain carbonyl. In contrast, the Spt6 nSH2 has a glycine at this position, G1264, which can make none of the same ligand interactions. This does not argue strongly against binding of pTyr by Spt6, however, because a number of other SH2 domains that bind pTyr ligands have a variety of substitutions at this position, including the PTPN11/SHPTP2/Syp phosphatase (PDB code 1ayc) that, like Spt6, has a Gly at αA2 and is known to bind a pTyr-containing peptide.[39] On the other side of a canonical pTyr ligand side chain, classical SH2 domains typically have a basic residue in the βD6 position. This residue is a lysine in Src but is an aspartate (D1306) in Spt6. Interestingly, in three of the five crystallographically independent Spt6 tSH2 molecules in our structures, the space typically occupied by the basic βD6 side chain is filled by R1286 in the BC3 position (Fig. 7c). Thus, Spt6 retains the ability to provide a positively charged basic group in this position, consistent with the potential to bind pTyr.

Ammonium sulfate was present in the crystallization solutions for both native and SeMet tSH2 domain structures, and all five crystallographically independent molecules displayed a sulfate ion at the putative nSH2 phosphate binding site, where it forms hydrogen bonds with R1282, S1284, and S1285 in the same manner as the phosphate of pTyr–SH2 complexes (Fig. 7d). Assignment of the density as sulfate was confirmed in anomalous difference Fourier maps for the native data, which showed peaks that were similar in size to those of cysteine and methionine sulfur atoms for some of the sulfates. This further suggests that nSH2 binds a phosphorylated ligand and that it might accommodate a pTyr side chain in a suboptimal binding pocket.

Typical SH2 domain ligands bind through a two-prong mechanism that, in addition to binding pTyr, also involves binding of the three side-chain residues C-terminal to the pTyr into the "specificity pocket."[38] Binding partner preference is typically defined in the specificity pocket by the BG and EF loops and the βD3 and βD5 residues, which usually favor binding of hydrophobic residues. In contrast, the Spt6 nSH2 fold predominantly displays charged and polar residues in this site, and there is no BG loop due to the extension of nSH2 αB to the cSH2 fold. Instead, cSH2 residues such as the DE loop and a βD side chain (K1411) protrude into the pocket, forming the top portion of the nSH2 specificity pocket (Fig. 7e). This indicates that, if the nSH2 fold binds substrate in the typical "two-pronged" manner common to SH2 domains, the cSH2 fold would make significant contributions to binding.

In contrast to nSH2, the cSH2 fold appears to be cryptic and unlikely to bind a phosphorylated ligand because residues critical for phosphate binding are substituted to display a very different chemical environment, and the Y1394 side chain fills the space where a phosphate would typically bind (Fig. 7f). Moreover, the region of the specificity pocket lacks even a shallow depression, as it is filled by bulky, aromatic side chains (F1397, Y1406, W1408, and F1434). Thus, in contrast to nSH2 and consistent with the lack of sequence conservation at cSH2 (Fig. 6c), it seems unlikely that cSH2 binds ligands in a manner reminiscent of SH2 domains.

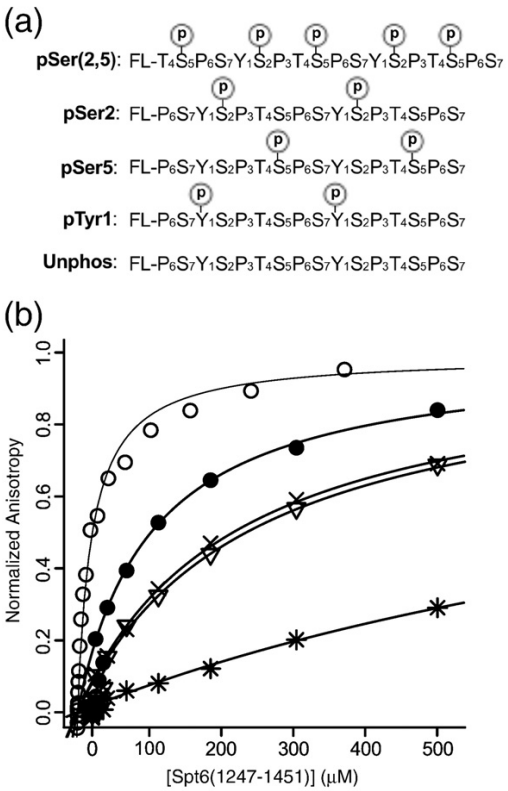
Binding of Spt6 tSH2 domain with phosphorylated peptides

The human Spt6 CTD has been reported to bind the heptad repeat sequences of the mammalian

RNAPII large subunit when the RNAPII CTD is treated with P-TEFb, a kinase that phosphorylates Ser2 during transcriptional elongation.[9] In order to investigate this interaction more quantitatively, we used fluorescence anisotropy (FA) to measure binding of *S. cerevisiae* Spt6(1247–1451) to di-heptad repeat peptides representing various phosphoisoforms of the RNAPII CTD (Fig. 8). Peptides tested include sequences representing phosphoserine (pSer) 2, pSer5, pTyr1, and pSer(2,5). pSer5 and pSer(2,5) peptides were included to test specificity and because these modifications also occur on the RNAPII heptad repeats. A pTyr1 peptide was included because SH2 domains typically bind pTyr peptides and this modification occurs in mammals,[40] although it has not been reported to occur in yeast.

All of the peptides assayed bound with affinities in the range of ~20–250 μM, which is similar to the affinity of other RNAPII CTD interactions with isolated binding domains,[41] but is ~10- to 100-fold weaker than is typically found for the interaction of SH2 domains with pTyr ligands.[42] Of the ligands we assayed, the pSer(2,5) peptide bound Spt6(1247–1451) with the highest affinity (23 μM). This may indicate that RNAPII CTD sequences phosphorylated on both Ser2 and Ser5 are the authentic *in vivo* ligands for the yeast Spt6 tSH2 domain, which would be consistent with the reports that Ser2-phosphorylated sequences are preferred[9] and that localization of Ser2 and Ser2/5 phosphorylation overlaps substantially in average transcription units.[43,44] On the other hand, interpretation is complicated by the fact that our pSer2/5 peptide has considerably more negative charge than the other peptides assayed and thus may be more prone to nonspecific effects. Interestingly, the pTyr1 peptide binds with a K_d of 110 μM, which is tighter than that of the pSer2 (197 μM) and pSer5 (245 μM) peptides bearing an equivalent number of phosphate groups. A higher affinity for pTyr1 peptide with the same overall charge as pSer2 or pSer5 peptide is consistent with the similarity between nSH2 and well-characterized pTyr binding SH2 domains, but the physiological relevance of this result is not clear given the lack of observed pTyr modification of the RNAPII CTD in yeast.

The conclusion that the Spt6 tSH2 domain interacts specifically with phospho-CTD peptides is reinforced by our observations that an unphosphorylated form of the RNAPII CTD showed negligible (K_d of >1000 μM) binding and that similar results were obtained when the assays were performed under a phosphate-buffered saline condition (data not shown). As a further test of specific interactions, we measured binding to a mutant form of Spt6(1247–1451) in which residues R1282 and S1284, which are important for phosphate binding in typical SH2 domains, were both substituted with



(c)

CTD peptide	K_D (μM)† WT	R1282A/S1284A
○ pSer(2,5)	23.0±2.2	90.9±14.2
● pTyr1	110.2±9.1	261.6±32.8
× pSer2	197.1±39.69	228.4±17.8
▽ pSer5	245.4±29.4	348.1±56.7
✳ Unphos	>1000	>1000

†Average K_D ± standard error for at least 3 independent measurements

Fig. 8. Spt6 tSH2 binds RNAPII CTD phosphopeptides. (a) Peptides used in binding studies with positions of pSer and pTyr residues indicated. (b) Representative FA binding isotherms for Spt6(1247–1451) binding to various peptides with symbols defined in (c). (c) Binding affinities for WT and R1282A/S1284A Spt6(1247–1451) proteins based on FA experiments.

alanine. Binding to the pSer(2,5) peptide was decreased by ~4-fold, while binding to the pTyr1, pSer2, and pSer5 peptides was decreased by 2.4-fold, 1.2-fold, and 1.4-fold, respectively. These modest effects are consistent with the putative nSH2 phosphate binding site contributing to the binding interaction but not performing a dominant role for interaction with the peptides assayed.

Deletion of the entire tSH2 region by truncation of the *SPT6* gene leads to defects in growth attributed

to suboptimal transcription elongation.[22,23] To examine the importance of tSH2 residues implicated in pTyr binding *in vivo* more carefully, we mutated the single genomic copy of *SPT6* to produce proteins with R1282H, S1284D, R1286A, Q1303E, EN1313/1314AA, or K1343E mutation (nSH2 domain) or P1390A or K1411E mutation (cSH2 domain). The effects of these mutations were quite mild, failing to recapitulate the severe defects caused by truncation of the gene (data not shown; see Materials and Methods for a list of phenotypes screened). The C-terminal region of Spt6 therefore appears to have some activity that is not interrupted when residues within the tSH2 domain expected to be important for binding phosphorylated substrates are mutated.

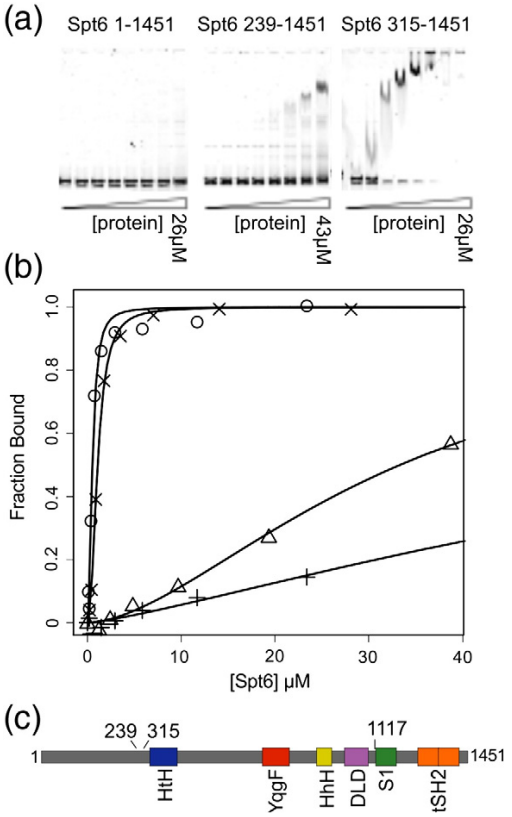


Fig. 9. dsDNA binding studies. (a) Representative gel shift assays for three different constructs of Spt6. (b) Representative binding isotherms used to calculate dissociation constants for various Spt6 constructs binding to the 177-bp Widom 601 dsDNA.[45] Symbols used to indicate isotherms for different constructs are as follows: (+) Spt6 (1–1451), $K_d = 106.7 \pm 38.9$ μM; (Δ) Spt6(239–1451), $K_d = 33.7 \pm 3.8$ μM; (×) Spt6(315–1117), $K_d = 1.08 \pm 0.06$ μM; (○) Spt6(315–1451), $K_d = 0.53 \pm 0.07$ μM. (c) Schematic diagram indicating endpoints for constructs used in DNA binding studies.

Our data are consistent with a recent report that concluded that pSer2 RNAPII CTD peptides bound the tSH2 domain of the *C. glabrata* Spt6 homolog with an affinity of 10 μM.[22] The tighter affinity observed in that study could reflect differences between the proteins but is more likely to be due to the very low (10 mM) concentration of NaCl used in the binding assays. Our data also extend the earlier work by showing that the *S. cerevisiae* Spt6 tSH2 domain displays little discrimination for binding CTD peptides with different modifications but does show a small preference for a peptide with a single phosphorylated tyrosine (pTyr1). Our findings suggest that Spt6 activities *in vivo* may be modulated by phosphorylation of binding partners and that the ligands of the tSH2 domain may include phosphorylated tyrosine residues.

Binding to dsDNA

To test whether Spt6 is capable of binding dsDNA, we performed electrophoretic mobility gel shift assays using a 177-bp dsDNA fragment. We tested several different Spt6 constructs and found that binding was tighter ($K_d \pm$ standard deviation) for Spt6(315–1451) (K_d of 0.53 ± 0.07 μM) than for Spt6 (239–1451) (K_d of 33.7 ± 3.8 μM) or Spt6(1–1451) (K_d of 106.7 ± 38.9 μM) (Fig. 9), demonstrating that the disordered and negatively charged N-terminal residues diminish dsDNA binding. Unlike Tex, which requires the S1 domain for nucleic acid binding,[21] an Spt6 construct lacking the S1 domain, Spt6(315–1117), retains the ability to bind dsDNA with a K_d of 1.08 ± 0.06 μM (Fig. 9b).

Overall structure and functional implications

Our composite model of Spt6 (Fig. 10) features a core region (residues ~298–1117) that has multiple recognizable domains whose packing in the crystal likely reflects, to a large extent, their organization in solution, at least in the absence of binding partners. The N-terminal residues 1–297 display considerable overall negative charge and are expected to be highly mobile, while residues 239–263 also comprise the Spn1/Iws1-binding determinant and overlap with the nucleosome binding site.[14] This high degree of mobility may provide a flexible tether for bridging binding partners, such as Spn1/IwsI, RNAPII, and nucleosomes,[9] and the negative charge may modulate histone and DNA interactions.

Inherent flexibility is also a feature of the C-terminal S1 domain, H44, and tSH2 domain. The S1 domain is loosely associated with the core, lacks density in the Spt6(236–1259) structure, and may be visible in the Spt6(239–1451) structure only because of ordering by a crystal lattice contact. Whereas the core of the distantly related Tex protein clearly resembles that of Spt6, the Tex and Spt6 S1 domains are displaced

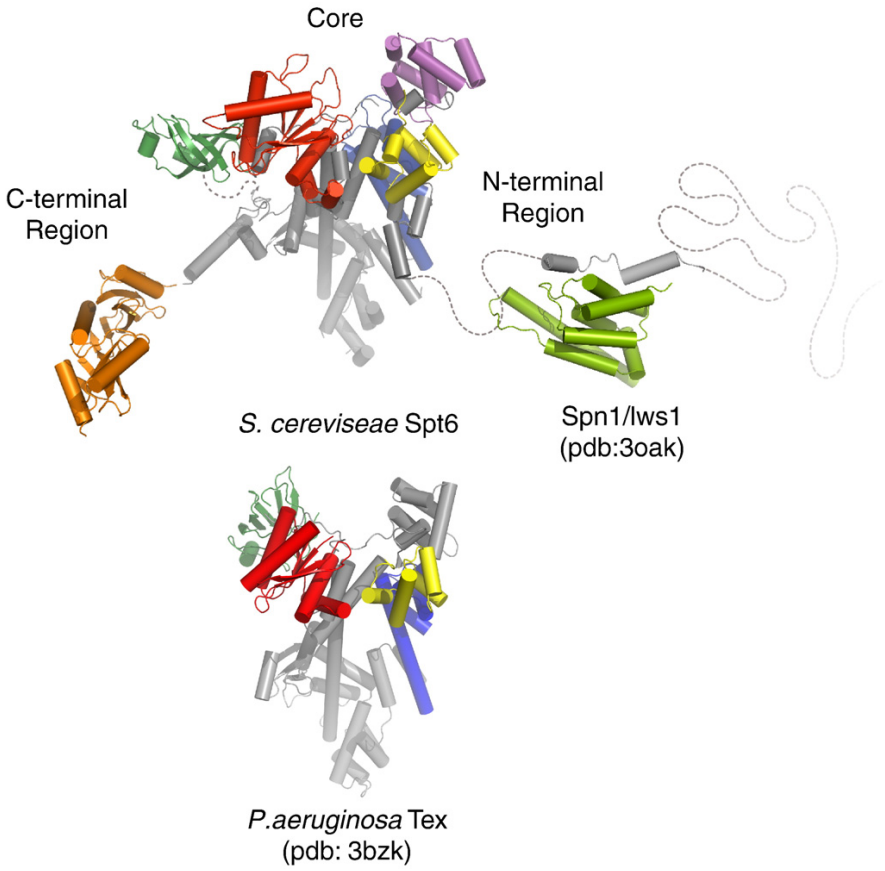


Fig. 10. Comparison of Spt6 and Tex structures. The overall structure of the Spt6 core resembles that of the prokaryotic Tex protein. This similarity implicates the Spt6 core in nucleosome-independent functions, such as transcriptional elongation on naked template DNA. The N-terminal region includes determinants essential for binding Spn1/Iws1 or nucleosomes, which appear to bind competitively with each other.[14] The negatively charged N-terminal region may also be important for modulating binding to nucleic acids. The C-terminal region has been implicated in binding RNAPII.[9]

by a rotation of ~80° and ~25- to 30-Å translation with respect to each other, and the S1 domain appears to be shifted by 14 Å in different crystal forms of Tex.[21] Therefore, it is likely that a highly mobile S1 domain is an important feature of both proteins, although one notable difference is that the Tex S1 binds nucleic acids, whereas our DNA binding data and consideration of surface amino acid residues indicate that the Spt6 S1 domain does not. Although the low conservation of residues in the putative binding cleft of the S1 domain suggests that this region is unlikely to have a highly conserved binding partner, the electrostatic characteristics of this surface (Fig. 5) are consistent with potential binding partners that are positively charged, such as histones.

H44 is visible only in the Spt6(236–1259) structure, where it also appears to be ordered by a lattice contact, and the C-terminal tSH2 domain is expected to be highly dynamic with respect to the rest of the protein. The leading model is that this domain binds RNAPII that is phosphorylated on Ser2 of its CTD.[9] Our binding data are consistent with this view, provided that other determinants contribute to binding/specificity, but also indicate the possibility that other binding partners might be functionally important. For example, Spt5 (discussed below) has been shown to co-localize with Spt6 and contains a phosphorylated C-terminal repeat domain similar to the RNAPII CTD,[46] which could be a physiological ligand for the Spt6 tSH2 domain.

Spt6 has been shown to be functionally associated with a large number of proteins involved in transcription elongation, chromatin maintenance, and RNA processing. The structure of Spt6 presented here will serve as a foundation for a more precise mapping of protein binding partners. Along with protein binding partners, Spt6 is also expected

to interact with nucleic acids in the transcription complex. DNA binding is probably important for nucleosome reassembly and potentially for transcription elongation. Association with nascent RNA transcripts could be important for enhancing the elongation rate or for organizing interactions with RNA modification/export factors such as REF/Aly. Future studies will be needed to see if Spt6 binds RNA and to further map nucleic acid binding to different domains of Spt6. For example, the YqgF or the $(HhH)_2$ domain may mediate binding to specific DNA structures such as four-way (Holliday) junction DNA, structures similar to those found at the DNA entry/exit points of nucleosomes.[47] The tSH2 domain may also contribute to nucleic acid interactions, as it is likely to bind negatively charged substrates containing phosphate groups. A simple electrostatic surface analysis indicates that each of these domains retains properties found in homologous domains with known functions. Furthermore, examination of the histone-binding activity of the various Spt6 domains will be of significant interest in furthering our understanding of nucleosome assembly/reassembly.

The structural similarity between Spt6 and the prokaryotic Tex protein is limited to the core and S1 domains. Consistent with the extent of structural similarity, N- and C-terminal regions that are unique to Spt6 are required for eukaryotic-specific interactions with nucleosomes,[11,13,14] hyperphosphorylated forms of the RNAPII CTD,[9] Spn1/IwsI,[14] and mRNA processing/export factors.[9] An attractive possibility is that the core region provides activities that are conserved among prokaryotes and eukaryotes. In this regard, it is striking that most of the Spt6 core domains belong to structural families whose members function in nucleic acid binding, an activity that is likely to be a key component of transcription factors such as Spt6 and Tex that are capable of stimulating elongation on nucleosome-free DNA templates.[9,16] Consistent with this possibility, our data indicate that the Spt6 core can bind dsDNA. Curiously, some of the putative nucleic-acid-binding surfaces of Spt6 domains are occluded in the structure, although conformational changes might displace residues 298–320 (H1 and H2) to expose a DNA-binding activity on the $(HhH)_2$ domain. Consistent with this model, we find that truncation of ~314 N-terminal residues leads to tighter dsDNA binding. One attractive model is that conformational changes of this nature are induced by binding partners such as histones or Spn1/Iws1. The relationship between Spt6 and Tex proposed here is reminiscent of that between another eukaryotic transcription elongation factor, Spt5, and its bacterial counterpart, NusG. These proteins also display similar core domains, while Spt5 has an acidic N-terminal extension and a C-terminal extension[48] that confers eukaryote-specific functions such as binding to RNAPII,[48] interaction with mRNA capping enzymes,[49] and extensive phosphorylation by RNAPII CTD kinases such as P-TEFb.[50] Therefore, like Spt5, Spt6 is likely to have built on the fundamental transcription activities of its core to accommodate the additional complexities of eukaryotic gene regulation.

Materials and Methods

Protein expression and purification

The protein constructs were expressed from pET151-D/TOPO vectors (Invitrogen) in BL21 codon plus (RIL) *E. coli* cells (Stratagene). Cultures were grown in autoinduction media[51] in baffled 1.8-l flasks at 37 °C with continuous shaking. After 4–8 h, the cultures were shifted to 23 °C and grown for an additional 16–24 h. Harvested cells were stored at −80 °C. Cells were thawed and lysed in buffer containing lysozyme and protease inhibitors, followed by sonication and centrifugation (25,000–30,000*g*). The soluble fraction was applied to nickel agarose resin (Qiagen) and eluted in buffer containing 300 mM imidazole and 100 mM NaCl, immediately followed by application to a heparin column (5-ml HiTrap Heparin; GE Healthcare Life Sciences) and elution over a NaCl gradient. Fractions containing Spt6 were pooled and processed overnight at room temperature in buffer containing tobacco etch virus protease. A nickel agarose column was used to remove the tagged tobacco etch virus protease and unprocessed Spt6 protein, and the flow-through was concentrated and loaded onto a size-exclusion column [Superdex 200 or S75 (for 1247–1451 constructs); GE Healthcare Life Sciences]. SeMet protein was expressed using an autoinduction protocol[51] for selenomethionine incorporation and purified by the same protocol as native protein. All crystals were grown in sitting drops, transferred briefly to a cryoprotection solution, suspended in a nylon loop, and plunged into liquid nitrogen (Table 1).

Crystal structure determinations and refinements

Data were processed with HKL2000.[52] Spt6(236–1259) was determined by the multiple-wavelength anomalous diffraction method. SOLVE[53] was used to locate selenium atoms, and RESOLVE[54] was used for density modification and preliminary model building. AutoSol in PHENIX[26] was used to determine the Spt6(1247–1451) structure by the single-wavelength anomalous diffraction method. The Spt6 (239–1451) structure was determined by molecular replacement using AutoMR in PHENIX[26] to a resolution of 3.3 Å. A homology model built by TASSER[55] of the Spt6 S1 domain was used as a guide for model building. The native Spt6 (1247–1451) structure was determined by molecular replacement using Phaser.[56] PHENIX[26] and TLSMD[57] were used for refinement, Coot[58] was used for model building, and MolProbity[27] was used for structure validation. The following residues were ordered and included in the refined models: Spt6(236–1259) 297–455, 464–484, 501–561, 567–1002, 1009–1128, and 1219–1248; Spt6(239–1451) 312–455, 464–489, 509–552, 567–649, 653–1001, and 1014–1210; and

Spt6(1247–1451) 1247–1440. Structural alignments were performed using Dali[59] and SSM.[60] PyMOL[61] was used to create the figures. Electrostatic surface representations were calculated using PDB2PQR and APBS tools[62,63] using the AMBER force field and colored from red (−5 kT/e) to blue (+5 kT/e).

DNA binding experiments

Widom 601 DNA (177 bp) with a 5′ Cy3 fluorophore was generated as described[45] followed by precipitation, gel purification, and electroelution. Electrophoretic mobility shift binding experiments were performed, and K_d values were determined as described previously.[21] In short, 2-fold serial dilutions of the respective purified protein construct were mixed with nucleic acid substrate at room temperature in binding buffer [15 mM Tris–HCl (pH 7.5), 100 mM NaCl, 10% glycerol, and 0.5 mM ethylenediaminetetraacetic acid] where the final concentration of dsDNA was 10- to 20-fold below the estimated K_d for the interaction. After incubation for 30 min, samples were run on 4–20% TBE native gels (Bio-Rad Laboratories) and imaged and quantified using a TYPHOON imaging system with ImageQuant software (GE Healthcare Life Sciences). The fraction bound was calculated by quantifying the DNA_{total} (total fluorescence in entire lane) and DNA_{free}. DNA of slower mobility than the DNA_{free} was considered bound. The fraction bound = $1-([DNA]_{free}/[DNA]_{total})$. Dissociation constants (K_d values) were calculated by plotting data points and curve fitting in the program R[64] using the Hill formalism where fraction bound = $1/(1+(K_d n/[P]n))$. In all cases, standard deviations are calculated from at least three measurements, except for the Spt6(315–1117) construct, which was repeated twice.

FA binding experiments

Peptides were synthesized by the University of Utah Core Facility or purchased commercially through AnaSpec Inc. (San Jose, CA), purified to >98% purity by HPLC, and confirmed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Purified Spt6(1247–1451) was titrated in 1.5- to 2.0-fold serial dilutions against a constant concentration of fluorescein-labeled peptide (10- to 20-fold below estimated K_d) in 20 mM Tris–Cl (pH 7.5), 100 mM NaCl, and 5% glycerol. Samples were incubated at room temperature for at least 15 min prior to reading. Parallel and perpendicular fluorescence intensity was measured in a multi-well format using a Tecan Infinite 200 microplate reader using excitation/emission wavelengths of 485 nm/535 nm. Anisotropy values were calculated, normalized, and plotted as a function of protein concentration. K_d values were determined by fitting the data using the equation[65] $A=((A_T\times([pro]/K_d))/(1+[pro]/K_d))$, where A is the measured anisotropy, A_T is the total change in anisotropy, and [pro] is the protein concentration.

Genetic analysis of tSH2 mutations

The following alleles of *SPT6* were screened for phenotypes in strains isogenic with the A364a genetic background: wild type (WT), *spt6-R1282H*, *spt6-S1284D*, *spt6-R1286A*, *spt6-Q1303E*, *spt6-E1313A*, *N1314A*, *spt6-K1343E*, *spt6-P1390A*, and *spt6-K1411E*. Mutations were introduced into the genomic copy of *SPT6* such that expression was from the native promoter at the normal locus except for the introduction of a *URA3* or *TRP1* marker downstream of the open reading frame. Strains were tested for growth on rich medium at 30 °C and 38 °C; on medium lacking lysine at 30 °C and 37 °C (all strains had the *lys2-128∂* allele; thus, growth would reveal an Spt^- phenotype); and on media containing 150 mM hydroxyurea, 75 μg/ml 6-azauracil, 0.6 μg/ml 4-nitroquinolone, 10 mM caffeine, 3% formamide, 1.2 M NaCl, 45 μg/ml mycophenolic acid, or 6% ethanol (all at 30 °C). None of the mutants were sensitive to any of the stress conditions relative to the WT strain. *spt6-S1284D* and *spt6-Q1303E* strains were somewhat more resistant to 3% formamide than the WT, and *spt6-Q1303E* strains displayed a very weak Spt^- phenotype (faint growth after 7 days). Mutants were not tested for a defect in cryptic initiation.

PDB accession numbers

Coordinates and structure factors for Spt6(236–1259), Spt6(239–1451), and SeMet Spt6(1247–1451) and native Spt6(1247–1451) have been deposited in the PDB with accession numbers 3psf, 3psi, 3psj, and 3psk, respectively.

Acknowledgements

We thank Hua Xin and Charisse Kettelkamp for technical assistance and Heidi Schubert for advice with the crystallographic analysis. Portions of this work were performed in Core Facilities at the University of Utah, which were supported by P30CA042014 from the National Cancer Institute. Some of the X-ray diffraction data for this study were measured at the National Synchrotron Light Source (NSLS). Financial support for NSLS comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the U.S. Department of Energy and from the National Center for Research Resources of the National Institutes of Health (NIH). Portions of this research were performed at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and the NIH, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. S.J.J. was supported by a postdoctoral fellowship (GM074368). This work was supported by NIH grant RO1 GM076242.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2011.03.002

References

1. Li, B., Carey, M. & Workman, J. L. (2007). The role of chromatin during transcription. *Cell*, **128**, 707–719.
2. Luna, R., Gaillard, H., Gonzalez-Aguilera, C. & Aguilera, A. (2008). Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. *Chromosoma*, **117**, 319–331.
3. Clark-Adams, C. D. & Winston, F. (1987). The *SPT6* gene is essential for growth and is required for delta-mediated transcription in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **7**, 679–686.
4. Kok, F. O., Oster, E., Mentzer, L., Hsieh, J. C., Henry, C. A. & Sirotkin, H. I. (2007). The role of the SPT6 chromatin remodeling factor in zebrafish embryogenesis. *Dev. Biol.* **307**, 214–226.
5. Ardehali, M. B., Yao, J., Adelman, K., Fuda, N. J., Petesch, S. J., Webb, W. W. & Lis, J. T. (2009). Spt6 enhances the elongation rate of RNA polymerase II *in vivo*. *EMBO J.* **28**, 1067–1077.
6. Nishiwaki, K., Sano, T. & Miwa, J. (1993). *emb-5*, a gene required for the correct timing of gut precursor cell division during gastrulation in *Caenorhabditis elegans*, encodes a protein similar to the yeast nuclear protein SPT6. *Mol. Gen. Genet.* **239**, 313–322.
7. Shen, X., Xi, G., Radhakrishnan, Y. & Clemmons, D. R. (2009). Identification of novel SHPS-1-associated proteins and their roles in regulation of insulin-like growth factor-dependent responses in vascular smooth muscle cells. *Mol. Cell. Proteomics*, **8**, 1539–1551.
8. Baniahmad, C., Nawaz, Z., Baniahmad, A., Gleeson, M. A., Tsai, M. J. & O'Malley, B. W. (1995). Enhancement of human estrogen receptor activity by SPT6: a potential coactivator. *Mol. Endocrinol.* **9**, 34–43.
9. Yoh, S. M., Cho, H., Pickle, L., Evans, R. M. & Jones, K. A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev.* **21**, 160–174.
10. Vanti, M., Gallastegui, E., Respaldiza, I., Rodriguez-Gil, A., Gomez-Herreros, F., Jimeno-Gonzalez, S. *et al.* (2009). Yeast genetic analysis reveals the involvement of chromatin reassembly factors in repressing HIV-1 basal transcription. *PLoS Genet.* **5**, e1000339.
11. Adkins, M. W. & Tyler, J. K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Mol. Cell*, **21**, 405–416.
12. Kaplan, C. D., Laprade, L. & Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
13. Bortvin, A. & Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science*, **272**, 1473–1476.
14. McDonald, S. M., Close, D., Xin, H., Formosa, T. & Hill, C. P. (2010). Structure and biological importance of the Spn1–Spt6 interaction, and its regulatory role in nucleosome binding. *Mol. Cell*, **40**, 725–735.
15. Yoh, S. M., Lucas, J. S. & Jones, K. A. (2008). The Iws1: Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev.* **22**, 3422–3434.
16. Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D. K., Aida, M. *et al.* (2004). Human Spt6 stimulates transcription elongation by RNA polymerase II *in vitro*. *Mol. Cell. Biol.* **24**, 3324–3336.
17. Andrulis, E. D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P. & Lis, J. T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*. *Nature*, **420**, 837–841.
18. Kaplan, C. D., Holland, M. J. & 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.
19. Pawson, T. (2004). Specificity in signal transduction: from phosphotyrosine–SH2 domain interactions to complex cellular systems. *Cell*, **116**, 191–203.
20. Maclennan, A. J. & Shaw, G. (1993). A yeast SH2 domain. *Trends Biochem. Sci.* **18**, 464–465.
21. Johnson, S. J., Close, D., Robinson, H., Vallet-Gely, I., Dove, S. L. & Hill, C. P. (2008). Crystal structure and RNA binding of the Tex protein from *Pseudomonas aeruginosa*. *J. Mol. Biol.* **377**, 1460–1473.
22. Sun, M., Lariviere, L., Dengl, S., Mayer, A. & Cramer, P. (2010). A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II CTD. *J. Biol. Chem.* **285**, 41597–41603.
23. Diebold, M. L., Loeliger, E., Koch, M., Winston, F., Cavarelli, J. & Romier, C. (2010). A non-canonical tandem SH2 enables interaction of elongation factor SPT6 with RNA polymerase II. *J. Biol. Chem.* **285**, 38389–38398.
24. Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217.
25. Ponting, C. P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic Acids Res.* **30**, 3643–3652.
26. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J. *et al.* (2008). Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426**, 419–435.
27. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
28. Dengl, S., Mayer, A., Sun, M. & Cramer, P. (2009). Structure and *in vivo* requirement of the yeast Spt6 SH2 domain. *J. Mol. Biol.* **389**, 211–225.
29. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635–645.
30. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). The many faces of the helix–turn–helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231–262.
31. Grant, R. P., Marshall, N. J., Yang, J. C., Fasken, M. B., Kelly, S. M., Harreman, M. T. *et al.* (2008). Structure of the N-terminal Mlp1-binding domain of the

Saccharomyces cerevisiae mRNA-binding protein, Nab2. *J. Mol. Biol.* **376**, 1048–1059.
32. Saito, A., Iwasaki, H., Ariyoshi, M., Morikawa, K. & Shinagawa, H. (1995). Identification of four acidic amino acids that constitute the catalytic center of the RuvC Holliday junction resolvase. *Proc. Natl Acad. Sci. USA*, **92**, 7470–7474.
33. Shao, X. & Grishin, N. V. (2000). Common fold in helix–hairpin–helix proteins. *Nucleic Acids Res.* **28**, 2643–2650.
34. Qiao, F. & Bowie, J. U. (2005). The many faces of SAM. *Sci. STKE*, **2005**, re7.
35. Park, H. H., Lo, Y. C., Lin, S. C., Wang, L., Yang, J. K. & Wu, H. (2007). The death domain superfamily in intracellular signaling of apoptosis and inflammation. *Annu. Rev. Immunol.* **25**, 561–586.
36. Theobald, D. L., Mitton-Fry, R. M. & Wuttke, D. S. (2003). Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 115–133.
37. Yu, E. Y., Wang, F., Lei, M. & Lue, N. F. (2008). A proposed OB-fold with a protein-interaction surface in *Candida albicans* telomerase protein Est3. *Nat. Struct. Mol. Biol.* **15**, 985–989.
38. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. & Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell*, **72**, 779–790.
39. Lee, C. H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E. & Kuriyan, J. (1994). Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure*, **2**, 423–438.
40. Duyster, J., Baskaran, R. & Wang, J. Y. (1995). Src homology 2 domain as a specificity determinant in the c-Abl-mediated tyrosine phosphorylation of the RNA polymerase II carboxyl-terminal repeated domain. *Proc. Natl Acad. Sci. USA*, **92**, 1555–1559.
41. Lunde, B. M., Reichow, S. L., Kim, M., Suh, H., Leeper, T. C., Yang, F. *et al.* (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* **17**, 1195–1201.
42. Ladbury, J. E., Lemmon, M. A., Zhou, M., Green, J., Botfield, M. C. & Schlessinger, J. (1995). Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. *Proc. Natl Acad. Sci. USA*, **92**, 3199–3203.
43. 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.
44. Phatnani, H. P. & Greenleaf, A. L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936.
45. Lowary, P. T. & Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42.
46. Liu, Y., Warfield, L., Zhang, C., Luo, J., Allen, J., Lang, W. H. *et al.* (2009). Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Mol. Cell. Biol.* **29**, 4852–4863.
47. Zlatanova, J. & van Holde, K. (1998). Binding to four-way junction DNA: a common property of architectural proteins? *FASEB J.* **12**, 421–431.
48. Guo, M., Xu, F., Yamada, J., Egelhofer, T., Gao, Y., Hartzog, G. A. *et al.* (2008). Core structure of the yeast spt4–spt5 complex: a conserved module for regulation of transcription elongation. *Structure*, **16**, 1649–1658.
49. Pei, Y. & Shuman, S. (2002). Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *J. Biol. Chem.* **277**, 19639–19648.
50. Ivanov, D., Kwak, Y. T., Guo, J. & Gaynor, R. B. (2000). Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol. Cell. Biol.* **20**, 2970–2983.
51. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* **41**, 207–234.
52. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In (Simon, M. I., Abelson, J. N., Carter, C. W. & Sweet, R. M., eds), Academic Press, New York, NY.
53. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **55**, 849–861.
54. Terwilliger, T. C. (2003). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **59**, 38–44.
55. Roy, A., Kucukural, A. & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738.
56. McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **63**, 32–41.
57. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **62**, 439–450.
58. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **60**, 2126–2132.
59. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
60. Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D. Biol. Crystallogr.* **60**, 2256–2268.
61. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* Delano Scientific, Palo Alto, CA.
62. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667.
63. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
64. Team, R. D. C.. (2010). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.
65. LiCata, V. J. & Wowor, A. J. (2008). Applications of fluorescence anisotropy to the study of protein–DNA interactions. *Methods Cell Biol.* **84**, 243–262.

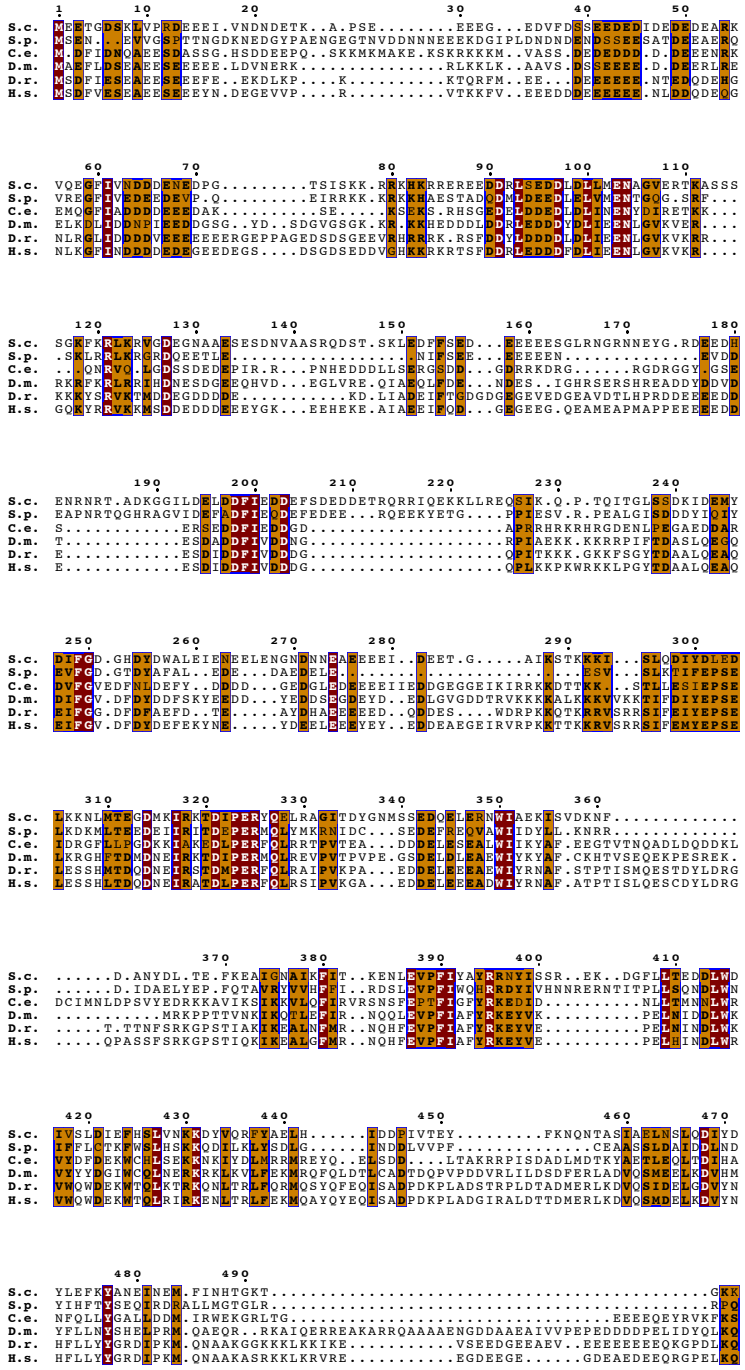
Supplemental Figure Legend

Figure S1. Spt6 Conservation.

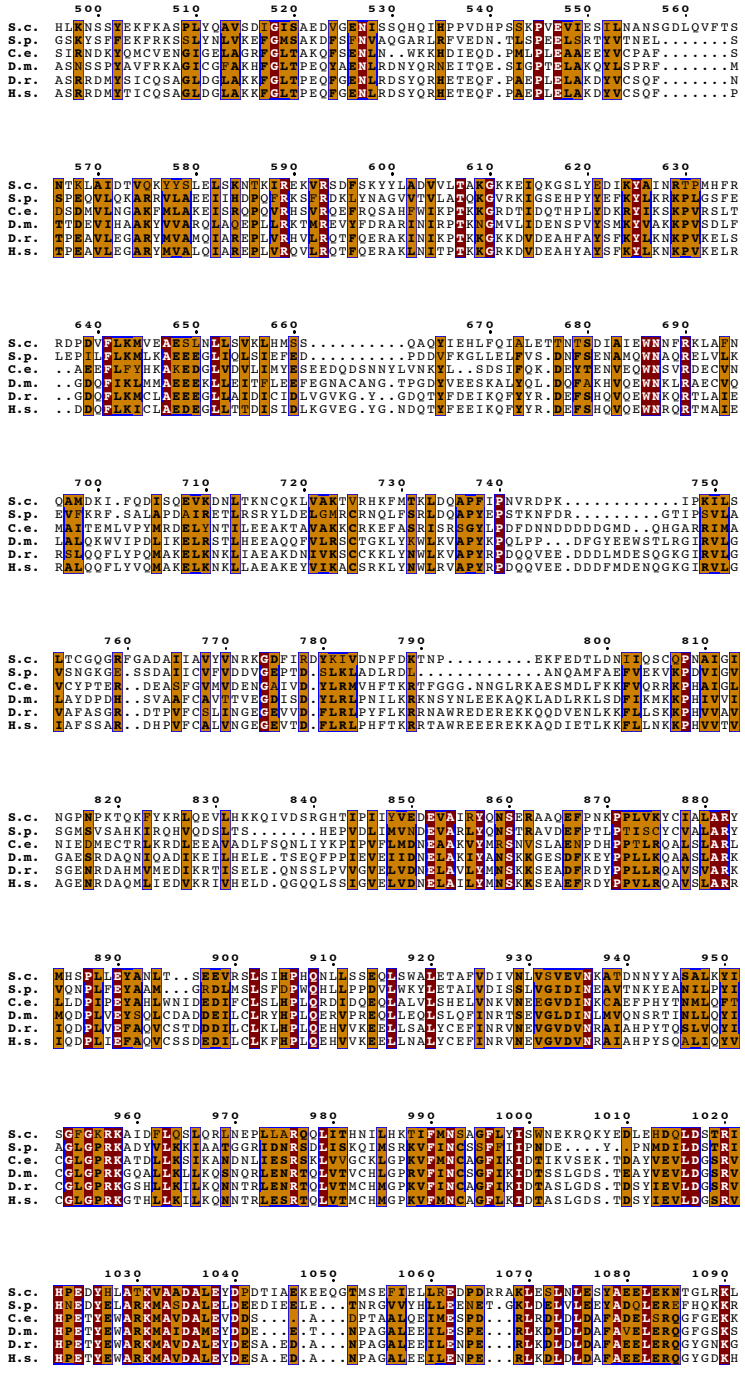
Spt6 protein from various eukaryotic species were aligned using T-Coffee[1] and ESPript[2] was used for visualization. Coloring of sequence represents degree of conservation, dark red background (invariant), orange font (conserved), in an alignment of proteins from *Saccharomyces cerevisiae* (S.c.), *Schizosaccharomyces pombe* (S.p.), *Caenorhabditis elegans* (C.e.), *Drosophila melanogaster* (D.m.), *Danio rerio* (D.r.), and *Homo sapiens* (H.s.).

1. Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-17.
2. Gouet, P., Robert, X. & Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research* **31**, 3320-3323.

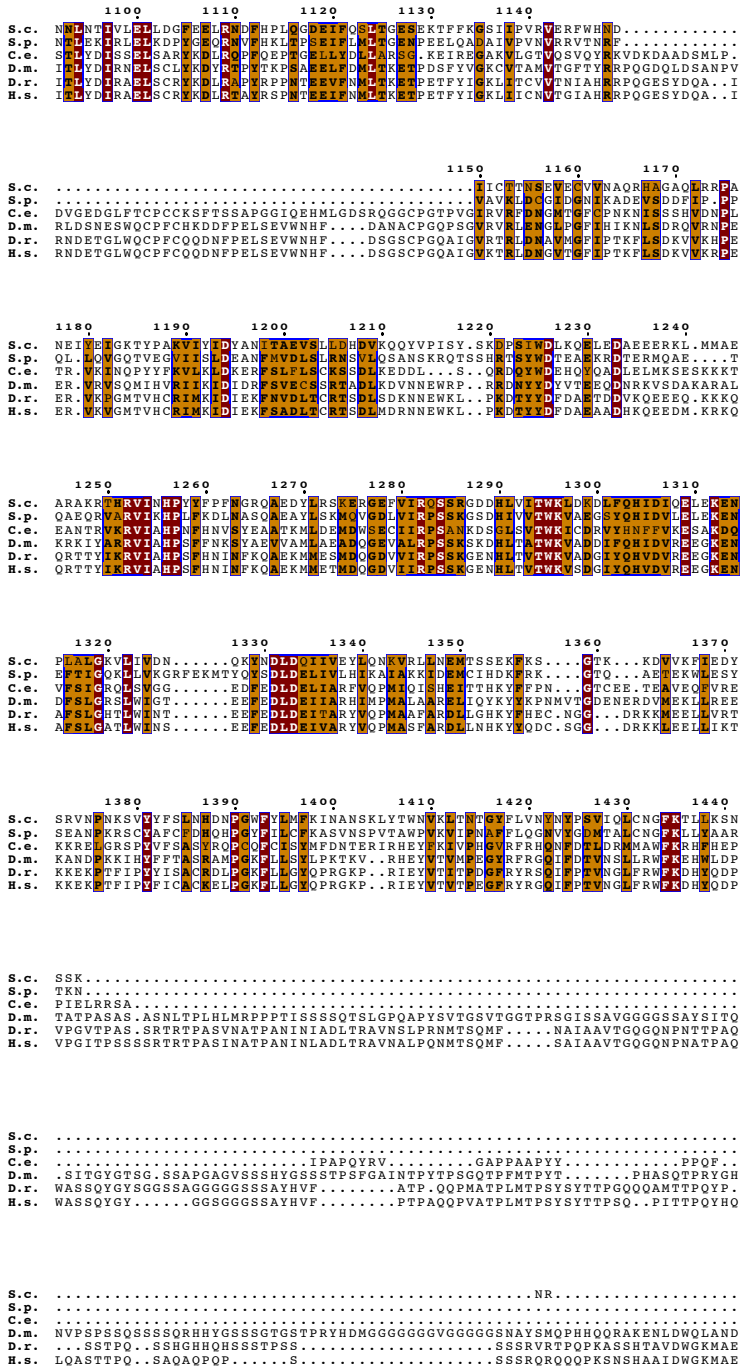
Supplemental Figure 1 Spt6 Crystal Structures Close, D et al



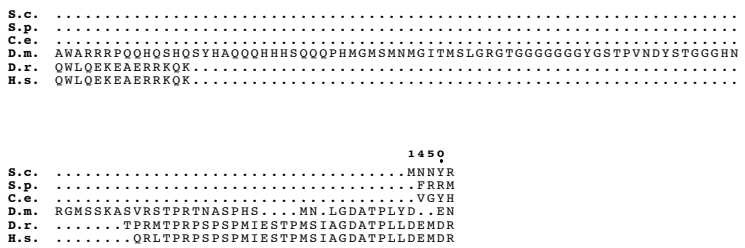
Supplemental Figure 1 Spt6 Crystal Structures Close, D et al



Supplemental Figure 1 Spt6 Crystal Structures Close, D et al



Supplemental Figure 1 Spt6 Crystal Structures Close, D et al



CHAPTER 3

STRUCTURE AND BIOLOGICAL IMPORTANCE OF THE SPN1-SPT6 INTERACTION, AND ITS REGULATORY ROLE IN NUCLEOSOME BINDING

Seth M. McDonald*, Devin Close*, Hua Xin,Tim Formosa,

and Christopher P. Hill (2010)

Molecular Cell

40, 725-735

*Authors contributed equally

Molecular Cell
Article

Structure and Biological Importance of the Spn1-Spt6 Interaction, and Its Regulatory Role in Nucleosome Binding

Seth M. McDonald,[1,2] Devin Close,[1,2,3] Hua Xin,[1] Tim Formosa,[1,*] and Christopher P. Hill[1,*]
[1]Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112-5650, USA
[2]These authors contributed equally to this work
[3]Present address: Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
*Correspondence: tim@biochem.utah.edu (T.F.), chris@biochem.utah.edu (C.P.H.)
DOI 10.1016/j.molcel.2010.11.014

SUMMARY

Eukaryotic transcription and mRNA processing depend upon the coordinated interactions of many proteins, including Spn1 and Spt6, which are conserved across eukaryotes, are essential for viability, and associate with each other in some of their biologically important contexts. Here we report crystal structures of the Spn1 core alone and in complex with the binding determinant of Spt6. Mutating interface residues greatly diminishes binding in vitro and causes strong phenotypes in vivo, including a defect in maintaining repressive chromatin. Overexpression of Spn1 partially suppresses the defects caused by an *spt6* mutation affecting the Spn1 interface, indicating that the Spn1-Spt6 interaction is important for managing chromatin. Spt6 binds nucleosomes directly in vitro, and this interaction is blocked by Spn1, providing further mechanistic insight into the function of the interaction. These data thereby reveal the structural and biochemical bases of molecular interactions that function in the maintenance of chromatin structure.

INTRODUCTION

Spn1 and Spt6 are transcription factors that interact with one another and are each essential for viability in yeast (Clark-Adams and Winston, 1987; Fischbeck et al., 2002). *S. cerevisiae* Spn1 is a 410 residue, 46 kDa protein with a central core domain (residues 140–300) that is flanked on both sides by regions that are predicted to be disordered (Ward et al., 2004). Spt6 is a 1451 residue, 168 kDa protein whose core (residues 300–1250) likely resembles the structure of the bacterial Tex protein (Johnson et al., 2008) with an acidic N-terminal extension that is expected to be unstructured (Ward et al., 2004) and a C-terminal domain (CTD) that adopts an SH2 fold (Dengl et al., 2009; Maclennan and Shaw, 1993). Spn1 and Spt6 interact stably with one another, and they and their interaction have been implicated in several aspects of gene expression (Krogan et al., 2002; Lindstrom et al., 2003; Yoh et al., 2007; Yoh et al., 2008).

Spt6 was originally identified in a screen for factors that alter normal initiation of transcription (Clark-Adams and Winston, 1987; Denis, 1984; Neigeborn et al., 1987; Simchen et al., 1984). Subsequently, Spt6 was implicated in a variety of biological processes in organisms ranging from yeasts to humans, including embryogenesis in Zebrafish (Keegan et al., 2002; Kok et al., 2007), multiple stages of development in *Drosophila* (Ardehali et al., 2009), gut morphogenesis in *C. elegans* (Nishiwaki et al., 1993), signal transduction in mammals (Baniahmad et al., 1995; Shen et al., 2009), and HIV transcription regulation and mRNA processing in human cells (Vanti et al., 2009; Yoh et al., 2007). The role in transcription initiation has been ascribed to the ability of Spt6 to chaperone histones to promote reassembly of nucleosomes in the wake of RNA polymerase II (RNAPII), thereby reestablishing the default repressive chromatin state that prevents inappropriate initiation of transcription (Adkins and Tyler, 2006; Bortvin and Winston, 1996; Cheung et al., 2008; Kaplan et al., 2003). While of profound importance, maintaining repressive chromatin appears to be just one of Spt6's roles. For example, Spt6 also promotes elongation by RNAPII (Hartzog et al., 1998; Kaplan et al., 2005; Kaplan et al., 2000; Lindstrom et al., 2003) on nucleosome-free DNA templates in vitro (Endoh et al., 2004; Hartzog et al., 1998; Keegan et al., 2002; Yoh et al., 2007), as well as on chromatin templates in vivo (Ardehali et al., 2009). Together, these data indicate that Spt6 plays a number of mechanistically distinct roles during transcription.

The *SPN1* gene was originally identified as a key regulator of transcription from genes that are regulated postrecruitment of RNAPII (Fischbeck et al., 2002). Spn1 was also identified as a protein that interacts with Spt6 and has been reported to bind with Spt6 in some but not all of Spt6's functional states (Lindstrom et al., 2003; Yoh et al., 2007; Zhang et al., 2008). For example, Spt6 can be coimmunopurified with three distinct Spt4/5-RNAPII complexes, whereas Spn1 is found in only two of these complexes (Lindstrom et al., 2003). The *CYC1* gene of *S. cerevisiae* provides an example of how the Spn1-Spt6 interaction contributes to postrecruitment regulation (Zhang et al., 2008). RNAPII is constitutively bound to the *CYC1* promoter, but is kept from elongating because it interacts with Spn1, which in turn inhibits the Swi/Snf nucleosome remodeling

complex from promoting transcription. During activation, Spt6 binds to Spn1, and repression of Swi/Snf recruitment is relieved.

Spn1 is also needed to achieve normal recruitment of the histone methyltransferase HYPB/Setd2 (Yoh et al., 2008) and the elongation factor TFIIS (Ling et al., 2006; Zhang et al., 2008) to RNAPII complexes traversing active genes. HYPB/Setd2 methylates histone H3K36, which in turn recruits Rpd3-type histone deacetylases to restore chromatin to the repressive hypoacetylated state and block inappropriate transcription (Yoh et al., 2008). In contrast to their antagonistic relationship in activating postrecruitment initiation, Spn1 and Spt6 each contribute toward restoration of repressive chromatin. Human Spn1/IWS1 also binds the protein arginine methyltransferase PRMT5, which methylates the elongation factor Spt5 and thereby regulates its interaction with RNAPII (Liu et al., 2007). Spn1 can additionally function through interactions with pathway-specific regulatory factors, such as the *Arabidopsis* steroid hormone responsive transcription factor BES1, which recruits Spn1 to the promoter and transcribed regions of activated genes (Li et al., 2010). Spn1 therefore contributes in several ways to the appropriate functioning of RNAPII.

In addition to their roles in regulating transcription, Spt6 and Spn1 also collaborate to promote mRNA processing and export. Spt6 is required for proper 3′ end formation by preventing premature 3′ processing at upstream polyadenylation signals (Bucheli and Buratowski, 2005; Kaplan et al., 2005). Further, mammalian Spt6 can bind the Ser2-phosphorylated RNAPII CTD, enhancing recruitment of RNA processing/export factors (Yoh et al., 2007, 2008), and *Drosophila* Spt6 copurifies with the RNA processing exosome complex (Andrulis et al., 2002). Both *SPN1* and *SPT6* have also been implicated in mRNA splicing in *S. cerevisiae* (Burckin et al., 2005), and binding of human Spn1/IWS1 to the RNA export factor REF1/Aly is important for recruitment of REF1/Aly to the body of the *c-Myc* gene during transcription (Yoh et al., 2007).

Spt6 and Spn1 and their interaction with one another therefore play pivotal roles in defining the composition of RNAPII elongation complexes, maintaining the structure of chromatin, and modulating the production of mature mRNA transcripts. To advance mechanistic understanding of their functions, we have determined the structural basis of the Spn1-Spt6 interaction. We also demonstrate the importance of this interface in vitro and in vivo, and show that Spn1 negatively regulates binding of Spt6 to nucleosomes.

RESULTS

Mapping of the Spn1-Spt6 Interface

Full-length *S. cerevisiae* Spt6 and Spn1 proteins were poorly behaved, but deletion of much of their presumably unstructured N-terminal regions (Ward et al., 2004) allowed us to observe coelution of a complex of recombinant Spn1(120–410) and Spt6(206–1451) by size-exclusion chromatography in sodium chloride concentrations up to 300 mM (data not shown). Spt6 (239–1451) also bound Spn1, whereas Spt6(315–1451) did not. Further truncations revealed that Spt6(239–268) is sufficient for Spn1 binding (Figures 1A and 1B). This 30 residue segment of Spt6 is predicted to be unstructured, and comes from a region that is N-terminal to the region expected to resemble the structure of the bacterial protein Tex (Johnson et al., 2008). Spn1 (148–293) includes most of the Spn1 residues that are predicted to be structured and retained the ability to bind Spt6. The slightly larger Spn1(141–305) fragment was previously shown to complement a deletion of *SPN1* (Fischbeck et al., 2002), indicating that this core domain provides the major function(s) of Spn1 in vivo. Isothermal titration calorimetry (ITC) was used to measure binding affinities of Spn1(148–293) for two different Spt6 constructs: Spt6(239–1117), the largest Spt6 construct that remained soluble at sufficient concentrations for these experiments, and Spt6(239–268), the smallest construct tried that retained full binding affinity. In both cases the binding displayed 1:1 stoichiometry and the mean binding constant (K_D) was 170 nM (Figures 1C and 1D and see Table S1 available online). This indicates that the 30 residue segment of Spt6, Spt6 (239–268), is sufficient to recapitulate the binding energy observed for larger Spt6 constructs.

Crystal Structures of the Spn1 Core

The crystal structure of Spn1(148–307) was determined by the single-wavelength anomalous diffraction method using data collected to 3.0 Å resolution from a selenomethionine-substituted crystal (Figure 2A, Table 1). This unrefined model was used in molecular replacement calculations with 2.15 Å data from a native crystal that belonged to a different space group, and the native structure was refined to R_{work}/R_{free} values of 18.5%/22.4%. Residues 148–295 were clearly observed in the electron density, as were four nonnative N-terminal residues that remained after TEV digestion. The 12 C-terminal residues were disordered and are not included in the final model.

Spn1(148–307) forms a right-handed superhelical bundle of eight helices (named H1–8, Figure 2A). Surprisingly, this structure resembles the domains of the RNA processing factors Pcf11 (*S. cerevisiae*) (Meinhart and Cramer, 2004) and SCAF8 (human) (Becker et al., 2008) that bind the RNAPII CTD (Figure S1A). Spn1 is reported to associate with RNAPII (Zhang et al., 2008), and the structural similarity suggested that Spn1 might bind the RNAPII CTD. We have not, however, observed binding in a fluorescence polarization assay between the Spn1 core and synthetic peptides, either phosphorylated or unphosphoryalated, that span more than two heptad repeats of the RNAPII CTD (data not shown).

Crystal Structure of an Spn1-Spt6 Complex

The structure of an Spn1(148–293):Spt6(239–268) complex was determined by molecular replacement and refined to R_{work}/R_{free} values of 18.6%/24.4% against 2.15 Å resolution data (Figure 2B, Figure S2, Table 1). There are two complexes in the asymmetric unit that superimpose closely with an rmsd of ~0.5 Å over all 170 native ordered Cα atoms. Residues 148–292 of Spn1 are clearly visible in the electron density, as are six nonnative N-terminal residues that remain after TEV digestion. Residues 239–263 of Spt6 are also clearly defined in the electron density as well as four nonnative N-terminal residues. There is no clear electron density for the five C-terminal residues of Spt6 and the one C-terminal residue of Spn1, and these are

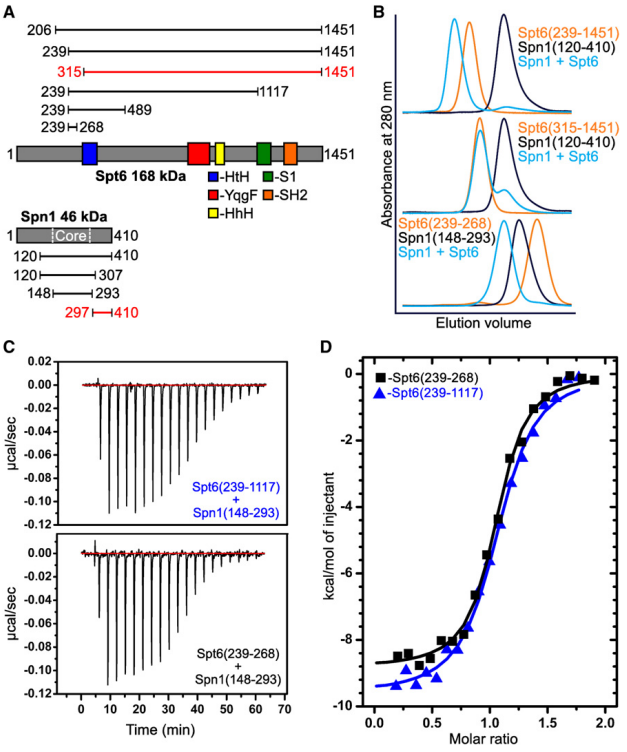


Figure 1. The Spn1 Core Binds 30 Residues in the Spt6 N-Terminal Region

(A) Spt6 and Spn1 domain organization and their interacting regions. Constructs assayed for binding are indicated above (Spt6) or below (Spn1) with their N- and C-terminal residues given. Constructs that showed binding are colored black, while those that did not show binding are colored red.

(B) Overlays of size-exclusion chromatograms with Spn1 in black, Spt6 in orange, and binding experiments in blue. Chromatograms are scaled on the y axis to allow comparison of the elution volume profiles of each protein or binding experiment. Spn1 has a 6-fold lower molar extinction coefficient, and so at equimolar concentrations, the Spn1 peak appears approximately six times smaller than the Spt6 peak (blue, middle panel). The x axis shows the 50–100 ml elution volume.

(C) Raw ITC data of the titration of Spn1(148–293) into Spt6(239–1117) (top) and the titration of Spt6 (239–268) into Spn1(148–293) (bottom).

(D) Representative binding isotherms of the titration experiments from (C). See text and Table S1 for thermodynamic parameters.

not included in the final model. Spn1 retains the same globular eight helix-bundle fold in the complex, showing no notable conformational changes upon binding Spt6 (rmsd ~0.5 Å on 135/145 pairs of Cα atoms). Binding of Spt6 to Spn1 does not mimic the binding of RNAPII CTD peptides to Pcf11/SCAF8, as it involves a face of Spn1 that is structurally distinct from the binding surface of Pcf11/SCAF8 (Figure S1B).

The structure of the Spn1(148–293):Spt6(239–268) complex reveals an extensive interface in which the Spt6 residues drape across the structured Spn1 core domain as two helices, H1 (residues 239–249) and H2 (256–265), that are connected by a short extended segment. There are multiple contacts along the length of the Spt6 segment that bury a total of 1790 Å^2 of accessible surface area upon complex formation. Spt6 H1 and the connecting segment, in particular, include a high fraction of conserved residues and contact a conserved patch on the Spn1 surface, thereby indicating that this interface is likely to be preserved across eukaryotes (Figures 2C–2E, Figures S1C and S1D).

Distinctive hydrophobic and polar interactions are made throughout the length of the Spt6 peptide (Figure 3). Starting at the Spt6 peptide N terminus, the H1 residues M245, I248, and F249 contact a hydrophobic pocket formed by Spn1 residues L256, G262, I266, I286, and W289. The C-terminal carbonyl groups of Spt6 H1 form water-mediated hydrogen bonds with Spn1 R263 guanidinium, whose extensive network of contacts also includes direct hydrogen bonds with the carboxylate of D254 in the extended region of Spt6. The aromatic side chain of Spt6 Y255, from the extended central region, makes van der Waals contacts with Spn1 P227, V264, and F267, while its hydroxyl group forms a hydrogen bond with the carboxylate of Spn1 E226. The Spt6 H2 residues W257, A258, L259, and I261 contact two hydrophobic patches formed by Spn1 P227, G231, V264, and F267. Finally, the Spt6 E262 carboxylate hydrogen bonds with Spn1 S271 and R273, and K272 forms polar contacts with the amide of Spt6 N263. These extensive interactions are consistent with the strong specific binding observed between Spn1 and Spt6.

Mutations that Disrupt Spn1-Spt6 Complex Formation In Vitro

To validate the relevance of the interface seen in the crystal structure, Spn1(148–293) and Spt6(239–268) variants were expressed and purified, and binding affinities were measured by ITC (Figure 4, Figure S3, Table S1). The Spt6-F249K protein bound Spn1 with a mean K_D of 11 μM, a reduction in affinity compared to the WT interaction of about 60-fold. The affinity of the Spn1-R263D protein for Spt6 (mean K_D of 30 μM) was reduced about 170-fold compared to WT. No binding was

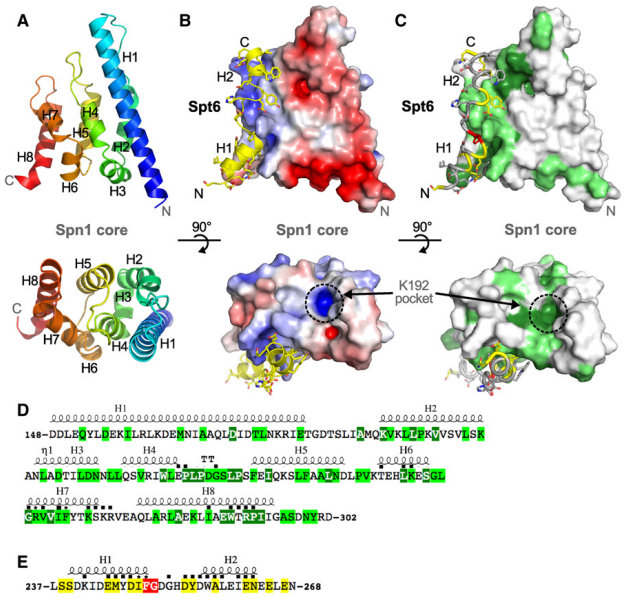


Figure 2. Structures of the Spn1 Core and Its Complex with Spt6

(A) The structure of the Spn1 core shown as a cartoon representation in two orthogonal views. The polypeptide chain is colored as a blue to red rainbow from N to C terminus. Secondary structures and the N and C termini are labeled.

(B) Structure of the Spn1-Spt6 complex. The Spn1 core is shown as a surface electrostatic (±5 kT/e) representation in the same views as (A). Spt6 is colored yellow and the N and C termini and the helices are labeled. Spn1 K192 has been implicated in interactions with RNAPII (Zhang et al., 2008) and lies in the basic conserved pocket indicated. Remnants of the affinity tags that remain in the crystallized Spn1 and Spt6 proteins are not shown in any of the figures and do not contribute to the Spn1 interface.

(C) Same as (B), except that the Spn1 surface and Spt6 peptide are colored according to conservation.

(D) Spn1 core amino acid sequence. Secondary structural elements are indicated above, and were defined with ESPript (Gouet et al., 2003). Residues that make direct contact across the interface are indicated with a black square or with an asterisk if they were mutated in this study. Numbering and residue identities shown here refer to the *S. cerevisiae* protein. Coloring represents degrees of conservation, dark green (high), light green (medium), in an alignment of proteins from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio*, and *Homo sapiens*. Amino acid sequences were aligned using the T-coffee multiple sequence alignment method (Notredame et al., 2000) and slightly adjusted manually in light of the structure (Figure S2).

(E) Same as (D) but for Spt6. High conservation (red), medium conservation (yellow).

detected for the Spn1-F267E protein at concentrations expected to give a reliable estimate of affinity around 40 μM. These observations validate the crystallographic interface and demonstrate that single point mutations are sufficient to significantly reduce binding affinity in vitro.

Mutating the Spn1-Spt6 Interface Causes Profound Effects In Vivo

To determine the physiological importance of the Spn1-Spt6 interaction, we introduced *spt6-F249K*, *spn1-R263D*, and *spn1-F267E* mutations into the genomes of yeast cells such that each mutant protein was expressed from its native promoter as the sole source of the affected protein. *spt6-F249K* caused a moderate growth defect at low temperatures, and this was strongly enhanced at elevated temperatures (Figure 5A). Strains with the *spn1-R263D* mutation grew normally at all temperatures, while those with the *spn1-F267E* mutation were normal at low temperatures but failed to grow at high temperatures. Each of the three mutations caused a defect in transcription initiation site selection due to defective chromatin repression, as indicated by the Spt$^-$ phenotype (growth of a strain with the *lys2-128∂* allele on medium lacking lysine (Simchen et al., 1984; Figures 5A–5C). The strength of the defect was different for each mutation, being weakest for *spt6-F249K*, stronger for *spn1-R263D*, and strongest for *spn1-F267E*. This order correlates precisely with the level of perturbation of binding observed with these mutant proteins in vitro (Figure 4, Table S1), strongly supporting the importance of the Spn1-Spt6 binding interface detected in our crystal structure in maintaining a repressive chromatin state in vivo.

Each mutation in the Spn1-Spt6 binding interface disturbed the interaction to a different extent in vitro (Figure 4, Table S1), but each individual mutation was tolerated in vivo (Figure 5). If the phenotypes caused by individual mutations result from partial disruption of binding, then combining the mutations should lead to enhanced defects. Consistent with this prediction, cells with both *spt6-F249K* and *spn1-R263D* mutations were viable but severely impaired for growth (Figure 5A), and cells with both *spt6-F249K* and *spn1-F267E* were inviable (Figure S4D). The severity of the defect caused by combining mutations therefore correlates with the level of disruption of binding by the individual mutations, suggesting that the Spn1-Spt6 interaction is essential for viability. We were unable to detect an interaction between Spt6 and Spn1-F267E in vitro (Figure 4), but the viability of the *spn1-F267E* strain suggests that this mutant retains some binding. Combining *spt6-F249K* with *spn1-K192N* was also lethal (Figure S4D); this mutation does not directly affect the Spn1-Spt6 interface (Figures 2B and 2C) but has been shown to decrease the

Table 1. Data Collection and Refinement Statistics

	Spn1 Core SeMet	Spn1 Core Native	Spn1-Spt6 Complex
Data Collection			
Beamline	SSRL 9-2	Home source	Home source
Space group	$P3_112$	$P3_212$	$P2_12_12$
Unit cell dimensions			
a, *b*, *c* (Å)	61.9, 61.9, 240.4	61.3, 61.3, 116.05	105.9, 68.7, 73.9
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90
Resolution (Å)	35–3.0	30–2.15	30–2.15
Wavelength (Å)	0.97923 (Peak)	1.54178	1.54178
I/σ_I	23.2 (2.8)	22 (4.1)	19.7 (3.1)
Completeness (%)	90.8 (92.3)	99.8 (98.4)	100.0 (100.0)
R_{sym} (%)	5.2 (47.6)	6.1 (31.9)	6.4 (54.4)
Redundancy	4.2 (4.1)	5.8 (3.8)	5.3 (5.0)
Refinement			
Resolution (Å)		29.6–2.15	27.7–2.15
Number of reflections		13,847	30,096
R_{work}/R_{free} (%)		18.5/22.4	18.6/24.4
Number of protein atoms		1,216	2,903
Number of solvent atoms		144	268
Rmsd bond lengths (Å)/angles (°)		0.007/1.001	0.012/1.279
φ/ψ most favored/allowed (%)		99.3/100.0	99.2/100.0

Values in parentheses are for the highest-resolution shell.

interaction between Spn1 and RNAPII (Zhang et al., 2008), and our results with the recombinantly expressed protein suggest that Spn1-K192N protein is unstable (data not shown). Synthetic growth defects therefore support the importance of the Spn1-Spt6 interaction, but defects outside this interface can also cause additive growth defects.

Another strategy for determining the importance of the Spn1-Spt6 interface is to test the effect of overexpressing one partner. If decreased affinity of Spt6-F249K protein for Spn1 is responsible for growth defects in vivo, these defects might be suppressed by increasing the level of Spn1 protein. We tested this by transforming an *spt6-F249K* strain with high-copy plasmids containing variants of *SPN1*. As shown in Figure 5C, increasing the level of normal *SPN1* suppressed the temperature sensitivity and partially corrected the Spt$^-$ phenotype caused by *spt6-F249K*. Consistent with retention of partial binding, overexpression of *spn1-R263D* also had an effect but suppressed less efficiently, rescuing growth at 36° but not at 38° and having no effect on the Spt$^-$ phenotype. Spn1-K192N is not active at elevated temperatures (Zhang et al., 2008; and Figure 5A), and overexpression of this allele also did not rescue the phenotypes caused by *spt6-F249K*. Elevated *SPN1* copy number did not suppress the temperature sensitivity or Spt$^-$ phenotypes caused by the *spt6-1004* allele (data not shown), which is a deletion of the helix-hairpin-helix domain within the Tex-like core of Spt6 (Kaplan et al., 2005). The suppression of *spt6* defects by increased Spn1 is therefore at least partly specific for a mutation that alters the Spn1-Spt6 interface, and supports the importance of this interaction in an essential function in vivo.

Spt6 Binds Nucleosomes Directly and Is Inhibited by Spn1

Spt6 has been shown to bind both $(H3\text{-}H4)_2$ tetramers and H2A-H2B dimers (Bortvin and Winston, 1996), leading to models in which Spt6 acts as a histone chaperone during nucleosome eviction and redeposition. Following our earlier prediction (Johnson et al., 2008) that the N-terminal region of Spt6 binds nucleosomes, we examined purified Spt6 constructs for nucleosome-binding activity in an electrophoretic mobility shift assay (EMSA). Interestingly, we found that intact Spt6 does bind nucleosomes in this assay (Figure 6), but does so only in the presence of the small HMGB family member Nhp6 (Stillman, 2010). This is similar to the requirement for addition of Nhp6 to observe complexes between nucleosomes and a different histone chaperone, FACT (Formosa et al., 2001). As with FACT, this suggests that Spt6 may form stable complexes with nucleosomes only after the nucleosome has been partially destabilized by Nhp6. This requirement appears to be physiologically relevant, because loss of Nhp6 in vivo exacerbated the growth defects caused by any of several mutations in Spt6 (Figure 5C, Figure S4A). Histone chaperones do not necessarily bind intact nucleosomes; in fact, the Asf1-(H3-H4) interaction is incompatible with histone contacts within the nucleosome (Antczak et al., 2006; English et al., 2006; Natsume et al., 2007). Binding to both free histones and to nucleosomes therefore might indicate that Spt6 makes multiple distinct contacts with nucleosomes and their components during different steps in chromatin maintenance. Spt6(239–1451) bound nucleosomes while Spt6 (315–1451) did not (Figure 6, compare lanes 2 and 3). Thus, the

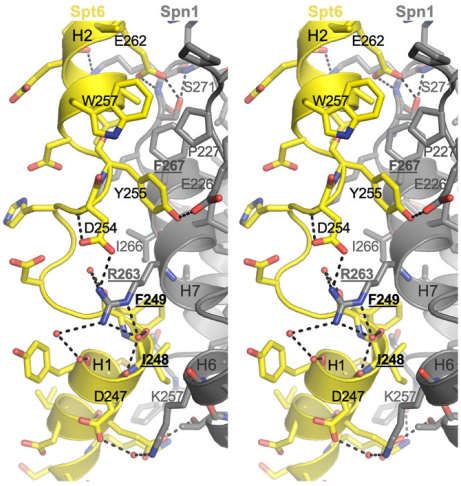


Figure 3. Details of the Spn1-Spt6 Interface
Stereoview of the Spn1-Spt6 interface. Residues mutated in this study are indicated in bold font and underlined. Water molecules are shown as red spheres, and polar interactions are indicated by black dashes.

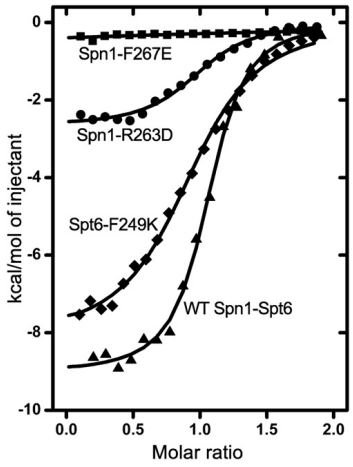


Figure 4. Mutations at the Spn1-Spt6 Interface Disrupt Binding In Vitro
Representative ITC binding isotherms for the indicated mutant Spt6(239–268) and Spn1(148–293) proteins, Spt6-F249K (diamonds), Spn1-R263D (circles), and Spn1-F267E (squares). A WT isotherm is shown for reference (triangles). See text and Table S1 for thermodynamic parameters.

region of Spt6 required for nucleosome binding (residues 239–314) contains the region required for Spn1 binding (239–268), suggesting that these two interactions may be mutually exclusive.

Consistent with this possibility, adding Spn1(120–410) to the nucleosome-binding assay inhibited the formation of Spt6-nucleosome complexes (Figure 6, compare lanes 2 and 5). Spn1 did not form stable complexes with nucleosomes itself, so the inhibition is unlikely to be caused by competition between Spn1 and Spt6 for a common binding site on nucleosomes. Spn1 did not interact with Nhp6 genetically (Figures S4B and S4C), and inhibition of Spt6-nucleosome complex formation by Spn1 could not be overcome by increasing the concentration of Nhp6 (data not shown), making it unlikely that the inhibition is caused by sequestering of Nhp6 by Spn1. Instead, Spn1 appears to prevent Spt6 from binding to nucleosomes directly by blocking the Spt6 binding domain. Supporting this interpretation, the Spn1-R263D variant with reduced affinity for Spt6 did not block formation of Spt6-nucleosome complexes efficiently (Figure 6, lane 7). Further, the Spt6-F249K mutation affects a region important for both nucleosome binding and Spn1 interaction (Figures 4 and 6, compare lanes 2 and 8). Spt6-F249K protein was impaired for nucleosome binding, but the residual binding was partially resistant to the addition of Spn1 (Figure 6, lanes 8 and 9). These results show that Spt6 residue F249 contributes to both nucleosome binding and to Spn1 binding, and that Spn1 binding can block an interaction between Spt6 and nucleosomes. The Spn1-Spt6 interaction can therefore provide a switch that controls the interaction of Spt6 with nucleosomes, and the proper functioning of this switch is important for maintaining normal chromatin structure.

DISCUSSION

We have determined a crystal structure of the ordered central domain of Spn1 and shown that it binds a 30 residue segment of Spt6. We have also determined a crystal structure of an Spn1-Spt6 complex that reveals that Spt6 binds in an extended/helical conformation that drapes the Spt6 residues along one face of Spn1. We further used site-directed mutations to disrupt this interaction and demonstrated that the interaction observed in solution depends on residues located at the interface seen in the crystal structure. Binding is not accompanied by conformational changes in Spn1. In contrast, Spt6(239–268) is almost certainly unstructured in isolation and becomes ordered upon binding Spn1. Indeed, the first 300 residues of Spt6 are likely to be unstructured in isolation (Ward et al., 2004), whereas much of the remainder of the protein appears to comprise multiple recognizable structural domains that likely fold against each other to form an elongated structure (Johnson et al., 2008). The extended and inherently flexible nature of the Spn1 binding sequence of Spt6 presumably explains why dramatic mutations in the interface substantially weaken but do not completely eliminate this interaction, as localized perturbations can be accommodated by conformational changes that do not propagate across the entire interface. Moreover, the ~35 residues separating the Spn1-binding residues from the ordered

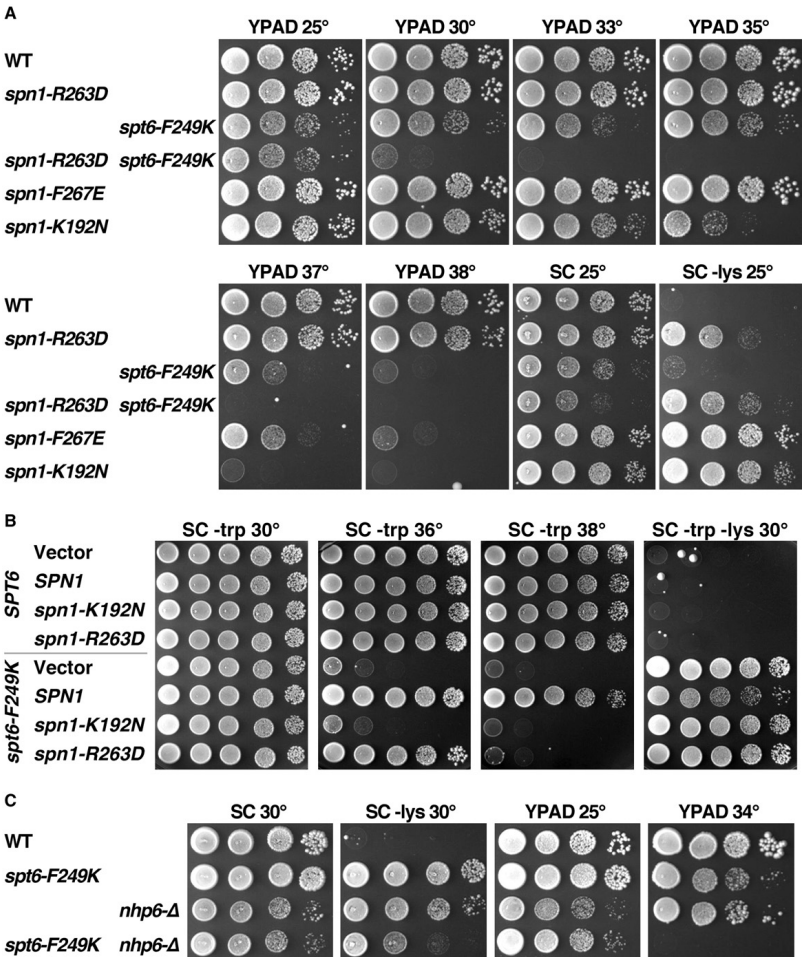


Figure 5. The Spn1-Spt6 Interaction Is Significant for the Essential Activities of Each Protein In Vivo

(A) Isogenic strains from the A364a genetic background and with the relevant genotypes indicated (Supplemental Information) were grown to saturation in rich medium, then aliquots of 10-fold serial dilutions were placed on solid medium and incubated as labeled. SC is synthetic medium (complete or lacking tryptophan or lysine as noted) and YPAD is rich medium. Growth on medium lacking lysine reveals the Spt$^-$ phenotype, reporting here on aberrant transcription initiation from the *lys2-128∂* allele. The strain with the *spt6-F249K* allele grows slowly on SC –lys at 25°, but the Spt$^-$ phenotype is more robust at 30° (B).

(B) WT and *spt6-F249K* strains were transformed with a high-copy number vector or the same vector carrying the version of *SPN1* noted (Supplemental Information). Multiple transformants were tested to insure that the phenotypes detected are typical, then one clone of each was grown to saturation in synthetic medium lacking tryptophan to select for retention of the plasmids. Aliquots of 10-fold serial dilutions were placed on solid synthetic medium as in (A) and incubated as indicated.

(C) As in (A). *nhp6-Δ* indicates deletion of both *NHP6A* and *NHP6B*.

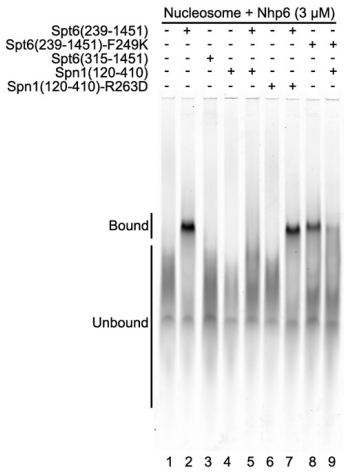


Figure 6. Spt6 Binds Nucleosomes and Is Competed by Spn1

EMSA visualizing the signal from the labeled DNA incorporated into the nucleosomes (Cy5, 633 nm). All reactions included 150 fmoles of nucleosomes, 3 μM Nhp6, and 2 μM of the indicated Spt6 or Spn1 proteins. Migration positions of bound and unbound nucleosomes are indicated. Addition of Spn1 reduced the amount of Spt6-nucleosome complex from 37% of the total signal in lane 2 to 3.2% in lane 5 (9% remaining), but Spn1-R263D only reduced it to 28% (76% remaining). Spt6-F249K formed a lower amount of stable complex (18%), but Spn1 was less effective at inhibiting this (6% bound, or 36% of the original signal remaining comparing lanes 8 and 9) than it was with WT Spt6 (9% of the original signal remaining).

region of Spt6 likely provide a tether whose flexibility may be required to allow the Spn1-Spt6 complex to form in multiple functional contexts.

Spt6 and Spn1 were found to copurify in high-throughput screens (Gavin et al., 2002; Krogan et al., 2002), and further studies suggested that their interaction promotes normal activation of genes regulated after recruitment of RNAPII (Fischbeck et al., 2002; Yoh et al., 2007; Zhang et al., 2008). However, Spt6 and Spn1 have also been implicated in several other distinct and overlapping roles, and each has been implicated in interactions with multiple factors. We have verified a functional interplay between Spt6 and Spn1 proteins in *S. cerevisiae* by demonstrating that cells cannot tolerate certain partial loss-of-function alleles of both genes simultaneously, and by showing that the defects caused by the *spt6-F249K* allele, which impairs binding with Spn1, can be suppressed by overexpressing *SPN1*. Our results therefore show that a short region of Spt6 is important for interacting with Spn1 during at least one of the essential processes mediated by these proteins.

One of the primary functions ascribed to Spt6 is as a histone chaperone that promotes the reassembly of nucleosomes following passage of RNAPII (Adkins and Tyler, 2006; Bortvin and Winston, 1996; Cheung et al., 2008; Kaplan et al., 2003). Spt6 was previously shown to bind histones and to promote nucleosome deposition in vitro (Bortvin and Winston, 1996), and we now show that it can also bind to intact nucleosomes. The dependence of this interaction upon the presence of Nhp6 is consistent with our genetic studies and with an equivalent dependence of the unrelated FACT histone chaperone for its interaction with intact nucleosomes (Formosa et al., 2001). We found that nucleosome binding requires a section of Spt6 that overlaps with the Spn1 binding site, that Spn1 antagonizes Spt6-nucleosome binding in vitro, that either *SPT6* or *SPN1* mutations that affect the Spn1-Spt6 interaction cause the Spt$^-$ phenotype, and that the effects of the *spt6-F249K* allele can be suppressed by overexpressing Spn1. These results show that the Spn1-Spt6 interaction disrupts Spt6-nucleosome binding and that this disruption has a positive role in maintaining normal chromatin. An attractive explanation is that Spn1 is actively engaged in the nucleosome reassembly process, possibly by disengaging Spt6 from nucleosomes to allow multiple rounds of reassembly (see the graphical abstract available online). Another possibility is that Spn1 binding influences the balance between Spt6's functional roles in nucleosome reassembly and mRNA processing. Regardless of the precise mechanistic details, Spn1 appears to be important for Spt6-mediated nucleosome reassembly in vivo.

Nucleosome reassembly is just one of several processes in which Spt6 and Spn1 have been implicated. For example, Spn1, in complex with Spt6, has been reported to interact with a variety of other factors that function in mRNA processing, nucleosome modification, or transcription, including REF1/Aly, Setd2, and RNAPII (Yoh et al., 2007, 2008; Zhang et al., 2008). Our structural data provide insight into how Spn1 might accommodate simultaneous interactions. Spt6 binds against a region of the Spn1 surface that is rich in conserved residues, but other regions of Spn1 display a similar level of conservation and are therefore good candidates for binding surfaces for other proteins (Figure 2C, Figure S1C). This includes residues that are immediately adjacent to the Spt6 binding surface but extend beyond contacts with Spt6, which may indicate that other factors can bind Spn1 cooperatively or competitively with Spt6. Finally, K192, whose mutation to asparagine impairs Spn1 function (Zhang et al., 2008) and appears to function in interactions with RNAPII, is located in a conserved pocket that is formed at the ends of H1, H2, and H4 (Figures 2B and 2C), suggesting another potential binding surface.

In summary, we have determined the structural basis for the interaction between Spn1 and Spt6, and shown that this interaction is important in vivo and that it regulates Spt6-nucleosome binding in vitro. Overall, our data indicate that Spn1 is important for Spt6-mediated nucleosome reassembly, perhaps by regulating the process or by providing a switch that drives disengagement. This does not, however, seem to encompass all of the functional roles in which these two proteins and their interaction with each other participate (Lindstrom et al., 2003; Yoh et al., 2007, 2008; Zhang et al., 2008). Sequence conservation indicates that other surfaces on the Spn1 core domain are good candidates for mediating functionally important interactions. Moreover, interactions with Setd2 and REF1/Aly have been mapped to the N- and C-terminal regions, respectively, which extend beyond the Spn1 core domain and are predicted to be

Molecular Cell
Spn1-Spt6 Structure/Function

unstructured (Ward et al., 2004; Yoh et al., 2007, 2008). This use of inherently flexible segments, including the N-terminal region of Spt6 that binds the Spn1 core, may provide a mechanism that allows flexibility in a crowded transcriptional environment.

EXPERIMENTAL PROCEDURES

See the Supplemental Information for protein expression and purification, strains used, strain construction, plasmid construction, and tests of the effects of combining *spn1* and *spt6* mutations.

Crystallization and Structure Determination

All crystals were grown at 20°C by sitting drop vapor diffusion. Se-Spn1 (148–307) drops comprised 2 μl of 8 mg.mL^{-1} protein with 2 μl of well solution (0.1 M Bis-Tris propane [pH 7.0], 1.4 M Li_2SO_4). Native Spn1(148–307) drops comprised two parts 7.5 mg.mL^{-1} protein and one part well solution (0.01 M $MgCl_2$, 0.05 M HEPES [pH 7.0], 1.6 M $[NH_4]_2SO_4$). Spn1(148–293)-Spt6(239–268) drops comprised two parts 13 mg.mL^{-1} protein and one part well solution (0.2 M $Mg[CH_3CO_2]_2$, 0.1 M MES [pH 6.5], 20% PEG 8000). Crystals were cryoprotected in a solution of the reservoir made up with 30% glycerol, and cooled by plunging into liquid nitrogen.

All data were processed using HKL2000 (Otwinowski and Minor, 1997). Phases were determined for Se-Spn1(148–307) by single-wavelength anomalous diffraction. Phenix (phenix.autosol) (Adams et al., 2010) located four out of six possible selenium positions and computed a map into which a model was built. This unrefined model was used in molecular replacement using PHASER (McCoy et al., 2005) to determine the structure of native Spn1(148–307) at 2.15 Å resolution. This subsequently refined model was used to determine the Spn1(148–293):Spt6(239–268) structure by molecular replacement. In all cases, model building, refinement,and validation were performed using Coot (Emsley et al., 2010), Phenix (Adams et al., 2010), and MolProbity (Chen et al., 2010), respectively. Refinement of the Spn1(148–293)-Spt6(239–268) complex included TLS restraints (Painter and Merritt, 2006).

Electrostatic potential surfaces were calculated using APBS (Baker et al., 2001). Figures of molecular structures were generated using PyMol (DeLano, 2002).

Size-Exclusion Chromatography Binding Assay

Purified recombinant proteins were mixed at equimolar concentrations (2 μM) and incubated for 2 hr at 4°C. The protein mixture was concentrated to 15 μM and chromatographed on a 120 ml Superdex 200 16/60 column (GE Healthcare) in 15 mM Tris (pH 7.5), 200 mM NaCl, 5% glycerol, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol.

Isothermal Titration Calorimetry

Amino acid substitutions were made by site-directed mutagenesis and verified by DNA sequencing. Purified recombinant proteins were dialyzed overnight at 4°C against 2 L of degassed ITC buffer (20 mM Tris [pH 7.5], 150 mM NaCl, 5% glycerol, 2 mM 2-mercaptoethanol, 0.5 mM EDTA). Titrations for all reactions were done at 25°C on an iTC200 (Microcal), including an initial injection of 0.4 μl (which was omitted from data analysis), and all injections were spaced 180 s apart. For Spt6(239–1117) reactions, the titrations were carried out with 18 injections of 1.8 μl 76 μM Spn1(148–293) into 8.2 μM Spt6(239–1117). For Spt6(239–268) reactions, the titrations were with 18 injections of 2 μl 74 μM Spt6(239–268) into 8 μM Spn1(148–293). For Spt6-F249K reactions, the titrations were with 18 injections of 1.8 μl 1.52 mM Spt6(239–268)-F249K into 168 μM Spn1(148–293). For Spn1-R263D reactions, the titrations were with 18 injections of 1.8 μl 3.5 mM Spt6(239–268) into 389 μM Spn1(148–293)-R263D. For Spn1-F267E reactions, the titrations were with 18 injections of 2.0 μl 5.2 mM Spt6(239–268) into 578 μM Spn1(148–293)-F267E). In all cases, three independent experiments were performed. Data were analyzed using Origin software (Microcal), and the stoichiometry (*N*), association constant (K_A), and change in enthalpy (ΔH) were obtained by fitting the isotherm to the one-site binding model. Other thermodynamic parameters were calculated using the following relationships:

$$K_A^{-1} = K_D \text{ and } -RT\ln K_A = \Delta H - T\Delta S.$$

Nucleosome Preparation and Gel Mobility Shift Binding Assay

A 146 bp sea urchin 5S rDNA fragment labeled with Cy5 was generated by PCR and gel purified. *Xenopus laevis* histone H2A-S113C was labeled with Oregon Green 488-maleimide and then assembled into nucleosome core particles as described (Xin et al., 2009). Reactions contained 15 nM nucleosomes, 2 μM Spt6 or Spn1 proteins, 100 mM NaCl, 0.8 mg/ml HSA, 9.7% glycerol, and 3 μM S. cerevisiae Nhp6a. Following incubation at 30°C for 15 min, samples were subjected to electrophoresis on native polyacrylamide gels (4.5% acrylamide [acr:bis, 37.5:1], 0.5 X TBE, 5% glycerol, 2 mM $MgCl_2$ at 160 V for 6 hr at 4°C). The gels were scanned using a Typhoon imager at 670 BP30/Red(633 nm) for Cy5-DNA and 520 BP40/Blue(488 nm) for Oregon Green 488-H2A, and the amount of signal in the bound form quantified with ImageQuant Software (GE Health Sciences).

ACCESSION NUMBERS

Coordinates and structure factor amplitudes have been deposited for the Spn1 and Spn1-Spt6 crystal structures have been deposited in the Protein Data Bank under ID codes 3o8z and 3oak, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes one table, four figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at doi:10.1016/j.molcel.2010.11.014.

ACKNOWLEDGMENTS

Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the U.S. National Institutes of Health (NIH), National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. We thank David Stillman and Warren Voth for helpful discussions and Charisse Kettelkamp and Laura McCullough for technical assistance. This work was supported by NIH grants to C.P.H. and T.F.

Received: May 26, 2010
Revised: August 11, 2010
Accepted: September 21, 2010
Published online: November 18, 2010

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Adkins, M.W., and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. Mol. Cell *21*, 405–416.

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. Nature *420*, 837–841.

Antczak, A.J., Tsubota, T., Kaufman, P.D., and Berger, J.M. (2006). Structure of the yeast histone H3-ASF1 interaction: implications for chaperone mechanism, species-specific interactions, and epigenetics. BMC Struct. Biol. *6*, 26.

Ardehali, M.B., Yao, J., Adelman, K., Fuda, N.J., Petesch, S.J., Webb, W.W., and Lis, J.T. (2009). Spt6 enhances the elongation rate of RNA polymerase II in vivo. EMBO J. *28*, 1067–1077.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Baniahmad, C., Nawaz, Z., Baniahmad, A., Gleeson, M.A., Tsai, M.J., and O'Malley, B.W. (1995). Enhancement of human estrogen receptor activity by SPT6: a potential coactivator. Mol. Endocrinol. *9*, 34–43.

Becker, R., Loll, B., and Meinhart, A. (2008). Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem. *283*, 22659–22669.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473–1476.

Bucheli, M.E., and Buratowski, S. (2005). Npl3 is an antagonist of mRNA 3′ end formation by RNA polymerase II. EMBO J. *24*, 2150–2160.

Burckin, T., Nagel, R., Mandel-Gutfreund, Y., Shiue, L., Clark, T.A., Chong, J.L., Chang, T.H., Squazzo, S., Hartzog, G., and Ares, M., Jr. (2005). Exploring functional relationships between components of the gene expression machinery. Nat. Struct. Mol. Biol. *12*, 175–182.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol. *6*, e277. 10.1371/journal.pbio.0060277.

Clark-Adams, C.D., and Winston, F. (1987). The SPT6 gene is essential for growth and is required for delta-mediated transcription in Saccharomyces cerevisiae. Mol. Cell. Biol. *7*, 679–686.

DeLano, W.L. (2002). The Pymol Molecular Graphics System (Palo Alto, CA: Delano Scientific).

Dengl, S., Mayer, A., Sun, M., and Cramer, P. (2009). Structure and in vivo requirement of the yeast Spt6 SH2 domain. J. Mol. Biol. *389*, 211–225.

Denis, C.L. (1984). Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics *108*, 833–844.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D.K., Aida, M., Inukai, N., Narita, T., Yamada, T., Furuya, A., et al. (2004). Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. Mol. Cell. Biol. *24*, 3324–3336.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495–508.

Fischbeck, J.A., Kraemer, S.M., and Stargell, L.A. (2002). SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant. Genetics *162*, 1605–1616.

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y., and Stillman, D.J. (2001). Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. EMBO J. *20*, 3506–3517.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. *31*, 3320–3323.

Hartzog, G.A., Wada, T., Handa, H., and Winston, F. (1998). Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. Genes Dev. *12*, 357–369.

Johnson, S.J., Close, D., Robinson, H., Vallet-Gely, I., Dove, S.L., and Hill, C.P. (2008). Crystal structure and RNA binding of the Tex protein from Pseudomonas aeruginosa. J. Mol. Biol. *377*, 1460–1473.

Kaplan, C.D., Morris, J.R., Wu, C., and Winston, F. (2000). Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. Genes Dev. *14*, 2623–2634.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096–1099.

Kaplan, C.D., Holland, M.J., and 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.

Keegan, B.R., Feldman, J.L., Lee, D.H., Koos, D.S., Ho, R.K., Stainier, D.Y., and Yelon, D. (2002). The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription in the zebrafish embryo. Development *129*, 1623–1632.

Kok, F.O., Oster, E., Mentzer, L., Hsieh, J.C., Henry, C.A., and Sirotkin, H.I. (2007). The role of the SPT6 chromatin remodeling factor in zebrafish embryogenesis. Dev. Biol. *307*, 214–226.

Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J.F. (2002). RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol. Cell. Biol. *22*, 6979–6992.

Li, L., Ye, H., Guo, H., and Yin, Y. (2010). Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. Proc. Natl. Acad. Sci. USA *107*, 3918–3923.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., 3rd, and Hartzog, G.A. (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.

Ling, Y., Smith, A.J., and Morgan, G.T. (2006). A sequence motif conserved in diverse nuclear proteins identifies a protein interaction domain utilised for nuclear targeting by human TFIIS. Nucleic Acids Res. *34*, 2219–2229.

Liu, Z., Zhou, Z., Chen, G., and Bao, S. (2007). A putative transcriptional elongation factor hIws1 is essential for mammalian cell proliferation. Biochem. Biophys. Res. Commun. *353*, 47–53.

Maclennan, A.J., and Shaw, G. (1993). A yeast SH2 domain. Trends Biochem. Sci. *18*, 464–465.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.

Meinhart, A., and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. Nature *430*, 223–226.

Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. Nature *446*, 338–341.

Neigeborn, L., Celenza, J.L., and Carlson, M. (1987). SSN20 is an essential gene with mutant alleles that suppress defects in SUC2 transcription in Saccharomyces cerevisiae. Mol. Cell. Biol. *7*, 672–678.

Nishiwaki, K., Sano, T., and Miwa, J. (1993). emb-5, a gene required for the correct timing of gut precursor cell division during gastrulation in Caenorhabditis elegans, encodes a protein similar to the yeast nuclear protein SPT6. Mol. Gen. Genet. *239*, 313–322.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. *302*, 205–217.

Otwinowski, Z., and Minor, W., eds. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode (New York: Academic Press).

Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr. D Biol. Crystallogr. *62*, 439–450.

Shen, X., Xi, G., Radhakrishnan, Y., and Clemmons, D.R. (2009). Identification of novel SHPS-1-associated proteins and their roles in regulation of insulin-like

growth factor-dependent responses in vascular smooth muscle cells. Mol. Cell. Proteomics *8*, 1539–1551.

Simchen, G., Winston, F., Styles, C.A., and Fink, G.R. (1984). Ty-mediated gene expression of the LYS2 and HIS4 genes of Saccharomyces cerevisiae is controlled by the same SPT genes. Proc. Natl. Acad. Sci. USA *81*, 2431–2434.

Stillman, D.J. (2010). Nhp6: a small but powerful effector of chromatin structure in Saccharomyces cerevisiae. Biochim. Biophys. Acta *1799*, 175–180.

Vanti, M., Gallastegui, E., Respaldiza, I., Rodriguez-Gil, A., Gomez-Herreros, F., Jimeno-Gonzalez, S., Jordan, A., and Chavez, S. (2009). Yeast genetic analysis reveals the involvement of chromatin reassembly factors in repressing HIV-1 basal transcription. PLoS Genet. *5*, e1000339. 10.1371/journal.pgen.1000339.

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

Xin, H., Takahata, S., Blanksma, M., McCullough, L., Stillman, D.J., and Formosa, T. (2009). yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement. Mol. Cell *35*, 365–376.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev. *21*, 160–174.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes Dev. *22*, 3422–3434.

Zhang, L., Fletcher, A.G., Cheung, V., Winston, F., and Stargell, L.A. (2008). Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. Mol. Cell. Biol. *28*, 1393–1403.

Note Added in Proof

An independent study (M.-L. Diebold, M. Koch, E. Loeliger, V. Cura, F. Winston, J. Cavarelli, and C. Romier, EMBO J, 10.1038/emboj.2010.272) has also determined a very similar structure of a Spn1-Spt6 complex; the results of the two studies are highly complementary.

Molecular Cell, Volume *40*

Supplemental Information

Structure and Biological Importance of the Spn1-Spt6 Interaction, and Its Regulatory Role in Nucleosome Binding

Seth M. McDonald, Devin Close, Hua Xin, Tim Formosa, and Christopher P. Hill

Supplemental Information Inventory

Table S1. Thermodynamic parameters of the Spn1-Spt6 interaction. Related to Figures 1, 4, and S3.

Figure S1. Structural similarity between Spn1 core and proteins that bind RNAPII CTD and conservation of amino acid sequences. Related to Figure 2.

Figure S2. Electron density surrounding the Spt6 residues. Related to Figure 3.

Figure S3. Mutating the Spn1-Spt6 interface disrupts binding *in vitro*. Related to Figure 4.

Figure S4. Nhp6 and Spn1 each support functions of Spt6 *in vivo*. Related to Figure 5.

Supplemental Experimental Procedures:

Protein Expression and Purification

Strains Used

Strain Construction

Plasmid Construction

Testing the effect of combining *spn1* and *spt6* mutations

Supplemental References

Table S1. Thermodynamic parameters of the Spn1-Spt6 interaction.

Related to Figures 1, 4, and S3.

Reaction:	Spt6(239-1117) + Spn1(148-293)	Spt6(239-268) + Spn1(148-293)	[a] Spt6-F249K	[a] Spn1-R263D
N (sites)	[b] 1.01 ± 0.04	1.03 ± 0.01	0.97 ± 0.06	0.98 ± 0.1
K_D (μM)	0.168 ± 0.036	0.169 ± 0.053	10.7 ± 1.5	29.7 ± 6.3
ΔH (kcal mol^{-1})	-8.9 ± 0.6	-9.4 ± 0.7	-8.0 ± 0.5	-2.8 ± 0.2
ΔS (cal mol^{-1} K^{-1})	-0.5 ± 2.5	1.06 ± 2.8	-4.14 ± 2.1	11.3 ± 1.0

[a] Reactions were titrations of Spt6(239-268) into Spn1(148-293) with indicated mutant proteins. Spn1-F267E protein produced no detectable binding at concentrations expected to give a reliable estimate of affinity around 40 μM.

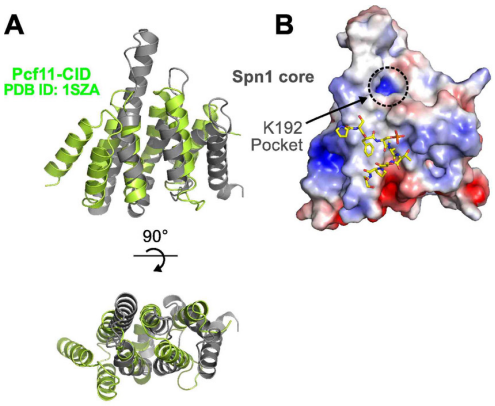
[b] Values reported as mean ± standard deviation and were calculated from three independent experiments of each indicated reaction.

Figure S1. Structural similarity between Spn1 core and proteins that bind RNAPII CTD and conservation of amino acid sequences. Related to Figure 2.

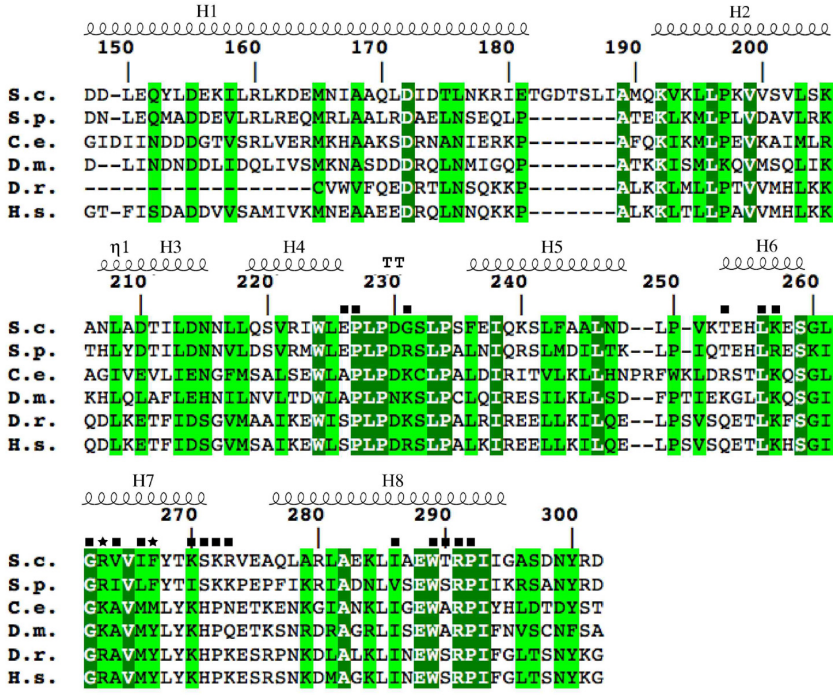
(A) An alignment of Spn1 (gray) and Pcf11 (yellow/green). Structures of the Spn1 core and Pcf11 (Meinhart and Cramer, 2004) and SCAF8 (Becker et al., 2008) were aligned with DALI Lite (Holm and Park, 2000). A population of overlaps of both proteins with Spn1 were obtained that typically displayed Z scores ~6 and RMSD values of 3-4 Å for overlap on ~100 residues. Sequence identity after structural alignment is just 5-10%. Alignments on SCAF8 are essentially identical.

(B) Spn1 surface with the Pcf11-bound RNAPII CTD peptide shown after overlap on the proteins. The Spn1 K192 pocket that has been implicated in interactions with RNAPII (Zhang et al., 2008) is indicated. As indicated in the main text, despite the structural similarity, we did not detect measurable binding of CTD peptides with Spn1.

(C-D) Alignments of homologs of (C) the Spn1 core and (D) the fragment of Spt6 that binds Spn1. S.c. *Saccharomyces cerevisiae*; S.p. *Schizosaccharomyces pombe*; C.e. *Caenorhabditis elegans*; D.m. *Drosophila melanogaster*; D.r. *Danio rerio*; H.s. *Homo sapiens*. Secondary structure is indicated above. Residues that make contact across the interface are indicated with a black square or with an asterisk if they were mutated in this study. Numbering refers to the *S. cerevisiae* proteins. Coloring represents degrees of conservation: dark green and red (high), light green and yellow (medium). Amino acid sequences were aligned using the T-coffee multiple sequence alignment method (Notredame et al., 2000) and slightly adjusted manually in light of the structure. Secondary structure designation was by ESPript (Gouet et al., 2003).



C Spn1 core region



D Spt6 N-terminal fragment

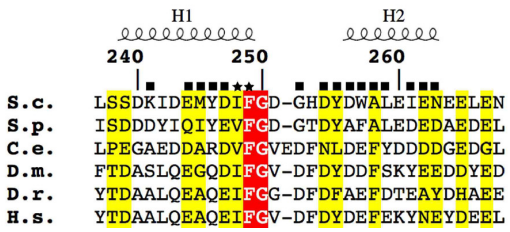


Figure S2. Electron density surrounding the Spt6 residues. Related to Figure 3.

Stereoview of the simulated annealing omit-electron density map at 1.5 σ displayed with the final refined model. Density is shown that lies within 4Å of atoms.

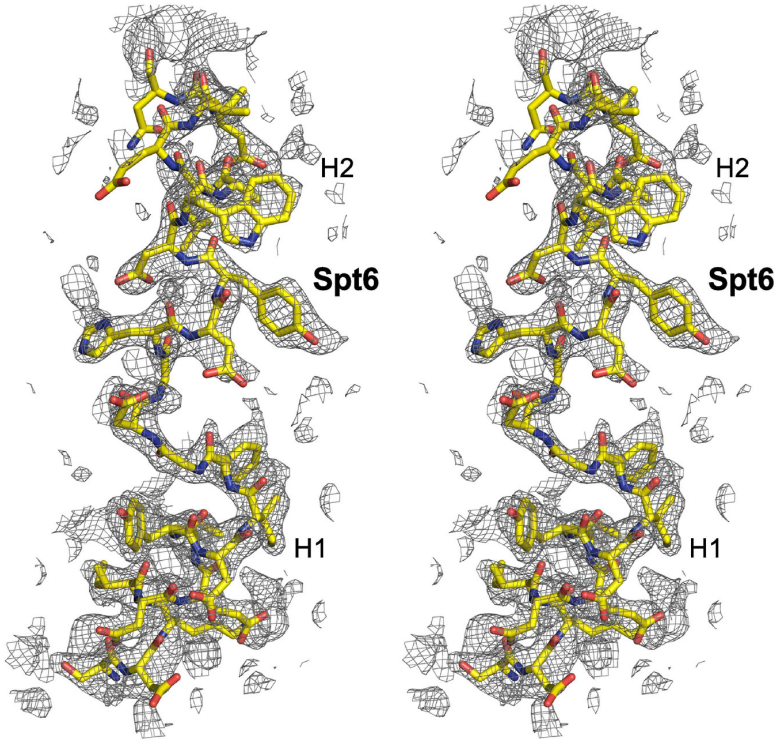


Figure S3. Mutations at the Spn1-Spt6 interface disrupt binding *in vitro*. Related to Figure 4.

Raw ITC data of the indicated titrations for the binding isotherms shown in Figure 4.

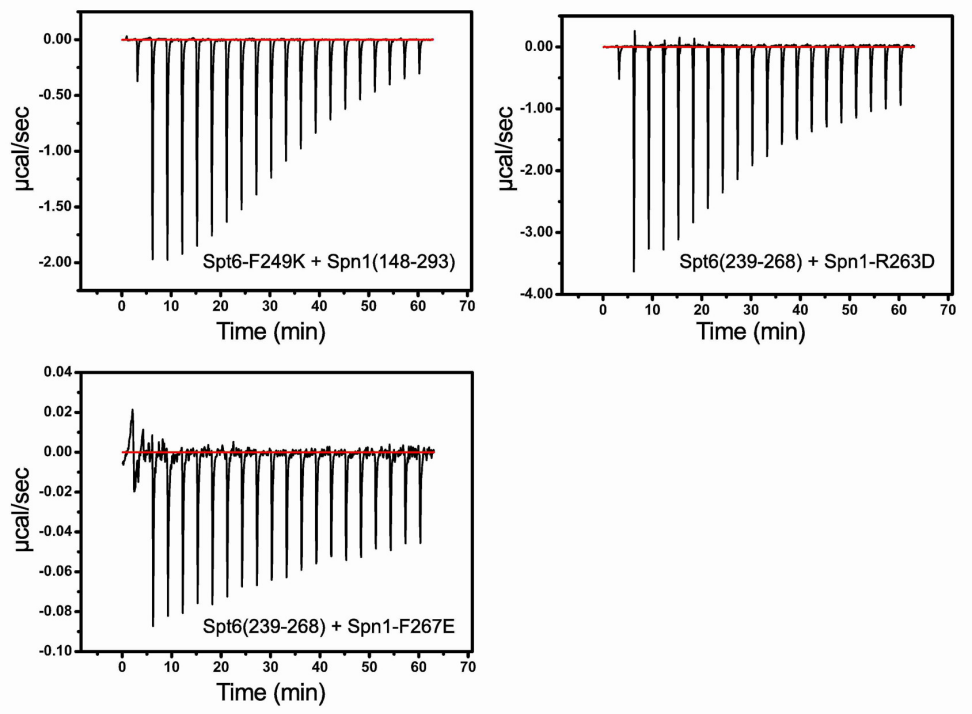
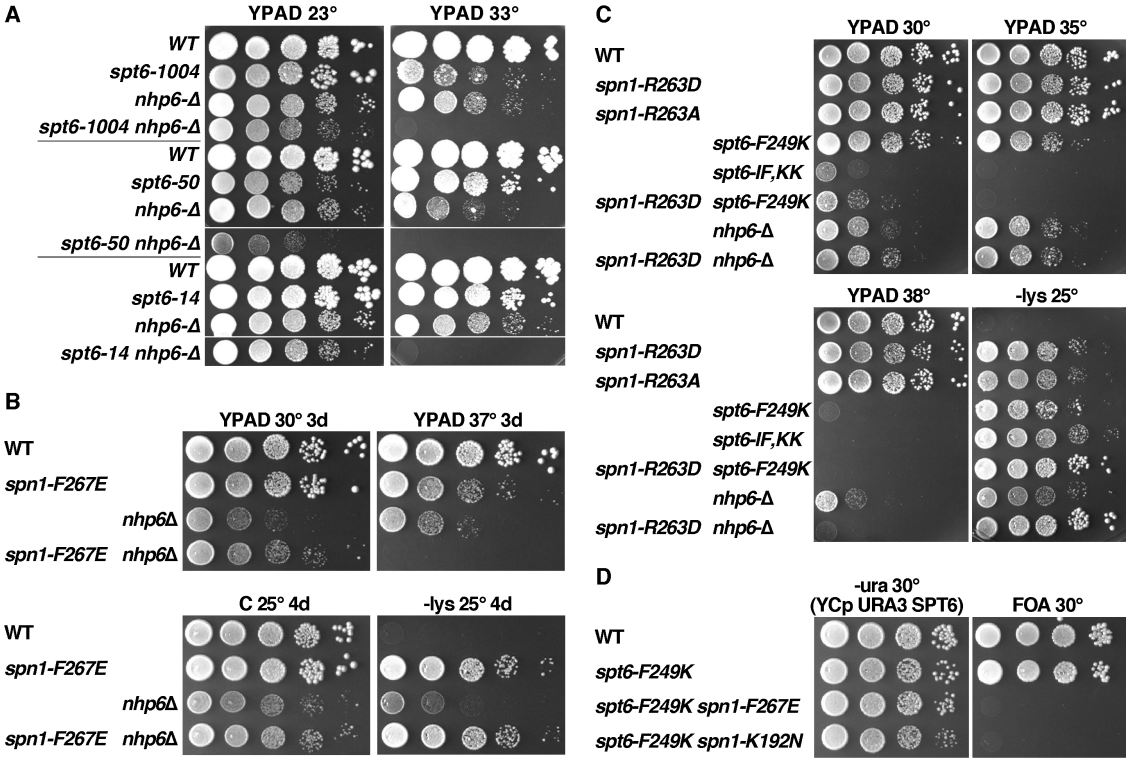


Figure S4. Nhp6 and Spn1 each support functions of Spt6 *in vivo.* Related to Figure 5.

(A) Loss of Nhp6 is detrimental to *spt6* mutants. Strains in the S288C background with the mutations indicated were grown to saturation in rich medium, then 10-fold serial dilutions were spotted to rich medium and incubated for 6 days at 23° C or 3 days at 33° C. Nhp6 is encoded by two very similar genes, *NHP6A* and *NHP6B*, so "*nhp6*-Δ" indicates deletion of both genes. *spt6-1004* is an internal deletion of the helix-hairpin-helix domain within the Tex-like core (Kaplan et al., 2005), *spt6-50* (K1274stop) is a premature termination of the C-terminal domain (deletion of 1274-1451), and spt6-14 is an S952F substitution (F Winston, personal communication); none of these alleles directly affect the Spn1-Spt6 interface. The growth defects caused by these three standard alleles of *SPT6* are all enhanced by loss of Nhp6, especially at elevated temperatures (rows 4, 8, 12), indicating that Nhp6 supports an essential function of Spt6 *in vivo*.

(B, C) As in panel A except the *spn1-F267E* and *spn1-R263D* alleles affecting the Spn1-Spt6 interface are tested in the A364a genetic background. Combining *nhp6*-Δ and *spn1* mutations has more subtle and variable effects, with *spn1-F267E* partially suppressing the growth defect caused by *nhp6*-Δ at 30°, but the double mutant being more impaired than single mutants at 37°. *spn1-R263D* also slightly suppressed *nhp6*-Δ at 30° but did not enhance the Ts- phenotype. These results suggest that Nhp6 does not directly support a function of Spn1, consistent with the model that Spn1 interacts with Spt6 only when Spt6 is not interacting with nucleosomes, and Nhp6 is needed only to support the role of Spt6 when it is interacting with nucleosomes. Panel C also shows the effect of the *spn1-R263A* allele (similar in effect to R263D), and the *spt6-I248K, F249K* double mutation (more severe than F249K alone).

(D) Combining partially defective alleles of *SPN1* and *SPT6* can cause synthetic lethality. Isogenic strains in the A364a genetic background with the *spn1-F267E* or *spn1-K192N* alleles were mated to a strain with the *spt6-F249K* mutation and the diploid was transformed with low copy plasmids with the *URA3* gene and either the *SPT6* or *SPN1* genes (YCp *URA3 SPT6* is shown). Haploid segregants with both *spn1* and *spt6* mutations were obtained after sporulation but were unable to grow on medium containing 5-FOA, indicating that either single mutation is tolerable but the combination of either *spn1* mutation with *spt6-F249K* is lethal. Similar tests also showed that *spn1-K192N* is lethal when combined with either *spt6-1004* or *spt6-F249K* in the S288C background (not shown).



SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Protein Expression and Purification

Coding sequences for full-length *S. cerevisiae* Spn1 and Spt6 proteins were amplified from genomic DNA and inserted into pET151-D-TOPO Amp expression vectors (Invitrogen) containing an N-terminal hexahistidine tag and TEV protease cleavage site. Site-directed mutagenesis was used to insert termination codons to generate C-terminally truncated constructs, while N-terminal truncations were cloned by amplification from the full-length plasmids and subsequent TOPO reaction. Proteins were expressed in *E. coli* BL21-codonplus-(DE3)-RIL cells (Stratagene) in ZY-5052 autoinduction media (Studier, 2005) at 37°C for 5 h and then at 19°C until saturated. For Spn1(148-307) and Spn1(148-293), expression was in LB media at 37°C with induction by 1 mM IPTG at an OD_{600} of 1.0, with subsequent incubation at 19°C for 5 hours. Harvested cell pellets were stored at -80°C.

Cell pellets were thawed, resuspended in lysis buffer (50 mM Tris pH 7.5, 500 mM NaCl, 5% glycerol, and 15 mM imidazole) in the presence of protease inhibitors and lysozyme, and sonicated. Following clarification by centrifugation (29,000xg, 45 min), soluble fractions were applied to Ni-NTA agarose resin (Qiagen) and eluted with 300 mM imidazole (100 mM NaCl). The N-terminal hexahistidine tags of the Spn1 proteins were then removed with TEV protease and any uncleaved protein removed by passage over an Ni-NTA column. Subsequently, Spt6(239-268) was applied to an anion-exchange column (5 mL HiTrap Q HP, GE Healthcare) and eluted with a NaCl gradient. Spn1 proteins were chromatographed on Heparin (5 mL HiTrap Heparin HP, GE Healthcare) and Q columns and collected in the flow through. The Spn1(148-293)-Spt6(239-268) complex was prepared by mixing equimolar amounts, removing

the N-terminal hexahistidine tag of Spt6(239-268) with TEV protease, and passing over a Ni-NTA column. The final purification step for Spn1(148-307) and for the complex was size-exclusion chromatography (Superdex 200 16/60, GE Healthcare) in 15 mM Tris pH 7.5, 100 mM NaCl, 5% glycerol, 0.5 mM EDTA, and 2 mM 2-mercaptoethanol. Selenomethionine-substituted Spn1(148-307) was expressed as described (Studier, 2005) and purified as for the native protein.

Strains Used

A364a background

Fig 5A	
8127-7-4	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂*
8824-1-1	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spn1-R263D(TRP1, +49)*
8791-13-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spt6-F249K(KanMX, -424)*
8834-3-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spn1-R263D(TRP1, +49) spt6-F249K(KanMX, -424)*
8824-1-2	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spn1-F267E(TRP1, +49)*
8768	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spn1-K192N*
Fig 5B	
8127-7-4	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂*
8762-5-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spt6-F249K(URA3, -424)*
Fig 5C	
8127-7-4	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂*
8762-5-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spt6-F249K(URA3, -424)*
8359-2-4	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3)*
8772-3-4	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3) spt6-F249K(URA3, -424)*
Fig S4B	
8151-1-1	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂*
8824-4-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spn1-F267E(TRP1, +49)*
8359-2-4	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3)*
8838-1-4	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3) spn1-F267E(TRP1, +49)*
Fig S4C	
8151-1-1	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂*
8824-6-1	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spn1-R263D(TRP1, +49)*
8824-5-2	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spn1-R263A(TRP1, +49)*
8791-9-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spt6-F249K(KanMX, -424)*
8763-7-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spt6-I248K, F249K(URA3, -424)*
8834-2-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spn1-R263D(TRP1, +49) spt6-F249K(KanMX, -424)*
8359-2-4	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3)*
8835-7-3	*MATa ura3 leu2 trp1 his3 lys2-128∂ nhp6a-Δ(::KanMX) nhp6b-Δ(::HIS3) spn1-R263D(TRP1, +49)*
Fig S4D	
8839-6-2	*MATα ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ pLK04 (YCp URA3 SPT6)*

8839-2-3	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spt6-F249K(KanMX, -424) pLK04 (YCp URA3 SPT6)*
8839-6-4	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his3 lys2-128∂ spn1-F267E(TRP1, +49) spt6-F249K(KanMX, -424) pLK04 (YCp URA3 SPT6)*
8791-6-4	*MATa ura3-Δ0 leu2-Δ0 trp1-Δ2 his7 lys2-128∂ spn1-K192N spt6-F249K(KanMX, -424) pLK04 (YCp URA3 SPT6)*

S288C background

Fig S4A	
8165-12-2	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂*
8165-1-1	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2 FLAG-spt6-1004*
8165-2-4	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂ nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*
8165-8-4	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂ FLAG-spt6-1004 nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*
8164-1-1	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂*
8164-3-3	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂ spt6-50*
8164-9-2	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂ nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*
8164-10-4	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂ spt6-50 nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*
8142-1-2	*MATa leu2 trp1-Δ63 ura3 his4-912∂ lys2-128∂*
8142-8-1	*MATa leu2 ura3 his4-912∂ spt6-14*
8142-7-2	*MATa leu2 ura3 his4-912∂ lys2-128∂ nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*
8142-1-4	*MATa leu2 ura3 his4-912∂ lys2 spt6-14 nhp6a-Δ(::URA3) nhp6b-Δ(::LEU2)*

Strain Construction

The *URA3* gene was inserted 424 bp upstream of the *SPT6* ORF, which is 60 bp downstream of *DAM1*. This site was chosen to minimize potential interference with either of these essential genes, and the insertion itself caused no detectable phenotype in any of our tests. Genomic DNA from this strain was then used as the template in a PCR amplification using one wild type primer 343 bp upstream of the insertion and one mutant primer downstream of the insertion containing the F249K mutation followed by 31 nucleotides of WT sequence. This PCR product was used to transform a WT strain. Ura+ transformants were screened for integration of the F249K mutation, indicating recombination between the PCR product and the genomic DNA in the 31 bp interval between the mutation and the end of the product. The transformation was performed in a diploid strain, which was then sporulated to derive haploids. All Ura+ derivatives displayed a severe growth defect initially in the A364a genetic background (but not in the S288C background), but this became more subtle during subsequent vegetative growth. This does not appear to be due to acquisition of a suppressor mutation, as only a mild growth defect was observed segregating 2:2 with the *URA3* marker in subsequent crosses.

Similarly, the *TRP1* gene was integrated 49 bp downstream of the *SPN1* ORF (which is also 161 bp downstream of the convergently transcribed gene *RPS23B*) without any detected phenotype, and genomic DNA from this strain was used as the template to transfer the *TRP1* gene along with the R263D and F267E mutations to a fresh diploid strain. After confirming the integration of the mutation by DNA sequencing, the diploids were sporulated to produce haploid strains with the desired mutation linked to the *TRP1* marker.

The *spn1-K192N* mutation was introduced into an integrating plasmid carrying the *SPN1* and *URA3* genes (described below), and this plasmid was integrated into a haploid yeast strain to

create a duplication of the *SPN1* locus. Recombinants that popped out the plasmid were obtained by selection on medium containing 5-FOA (Boeke et al., 1987), then these were screened for temperature sensitivity. Successful mutation of the *SPN1* gene without other changes was confirmed by DNA sequencing.

Other strains were derived by standard crosses within isogenic backgrounds.

Plasmid Construction

The integrating vector YIplac211 (Gietz and Sugino, 1988) was digested with BamHI and SacI, then a PCR product containing the *SPN1* ORF along with 332 bp of sequence upstream and 50 bp downstream was inserted. The WT sequence was verified by sequencing, and variants with the K192N and R263D mutations were derived using the Quikchange strategy (Stratagene). The mutated plasmids were digested with KpnI to target integration of the K192N mutation to the *SPN1* locus. The inserts from these plasmids were transferred to the high copy vector YEplac112 (Gietz and Sugino, 1988) for the experiment shown in Figure 5B.

pLK04 containing the *SPT6* gene was constructed by inserting DNA fragments flanking the *SPT6* locus into a low copy vector, then using gap repair from a WT strain to fill in the remainder of the gene. This resulted in a plasmid with 495 bp upstream (including the final 11 bp of the adjacent *DAM1* locus to insure inclusion of the entire intergenic region) and 161 bp downstream of the *SPT6* ORF. The sequence was confirmed by DNA sequencing; pLK04 complemented the lethality of an *spt6*-Δ deletion and the temperature sensitivity of all *spt6* alleles tested, but either *spt6* mutant or *SPT6* WT strains with this low copy version of *SPT6* displayed the Spt$^-$ phenotype. Expression of *SPT6* from the plasmid context is therefore aberrant and is sufficient to cause a defect in regulation of transcription.

Testing the effect of combining *spn1* and *spt6* mutations

Isogenic strains in the A364a genetic background with the *spn1-K192N* and *spt6-F249K* alleles were mated and the diploid was transformed with low copy plasmids with the *URA3* gene and either the *SPT6* or *SPN1* genes. Haploid segregants with both mutations were obtained after sporulation in both cases but were unable to grow on medium containing 5-FOA, indicating that either single mutation is tolerable but the combination of *spn1-K192N* and *spt6-F249K* is lethal. Similar tests also showed that *spn1-K192N* is lethal when combined with either *spt6-1004* or *spt6-F249K* in the S288C background.

SUPPLEMENTAL REFERENCES

Becker, R., Loll, B., and Meinhart, A. (2008). Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. J Biol Chem *283*, 22659-22669.

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol *154*, 164-175.

Gietz, R.D., and Sugino, A. (1988). New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. Gene *74*, 527-534.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res *31*, 3320-3323.

Holm, L., and Park, J. (2000). DaliLite workbench for protein structure comparison. Bioinformatics *16*, 566-567.

Kaplan, C.D., Holland, M.J., and 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.

Meinhart, A., and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. Nature *430*, 223-226.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif *41*, 207-234.

Zhang, L., Fletcher, A.G., Cheung, V., Winston, F., and Stargell, L.A. (2008). Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. Mol Cell Biol *28*, 1393-1403.

CHAPTER 4

MECHANISM OF HISTONE CHAPERONE ACTIVITY OF SPT6 AND ITS REGULATION BY SPN1

Summary

Spt6 is a histone chaperone that binds and reassembles nucleosomes in the wake of elongating RNAPII. This activity is essential to restore the default repressive chromatin state and to prevent inappropriate transcription initiation. Nucleosome binding by Spt6 requires a region of the protein that overlaps with the Spn1-binding site, and the Spn1-Spt6 interaction antagonizes Spt6-nucleosome binding *in vitro*. In an effort to further understanding of the molecular details of the histone chaperone activity of Spt6, we have performed a number of direct binding and competition assays. These data reveal that Spt6 is able to bind both histone H3-H4 tetramers [$(H3\text{-}H4)_2$] and H2A-H2B dimers, although the mode of interaction between Spt6 and each of the two histone subcomplexes appears to be different. The interaction with $(H3\text{-}H4)_2$ is primarily hydrophobic in nature while the interaction with H2A-H2B is largely electrostatic. Competition experiments demonstrated that the Spn1-Spt6 and Spt6-$(H3\text{-}H4)_2$ interactions are mutually exclusive. We further demonstrated that the Spn1-binding-motif of

Spt6 (Spt6-SBM, residues 239-268) is necessary and sufficient for histone binding, that Spt6 recognizes the ordered histone core of both subcomplexes, and that Spt6 competes with DNA for histone binding. Overall, the data presented in this chapter strongly support our previous hypothesis that Spn1 plays a regulatory role in the histone chaperone activity of Spt6, and suggest a mechanism by which Spt6 chaperones histones by direct competition with DNA.

Introduction

Eukaryotic organisms have evolved an effective DNA packaging system, collectively known as chromatin, which consists of four core histone proteins (H2A, H2B, H3 and H4) as well as linker histones and other factors (Li and Reinberg, 2011). This system allows organisms to compact their large genomes and fit them into the nucleus of each cell. It is not simply enough to package efficiently, however, because much of the DNA needs to be accessed on a regular basis, and the packaging therefore must be dynamic. For example, every time a cell divides the entire genome must be replicated. Similarly, the synthesis of RNA and protein molecules that drive cellular processes requires accessing, processing, and repackaging relevant segments of the genome. Thus, packaging, accessing, and repackaging of genomic DNA is intimately linked to biological function and is highly regulated.

The fundamental DNA packaging unit in eukaryotic cells is called the nucleosome, which consists of nearly 150 base pairs of DNA wrapped twice

around a protein core made of two copies each of the four core histone proteins (Luger et al., 1997). Individual nucleosomes are separated by a short DNA linker, which allows the nucleosomes to further associate with protein partners and wind around each other to establish a higher level of compaction (Li and Reinberg, 2011; Saunders et al., 2006).

In a simplistic transcription model, there are two different classes of proteins or protein complexes that mediate access to DNA by RNA polymerase II (RNAPII). In one case, enzymatic protein complexes called chromatin remodelers use the energy of ATP binding and hydrolysis to drive a motor-like ratchet system that can translocate DNA with respect to the nucleosomal histones or even eject histones completely (Cairns, 1998; Clapier and Cairns, 2009). In the other case, histone chaperone proteins use inherent binding energy to displace histones from intimate contact with the DNA and/or reassemble nuclesomes after the DNA has been accessed (Andrews et al., 2010; Winkler et al., 2011). The Spt6 protein falls into the latter category of histone chaperones.

The essential and conserved Spt6 protein is implicated in a multitude of chromatin-associated events including transcription initiation (Adkins and Tyler, 2006; Cheung et al., 2008; Kaplan et al., 2003), transcription elongation by RNA polymerase I (Beckouet et al., 2011) (RNAPI) and RNAPII (Ardehali et al., 2009; Endoh et al., 2004; Hartzog et al., 1998; Kaplan et al., 2005; Kaplan et al., 2000; Keegan et al., 2002; Lindstrom et al., 2003; Yoh et al., 2007), silencing of

heterochromatin (Kiely et al., 2011), class-switch recombination of Ig genes (Okazaki et al., 2011), various transcription-directed signaling pathways (Baniahmad et al., 1995; Kok et al., 2007; Li et al., 2010; Nishiwaki et al., 1993; Shen et al., 2009; Widiez et al., 2011), and mRNA maturation and export (Andrulis et al., 2002; Bucheli and Buratowski, 2005; Burckin et al., 2005; Kaplan et al., 2005; Yoh et al., 2007). Spt6 appears to accomplish these feats through direct and indirect binding events with an array of factors, including RNAPI (Beckouet et al., 2011) and RNAPII (Diebold et al., 2010; Lindstrom et al., 2003; Sun et al., 2010; Yoh et al., 2007; Yoh et al., 2008), histones H3 and H4 (Bortvin and Winston, 1996), nucleosomes (McDonald et al., 2010), DNA (Close et al., 2011), and various transcription factors (Andrulis et al., 2002; Gavin et al., 2002; Lindstrom et al., 2003). The *Saccharomyces cerevisiae* Spt6 protein is large (~168 kDa) and folds into six recognizable structural domains with homology to known DNA binding and protein-protein interaction motifs, and has an acidic, disordered N-terminal region of ~300 amino acid residues (Close et al., 2011). This makes it easy to imagine Spt6 serving as a docking platform in elongation complexes, where its multiple domains are responsible for recruiting a variety of factors and binding RNAPII. The inherent disorder of the acidic N-terminal region of Spt6, including the segment required for binding Spn1, may provide the necessary flexibility required for function in a crowded transcriptional environment. Delineation of the intricacies of the various Spt6 interaction

networks and correlation with the numerous regulatory and transitional steps in gene expression is an important remaining obstacle in this field.

Nucleosome reassembly mediated by Spt6 has profound regulatory effects at both intergenic and intragenic start sites (Cheung et al., 2008; Kaplan et al., 2003) and is likely to be a result of direct interactions with histones and perhaps DNA. *In vitro*, Spt6 promotes histone deposition on plasmid DNA, and directly binds DNA, nucleosomes, and free histones with a preference for binding the histone H3-H4 tetramer (Bortvin and Winston, 1996; Close et al., 2011; McDonald et al., 2010). Our previous biochemical and genetic studies on this process (see Chapter 3) have found that the ability of Spt6 to bind nucleosomes is dependent on the small HMGB family member Nhp6, which is thought to bind DNA and weaken the interactions between the DNA and histones within a nucleosome (Stillman, 2010). The ability of the unrelated histone chaperone FACT to make complexes with nucleosomes is also dependent on Nhp6 (Formosa et al., 2001). This implies that the histone chaperone activity of Spt6 is likely to involve direct interactions with histones, that the interface may at least partially overlap with the histone-DNA interface, and that this may be a common mode of action for other histone chaperones. Nucleosome binding by Spt6 requires a region of the protein that overlaps with the Spn1-binding site, and the Spn1-Spt6 interaction antagonizes Spt6-nucleosome binding in vitro. Mutations in *SPT6* or *SPN1* that disrupt the Spn1-Spt6 interaction cause the Spt$^-$ phenotype, and the effects of the spt6-F249K allele are suppressed by

overexpressing Spn1. This shows that the Spn1-Spt6 interaction disrupts Spt6-nucleosome binding and that this disruption has a positive role in maintaining normal chromatin. An attractive explanation is that Spn1 is actively engaged in the nucleosome reassembly process, possibly by disengaging Spt6 from nucleosomes to allow multiple rounds of reassembly. Another possibility is that Spn1 binding influences the balance between Spt6's functional roles in nucleosome reassembly, transcription elongation, and mRNA processing. Regardless of the precise mechanistic details, Spn1 appears to be important for Spt6-mediated nucleosome reassembly in vivo.

In order to gain insight into the molecular details of the histone chaperone activity of Spt6, and to assess the potential regulatory role of Spn1 in this process, we utilized a number of biochemical binding and competition assays. The data presented in this chapter confirm that Spt6 binds directly to both histone H3-H4 tetramers [$(H3\text{-}H4)_2$] and H2A-H2B dimers, that Spn1 directly competes with histones for Spt6 binding, and that Spt6 competes directly with DNA for histone binding. Overall, these data strongly support our previous hypothesis that Spn1 plays a regulatory role in the histone chaperone activity of Spt6, and demonstrate that Spt6 utilizes a direct DNA-competition mechanism for chaperoning histones similar to other well-characterized histone chaperones Nap1 and FACT.

Results and Discussion

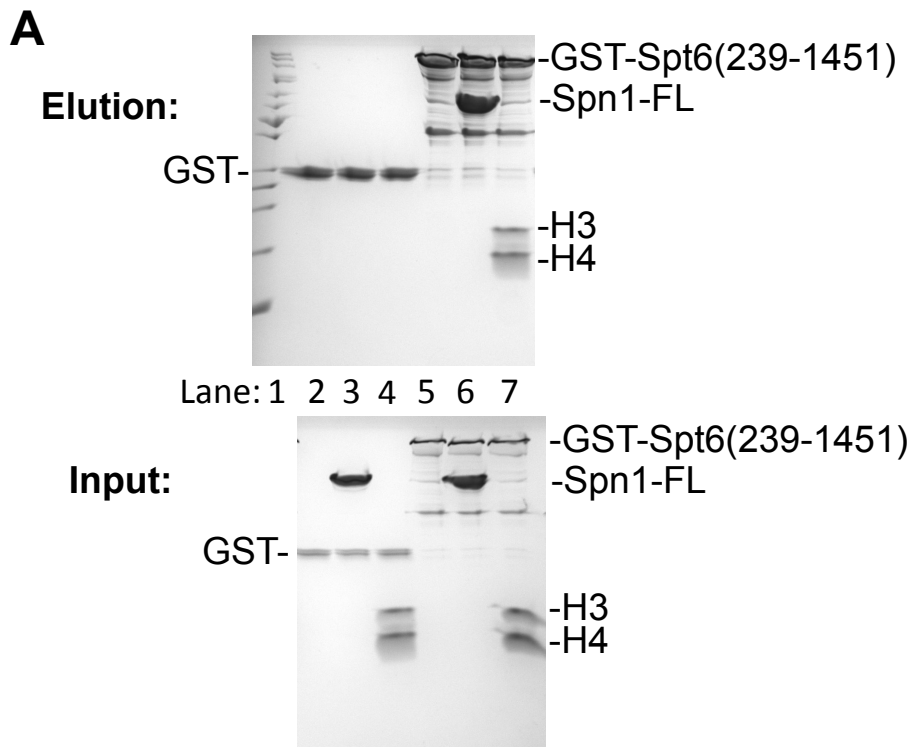
Spt6 binds both histones $(H3\text{-}H4)_2$ and H2A-H2B

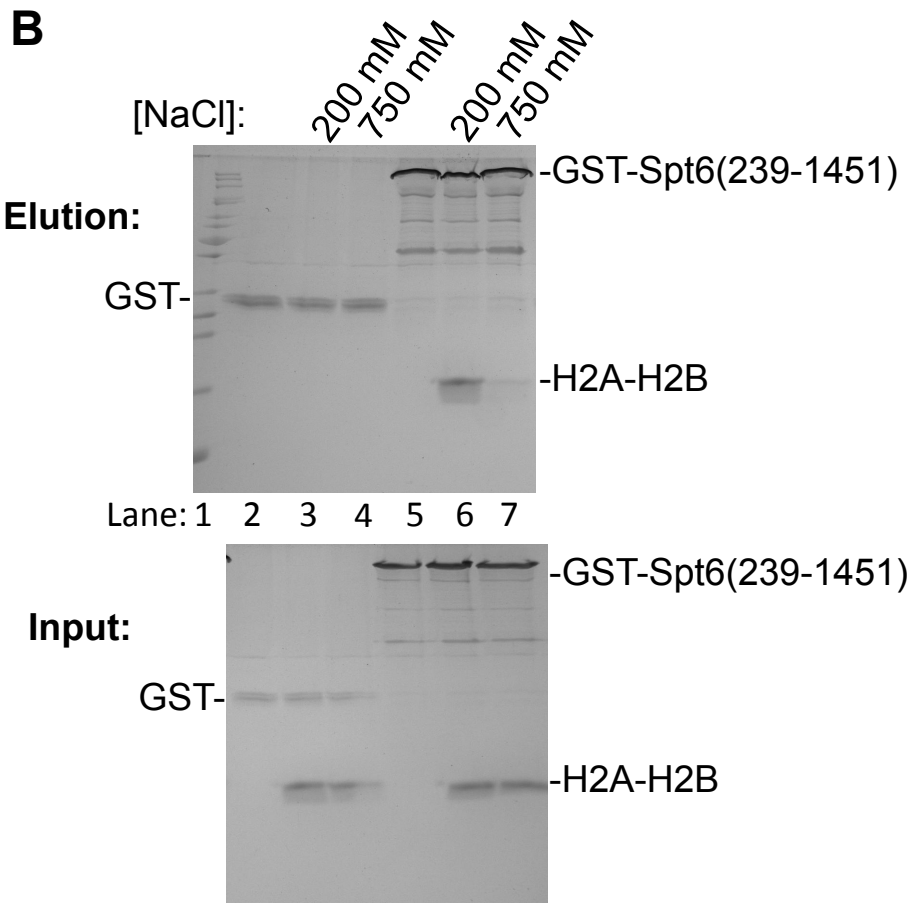
Full-length Spt6 protein with an N-terminal GST fusion does not express in bacterial cells; therefore, we utilized an N-terminal GST fusion to an N-terminal truncation of Spt6 (residues 239-1451), which retains nucleosome binding, to assess binding to the two different histone subcomplexes, histone H3-H4 tetramers [$(H3\text{-}H4)_2$] and H2A-H2B dimers. It is important to note that the different recombinant histone subcomplexes at concentrations typical of biochemical assays (1-50 μM) differ in their stability and solubility; H2A-H2B is tolerant of low salt buffers that mimic physiological conditions while $(H3\text{-}H4)_2$ begins to precipitate out of solution in buffers below 500 mM sodium chloride. We performed GST-pulldown assays and found that GST-Spt6(239-1451) binds both histone subcomplexes (Figure 4-1A and B, top panels are elutions from the GST beads after washing). Binding to $(H3\text{-}H4)_2$ was assessed under high salt conditions (750 mM NaCl) in order to prevent nonspecific binding (Figure 4-1A, lane 4), as well as to keep $(H3\text{-}H4)_2$ stable and in solution during the binding reactions. Binding of GST-Spt6(239-1451) to Spn1 was used as a positive control under the same high salt conditions (Figure 4-1A, lane 6). In Figure 4-1A, lanes 3 and 4 clearly show that neither $(H3\text{-}H4)_2$ or Spn1 bind nonspecifically to GST or the glutathione resin (compare top panel elutions to bottom panel inputs). GST-Spt6(239-1451) clearly bound both Spn1 and $(H3\text{-}H4)_2$ in separate reactions (Figure 4-1A, lanes 6 and 7, top panel), indicating that the Spn1-Spt6

Figure 4-1. Spt6 binds directly to both histone tetramers and dimers.

A) GST-pulldowns with GST-Spt6(239-1451) show direct binding to tetrameric histones $(H3\text{-}H4)_2$. Top panels are elution samples after washing with high salt (750 mM NaCl) binding buffer. Bottom panels are input samples of the reactions prior to addition of glutathione resin. Migration positions of the different proteins are indicated. Lane 1, protein standards; Lane2, GST alone; Lane 3, GST binding negative control with Spn1; Lane 4, GST binding negative control with $(H3\text{-}H4)_2$; Lane 5, GST-Spt6(239-1451) alone; Lane 6, GST-Spt6 binding positive control to Spn1; Lane 7, GST-Spt6 binding to $(H3\text{-}H4)_2$.

B) As in (A), GST-pulldowns with GST-Spt6(239-1451) show direct binding to dimeric histones H2A-H2B. Lane 1, protein standards; Lane2, GST alone; Lane 3, GST binding negative control with H2A-H2B in low salt buffer; Lane 4, GST binding negative control with H2A-H2B in high salt buffer; Lane 5, GST-Spt6(239-1451) alone; Lane 6, GST-Spt6 binding to H2A-H2B in low salt buffer; Lane 7, GST-Spt6 binding to H2A-H2B in high salt buffer.


A
Elution:
GST-
-GST-Spt6(239-1451)
-Spn1-FL
-H3
-H4
Lane: 1 2 3 4 5 6 7
Input:
GST-
-GST-Spt6(239-1451)
-Spn1-FL
-H3
-H4


B
[NaCl]:
200 mM
750 mM
200 mM
750 mM
Elution:
GST-
-GST-Spt6(239-1451)
-H2A-H2B
Lane: 1 2 3 4 5 6 7
Input:
GST-
-GST-Spt6(239-1451)
-H2A-H2B

interaction and the Spt6-$(H3\text{-}H4)_2$ interaction are both highly specific, robust, and strongly hydrophobic. Histone H2A-H2B dimers showed no nonspecific background binding to free GST or the glutathione resin in low salt (200 mM NaCl) or high salt buffers (Figure 4-1B, lanes 3 and 4). GST-Spt6(239-1451) bound H2A-H2B in low salt buffer; however, the interaction was severely perturbed in the high salt buffer (Figure 4-1B, compare lanes 6 and 7, top panel), indicating the interaction is primarily of an electrostatic nature. These data support a model in which Spt6 can chaperone both histone subcomplexes, but indicate that Spt6 utilizes a different mode of interaction with each subcomplex, perhaps providing some specificity in nucleosome reassembly during transcription elongation. Consistent with this, Spt6 also promotes deposition of histones onto plasmid DNA (Bortvin and Winston, 1996), therefore, Spt6 may be capable of chaperoning both histone subcomplexes. Further, other histone chaperones such as Nap1, Vps75, and FACT have been found to bind and promote the DNA deposition of both histone subcomplexes (Andrews et al., 2010; Bowman et al., 2011; Formosa, 2012; Su et al., 2011), suggesting that a common mechanism may be used by different histone chaperones that function in the dynamic chromatin environment.

Spn1 directly competes with histones for Spt6 binding

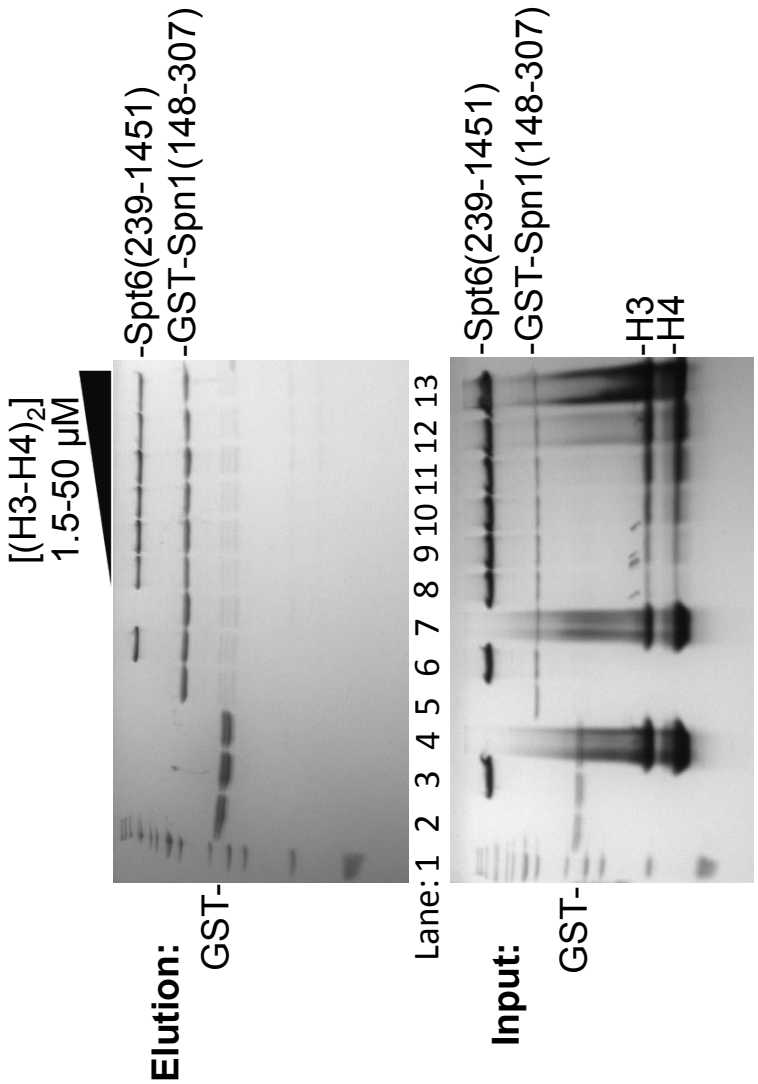
We have shown previously that an Spt6 construct with a deletion of the first 238 N-terminal residues (Spt6(239-1451)) can bind nucleosomes in the

presence of Nhp6 (McDonald et al., 2010). An Spt6 construct with a deletion of the first 314 N-terminal residues (Spt6(315-1451)) was unable to bind nucleosomes, suggesting that Spt6 may utilize a small region of its acidic N-terminus to bind both Spn1 and histones. Indeed, Spn1 blocked the nucleosome-Spt6 interaction in nucleosome binding assays.

We have extended this observation by performing a direct competition GST-pulldown assay between GST-labeled Spn1(148-307), Spt6(239-1451) and $(H3\text{-}H4)_2$ (Figure 4-2). In this assay, a high salt buffer (750 mM NaCl) was again used to minimize nonspecific binding of histones to GST and the glutathione resin and maintain the stability and solubility of $(H3\text{-}H4)_2$ (Figure 4-2, lane 4). In the competition reactions, the concentrations of GST-Spn(148-307) and Spt6(239-1451) are maintained at a constant value of 5 μM, while the concentration of $(H3\text{-}H4)_2$ was increased in 2-fold increments from 1.5 μM to 50 μM in lanes 8 through 13. There were no nonspecific interactions observed between GST or the glutathione resin and either Spt6(239-1451) or the highest concentrations of histones (Figure 4-2, lanes 3 and 4). GST-Spn1(148-307) bound Spt6 in each reaction of this assay (Figure 4-2, lanes 6 and 8 through 13), but failed to make a ternary complex with Spt6 and histones (Figure 4-2, lanes 8 through 13), even at the highest histone concentration tested, which was 10-fold greater than the Spt6 and Spn1 concentrations (Figure 4-2, lane 13, 50 μM histones vs. 5 μM Spt6 and Spn1). Spn1 appears to prevent the histones from binding Spt6 in this assay, which strongly implies that the Spn1-Spt6 interaction

Figure 4-2. Spn1 blocks histones $(H3\text{-}H4)_2$ from binding Spt6

GST-pulldowns were performed as in figure 4-1A (high salt buffer) using GST-Spn1(148-307) and untagged Spt6(239-1451) and $(H3\text{-}H4)_2$. Lanes 1 through 5 are controls as in figure 4-1A; Lane 6, GST-Spn1 positive control binding to Spt6(239-1451); Lane 7, GST-Spn1 negative control binding to $(H3\text{-}H4)_2$; Lanes 8-13; GST-Spn1 binding to Spt6 in the presence of 2-fold incremental increases in the concentration of $(H3\text{-}H4)_2$.


[(H3-H4)2]
1.5-50 µM
-Spt6(239-1451)
-GST-Spn1(148-307)
Elution:
GST-
Lane: 1 2 3 4 5 6 7 8 9 10 11 12 13
-Spt6(239-1451)
-GST-Spn1(148-307)
-H3
-H4
Input:
GST-

and the Spt6-$(H3\text{-}H4)_2$ interaction are mutually exclusive. This also indicates the Spt6 likely utilizes its Spn1-binding-motif (SBM, residues 239-268) to bind both histones and Spn1. These data provide strong supportive evidence that Spn1 does indeed play a regulatory role in the histone chaperone activity of Spt6 by directly competing for Spt6 binding with histones.

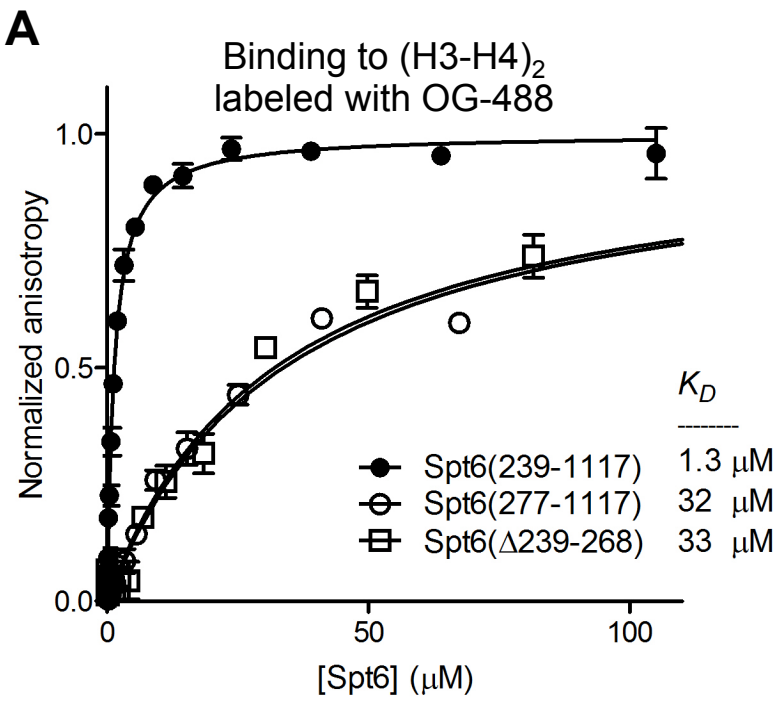
The Spt6-SBM is necessary and sufficient for binding $(H3\text{-}H4)_2$

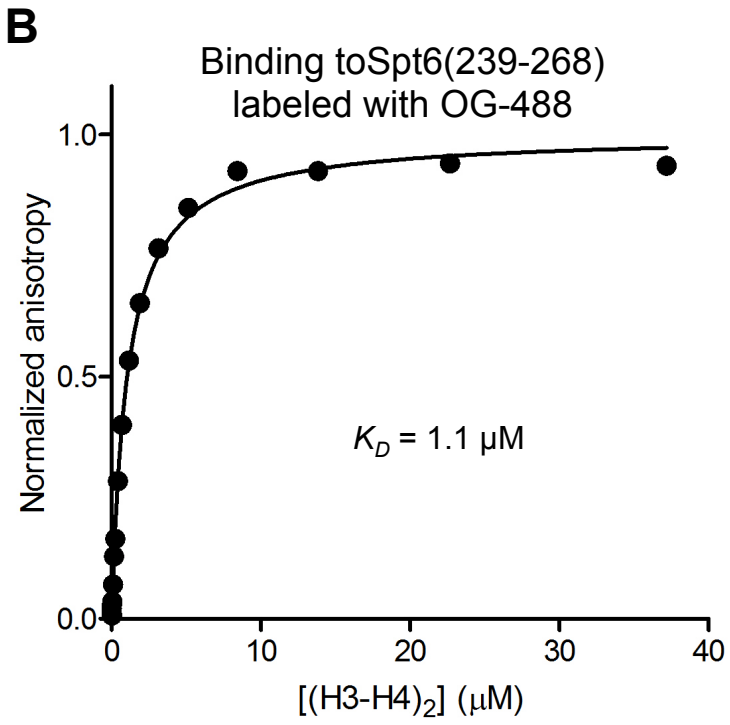
In order to further address the importance of the Spt6-SBM in binding histones, we determined that the Spt6-SBM was necessary for binding histones using three different Spt6 constructs (Figure 4-3A and B). Two of these constructs were N-terminal truncations that removed either the first 238 residues (residues 239-1117, keeping the SBM intact), or the first 276 residues of the N-terminus (residues 277-1117, removing the SBM). The other was a deletion construct within the context of the full-length protein, Spt6(Δ239-268), which specifically deletes only the SBM. We were unable to test binding to the WT full-length Spt6 protein due to lack of solubility at the high concentrations needed for these assays. We measured the equilibrium binding constants, K_D, between each Spt6 construct and fluorescently labeled $(H3\text{-}H4)_2$ using a slightly modified version of the fluorescence assay developed in the Luger laboratory. Instead of just measuring a change in fluorescence intensity upon titration of the unlabeled binding partner, we measured parallel and perpendicular intensities and calculated the change in anisotropy (LiCata and Wowor, 2008) to quantify binding

Figure 4-3. The spn1-binding-motif (SBM, residues 239-268) of Spt6 is necessary and sufficient to bind histones $(H3\text{-}H4)_2$.

A) Fluorescence anisotropy measurements under high salt conditions (750 mM NaCl) indicate binding between Spt6 and fluorescently labeled $(H3\text{-}H4)_2$ requires the presence of the Spt6-SBM. Binding isotherms are shown for each Spt6 protein construct tested, Spt6(239-1117) closed circles, Spt6(277-1117) open circles, and Spt6(Δ239-269) open squares. Error bars are ± standard error from at least three independent measurements. Affinity values are indicated.

B) As in (A), fluorescence anisotropy measurements with fluorescently labeled Spt6-SBM (Spt6(239-268)) binding to unlabeled $(H3\text{-}H4)_2$ indicate that the Spt6-SBM is suffucuent to recapitulate the binding energy of larger Spt6 constructs.


A
Binding to (H3-H4)2
labeled with OG-488
Normalized anisotropy
1.0
0.5
0.0
0
50
100
[Spt6] (μM)
KD
Spt6(239-1117) 1.3 μM
Spt6(277-1117) 32 μM
Spt6(Δ239-268) 33 μM


B
Binding toSpt6(239-268)
labeled with OG-488
Normalized anisotropy
1.0
0.5
0.0
0
10
20
30
40
[(H3-H4)2] (μM)
KD = 1.1 μM

of Spt6 to histones. Again, a high salt (750 mM NaCl) buffer was used to prevent any nonspecific histone interactions and to maintain the stability and solubility of the histones. Binding between the two Spt6 constructs lacking the SBM (Spt6(277-1117) and Spt6(Δ239-268)) and $(H3\text{-}H4)_2$ was reduced at least 20-fold compared to the Spt6 construct that contained the SBM (Spt6(239-1117), figure 4-3A, ~ 30 μM K_D vs. 1.3 μM K_D). These data show that the Spt6-SBM is necessary to retain full binding to $(H3\text{-}H4)_2$.

We next sought to address whether the Spt6-SBM is sufficient for binding histones by fluorescently labeling recombinant Spt6-SBM and measuring binding affinity to unlabeled histones (Figure 4-3B). In these assays (high salt buffer), the binding affinity between Spt6-SBM and $(H3\text{-}H4)_2$ was essentially identical to the affinity between Spt6(239-1117) and $(H3\text{-}H4)_2$ (K_D = 1.1 μM vs. 1.3 μM for Spt6(239-1117)). These data indicate that the Spt6-SBM is sufficient to recapitulate the binding energy of larger Spt6 constructs. Taken together, this provides strong evidence that the Spt6-SBM is required for histone binding, and is consistent with a model in which Spn1 binding influences the nucleosome reassembly activity of Spt6.

Spt6 binds the core ordered region of both histone subcomplexes

Each of the four core histone proteins has flexible N-terminal tails that are targets for a number of posttranslational modifications (Luger and Richmond, 1998; Strahl and Allis, 2000). As discussed in Chapter 1, these posttranslation

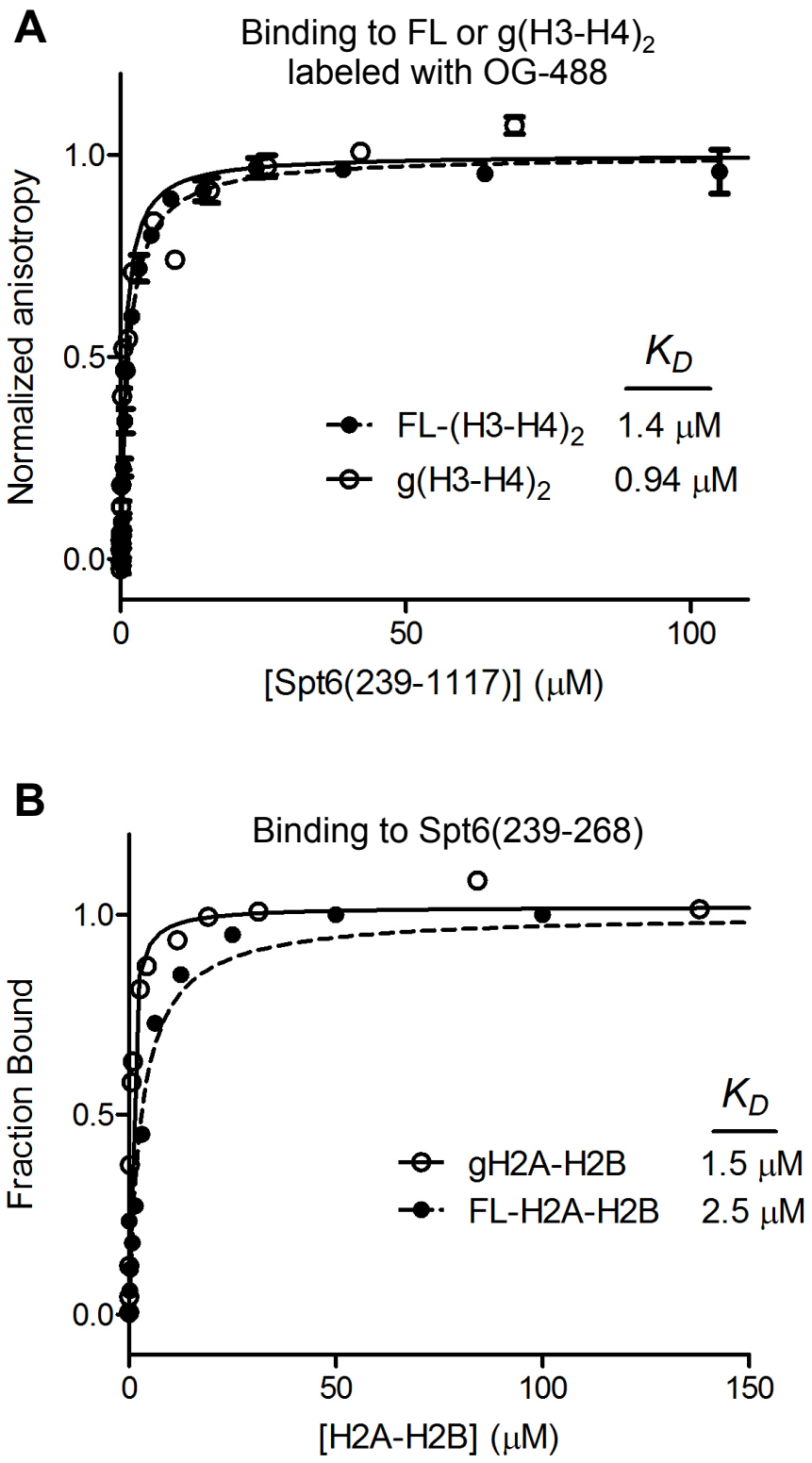
modifications are a result of various intracellular signaling pathways that ultimately provide a code determining the state of any given gene. Certain modification combinations can yield an actively transcribed gene, while others suppress transcription (Strahl and Allis, 2000). A number of protein domains are known to specifically recognize different chemical modifications on amino-acid sidechains, and thus, the code can be discerned and the signal relayed to the transcription machinery (Jenuwein and Allis, 2001; Ruthenburg et al., 2007; Taverna et al., 2007). We were curious to know if the histone tails contributed to the interaction with Spt6, although primary sequence and secondary structural analyses did not indicate the presence of any known histone modification recognition domains (Maclennan and Shaw, 1993; McGuffin et al., 2000).

Tailless versions of all four histones were made that included H2A residues 14-118, H2B residues 24-122, H3 residues 27-135, and H4 residues 20-102, and we assessed binding to Spt6 using the fluorescence anisotropy assay for $(H3\text{-}H4)_2$ (Figure 4-4A) and a native gel shift binding assay for H2A-H2B (Figure 4-4B). The tailless, or globular, H3-H4 tetramers [$g(H3\text{-}H4)_2$] were fluorescently labeled as for the WT proteins, and unlabeled Spt6(239-1117) was titrated, with the reactions again performed in the high salt buffer (Figure 4-4A). The measured K_D was 0.94 μM, which is nearly identical to the K_D measured for full-length histone proteins (1.4 μM). These results indicate that the tails of $(H3\text{-}H4)_2$ do not contribute to the binding of Spt6.

Figure 4-4. Histone tails do not contribute to the interaction with Spt6.

A) Binding isotherms derived from fluorescence anisotropy measurements in high salt conditions (750 mM NaCl) indicate no loss of binding affinity between Spt6 and fluorescently labeled 'tailless' or g(H3-H4)$_2$ (open circles) compared to FL-(H3-H4)$_2$ (closed circles, dashed line). Error bars are ± standard error from at least three independent measurements. Affinity values are indicated.

B) Binding isotherms and affinity estimations from EMSA assays with Spt6-SBM and H2A-H2B (full-length, closed circles, dashed line; 'tailless' or gH2A-H2B, open circles) show no loss of binding affinity upon tail removal. Affinity values are indicated.


A
Binding to FL or g(H3-H4)2
labeled with OG-488
Normalized anisotropy
1.0
0.5
0.0
0
50
100
[Spt6(239-1117)] (μM)
KD
FL-(H3-H4)2 1.4 μM
g(H3-H4)2 0.94 μM
B
Binding to Spt6(239-268)
Fraction Bound
1.0
0.5
0.0
0
50
100
150
[H2A-H2B] (μM)
KD
gH2A-H2B 1.5 μM
FL-H2A-H2B 2.5 μM

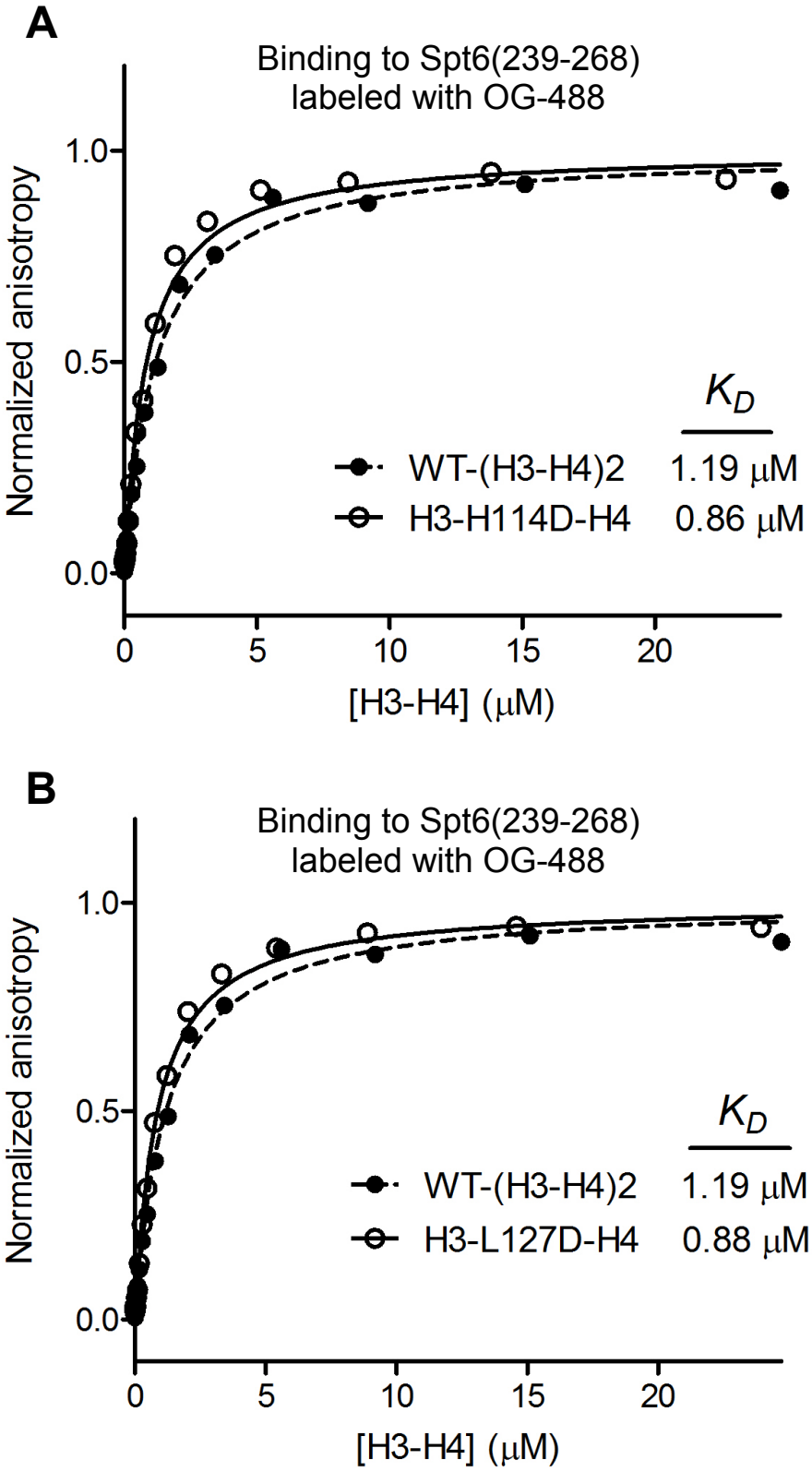
Binding to tailless, or globular, H2A-H2B (gH2A-H2B) was assessed in low salt buffer (Figure 4-4B). We attempted to use the fluorescence anisotropy assay; however, the results did not correlate with other quantitative binding assays using identical buffers. We therefore used an electrophoretic mobility shift assay (EMSA) because the measured K_D values using this assay with full-length H2A-H2B correlated well with those from isothermal titration calorimetry (ITC, see below) under identical conditions. It is important to note that results from the fluorescence anisotropy assay with $(H3\text{-}H4)_2$ using the high salt buffer did correlate precisely with results from ITC experiments under identical buffer conditions (see below). The measured K_D value for the interaction between Spt6 and gH2A-H2B in the EMSA assay was 1.5 μM (Figure 4-4B, open circles), which is only marginally tighter than the K_D measured for full-length H2A-H2B using the same assay and conditions (2.5 μM, Figure 4-4B, closed circles). We conclude that histone tails do not contribute to binding Spt6 and that this may be a necessary physiological requirement if Spt6 is needed to chaperone histones and reassemble nucleosomes that potentially have an assortment of chemical modifications. One assumption behind this reasoning is that cells probably do not want to expend the energy to reestablish the chemical modification code for each gene after each passage of the transcription machinery. Chaperones, therefore, likely reassemble each nucleosome with the same histones that were evicted prior to RNAPII passage, thereby maintaining the chemical code.

Spt6 does not bind the H3:H3 tertramerization interface

Structures of an unrelated histone chaperone called Asf1 were solved in complex with histones H3-H4 (Antczak et al., 2006; English et al., 2006; Natsume et al., 2007), revealing details of how other chaperones may interact with histones. The most striking finding was that Asf1 bound H3-H4 at the H3:H3 interface, thereby preventing tetramer formation and maintaining H3-H4 in a dimeric state. The physiological significance of dimeric H3-H4 and the potential impact on proposed steps in the nucleosome (re)assembly pathway remain unclear; however, we wanted to assess whether Spt6 bound to H3-H4 in the same fashion as Asf1. We made separate mutations in histone H3 (H114D and L127D) that prevent tetramer formation as assessed by size-exclusion chromatography (data not shown). We used the fluorescence anisotropy assay to measure the K_D values for the interaction between labeled Spt6-SBM and unlabeled dimeric H3-H4 (Figure 4-5A and B). These results indicate that Spt6 does not bind to the H3:H3 interface, as these mutations had no effect on binding affinity between Spt6 and H3-H4 as compared to WT $(H3\text{-}H4)_2$ (WT K_D = 1.3 μM vs. ~ 0.9 μM for the H3 mutants). We conclude that although Asf1 may share a common function of chaperoning histones, Spt6 accomplishes its function by binding histones in a different manner. Taken together, these results imply that Spt6 recognizes the ordered core region of the two different histone subcomplexes.

Figure 4-5. Spt6 does not bind the H3-H3 tetramer interface.

Binding isotherms derived from fluorescence anisotropy measurements of the interactions of two separate H3 mutants that prevent tetramer formation in solution (A, H3-H114D; B, H3-L127D) and fluorescently labeled Spt6-SBM. Affinity estimates are indicated.


A
Binding to Spt6(239-268)
labeled with OG-488
Normalized anisotropy
1.0
0.5
0.0
0
5
10
15
20
[H3-H4] (μM)
K_D
WT-(H3-H4)2 1.19 μM
H3-H114D-H4 0.86 μM
B
Binding to Spt6(239-268)
labeled with OG-488
Normalized anisotropy
1.0
0.5
0.0
0
5
10
15
20
[H3-H4] (μM)
K_D
WT-(H3-H4)2 1.19 μM
H3-L127D-H4 0.88 μM

Spt6 binds both histone subcomplexes with a 1:1 stoichiometry

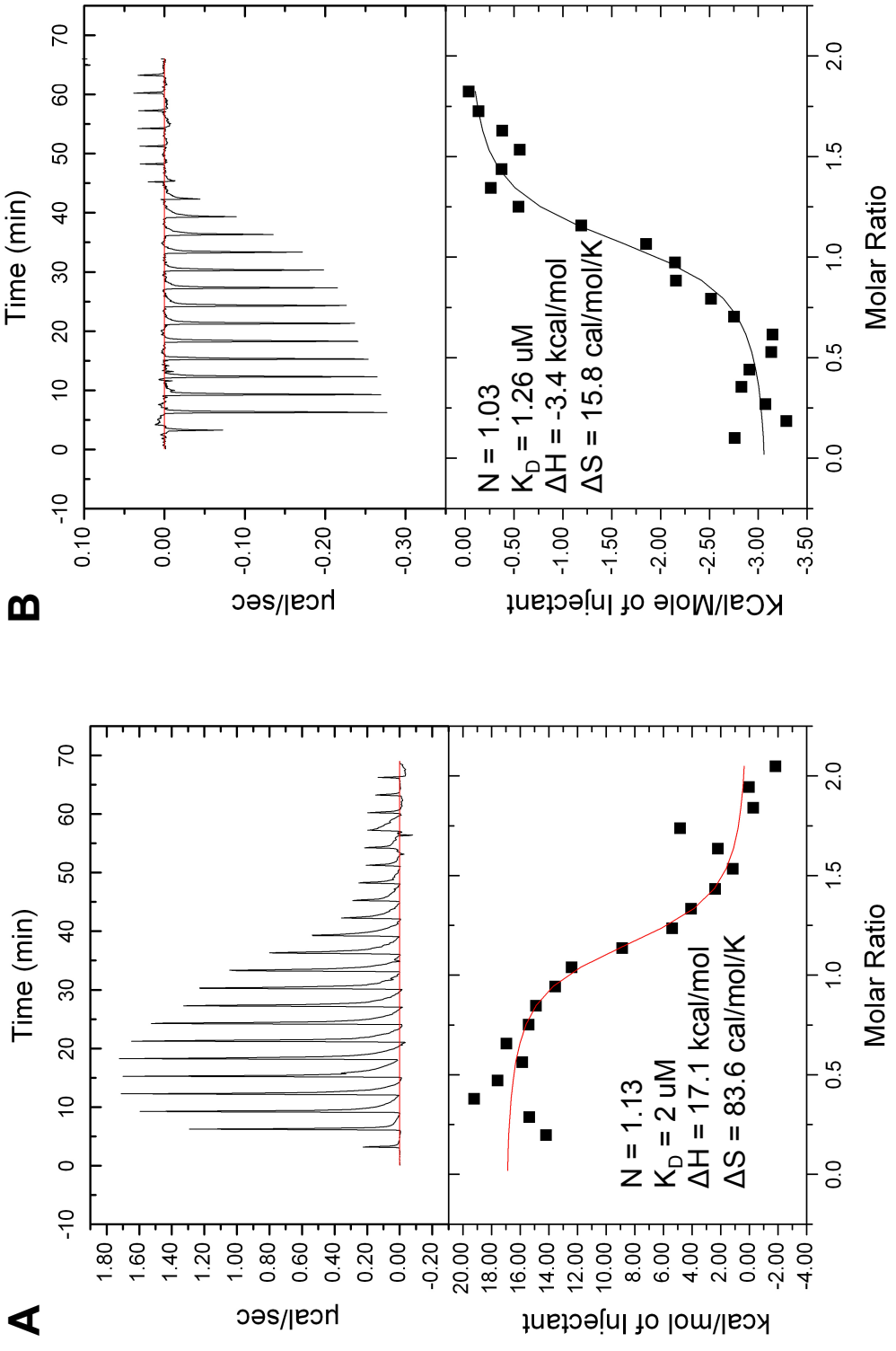
We assessed the stoichiometry of the Spt6-$(H3\text{-}H4)_2$ interaction and the Spt6-H2A-H2B interaction using ITC. This quantitative binding assay determines the thermodynamic nature of an interaction, and reveals stoichiometries by measuring the heat released (or absorbed) in any amenable protein-protein binding reaction (Velazquez-Campoy et al., 2004). Stoichiometry is measured by plotting the molar ratio of the protein being titrated into a cell containing the other binding partner. This ratio, when plotted against the heat change of each titration point gives a sigmoidal relationship, which when fitted with a nonlinear regression curve allows for the measurement of not only the equilibrium binding constant and stoichiometry, but also the overall change in enthalpy of the reaction, allowing the calculation of entropic contributions to the binding reaction. Calorimetry is therefore the most powerful and comprehensive assay for measuring the thermodynamics of a binding reaction, and also benefits from the fact that no labels or tags need be included in the protein samples.

Using this assay, we measured the K_D and stoichiometries of the reaction between Spt6(239-489) and $(H3\text{-}H4)_2$ under high salt conditions (Figure 4-6A), and the reaction between Spt6-SBM and H2A-H2B under low salt conditions (Figure 4-6B). The measured K_D values in the ITC experiments correlate very well with those measured in the high salt fluorescence anisotropy assay with $(H3\text{-}H4)_2$ (1.3 μM with anisotropy and 2.0 μM with ITC), and the low salt EMSA assay with H2A-H2B (2.1 μM with EMSA and 1.26 μM with ITC). The results of these

Figure 4-6. Spt6 binds with a 1:1 stoichiometry to histone tetramers and dimers.

A) ITC measurements in high salt buffer (750 mM NaCl) of the interaction between Spt6(239-489) and $(H3\text{-}H4)_2$. Raw ITC data are shown in the top panel and the binding isotherm in the bottom panel. Thermodynamic parameters derived from ITC measurements, including stoichiometry (*N*), are given.

B) As in (A), ITC measurements in low salt buffer (200 mM NaCl) of the interaction between Spt6-SBM and H2A-H2B.


A
Time (min)
µcal/sec
kcal/mol of Injectant
Molar Ratio
N = 1.13
K_D = 2 uM
ΔH = 17.1 kcal/mol
ΔS = 83.6 cal/mol/K
B
Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio
N = 1.03
K_D = 1.26 uM
ΔH = -3.4 kcal/mol
ΔS = 15.8 cal/mol/K

reactions indicate that Spt6 binds to dimeric H2A-H2B in a 1:1 stoichiometry (one Spt6 molecule per H2A-H2B dimer). Spt6 also binds to tetrameric $(H3\text{-}H4)_2$ with a 1:1 stoichiometry (one Spt6 molecule per $(H3\text{-}H4)_2$ tetramer). The result of the stoichiometry between Spt6 and $(H3\text{-}H4)_2$ is likely to be physiologically relevant even though the $(H3\text{-}H4)_2$ tetramer is symmetric, and there likely exists two identical binding sites for Spt6 on a single tetramer. Our preferred model is that the ordered core region of Spt6 prevents a second Spt6 molecule from interacting with the second binding site by steric occlusion.

The thermodynamic parameters of the Spt6-histone interactions in Figure 4-6 are consistent with what we would expect given the buffer conditions and the measured affinities. For the reactions with $(H3\text{-}H4)_2$, the relatively high enthalpy values and the endothermic heat absorption profile imply a strong hydrophobic interaction. This is expected given that we see binding in very high salt concentrations. The large entropic value is likely a result of displacement of ordered salt and water molecules from the histone surface upon Spt6 binding. Another potential explanation of a large entropic value is that the Spt6 peptide might be largely ordered in the high salt buffer. This would eliminate any decrease in entropy caused by the ordering of the peptide by binding to histones.

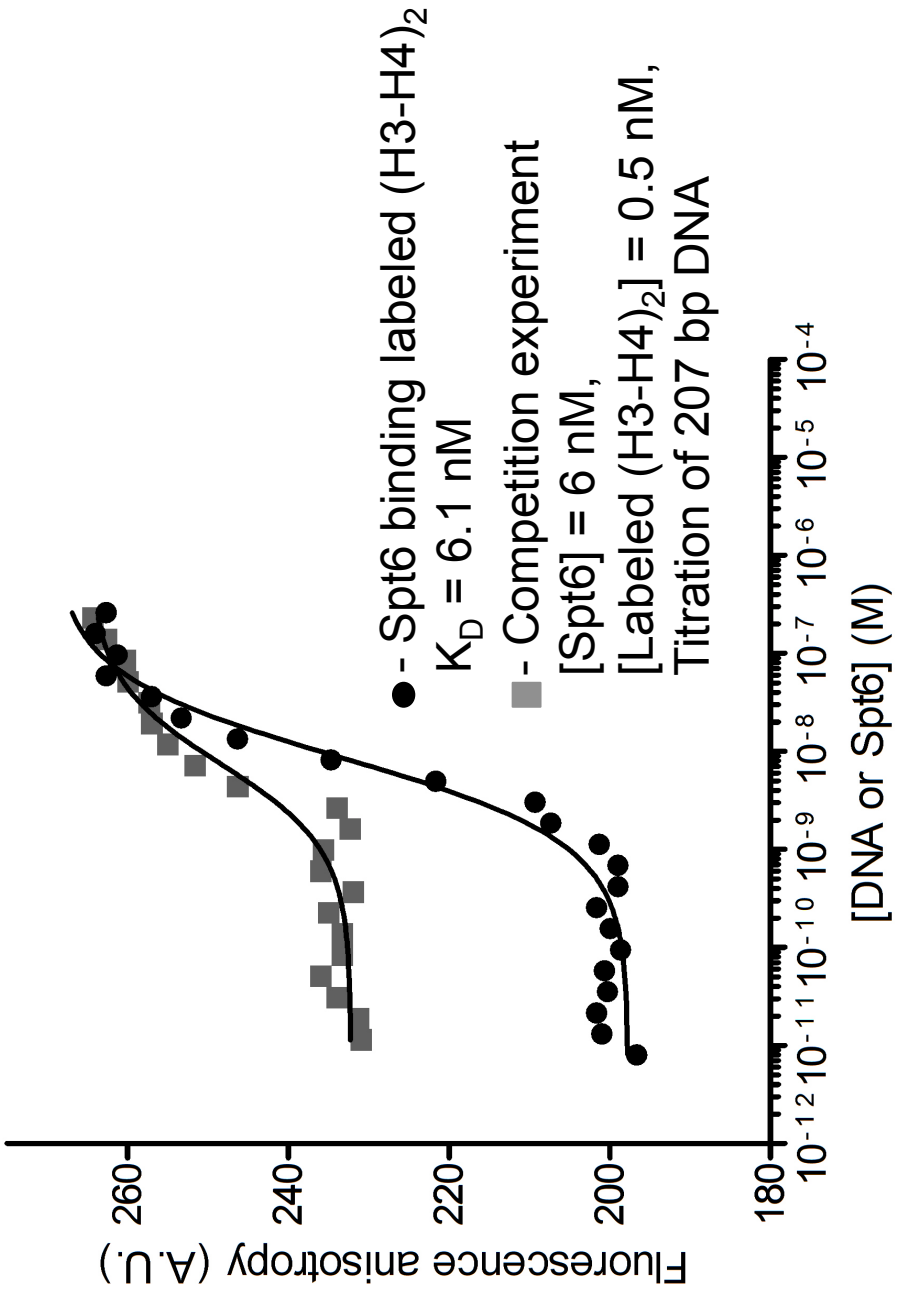
Spt6 directly competes with DNA for histone binding

Recent work from the Luger laboratory indicated that at least two histone chaperones, FACT and Nap1, utilize a direct competition mechanism to

chaperone histones whereby they prevent nonspecific histone-DNA contacts (Andrews et al., 2010; Winkler et al., 2011). This simple yet effective mechanism is an attractive model for all histone chaperones. We tested Spt6 for its ability to compete with DNA for binding $(H3\text{-}H4)_2$ using a heterologous equilibrium competition fluorescence anisotropy assay. In these assays, concentrations of labeled $(H3\text{-}H4)_2$ (0.5 nM) and full-length Spt6 were held constant, while a 207 bp DNA fragment was titrated. The concentration of Spt6 was held at 6 nM, which corresponds to the K_D of Spt6 for $(H3\text{-}H4)_2$ under low salt conditions in the fluorescence anisotropy assay (Figure 4-7, closed circles). The low salt conditions were used to reproduce the results from the Luger laboratory which indicated that a 207 bp DNA fragment, containing the Widom '601' nucleosome positioning sequence (NPS), binds $(H3\text{-}H4)_2$ with a K_D of 1 nM under low salt conditions (Andrews et al., 2008). Our results with titration of a similar 207 bp DNA fragment (also containing the '601' NPS) into labeled $(H3\text{-}H4)_2$ were essentially identical to the previously published results (K_D of 2 nM). Interestingly, we found that Spt6 does indeed compete for histone binding with DNA; however, $(H3\text{-}H4)_2$ prefers DNA binding under conditions where the concentration of DNA exceeds the K_D of Spt6 for $(H3\text{-}H4)_2$ (Figure 4-7, grey squares). These data reveal that Spt6 is capable of binding $(H3\text{-}H4)_2$ in the presence of low concentrations of DNA as indicated by the raw anisotropy signal being halfway between the signal of histones saturated with Spt6 and that of unbound histones (compare the unbound and saturated signal for the closed

Figure 4-7. Spt6 and DNA compete to bind $(H3\text{-}H4)_2$.

Binding isotherms derived from fluorescence anisotropy measurements of the interaction between Spt6 and $(H3\text{-}H4)_2$ (closed circles) and the equilibrium competition reaction (grey squares) in low salt buffer (150 mM KCl). Competition reactions contain constant concentrations of Spt6 and labeled $(H3\text{-}H4)_2$ as indicated such that half of the histones are bound to Spt6 and a 1.6-fold dilution series of DNA was titrated into these reactions.


Fluorescence anisotropy (A.U.)
260
240
220
200
180
10-12 10-11 10-10 10-9 10-8 10-7 10-6 10-5 10-4
[DNA or Spt6] (M)
● - Spt6 binding labeled (H3-H4)2
KD = 6.1 nM
■ - Competition experiment
[Spt6] = 6 nM,
[Labeled (H3-H4)2] = 0.5 nM,
Titration of 207 bp DNA

circles and the grey squares). In these reactions, half of the histones are bound to Spt6. Once the concentration of DNA in the titration exceeds the equilibrium binding constant of Spt6 for $(H3\text{-}H4)_2$, the DNA effectively out competes Spt6 for histone binding and we see all of the histones become saturated with DNA. Therefore, Spt6 likely utilizes a direct competition mechanism similar to other histone chaperones to chaperone $(H3\text{-}H4)_2$. In a physiological context, this model also makes sense in that histones should be the preferred binding partner for genomic DNA and histones only need to be chaperoned when the DNA has somehow become disengaged; for example, when the effective local DNA concentration drops significantly, as it must during processes such as transcription.

Functional implications for the histone chaperone activity of Spt6

The structural, genetic, and biochemical investigation of the Spn1-Spt6 interface reported in Chapter 3 indicated that the interaction is important for maintaining a repressive chromatin state. That study extended understanding of the histone chaperone activity of Spt6 by demonstrating that the region of the Spt6 protein required for binding Spn1 is also needed for interactions with nucleosomes and Spn1 antagonizes Spt6-nucleosome interactions in vitro. The work in this chapter also strongly supports a role for Spn1 in Spt6-mediated nucleosome reassembly, perhaps by regulating the process or acting as a switch that drives disengagement. Spt6 was also shown to compete with DNA for

binding to histones, indicating that Spt6 utilizes a direct-competition mechanism, similar to other histone chaperones (Andrews et al., 2010; Winkler et al., 2011), to accomplish the chaperoning of histones required for the proper reassembly of nucleosomes.

Experimental Procedures

Protein expression and purification

Spt6 proteins were cloned, expressed, and purified as previously described (Close et al., 2011; McDonald et al., 2010) with one modification; Spt6 constructs with C-terminal truncations at or before residue 489 (including Spt6-SBM) were purified via anion-exchange chromatography (5 mL Q HP, GE Healthcare) in place of a Heparin cation-exchange column after the Ni affinity purification. *Xenopus laevis* histones were expressed recombinantly and purified from inclusion-bodies as described elsewhere (Luger et al., 1999). 'Tailless' histones were prepared by inserting a PreScission Protease site by PCR such that upon protease treatment (after refolding) a GP dipeptide remains N-terminally fused to the following residues: H2A residues 14-118 (the C-terminal tail was removed by inserting two stop codons via site-directed mutagenesis reactions), H2B residues 24-122, H3 residues 27-135, and H4 residues 20-102. Purification of tailless histones was the same as for full-length histones, apart from protease treatment done in the presence of glutathione resin to remove the GST-labeled protease, and subsequent purification on a 120mL Superdex 200

16/60 (GE Healthcare) size exclusion column. Fluorescent labeling of histone H4 (T71C) and H2B (T112C) with Oregon Green-488 maleimide (Invitrogen) was performed according to manufacturers instructions prior to histone refolding and purification on a size exclusion column. The Spt6-SBM protein was expressed recombinantly, purified as above, labeled with Oregon Green-488 maleimide (Invitrogen) on a Cysteine residue introduced by site-directed mutagenesis in the linker sequence that remains after TEV protease treatment, and purified from free dye on a size exclusion column.

GST-Pulldown binding and competition assays

N-terminal GST-Spn1(148-307) and GST-Spt6(239-1451) fusion proteins were cloned by LR reaction from a pENTR clone into the pDEST15 vector as per manufacturer's instructions (Invitrogen). GST-tagged proteins were purified as described previously (McDonald et al., 2010) substituting a GST-sepharose column (GSTrap HP, GE Healthcare) for the Ni-NTA column. Purified recombinant proteins were mixed at equimolar concentrations (2 μM) in a total volume of 300 μL and incubated for 1 hour at 4 °C. Competition reactions contained 2 μM GST-Spn1(148-307) and Spt6(239-1451), and differing amounts of $(H3\text{-}H4)_2$ (1.5-50 μM). After the initial incubation, input samples were taken and 50 μL of washed glutathione sepharose beads (4 Fast Flow, GE Healthcare) were added to each reaction and incubated for another hour at 4 °C. Beads were then washed three times (500 μL for each wash) with either low salt (200 mM

NaCl) or high salt (750 mM NaCl) binding buffer (15 mM Tris pH 7.5, 200 or 750 mM NaCl, 5 % Glycerol, 2 mM 2-Mercaptoethanol, 0.5 mM EDTA, and 0.1 % Tween-20). Proteins were eluted from beads by boiling in SDS-PAGE loading buffer for 10 min (Elution samples). Samples were run on 15% SDS polyacrylamide gels using a Tris-Glycine running buffer and visualized with Coomasie staining.

Fluorescence anisotropy

All fluorescence polarization assays were incubated for 1 hr and measured at 25°C in 384-well flat bottom black (nonbinding surface) microplates (Corning). Purified Spt6 proteins were titrated in 1.6- to 1.8-fold serial dilutions against a constant concentration (0.5 nM) of refolded and labeled histones in either low salt buffer (20 mM Tris pH 7.5, 150 mM KCl, 2 mM MgCl2, 5 % glycerol, and 1 mM DTT) or high salt buffer (20 mM Tris pH 7.5, 750 mM NaCl, 5% glycerol, and 1 mM DTT). Anisotropy measurements and equilibrium dissociation constants (K_D) were determined as described (Close et al., 2011) and plotted with GraphPad Prism software. Competition assays were performed by titration of 1.64-fold serial dilutions of a 207 bp DNA fragment against a constant concentration of both labeled histones (0.5 nM) and Spt6 corresponding to the measured K_D for $(H3H4)_2$ (6 nM).

DNA preparation

The 207 bp '601' nucleosome positioning sequence DNA fragment was amplified by PCR and purified by precipitation and gel extraction/electroelution.

Native electrophoretic mobility shift assay

Purified recombinant proteins were incubated at room temperature for one hour in EMSA reaction buffer (15 mM Tris pH 7.5, 200 mM NaCl, 5% Glycerol, 0.5 mM EDTA, 2 mM BME). A 2-fold dilution series of H2A-H2B starting at 200 uM concentration was made such that each reaction contained a constant concentration of Spt6-SBM (20 uM), and various concentrations of H2A-H2B. After incubation at room temperature, the reactions were subjected to electrophoresis on native 4-15% TGX gels (Bio-Rad) in 0.4x TBE at 110v for 40 min at room temperature. The gels were stained with Coomassie and band intensities were quantified using an Odyssey Imaging System (LiCor Biosciences). The fraction bound was calculated by quantifying the $Spt6_{free}$ (total band intensity of reaction containing no histones). Spt6 bound to histones does not enter a native gel and thus a reduction in $Spt6_{free}$ indicates binding to histones. The fraction bound=$1-([Spt6]_{free}/\Delta[Spt6]_{free})$. Equilibrium binding constants (K_D) were calculated by plotting data points and curve fitting using GraphPad Prism software and the full quadratic expansion non-linear regression equation (LiCata and Wowor, 2008).

Isothermal titration calorimetry

Purified recombinant proteins were dialyzed at 4°C overnight against 2 L of degassed high salt ITC buffer (20 mM HEPES pH 7.5, 750 mM NaCl, 0.5 mM EDTA, 5 % glycerol, and 2 mM 2-mercaptoethanol) for reactions with $(H3\text{-}H4)_2$, or low salt ITC buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 0.5 mM EDTA, 5 % glycerol, and 2 mM 2-mercaptoethanol) for reactions with H2A-H2B. Titrations were done at 25°C on an iTC200 (Microcal) and spaced 180 s apart. Spt6(239-489) titrations were done with 1.8 μL injections of 1.25 mM Spt6 into 0.125 mM $(H3H4)_2$. H2A-H2B titrations were done with 1.8 μL injections of 0.075 mM H2A-H2B into 0.6 mM Spt6-SBM. Origin software (Microcal) was used for data analysis and determination of thermodynamic parameters as previously described (McDonald et al., 2010).

References

Adkins, M.W., and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. Mol Cell *21*, 405-416.

Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A., and Luger, K. (2010). The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. Mol Cell *37*, 834-842.

Andrews, A.J., Downing, G., Brown, K., Park, Y.J., and Luger, K. (2008). A thermodynamic model for Nap1-histone interactions. J Biol Chem *283*, 32412-32418.

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. Nature *420*, 837-841.

Antczak, A.J., Tsubota, T., Kaufman, P.D., and Berger, J.M. (2006). Structure of the yeast histone H3-ASF1 interaction: implications for chaperone mechanism, species-specific interactions, and epigenetics. BMC Struct Biol *6*, 26.

Ardehali, M.B., Yao, J., Adelman, K., Fuda, N.J., Petesch, S.J., Webb, W.W., and Lis, J.T. (2009). Spt6 enhances the elongation rate of RNA polymerase II in vivo. EMBO J *28*, 1067-1077.

Baniahmad, C., Nawaz, Z., Baniahmad, A., Gleeson, M.A., Tsai, M.J., and O'Malley, B.W. (1995). Enhancement of human estrogen receptor activity by SPT6: a potential coactivator. Mol Endocrinol *9*, 34-43.

Beckouet, F., Mariotte-Labarre, S., Peyroche, G., Nogi, Y., and Thuriaux, P. (2011). Rpa43 and its partners in the yeast RNA polymerase I transcription complex. FEBS Lett *585*, 3355-3359.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

Bowman, A., Ward, R., Wiechens, N., Singh, V., El-Mkami, H., Norman, D.G., and Owen-Hughes, T. (2011). The histone chaperones Nap1 and Vps75 bind histones H3 and H4 in a tetrameric conformation. Mol Cell *41*, 398-408.

Bucheli, M.E., and Buratowski, S. (2005). Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. Embo J *24*, 2150-2160.

Burckin, T., Nagel, R., Mandel-Gutfreund, Y., Shiue, L., Clark, T.A., Chong, J.L., Chang, T.H., Squazzo, S., Hartzog, G., and Ares, M., Jr. (2005). Exploring functional relationships between components of the gene expression machinery. Nat Struct Mol Biol *12*, 175-182.

Cairns, B.R. (1998). Chromatin remodeling machines: similar motors, ulterior motives. Trends Biochem Sci *23*, 20-25.

Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol *6*, e277.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Close, D., Johnson, S.J., Sdano, M.A., McDonald, S.M., Robinson, H., Formosa, T., and Hill, C.P. (2011). Crystal structures of the S. cerevisiae Spt6 core and C-terminal tandem SH2 domain. J Mol Biol *408*, 697-713.

Diebold, M.L., Loeliger, E., Koch, M., Winston, F., Cavarelli, J., and Romier, C. (2010). Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II. J Biol Chem *285*, 38389-38398.

Endoh, M., Zhu, W., Hasegawa, J., Watanabe, H., Kim, D.K., Aida, M., Inukai, N., Narita, T., Yamada, T., Furuya, A.*, et al.* (2004). Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. Mol Cell Biol *24*, 3324-3336.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495-508.

Formosa, T. (2012). The role of FACT in making and breaking nucleosomes. Biochim Biophys Acta *1819*, 247-255.

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y., and Stillman, D.J. (2001). Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. EMBO J *20*, 3506-3517.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M.*, et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Hartzog, G.A., Wada, T., Handa, H., and Winston, F. (1998). Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. Genes Dev *12*, 357-369.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Kaplan, C.D., Holland, M.J., and 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.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

Kaplan, C.D., Morris, J.R., Wu, C., and Winston, F. (2000). Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. Genes Dev *14*, 2623-2634.

Keegan, B.R., Feldman, J.L., Lee, D.H., Koos, D.S., Ho, R.K., Stainier, D.Y., and Yelon, D. (2002). The elongation factors Pandora/Spt6 and Foggy/Spt5 promote transcription in the zebrafish embryo. Development *129*, 1623-1632.

Kiely, C.M., Marguerat, S., Garcia, J.F., Madhani, H.D., Bahler, J., and Winston, F. (2011). Spt6 is required for heterochromatic silencing in the fission yeast Schizosaccharomyces pombe. Mol Cell Biol *31*, 4193-4204.

Kok, F.O., Oster, E., Mentzer, L., Hsieh, J.C., Henry, C.A., and Sirotkin, H.I. (2007). The role of the SPT6 chromatin remodeling factor in zebrafish embryogenesis. Dev Biol *307*, 214-226.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Curr Opin Genet Dev *21*, 175-186.

Li, L., Ye, H., Guo, H., and Yin, Y. (2010). Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. Proc Natl Acad Sci U S A *107*, 3918-3923.

LiCata, V.J., and Wowor, A.J. (2008). Applications of fluorescence anisotropy to the study of protein-DNA interactions. Methods Cell Biol *84*, 243-262.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., 3rd, and Hartzog, G.A. (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.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol Biol *119*, 1-16.

Luger, K., and Richmond, T.J. (1998). The histone tails of the nucleosome. Curr Opin Genet Dev *8*, 140-146.

Maclennan, A.J., and Shaw, G. (1993). A yeast SH2 domain. Trends Biochem Sci *18*, 464-465.

McDonald, S.M., Close, D., Xin, H., Formosa, T., and Hill, C.P. (2010). Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. Mol Cell *40*, 725-735.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404-405.

Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. Nature *446*, 338-341.

Nishiwaki, K., Sano, T., and Miwa, J. (1993). emb-5, a gene required for the correct timing of gut precursor cell division during gastrulation in Caenorhabditis elegans, encodes a protein similar to the yeast nuclear protein SPT6. Mol Gen Genet *239*, 313-322.

Okazaki, I.M., Okawa, K., Kobayashi, M., Yoshikawa, K., Kawamoto, S., Nagaoka, H., Shinkura, R., Kitawaki, Y., Taniguchi, H., Natsume, T.*, et al.* (2011). Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. Proc Natl Acad Sci U S A *108*, 7920-7925.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol *8*, 983-994.

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

Shen, X., Xi, G., Radhakrishnan, Y., and Clemmons, D.R. (2009). Identification of novel SHPS-1-associated proteins and their roles in regulation of insulin-like growth factor-dependent responses in vascular smooth muscle cells. Mol Cell Proteomics *8*, 1539-1551.

Stillman, D.J. (2010). Nhp6: a small but powerful effector of chromatin structure in Saccharomyces cerevisiae. Biochim Biophys Acta *1799*, 175-180.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Su, D., Hu, Q., Zhou, H., Thompson, J.R., Xu, R.M., Zhang, Z., and Mer, G. (2011). Structure and histone binding properties of the Vps75-Rtt109 chaperone-lysine acetyltransferase complex. J Biol Chem *286*, 15625-15629.

Sun, M., Lariviere, L., Dengl, S., Mayer, A., and Cramer, P. (2010). A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). J Biol Chem *285*, 41597-41603.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Velazquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S., and Freire, E. (2004). Isothermal titration calorimetry. Curr Protoc Cell Biol *Chapter 17*, Unit 17 18.

Widiez, T., El Kafafi el, S., Girin, T., Berr, A., Ruffel, S., Krouk, G., Vayssieres, A., Shen, W.H., Coruzzi, G.M., Gojon, A.*, et al.* (2011). High nitrogen insensitive 9 (HNI9)-mediated systemic repression of root NO3- uptake is associated with changes in histone methylation. Proc Natl Acad Sci U S A *108*, 13329-13334.

Winkler, D.D., Muthurajan, U.M., Hieb, A.R., and Luger, K. (2011). Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events. J Biol Chem *286*, 41883-41892.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev *21*, 160-174.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes & Development *22*, 3422-3434.

CHAPTER 5

CONCLUSIONS AND ONGOING RESEARCH

Summary

Spt6 and Spn1 are transcription factors that bind one another and are each essential for yeast viability. Although a broad role for both Spt6 and Spn1 in transcription is clear, little is known of the precise mechanistic details for any of their proposed functions. The structural and biochemical work reported in Chapter 2 established an atomic model of the five core domains and the C-terminal domain of Spt6, and began to analyze potential binding partners. The work in Chapter 3 extended our understanding of the structure of Spn1, structurally, biochemically, and genetically characterized the Spn1-Spt6 interaction, and determined that the interaction is important biologically for maintaining a repressive chromatin state. Chapter 4 further investigated the direct physical interaction between Spt6 and histones and established that Spn1 indeed competes with histone for Spt6 binding, supporting a model in which the histone chaperone/ nucleosome reassembly activity of Spt6 is influenced by Spn1 binding. This chapter will describe ongoing preliminary structural and biochemical studies on Spt6 and will discuss potential future approaches aimed

at furthering our understanding of Spt6 and Spn1 function. Our preliminary structural studies on the Spt6-H2A-H2B complex, which are being undertaken by Matt Sdano, have yielded protein-containing crystals that diffract to low (~7 Å) resolution. Matt has also taken over crystallization trials on the Spt6-$(H3\text{-}H4)_2$ complex, which have thus far failed to yield protein crystals. Matt Sdano has also developed a small-scale proteomics approach that will be exploited to identify and characterize new Spt6 and Spn1 binding partners from yeast lysates. Preliminary studies using this approach reveal interactions between the Spt6 C-terminal domain and several factors involved in transcription, including Yra1, an mRNA export adaptor protein, and Tom1, an E3-ubiquitin ligase of the HECT-class that is important for mRNA export and excess histone degradation. Long-term goals include structural, biochemical, and genetic characterization of confirmed direct binding partners. Work in this chapter has been done in close collaboration with Matt Sdano and Devin Close, who have each contributed extensively to the search for new binding partners and the characterization of Spt6-histone interactions.

Functional Implications For Spt6 and Spn1

The Spt6 core and C-terminal domain structures

The crystal structures reported in Chapter 2 provide an atomic model for nearly all of the Spt6 protein that is predicted to be ordered. The structures of the core of Spt6 revealed five domains with structural homology to other known and

well-characterized proteins. It is likely that these domains contribute to the function of Spt6 by serving as recruiting modules during the transcription cycle of the many protein-coding genes that Spt6 influences. Spt6 is required for recruiting factors that affect the maturation and export of the mRNA molecules produced during transcription (Andrulis et al., 2002; Bucheli and Buratowski, 2005; Burckin et al., 2005; Kaplan et al., 2005; Yoh et al., 2007), and the Spt6 core domains, which resemble nucleic acid-binding modules, appear poised to execute this function. However, the contribution of each of the five core domains to the function of Spt6 remains enigmatic, and will be the focus of future research. The C-terminal domain of Spt6 folds into a novel tandem SH2 domain, the only predicted SH2 domain in yeast (Maclennan and Shaw, 1993), and may bind directly to the phosphorylated tail of RNAPII (Diebold et al., 2010b; Sun et al., 2010; Yoh et al., 2007; Yoh et al., 2008). It is possible, even likely, that each domain of Spt6 has multiple binding partners as the composition of the transcription elongation complex varies throughout the length of a gene. Further, nearly every individual gene is regulated in a unique fashion (Martinez-Rucobo and Cramer, 2012; Treutlein et al., 2012; Yoh et al., 2007; Zhang et al., 2008) and Spt6 impacts a large number of genes (Mayer et al., 2010). This means that interactions important for the proper transcription of one gene may not be required at other genes, creating a complicated, dynamic, and transient web of important interactions throughout the genome. Teasing apart those interactions

that are physiologically relevant will be difficult and tedious, but clearly important to further our understanding of transcription and life.

The Spn1 core structure

The structure of the ordered, central core of Spn1 revealed a novel eight-helix-bundle fold. The superhelical orientation of the fold is common among proteins with helical repeats, which come in multiple varieties including HEAT-repeats, armadillo repeats, and TPR repeats (Andrade et al., 2001a; Andrade et al., 2001b). Proteins with each of these repeat folds are found in the nucleus and include transcription factors such as ß-catenin (Huber et al., 1997), and the Swi2/Snf2 chromatin remodeler Mot1 (Wollmann et al., 2011). The orientation of the last four helices of the Spn1 structure closely resemble the structure of the transcription factor TFIIS (PDB ID: 1wjt, RIKEN structural genomics/proteomics initiative), and therefore may be a common protein binding module utilized in various transcription complexes. The structure also revealed the presence of a highly conserved, positively charged, deep pocket formed at the ends of helices 1, 2, and 4 (Figure 3-2B and C). The conservation suggests that this pocket is an important interaction surface and identifying other Spn1-interacting partners will be another focus of future research.

The structure of the Spn1-Spt6 complex

The structure of the complex between Spt6 and Spn1 provided atomic details of how the disordered N-terminal segment of Spt6 becomes ordered upon binding to the Spn1 core, and the extensive interface seen in the structure is consistent with the strong specific binding observed in solution. One striking feature of the interface is the deep hydrophobic pocket formed by the last three helices at the C-terminus of Spn1, which contacts the highly conserved hydrophobic Spt6 residues I248 and F249. This strong hydrophobic portion of the interface appears to be conserved in sequence and utilized by other transcription factors, with helical arrangements and hydrophobic pockets similar to that of Spn1 found in TFIIS, Elongin A, and Med26 (Diebold et al., 2010a), and the IF motif of Spt6 that binds the hydrophobic pocket of Spn1 is found and conserved in multiple subunits of the SAGA and Mediator complexes (Diebold et al., 2010a).

The extended and inherently flexible nature of the Spn1 binding sequence of Spt6 presumably explains why dramatic mutations in the interface substantially weaken but do not completely eliminate this interaction, as localized perturbations can be accommodated by conformational changes that do not propagate across the entire interface. Moreover, the approximately 35 residues separating the Spn1-binding residues from the ordered core of Spt6 likely provide a tether whose flexibility may be required to allow the Spn1-Spt6 complex to form in multiple functional contexts. The use of the inherently flexible N-terminal

segment of Spt6 that binds the Spn1 core may provide a mechanism that allows flexibility in a crowded transcriptional environment.

Histone chaperone activity of Spt6

The structural, genetic, and biochemical investigation of the Spn1-Spt6 interface reported in Chapter 3 indicated that the interaction is important for maintaining a repressive chromatin state. Our studies have extended understanding of the histone chaperone activity of Spt6 by demonstrating that the region of the protein required for binding Spn1 is also needed for interactions with nucleosomes and free histones. Our data presented in Chapter 4 also strongly supports a role for Spn1 in Spt6-mediated nucleosome reassembly, perhaps by regulating the process or acting as a switch that drives disengagement. Spt6 was also shown to compete with DNA for binding to histones, indicating that Spt6 utilizes a direct-competition mechanism, similar to other histone chaperones (Andrews et al., 2010; Winkler et al., 2011), to accomplish the chaperoning of histones required for the proper reassembly of nucleosomes.

Biochemical Analysis of Spt6

Preliminary crystallization of Spt6 in complex with histones H2A-H2B

In studies complimentary to our analysis of Spt6-histone interactions, Matt Sdano has begun crystallization trials with the Spt6-SBM and the two different histone subcomplexes. Preliminary trials with $(H3\text{-}H4)_2$ have failed to produce

protein crystals, although many options remain to be tested, including using 'tailless' histones, combinations of full-length and 'tailless' histones, and varying the amount of the Spt6 peptide. Preliminary trials with H2A-H2B have yielded protein-containing crystals grown from drops that contained equimolar concentrations of the Spt6-SBM and full-length H2A-H2B (Figure 5-1A). These crystals diffract poorly to a resolution of about 7 Å (Figure 5-1B), and efforts to determine the crystalline lattice space group have been unsuccessful. Mass spectrometry analysis of the crystals suggest that the histone proteins are present in the crystals, but the presence of the Spt6-SBM could not be determined due to the presence of similar molecular weight polyethylene glycol polymers in the crystallization buffer. Efforts to improve crystal packing and diffraction are ongoing. Structural data may also be obtained by using NMR, and the studies by Yawen Bai and colleagues (Zhou et al., 2008; Zhou et al., 2011) provide precedent that histones are amenable to NMR experiments, including the necessary isotope labeling by expression in minimal media.

Identifying new Spt6 and Spn1 binding partners

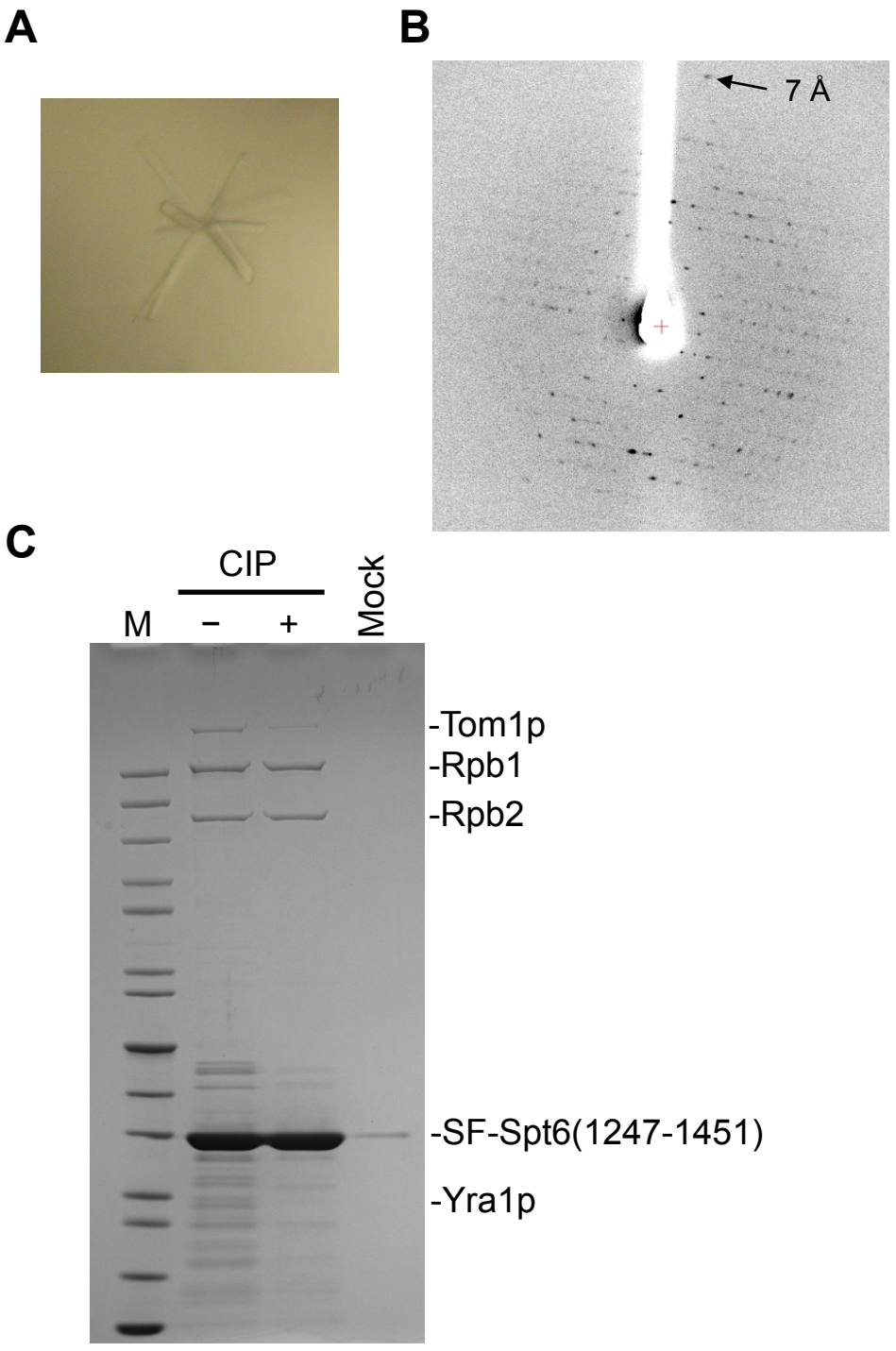
A small-scale proteomics approach has been developed by Matt Sdano to screen for new Spt6 and Spn1 binding partners in yeast. This approach takes advantage of the inducible *GAL1* promoter to drive the overexpression of proteins with an N-terminal tandem strep and flag-epitope tag. This allows for easy tandem affinity purifications where the eluates are separated by SDS-PAGE and

Figure 5-1. Preliminary structural work and identification of new Spt6 binding partners

A) Putative crystals of an Spt6-H2A-H2B complex. Crystals were grown at 20 °C in drops comprised of 0.3 µL of 5.7 mg.mL^{-1}protein (a 1.1: 1 molar ratio of Spt6-SBM:H2A-H2B) and 0.3 µL of well solution (0.2 M $[NH^4]_2SO_4$, 0.1 M Bis-tris [pH 5.5], 25% PEG 3350).

B) X-ray diffraction image collected from a crystal in (A).

C) SDS-PAGE of the TAP purification of the Spt6 tSH2 domain (residues 1247-1451) expressed in yeast. The bands of putative interacting partners are indicated and identified by mass spectrometry. TAP-purification lanes were either untreated (-) or treated (+) with calf-intestinal phosphatase (CIP) to identify partners that bind in a phosphorylation dependent manner. The previously identified Spt6-interacting subunits of RNA polymerase II (Rpb1 and Rpb2) are indicated. Bands were visualized by coomassie staining.


A
B
7 Å
C
CIP
M
−
+
Mock
-Tom1p
-Rpb1
-Rpb2
-SF-Spt6(1247-1451)
-Yra1p

interacting protein bands are identified using mass spectrometry. Matt has used this approach to identify binding partners of the Spt6 C-terminal tSH2 domain (Figure 5-1C), including Tom1, a HECT domain containing E3 ubiquitin-ligase important for mRNA export and excess histone degradation (Saleh et al., 1998), and the mRNA export adapter Yra1 (Strasser and Hurt, 2000). Further confirmation of direct interactions will employ techniques such as surface plasmon resonance and ITC.

In order to assess the amenability of proteins or individual protein domains for use in this system, we test for expression and solubility in bacteria, and employ standard purification techniques including size exclusion chromatography to ensure monodispersity in solution. This is necessary to avoid overexpressing proteins or domains that tend to aggregate, or form soluble aggregates that would yield unreliable or misleading results. If necessary, CD spectroscopy can also be used to assess whether the expressed protein appears to have folded properly. Alternative approaches to this system will employ genomic tagging of proteins to allow for purification and identification of binding partners of proteins expressed from their native promoters at endogenous levels. To this end we have successfully cloned, expressed and purified Spn1, the Spn1 core, and the Spt6 HtH domain from bacteria. We are currently cloning and testing the other individual domains of Spt6. The overall goal will be to try to identify physiologically relevant binding partners for Spn1 and each Spt6 domain. Once viable candidates have been identified, we will thoroughly characterize the

importance of these interactions using the structural, biochemical, and genetic approaches at our disposal.

Conclusions

The multiple functional roles of Spt6 and Spn1 and their requirement for yeast viability make each exciting targets for precise mechanistic analyses. The work in this dissertation has expanded upon the foundational structural and biochemical work on Spt6 of Sean Johnson and Devin Close, both former members of the Hill lab. The primary focus has been the flexible, disordered, and highly negatively charged N-terminal region of Spt6, which contributes to interactions with both Spn1 and histones. We have provided evidence of a novel functional role for Spn1 in the regulation of the histone chaperone and nucleosome reassembly activity of Spt6. Further, we have determined that Spt6 chaperones histones by direct-competition with DNA, which provides mechanistic details about the Spt6-histone interface and indicates how Spt6 accomplishes nucleosome reassembly. The precise mechanistic details of how Spn1 influences nucleosome reassembly in vivo remain an important obstacle and focus of future work. Ongoing work includes structural, biochemical, and genetic studies to more accurately dissect and define mechanisms that contribute to the multiple functions of both Spn1 and Spt6. This includes a more accurate description of the interfaces formed upon Spt6 binding to the two different histone subcomplexes.

Refereneces

Andrade, M.A., Perez-Iratxeta, C., and Ponting, C.P. (2001a). Protein repeats: structures, functions, and evolution. J Struct Biol *134*, 117-131.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Muller, C.W., and Bork, P. (2001b). Comparison of ARM and HEAT protein repeats. J Mol Biol *309*, 1-18.

Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A., and Luger, K. (2010). The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. Mol Cell *37*, 834-842.

Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P., and Lis, J.T. (2002). The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. Nature *420*, 837-841.

Bucheli, M.E., and Buratowski, S. (2005). Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. Embo J *24*, 2150-2160.

Burckin, T., Nagel, R., Mandel-Gutfreund, Y., Shiue, L., Clark, T.A., Chong, J.L., Chang, T.H., Squazzo, S., Hartzog, G., and Ares, M., Jr. (2005). Exploring functional relationships between components of the gene expression machinery. Nat Struct Mol Biol *12*, 175-182.

Diebold, M.L., Koch, M., Loeliger, E., Cura, V., Winston, F., Cavarelli, J., and Romier, C. (2010a). The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. EMBO J *29*, 3979-3991.

Diebold, M.L., Loeliger, E., Koch, M., Winston, F., Cavarelli, J., and Romier, C. (2010b). Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II. J Biol Chem *285*, 38389-38398.

Huber, A.H., Nelson, W.J., and Weis, W.I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. Cell *90*, 871-882.

Kaplan, C.D., Holland, M.J., and 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.

Maclennan, A.J., and Shaw, G. (1993). A yeast SH2 domain. Trends Biochem Sci *18*, 464-465.

Martinez-Rucobo, F.W., and Cramer, P. (2012). Structural basis of transcription elongation. Biochim Biophys Acta.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol *17*, 1272-1278.

Saleh, A., Collart, M., Martens, J.A., Genereaux, J., Allard, S., Cote, J., and Brandl, C.J. (1998). TOM1p, a yeast hect-domain protein which mediates transcriptional regulation through the ADA/SAGA coactivator complexes. J Mol Biol *282*, 933-946.

Strasser, K., and Hurt, E. (2000). Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. EMBO J *19*, 410-420.

Sun, M., Lariviere, L., Dengl, S., Mayer, A., and Cramer, P. (2010). A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). J Biol Chem *285*, 41597-41603.

Treutlein, B., Muschielok, A., Andrecka, J., Jawhari, A., Buchen, C., Kostrewa, D., Hog, F., Cramer, P., and Michaelis, J. (2012). Dynamic architecture of a minimal RNA polymerase II open promoter complex. Mol Cell *46*, 136-146.

Winkler, D.D., Muthurajan, U.M., Hieb, A.R., and Luger, K. (2011). Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events. J Biol Chem *286*, 41883-41892.

Wollmann, P., Cui, S., Viswanathan, R., Berninghausen, O., Wells, M.N., Moldt, M., Witte, G., Butryn, A., Wendler, P., Beckmann, R.*, et al.* (2011). Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. Nature *475*, 403-407.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. (2007). The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev *21*, 160-174.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes & Development *22*, 3422-3434.

Zhang, L., Fletcher, A.G., Cheung, V., Winston, F., and Stargell, L.A. (2008). Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. Mol Cell Biol *28*, 1393-1403.

Zhou, Z., Feng, H., Hansen, D.F., Kato, H., Luk, E., Freedberg, D.I., Kay, L.E., Wu, C., and Bai, Y. (2008). NMR structure of chaperone Chz1 complexed with histones H2A.Z-H2B. Nat Struct Mol Biol *15*, 868-869.

Zhou, Z., Feng, H., Zhou, B.R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C.*, et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature *472*, 234-237.