Architecture of the *Saccharomyces cerevisiae* SAGA transcription coactivator complex

Yan Han

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Steven Hahn, Chair

Jeff Ranish

Toshio Tsukiyama

Program Authorized to Offer Degree:

Department of Biochemistry

University of Washington

Abstract

Architecture of the *Saccharomyces cerevisiae* SAGA transcription coactivator complex

Yan Han

Chair of the Supervisory Committee:
Affiliate Professor, Steven Hahn
Department of Biochemistry

The evolutionarily conserved transcription coactivator SAGA (Spt-Ada-Gcn5 Acetyltransferase) is a multi-subunit complex with a modular structure, which has several distinct activities that are used to regulate activator-dependent transcription. SAGA covalently modifies histones using its histone acetyltransferase (HAT) and deubiquitination (DUB) modules. It also directly regulates the formation of the transcription preinitiation complex (PIC) through direct interactions with both transcriptional activators and the TATA-box Binding Protein (TBP). Despite SAGA's important roles in regulating transcription, its overall architecture and structural organization of its modules remain unclear. The large size and complex subunit composition of SAGA make it difficult to study its structure using high-resolution approaches such as X-ray crystallography. Using an alternative approach, I combined chemical crosslinking with mass spectrometry (CXMS) to investigate the architecture of SAGA. In Chapter 2, I describe the results of my efforts using this approach, finding that the SAGA Taf and Taf-like subunits form a TFIID-like core complex at the center of

SAGA that makes extensive interactions with all other SAGA modules. In Chapter 3, I show that the HAT and DUB modules are in close proximity, and the DUB module modestly stimulates HAT function. In Chapter 4, I describe the finding that SAGA-TBP binding involves a network of interactions between subunits Spt3, Spt8, Spt20, and Spt7, and the attempts I have made toward solving the crystal structure of Spt8 in complex with TBP. Finally, in Chapter 5, I combine all of the data and derive a model for the molecular architecture of the SAGA complex. My results provide new insight into SAGA function in gene regulation, its structural similarity with TFIID, and functional interactions between the SAGA modules.

Table of content

List of figures

List of tables

Acknowledgement

I would like to thank my advisor Dr. Steve Hahn for his guidance and support throughout my graduate study. I would also like to thank my thesis committee members, Dr. Steve Hahn, Dr. Julian Simon, Dr. Jeff Ranish, and Dr. Toshio Tsukiyama, for their advises and comments on my work and dissertation. Many thanks go to all of the members of the Hahn lab and the Ranish lab for their expertise and advice on my thesis work. It has always been a pleasure to work with such wonderful scientists. I also want to thank my collaborator Dr. Ning Zheng, and members of the Zheng lab. This dissertation would never have been accomplished without the support and encouragement from my family and all of my friends. At last, I am especially grateful for my wife and all the sacrifices she has made.

Dedication

To my beloved family.

Chapter 1 Introduction to the SAGA complex

Transcription is the process whereby genetic information is passed from DNA to RNA. RNA synthesis reactions are catalyzed by a group of enzymes named RNA polymerases, and generally include three major steps: initiation, elongation, and termination. In eukaryotes, at least three classes of RNA polymerases generate RNA transcripts using DNA as a template (Vannini & Cramer, 2012). Of the three classes of RNA polymerases, RNA polymerase II (Pol II) transcribes all of the protein coding genes, and its regulation is one of the most important mechanisms for control of gene expression (Vannini & Cramer, 2012; Fuda *et al*, 2009; Hahn & Young, 2011).

1.1 Regulation of transcription initiation by RNA Polymerase II

Transcription initiation by Pol II is a highly regulated process that requires a large number of protein factors. To initiate transcription, Pol II must form a preinitiation complex (PIC) at promoter regions together with several general transcription factors (GTFs) that include TATA binding protein (TBP), transcription factor (TF) IIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Hahn & Young, 2011; Liu *et al*, 2013; Fuda *et al*, 2009). The primary step in PIC formation is the binding of TBP specifically to TATA or TATA-like sequences at the promoter region of DNA (Hahn & Young, 2011; Huisinga & Pugh, 2004; Pugh, 2000). This binding induces a sharp 90° bend in DNA (Hahn & Young, 2011; Pugh, 2000; Rhee & Pugh, 2012). Pol II within the PIC then is able to open the double-stranded DNA, select a start site to initiate transcription and escape the promoter region with the help of the GTFs (Hahn & Young, 2011; Fuda *et al*, 2009; Liu *et al*, 2013).

Transcription activation is one of the key mechanisms of regulating gene expression. Transcriptional activators typically turn on transcription by targeting and recruiting coactivators to the promoters (Fuda *et al*, 2009). Transcriptional coactivators, which are usually multiprotein complexes and highly conserved, facilitate transcription activation by recruiting the basal transcription machinery, and/or by modifying chromatin structure either by covalent modifications of histones or ATP-dependent remodeling (Hahn & Young, 2011). The focus of my thesis, the SAGA (Spt-Ada-Gcn5 Acetyltransferase) complex, is one important example of a conserved transcriptional coactivator.

1.2 SAGA is a multifunctional complex

1.2.1 SAGA is composed of different modules

Yeast SAGA and the closely related mammalian orthologs STAGA, PCAF and TFTC are broadly conserved coactivator complexes that regulate the transcription of many inducible and developmentally regulated genes (Rodríguez-Navarro, 2009; Weake & Workman, 2012; Helmlinger, 2012; Samara & Wolberger, 2011). In yeast, SAGA is required for the transcription of about 10% of the genes, which are normally TATA-containing, stress-regulated and highly inducible (Bhaumik, 2011; Huisinga & Pugh, 2004). SAGA is a multifunctional complex, containing at least five separate activities: nucleosomal histone H3 acetyltransferase (HAT), histone H2B deubiquitinase (DUB), TBP binding, nucleosome binding, and activator binding.

SAGA from *S. cerevisiae* is a 1.8 MDa complex comprised of 19-20 subunits with its various coactivator activities separated into distinct modules (Table 1.1). Of the 19-20 subunits, Spt7, Ada1, and Spt20 were shown to be essential for the integrity of the

complex, as deletion of these subunits leads to dissociation of SAGA (Koutelou *et al*, 2010; Sterner *et al*, 1999). SAGA also shares a set of subunits with the coactivator TFIID, which likely play a structural role in SAGA architecture (Grant *et al*, 1998a). Within TFIID, two copies each of the histone fold pair-containing subunits Taf6-Taf9 and Taf4-Taf12, together with two copies of Taf5, form a symmetric complex (Bieniossek *et al*, 2013). This complex is bound by the Taf8-Taf10 heterodimer, generating an asymmetric core structure upon which the other TFIID-specific Tafs assemble. Likewise, SAGA contains a similar set of Tafs but with Taf4 and Taf8 replaced by the histone fold domain containing SAGA-specific subunits Ada1 and Spt7 (Gangloff *et al*, 2001; 2000).

Table 1.1 SAGA subunits and functions.

SAGA subunit	Function
Gcn5	Histone Acetyltransferase (HAT)
Ada2	
Ada3	
Sgf29	
Ubp8	Deubiquitination (DUB)
Sus1	
Sgf11	
Sgf73	
Spt7	Structural integrity
Spt20	
Ada1	
Tra1	Activator interaction
Spt3	TBP interaction
Spt8	
Taf5	TBP associated factors (TAFs)
Taf6	
Taf9	
Taf10	
Taf12	
Chd1	Chromatin remodeler

Subunits carrying out catalytic reactions are underlined.

As a transcriptional coactivator, SAGA bridges activators and the basal transcription machinery through different subunits (Figure 1.1). The Tra1 subunit of SAGA has been shown to directly contact acidic activators, such as Gcn4 and Gal4 (Grant *et al*, 1998b; Brown, 2001; Knutson & Hahn, 2011). In addition to Tra1, Taf12 has also been shown to interact with transcription activators (Reeves & Hahn, 2005). Both genetic and biochemical studies have shown that Spt3 and Spt8 subunits of SAGA interact with and recruit TBP to SAGA-dependent promoters (Sterner *et al*, 1999; Bhaumik & Green, 2002; Mohibullah & Hahn, 2008; Yu *et al*, 2003; Warfield *et al*, 2004; Laprade *et al*, 2007; Dudley *et al*, 1999; Sermwittayawong & Tan, 2006), however, the mechanism of these interactions and whether both subunits can simultaneously bind TBP are unclear.

The HAT and DUB enzymatic activities of SAGA are catalyzed by Gcn5 and Ubp8, respectively. The activity and specificity of these catalytic functions are dictated by non-catalytic subunits within each module. For example, Ubp8 is inactive unless associated with the three other subunits of the DUB module (Lang *et al*, 2011; Köhler *et al*, 2008). The X-ray structure of this module shows a complex intertwining of the DUB module subunits and provides an explanation for the interdependence of the subunits in DUB function (Samara *et al*, 2010; Köhler *et al*, 2010). Similarly, the substrate specificity of Gcn5 HAT activity is altered by its associated subunits, Ada2, Ada3, and Sgf29, to acetylate multiple residues within the nucleosomal histone H3 tail (Bian *et al*, 2011; Grant, 1999; Balasubramanian, 2001).

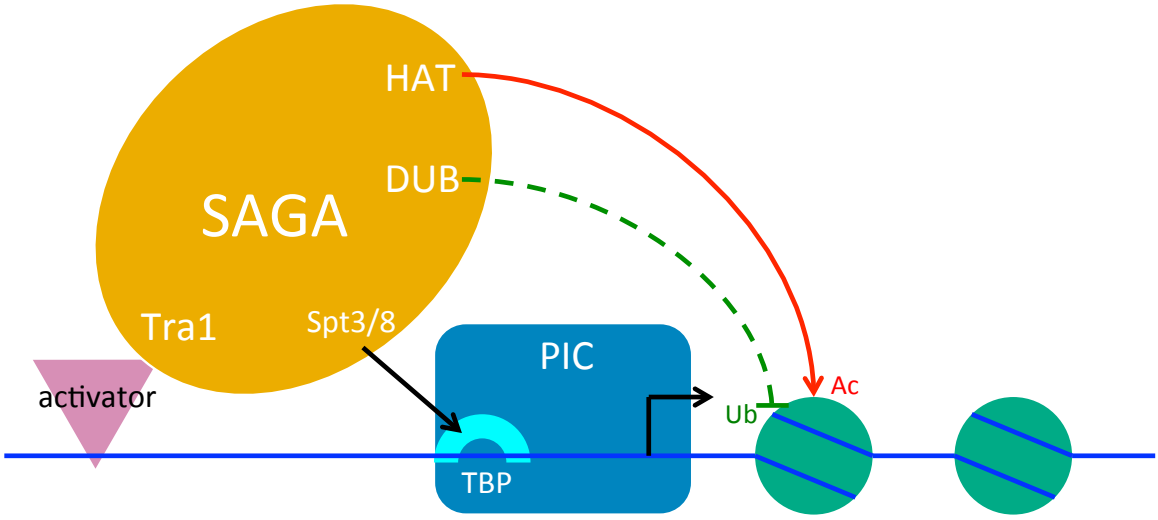

Figure 1.1 SAGA functions as a transcription coactivator. Adapted from Fuda et al, 2009 and Koutelou et al, 2010.

The chromatin remodeler Chd1 has also been suggested to be a component of the SAGA complex (Pray-Grant *et al*, 2005). The second chromodomain of Chd1 was shown to bind H3 peptides with K4 di- or trimethylated, thus linking histone methylation to acetylation within SAGA complex (Pray-Grant *et al*, 2005). However, recent purifications of SAGA from yeast did not recover Chd1 (Lee *et al*, 2011). This may reflect the different purification scheme, or strain backgrounds used in these studies, or indicate that Chd1 is not a specific subunit of SAGA.

1.2.2 SAGA functions outside of transcription

Besides its roles in transcription, SAGA has also been implicated in other cellular processes such as nucleotide excision repair (NER) and mRNA export (Rodríguez-Navarro, 2009; Baker & Grant, 2007) (Figure 1.2). Ultraviolet light (UV) can induce damage in DNA that can be fixed through the NER pathway. The mechanism of NER is complex, and involves many protein factors, including the general transcription factor TFIIH (de Laat *et al*, 1999). Gcn5, one of the HAT module subunits of SAGA, is responsible for the hyperacetylation of histone H3 at the repressed *MFA2* promoter (Yu *et al*, 2005), and deletion of Gcn5 slowed NER at this locus. At the *MET16* gene locus, both Gcn5 and Ada2 were implicated in affecting the repair process (Ferreiro, 2006). Thus, the HAT module of SAGA links the SAGA complex to NER at some gene loci.

Eukaryotic mRNAs are transcribed in the nucleus and must be transported into the cytoplasm to serve as a template for protein translation. Before being exported, mRNAs undergo several steps of processing, including 5'-capping, splicing, and 3'-polyadenylation. The resulting mRNPs (messenger ribonucleoproteins) are then exported

out of the nucleus through the nuclear pore complex (NPC). The SAGA DUB module subunit Sus1 is also a component of the RNA exporting complex TREX-2, which is composed of Cdc31, Sac3, Thp1 and Sus1 (Pascual-García & Rodríguez-Navarro, 2009; Fischer *et al*, 2004). The deletion of *SUS1* in yeast showed a defect in targeting the TREX-2 subunits Sac3 and Thp1 to the NPC, and blocked mRNA export (Pascual-García & Rodríguez-Navarro, 2009; Köhler *et al*, 2008). Furthermore, another DUB module subunit of SAGA, Sgf73, was also found to affect the integrity and NPC targeting of the TREX-2 complex (Köhler *et al*, 2008). Thus, the DUB module of SAGA is also linked to mRNA export. In addition, the DUB module has also been shown to modulate yeast replication lifespan through its interaction with Sir2 (McCormick *et al*, 2014). The C-terminus of Sgf73 has also been found to be essential for heterochromatin boundary formation (Kamata *et al*, 2013).

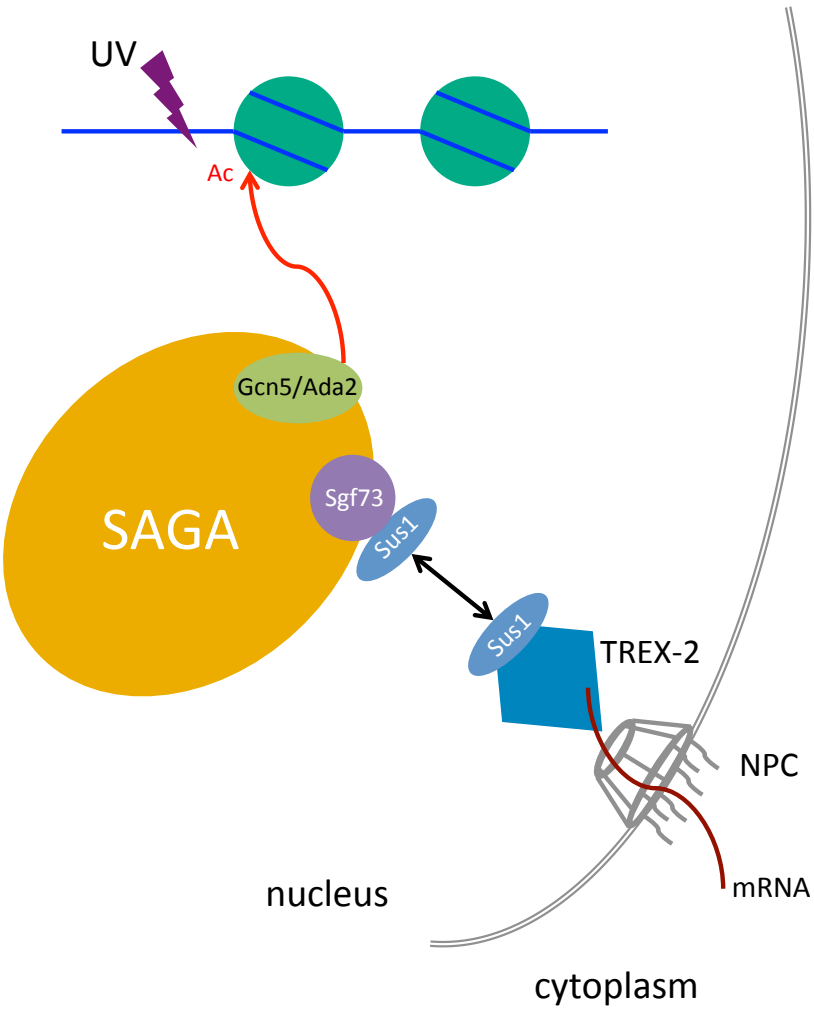

Figure 1.2 SAGA functions in NER and mRNA export. Adapted from Rodriguez-Navarro, 2009.

1.2.3 SAGA in development and disease

In metazoans, many lines of evidence have led to the conclusion that the SAGA complex is essential for development. First, histone H2B deubiquitination by SAGA is required for photoreceptor axon targeting in *Drosophila melanogaster* (Weake *et al*, 2008). Second, deletion of Gcn5 is detrimental to mouse embryogenesis, showing defects in mesoderm development (Xu *et al*, 2000). Third, loss of the acetyltransferase activity of Gcn5 leads to defects in neuronal development in mouse embryos (Bu *et al*, 2007). Fourth, both Gcn5 and USP22 (the metazoan homolog of yeast Ubp8) have been implicated in regulating the pluripotency and differentiation of embryonic stem (ES) cells (Wang & YR Dent, 2014). Thus, metazoan SAGA regulates a group of genes that are essential for development and viability.

SAGA has also been implicated in human diseases (Baker & Grant, 2007; Wang & YR Dent, 2014; Koutelou *et al*, 2010). Polyglutamine expansion in Ataxin-7, the human homolog of yeast Sgf73, causes spinocerebellar ataxia type 7 (SCA7), a neurodegenerative disease that leads to blindness and ataxia (Mohan *et al*, 2014). The polyglutamine expanded Ataxin-7 can incorporate into both SAGA and STAGA (human homolog of SAGA), and disrupt the HAT activity in a dominant-negative manner (McMahon *et al*, 2005; Palhan *et al*, 2005). Moreover, STAGA was shown to be recruited to c-Myc target genes by direct interactions between c-Myc and STAGA subunits TRRAP (human homolog of yeast Tra1) and Gcn5 (McMahon *et al*, 2000; Liu *et al*, 2003). Furthermore, USP22 was reported to act as an oncogene (Liu *et al*, 2012), and the aberrant expression of USP22 has been associated with many cancers (Wang & YR Dent, 2014). These observations, together with SAGA's role in NER, suggests that

misregulation of SAGA functions may lead to human cancers (Wang & YR Dent, 2014; Baker & Grant, 2007).

1.2.4 The SLIK/SALSA complex

SLIK (SAGA-like)/SALSA (SAGA altered, Spt8 absent) complex is similar to SAGA in subunit composition with two differences (Sterner *et al*, 2002; Pray-Grant *et al*, 2002). SLIK/SALSA contains a C-terminal truncated version of Spt7, which cannot associate with Spt8 (Sterner *et al*, 2002; Wu & Winston, 2002), one of the TBP-interacting subunits. Moreover, SLIK/SALSA contains a specific subunit Rtg2, implicating that the complex also functions in the retrograde response pathway (Pray-Grant *et al*, 2002; Daniel & Grant, 2007).

1.3 Structural studies of SAGA complex

Because of its complex composition, we know little about the overall structure of SAGA and how the different modules are organized. However, previous studies have solved the structures of smaller components of the complex. For example, the crystal structure of the DUB module shows the intertwined subunit organization (Samara *et al*, 2010; Köhler *et al*, 2010). Structures of the HAT module subunits have also been reported: the structure of the Gcn5 HAT domain reveals the substrate and co-enzyme binding sites (Trievel *et al*, 1999); the crystal structure of the Gcn5 bromodomain has also been solved (Owen *et al*, 2000), and a recent report provides evidence that the bromodomain regulates the lysine specificity of histone H3 acetylation (Cieniewicz *et al*, 2014); the structure of the Sgf29 Tudor domain shows the H3K4me2/3 binding activity, which is required for SAGA recruitment to promoters and histone H3 acetylation (Bian *et al*, 2011). Recently,

the core structure of TFIID containing two copies of each Taf5, 6, and 9 has been determined by cryo-electron microscopy (EM), revealing the symmetric organization of these Tafs (Bieniossek *et al*, 2013). Since these Taf proteins are shared between SAGA and TFIID, it seems very likely that they also form a similar structure within the SAGA complex.

Although no structure has been determined for Spt3, sequence analysis showed that it contains N- and C-terminal histone fold domain homologous to Taf13 and Taf11 (Birck *et al*, 1998), respectively. The histone folds of Taf11 and Taf13 form a heterodimer, and a structural model of Spt3 has been proposed based on the crystal structure of the heterodimer of Taf11 and Taf13, with the N- and C-terminus folding together (Birck *et al*, 1998). The Taf11 histone fold domain has been shown to bind directly with TBP (Mengus *et al*, 1995). Interestingly, a mutation within Spt3, E240K, which was suggested to result in a more stable SAGA/TBP interaction compared to wild type SAGA (Laprade *et al*, 2007), is also located in the homologous histone fold domain. Moreover, bioinformatic analysis has also provided some structural insight for the largest SAGA subunit Tra1, revealing four different α-helical domains flanking the inactive PI3K domain (Knutson & Hahn, 2011).

Together, these initial findings suggest a series of complex interactions between different SAGA subunits and modules, perhaps somewhat analogous to the complex interactions between subunits of the DUB module. In 2004, Wu *et al* reported a negative-stained electron microscopy model of the yeast SAGA complex (Wu *et al*, 2004). They used

antibody-labeling and localized a subset of subunits within the SAGA EM model. A recent systematic mass spectrometry analysis of SAGA complex purified from strains lacking individual non essential SAGA subunits led to a model for the composition of each module and organization of the complex (Lee *et al*, 2011). However, the relative locations of the Taf, Spt, and Tra1 subunits are different in the two models.

To understand how the diverse functions of SAGA are coordinated to modulate gene expression, it is important to determine the overall architecture of SAGA, the arrangement of the different modules, and whether these modules interact. In my thesis, I have used a combination of chemical crosslinking and mass spectrometry, together with biochemical and genetic analyses to study the molecular architecture of the SAGA complex and its interaction with TBP.

Chapter 2 Chemical crosslinking and mass spectrometry analysis of SAGA

2.1 Introduction

Determining the overall structure and the subunit organization of SAGA is important for further understanding the multiple roles it plays in various cellular processes. However, its large size and dynamic composition have hindered structural studies using X-ray crystallography and nuclear magnetic resonance (NMR) analyses. One decade ago, using EM methodology, Wu *et al* determined the overall structure of SAGA at 3.1nm resolution (Wu *et al*, 2004). Although this was a breakthrough in determining SAGA's structure, more needs to be done to comprehensively understand SAGA's function and structure. For example, the DUB module had not been identified as an important SAGA component in 2004. Therefore, an alternative approach was required to resolve the molecular architecture of SAGA.

Chemical crosslinking combined with mass spectrometry (CXMS) is an alternative method to X-ray crystallography and NMR for determining the architecture of protein complexes (Back *et al*, 2003; Sinz, 2006; Walzthoeni *et al*, 2013). Covalently linking residues by chemical crosslinkers within protein complexes under near-physiological conditions provides useful information about the structural organization of the complexes. For example, in vivo incorporated photocrosslinker BPA (ρ-benzoyl-phenylalanine) at specific surface positions of TBP identified several TBP-interacting factors within the PIC (Mohibullah & Hahn, 2008). However, in such experiments, it is difficult to use the distance restrains provided by the crosslinkers to deduce the three-

dimensional organization of the protein complex, simply because it is difficult to identify the crosslinked residue of the target protein. In addition, because this approach involves site-specific incorporation of the crosslinker into a bait protein, it is difficult to generate comprehensive crosslinking results especially for large complexes. CXMS takes advantage of mass spectrometry (MS) to identify the crosslinked residues within proteins, and has become a popular method to study the architecture of many multiprotein complexes (Walzthoeni *et al*, 2013). MS analysis is normally very fast and requires very little amount of samples, and CXMS can shed light on the structural interactions within and between flexible regions (Sinz, 2006). By applying this approach, several groups have deciphered the interaction network of protein phosphatase 2A (Herzog *et al*, 2012), and the architecture of the RNA Pol II transcription apparatus (Chen *et al*, 2010; Mühlbacher *et al*, 2014). CXMS has also been applied in combination with other structural approaches to model the architecture of large multiprotein assemblies. For example, in combination with EM analysis, CXMS was used to map the subunit interactions within the INO80 chromatin remodeling complex (Tosi *et al*, 2013), the nuclear pore complex (Bui *et al*, 2013), and the yeast PIC (Murakami *et al*, 2013). In addition, homology modeling together with CXMS has generated structural models for the TRiC/CCT chaperonin complex (Kalisman *et al*, 2012), and the yeast RNA Pol I core factor complex (Knutson *et al*, 2014).

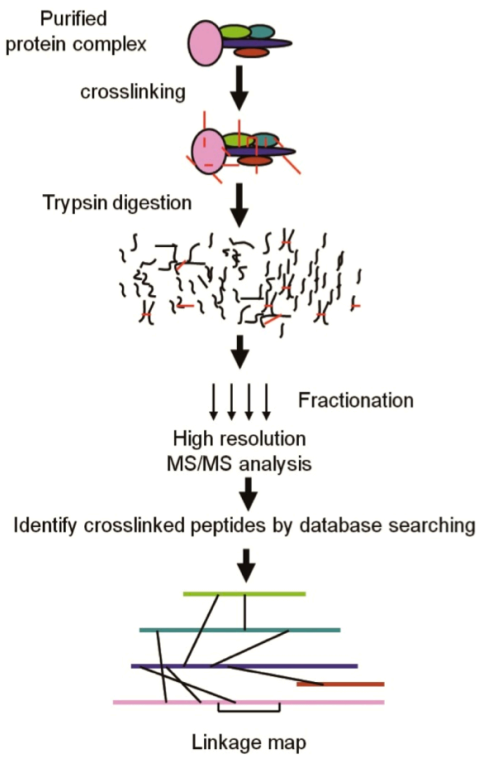

Figure 2.1 Schematic of a CXMS approach. Adapted from Luo *et al*, 2012.

Figure 2.1 shows a schematic of a commonly used CXMS approach. Briefly, the protein complex of interest is purified or recombinantly reconstituted. This may be the limiting step for many other types of structural analyses, because the yield of large protein complexes is not always adequate. However, the sensitivity of the MS approach can overcome this issue (Aebersold & Mann, 2003). Next, the protein complex is crosslinked in solution under a condition that is close to the physiological situation. In this step, many chemical crosslinkers are available with different reaction groups and spacer distances (Sinz, 2006). It is worth noting that MS labile crosslinkers have been developed to facilitate the database searching and peptide identification (Sinz, 2006; Luo *et al*, 2012). The crosslinked protein complex is then digested with a protease such as trypsin and the resulting peptides are analyzed by electrospray ionization MS using a high resolution Orbitrap mass spectrometer. During MS analysis, the m/z ratio of the peptides, including crosslinked peptides, are measured. Next, the MS selects peptides for MS2 analysis during which the peptides are fragmented by collision with inert gas. This generates an MS2 spectrum that contains information about the peptide sequence. Database searching algorithms use this information to identify the crosslinked peptides and the sites of crosslinking, and these identifications are then combined together to generate a site-specific linkage map of the protein complex.

It is also worth noting that there are limitations of the CXMS approach. For example, many crosslinkers react with primary amine groups of lysine side chains, and the commonly used protease trypsin will not cleave the peptide bond at the C-terminal side of the modified lysine residues, resulting in large peptides due to missed cleavage (Sinz,

2006). These large peptides may not be detected during MS analysis if their m/z ratios exceed the maximum measured m/z ratio. Moreover, crosslinked peptides with low charge state due to the modification of lysine residues may not be detected during MS analysis (Sinz, 2006). In the following sections of this chapter, I will describe my experiments using the CXMS approach to map the subunit interactions within the SAGA complex.

2.2 Formation of the SAGA-TBP complex

To determine the molecular architecture of the SAGA complex and its interactions with TBP, I first purified yeast SAGA using a TAP tag at the C-terminus of subunit Spt7. This approach selectively purifies SAGA and not the related SLIK complex, which lacks the C-terminus of Spt7 (Wu & Winston, 2002). Purified SAGA was visualized on a silver stained protein gel (Figure. 2.2A lane 3). Mass spectrometry analysis of this preparation identified all known SAGA subunits with the exception of Chd1 (Table 2.1), a chromatin remodeler that reportedly associates with SAGA (Pray-Grant *et al*, 2005). In agreement with most previous findings, TBP did not copurify with SAGA, indicating a weak association. As a first attempt to generate the SAGA-TBP complex, I purified SAGA from a strain with the *spt3-401* allele, containing the Spt3 mutation E240K, which was suggested to bind TBP more stably compared to wild type SAGA (Laprade *et al*, 2007). Under the purification conditions, however, this SAGA derivative was not stably associated with TBP (Figure 2.2A, lane 5). As an alternative approach for finding conditions to generate the SAGA-TBP complex, purified SAGA was immobilized to calmodulin beads via the calmodulin-binding tag on Spt7. Immobilized SAGA was incubated with recombinant TBP for 90 min, washed, and then eluted from the beads

using EGTA. Both wild type and the Spt3-401 SAGA derivative specifically bound TBP (Figure 2.2A, lanes 4,6,7). Quantitation of the TBP/Ada1 ratio for the two preparations showed that the Spt3-401 SAGA bound about 4-fold more TBP compared to wild type SAGA (Figure 2.2B). Since wild-type SAGA bound TBP under these conditions, TBP and wild type SAGA were combined in a 1:1.2 molar ratio and used as a substrate for subsequent crosslinking experiments.

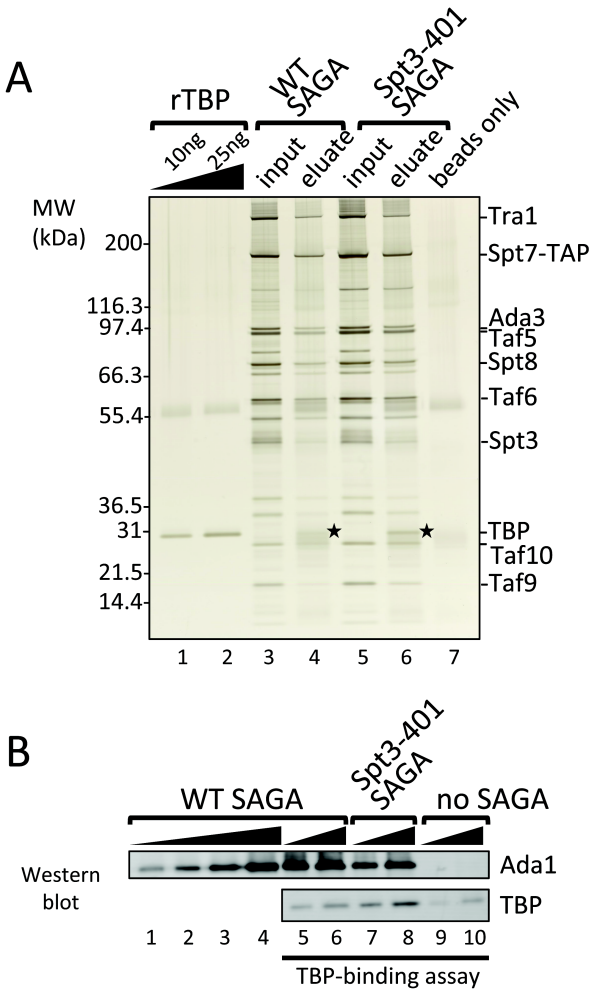

Figure 2.2 TAP purified SAGA complex binds TBP. (**A**) A silver stained protein gel showing the TBP binding assay of wild type and Spt3-401 (E240K) SAGA complexes. 3 μg SAGA was immobilized to Calmodulin beads and incubated with 0.5 μg recombinant TBP (rTBP). EGTA eluted samples were analyzed for TBP binding. Stars denote bound TBP. (**B**) Western blot analysis of the above TBP binding assay probing for Ada1 and TBP.

Table 2.1 Sequence coverage of WT SAGA complex.

SAGA subunit	Sequence coverage [%]					
	a	b	c	d	2mM	5mM
Gcn5	81.8	81.8	81.8	81.5	34.4	21.9
Ada2	77.6	77.6	77.6	77.4	26.3	24.4
Ada3	77.6	78.8	74.9	78.8	48.3	31.6
Sgf29	81.9	81.9	78.8	81.9	53.3	49
Ubp8	82	82.8	80.7	82.8	29.9	16.6
Sgf11	77.8	77.8	77.8	77.8	0	0
Sgf73	50.5	50.5	50.5	50.5	27.7	22.2
Sus1	85.4	85.4	85.4	85.4	21.9	0
Spt7	84.2	86.9	85.7	86.4	40.9	33.1
Spt8	77.1	77.1	77.1	77.1	44	35
Spt20	73.8	74	73.8	73.7	47	45.2
Spt3	62.9	62.9	62.9	62.9	27.3	27
Ada1	72.1	74.2	74.2	71.9	24.4	24.4
Tra1	77.8	78	78.6	79.6	26.1	21.9
Taf10	64.1	67.5	67.5	67.5	35.4	24.3
Taf12	72	75.3	67.5	82.4	33.6	31.9
Taf9	77.7	79	79	79	52.2	52.9
Taf6	68.4	68.6	68.6	68.6	46.3	39.1
Taf5	71.9	73.9	79.9	79.9	45.4	48.1
Chd1	0	0	0.8	0.7	0	0

Sequence coverage was calculated using MaxQuant. Subscript a-d denotes uncrosslinked SAGA preparations that were the same data set as in the iBAQ analysis. 2mM and 5mM indicate the crosslinked complexes in 2mM and 5mM BS3 crosslinking experiments, respectively.

2.3 Activity test of purified SAGA complex

If a protein complex is functional after purification, it is likely that it retains an *in vivo* relevant conformation. SAGA is a multisubunit complex that has many functions. One of the important functions that SAGA possesses is HAT activity. Gcn5 catalyzes the acetylation reaction with histone H3 as its substrate (Grant *et al*, 1997). In order to make sure that the SAGA complex I purified was in an active state, I decided to first test the HAT activity of my purified SAGA complex in an *in vitro* histone H3 acetylation assay. Figure 2.3 shows that both WT and Spt3-401 SAGA (SAGA bearing the Spt3-401 mutant) are able to acetylate histone H3 *in vitro*. However, it is worth noting here that the HAT activity observed in Figure 2.3 can only suggest that Gcn5 copurified with TAP-tagged Spt7, because Gcn5 alone is able to acetylate histone H3. To further test the HAT activity of SAGA, it is necessary to do similar experiments using a chromatin template, as Gcn5 can only acetylate a chromatin substrate when it is incorporated into SAGA (Grant *et al*, 1997). Indeed, the purified SAGA complex was able to acetylate histone H3 in the context of a mononucleosome (see Chapter 3), showing that the purified SAGA complex retained HAT activity.

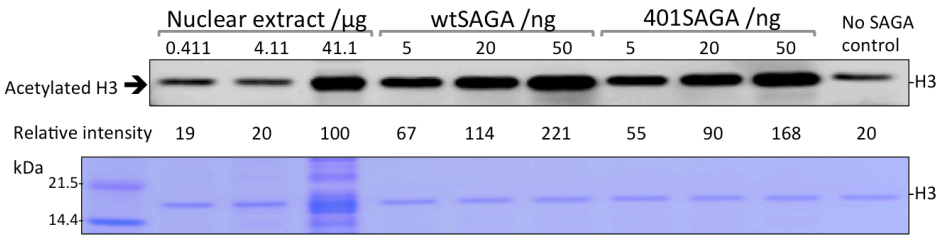

Figure 2.3 SAGA is active as HAT. TAP purified SAGA complexes can acetylate histone H3 *in vitro* (western blot in upper panel). Numbers under the western blot indicate the relative band intensity, with 41.1µg nuclear extract control set to 100. The Coomassie blue stained gel in the lower panel shows that equivalent amount of histone H3 was added in each reaction.

2.4 Chemical crosslinking and mass spectrometry of SAGA

The SAGA-TBP complex was crosslinked using the bifunctional crosslinking reagent BS3, digested with trypsin, and the peptides analyzed by mass spectrometry. The BS3 crosslinker reacts with accessible primary amine groups found in lysine side chains and at the N-termini of polypeptides (Figure 2.4A). Each SAGA subunit, except Sgf11, contains many lysine residues that are distributed throughout the polypeptides (Figures 2.5, 2.6), making SAGA a good substrate for this crosslinking strategy. A database search algorithm pLink (Yang *et al*, 2012) was used to identify the crosslinked peptides, and the identifications were then used to assemble site-specific linkage maps of all the crosslinked residues between and within the SAGA subunits and TBP (Figures 2.5, 2.6).

SAGA-TBP was crosslinked with either 2 or 5 mM BS3 (Figure 2.4B) and the reactions were independently analyzed by mass spectrometry for peptide crosslinks. Results from both experiments were combined. Altogether, I identified 199 intermolecular crosslinks (between two different subunits) and 240 intramolecular crosslinks (within a single subunit) (Figures 2.4C, 2.5, 2.6, Table A.1). To validate the crosslinking approach, I mapped the crosslinked lysine pairs within known structured domains of SAGA and TBP. BS3 has a theoretical Cα- Cα crosslinking distance of 5~30Å between two lysine residues (Chen *et al*, 2010; Kalkhof & Sinz, 2008; Mädler *et al*, 2009; Müller & Sinz, 2012; Merkley *et al*, 2014). Within domains for which high-resolution structural information is available, 17 out of the 19 crosslinks were within the theoretical crosslinking distance for BS3 (Table 2.2).

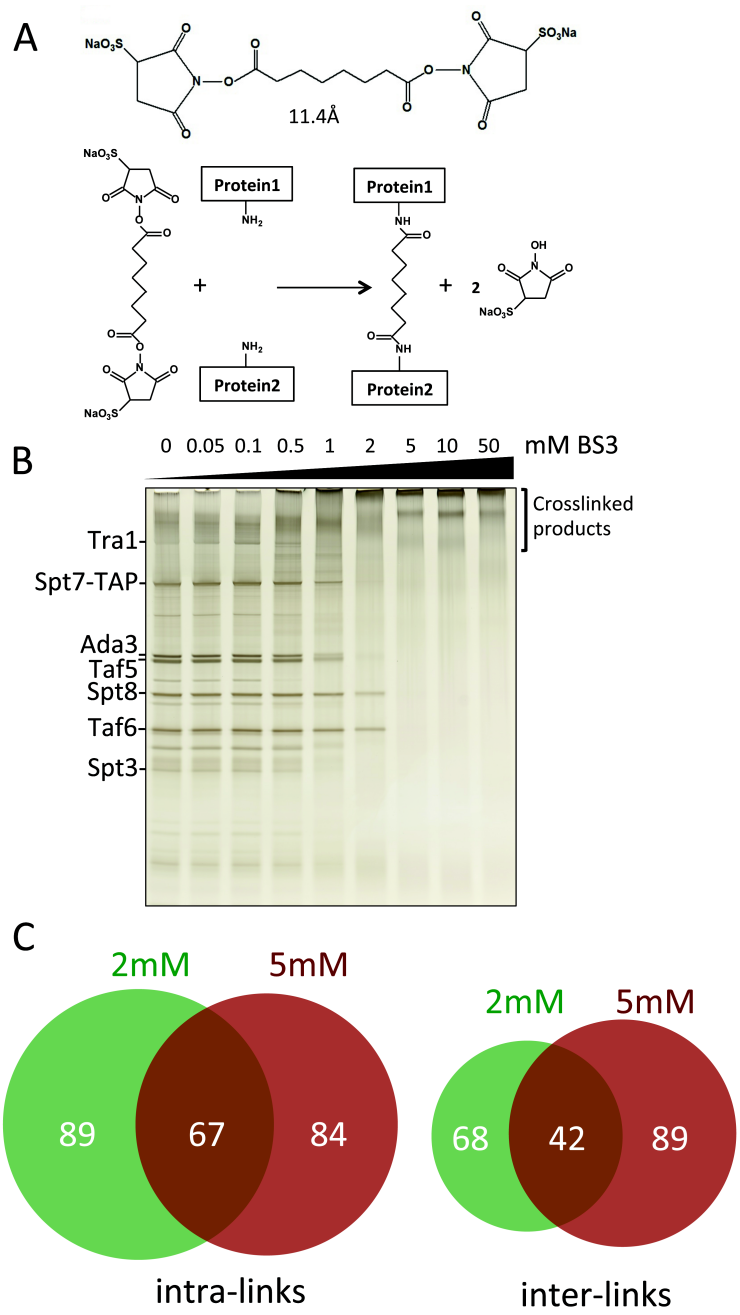

Figure 2.4 BS3 crosslinking of SAGA complex. (**A**) Chemical structure of BS3 (upper panel) and crosslinking mechanism (lower panel). (**B**) Silver stained protein gel showing the crosslinking of SAGA complex with increasing amounts of BS3. (**C**) Venn diagram showing the number of crosslinks identified in experiments with 2 mM or 5 mM BS3. 199 total intermolecular and 240 intramolecular crosslinks were identified. Data is from Table A.1.

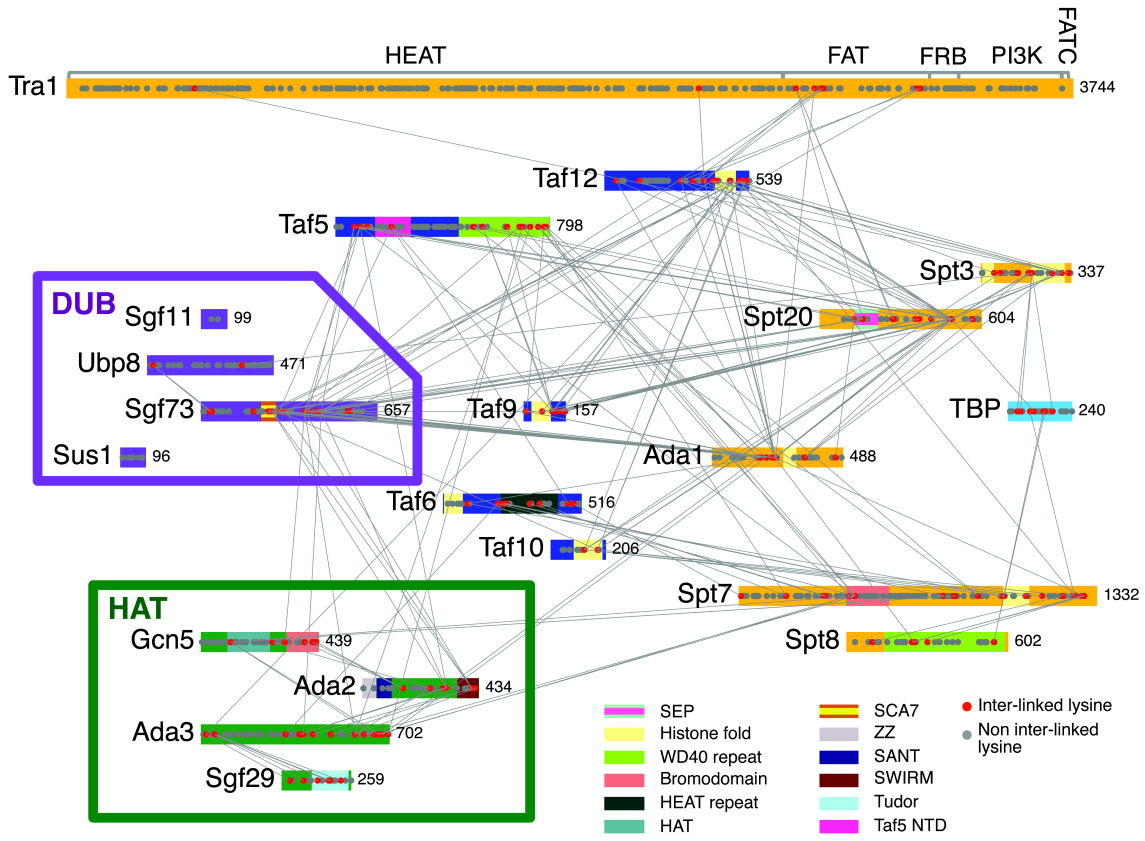

Figure 2.5 BS3 intermolecular crosslinking map of SAGA-TBP complex. SAGA subunits are color-coded according to the different modules they belong to. Domains are highlighted as indicated based on previous studies or HHpred predictions (with >95% probability score). Grey and red dots represent lysine residues (not crosslinked and crosslinked respectively), while grey lines connecting red dots denote intermolecular crosslinked lysine pairs. SEP: Shp1, Eyc and p47; SCA7: spinocerebellar ataxia type 7; ZZ: zinc-binding domain; SANT: SWI3, ADA2, N-CoR and TFIIIB" DNA-binding domain; SWIRM: SWI3, RSC8 and MOIRA; Taf5 NTD: Taf5 N-terminal domain; FAT: FRAP, ATM and TRRAP; FRB: FKBP12 rapamycin binding; FATC: FAT C-terminal.

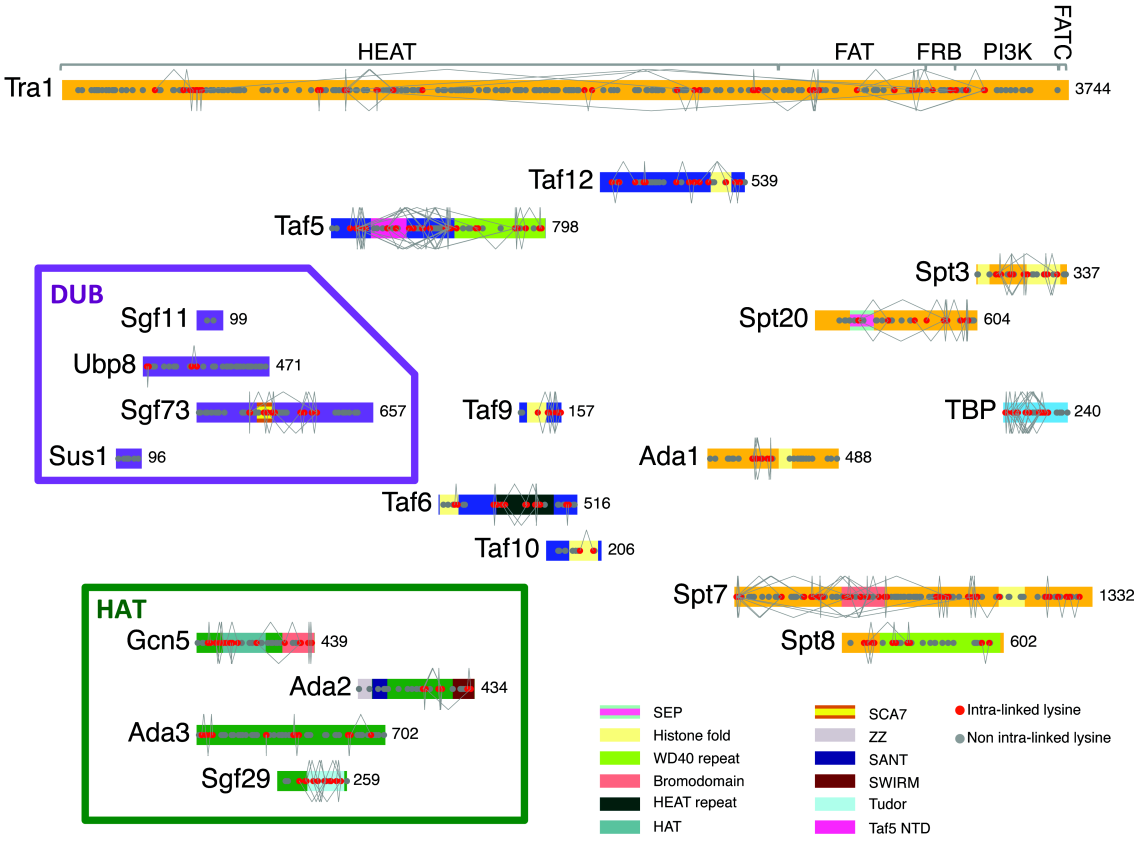

Figure 2.6 SAGA intramolecular crosslinking map. Similar to Figure 2.5, but only the identified intramolecular crosslinks are shown.

Table 2.2 Distance measurements of crosslinked residues in known 3D structures.

Subunit (PDB ID)	Residue 1	Residue 2	Cα distance (Å)
Ubp8 (3MHH)	15	22	10.5
	181	T200	23.7*
Sgf29 (3MP8)	134	181	8.8
	151	177	9.1
	200	237	16.1
	220	237	16.1
	134	237	25.6
	177	237	36**
Gcn5 bromodomain (1E6I)	422	429	10.5
	415	422	10.8
	R329	385	13.1*
Gcn5 HAT domain (1YGH)	126	133	8.9
	111	153	16.6
TBP (1YTB)	133	156	8.7
	83	97	12.8
	83	145	15.2
	83	167	49.6**
DUB (3MHH)	Ubp8 K22	Sgf73 K33	11.8
	Ubp8 K22	Sgf73 K40	15

* One of the lysines is missing in the structure. Distance to the Cα of the adjacent residue (R329 instead of K328 in Gcn5 bromodomain; T200 instead of K201 in Ubp8) was measured.
** Distances that are larger than the 30Å cut-off.

2.5 A TFIID-like core complex at the center of SAGA

SAGA and TFIID share a subset of Tafs that are important for the function of both complexes. A recent study reported the cryo-EM structure of a human TFIID subcomplex, comprised of two copies each of Tafs 4, 5, 6, 9 and 12 (Bieniossek *et al*, 2013). Addition of the Taf8-Taf10 heterodimer to this complex generates an asymmetric TFIID core complex that acts as a scaffold to assemble the remaining TFIID subunits. All of the above-mentioned Tafs are also subunits of SAGA, with the exceptions of Taf4 and Taf8. The SAGA subunit Ada1 has a histone fold domain homologous to that of Taf4 and forms a histone fold pair with Taf12 (Gangloff *et al*, 2000), while Spt7 has a histone fold domain homologous to that of Taf8 and pairs with Taf10 (Gangloff *et al*, 2001). Indeed, I observed crosslinks between Taf12 and Ada1, and between Taf10 and Spt7, close to the histone fold domains (Figure 2.5), indicating that these proteins likely form heterodimers through histone fold pairing within the SAGA complex. I also observed crosslinks between all of the other Taf subunits except for Taf6 and Taf9, which are known to form a histone fold pair within TFIID (Bieniossek *et al*, 2013). Together with previous findings on the TFIID core complex, the combined results suggest that SAGA contains a TFIID-like core. As described in more detail below, this TFIID-like core complex crosslinks to all other SAGA modules (Figure 2.7A). Thus, similar to its role in TFIID, this TFIID-like core is likely positioned at the center of SAGA where it functions to nucleate assembly of other SAGA modules.

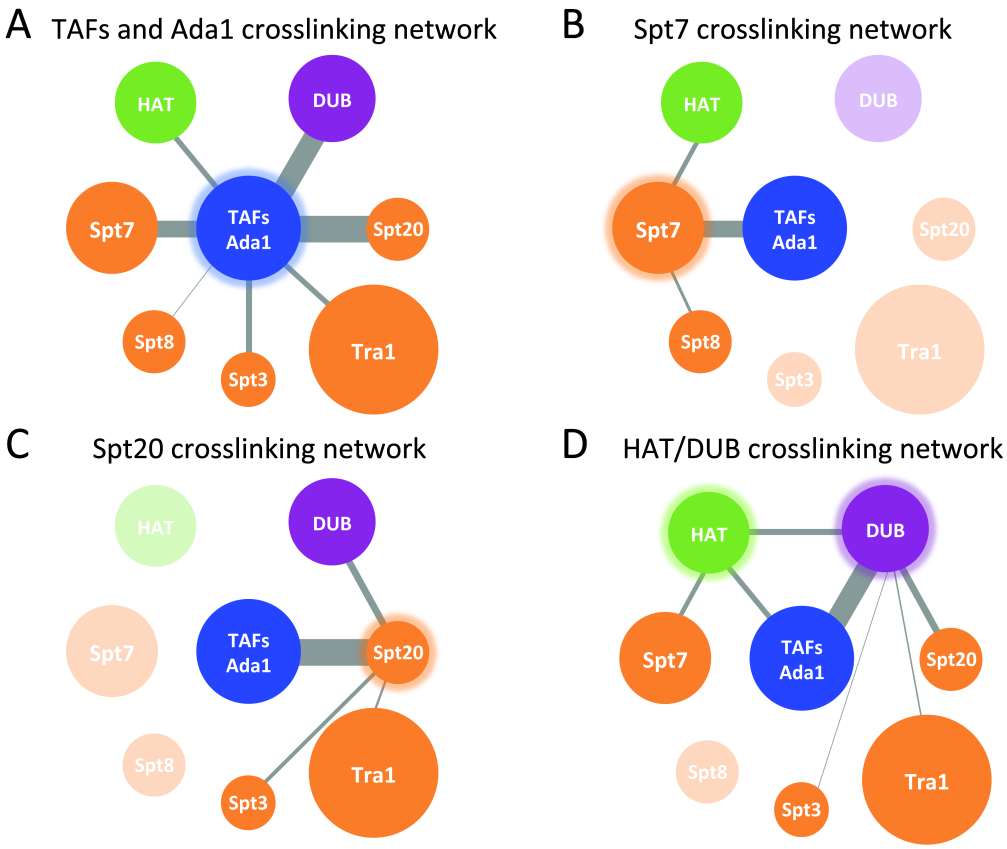

Figure 2.7 Crosslinking network of the SAGA complex. Crosslinking network for (**A**) TAFs/Ada1, (**B**) Spt7, (**C**) Spt20, (**D**) HAT and DUB modules. SAGA subunits or modules are represented as spheres, and are color-coded similar to Figure 2. Grey lines connect crosslinked SAGA components, with the width of the lines proportional to the number of intermolecular crosslinks. Transparent spheres in (**B**), (**C**) and (**D**) denote the subunits or modules that do not crosslink to the highlighted components.

2.6 Subunits Spt7 and Spt20 crosslink to distinct SAGA modules

The SAGA subunits Spt7, Spt20, and Ada1 were initially proposed to be components of a core module (Wu & Winston, 2002), as mutation of any of these subunits disrupted conventional purification of SAGA (Sterner *et al*, 1999). This view was supported by systematic deletion analysis, suggesting that all Spt subunits were part of a common module (Lee *et al*, 2011). In contrast, mapping Spt7 and Spt20 within the SAGA EM structure led to an alternative view suggesting that these two subunits are positioned at different ends of a large structural module (Wu *et al*, 2004). Consistent with this latter view, the crosslinking results suggest distinct roles for these two subunits in SAGA assembly as Spt7 and Spt20 crosslink to different sets of subunits but do not crosslink to each other.

I found that Spt7, part of the TFIID-like module, crosslinks to Spt8, one of the TBP binding subunits (Figure 2.7B). This interaction occurs primarily between the C-terminus of Spt7 and WD40 repeats at the center of Spt8. This is in perfect agreement with previous biochemical results, which found that the 152 C-terminal residues of Spt7 are required for Spt8 binding to SAGA (Wu & Winston, 2002). Spt7 also crosslinks to Gcn5 and Ada3, showing close positioning of Spt7 to the HAT module. In contrast, Spt20 crosslinks to neither Spt7, Spt8, nor the HAT module (Figure 2.7C). Instead, I observed extensive crosslinking between Spt20 and the TFIID-like core domain, the DUB module, and the TBP binding subunit Spt3.

Tra1, an activator-binding subunit of both SAGA and the coactivator NuA4, crosslinks to Spt20 at several sites within its FAT domain. The FAT domain also crosslinks to TAF12,

in a region of Taf12 that also crosslinks to Spt20. Thus, Tra1 seems to primarily associate with SAGA via interactions between its FAT domain and segments of Taf12 and Spt20 that are in close proximity (Figures 2.4, 2.7). While the sparse Tra1-SAGA crosslinking outside of the FAT domain does not rule out other Tra1 interactions, the observation that the major Tra1-SAGA crosslinking is localized within the relatively small Tra1 FAT domain, together with the findings that the FAT domain crosslinks with two closely positioned SAGA subunits, suggests a peripheral position for the large Tra1 subunit within the SAGA complex, in agreement with a previous model (Wu *et al*, 2004).

Interestingly, I also observed a number of crosslinks between the DUB and HAT modules (Figures 2.5, 2.7D), suggesting an interaction between the two chromatin modifying activities within SAGA complex. I will focus on this interaction in the following chapter.

2.7 Materials and methods

2.7.1 TAP purifications.

Tandem affinity purifications were performed as previously described but with modifications (Wu *et al*, 2004). Briefly, 12 liters of yeast were grown to an optical density at 600nm (OD_{600}) of 2-5 in YPD (3% glucose). Next, cells were harvested by centrifugation and washed with 200 ml of cold TAP extraction buffer (40 mM HEPES pH 7.5, 350 mM NaCl, 10% glycerol, 0.1% Tween 20, 1 mM dithiothreitol [DTT], 1 mM phenylmethylsulfonyl fluoride [PMSF], 0.31 mg/ml benzamidine, 0.3 μg/ml leupeptin, 1.4 μg/ml pepstatin, 2 μg/ml chymostatin). Cells were resuspended in 150 ml cold TAP extraction buffer and lysed in a BeadBeater (Biospec Products). Cell debris was removed by centrifugation at 15,000 ×g at 4°C for 30 min. The whole cell extracts (WCE) were

then cleared by ultracentrifugation at 150,000 ×g at 4°C for 90 min. For TAP-tag purifications, 2.5 ml IgG Sepharose beads (GE Healthcare) were incubated with the WCE at 4°C overnight. The beads were next washed and resuspended in 4 ml cold TEV (tobacco etch virus) cleavage buffer (10 mM Tris pH8, 150 mM NaCl, 0.1% NP-40, 0.5 mM EDTA, 10% glycerol). 25 μg of TEV protease (TEVp) was added to the IgG beads, and the cleavage was performed at 4°C overnight. On day three, the TEVp-cleaved products were collected, and the IgG beads were washed with 3 column volumes (~7.5 ml total) cold Calmodulin binding buffer (15 mM HEPES pH8, 1 mM MgOAc, 1 mM imidazole, 2 mM $CaCl_2$, 0.1% NP-40, 10% glycerol, 300 mM NaCl). $CaCl_2$ was added to the combined eluate at a final concentration of 2 mM and incubated with 1 ml Calmodulin affinity resin (Agilent Technologies) at 4°C for 4 hours. After incubation, the beads were washed with cold Calmodulin binding buffer and cold Calmodulin wash buffer (same as Calmodulin binding buffer, but containing 0.01% NP-40), and bound proteins were eluted in eleven 0.5 ml fractions with Calmodulin elution buffer (15 mM HEPES pH 8, 1 mM MgOAc, 1 mM imidazole, 2 mM EGTA, 10% glycerol, 300 mM NaCl, 0.01% NP-40) at room temperature. TAP purified complexes were analyzed on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (4-12% acrylamide) gradient gel (Invitrogen) in MOPS buffer. For Spt3 E240K SAGA purification, the Calmodulin binding buffer, Calmodulin wash buffer, and Calmodulin elution buffer contained only 150 mM NaCl. For the purifications of complexes bearing mutations in Sgf73 or Ada3 (see Chapter 3), only 4 liters of yeast cells were grown. The amount of buffer and beads used was adjusted accordingly. Protein concentrations of SAGA complexes were determined by absorbance at 280 nm.

2.7.2 TBP binding assay.

EGTA eluted SAGA complexes (~3 μg) were diluted in 500 μl Calmodulin binding buffer. The diluted complexes were then incubated with 20 μl Calmodulin affinity resin (Agilent Technologies) for 3 hr at 4°C. Beads were washed and blocked with 10% bovine serum albumin (BSA) for 2 hr at 4°C. After blocking, the protein-bound beads were incubated with 0.5 μg recombinant yeast TBP for 1.5 hr at 4°C. Proteins were eluted with 40 μl Calmodulin elution buffer. The eluates were analyzed by SDS PAGE, and visualized by silver stain and Western blot using rabbit polyclonal antibodies against yeast TBP and Ada1. TBP and Ada1 were quantified using the Odyssey infrared imaging system (LiCor Biosciences).

2.7.3 BS3 crosslinking.

Wild type SAGA complex was mixed with 20% molar excess of recombinant TBP. 50 μg of this complex was incubated with 2 mM or 5 mM BS3 (Thermo Scientific) in 200 μl total volume in Calmodulin elution buffer at room temperature for 2 hr. Reactions were quenched with 10 μl 1 M Tris pH7.5. Small-scale titrations were performed with ~3 μg of SAGA complex in 20 μl total volume with various amount of BS3 as shown in Figure 2.4B. Reactions were analyzed by SDS-PAGE and silver staining (Invitrogen).

2.7.4 MS sample preparation and data analysis.

Protein samples were denatured in 50% trifluoroethanol (TFE) at 60°C for 30 min and reduced by addition of 5 mM TCEP at 37 °C for 30 min. The samples were then alkylated in the dark at room temperature with 10 mM iodoacetamide for 30 min. The samples were diluted 10-fold with 100 mM ammonium bicarbonate and digested with

trypsin (3:50 w/w) overnight at 37 °C. Tryptic peptides were then purified by C18 chromatography (Waters), and dried in a speedvac. For crosslinked samples, dried peptides were resuspended in 30 µl buffer A (25 mM ammonium formate, 20% acetonitrile [ACN], pH 2.8), and then fractionated by microcapillary partisphere strong cation exchange (SCX, 5 µm, 200 Å, GE Healthcare) For separation of crosslinked peptides by HPLC, peptides were loaded onto the capillary column equilibrated in 5% ACN/0.1% trifluoroacetic acid (TFA) and washed with 20% ACN/0.1% formic acid (FA). Bound peptides were eluted with 50 µl of buffer A with 10%, 30%, 50%, 70%, 90% buffer B (800 mM ammonium formate, 20% ACN, pH 2.8), followed by 50 µl 100% buffer B and 50 µl buffer C (1 M ammonium acetate, 10% ACN, 0.1% FA, pH8). All fractions were then dried in a speedvac. Most crosslinked peptides were eluted in fractions with 50-90% buffer B.

All dried peptides were analyzed on a Thermo Scientific Orbitrap Elite at the Proteomics facility at the Fred Hutchinson Cancer Research Center (FHCRC, Seattle, Washington, USA) with HCD fragmentation and serial MS events that included one FTMS1 event at 30,000 resolution followed by 10 FTMS2 events at 15,000 resolution. Other instrument settings included: Charge state rejection: +1, +2, +3 and unassigned charges; Monoisotopic precursor selection enabled; Dynamic exclusion enabled: repeat count 1, exclusion list size 500, exclusion duration 30s; HCD normalized collision energy 35%, isolation width 3Da, minimum signal count 5000; FTMS MSn AGC target 50,000, FTMS MSn Max ion time 250ms. Online microcapillary C18 HPLC was performed using a 90 min gradient from 5% ACN to 40% ACN for crosslinked fractions and a 60 min gradient

from 5% ACN to 35% ACN for normal non-crosslinked peptides. HPLC buffers contain 0.1% FA.

Crosslinking data were analyzed using pLink (Yang *et al*, 2012) with default settings (precursor monoisotopic mass tolerance: ±10 ppm; fragment mass tolerance: ±20 ppm; up to 4 isotopic peaks; max evalue 1; static modification on Cysteines; 57. 0215 Da; differential oxidation modification on Methionines; 15. 9949 Da) against a database containing 22 proteins (21 SAGA subunits and TBP), and the search engine outputs results with 5% FPR. Each spectrum was manually evaluated for the quality of the match to each peptide using the COMET/Lorikeet Spectrum Viewer (Trans-Proteomic Pipeline, TPP). The crosslinked peptides were considered confidently identified if at least 3 consecutive b or y ions for each peptide were observed with minimum peptide length >=4 and the majority of the observed ions are accounted for. 14% of spectra were removed after manual inspection.

Chapter 3 DUB modestly enhances the HAT activity

3.1 Introduction

Both the HAT and DUB modules function to enhance transcription (Weake & Workman, 2010). Recruitment of SAGA to promoters allows targeted acetylation of the H3 tail by Gcn5, promoting the binding of other regulatory factors and generating a permissive environment for chromatin remodeling. The DUB module removes ubiquitin from H2B, a histone modification that transiently increases at the 5' end of transcribed regions upon gene induction and is linked to transcription elongation. Ubiquitin removal allows subsequent recruitment of the CTD Ser2 kinase CTK1 to coding regions (Wyce *et al*, 2007), promoting efficient transcription elongation and the placement of additional downstream chromatin marks.

The crystal structure of the DUB module reveals the intertwined nature of the subunits within the module (Samara *et al*, 2010; Köhler *et al*, 2010). Although Ubp8 requires the other three DUB subunits to be fully active (Lang *et al*, 2011; Köhler *et al*, 2008), Sgf73 is required to connect the DUB module to the rest of the SAGA complex (Köhler *et al*, 2008; Lee *et al*, 2009). Both Ada2 and Ada3 are required for linking Gcn5 to SAGA (Horiuchi *et al*, 1995; Candau & Berger, 1996). In addition, several lines of evidence show potential crosstalk between the DUB and HAT modules. For example, a mutant Poly Q-containing form of the DUB module subunit Ataxin7 (the human ortholog of yeast Sgf73), was shown to inhibit HAT function (McMahon *et al*, 2005; Palhan *et al*, 2005). Furthermore, the Poly Q-containing, pathogenic form of Ataxin7 can associate with yeast SAGA, but diminishes the presence of other subunits, including Ada2, Ada3,

and Taf12 (McMahon *et al*, 2005; Baker & Grant, 2007). Moreover, Gcn5 can also enhance the stability of the DUB module in the related mammalian complex STAGA (Atanassov *et al*, 2009).

In this chapter, I will describe my observation that the DUB and HAT modules crosslink to each other in the SAGA complex, together with other biochemical experiments, revealing a functional interaction between these two modules.

3.2 The crosslinking network of DUB and HAT modules

The crosslinking data showed a surprising network of interactions between the HAT, DUB and TFIID-like core modules (Figures 2.5, 2.7D, 3.1). Within the DUB module, the C-terminal region of Sgf73 makes extensive crosslinks with the TFIID-like core, Spt20, and the HAT module subunits Ada2 and Ada3. This agrees with previous results showing that Sgf73 is responsible for linking the DUB module to SAGA (Lee *et al*, 2009; Köhler *et al*, 2008). The HAT module subunit Ada3 crosslinks to the TFIID-like core through Taf5, Taf6, Spt7, and Ada1. Together, these results demonstrate that the two enzymatic modules are positioned near each other and raise the possibility that they functionally interact.

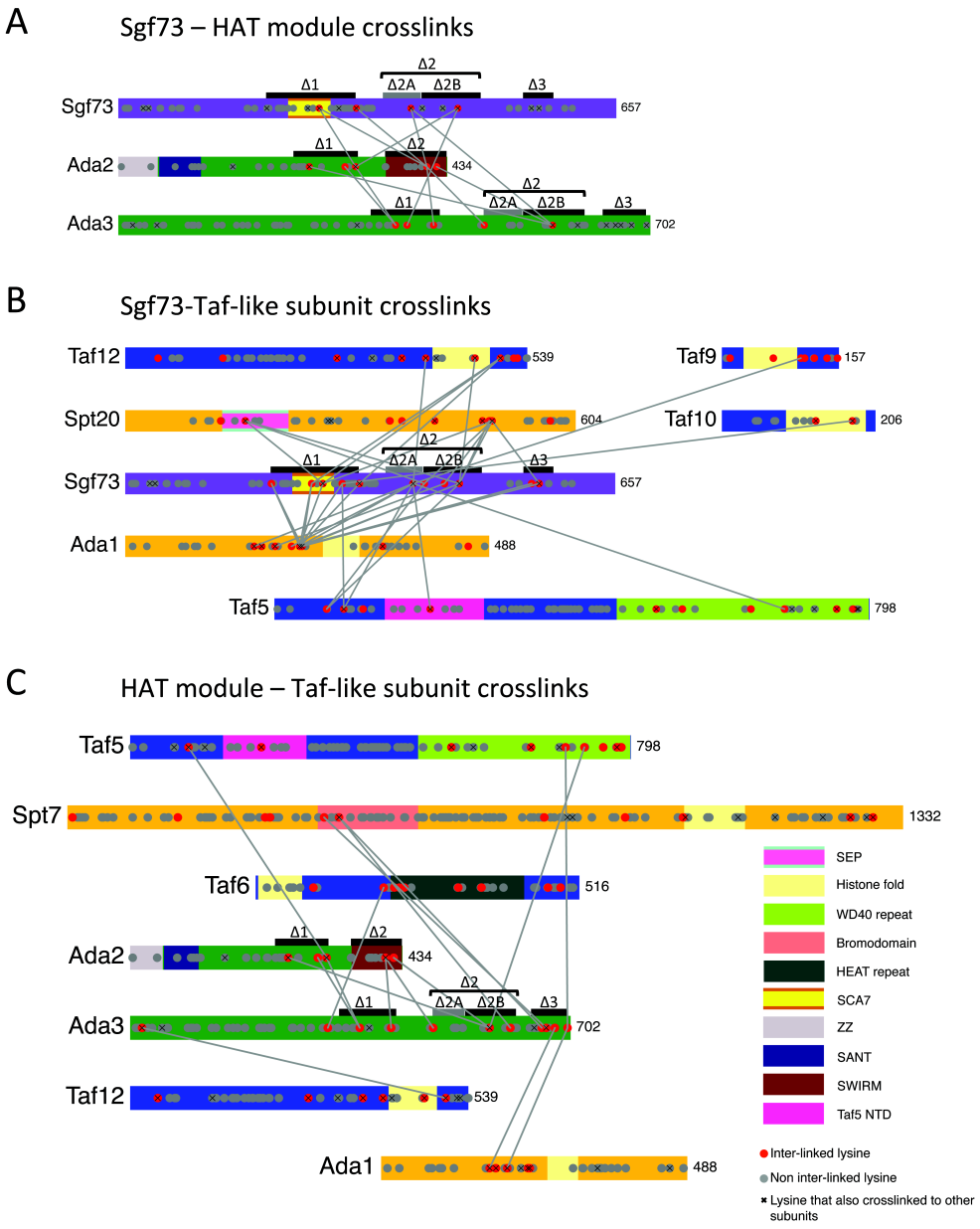

Figure 3.1 Crosslinks involving Sgf73, Ada2 and Ada3. (**A**) Intermolecular crosslinking map of Sgf73, Ada2 and Ada3. (**B**) Intermolecular crosslinking map of Sgf73 and Taf or core subunits of SAGA. (**C**) Intermolecular crosslinking map of Ada2/3 and Taf or core subunits of SAGA. Also shown in all three maps are the deletions designed in Sgf73, Ada2 and Ada3 (see following sections in this chapter).

3.3 Immunoprecipitation of Sgf73, Ada2, and Ada3 mutants

I next determined the regions of Sgf73, Ada2 and Ada3 required for the integrity of each module and their association with SAGA using the crosslinks as a guide to test specific regions of these subunits. Internal deletions were designed within subunits Sgf73, Ada2, and Ada3, and complex integrity was monitored by immunoprecipitation (IP) experiments (Figure 3.2). Within Sgf73, deletions were generated between residues 196-575 in segments that crosslink to Spt20, Ada2, Ada3, or the TFIID-like core (Figures 3.1A, B, 3.2A, B). An N-terminal Sgf73 deletion (Δ2-104) that was shown to block association of the other three DUB subunits (Köhler *et al*, 2008) was also included in the analysis. Whole cell extracts containing the Sgf73 derivatives were assayed by IP using either a 3x Flag tag at the C-terminus of Sgf73 or via the Spt7 TAP-tag (Figures 3.2A, B). The results showed that deletion of Sgf73 residues 196-314 (Δ1; encompassing the nucleosome-binding SCA7 domain) or residues at the C-terminus (401-479 and 535-575; Δ2B and Δ3) had no effect on integrity of SAGA or the DUB module (Figure 3.2B, lanes 3, 6, 7). In contrast, deletion of the Sgf73 N-terminus (Δ4) caused dissociation of Ubp8 as expected (Figure 3.2B, lane 10). Deletion of Sgf73 residues 350-400 (Δ2A) had the most severe effect, resulting in release of Sgf73 and the DUB module from SAGA, as Sgf73 failed to copurify with the TFIID-like core and the HAT modules (Figure 3.2B, lane 5). In a yeast growth assay, this latter mutant had a slow growth rate similar to a strain with a complete deletion of *SGF73* (Figure 3.3A). While this important segment overlaps with the region of Sgf73 that crosslinks to the HAT module, Spt20 and the TFIID-like core, deletion of adjacent regions in Sgf73 that also crosslink to these subunits had no effect on

DUB-SAGA association. I speculate that these adjacent regions of Sgf73 are in close proximity to other SAGA subunits but do not make direct protein-protein contacts.

Using the same strategy, I designed internal deletions surrounding the crosslinked lysine residues within HAT subunits Ada2 and Ada3 (Figures 3.2C-F). For Ada2, the IP assays of these mutant proteins agreed with the crosslinks and showed that the regions required for association with Ada3 and the TFIID-like core precisely overlap. Ada2 deletion Δ1, removing residues 232-317, retained association with Gcn5 but was released from its interactions with Ada3 and the TFIID-like core (Figure 3.2D, lane 3). This strain has a slow growth phenotype, equivalent to deletion of the *ADA2* gene (Figure 3.3B). Deletion of Ada2 residues 353-434 (Δ2) retained association with other HAT module subunits, but its binding to the TFIID-like core was diminished (Figure 3.2D, lane 4). This mutation has only a slight slow-growth phenotype (Figure 3.3B), suggesting that the HAT-SAGA interaction is not completely disrupted *in vivo*. Within Ada3, deletion of residues 334-425 (Δ1) and 483-534 (Δ2A) showed similar biochemical defects with an intact HAT module that was released from the TFIID-like core (Figure 3.2F, lanes 3, 5). In contrast, deletion of Ada3 residues 535-616 (Δ2B) disrupted the entire HAT module (Figure 3.2F, lane 6). All these mutant strains showed a slow growth phenotype (Figure 3.3C). Interestingly, Ada3 deletion Δ3 (residues 640-697) was not affected for SAGA or HAT stability (Figure 3.2F, lane 7) but still showed a slow growth phenotype equivalent to deletion of the *ADA3* gene (Figure 3.3C). This suggests that the Ada3 C-terminus may play a role in modulating Gcn5 activity, perhaps related to its interactions with the TFIID-like core and Spt7.

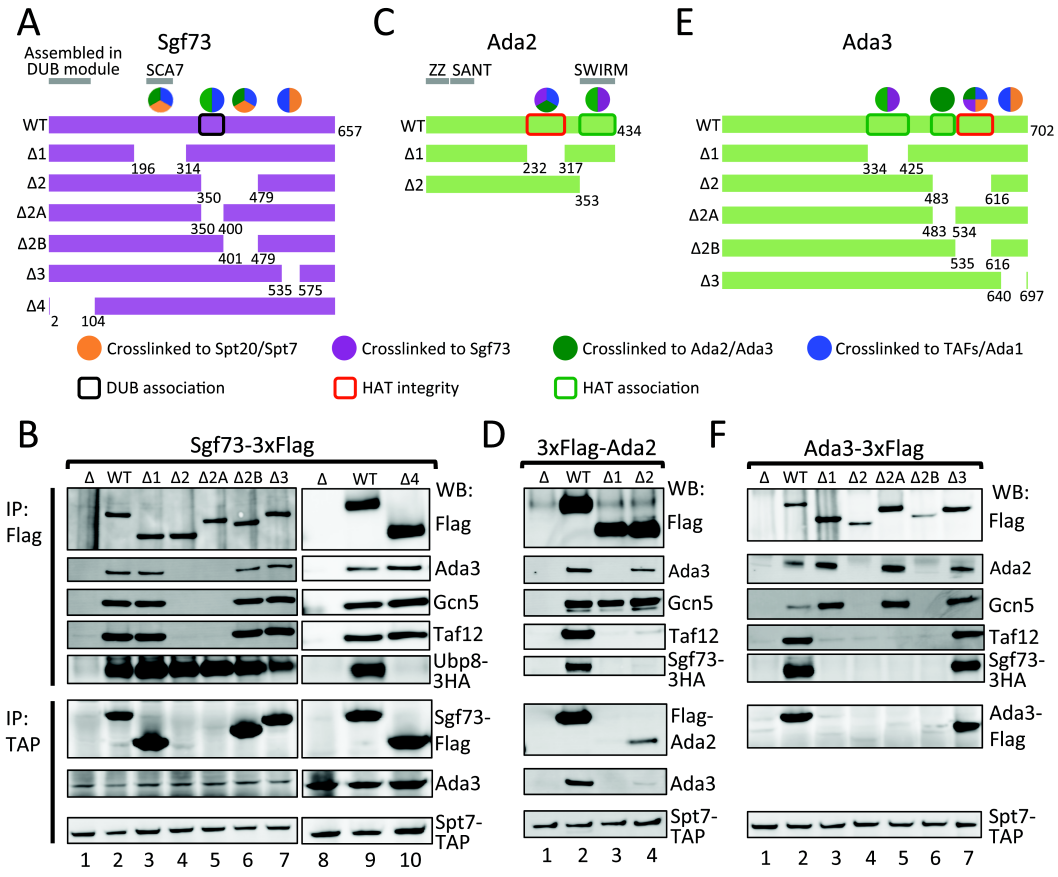

Figure 3.2 Interactions linking SAGA with the HAT and DUB modules. Deletion schematics for (**A**) Sgf73, (**C**) Ada2 and (**E**) Ada3. Numbers indicate deleted residues. Sgf73 is colored purple, and Ada2 and Ada3 are colored green. Colored circles above wild type (WT) constructs denote the crosslinking partners to corresponding regions: orange, Spt7/Spt20; purple, Sgf73; dark green, Ada2/Ada3; blue, Tafs/Ada1. Sgf73 and Ada3 constructs have a 3x Flag epitope tag at the C-terminus; Ada2 constructs have a 3x Flag epitope tag at the N-terminus. Boxes summarize the IP experiments in the lower panel: black, required for the association of the DUB module; red, required for the integrity of the HAT module; dark green, required for the association of the HAT module. IP experiments for (**B**) Sgf73 mutants, (**D**) Ada2 mutants and (**F**) Ada3 mutants are shown. Upper panels: anti-Flag IP; lower panels: anti-Spt7-TAP IP. Proteins are visualized in Western blot (WB) with corresponding antibodies labeled on the right of the blots.

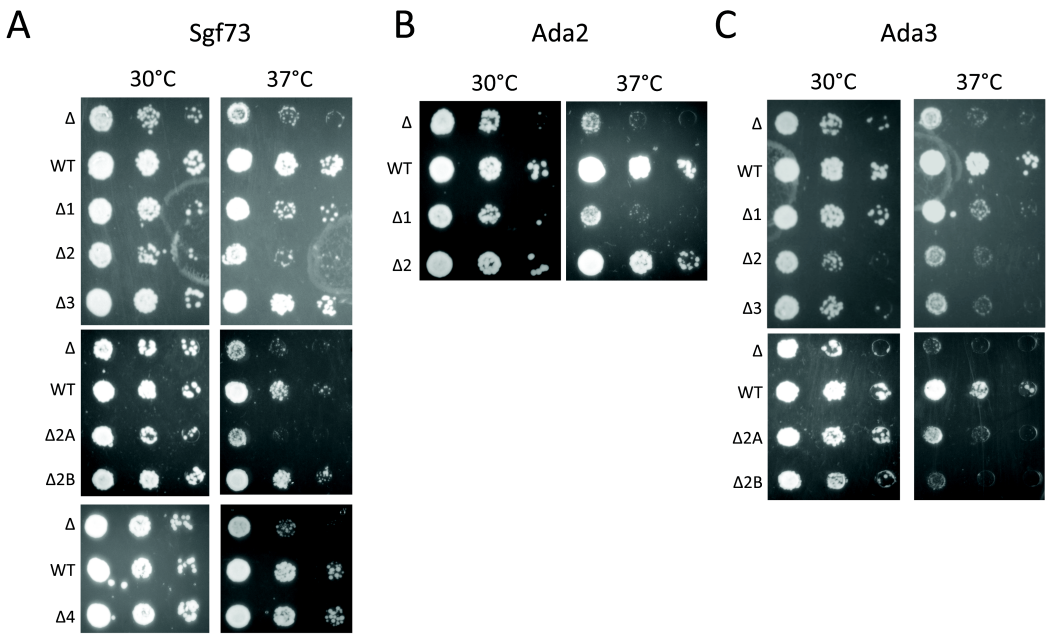

Figure 3.3 Growth phenotypes of DUB and HAT mutants. Serial dilutions of yeast strains bearing mutations within (**A**) *SGF73*, (**B**) *ADA2*, and (**C**) *ADA3* were spotted onto minimal media and incubated at 30°C and 37°C for two days. WT, wild type.

3.4 DUB stabilizes the association of the HAT module with SAGA

While the IP assays did not reveal SAGA-HAT association defects resulting from mutations within the DUB module or vice versa, the fact that these two modules are in close proximity raised the possibility of a functional interaction. To examine whether one of these enzymatic modules could affect the other, I first measured subunit association in the mutant strains using a more stringent assay. Three mutations were selected based on the above IP experiments: (i) Sgf73 Δ4, which results in specific dissociation of DUB module subunits but leaves Sgf73-SAGA association intact, (ii) Sgf73 Δ2A, which results in release of Sgf73 and its associated DUB module subunits from SAGA, and (iii) Ada3 Δ2A, which results in release of the SAGA HAT module.

SAGA was TAP-tag purified from each of these mutant strains and Figure 3.4A shows a representative silver stained gel of the preparations. The purified complexes were digested in-solution with trypsin and analyzed by mass spectrometry. The iBAQ (intensity-based absolute quantification) values (Schwanhäusser *et al*, 2011) were calculated using MaxQuant (Cox & Mann, 2008) to determine the relative abundance of each SAGA subunit in the purified complexes. The iBAQ value of a given protein is the sum of intensities of all tryptic peptides identified for that protein divided by the number of theoretical observable peptides. Because all complexes were purified through the tagged Spt7 subunit, the iBAQ value for each individual subunit was then normalized to the value of Spt7 (Figure 3.4B, Table 3.1).

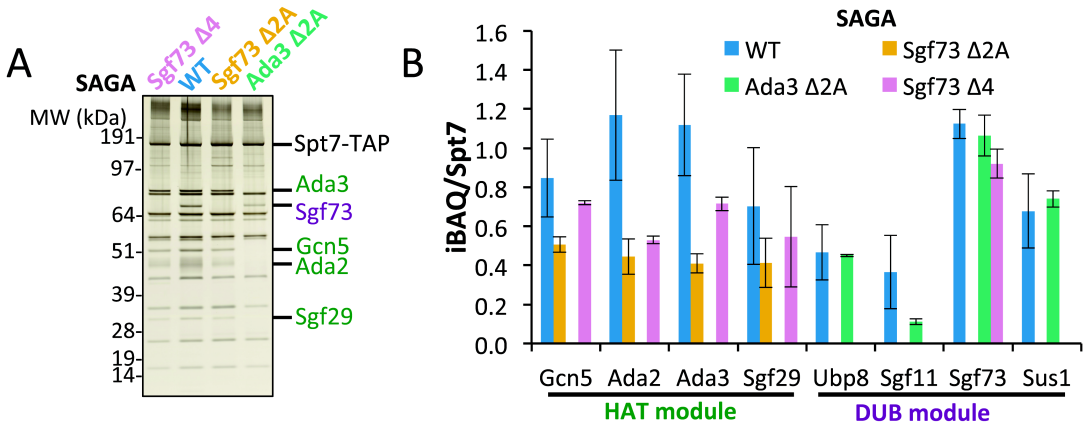

Figure 3.4 DUB stabilizes the association of HAT module with SAGA. (**A**) A silver stained protein gel showing the purification of Sgf73 Δ4, wild type (WT), Sgf73 Δ2A, and Ada3 Δ2A SAGA complexes. DUB subunits are colored purple; HAT subunits are colored green; Sgf73 Δ4 complex is colored pink; WT complex is colored blue; Sgf73 Δ2A complex is colored orange; Ada3 Δ2A complex is colored light green. (**B**) iBAQ analysis of WT, Sgf73 Δ2A, Ada3 Δ2A, and Sgf73 Δ4 SAGA complexes. Only iBAQ values normalized to Spt7 for DUB and HAT subunits are shown. See Figure 3.5 and Table 3.1 for iBAQ values for other SAGA subunits. Error bars represent standard deviation from 4 replicates for WT complex, 2 replicates for all three mutant complexes. Color scheme is the same as in panel A.

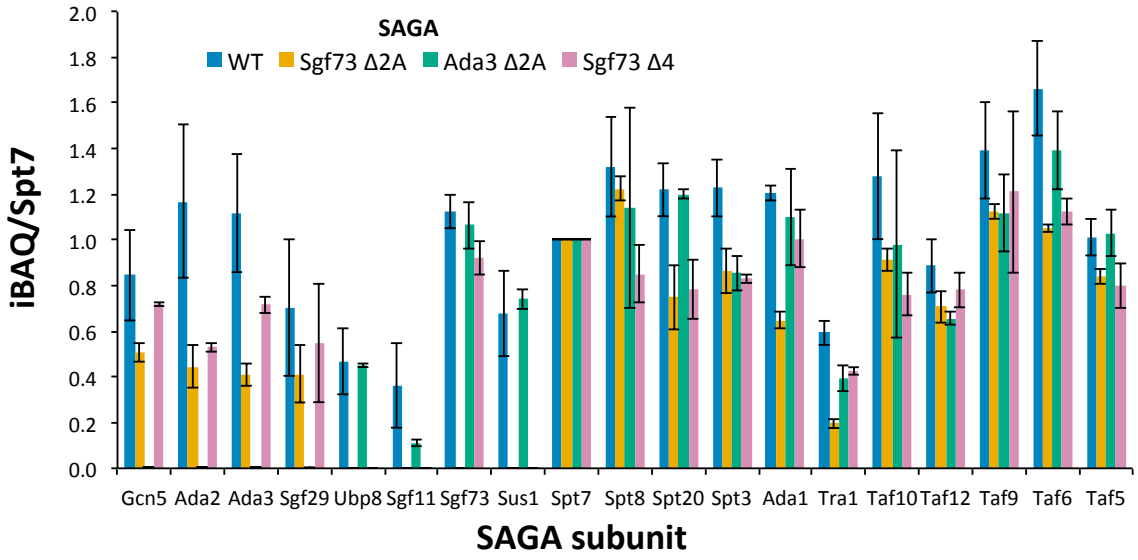

Figure 3.5 iBAQ analysis of SAGA subunits. Color coding is same as in Figure 3.4B. See also Table 3.1.

Table 3.1 List of iBAQ values for SAGA subunits.

SAGA subunits	iBAQ									
	Sgf73 Δ4 a	Sgf73 Δ4 b	WT a	WT b	WT c	WT d	Sgf73 Δ2A a	Sgf73 Δ2A b	Ada3 Δ2A a	Ada3 Δ2A b
Gcn5	203370000	114760000	443460000	403580000	226960000	410860000	151880000	81262000	73338	63005
Ada2	153200000	89191000	685210000	463620000	316480000	583770000	144830000	64448000	79984	139100
Ada3	206720000	122900000	550030000	544240000	287210000	588440000	126290000	63638000	24232	16242
Sgf29	188050000	65012000	398350000	322340000	105170000	428040000	142870000	55002000	33838	13386
Ubp8	0	0	235360000	220890000	98801000	272620000	0	0	27890000	28046000
Sgf11	0	0	122330000	262880000	67732000	192490000	0	0	7523000	6380800
Sgf73	179610000	140010000	457780000	528720000	428590000	542740000	0	0	69686000	62334000
Sus1	0	0	389590000	288680000	177840000	332870000	0	0	47205000	44718000
Spt7	267550000	170720000	423730000	427650000	388930000	503670000	284040000	169770000	61263000	62933000
Spt8	212680000	150780000	654970000	438390000	526350000	681380000	336920000	214500000	50883000	91161000
Spt20	181160000	152700000	568340000	522790000	413690000	630780000	241750000	110370000	72579000	76491000
Spt3	228660000	126070000	469390000	492100000	492210000	695690000	264560000	135080000	49056000	56984000
Ada1	252000000	190750000	515640000	523410000	448630000	616410000	191500000	105860000	76584000	59869000
Tra1	118940000	70123000	268810000	237470000	210280000	323930000	51849000	36005000	21697000	27141000
Taf10	231070000	123840000	669010000	485330000	378360000	719000000	269900000	148840000	77929000	43449000
Taf12	227460000	168740000	356020000	450740000	337630000	395390000	215100000	111780000	41479000	39914000
Taf9	268530000	209860000	520150000	727490000	496080000	689760000	326140000	187150000	75973000	62915000
Taf6	299440000	192980000	832190000	656430000	590600000	820680000	302720000	176200000	92558000	79915000
Taf5	203640000	155010000	470990000	445820000	366540000	479960000	246100000	138390000	58762000	69340000
Chd1	0	0	0	0	9693.3	0	9206.8	6482.9	0	0

Table 3.2 List of adjNSAF values of SAGA subunits.

SAGA subunits	adjNSAF									
	Sgf73 Δ4 a	Sgf73 Δ4 b	WT a	WT b	WT c	WT d	Sgf73 Δ2A a	Sgf73 Δ2A b	Ada3 Δ2A a	Ada3 Δ2A b
Gcn5	0.02589699	0.01735487	0.02577707	0.02352921	0.01408449	0.02140644	0.02329172	0.01916568	0.00049374	0.00093988
Ada2	0.02684144	0.01788247	0.03063988	0.03030305	0.01858707	0.03398785	0.02891386	0.01671689	0.00033294	0.00047534
Ada3	0.03171692	0.02070928	0.03217602	0.03304939	0.0186637	0.03231774	0.03188945	0.01921033	0.0011022	0.00052454
Sgf29	0.0243476	0.01235178	0.02365979	0.0182888	0.00763474	0.01954024	0.01814067	0.01269165	0.00027852	0.00053019
Ubp8	0	0	0.02837378	0.02563218	0.01425793	0.02479277	0	0	0.03267869	0.02380224
Sgf11	0	0	0.01858279	0.01745539	0.01355632	0.02169186	0	0	0.02172443	0.0144741
Sgf73	0.02471318	0.0189023	0.02239829	0.0240581	0.01604037	0.02368372	0	0	0.02509204	0.01843567
Sus1	0	0	0.04029684	0.04157522	0.01996513	0.05884933	0	0	0.08659909	0.06608203
Spt7	0.04078837	0.02896392	0.035909	0.03531085	0.02625983	0.03580052	0.05093426	0.0345652	0.05204302	0.04296779
Spt8	0.03627689	0.02662904	0.0309235	0.03094396	0.02770037	0.03502657	0.05166282	0.04185338	0.04815642	0.04252044
Spt20	0.04813179	0.03786517	0.05094512	0.04218344	0.02992954	0.04510096	0.0493746	0.03514971	0.06331803	0.04534191
Spt3	0.02913388	0.02153643	0.03583092	0.03668445	0.02621309	0.0393507	0.04479177	0.02946923	0.05013329	0.0352778
Ada1	0.03207621	0.02641557	0.03459457	0.02855687	0.02396019	0.03046807	0.0302459	0.02161897	0.03568907	0.02678039
Tra1	0.02727105	0.02107626	0.02727404	0.02312813	0.01742697	0.02508946	0.01788083	0.0163009	0.02496327	0.02186436
Taf10	0.02990194	0.02275432	0.03714701	0.03053166	0.02385688	0.03391939	0.03234961	0.02833993	0.04512804	0.03362971
Taf12	0.0293073	0.02563883	0.02717407	0.03143307	0.02053173	0.02981045	0.02922248	0.02184044	0.03325715	0.02476168
Taf9	0.0378398	0.0257459	0.03325597	0.03733365	0.02144988	0.03179353	0.04238209	0.02513935	0.05637352	0.02879115
Taf6	0.03401234	0.02581317	0.02850691	0.02980444	0.0195722	0.02771382	0.0574409	0.03155866	0.05406595	0.04439416
Taf5	0.02500107	0.02134722	0.02518232	0.0240312	0.01696661	0.02200479	0.04000987	0.02271473	0.03752154	0.02656905
Chd1	0	0	0	0	0.00003296	0.00003886	0	0.0001248	0	0

For Sgf73 Δ4, the MS analysis confirmed that SAGA lacked the DUB subunits Ubp8, Sgf11, and Sus1, while retaining wild type levels of Sgf73 (Figure 3.4B). As expected from the IP results, the more severe mutation Sgf73 Δ2A caused dissociation of all DUB associated subunits including Sgf73 (Figure 3.4B). In contrast, mutation of the HAT module subunit Ada3 (Δ2A) had no effect on the stability of the DUB module apart from a surprising 70% decrease in association of Sgf11 (Figure 3.4B). Finally, I found that both Sgf73 mutations affected the stability of the HAT module, consistent with crosslinks observed between Sgf73 and HAT subunits. SAGA purified from the mutants Sgf73 Δ2A and Δ4 showed a consistent 15-63% decrease in the amount of the associated HAT subunits (Figure 3.4B). I obtained similar results by analyzing the same data set and calculating the adjusted normalized spectral abundance factor (adjNSAF) using Abacus (Fermin *et al*, 2011) (Table 3.2). Together, the analysis of subunit composition in the HAT and DUB mutants suggest that the DUB module plays a modest role in stabilizing association of the HAT module with SAGA.

3.5 Mutations in the DUB module subunit Sgf73 affect SAGA nucleosomal HAT activity

Since the above analysis showed an interaction between the DUB and HAT modules, I tested whether mutations in Sfg73 affect HAT function using both recombinant histone H3 and mono-nucleosome substrates. The amounts of purified wild type and mutant SAGA added to the HAT assays were normalized based on the amounts of added Gcn5. Histone H3 acetylation was monitored by Western blot assay using an antibody against acetylated H3K9. Unexpectedly, we observed that the two Sgf73 mutations have opposite effects on HAT activity.

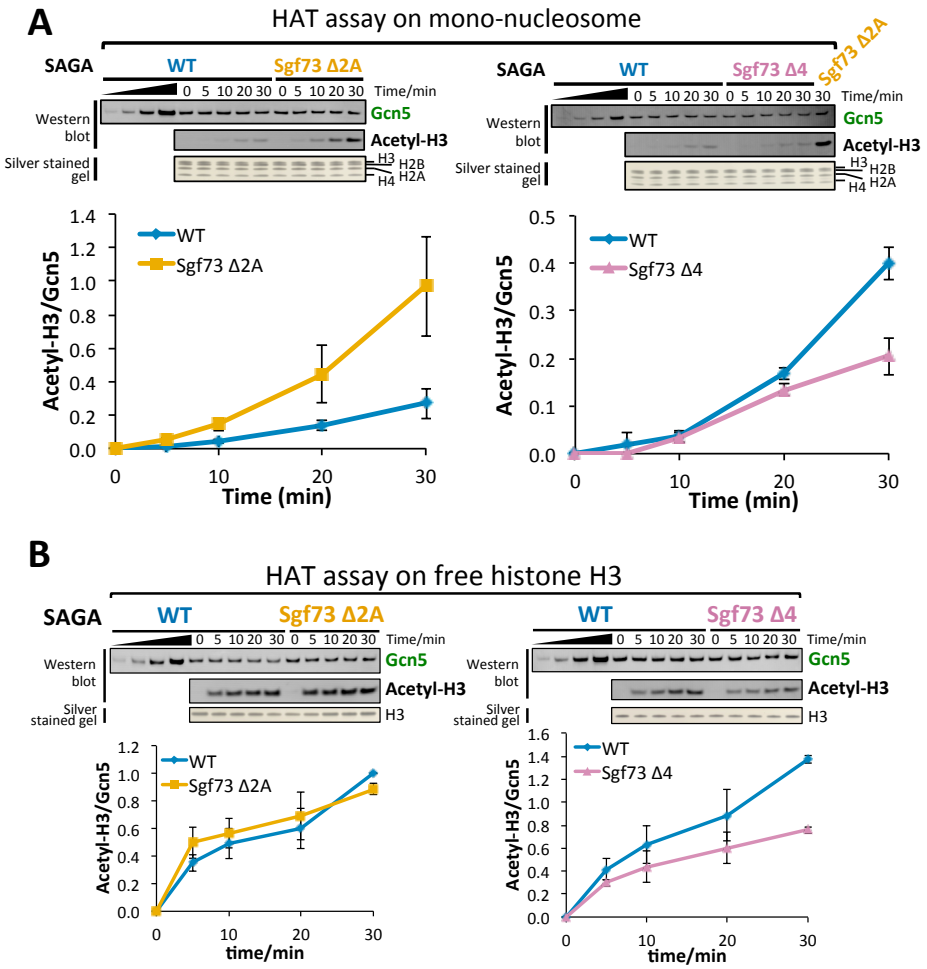

Figure 3.6 DUB module affects the nucleosomal HAT activity. (**A**) WT, Sgf73 Δ2A, and Sgf73 Δ4 SAGA complexes were analyzed in an *in vitro* HAT assay using mono-nucleosome as substrate. Equivalent amounts of Gcn5 and all 4 histones were present at each time point as visualized by Western blot and silver staining (upper panel). Reactions were monitored by Western blot using antibody against acetylated H3K9 (Acetyl-H3). The acetylation signal was normalized to the amount of Gcn5 in the reaction at each time point (lower panel). Error bars represent standard deviation from 4 experiments for the left panel (two biological replicates with two technical repeats), and 2 biological replicates for the right panel. Color scheme is the same as in Figures 3.4, 3.5. Sample from Sgf73 Δ2A complex at 30 min time point was also loaded as a control on the right panel. **(B)** *in vitro* HAT assays. Same as in A, but with free histone H3 as substrate. Color scheme is the same as in A. Error bars represent standard deviations from two biological replicates.

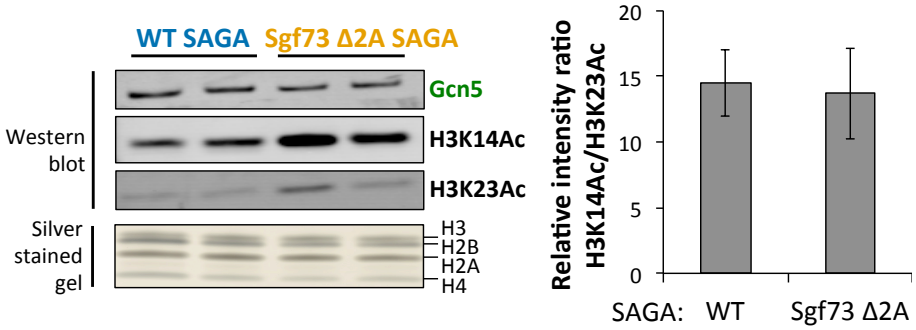

Figure 3.7 The DUB module does not affect the lysine specificity of Gcn5. An *in vitro* HAT assay using mono-nucleosome as substrate was performed at 30°C for 30 min with both wild type (WT) and Sgf73 Δ2A SAGA complexes. Reactions were analyzed by SDS-PAGE gel and Western blot or silver staining. Equivalent amounts of Gcn5 and all 4 histones were present in each sample, as shown in the top and bottom panels. Histone acetylation was monitored using antibodies against acetylated H3K14 (H3K14Ac) or H3K23 (H3K23Ac). Quantification of the relative intensity ratio is shown as bar graph on the right. Error bars in the histogram on the left represent standard deviation from the two independent experiments for both complexes.

SAGA purified from the more severe Sgf73 Δ2A mutation had HAT activity equivalent to wild type SAGA using the histone H3 substrate, but showed ~4-fold higher acetylation of the mono-nucleosome substrate after 30 min of incubation (Figure 3.6, left panels). In contrast, SAGA purified from the less severe Sgf73 Δ4 mutation shows about 2-fold lower activity compared to wild type SAGA on both the H3 and mono-nucleosome substrates (Figure 3.6, right panels). Since this latter mutation only affects stability of the DUB module, it suggests that the DUB modestly enhances enzymatic activity of the HAT module. Because Sgf73 makes numerous protein-protein interactions with other SAGA modules, complete dissociation of Sgf73 in the Δ2A mutant may indirectly affect SAGA HAT activity due to changes in other SAGA-HAT interactions. I also tested whether the DUB module affected the specificity of H3 acetylation. The products of nucleosome acetylation were probed using antibodies against acetylated H3K14 and H3K23 (Figure 3.7). While the Sgf73 Δ2A SAGA derivative showed higher levels of acetylation for both H3K14 and H3K23 compared to the wild type complex, there was no change in the relative ratio of these two acetylated residues. The combined analysis of SAGA indicates a close proximity between the DUB and HAT modules and that the DUB module modestly enhances the enzymatic activity of the HAT module.

3.6 Materials and methods

3.6.1 Yeast strains, plasmids, and spot test assay.

Yeast strains and plasmids used are listed in Tables 3.3 and 3.4. All strains are derivatives of BY4705 (Brachmann *et al*, 1998). Plasmid pYH1 is based on the pRS426 vector; all remaining plasmids are based on pRS316 vector (Brachmann *et al*, 1998), both

containing the *URA3* selectable marker. All internal deletions, except that in pYH59, contain a GSGSGS linker at the internal deletion junction. To examine growth phenotype, serial dilutions were spotted on glucose minimal medium lacking uracil and grown for 2 days at 30°C and 37°C (Figure 3.3).

3.6.2 TAP purifications.

Same as described in **2.7.1**.

3.6.3 MS sample preparation and data analysis.

Sample preparation is the same as described in **2.7.4**. For iBAQ analysis, MS raw files were loaded to MaxQuant (Cox & Mann, 2008) (version 1.3.0.5), and searched against the yeast whole genome database with deletions in Sgf73 or Ada3 under default settings with the following exceptions: Group-specific parameters tab, Multiplicity=1, Enzyme=trypsin; Identification & quantification tab, Min. ration count=1; Misc. tab, check Re-quantify and Keep low-scoring versions of identified peptides Only within parameter groups, check Match between runs, check Label-free quantification, LFQ min. ratio count=1, uncheck Fast LFQ, check iBAQ and Log fit, check Calculate peak properties. Alternatively, MS data were first searched against the yeast whole genome database using X!Tandem on the computer cluster at the Institute for Systems Biology (ISB, Seattle, Washington, USA). The search results were then analyzed using PeptideProphet and ProteinProphet in TPP on the computer cluster at ISB. Adjusted NSAF (adjNSAF) values were calculated using Abacus (Fermin *et al*, 2011) with default parameters.

Table 3.3 Yeast strains used.

Strain	Plasmid	Strain background
SHY627	none	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1*
YHY8	pYH1	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δspt3::HPH*
YHY28	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY29	pYH28	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY30	pYH31	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY31	pYH32	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY32	pYH33	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY33	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada2::KanMX*
YHY37	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY38	pYH30	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY39	pYH36	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY40	pYH37	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY41	pYH38	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY52	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada2::KanMX*
YHY56	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY57	pYH30	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY58	pYH36	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY59	pYH37	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY60	pYH38	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY62	pYH39	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY63	pYH40	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY66	pYH41	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY67	pYH42	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada3::KanMX*
YHY68	pYH41	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY69	pYH42	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada3::KanMX*
YHY71	pRS316	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY72	pYH28	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3*

		Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX
YHY73	pYH31	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY74	pYH32	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY75	pYH39	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY76	pYH40	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY77	pYH33	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*
YHY78	pYH57	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada2::KanMX*
YHY79	pYH58	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada2::KanMX*
YHY80	pYH59	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Sgf73-3HA::His3MX6, Δada2::KanMX*
YHY81	pYH57	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada2::KanMX*
YHY82	pYH58	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada2::KanMX*
YHY83	pYH59	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δada2::KanMX*
YHY85	pYH67	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, SPT7-(Flag1)-TAP::TRP1, Δsgf73::KanMX*
YHY86	pYH67	*matα, Δade2::hisG, his3 Δ200, leu2 Δ0, lys2 Δ0, met15 Δ0, trp1 Δ63, ura3 Δ0, Ubp8-3HA::His3MX6, Δsgf73::KanMX*

Table 3.4 Plasmids used.

Plasmid name	Insert
pYH1	Spt3 ORF with E240K (401), 608bp promoter sequence, 726bp of sequence past stop codon
pYH28	Sgf73 ORF with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH30	Ada3 ORF with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH31	Sgf73 ORF Δ residues 196-314 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH32	Sgf73 ORF Δ residues 350-479 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH33	Sgf73 ORF Δ residues 535-575 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH36	Ada3 ORF Δ residues 334-425 with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH37	Ada3 ORF Δ residues 483-616 with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH38	Ada3 ORF Δ residues 640-697 with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH39	Sgf73 ORF Δ residues 350-400 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH40	Sgf73 ORF Δ residues 401-479 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon
pYH41	Ada3 ORF Δ residues 483-534 with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH42	Ada3 ORF Δ residues 535-616 with 3xFlag on C-terminus, 500bp promoter sequence, 106bp of sequence past stop codon
pYH57	Ada2 ORF with 3xFlag on N-terminus, 491bp promoter sequence, 100bp of sequence past stop codon
pYH58	Ada2 ORF Δ residues 232-317 with 3xFlag on N-terminus, 491bp promoter sequence, 100bp of sequence past stop codon
pYH59	Ada2 ORF Δ residues 353-434 with 3xFlag on N-terminus, 491bp promoter sequence, 100bp of sequence past stop codon
pYH67	Sgf73 ORF Δ residues 2-104 with 3xFlag on C-terminus, 500bp promoter sequence, 100bp of sequence past stop codon

3.6.4 IP experiments.

For IP assays, small scale whole cell extracts were prepared by bead beating as described (Knutson & Hahn, 2011), except that TAP extraction buffer was used for TAP IP experiments. Protein concentrations of extracts were determined by Bradford assays. For IP experiments, 2-4 mg of whole cell extracts were incubated with 25 µl IgG Sepharose beads (for TAP IP), or with 25 µl Anti-FLAG M2 Affinity Gel (Sigma; for Flag IP) at 4°C overnight with gentle rocking. For TAP IPs, bound proteins were eluted by incubating beads with 20 µl TEV cleavage buffer containing 0.3 µg TEV protease at room temperature for 2.5 hr. The elution was repeated but for 1.5 hr, and eluates were combined. For Flag IPs, bound proteins were eluted by incubating beads with 25 µl 0.5 mg/ml 3×Flag peptide (Sigma) for 25 min at room temperature. The elution was repeated with 15µl 0.5mg/ml 3×Flag peptide for 15 min, and eluates were combined. Alternatively, bound proteins were eluted in 40 µl 1× NuPAGE LDS sample buffer (Invitrogen) at 95°C for 5 min. Eluted proteins were run on an SDS-PAGE and analyzed by Western blot using the following antibodies: M2-Flag antibody (Sigma), Ada2 antibody (yC-20; Santa Cruz, sc-6651), Ada3 antibody (yN-19; Santa Cruz, sc-6652), Gcn5 antibody (yV-19; Santa Cruz, sc-6305), HA-probe antibody (F-7; Santa Cruz, sc-7392X), and rabbit polyclonal antibody against Taf12. Spt7 was visualized by the Protein A epitope at the C-terminus of the TAP-tag in heat eluted samples.

3.6.5 HAT assays.

Histone acetyltransferase (HAT) assays were performed as previously described (Knutson & Hahn, 2011), but with modifications. Briefly, 60-100 ng of wild type or Sgf73 mutant SAGA complexes normalized by Gcn5 content were incubated with 2 µg

recombinant histone H3 or mono-nucleosome (a kind gift from Steven Henikoff at FHCRC) in 100 µl HAT buffer (50 mM Tris pH8, 50 mM KCl, 0.2 mM EDTA, 20 mM sodium butyrate, 5% glycerol, 2 mM DTT, 2 mM PMSF, 0.1% NP-40) at 30°C. After removing a sample at the zero time point, acetyl-CoA (Sigma) was added to a final concentration of 10 µM in each reaction. Equal volume of samples were taken out at 5 min, 10 min, 20 min, and 30 min time points, and were mixed with SDS sample buffer to stop acetylation. Samples were split into two parts, and were analyzed by silver stained SDS-PAGE and by Western blot with Gcn5 antibody, H3K9Ac antibody (Millipore, 07-352), H3K14Ac antibody, and H3K23Ac antibody. The latter two H3 acetylation antibodies were gifts from Toshio Tsukiyama at FHCRC. Gcn5 and acetylated H3 were quantified using the Odyssey infrared imaging system (LiCor Biosciences).

Chapter 4 Interaction between SAGA and TBP

4.1 Introduction

The two TBP binding subunits within SAGA, Spt3 and Spt8, play complex roles in gene regulation and, depending on the regulated gene, can have either positive or negative effects on transcription (Bhaumik, 2011). For example, deletion of Spt8 causes an increase in basal transcription at some genes (Belotserkovskaya *et al*, 2000), while at others Spt3 and Spt8 act to promote transcription initiation (Sterner *et al*, 1999).

While both biochemical and genetic analyses have implicated Spt3 and Spt8 in TBP binding, Spt8 is the only subunit known to bind to TBP in the absence of other SAGA components (Sermwittayawong & Tan, 2006; Warfield *et al*, 2004). In pull-down experiments, Spt8 was shown to compete with DNA (Sermwittayawong & Tan, 2006) and TFIIA (Warfield *et al*, 2004) for binding to TBP. These competitions might explain the inhibitory effects Spt8 has on the basal transcription of some specific genes (Belotserkovskaya *et al*, 2000), but the mechanism of the competition and whether it plays different roles at other promoters remain unclear. Spt3 contains both halves of an interacting histone fold pair, located at the N and C-terminal ends of the protein (Birck *et al*, 1998). The *spt3-401* allele (E240K), a suppressor of a TBP mutation in the Spt3-binding surface, lies within the C-terminal histone fold and implicates this region in TBP binding (Laprade *et al*, 2007). However, the fact that Spt3 alone does not bind TBP raises the possibility that other SAGA subunits might facilitate this interaction. Interestingly, previous biochemical and mutational analyses also suggest that deletion of Sgf29, one of

the HAT module subunits, affects TBP recruitment to the core promoter of the GAL1 gene (Shukla *et al*, 2012).

In this chapter, I will focus on the biochemical and structural experiments I did based on the crosslinking results described in Chapter 2. The results show a complex interaction network between several SAGA subunits and TBP and shed light on the mechanism of the competition between TFIIA and Spt8 for binding to TBP.

4.2 A network of interactions between TBP and four SAGA subunits

Crosslinking of the SAGA-TBP complex revealed four subunits positioned close to TBP: Spt3, 8, 7 and 20 (Figures 2.5, 4.1A). While the crosslinks between Spt3, Spt8 and TBP are consistent with previous findings, the additional crosslinks of TBP with Spt20 and Spt7 suggest that these subunits may function to facilitate TBP binding. The lysine residues of TBP that crosslink to Spt3 and Spt8 are located on opposite ends of the TBP surface (Figure 4.1B). Previous studies using the photocrosslinker BPA (ρ-benzoyl-phenylalanine) positioned on TBP identified two TBP surfaces that are close to Spt3 and Spt8 in the preinitiation complex (Mohibullah & Hahn, 2008). Consistent with this, I found that the TBP lysine residues that crosslink to Spt3 or Spt8 are close to the previously identified BPA crosslinks, further supporting the existence of two separate SAGA binding interfaces on TBP. The region of TBP that crosslinks to Spt3 also agrees with the location of mutations in TBP that suppress Spt3 mutations (Figure 4.1B) (Laprade *et al*, 2007; Eisenmann *et al*, 1992).

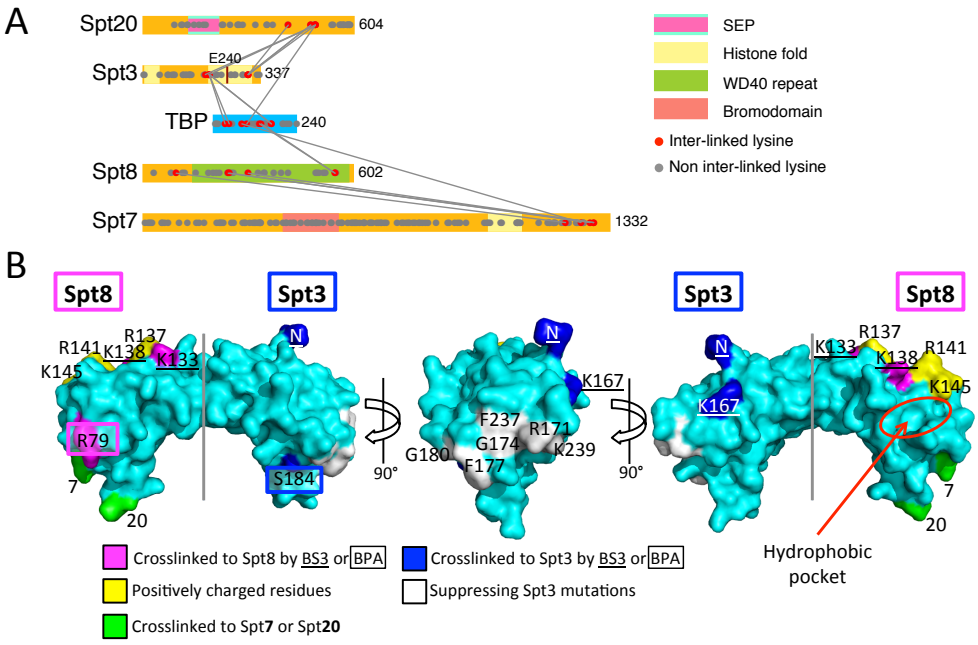

Figure 4.1 Two distinct surface regions on TBP that interact with Spt3 and Spt8. (**A**) A mini-crosslinking map involving TBP and SAGA subunits. (**B**) Surface representation of TBP (PDB ID 1TBP). TBP residues that crosslink to Spt8 and Spt3 are colored magenta and blue, respectively, with BS3 crosslinked residues underlined and BPA crosslinked residues boxed. Residues crosslinked to Spt7 (7) and Spt20 (20) are colored green. A conserved positively charged groove on the TBP surface is colored yellow. A hydrophobic pocket on TBP is shown as a red circle. Residues on the TBP surface that have been shown to suppress Spt3 mutations are colored white (Laprade *et al*, 2007; Eisenmann *et al*, 1992).

Spt3 lysine residue 190, located at one end of the C-terminal histone fold, which also harbors the Spt3-401 mutation, E240K (Figure 4.1A), crosslinks to the C-terminal half of TBP. Spt3 K190 also crosslinks to two residues within the disordered and non-conserved TBP N-terminal region (Figure 4.1A). This latter interaction is probably not functionally important, as the N-terminal region of TBP can be deleted with no obvious phenotype (Reddy & Hahn, 1991). Interestingly, the C-terminal region of Spt20 crosslinks to both ends of the Spt3 C-terminal histone fold and to the opposite side of TBP near the Spt8-interacting surface (Figures 4.1A, B). Since Spt3 alone cannot stably bind TBP, these findings suggest that Spt20 may play a complex role in SAGA-TBP binding by modulating the Spt3-TBP interaction and possibly by direct interactions with TBP.

The C-terminal region of Spt7 crosslinks to both Spt8 and TBP (Figure 4.1A), suggesting that this region of Spt7 may be directly involved in TBP binding. Although recombinant Spt8 alone can bind TBP (Figure 4.2), Spt7 may modulate this interaction. Interestingly, both Spt7 and Spt20 crosslink to nearby positions on the TBP N-terminal stirrup where they could potentially interact (Figure 4.1B). Thus, the crosslinking analysis suggests that Spt3, Spt20, Spt8 and Spt7 have a complex network of interactions with TBP.

4.3 Spt8 occupies a positively charged groove on the surface of TBP

TBP residues K133 and K138 crosslink to the Spt8 WD40 repeats and are located in a conserved positively charged groove on the convex surface of TBP (Figure 4.1B). These two lysine residues, together with K145 on TBP helix 2, are essential for interactions within Pol I, Pol II, and Pol III preinitiation complexes. For example, Mot1, the Taf1 TAND2 domain and TFIIA all interact with this positively charged groove (Buratowski

& Zhou, 1992; Kim & Roeder, 1994; Wollmann *et al*, 2011; Anandapadamanaban *et al*, 2013). Both the Taf1 TAND2 domain and Spt8 compete with TFIIA for binding TBP (Kokubo *et al*, 1998; Warfield *et al*, 2004; Bagby *et al*, 2000), suggesting that Spt8 might also bind TBP via the positively charged groove. To test this model, I first examined the sequence of Spt8, which has an acidic N-terminal region followed by a predicted WD40 repeat domain (Figure 4.2A). Sequence alignment shows a good match between residues 383-398 of Spt8 and residues 55-66 of Taf1 TAND2 domain (Figure 4.2B). This region in Taf1 TAND2 domain was shown to bind the positively charged groove on the surface of TBP, with the phenylalanine residue interacting with the hydrophobic pocket (Figure 4.1B) (Anandapadamanaban *et al*, 2013). Interestingly, Spt8 residues 383-398 also show such features, with an aromatic residue followed by a series of acidic residues (Figure 4.2B), suggesting Spt8 might also use this motif to bind to the positively charged groove of TBP. It is also worth noting here that the motif might be reversed in the primary sequence, as was shown for Mot1 (Anandapadamanaban *et al*, 2013). Taking this into consideration, I tested this hypothesis by making deletion mutations of potential motifs within Spt8 (Figure 4.2A). The mutant Spt8 constructs were purified form *E. coli* as GST fusion proteins, and TBP-binding was tested in a GST pull-down experiment (Figure 4.2C). Surprisingly, none of the Spt8 mutants show decreased TBP binding (Figure 4.2C), suggesting that Spt8 might not interact with TBP in the same way as Taf1 TAND2 domain does. Furthermore, Spt8 also competes with DNA for binding to TBP (Sermwittayawong & Tan, 2006), suggesting the major interaction interface between Spt8 and TBP may involve the DNA binding surface of TBP.

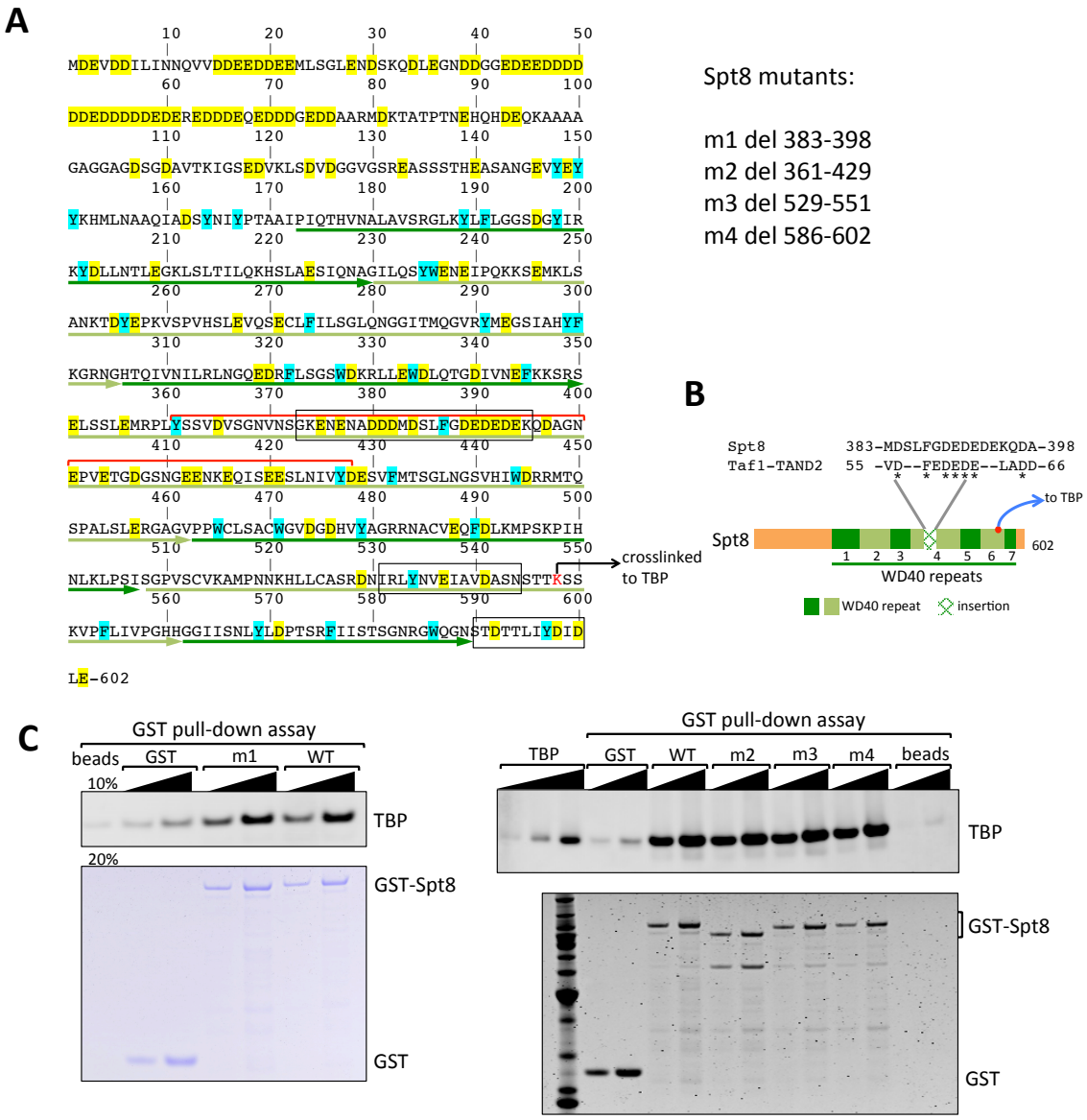

Figure 4.2 Spt8 binds TBP in a GST pull-down assay. (**A**) Sequence of Spt8. Acidic residues are shaded yellow, and aromatic residues are shaded cyan. The lysine residue K548 that crosslinks to TBP is colored red. The 7 WD40 repeats are indicated by green or light green arrows underneath the sequence. The insertion within WD40 repeat 4 is shown by a red bracket above the sequence. The sequences that match the sequence pattern in Taf1 TAND2 domain are shown in black boxes. Spt8 deletion designs are shown to the right of the sequence. (**B**) Cartoon representation of Spt8 with the sequence alignment between Spt8 and Taf1 TAND2 domain shown. (**C**) GST pull-down experiments between TBP and various Spt8 constructs. Upper panels are Western blot (WB) using an antibody against TBP; lower panels are Coomassie blue stained (CBS) gels. For WB, 5% and 10% eluates were loaded for each sample; for CBS, 10% and 20% eluates were loaded.

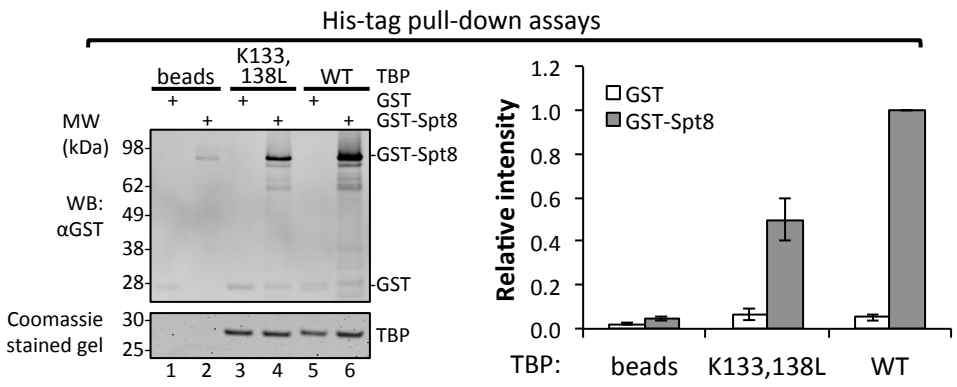

Figure 4.3 Spt8 physically contacts the positively charged groove on TBP. Representative Western blot and Coomassie stained gel of 6His-TBP pull-down experiments are shown on the left. 10% eluates were loaded each lane. Relative band intensity for GST or GST-Spt8 is shown as bar graph on the right. Intensity of GST-Spt8 bound with WT TBP (lane 6 on the left) was set to 1. Error bars represent standard deviation from three independent experiments.

To further test whether or not Spt8 interacts with the positively groove of TBP, I introduced a double leucine substitution to TBP (K133, 138L) that is defective in forming the TBP-Mot1-DNA complex (Auble & Hahn, 1993), and in interacting with TFIIA (Buratowski & Zhou, 1992). These two lysine residues both crosslink to Spt8 with BS3 (Figure 4.1). To test the effect of these TBP mutations on Spt8 binding, I performed another pull-down assay where His-tagged wild type or K133,138L TBP were immobilized on Ni Sepharose resin and then incubated with GST or GST-tagged Spt8 (Figure 4.3). The results showed that the K133,138L mutation reduced TBP-Spt8 binding by 2-fold (Figure 4.3), suggesting that the positively charged groove of TBP is indeed one binding interface between Spt8 and TBP. Taken together, both crosslinking and biochemical analyses suggest that Spt8 physically contacts the positively charged groove on TBP and that this interaction contributes to TBP-Spt8 binding.

4.4 Attempts to crystallize Spt8 and TBP

Although the aforementioned biochemical experiments reveal that Spt8 interacts with TBP partially through binding to the positively charged groove on TBP, many questions remain unanswered. First, we don't know the structure of Spt8, let alone the regions in Spt8 that are responsible for contacting the positively charged groove on TBP. The mutagenesis experiments I performed (Figure 4.2) failed to identify such regions. Furthermore, replacement of the key lysine residues within the groove on TBP only partially affected Spt8 binding (Figure 4.3), suggesting other regions on TBP are also contacted by Spt8 when the two bind to each other. The DNA binding surface of TBP might be one candidate as suggested by the competition between Spt8 and DNA for binding to TBP (Sermwittayawong & Tan, 2006). However, the exact mechanism

requires further investigation. To answer these questions, I attempted to solve the crystal structure of Spt8 in complex with TBP.

Both previous studies (Warfield *et al*, 2004; Sermwittayawong & Tan, 2006) and my experiments (Figures 4.2, 4.3) show Spt8 alone can bind to TBP *in vitro*. To study the structure of Spt8/TBP complex, I first tried to copurify the two proteins. Based on sequence alignment and secondary structure prediction (Figure 4.4), one truncated Spt8 construct (145-602), namely Spt8c01, was chosen together with the full length protein. As for TBP, the N-terminal 60 residues were removed as no phenotype was observed when deleted (Reddy & Hahn, 1991). The remaining TBP core domain displays more than 80% sequence identity between yeast and human, and has been successfully crystallized previously either alone (Nikolov *et al*, 1992; Chasman *et al*, 1993), in complex with DNA (Kim *et al*, 1993b; 1993a), or in a tertiary complex with DNA and TFIIA (Tan *et al*, 1996; Geiger *et al*, 1996) or TFIIB (Nikolov *et al*, 1995). Although many attempts were made, I failed to copurify either full length Spt8 with TBP core, or Spt8c01 with TBP core, using either the *E.coli* expression system or the baculovirus expression system in insect cells.

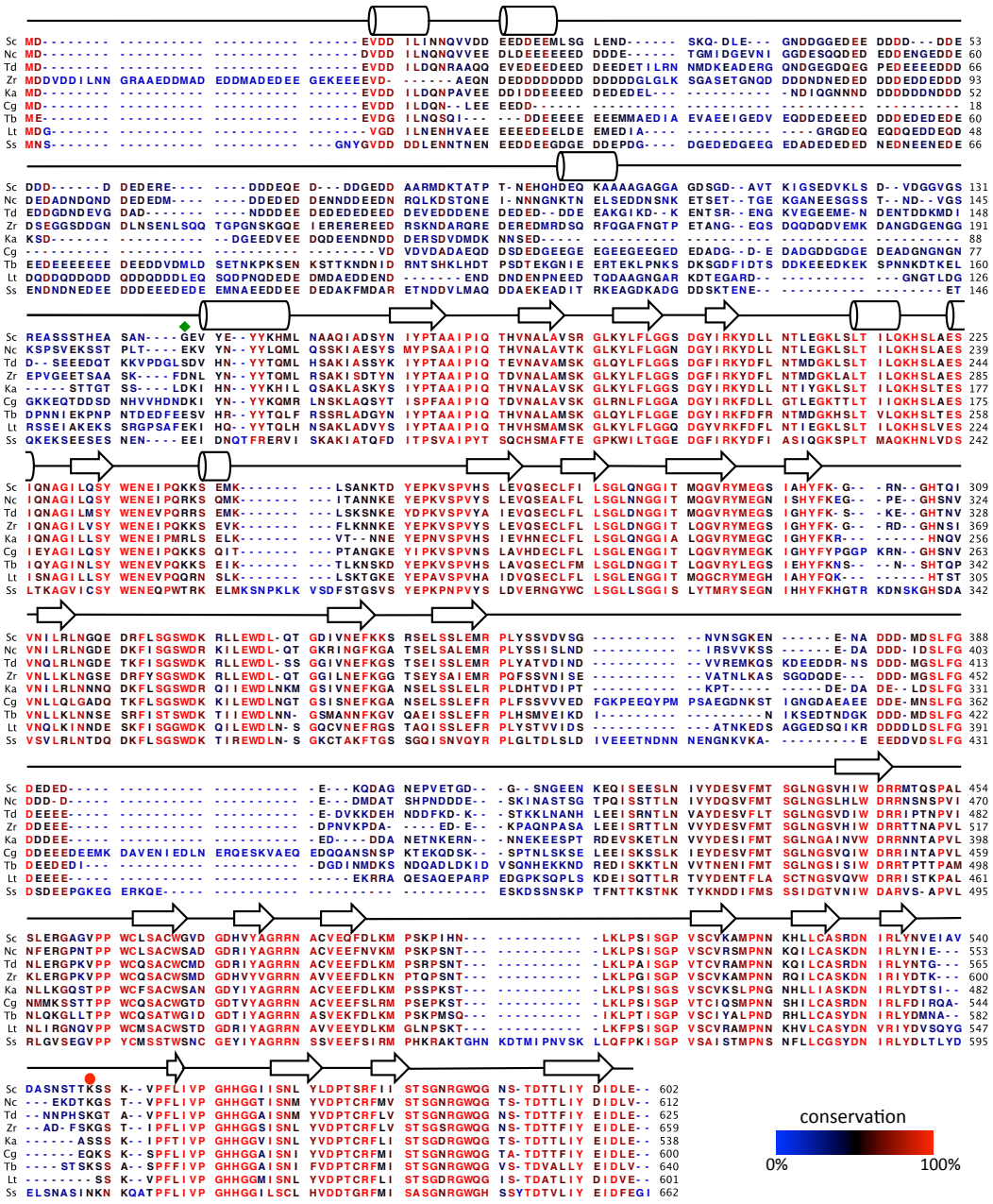

Figure 4.4 Sequence alignment of Spt8 from nine yeast species and secondary structure prediction. Spt8 sequences from *Saccharomyces cerevisiae* (Sc), *Naumovozyma castellii* (Nc), *Torulaspora delbrueckii* (Td), *Zygosaccharomyces rouxii* (Zr), *Kazachstania Africana* (Ka), *Candida glabrata* (Cg), *Tetrapisispora blattae* (Tb), *Lachancea thermotolerans* (Lt), and *Scheffersomyces stipites* (Ss) were aligned using CLC Sequence Viewer (Version 6.8.2). Secondary structure prediction of Sc Spt8 was performed using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), and is shown on top of the alignments, with α-helix, β-stand, and disordered regions represented by cylinder, arrow, and solid lines, respectively. Red dot indicates the lysine residue that crosslinked to TBP. Green diamond indicates the starting residue of Spt8c01.

My next attempt was to purify Spt8 and TBP separately. TBP core domain was successfully purified from *E.coli* through a SUMO tag at the N-terminus, followed by cation exchange chromatography and gel filtration. Spt8c01 was purified as a GST fusion protein using the baculovirus expression system. The GST tag was then removed by TEV protease digestion, as there is a TEV protease recognition site following the GST tag. Next, I tested the interaction between Spt8c01 and TBP core domain in a gel filtration experiment. Unexpectedly, Spt8c01 did not stably bind to TBP core as only two elution peaks were observed (Figure 4.5). Thus, Spt8c01 is not suitable to study the structure of the Spt8/TBP complex.

While my initial attempts to obtain a stable Spt8/TBP complex *in vitro* produced only negative results, more work screening different Spt8 constructs, including the full length protein is under way. In addition, Spt7 constructs are also being screened, as the 152 C-terminal region of Spt7 is required for Spt8 binding to SAGA (Wu & Winston, 2002), and the same region also crosslinked to Spt8 (Figures 2.5, 4.1A).

4.5 Materials and methods

4.5.1 Pull-down experiments.

Recombinant proteins used in pull-down assays were purified from *E.coli* strain BL21 (DE3) RIL. Spt8 was GST-tagged using the pGEX-4T-3 vector, and was purified using Glutathione Sepharose 4 Fast Flow resin (GE Healthcare). TBP and mutant K133,138L were His-tagged using the pET21a vector, and were purified using Ni Sepharose High Performance resin (GE Healthcare). Recombinant TBP proteins were further purified using Source 15S resin (Pharmacia Biotech) in batch.

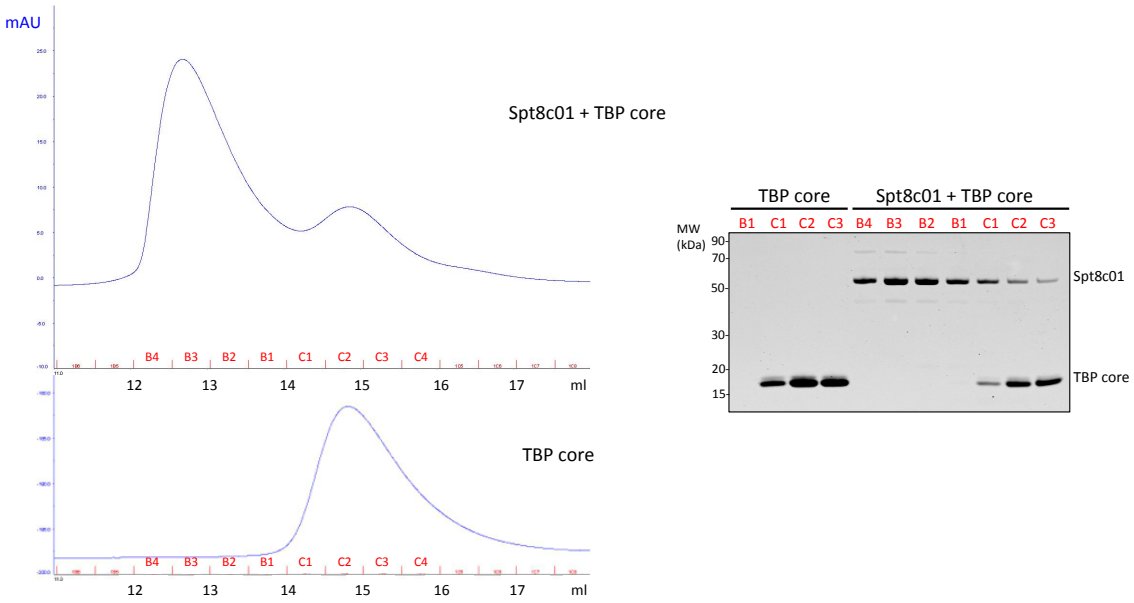

Figure 4.5 Gel filtration experiment shows no interaction between Spt8c01 and TBP core. Chromatograms (280nm absorption) for the gel filtration runs are shown on the left. A Coomassie Blue stained protein gel of the peak fractions are shown on the right.

For GST pull-down assays, 4 μg GST or GST fusion Spt8 WT or mutants were immobilized on 20 μl Glutathione Sepharose resin in 400 μl pull-down buffer A (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM PMSF, 1 mM DTT, 50 μg/ml BSA) for 1 hr at 4°C. Protein bound resin was then washed 3 times with 200 μl pull-down buffer A. Next, 1 μg TBP was incubated with GST protein bound resin in 400 μl pull-down buffer A for 1 hr at 4°C. TBP bound resin was washed 3 times with 500 μl pull-down buffer A, followed by two washes with 500 μl buffer containing 20 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol. All bound proteins were eluted using 20 μl 10 mM glutathione in 20 mM HEPES pH 8, 150 mM NaCl, 10% glycerol, 1 mM PMSF, 1 mM DTT for 10 min at room temperature with gentle shaking. The elution was repeated, and both eluates were combined with a total volume of 40 μl. Eluates were analyzed on SDS-PAGE, and were visualized by Coomassie Blue staining or Western blotting using a polyclonal antibody against TBP.

For His-tag pull-down assays, 4 μg WT or K133,138L TBP proteins were bound with 20 μl Ni Sepharose beads in 400 μl pull-down buffer B (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 0.001% NP-40, 1 mM PMSF, 1 mM DTT, 100 μg/ml BSA) for 30 min at 4°C, then washed three times. 0.5 μg GST or GST-Spt8 were incubated with TBP-bound resin or beads only in 400 μl pull-down buffer B for 1 hr at 4°C, washed three times with Ni wash buffer (20 mM HEPES pH 8, 300 mM NaCl, 10% glycerol, 20 mM imidazole), and eluted in 40 μl Ni elution buffer (20 mM HEPES pH 8, 300 mM NaCl, 10% glycerol, 200 mM imidazole) with agitation for 7 min at room temperature. 4 μl of eluates (10%) were analyzed on SDS-PAGE, and were visualized by Coomassie Blue

staining or Western blot using GST antibody (Z-5, Santa Cruz, sc-459). GST or GST-Spt8 were quantified using the Odyssey infrared imaging system (LiCor Biosciences).

4.5.2 Protein purifications.

Spt8c01 was cloned into pFastBac GTE vector with an N-terminal GST tag followed by a TEV protease recognition site. The Bac-to-Bac system from Invitrogen was used to generate recombinant baculovirus. Next, the baculovirus bearing yeast Spt8c01 gene was used to infect monolayer High Five insect cells for protein expression. After 72 hrs, cells were harvested by centrifugation, washed and resuspended in Buffer S (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1mM PMSF, 1mM DTT). Cells were lysed by sonication, and the lysate was clarified by centrifugation. Cleared WCE was then incubated with 2 ml Glutathione Sepharose resin for 1.5 hr at 4°C. After the incubation, resin was transferred into a 10 ml disposable column (Bio-Rad Laboratories, Inc.), and washed with 10-20 column volumes (CV) of Buffer S. To remove the GST tag, 1 mg TEV protease was added, and on-column cleavage was carried out overnight in 3 CV Buffer S at 4°C. The next day, the flow through fraction of TEV cleavage was collect, and the resin was washed with 2 CV Buffer S. Flow through and wash were combined, as they both contained Spt8c01. TEV protease with an N-terminal His-tag was removed by passing the protein solution through Ni Sepharose beads.

TBP core domain was purified as an N-terminal SUMO tagged protein (6His-SUMO-TBP core in pET21a) as described (Kamenova *et al*, 2014) but with modifications. Briefly, 6 L *E.coli* BL21 (RIL) DE3 cells were grown to OD600 around 0.6-0.7, and induced with 0.4 mM IPTG for 3 hrs at 37°C. After induction, cells were harvested by

centrifugation and washed with 30 ml TBPc lysis buffer (30 mM Tris-HCl pH 7.5, 500 mM KCl, 40 mM imidazole, 10% glycerol, 1 mM PMSF, 1 mM DTT). Cells were then pelleted and quick frozen in liquid nitrogen, and stored in -70°C until further use. Thawed cells were resuspended in 120 ml TBPc lysis buffer, and lysed by sonication. Lysate was then clarified by centrifugation, and the resulting WCE was loaded onto a column containing 4 ml Ni-Sepharose beads equilibrated with TBP lysis buffer (PMSF and DTT were left out), and drained by gravity at 4°C. Protein bound beads were washed 4 times with 10 ml TBPc lysis buffer, and His-tagged protein was eluted 4 times with TBPc elution buffer (30 mM Tris pH 7.5, 500 mM KCl, 500 mM imidazole, 10% glycerol, 1 mM PMSF, 1 mM DTT). Eluates were checked by SDS-PAGE and Coomassie Blue staining, and the fractions with the most protein were combined and dialyzed against SUMO buffer (30 mM Tris-HCl pH 7.5, 300 mM KCl, 10% glycerol, 0.5 mM DTT) overnight at 4°C. After dialysis, the protein solution was cleared by centrifugation to remove precipitates formed overnight. Next, 100 μg SUMO protease was added, and SUMO cleavage was performed for 3 hr at 4°C. SUMO tag was removed by passing the solution through Ni-Sepharose beads equilibrated in SUMO buffer (DTT was left out). Beads were washed with TBPc lysis buffer for 5 CV, and the washes were combined with the flow through. The combined protein solution was dialyzed against S Buffer A (20 mM Tris-HCl pH 7.8, 150 mM KCl, 10% glycerol, 1 mM PMSF, 1 mM DTT) overnight at 4°C. Dialyzed protein was loaded onto a Source 15S column equilibrated in S Buffer A, and eluted with a gradient from 0% to 50% S Buffer B (20 mM Tris-HCl pH 7.8, 1 M KCl, 10% glycerol, 1 mM PMSF, 1 mM DTT). Peak fractions were combined, concentrated, and further purified via gel filtration in a buffer containing

20 mM Tris-HCl pH 7.8, 300 mM KCl, 10% glycerol, 1 mM PMSF, 1 mM DTT. Peak fractions of gel filtration were checked by SDS-PAGE and Coomassie Blue staining, and then combined, and concentrated in Amicon Ultra concentrators with 10,000 molecular weight cutoff (Millipore).

Chapter 5 Model for the molecular architecture of SAGA

SAGA is a multifunctional transcription coactivator that preferentially acts at inducible promoters where its multiple activities stimulate steps during transcription initiation and early elongation. The complexity of SAGA has led to several different models for the organization of its subunits and the architecture of the complex. To resolve these differences, I used an integrated approach to define molecular interactions between subunits and to determine how the different SAGA modules are arranged in the complex. My work has led to a model for SAGA architecture, identification of molecular interactions between the different modules and new insight into the mechanisms and regulation of SAGA-TBP binding.

Previous studies revealed that SAGA contains a subset of the Taf subunits from the coactivator TFIID and two SAGA-specific subunits (Ada1 and Spt7) with Taf-like histone fold domains. Within TFIID, this set of Taf proteins forms an asymmetric core complex at the center of TFIID, which acts as a scaffold for assembly of the other Tafs. My results are consistent with an analogous structure at the center of SAGA. First, the SAGA Taf and Taf-like subunits crosslink to each other as expected from an analogous complex within SAGA. Second, previous studies showed disruption of the SAGA complex upon inactivation of the SAGA Taf and Taf-like proteins. The crosslinking data suggest that these subunits are at the center of SAGA, as they are a unique group of subunits that crosslink to all other SAGA modules. Within the TFIID-like core, SAGA-specific regions within Spt7 and Ada1 outside of their histone fold domains play an important role in bridging other structural modules with the TFIID-like core.

5.1 A Model for the architecture of SAGA

Figure 5.1 shows a model for the architecture of SAGA with the TFIID-like structure at the center. Based on the proposed arrangement of the TFIID subunits, the central core of the SAGA model contains two copies each of the histone-fold containing pairs Taf6-Taf9, and Ada1-Taf12 together with two copies of Taf5. This core structure contains extra SAGA-specific density (not shown) since Ada1 has over 170 residues both N- and C-terminal to its histone fold domain that are involved in interactions with other SAGA subunits. One copy of Spt7-Taf10 is predicted to bind this complex, analogous to the binding of Taf8-Taf10 in TFIID. Spt7 has a broad interface within this structure, as it crosslinks to the widely separated Taf5 N and C-terminal domains. Taf10, which requires only a 91-residue segment containing the histone fold domain for function (Kirschner *et al*, 2002), crosslinks to the histone fold domain of Taf12 in the lower half of the core.

The crosslinking data also reveal the interactions this TFIID-like core makes with the other SAGA modules and more peripheral subunits. For example, in addition to its crosslinks with the Taf5 N-terminal domain, the C-terminus of Spt7 also crosslinks to the Spt8 WD40 repeats, likely positioning Spt8 in the lower half of the TFIID-like core. Sgf73 links the DUB module to the TFIID-like core, demonstrated by its extensive crosslinks with the histone fold domains of Ada1 and Taf12 as well as the Taf5 N-terminus. These crosslinks position the DUB module on the lower half of the TFIID-like core. Spt20, which crosslinks to the DUB module, Tra1 and Spt3, displays numerous crosslinks with the histone fold domains of Ada1, Taf9, Taf10, Taf12 and the N-terminal domain of Taf5. These results locate Spt20 in the center of the TFIID-like complex and explain how Spt20 and Sgf73 can extensively crosslink to each other. Ada3, the HAT

module subunit which links this module to SAGA, crosslinks to the histone fold domains of Ada1, and Taf12, the WD40 repeats in Taf5 and near the HEAT repeat of Taf6. This suggests that the HAT module is situated in the central and upper half of the TFIID-like core and in a position to crosslink with the C-terminus of Sgf73. The TBP-binding subunit, Spt3 crosslinks extensively to the same region of Spt20 that makes extensive crosslinks to the Taf12 histone fold domain and the Taf5 N-terminus. This maps Spt3 to the lower half of the TFIID-like complex, and potentially situated so that both Spt8 and Spt3 can simultaneously interact with TBP.

Wu *et al* reported an EM model of the SAGA complex, in which they mapped the positions of several SAGA subunits through immune labeling and electron microscopic analysis (Wu *et al*, 2004). The results of most of these mapping studies are in good agreement with the crosslinking results, with TAFs in a central position, Spt20 connecting with Spt3, and Spt7 close to Gcn5. The only discrepancy is their mapping of Tra1 and Spt20 at opposite ends of SAGA, in contrast to the crosslinking data, which indicate that Tra1 and Spt20 are in close proximity. I reason that this difference is likely due to the extended structure of Tra1, as Wu *et al* used Myc-antibody to probe the N-terminus of Tra1, and I observed Spt20 crosslinks near the Tra1 C-terminus (Figure 2.5). The model shown in Figure 5.1 is also distinct from another model, based on combinatorial depletion analysis, which linked all Spt subunits, Ada1 and Tra1 in a single group and positioned the Taf subunits in a more peripheral part of the protein-protein interaction network (Lee *et al*, 2011).

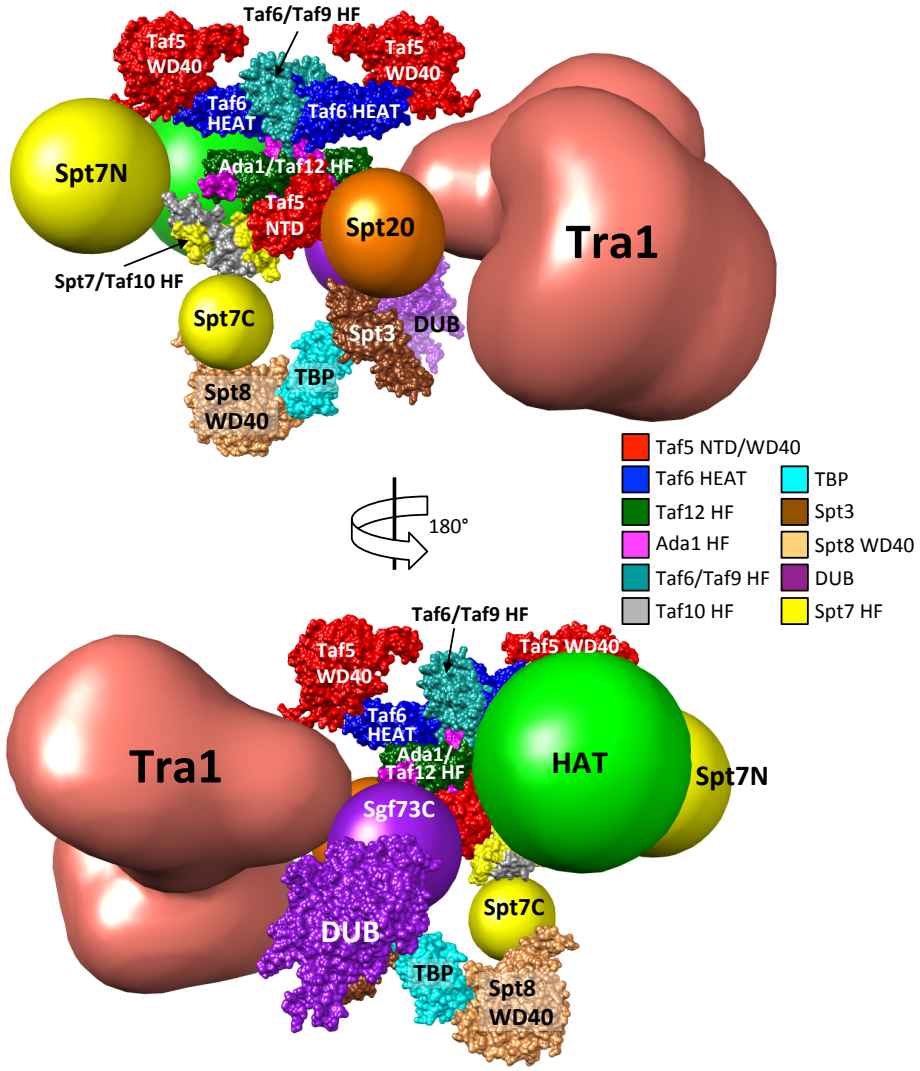

Figure 5.1 Model for the molecular architecture of the SAGA complex. SAGA subunits were positioned around the TFIID core EM model. Homology models of yeast Taf proteins and Ada1 were superimposed onto the coordinates kindly provided by Patrick Schultz (Bieniossek *et al*, 2013) and color coded as indicated. Spt7/Taf10 HF (Histone Fold), Spt3, Spt8 WD40 are homology models. Crystal structures used: TBP, 1TBP; DUB, 3M99; Taf5 NTD (N-terminal domain), 2J49. HF, histone fold domain; Spt7N/C, Spt7 N-terminal/C-terminal region; Sgf73C, Sgf73 C-terminal region. Images were prepared using UCSF Chimera (Pettersen *et al*, 2004).

5.2 A surprisingly small interface between Tra1 and SAGA

Crosslinks between SAGA and the large activator-binding subunit Tra1 are primarily between the Tra1 FAT domain and a region of Taf12 that is N-terminal to the histone fold domain. While the absence of other crosslinks does not rule out additional Tra1-SAGA interactions, this limited region of crosslinking suggests a peripheral position for the majority of Tra1, perhaps on the opposite side of the TFIID-like core from the other modules. Interestingly, the FAT domain also crosslinks to a small segment of Sgf73, which in turn interacts with the HAT module subunits Ada2, Ada3 and Gcn5. It was previously shown that a small deletion within the Tra1 FAT domain caused inhibition of HAT activity (Knutson & Hahn, 2011). From the crosslinking results, this inhibition is likely an indirect effect. One explanation for this inhibitory effect is that the FAT domain mutation alters its interaction with Sgf73, causing an aberrant interaction between Sgf73 and the HAT module. It is also interesting that many small deletions and other mutations throughout Tra1 cause dissociation of Tra1 from other SAGA components (Knutson & Hahn, 2011). Given the domain topology derived from the crosslinking results, most of these mutations likely affect the FAT domain-SAGA interaction. Intramolecular Tra1 crosslinks show interactions between the FAT and HEAT repeat domains as well as interactions between the Tra1 PI3K, FRB and FAT domains (Figure 2.6). I propose that disrupting these intramolecular FAT domain interactions alters the conformation of the FAT domain and disrupts Tra1-SAGA contacts.

5.3 TBP-SAGA interactions

Spt3 and Spt8 both are implicated in binding TBP and can have either positive or negative effects on transcription depending on the regulated gene (Bhaumik, 2011). For

example, deletion of Spt8 causes an increase in basal transcription at some genes (Belotserkovskaya *et al*, 2000) and Spt8 can compete with TFIIA for binding to TBP *in vitro* (Warfield *et al*, 2004). My crosslinking and biochemical results provide an explanation for this competition, as both the unstructured region of the TFIIA large subunit and Spt8 bind a positively charged groove on TBP. Since DNA was found to compete with Spt8 for TBP binding (Sermwittayawong & Tan, 2006), it is likely that the isolated Spt8 subunit additionally interacts with TBP near its DNA binding surface. Although Spt3-TBP binding is genetically the most important SAGA-TBP interaction for gene activation, Spt3 alone cannot bind TBP. The crosslinking data shows that Spt20 is positioned on either side of the Spt3 histone fold domain involved in TBP binding. This finding suggests that Spt20 may enhance Spt3-TBP binding by promoting a conformation of Spt3 that is active for TBP binding. My combined results show a complex network of interactions between TBP and Spt3, 7, 8, and 20.

5.4 SAGA DUB-HAT interactions

The human DUB module subunit Ataxin7 (orthologous to yeast Sgf73) contains a region that can undergo poly-glutamine expansion, causing the disease spinocerebellar ataxia type 7. Human Ataxin7 can partially complement a *SGF73* deletion but the Poly Q form inhibits HAT activity *in vivo* and *in vitro* (Palhan *et al*, 2005; McMahon *et al*, 2005; Burke *et al*, 2013). Here, I have shown that mutations in the DUB module subunit Sgf73 have modest effects on SAGA nucleosomal HAT activity and stability of the HAT module within SAGA. This relationship is consistent with the observed protein crosslinking results as well as the inhibition of HAT activity by the mutant form of Ataxin7.

Taken together, my results have led to a model for SAGA architecture and revealed new and unexpected interactions between the SAGA subunits and modules. My model reveals insights into how the disparate functions of SAGA are coordinated to activate a wide range of genes with different coactivator requirements and suggest new directions for understanding the molecular interactions of SAGA with transcription regulators and components of the transcription machinery.

5.5 Materials and methods

Structural modeling.

Domain prediction and structure similarity searches were performed by HHpred (http://toolkit.tuebingen.mpg.de/hhpred) using HMM PDB under default settings and thresholds (Söding *et al*, 2005). HHpred alignments were then used to manually generate PIR-alignments between query and target sequences. Templates were selected based on the highest probability score and/or orthologous proteins. PIR-alignment files were used as input to generate structural homology models using Modeller 9.10 with default parameters (Sali & Blundell, 1993). Twenty models of each alignment were generated, and the best scoring model was chosen. Sequences and templates used for generating homology models: Taf5 WD40 residues 460-798 were aligned with human platelet-activating factor acetylhydrolase IB alpha subunit residues 102-402 (PDB ID 1VYH_C); Taf6 HEAT residues 216-429 were aligned with *Antonospora locustae* Taf6 residues 161-349 (PDB ID 4ATG_A); Taf9/Taf6 histone fold residues 30-102 (Taf9) and 5-75 (Taf6) were aligned with *Drosophila melanogaster* Taf9 19-86 (PDB ID 1TAF_A) and Taf6 1-70 (PDB ID 1TAF_B); Ada1/Taf12 histone fold residues 266-315 (Ada1) and 413-490 (Taf12) were aligned with human Taf4 870-918 (PDB ID 1H3O_A) and Taf12

55-128 (PDB ID 1H3O_B); Spt7/Taf10 histone fold residues 984-1030+GSGSGS linker+1065-1081 (Spt7) and 87-135+GSGSGS linker+179-194 (Taf10) were aligned with *Methanothermus fervidus* histone HMfA 6-69 (PDB ID 1B67_A) and 3-67 (PDB ID 1B67_B); Spt8 WD40 residues 142-590 were aligned with yeast Rpn14 residues 20-416 (PDB ID 3VL1_A); Spt3 histone fold residues 6-50 and 188-313 were aligned with human Taf13 31-75 (PDB ID 1BH9_A) and Taf11 113-195 (PDB ID 1BH9_B).

Appendix A: Summary of the BS3 crosslinks within the SAGA complex

I list the crosslinked peptide pairs identified in my thesis study in the following table, with the BS3 modified lysine in each peptide indicated by an asterisk (*)."position" indicates the residue number of the BS3 modified lysine in the primary sequence of the protein. "type" is either 1 (intra-links) or 2 (inter-links). "experiment" indicates the crosslinks identified in either 2mM BS3 experiment or 5mM BS3 experiment or both experiments. Number of spectra is the total number of spectra used to identify each crosslink.

Table A.1 Summary of the BS3 crosslinks within SAGA complex.

protein 1	peptide1	position 1	protein 2	peptide2	position 2	type	experiment	spectra number
Spt7	K.ADPTTTVNAK*VGAENDGDSSLFLR.T	1284	Ada1	K.IPIVTNPESLK*R.V	234	2	5mM	1
Spt7	K.ADPTTTVNAK*VGAENDGDSSLFLR.T	1284	Spt8	R.YMEGSIAHYFK*GR.N	301	2	5mM	2
Spt3	K.AGGGEKDEK*DGGNMMK.V	126	Spt3	K.K*SQIK.L	136	1	2mM	1
Tra1	K.AQQLYEVAQVK*AR.S	2713	Spt20	K.K*ASFK.R	272	2	2mM	1
Tra1	K.AQQLYEVAQVK*AR.S	2713	Spt20	K.ASFK*RPR.V	276	2	2mM	1
Tra1	K.AQQLYEVAQVK*AR.S	2713	Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	2	both	2
Spt7	K.ARLPPTGK*ISTTYK.K	1248	Taf12	R.K*WNPSQNYNQK.L	503	2	both	5
Taf5	K.ATTEPSAEPDEPFIGYLGDVTASINQDIK*EYGR.R	754	Spt7	K.HLLSSIQQK*K.S	409	2	5mM	1
Taf5	K.ATTEPSAEPDEPFIGYLGDVTASINQDIK*EYGR.R	754	Spt7	R.SK*WTSDER.L	433	2	5mM	2
Sgf29	K.AVGK*VGR.S	109	Sgf29	R.GSADGEWIQCEVLK*VVADGTR.F	151	1	2mM	1
Sgf29	K.AVGK*VGR.S	109	Sgf29	R.K*ELLLIPPGFPTK.N	181	1	2mM	1
Spt7	K.AYIAK*QSK.S	1204	Spt8	K.TATPTNEHQHDEQK*AAAAGAGGAGDSGDAVTK.L	96	2	2mM	1
Ubp8	K.DGVLK*TCNAAR.Y	22	Ubp8	M.SICPHIQQVFQNEK*SK.D	15	1	2mM	1
Ubp8	K.DGVLK*TCNAAR.Y	22	Sgf73	K.VIEEYSLSQGSGPSNDSWK*SLMSSAK.D	33	2	2mM	2
Ada3	K.DLSDDNLK*FLK.M	315	Taf6	R.GAIVTALNDNSLQTPVTSTTASASVTDTGASQHLSNVK*PGQNTEVKPLVK.H	204	2	2mM	1

Ada3	K.DMDQEEDEDEADVFAQNTNK*DVELN.-	697	Ada1	K.MITK*EAGK.R	201	2	both	3
Gcn5	K.EAGWTPEMDALAQRPK*R.G	328	Gcn5	K.LESNK*YQK.M	385	1	5mM	1
TBP	K.EANK*IVFDPNTR.Q	15	TBP	R.DGTKPATTFQSEEDIK*R.A	47	1	2mM	1
TBP	K.EANK*IVFDPNTR.Q	15	TBP	R.LK*EFK.E	8	1	both	2
Taf5	K.ENEVASAFQSHK*YR.L	241	Taf5	K.TALDLK*LEIQK.V	432	1	2mM	2
Gcn5	K.ENK*GKFEK.E	51	Gcn5	R.VKLENNVEEIQPEQAETNK*QEGTDK.E	42	1	2mM	1
Gcn5	K.ERPSVVEENEGK*IEFR.V	99	Gcn5	R.IGGSEVVTDVEK*GIVK.F	72	1	both	5
Taf9	K.EWDLEDPK*SM.-	155	Spt20	K.K*SFIHEHR.A	355	2	both	3
Tra1	K.FEDIEIPGQYLLNK*DNNVHFIK.L	3362	Tra1	K.SLSK*NVETR.R	3432	1	2mM	2
Gcn5	K.FEFDGVEYTFK*ERPSVVEENEGK.L	87	Gcn5	R.IGGSEVVTDVEK*GIVK.F	72	1	5mM	1
Spt7	K.FMHK*NISK.V	985	Spt7	R.FLANK*DLGLTPK.M	897	1	2mM	2
Tra1	K.FNADFIDNK*PDYETYIKR.L	3304	Tra1	R.IGIEK*NFDLEPTVNKR.D	1059	1	5mM	1
Taf6	K.FPQAYK*SLKPR.V	366	Taf6	R.DFAASLLDYVLK*K.F	359	1	5mM	3
Taf10	K.FVSDIAK*DAYEYSR.L	126	Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	1	5mM	1
Taf10	K.FVSDIAK*DAYEYSR.L	126	Spt7	K.FMHK*NISK.V	985	2	5mM	1
Sgf73	K.HTQQQK*QGQR.S	314	Ada1	K.IPIVTNPESLK*R.V	234	2	2mM	1
Sgf73	K.HTQQQK*QGQR.S	314	Taf12	R.K*WNPSQNYNQK.L	503	2	2mM	1
Taf6	K.HVLSK*ELQIYFNK.V	221	Spt7	K.NGK*LNSDSEAFLK.N	760	2	both	10
Spt7	K.IEQNSIMK*NGFGTVLK.Q	805	Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	2	2mM	1
Ada3	K.IGPLFDK*PEIMK.R	663	Spt7	K.HLLSSIQQK*K.S	409	2	2mM	3
Ada3	K.IGPLFDK*PEIMKR.L	663	Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	2	5mM	1
Spt8	K.IGSEDVK*LSDVDGGVGSR.E	121	Spt8	K.TATPTNEHQHDEQK*AAAAGAGGAGDSGDAVTK.L	96	1	both	3
Taf5	K.IK*DDQEK.Q	339	Taf5	K.TALDLK*LEIQK.V	432	1	both	4
Taf5	K.IK*DDQEKQLNQQTAGDNYSGANNR.T	339	Taf5	K.VK*ESR.D	439	1	2mM	2
Taf5	K.IK*DDQEKQLNQQTAGDNYSGANNR.T	339	Taf5	R.FSPDFK*DFHGSEINR.L	209	1	5mM	1
Spt3	K.IK*QR.E	253	Spt3	K.DK*SQQSSQDNTNFEFASSTLHR.K	265	1	2mM	1
Spt3	K.IK*QR.E	253	Spt3	R.LFDGPENVINPLK*PR.H	301	1	5mM	1
Ada1	K.IPIVTNPESLK*R.V	234	Sgf73	R.AK*QQELQK.L	291	2	2mM	1
Ada1	K.IPIVTNPESLK*R.V	234	Sgf73	K.SQDTGLTPLEIQSQQQK*LR.Q	545	2	2mM	1
Ada1	K.IPIVTNPESLK*R.V	234	Taf12	R.K*ISSSNSTEIPSVTGPDALK.S	284	2	2mM	2
Ada1	K.IPIVTNPESLK*R.V	234	Sgf73	R.STSMESANTPNMDTK*R.S	196	2	5mM	1
Ada1	K.IPIVTNPESLK*R.V	234	Ada1	R.VK*SNNLK.T	237	1	both	13
Ada1	K.IPIVTNPESLK*R.V	234	Sgf73	R.TK*YFR.M	386	2	both	2
Ada1	K.IPIVTNPESLK*R.V	234	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	both	4
Ada1	K.IPIVTNPESLK*R.V	234	Taf12	R.K*SDNLEAR.D	469	2	both	4

TBP	K.IQNIVGSCDVK*FPIR.L	167	TBP	R.LDLK*TVALHAR.N	83	1	2mM	1
TBP	K.IQNIVGSCDVK*FPIR.L	167	TBP	R.DGTK*PATTFQSEEDIKR.A	35	1	both	3
TBP	K.IQNIVGSCDVK*FPIR.L	167	Spt3	R.NMTK*EEYVHWSDCR.Q	190	2	both	4
Ubp8	K.IVHELYGALNTK*QASSSSTSTNR.Q	201	Ubp8	R.SPDK*CFSCALDK.L	181	1	2mM	3
Taf5	K.IWSLDGSSLNNPNIALNNNDK*DEDPTCK.T	512	Spt7	R.MLQNGINK*QSR.F	889	2	both	7
Spt7	K.IYHSFEYDK*ETMIK.R	315	Spt7	R.LK*LEESDK.M	323	1	2mM	2
Spt7	K.IYHSFEYDKETMIK*R.L	320	Spt7	K.RLK*LEESDK.M	323	1	5mM	1
Taf9	K.K*ALPQVMGTWGVR.L	124	Spt20	R.EGTK*GVVGHIEER.D	371	2	5mM	1
Ada3	K.K*EGGDNTPSK.L	19	Sgf29	K.K*LNFLNMSK.D	32	2	2mM	1
Spt7	K.K*ITK.G	1172	Spt7	K.K*IQPEESDSIVYK.K	1159	1	both	2
Ada1	K.K*IVMSLPLNDR.N	183	Taf6	R.VLNVEPLYGYYDGSEVNK*AVSFSK.V	92	2	5mM	1
Ada1	K.K*IVMSLPLNDR.N	183	Taf9	K.NVLNK*NSVGSVSEVGPDSTQEETPR.D	11	2	5mM	1
Ada1	K.K*IVMSLPLNDR.N	183	Ada1	R.INK*HNSQIEVYK.K	173	1	both	5
Taf6	K.K*LGGSPK.D	323	Taf6	R.TLLK*TFLDINR.V	378	1	5mM	1
Taf6	K.K*LGGSPK.D	323	Spt7	K.IYHSFEYDK*ETMIK.R	315	2	both	2
Spt7	K.K*LQDIK.Q	1069	Taf10	K.FVSDIAK*DAYEYSR.L	126	2	both	3
Ada3	K.K*PSPYSNDASTILPK.K	366	Ada3	K.LGPLYTDVWFKDENDK*NSAYK.K	360	1	2mM	1
Ada3	K.K*PSPYSNDASTILPK.K	366	Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	2	2mM	2
Gcn5	K.K*QGFTK.E	223	Gcn5	K.EAGWTPEMDALAQRPK*R.G	328	1	2mM	2
Gcn5	K.K*QGFTK.E	223	Gcn5	K.YQK*MEDFIYDAR.L	388	1	2mM	1
Ada3	K.K*SANELDDNALESGSISCGPLLSR.L	381	Ada3	K.K*PSPYSNDASTILPK.K	366	1	both	2
Spt8	K.K*SEMK.L	244	Spt8	K.LSANK*TDYEPK.V	253	1	2mM	7
Spt8	K.K*SEMK.L	244	Spt7	K.ADPTTTVNAK*VGAENDGDSSLFLR.T	1284	2	2mM	1
Spt20	K.K*SFIHEHR.A	355	Taf5	K.IWSLDGSSLNNPNIALNNNDK*DEDPTCK.T	512	2	2mM	1
Spt7	K.K*SQLGISDYELK.H	410	Spt7	R.QK*IEQNSIMK.N	797	1	2mM	1
Ada3	K.K*TLLELQSQLDSLNEILGTIAR.G	60	Ada3	R.DIGMLNGK*SVR.S	51	1	5mM	1
Taf5	K.K*TPVFK.V	776	Ada1	R.INK*HNSQIEVYK.K	173	2	2mM	1
Ada3	K.K*YNVASYPTNDLK.D	260	Ada3	R.AAAALGLFNEEGLESTGEDFLK*K.K	258	1	5mM	1
Tra1	K.K*YQYNLANGLLFVLK.D	1232	Tra1	K.TAVNSIK*LER.L	1051	1	both	2
Gcn5	K.LESNK*YQK.M	385	Ada2	R.K*PMASVPSCHEVQGFMPGR.L	151	2	2mM	1
Tra1	K.LK*NSIQDNDKESEEFMR.K	504	Tra1	K.EK*AEK.L	499	1	2mM	2
Spt7	K.LK*SFQYDSK.Q	496	Spt7	R.GHAIAMQK*K.S	537	1	2mM	1
Spt7	K.LK*SFQYDSK.Q	496	Spt7	R.SK*WTSDER.L	433	1	both	2

Spt7	K.LK*TNNVEEIMGNWNK.L	293	Spt7	R.LVLNISISK*ETLSK.L	286	1	5mM	1
Ada3	K.LLSSMLK*TLDLTFER.D	35	Ada3	R.DIGMLNGK*SVR.S	51	1	5mM	3
Ada3	K.LLSSMLK*TLDLTFER.D	35	Ada3	K.K*EGGDNTPSK.L	19	1	both	5
Spt7	K.LNSDSEAFLK*NPQR.M	770	Spt7	K.RTEYFK*NGK.L	757	1	both	3
Spt8	K.LSANK*TDYEPK.V	253	Ada1	K.IPIVTNPESLK*R.V	234	2	2mM	1
Ada1	K.MITK*EAGK.R	201	Taf5	R.IFAGHLNDVDCVSFHPNGCYVFTGSSDK*TCR.M	639	2	5mM	4
Ada1	K.MITK*EAGKR.G	201	Sgf73	R.TTDYK*FR.V	428	2	both	3
Gcn5	K.MPK*EYIAR.L	133	Gcn5	K.NIFQK*QLPK.M	126	1	both	4
Spt3	K.NAK*DQDASAGVASGTGNPGAGGEDDLK.K	92	Spt3	R.LK*MADDR.T	179	1	5mM	1
Spt3	K.NAK*DQDASAGVASGTGNPGAGGEDDLKK.A	92	Spt3	K.K*SQIK.L	136	1	2mM	1
Spt7	K.NGK*LNSDSEAFLKNPQR.M	760	Taf6	K.ELQIYFNK*VISTLTAK.S	229	2	both	4
Tra1	K.NSIQDNDK*ESEEFMR.K	512	Tra1	R.K*IVNMSR.S	2351	1	5mM	2
Spt7	K.NVDDPPK*NLDSISSSNIEIDDERR.L	260	Spt7	R.TNAK*ALLK.L	16	1	both	2
Taf9	K.NVLNK*NSVGSVSEVGPDSTQEETPR.D	11	Taf5	K.K*TPVFK.V	776	2	2mM	1
Sgf29	K.NYPPGTK*VLAR.Y	200	Sgf29	R.LRFDGEEEVDK*ETEVTR.R	237	1	5mM	1
Spt7	K.PIASAFILPEEDLENDVK*ADPTTTVNAK.V	1274	Spt8	K.SEMK*LSANK.T	248	2	2mM	1
Gcn5	K.PIDPMTIPGLK*EAGWTPEMDALAQRPK.R	312	Sgf73	R.TK*YFR.M	386	2	2mM	1
Taf6	K.PLVK*HVLSK.E	216	Taf6	R.GAIVTALNDNSLQTPVTSTTASASVTDTGASQHLSNVK*PGQNTEVK.P	204	1	2mM	1
Tra1	K.QFLLSPIFAK*PLR.A	1340	Tra1	R.TTK*EDFAVIQR.Q	3175	1	2mM	1
Ubp8	K.QLGIHK*LPSVLVLQLK.R	353	Spt3	K.K*SQIK.L	136	2	2mM	1
Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	Ada1	K.IPIVTNPESLK*R.V	234	2	2mM	1
Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	Tra1	R.K*IVNMSR.S	2351	2	5mM	1
Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	Tra1	R.K*LCDEGIQLSLIK.W	2815	2	both	4
TBP	K.SEDDSK*LASR.K	133	TBP	K.FTDFK*IQNIVGSCDVK.F	156	1	2mM	1
TBP	K.SEDDSK*LASR.K	133	Spt8	R.LYNVEIAVDASNSTTK*SSK.V	548	2	2mM	3
TBP	K.SEDDSK*LASR.K	133	TBP	R.DGTKPATTFQSEEDIK*R.A	47	1	both	3
Taf6	K.SLK*PR.V	369	Taf6	K.SQADEAAQHMK*QAALTSLR.T	248	1	2mM	1
Sgf73	K.SLMSSAK*DTPLQYDHMNR.E	40	Ubp8	K.DGVLK*TCNAAR.Y	22	2	both	8
Tra1	K.SLSK*NVETR.R	3432	Tra1	R.K*HNMPDVCISQLAR.L	2956	1	both	4
Spt7	K.SMDLNTVLK*K.L	493	Spt7	K.LK*SFQYDSK.Q	496	1	2mM	2
Ada1	K.SNIDDLPDFLNEK*PTFTPLDER.N	460	Spt20	R.LYNFVEDADSILK*K.Y	128	2	5mM	1
Spt7	K.SQLGISDYELK*HLIMDVR.K	421	Spt7	R.SK*WTSDER.L	433	1	5mM	1
Ada1	K.SVSEMAADK*R.D	345	Spt3	R.LK*MADDR.T	179	2	5mM	2
Ada1	K.SVSEMAADK*R.D	345	Spt3	K.DK*SQQSSQDNTNFEFASSTLHR.K	265	2	5mM	1

Taf5	K.TAK*PISNPTNLSSK.R	119	Spt20	R.LYNFVEDADSILK*K.Y	128	2	2mM	1
Taf5	K.TAK*PISNPTNLSSKR.D	119	Taf5	R.TLTPQNK*QSPANTK.T	93	1	2mM	2
Taf5	K.TAK*PISNPTNLSSKR.D	119	Taf5	K.TGK*FPEQSSIPPNPGK.T	103	1	both	20
Taf5	K.TAKPISNPTNLSSK*R.D	130	Taf5	K.TGK*FPEQSSIPPNPGK.T	103	1	both	4
Taf5	K.TAKPISNPTNLSSK*R.D	130	Taf5	K.VK*ESR.D	439	1	both	4
Taf5	K.TALDLK*LEIQK.V	432	Taf5	K.IKDDQEK*QLNQQTAGDNYSGANNR.T	344	1	2mM	1
Taf5	K.TALDLK*LEIQK.V	432	Taf5	R.TLLQEYK*AMNNEK.F	368	1	2mM	2
Taf5	K.TALDLK*LEIQK.V	432	Taf5	K.TAKPISNPTNLSSK*R.D	130	1	both	2
Taf5	K.TALDLK*LEIQK.V	432	Taf5	K.VK*ESR.D	439	1	both	13
Spt8	K.TATPTNEHQHDEQK*AAAAGAGGAGDSGDAVTK.L	96	Spt8	K.K*SEMK.L	244	1	2mM	1
Tra1	K.TAVNSIK*LER.L	1051	Tra1	R.IGIEK*NFDLEPTVNKR.D	1059	1	both	6
Ada2	K.TGGNLSK*SACR.E	407	Ada3	R.LLSAVLKDDNDK*SELQSSK.L	416	2	2mM	1
Ada2	K.TGGNLSK*SACR.E	407	Sgf73	K.HTQQQK*QGQR.S	314	2	2mM	1
Ada2	K.TGGNLSK*SACR.E	407	Ada3	K.LAENK*GSNGGTTSTLPQQIGWITNGINLDYPTFEER.L	483	2	5mM	5
Ada2	K.TGGNLSK*SACR.E	407	Sgf73	R.SK*PYDVLLADYHR.E	265	2	both	2
Taf5	K.TGKFPEQSSIPPNPGK*TAK.P	116	Taf5	K.PISNPTNLSSK*R.D	130	1	2mM	1
Spt7	K.TLDK*MEDASVDR.M	873	Spt7	R.MLQNGINK*QSR.F	889	1	5mM	1
Taf5	K.TLVGHSGTVYSTSFSPDNK*YLLSGSEDK.T	538	Taf5	R.LK*QMR.G	687	1	5mM	1
Taf5	K.TPVFK*VK.F	781	Taf5	R.TVIPTSDLVASFYTK*K.T	775	1	2mM	1
Taf5	K.TPVFK*VK.F	781	Spt7	R.FIGNLSLK*IR.Y	176	2	5mM	1
Spt7	K.TSMTSTEDNSFALLEEDQFVSK*K.T	1235	Spt7	K.AYIAK*QSK.S	1204	1	2mM	1
Taf6	K.VISTLTAK*SQADEAAQHMK.Q	237	Taf6	K.LGGSPK*DDSPQEIHEFLER.T	329	1	both	4
Taf6	K.VISTLTAK*SQADEAAQHMK.Q	237	Spt7	K.NGK*LNSDSEAFLKNPQR.M	760	2	both	20
Gcn5	K.VK*EIPEYSHLID.-	429	Gcn5	R.LEK*FFNNK.V	422	1	both	2
Taf5	K.VK*ESR.D	439	Taf5	R.TLTPQNK*QSPANTK.T	93	1	both	10
Spt7	K.VQAK*K.I	1158	Spt7	K.IQPEESDSIVYK*K.L	1171	1	2mM	1
Taf6	K.VTDEDK*EKLLER.C	484	Taf6	K.DLPDLYEGK*GEK.V	475	1	5mM	2
Taf6	K.VTDEDKEK*LLER.C	486	Spt7	R.IPIK*NYQR.T	8	2	both	2
Taf12	K.VTNVNATASMLNNISSSK*SAIFK.Q	351	Taf12	R.GNVNTSQTEQSK*AK.V	331	1	2mM	1
Sgf29	K.VYK*CNR.K	177	Sgf29	R.LRFDGEEEVDK*ETEVTR.R	237	1	2mM	1
Sgf29	K.VYK*CNR.K	177	Sgf29	R.GSADGEWIQCEVLK*VVADGTR.F	151	1	both	9
Taf9	K.YCLTAK*EWDLEDPKSM.-	147	Taf9	K.K*ALPQVMGTWGVR.L	124	1	5mM	1
Taf5	K.YLLSGSEDK*TVR.L	547	Taf5	R.TLTPQNK*QSPANTK.T	93	1	5mM	1
Taf5	K.YLLSGSEDK*TVR.L	547	Taf9	R.TQYQFKPTAPK*ELMLQLAAER.N	111	2	5mM	1

Taf5	K.YLLSGSEDK*TVR.L	547	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	both	2
TBP	R.AAPESEK*DTSATSGIVPTLQNIVATVTLGCR.L	55	TBP	K.EANK*IVFDPNTR.Q	15	1	5mM	2
TBP	R.AAPESEK*DTSATSGIVPTLQNIVATVTLGCR.L	55	TBP	R.K*YAR.I	138	1	5mM	1
TBP	R.AAPESEK*DTSATSGIVPTLQNIVATVTLGCR.L	55	TBP	K.IQNIVGSCDVK*FPIR.L	167	1	both	7
Sgf73	R.AK*QQELQK.L	291	Sgf73	R.EHQTK*IGAAAEK.R	281	1	2mM	1
Sgf73	R.AK*QQELQK.L	291	Sgf73	R.SLTCK*SHSMGAK.R	250	1	both	2
Spt3	R.ALTNFK*GGR.L	326	Taf12	R.K*SDNLEAR.D	469	2	5mM	1
Tra1	R.AWVTLFPQVYK*SIPK.N	2569	Tra1	K.NEK*YGFVR.S	2576	1	2mM	1
Taf5	R.DAIK*LDNLQLALPSVCMYTFQNTNK.D	446	Taf5	K.VK*ESR.D	439	1	5mM	1
Taf5	R.DAIK*LDNLQLALPSVCMYTFQNTNK.D	446	Taf5	R.TLTPQNK*QSPANTK.T	93	1	both	6
Taf6	R.DFAASLLDYVLK*K.F	359	Taf6	K.SLK*PR.V	369	1	both	3
TBP	R.DGTK*PATTFQSEEDIKR.A	35	TBP	R.AAPESEK*DTSATSGIVPTLQNIVATVTLGCR.L	55	1	5mM	3
TBP	R.DGTK*PATTFQSEEDIKR.A	35	TBP	R.NAEYNPK*R.F	97	1	5mM	1
TBP	R.DGTK*PATTFQSEEDIKR.A	35	TBP	K.SEDDSK*LASR.K	133	1	5mM	1
TBP	R.DGTK*PATTFQSEEDIKR.A	35	Spt3	R.NMTK*EEYVHWSDCR.Q	190	2	5mM	2
TBP	R.DGTK*PATTFQSEEDIKR.A	35	TBP	K.EANK*IVFDPNTR.Q	15	1	both	2
TBP	R.DGTKPATTFQSEEDIK*R.A	47	TBP	K.IQNIVGSCDVK*FPIR.L	167	1	2mM	1
TBP	R.DGTKPATTFQSEEDIK*R.A	47	TBP	R.K*YAR.I	138	1	5mM	1
Ada3	R.DIGMLNGK*SVR.S	51	Sgf29	R.FEVRDPEPDELGNSGK*VYK.C	174	2	2mM	1
Ada3	R.DIGMLNGK*SVR.S	51	Sgf29	R.SNLSLMLNQSREEK*SEENTEDAEEGEGTR.M	82	2	both	7
Taf5	R.DILPLPPK*TALDLK.L	426	Taf5	K.LEIQK*VK.E	437	1	5mM	1
Ada2	R.DK*QAR.I	304	Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	1	2mM	1
Taf6	R.DVLTTDDVSK*ALR.V	71	Taf6	R.HSK*R.D	60	1	5mM	1
Spt3	R.EANIVTLK*R.L	176	Spt3	R.LK*MADDRTR.N	179	1	both	3
Sgf29	R.EEK*SEENTEDAEEGEGTR.M	82	Sgf29	K.VYK*CNR.K	177	1	5mM	4
Spt20	R.EGTK*GVVGHIEER.D	371	Spt20	R.K*QQALSANINPTPFNAR.L	492	1	2mM	1
Spt20	R.EGTK*GVVGHIEER.D	371	Taf9	R.LPPEK*YCLTAK.E	141	2	both	4
Spt7	R.ELGLEK*EFGVLSSSVPLQLLTTQFQTVDGETK.V	1128	Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	2	5mM	2
Spt7	R.ELIWK*FMHK.N	981	Taf6	R.VLNVEPLYGYYDGSEVNK*AVSFSK.V	92	2	5mM	1
Ada2	R.ELLNIDPIK*ANR.L	420	Ada2	K.TGGNLSK*SACR.E	407	1	5mM	1
Ada2	R.ELLNIDPIK*ANR.L	420	Ada3	R.K*AQLIPLVER.Q	573	2	5mM	2
Spt3	R.EQVLQTQK*DKSQQSSQDNTNFEFASSTLHR.K	263	Spt3	R.LFDGPENVINPLK*PR.H	301	1	5mM	2
Tra1	R.EVGVLAYK*R.L	1169	Tra1	R.IYEK*SCLIYGEELALSHSFIPELAK.Q	1174	1	5mM	1

Sgf29	R.FDGEEEVDK*ETEVTR.R	237	Sgf29	R.SYWTSEYNPNAPILVGSEVAYK*PR.R	134	1	5mM	1
Spt7	R.FDQLFLEYK*EQK.A	786	Spt7	K.LNSDSEAFLK*NPQR.M	770	1	5mM	1
Spt7	R.FIGNLSLK*IR.Y	176	Spt7	R.TNAK*ALLK.L	16	1	2mM	1
Spt7	R.FIGNLSLK*IR.Y	176	Taf6	K.K*FPQAYK.S	360	2	5mM	1
Tra1	R.FK*STTDEDLFR.L	3238	Tra1	R.TTK*EDFAVIQR.Q	3175	1	2mM	1
Tra1	R.FK*TLNR.Q	476	Tra1	K.LGK*ENPQEAPR.A	451	1	5mM	1
Tra1	R.FK*TLNR.Q	476	Spt3	R.LFDGPENVINPLK*PR.H	301	2	5mM	1
Tra1	R.FK*TLNR.Q	476	Tra1	R.QYDTIMK*YYGR.Y	487	1	both	3
Spt7	R.FLANK*DLGLTPK.M	897	Spt7	K.TLDK*MEDASVDR.M	873	1	2mM	1
Spt7	R.FLQEYDISNAIPDIVYEGVNTK*TLDKMEDASVDR.M	869	Spt7	R.MLQNGINK*QSR.F	889	1	5mM	1
Taf5	R.FSPDFK*DFHGSEINR.L	209	Taf5	R.EK*LADGIK.V	290	1	2mM	2
Taf5	R.FSPDFK*DFHGSEINR.L	209	Taf5	K.VLSDSENGNGK*QNLEMNSVPVK.L	307	1	2mM	1
Taf5	R.FSPDFK*DFHGSEINR.L	209	Taf5	K.IKDDQEK*QLNQQTAGDNYSGANNR.T	344	1	5mM	1
Taf5	R.FSPDFK*DFHGSEINR.L	209	Taf5	K.VK*ESR.D	439	1	5mM	1
Taf5	R.FSPDFK*DFHGSEINR.L	209	Taf6	K.K*DLPDLYEGKGEK.V	466	2	5mM	5
Taf6	R.GAIVTALNDNSLQTPVTSTTASASVTDTGASQHLSNVK*PGQNTEVKPLVK.H	204	Taf6	K.PGQNTEVK*PLVK.H	212	1	2mM	1
Spt7	R.GHAIAMQK*K.S	537	Spt7	K.LK*TNNVEEIMGNWNK.L	293	1	both	7
Taf5	R.GHGK*NAIYSLSYSK.E	694	Taf5	K.VK*FSR.S	783	1	2mM	2
Taf5	R.GHGK*NAIYSLSYSK.E	694	Taf5	R.VWDLK*K.A	724	1	5mM	1
Taf12	R.GKELMFQAAK*IK.Q	71	Taf12	R.QLQEMAAK*FR.T	44	1	2mM	1
Taf12	R.GNVNTSQTEQSK*AK.V	331	Tra1	R.K*GDQEVR.K	2808	2	2mM	1
Spt3	R.HDK*AK.V	72	Taf10	K.FVSDIAK*DAYEYSR.L	126	2	5mM	1
Taf12	R.HTSIK*EK.E	163	Taf12	K.ETYLK*QNIDR.L	170	1	5mM	1
Tra1	R.IAK*SYPQALHFQLR.T	3161	Taf12	R.GNVNTSQTEQSK*AK.V	331	2	2mM	1
Tra1	R.IAK*SYPQALHFQLR.T	3161	Tra1	R.TTK*EDFAVIQR.Q	3175	1	both	2
Tra1	R.IGIEK*NFDLEPTVNKR.D	1059	Tra1	K.K*YQYNLANGLLFVLK.D	1232	1	5mM	5
Tra1	R.IGIEK*NFDLEPTVNKR.D	1059	Tra1	R.LENK*LDR.A	3324	1	both	7
Ada2	R.IK*PFAR.V	252	Ada2	R.ELLNIDPIK*ANR.L	420	1	5mM	2
Tra1	R.ILGK*LGGR.N	951	Tra1	R.QFLK*PPTDLTEKTELDIDAIADFK.L	961	1	5mM	3
Spt20	R.INK*EK.R	488	Sgf73	R.TPQPINHLTNQNLNPK*QIQR.L	448	2	2mM	1
Spt20	R.INK*EKR.K	488	Taf9	R.TQYQFK*PTAPK.E	106	2	2mM	1
Ada1	R.INK*HNSQIEVYK.K	173	Sgf73	R.TK*YFR.M	386	2	2mM	1

Ada1	R.INK*HNSQIEVYK.K	173	Taf9	K.NVLNK*NSVGSVSEVGPDSTQEETPR.D	11	2	2mM	1
Ada1	R.INK*HNSQIEVYK.K	173	Taf5	K.ATTEPSAEPDEPFIGYLGDVTASINQDIK*EYGR.R	754	2	5mM	2
Ada1	R.INK*HNSQIEVYK.K	173	Ada1	K.MITK*EAGK.R	201	1	both	4
Spt20	R.INKEK*R.K	490	Taf12	R.K*WNPSQNYNQK.L	503	2	2mM	1
Spt7	R.IPIK*NYQR.T	8	Spt7	K.ALLK*LTEK.L	20	1	2mM	1
Spt7	R.IPIK*NYQR.T	8	Spt7	K.LTEK*LFNK.N	24	1	2mM	1
Spt7	R.IPIK*NYQR.T	8	Spt7	K.NVDDPPK*NLDSISSSNIEIDDERR.L	260	1	5mM	1
Spt7	R.IPIK*NYQR.T	8	Spt7	R.AEICLK*R.T	751	1	5mM	1
Spt7	R.IPIK*NYQR.T	8	Spt7	R.TNAK*ALLK.L	16	1	both	8
Spt7	R.IPIK*NYQR.T	8	Spt7	R.FIGNLSLK*IR.Y	176	1	both	6
Spt7	R.IPIK*NYQR.T	8	Spt7	R.SK*WTSDER.L	433	1	both	2
Ada3	R.IPNESVFK*DMDQEEDEDEADVFAQNTNKDVELN.-	677	Ada1	R.INK*HNSQIEVYK.K	173	2	5mM	1
Ada3	R.IPNESVFKDMDQEEDEDEADVFAQNTNK*DVELN.-	697	Taf5	R.GHGK*NAIYSLSYSK.E	694	2	5mM	1
Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	Sgf73	R.SK*PYDVLLADYHR.E	265	2	5mM	1
Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	Taf5	R.TLTPQNK*QSPANTK.T	93	2	5mM	2
Taf5	R.IVLEYLNK*K.G	70	Spt20	R.LK*FIEQWR.L	479	2	5mM	1
Taf5	R.IVLEYLNK*K.G	70	Sgf73	R.EMFASSFSVK*PGYTSPGYGAIHSR.V	401	2	both	6
Taf5	R.IVLEYLNK*K.G	70	Sgf73	R.TPQPINHLTNQNLNPK*QIQR.L	448	2	both	5
Ada3	R.K*AQLIPLVER.Q	573	Ada3	R.ELQGTLK*QVTKK.N	563	1	5mM	1
Ada3	R.K*AQLIPLVER.Q	573	Ada3	K.SLLDK*R.Q	649	1	5mM	2
Ada3	R.K*AQLIPLVER.Q	573	Ada2	R.IK*PFAR.V	252	2	5mM	2
Ada3	R.K*AQLIPLVER.Q	573	Sgf73	R.TK*YFR.M	386	2	5mM	1
Ada3	R.K*AQLIPLVER.Q	573	Taf5	R.VWDLK*K.A	724	2	5mM	1
Tra1	R.K*ELLHATR.H	345	Tra1	R.QYDTIMK*YYGR.Y	487	1	5mM	2
Sgf29	R.K*ELLLIPPGFPTK.N	181	Ada3	R.DIGMLNGK*SVR.S	51	2	2mM	1
Tra1	R.K*GDQEVR.K	2808	Tra1	R.QMFK*TFLALQNFAESR.K	2795	1	2mM	1
Tra1	R.K*GDQEVR.K	2808	Tra1	R.SIITLLSK*PYHTR.Q	2589	1	5mM	1
Tra1	R.K*GDQEVR.K	2808	Tra1	R.K*LCDEGIQLSLIK.W	2815	1	both	8
Tra1	R.K*HNMPDVCISQLAR.L	2956	Tra1	K.NSK*IR.E	3096	1	2mM	1
Taf12	R.K*ISSSNSTEIPSVTGPDALK.S	284	Taf12	R.GNVNTSQTEQSK*AK.V	331	1	2mM	3
Taf12	R.K*ISSSNSTEIPSVTGPDALK.S	284	Tra1	R.TTK*EDFAVIQR.Q	3175	2	2mM	1
Tra1	R.K*IVNMSR.S	2351	Tra1	K.LCK*DHLSISQPK.D	2287	1	2mM	1
Tra1	R.K*LCDEGIQLSLIK.W	2815	Tra1	R.VADWNSDRDALEQSVK*SVMDVPTPR.R	2781	1	5mM	1
Spt20	R.K*NAK.K	570	Spt20	R.SGNNATSNNNN	564	1	2mM	2

				NNNNLDKPK*VK.R				
Spt20	R.K*NAK.K	570	Taf9	R.YTQGVLK*DALVYNDYAGSGNSAGSGLGVEDIR.L	69	2	2mM	1
Ada2	R.K*PMASVPSCHEVQGFMPGR.L	151	Gcn5	K.VK*EIPEYSHLID.-	429	2	5mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Spt20	R.INK*EKR.K	488	1	2mM	3
Spt20	R.K*QQALSANINPTPFNAR.L	492	Spt20	R.SGNNATSNNNNNNNLDK*PK.V	562	1	2mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Ada1	K.SVSEMAADK*R.D	345	2	2mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Sgf73	R.QQQLQQQK*FEAAASYLANATK.L	555	2	2mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf12	R.K*ISSSNSTEIPSVTGPDALK.S	284	2	2mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf9	K.K*ALPQVMGTWGVR.L	124	2	2mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Sgf73	R.EMFASSFSVK*PGYTSPGYGAIHSR.V	401	2	5mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Spt3	R.LK*MADDR.T	179	2	5mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Spt3	R.NMTK*EEYVHWSDCR.Q	190	2	5mM	2
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf10	K.FVSDIAK*DAYEYSR.L	126	2	5mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	2	5mM	1
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	2	5mM	2
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf12	R.K*WNPSQNYNQK.L	503	2	5mM	2
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf9	R.YTQGVLK*DALVYNDYAGSGNSAGSGLGVEDIR.L	69	2	5mM	2
Spt20	R.K*QQALSANINPTPFNAR.L	492	Spt20	R.LK*FIEQWR.L	479	1	both	9
Spt20	R.K*QQALSANINPTPFNAR.L	492	Taf9	R.TQYQFK*PTAPK.E	106	2	both	4
Taf12	R.K*SDNLEAR.D	469	Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	1	5mM	1
Taf12	R.K*SDNLEAR.D	469	Sgf73	R.TPQPINHLTNQNLNPK*QIQR.L	448	2	5mM	1
Taf12	R.K*SDNLEAR.D	469	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	5mM	1
Taf12	R.K*SDNLEAR.D	469	Taf10	K.FVSDIAK*DAYEYSR.L	126	2	5mM	1
Taf12	R.K*SDNLEAR.D	469	Taf9	R.TQYQFK*PTAPK.E	106	2	5mM	1
Taf12	R.K*SDNLEAR.D	469	Taf9	R.LPPEK*YCLTAK.E	141	2	5mM	1
Taf12	R.K*SDNLEAR.D	469	Taf12	R.K*WNPSQNYNQK.L	503	1	both	2
Tra1	R.K*VLEPSDDDHLMPQPK.K	520	Tra1	K.NSIQDNDK*ESEEFMR.K	512	1	5mM	2
Taf12	R.K*WNPSQNYNQK.L	503	Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	1	2mM	1
Taf12	R.K*WNPSQNYNQK.L	503	Taf12	K.LQSITSDK*VAAAK.N	521	1	2mM	1
Taf12	R.K*WNPSQNYNQK.L	503	Ada3	K.K*EGGDNTPSK.L	19	2	2mM	1
Taf12	R.K*WNPSQNYNQK.L	503	Spt3	R.GQVIEILLQSNK*TAHLR.G	49	2	2mM	1
Taf12	R.K*WNPSQNYNQK.L	503	Sgf73	R.SK*PYDVLLADYHR.E	265	2	5mM	1

Taf12	R.K*WNPSQNYNQK.L	503	Spt20	R.LK*FIEQWR.L	479	2	both	2
TBP	R.K*YAR.I	138	TBP	K.EANK*IVFDPNTR.Q	15	1	2mM	1
TBP	R.K*YAR.I	138	Spt8	R.LYNVEIAVDASNSTTK*SSK.V	548	2	2mM	1
Spt8	R.K*YDLLNTLEGK.L	201	Spt8	R.GLK*YLFLGGSDGYIR.K	188	1	5mM	2
TBP	R.LDLK*TVALHAR.N	83	TBP	R.IIQK*IGFAAK.F	145	1	2mM	1
TBP	R.LDLK*TVALHAR.N	83	TBP	R.DGTK*PATTFQSEEDIKR.A	35	1	5mM	1
Spt3	R.LFDGPENVINPLK*PR.H	301	Taf12	R.K*ISSSNSTEIPSVTGPDALK.S	284	2	2mM	2
Spt3	R.LFDGPENVINPLK*PR.H	301	Taf12	R.VMSK*R.K	417	2	2mM	2
Spt3	R.LFDGPENVINPLK*PR.H	301	Spt3	K.AK*VNR.L	74	1	5mM	1
Spt3	R.LFDGPENVINPLK*PR.H	301	Spt3	R.EQVLQTQKDK*SQQSSQDNTNFEFASSTLHR.K	265	1	5mM	4
Spt3	R.LFDGPENVINPLK*PR.H	301	Spt20	R.TTTITNSTFAVSLTK*NAMEIASSSSNGVR.G	415	2	5mM	2
Spt3	R.LFDGPENVINPLK*PR.H	301	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	5mM	1
Spt3	R.LFDGPENVINPLK*PR.H	301	Spt20	R.LK*FIEQWR.L	479	2	both	8
Tra1	R.LFNK*SLSK.N	3428	Tra1	R.K*HNMPDVCISQLAR.L	2956	1	2mM	4
Spt20	R.LK*FIEQWR.L	479	Taf12	R.QLQEMAAK*FR.T	44	2	5mM	1
Spt20	R.LK*FIEQWR.L	479	Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	1	both	3
Spt20	R.LK*FIEQWR.L	479	Ada1	K.IPIVTNPESLK*R.V	234	2	both	3
Spt20	R.LK*FIEQWR.L	479	Taf9	R.TQYQFK*PTAPK.E	106	2	both	2
Spt20	R.LK*FIEQWR.L	479	Taf9	R.LPPEK*YCLTAK.E	141	2	both	2
Spt7	R.LK*LEESDKMIEK.G	323	Taf6	K.K*LGGSPK.D	323	2	2mM	1
Spt3	R.LK*MADDR.T	179	Spt3	R.NMTK*EEYVHWSDCR.Q	190	1	2mM	2
Taf5	R.LK*QMR.G	687	Taf5	R.TLTPQNK*QSPANTK.T	93	1	5mM	1
Taf5	R.LK*QMR.G	687	Taf5	R.GHGK*NAIYSLSYSK.E	694	1	both	5
TBP	R.LKEFK*EANK.L	11	TBP	K.SEDDSK*LASR.K	133	1	2mM	1
Spt7	R.LKLEESDK*MIEK.G	329	Spt7	K.LK*SFQYDSK.Q	496	1	2mM	1
Ada1	R.LNNIPK*IPIVTNPESLKR.V	223	Ada1	R.VK*SNNLK.T	237	1	2mM	1
Ada1	R.LNNIPK*IPIVTNPESLKR.V	223	Sgf73	R.TK*YFR.M	386	2	both	2
Taf9	R.LPPEK*YCLTAK.E	141	Taf9	K.K*ALPQVMGTWGVR.L	124	1	both	5
Spt7	R.LPPTGK*ISTTYK.K	1248	Spt7	K.ADPTTTVNAK*VGAENDGDSSLFLR.T	1284	1	2mM	1
Spt7	R.LPPTGK*ISTTYK.K	1248	TBP	R.LDLK*TVALHAR.N	83	2	2mM	1
Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	2	5mM	1
Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	Taf5	R.IVLEYLNK*K.G	70	2	5mM	1
Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	Spt20	K.K*ASFK.R	272	1	both	2
Spt20	R.LQQQQK*QPELTSDGL	161	Spt20	R.K*QQALSANINP	492	1	both	3

	ILTK.N			TPFNAR.L				
Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	Sgf73	R.EMFASSFSVK*PGYTSPGYGAIHSR.V	401	2	both	8
Sgf29	R.LRFDGEEEVDK*ETEVTR.R	237	Sgf29	R.YPETTTFYPAIVIGTK*R.D	220	1	both	14
Spt7	R.LVLNISISK*ETLSK.L	286	Spt7	R.GHAIAMQK*K.S	537	1	2mM	1
Spt8	R.LYNVEIAVDASNSTTK*SSK.V	548	Spt8	K.AMPNNK*HLLCASR.D	521	1	2mM	1
Spt8	R.MDK*TATPTNEHQHDEQK.A	82	Spt8	K.AAAAGAGGAGDSGDAVTK*IGSEDVK.L	114	1	2mM	2
Spt7	R.MLQNGINK*QSR.F	889	Spt7	K.DLGLTPK*MNQNITLIQQIR.H	904	1	2mM	1
Spt7	R.MLQNGINK*QSR.F	889	Spt7	R.IPIK*NYQR.T	8	1	5mM	1
Spt7	R.MLQNGINK*QSR.F	889	Spt7	R.FLANK*DLGLTPK.M	897	1	5mM	1
Spt7	R.MLQNGINK*QSR.F	889	Taf5	K.VK*FSR.S	783	2	5mM	3
Tra1	R.MNAAGTPDTWINWVK*R.V	1971	Tra1	R.K*IVNMSR.S	2351	1	5mM	1
Gcn5	R.MYNGENTSYYK*YANR.L	415	Ada2	R.K*PMASVPSCHEVQGFMPGR.L	151	2	5mM	2
Gcn5	R.MYNGENTSYYK*YANR.L	415	Gcn5	R.LEK*FFNNK.V	422	1	both	11
TBP	R.NAEYNPK*R.F	97	TBP	R.DGTKPATTFQSEEDIK*R.A	47	1	5mM	1
TBP	R.NAEYNPK*R.F	97	TBP	R.AAPESEK*DTSATSGIVPTLQNIVATVTLGCR.L	55	1	5mM	2
TBP	R.NAEYNPK*R.F	97	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	5mM	1
TBP	R.NAEYNPK*R.F	97	TBP	R.LDLK*TVALHAR.N	83	1	both	12
Ada1	R.NGNESSWGFGNGSNNPNNK*LK.R	167	Ada1	K.MITK*EAGKR.G	201	1	2mM	1
Spt3	R.NMTK*EEYVHWSDCR.Q	190	TBP	R.DGTKPATTFQSEEDIK*R.A	47	2	5mM	1
Taf12	R.NNSNK*FSNMIK.Q	131	Taf12	R.QLQEMAAK*FR.T	44	1	2mM	1
Taf12	R.NNSNK*FSNMIK.Q	131	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	2mM	1
Spt7	R.NRADLEK*EIEDMEK.D	559	Spt7	K.DK*DYELDEEEEVAGSGR.K	568	1	both	7
Spt7	R.NYTEHSTPFLNK*VSKR.E	469	Spt7	K.RLK*LEESDK.M	323	1	5mM	1
Gcn5	R.PSVVEENEGK*IEFR.V	99	Gcn5	K.FEFDGVEYTFK*ER.P	87	1	5mM	1
Spt7	R.QK*IEQNSIMK.N	797	Spt7	R.FDQLFLEYK*EQK.A	786	1	2mM	1
Spt7	R.QK*IEQNSIMK.N	797	Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	2	2mM	2
Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	Sgf73	R.AK*QQELQK.L	291	2	2mM	1
Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	Taf12	K.VAAAK*NNGNNVASLNTKK.-	526	2	2mM	6
Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	Spt20	R.LK*FIEQWR.L	479	2	5mM	2
Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	Spt7	R.LPPTGK*ISTTYK.K	1248	2	both	3
Taf10	R.QLLQGQQQPGVQQISQQQHQQNEK*TTASK.V	175	Taf12	K.LQSITSDK*VAAAK.N	521	2	both	2
Taf12	R.QLQEMAAK*FR.T	44	Ada1	R.VK*SNNLK.T	237	2	2mM	1
Spt20	R.QQQALEQQQNGGAMKNANK*R.S	543	Spt20	K.RK*QQALSANINPTPFNAR.L	492	1	5mM	2
Sgf73	R.QQQLQQQK*FEAAASYLANATK.L	555	Ada1	R.VK*SNNLK.T	237	2	2mM	1

Tra1	R.QTMAVMGDK*PDTNDR.N	3192	Tra1	R.IAK*SYPQALHFQLR.T	3161	1	5mM	1
Tra1	R.QTMAVMGDK*PDTNDRNGR.R	3192	Tra1	R.TTK*EDFAVIQR.Q	3175	1	both	5
Spt20	R.SGNNATSNNNNNNNLDK*PK.V	562	Spt20	K.NANK*R.S	543	1	2mM	1
Spt20	R.SGNNATSNNNNNNNLDK*PK.V	562	Spt20	R.K*NAK.K	570	1	2mM	1
Spt20	R.SGNNATSNNNNNNNLDK*PK.V	562	Spt20	K.VK*RPR.K	566	1	both	3
Ada2	R.SK*EAK.E	242	Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	1	2mM	5
Sgf73	R.SK*PYDVLLADYHR.E	265	Sgf73	R.STSMESANTPNMDTK*R.S	196	1	5mM	1
Sgf73	R.SK*PYDVLLADYHR.E	265	Sgf73	R.TK*YFR.M	386	1	5mM	1
Sgf73	R.SK*PYDVLLADYHR.E	265	Ada1	K.IPIVTNPESLK*R.V	234	2	5mM	1
Sgf73	R.SK*PYDVLLADYHR.E	265	Spt20	R.LQQQQK*QPELTSDGLILTK.N	161	2	5mM	1
Sgf73	R.SK*PYDVLLADYHR.E	265	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	5mM	2
Spt7	R.SK*WTSDER.L	433	Spt7	K.SMDLNTVLK*K.L	493	1	2mM	1
Spt7	R.SK*WTSDER.L	433	Ada3	K.QIDQAYVK*R.L	606	2	5mM	1
Spt7	R.SK*WTSDER.L	433	Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	2	5mM	1
Spt7	R.SK*WTSDER.L	433	Taf5	R.TLTPQNK*QSPANTK.T	93	2	5mM	1
Spt7	R.SK*WTSDER.L	433	Taf5	R.GHGK*NAIYSLSYSK.E	694	2	5mM	1
Spt7	R.SK*WTSDER.L	433	Taf6	R.GAIVTALNDNSLQTPVTSTTASASVTDTGASQHLSNVK*PGQNTEVKPLVK.H	204	2	5mM	1
Ada2	R.SKEAK*ELYNR.L	245	Ada2	R.IK*PFAR.V	252	1	2mM	2
Ada2	R.SKEAK*ELYNR.L	245	Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	1	5mM	2
Sgf73	R.SLTCK*SHSMGAK.R	250	Sgf73	R.SK*PYDVLLADYHR.E	265	1	2mM	1
Sgf73	R.SLTCK*SHSMGAK.R	250	Ada1	K.IPIVTNPESLK*R.V	234	2	2mM	1
Sgf73	R.SLTCK*SHSMGAK.R	250	Ada1	R.VK*SNNLK.T	237	2	2mM	2
Sgf73	R.SLTCK*SHSMGAK.R	250	Taf12	R.K*WNPSQNYNQK.L	503	2	2mM	1
Ada2	R.SNGLTTLEAGLK*YER.D	299	Ada2	R.ISSFEK*FGASTAASLSEGNSR.Y	313	1	2mM	1
Ada2	R.SNGLTTLEAGLK*YER.D	299	Ada3	K.K*PSPYSNDASTILPK.K	366	2	2mM	1
Ada2	R.SNGLTTLEAGLK*YER.D	299	Ada2	R.DK*QAR.I	304	1	5mM	2
Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	Ada1	K.IPIVTNPESLK*R.V	234	2	5mM	1
Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	Sgf73	R.TK*YFR.M	386	2	5mM	1
Taf12	R.SNRPTITGGSAMNASALNTPATTK*LPPYEMDTQR.V	403	Spt3	R.LFDGPENVINPLK*PR.H	301	2	5mM	2
Sgf73	R.STSMESANTPNMDTK*R.S	196	Sgf73	R.SK*TGTPQTFSSSIK.K	199	1	both	3
Tra1	R.SWIFNTEIFPTVK*EK.A	2370	Tra1	R.K*IVNMSR.S	2351	1	2mM	1
Sgf29	R.SYWTSEYNPNAPILVGSEVAYK*PR.R	134	Sgf29	R.MALSQGK*K.A	104	1	2mM	1
Sgf29	R.SYWTSEYNPNAPILVG	134	Sgf29	R.K*ELLLIPPGFPT	181	1	both	3

	SEVAYK*PR.R			K.N				
Sgf29	R.SYWTSEYNPNAPILVG SEVAYK*PR.R	134	Ada3	R.DIGMLNGK*SV R.S	51	2	both	3
Sgf73	R.TK*YFR.M	386	Sgf73	R.EMFASSFSVK*P GYTSPGYGAIHSR. V	401	1	5mM	1
Sgf73	R.TK*YFR.M	386	Ada2	K.TGGNLSK*SACR .E	407	2	5mM	1
Sgf73	R.TK*YFR.M	386	Taf5	R.FSPDFK*DFHGS EINR.L	209	2	5mM	1
Sgf73	R.TK*YFR.M	386	Taf5	R.WLSTGSEDGIIN VWDIGTGK*R.L	684	2	5mM	1
Sgf73	R.TK*YFR.M	386	Taf9	R.TQYQFK*PTAPK .E	106	2	5mM	1
Taf5	R.TLLQEYK*AMNNEK.F	368	Taf5	K.IKDDQEK*QLN QQTAGDNYSGAN NR.T	344	1	2mM	1
Taf5	R.TLLQEYK*AMNNEK.F	368	Taf5	K.IK*DDQEKQLN QQTAGDNYSGAN NR.T	339	1	5mM	1
Taf5	R.TLTPQNK*QSPANTK. T	93	Taf5	K.K*GYHR.T	71	1	2mM	1
Taf5	R.TLTPQNK*QSPANTK. T	93	Sgf73	R.AK*QQELQK.L	291	2	2mM	1
Taf5	R.TLTPQNK*QSPANTK. T	93	Taf5	R.DAIKLDNLQLAL PSVCMYTFQNTNK *DMSCLDFSDDCR .L	467	1	5mM	2
Taf5	R.TLTPQNK*QSPANTK. T	93	Sgf73	R.TK*YFR.M	386	2	5mM	2
Taf5	R.TLTPQNK*QSPANTK. T	93	Spt20	R.K*QQALSANINP TPFNAR.L	492	2	5mM	1
Taf5	R.TLTPQNK*QSPANTK. T	93	Taf5	K.TGK*FPEQSSIPP NPGK.T	103	1	both	10
Taf5	R.TLTPQNK*QSPANTK. T	93	Taf5	K.TAKPISNPTNLS SK*R.D	130	1	both	2
Taf5	R.TLTPQNK*QSPANTK. T	93	Taf5	R.WLSTGSEDGIIN VWDIGTGK*R.L	684	1	both	2
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Sgf73	R.AK*QQELQK.L	291	1	2mM	2
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Sgf73	R.TK*YFR.M	386	1	2mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Ada1	K.IPIVTNPESLK*R. V	234	2	2mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Spt20	R.K*QQALSANINP TPFNAR.L	492	2	2mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Sgf73	R.STSMESANTPN MDTK*R.S	196	1	5mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Ada2	R.ISSFEK*FGASTA ASLSEGNSR.Y	313	2	5mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Ada3	K.K*SANELDDNA LESGSISCGPLLSR. L	381	2	5mM	1
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Ada1	R.VK*SNNLK.T	237	2	both	7
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Spt20	R.LQQQQK*QPELT SDGLILTK.N	161	2	both	12
Sgf73	R.TPQPINHLTNQNLNPK *QIQR.L	448	Tra1	R.K*GDQEVR.K	2808	2	both	3
Taf9	R.TQYQFK*PTAPK.E	106	Taf9	R.YTQGVLK*DAL VYNDYAGSGNSA GSGLGVEDIR.L	69	1	5mM	1
Taf9	R.TQYQFK*PTAPK.E	106	Taf5	R.FSPDFK*DFHGS EINR.L	209	2	5mM	1
Taf9	R.TQYQFK*PTAPK.E	106	Taf5	R.IFAGHLNDVDC VSFHPNGCYVFTG SSDK*TCR.M	639	2	5mM	1
Taf9	R.TQYQFK*PTAPK.E	106	Taf9	K.K*ALPQVMGTW GVR.L	124	1	both	5

Taf9	R.TQYQFK*PTAPK.E	106	Taf5	K.YLLSGSEDK*TVR.L	547	2	both	3
Taf9	R.TQYQFKPTAPK*ELMLQLAAER.N	111	Taf9	K.EWDLEDPK*SM.-	155	1	5mM	1
Taf9	R.TQYQFKPTAPK*ELMLQLAAER.N	111	Taf12	R.K*SDNLEAR.D	469	2	5mM	1
Sgf73	R.TTDYK*FR.V	428	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	2mM	1
Sgf73	R.TTDYK*FR.V	428	Sgf73	R.TPQPINHLTNQNLNPK*QIQR.L	448	1	both	9
Tra1	R.TTK*EDFAVIQR.Q	3175	Sgf73	R.QQQLQQQK*FEAAASYLANATK.L	555	2	2mM	1
Spt20	R.TTTITNSTFAVSLTK*NAMEIASSSSNGVR.G	415	Spt20	R.K*QQALSANINPTPFNAR.L	492	1	5mM	2
Spt20	R.TTTITNSTFAVSLTK*NAMEIASSSSNGVR.G	415	Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	2	5mM	2
Spt3	R.TYLSWK*DLR.K	85	Spt3	K.NAK*DQDASAGVASGTGNPGAGGEDDLKK.A	92	1	both	5
Spt3	R.TYLSWK*DLR.K	85	Spt3	R.LK*MADDR.T	179	1	both	6
Tra1	R.VADWNSDRDALEQSVK*SVMDVPTPR.R	2781	Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	2	5mM	3
Tra1	R.VADWNSDRDALEQSVK*SVMDVPTPR.R	2781	Tra1	R.K*GDQEVR.K	2808	1	both	18
Sgf73	R.VGCLDLDRTTDYK*FR.V	428	Sgf73	R.TK*YFR.M	386	1	5mM	1
Ada1	R.VK*SNNLK.T	237	Sgf73	R.STSMESANTPNMDTK*R.S	196	2	2mM	1
Ada1	R.VK*SNNLK.T	237	Taf12	K.QTEPAIPISENISTK*TPAPVAYR.S	371	2	2mM	1
Ada1	R.VK*SNNLK.T	237	Spt20	R.K*QQALSANINPTPFNAR.L	492	2	5mM	1
Ada1	R.VK*SNNLK.T	237	Tra1	R.VADWNSDRDALEQSVK*SVMDVPTPR.R	2781	2	5mM	1
Ada1	R.VK*SNNLK.T	237	Spt20	R.LK*FIEQWR.L	479	2	both	4
Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	Ada3	K.AANSSLK*SLLDKR.Q	644	2	2mM	1
Gcn5	R.VVNNDNTK*ENMMVLTGLK.N	111	Gcn5	R.K*PLTVVGGITYRPFDK.R	153	1	both	4
Taf5	R.VWDLK*K.A	724	Taf5	R.WLSTGSEDGIINVWDIGTGK*R.L	684	1	5mM	1
Ada3	R.WINK*IGPLFDKPEIMK.R	656	Spt7	R.SK*WTSDER.L	433	2	5mM	1
Taf5	R.WLSTGSEDGIINVWDIGTGK*R.L	684	Taf5	R.LK*QMR.G	687	1	both	6
Gcn5	R.YLDAGK*ILLLQEAALR.R	265	Ada2	R.IK*PFAR.V	252	2	5mM	1
Tra1	R.YLVK*QSLDVLTPVLHER.M	1943	Tra1	R.MNAAGTPDTWINWVK*R.V	1971	1	5mM	1
Sgf29	R.YPETTTFYPAIVIGTK*R.D	220	Ada3	R.DIGMLNGK*SVR.S	51	2	both	3
Taf9	R.YTQGVLK*DALVYNDYAGSGNSAGSGLGVEDIR.L	69	Taf9	K.K*ALPQVMGTWGVR.L	124	1	5mM	1
Taf9	R.YTQGVLK*DALVYNDYAGSGNSAGSGLGVEDIR.L	69	Taf12	R.K*WNPSQNYNQK.L	503	2	both	6

References

Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* **422:** 198–207

Anandapadamanaban M, Andresen C, Helander S, Ohyama Y, Siponen MI, Lundström P, Kokubo T, Ikura M, Moche M & Sunnerhagen M (2013) High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nature Structural & Molecular Biology* **20:** 1008–1014

Atanassov BS, Evrard YA, Multani AS, Zhang Z, Tora L, Devys D, Chang S & Dent SYR (2009) Gcn5 and SAGA Regulate Shelterin Protein Turnover and Telomere Maintenance. *MOLCEL* **35:** 352–364

Auble DT & Hahn S (1993) An ATP-dependent inhibitor of TBP binding to DNA. *Genes & Development* **7:** 844–856

Back JW, de Jong L, Muijsers AO & de Koster CG (2003) Chemical Cross-linking and Mass Spectrometry for Protein Structural Modeling. *Journal of Molecular Biology* **331:** 303–313

Bagby S, Mal TK, Liu D, Raddatz E, Nakatani Y & Ikura M (2000) TFIIA-TAF regulatory interplay: NMR evidence for overlapping binding sites on TBP. *FEBS Letters* **468:** 149–154

Baker SP & Grant PA (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* **26:** 5329–5340

Balasubramanian R (2001) Role of the Ada2 and Ada3 Transcriptional Coactivators in Histone Acetylation. *Journal of Biological Chemistry* **277:** 7989–7995

Belotserkovskaya R, Sterner DE, Deng M, Sayre MH, Lieberman PM & Berger SL (2000) Inhibition of TATA-binding protein function by SAGA subunits Spt3 and Spt8 at Gcn4-activated promoters. *Molecular and Cellular Biology* **20:** 634–647

Bhaumik SR (2011) Distinct regulatory mechanisms of eukaryotic transcriptional activation by SAGA and TFIID. *Biochim. Biophys. Acta* **1809:** 97–108

Bhaumik SR & Green MR (2002) Differential Requirement of SAGA Components for Recruitment of TATA-Box-Binding Protein to Promoters In Vivo. *Molecular and Cellular Biology* **22:** 7365–7371

Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu M, Zhou BO, Fleharty BE, Paulson A, Allali-Hassani A, Zhou J-Q, Mer G, Grant PA, Workman JL, Zang J & Min J (2011) Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *The EMBO Journal* **30:** 2829–2842

Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P & Berger I (2013) The architecture of human general transcription factor TFIID core complex. *Nature* **493:** 699–702

Birck C, Poch O, Romier C, Ruff M, Mengus G, Lavigne AC, Davidson I & Moras D (1998) Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* **94:** 239–249

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P & Boeke JD (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14:** 115–132

Brown CE (2001) Recruitment of HAT Complexes by Direct Activator Interactions with the ATM-Related Tra1 Subunit. *Science* **292:** 2333–2337

Bu P, Evrard YA, Lozano G & Dent SYR (2007) Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Molecular and Cellular Biology* **27:** 3405–3416

Bui KH, Appen von A, DiGuilio AL, Ori A, Sparks L, Mackmull M-T, Bock T, Hagen W, Andrés-Pons A, Glavy JS & Beck M (2013) Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. *Cell* **155:** 1233–1243

Buratowski S & Zhou H (1992) Transcription factor IID mutants defective for interaction with transcription factor IIA. *Science* **255:** 1130–1132

Burke TL, Miller JL & Grant PA (2013) Direct Inhibition of Gcn5 Protein Catalytic Activity by Polyglutamine-expanded Ataxin-7. *Journal of Biological Chemistry* **288:** 34266–34275

Candau R & Berger SL (1996) Structural and functional analysis of yeast putative adaptors. Evidence for an adaptor complex in vivo. *J. Biol. Chem.* **271:** 5237–5245

Chasman DI, Flaherty KM, Sharp PA & Kornberg RD (1993) Crystal structure of yeast TATA-binding protein and model for interaction with DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90:** 8174–8178

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills J-C, Nilges M, Cramer P & Rappsilber J (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *The EMBO Journal* **29:** 717–726

Cieniewicz AM, Moreland L, Ringel AE, Mackintosh SG, Raman A, Gilbert TM, Wolberger C, Tackett AJ & Taverna SD (2014) The bromodomain of Gcn5 regulates site-specificity of lysine acetylation on histone H3. *Mol. Cell Proteomics***:** mcp.M114.038174

Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26:** 1367–1372

Daniel JA & Grant PA (2007) Multi-tasking on chromatin with the SAGA coactivator complexes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **618:** 135–148

de Laat WL, Jaspers NG & Hoeijmakers JH (1999) Molecular mechanism of nucleotide excision repair. *Genes & Development* **13:** 768–785

Dudley AM, Rougeulle C & Winston F (1999) The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo. *Genes & Development* **13:** 2940–2945

Eisenmann DM, Arndt KM, Ricupero SL, Rooney JW & Winston F (1992) SPT3 interacts with TFIID to allow normal transcription in Saccharomyces cerevisiae. *Genes & Development* **6:** 1319–1331

Fermin D, Basrur V, Yocum AK & Nesvizhskii AI (2011) Abacus: A computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis. *Proteomics* **11:** 1340–1345

Ferreiro JA (2006) Roles for Gcn5p and Ada2p in transcription and nucleotide excision repair at the Saccharomyces cerevisiae MET16 gene. *Nucleic Acids Research* **34:** 976–985

Fischer T, Rodríguez-Navarro S, Pereira G, Rácz A, Schiebel E & Hurt E (2004) Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. *Nat Cell Biol* **6:** 840–848

Fuda NJ, Ardehali MB & Lis JT (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461:** 186–192

Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L & Davidson I (2001) Histone Folds Mediate Selective Heterodimerization of Yeast TAFII25 with TFIID Components yTAFII47 and yTAFII65 and with SAGA Component ySPT7. *Molecular and Cellular Biology* **21:** 1841–1853

Gangloff YG, Werten S, Romier C, Carre L, Poch O, Moras D & Davidson I (2000) The Human TFIID Components TAFII135 and TAFII20 and the Yeast SAGA Components ADA1 and TAFII68 Heterodimerize to Form Histone-Like Pairs. *Molecular and Cellular Biology* **20:** 340–351

Geiger JH, Hahn S, Lee S & Sigler PB (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272:** 830–836

Grant PA (1999) Expanded Lysine Acetylation Specificity of Gcn5 in Native Complexes.

Journal of Biological Chemistry **274:** 5895–5900

Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL & Workman JL (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & Development* **11:** 1640–1650

Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR & Workman JL (1998a) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94:** 45–53

Grant PA, Schieltz D, Pray-Grant MG, Yates JR & Workman JL (1998b) The ATM-related cofactor Tra1 is a component of the purified SAGA complex. *Molecular Cell* **2:** 863–867

Hahn S & Young ET (2011) Transcriptional Regulation in Saccharomyces cerevisiae: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. *Genetics* **189:** 705–736

Helmlinger D (2012) New insights into the SAGA complex from studies of the Tra1 subunit in budding and fission yeast. *transcription* **3:** 13–18

Herzog F, Kahraman A, Boehringer D, Mak R, Bracher A, Walzthoeni T, Leitner A, Beck M, Hartl F-U, Ban N, Malmström L & Aebersold R (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **337:** 1348–1352

Horiuchi J, Silverman N, Marcus GA & Guarente L (1995) ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Molecular and Cellular Biology* **15:** 1203–1209

Huisinga KL & Pugh BF (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. *Molecular Cell* **13:** 573–585

Kalisman N, Adams CM & Levitt M (2012) Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *PNAS* **109:** 2884–2889

Kalkhof S & Sinz A (2008) Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Anal Bioanal Chem* **392:** 305–312

Kamata K, Hatanaka A, Goswami G, Shinmyozu K, Nakayama J-I, Urano T, Hatashita M, Uchida H & Oki M (2013) C-terminus of the Sgf73 subunit of SAGA and SLIK is important for retention in the larger complex and for heterochromatin boundary function. *Genes Cells* **18:** 823–837

Kamenova I, Warfield L & Hahn S (2014) Mutations on the DNA binding surface of TBP discriminate between yeast TATA and TATA-less gene transcription. *Molecular and Cellular Biology* **34:** 2929–2943

Kim JL, Nikolov DB & Burley SK (1993a) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365:** 520–527

Kim TK & Roeder RG (1994) Involvement of the basic repeat domain of TATA-binding protein (TBP) in transcription by RNA polymerases I, II, and III. *Journal of Biological Chemistry* **269:** 4891–4894

Kim Y, Geiger JH, Hahn S & Sigler PB (1993b) Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365:** 512–520

Kirschner DB, Baur vom E, Thibault C, Sanders SL, Gangloff Y-G, Davidson I, Weil PA & Tora L (2002) Distinct mutations in yeast TAF(II)25 differentially affect the composition of TFIID and SAGA complexes as well as global gene expression patterns. *Molecular and Cellular Biology* **22:** 3178–3193

Knutson BA & Hahn S (2011) Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Molecular and Cellular Biology* **31:** 818–831

Knutson BA, Luo J, Ranish J & Hahn S (2014) Architecture of the Saccharomyces cerevisiae RNA polymerase I Core Factor complex. *Nature Structural & Molecular Biology* **21:** 810–816

Kokubo T, Swanson MJ, Nishikawa JI, Hinnebusch AG & Nakatani Y (1998) The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Molecular and Cellular Biology* **18:** 1003–1012

Koutelou E, Hirsch CL & Dent SY (2010) Multiple faces of the SAGA complex. *Current Opinion in Cell Biology* **22:** 374–382

Köhler A, Schneider M, Cabal GG, Nehrbass U & Hurt E (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* **10:** 707–715

Köhler A, Zimmerman E, Schneider M, Hurt E & Zheng N (2010) Structural Basis for Assembly and Activation of the Heterotetrameric SAGA Histone H2B Deubiquitinase Module. *Cell* **141:** 606–617

Lang G, Bonnet J, Umlauf D, Karmodiya K, Koffler J, Stierle M, Devys D & Tora L (2011) The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. *Molecular and Cellular Biology* **31:** 3734–3744

Laprade L, Rose D & Winston F (2007) Characterization of new Spt3 and TATA-binding

protein mutants of Saccharomyces cerevisiae: Spt3 TBP allele-specific interactions and bypass of Spt8. *Genetics* **177:** 2007–2017

Lee KK, Sardiu ME, Swanson SK, Gilmore JM, Torok M, Grant PA, Florens L, Workman JL & Washburn MP (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Molecular Systems Biology* **7:** 1–12

Lee KK, Swanson SK, Florens L, Washburn MP & Workman JL (2009) Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics Chromatin* **2:** 2

Liu X, Bushnell DA & Kornberg RD (2013) RNA polymerase II transcription: structure and mechanism. *Biochim. Biophys. Acta* **1829:** 2–8

Liu X, Tesfai J, Evrard YA, Dent SYR & Martinez E (2003) c-Myc transformation domain recruits the human STAGA complex and requires TRRAP and GCN5 acetylase activity for transcription activation. *J. Biol. Chem.* **278:** 20405–20412

Liu Y-L, Jiang S-X, Yang Y-M, Xu H, Liu J-L & Wang X-S (2012) USP22 acts as an oncogene by the activation of BMI-1-mediated INK4a/ARF pathway and Akt pathway. *Cell Biochem. Biophys.* **62:** 229–235

Luo J, Fishburn J, Hahn S & Ranish J (2012) An integrated chemical cross-linking and mass spectrometry approach to study protein complex architecture and function. **11:** M111.008318

Mädler S, Bich C, Touboul D & Zenobi R (2009) Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* **44:** 694–706

McCormick MA, Mason AG, Guyenet SJ, Dang W, Garza RM, Ting MK, Moller RM, Berger SL, Kaeberlein M, Pillus L, La Spada AR & Kennedy BK (2014) The SAGA histone deubiquitinase module controls yeast replicative lifespan via Sir2 interaction. *Cell Rep* **8:** 477–486

McMahon SB, Wood MA & Cole MD (2000) The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. *Molecular and Cellular Biology* **20:** 556–562

McMahon SJ, Pray-Grant MG, Schieltz D, Yates JR III & Grant PA (2005) Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *PNAS* **102:** 8478–8482

Mengus G, May M, Jacq X, Staub A, Tora L, Chambon P & Davidson I (1995) Cloning and characterization of hTAFII18, hTAFII20 and hTAFII28: three subunits of the human transcription factor TFIID. *The EMBO Journal* **14:** 1520–1531

Merkley ED, Rysavy S, Kahraman A, Hafen RP, Daggett V & Adkins JN (2014)

Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23:** 747–759

Mohan RD, Abmayr SM & Workman JL (2014) Pulling complexes out of complex diseases: Spinocerebellar Ataxia 7. *rarediseases* **2:** e28859

Mohibullah N & Hahn S (2008) Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. *Genes & Development* **22:** 2994–3006

Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, Azubel M, Elmlund D, Levi-Kalisman Y, Liu X, Gibbons BJ, Levitt M & Kornberg RD (2013) Architecture of an RNA Polymerase II Transcription Pre-Initiation Complex. *Science* **342:** 1238724–1238724

Mühlbacher W, Sainsbury S, Hemann M, Hantsche M, Neyer S, Herzog F & Cramer P (2014) Conserved architecture of the core RNA polymerase II initiation complex. *Nature Communications* **5:** 4310

Müller MQ & Sinz A (2012) Chemical cross-linking and high-resolution mass spectrometry to study protein-drug interactions. *Methods Mol. Biol.* **803:** 205–218

Nikolov DB, Chen H, I-lalay ED & Usheva AA (1995) Crystal structure of a TFIlB-TBP-TATA-element ternary complex. *Nature* **377:** 119–128

Nikolov DB, Hu SH, Lin J, Gasch A, Hoffmann A, Horikoshi M, Chua NH, Roeder RG & Burley SK (1992) Crystal structure of TFIID TATA-box binding protein. *Nature* **360:** 40–46

Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P & Travers AA (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *The EMBO Journal* **19:** 6141–6149

Palhan VB, Chen S, Peng G-H, Tjernberg A, Gamper AM, Fan Y, Chait BT, La Spada AR & Roeder RG (2005) Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 8472–8477

Pascual-García P & Rodríguez-Navarro S (2009) A tale of coupling, Sus1 function in transcription and mRNA export. *RNA Biol* **6:** 141–144

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612

Pray-Grant MG, Daniel JA, Schieltz D, Yates JR & Grant PA (2005) Chd1

chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433:** 434–438

Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates JR & Grant PA (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Molecular and Cellular Biology* **22:** 8774–8786

Pugh BF (2000) Control of gene expression through regulation of the TATA-binding protein. *Gene* **255:** 1–14

Reddy P & Hahn S (1991) Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region. *Cell* **65:** 349–357

Reeves WM & Hahn S (2005) Targets of the Gal4 Transcription Activator in Functional Transcription Complexes. *Molecular and Cellular Biology* **25:** 9092–9102

Rhee HS & Pugh BF (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483:** 295–301

Rodríguez-Navarro S (2009) Insights into SAGA function during gene expression. *EMBO reports* **10:** 843–850

Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234:** 779–815

Samara NL & Wolberger C (2011) A new chapter in the transcription SAGA. *Curr. Opin. Struct. Biol.* **21:** 767–774

Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE & Wolberger C (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **328:** 1025–1029

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W & Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473:** 337–342

Sermwittayawong D & Tan S (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. *The EMBO Journal* **25:** 3791–3800

Shukla A, Lahudkar S, Durairaj G & Bhaumik SR (2012) Sgf29p Facilitates the Recruitment of TATA Box Binding Protein but Does Not Alter SAGA's Global Structural Integrity in Vivo. *Biochemistry* **51:** 706–714

Sinz A (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrom. Rev.* **25:** 663–682

Söding J, Biegert A & Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33:** W244–8

Sterner DE, Belotserkovskaya R & Berger SL (2002) SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc. Natl. Acad. Sci. U.S.A.* **99:** 11622–11627

Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, Winston F, Workman JL & Berger SL (1999) Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Molecular and Cellular Biology* **19:** 86–98

Tan S, Hunziker Y, Sargent DF & Richmond TJ (1996) Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381:** 127–151

Tosi A, Haas C, Herzog F, Gilmozzi A, Berninghausen O, Ungewickell C, Gerhold CB, Lakomek K, Aebersold R, Beckmann R & Hopfner K-P (2013) Structure and Subunit Topology of the INO80 Chromatin Remodeler and Its Nucleosome Complex. *Cell* **154:** 1207–1219

Trievel R, Rojas JR, Sterner DE, Venkataramani RN, Wang L, Zhou J, Allis CD, Berger SL & Marmorstein R (1999) Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. *PNAS* **96:** 8931–8936

Vannini A & Cramer P (2012) Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *MOLCEL* **45:** 439–446

Walzthoeni T, Leitner A, Stengel F & Aebersold R (2013) Mass spectrometry supported determination of protein complex structure. *Curr. Opin. Struct. Biol.* **23:** 252–260

Wang L & YR Dent S (2014) Functions of SAGA in development and disease. *Epigenomics* **6:** 329–339

Warfield L, Ranish JA & Hahn S (2004) Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes & Development* **18:** 1022–1034

Weake VM & Workman JL (2010) Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet* **11:** 426–437

Weake VM & Workman JL (2012) SAGA function in tissue-specific gene expression. *Trends in Cell Biology* **22:** 177–184

Weake VM, Lee KK, Guelman S, Lin C-H, Seidel C, Abmayr SM & Workman JL (2008) SAGA-mediated H2B deubiquitination controls the development of neuronal connectivity in the Drosophila visual system. *The EMBO Journal* **27:** 394–405

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

Wu P-YJ & Winston F (2002) Analysis of Spt7 function in the Saccharomyces cerevisiae SAGA coactivator complex. *Molecular and Cellular Biology*

Wu P-YJ, Ruhlmann C, Winston F & Schultz P (2004) Molecular Architecture of the S. cerevisiae SAGA Complex. *Molecular Cell* **15:** 199–208

Wyce A, Xiao T, Whelan KA, Kosman C, Walter W, Eick D, Hughes TR, Krogan NJ, Strahl BD & Berger SL (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *MOLCEL* **27:** 275–288

Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR & Roth SY (2000) Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat. Genet.* **26:** 229–232

Yang B, Wu Y-J, Zhu M, Fan S-B, Lin J, Zhang K, Li S, Chi H, Li Y-X, Chen H-F, Luo S-K, Ding Y-H, Wang L-H, Hao Z, Xiu L-Y, Chen S, Ye K, He S-M & Dong M-Q (2012) Identification of cross-linked peptides from complex samples. *Nature Methods* **9:** 904–906

Yu Y, Eriksson P, Bhoite LT & Stillman DJ (2003) Regulation of TATA-Binding Protein Binding by the SAGA Complex and the Nhp6 High-Mobility Group Protein. *Molecular and Cellular Biology* **23:** 1910–1921

Yu Y, Teng Y, Liu H, Reed SH & Waters R (2005) UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 8650–8655