DNA-binding independent mechanism of TATA-binding protein at TATA-less genes

Ivanka Kamenova

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2013

Reading Committee:
Steven Hahn, Chair
Stanley Fields
Toshio Tsukiyama

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

DNA-binding independent mechanism of TATA-binding protein at TATA-less genes

Ivanka Kamenova

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

The core promoter is defined as the minimal sequence of DNA surrounding the transcription start site that is required for transcription of mRNA genes by the eukaryotic RNA Polymerase II (Pol II). Multiple core promoter elements exist in higher eukaryotes, and these DNA elements make distinct contributions to the combinatorial control of gene expression. The TATA box is the most ancient and conserved core promoter element, yet it is present in no more than 20% of Pol II promoters. To address the need for *in vitro* models of TATA-less genes, I used the *Saccharomyces cerevisiae* Ribosomal Protein gene promoter *RPS5*, which lacks a TATA box, to develop an *in vitro* transcription system. Using this system to characterize the requirements of yeast TATA-less promoters, I found that the *RPS5* promoter was dependent on the coactivactor complex TFIID in addition to requiring the TATA-binding protein TBP for transcription. The general transcription factors (IIA, TBP/TFIID, IIB, IIF, IIE, IIH) and Pol II were insufficient to drive transcription from the *RPS5* promoter, suggesting that additional proteins are needed for transcription of TATA-less genes. I further examined the requirement for DNA binding by TBP both *in vitro* and *in vivo*, demonstrating that TATA-less promoters can be transcribed by TBP deficient in DNA binding. TBP binds to two distinct sites on the *RPS5* promoter. Mutational analysis of these binding sites revealed that they are dispensable for *in vitro* transcription, while replacement with GC-rich DNA is detrimental. My work has established an *in vitro* system with a TFIID-dependent TATA-less promoter and implicated a more relaxed mechanism of TBP binding on this type of promoter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GENERAL INTRODUCTION

Pre-initiation complex formation at the Pol II core promoter

Transcription of mRNA in eukaryotic cells is an intricately regulated process that is initiated by the binding of sequence-specific activator proteins to unique regulatory DNA elements. Transcription activators orchestrate downstream events leading to mRNA synthesis by recruiting chromatin modifying enzymes, coactivator complexes and ultimately RNA Polymerase II and its accessory factors termed General Transcription Factors (GTFs). The transcription cycle progresses through many steps: chromatin remodeling and nucleosome eviction, preinitiation complex (PIC) formation, promoter melting, promoter escape, productive elongation, and termination (Fuda et al. 2009). All of these events are the subject of tight regulatory control.

The recruitment of Pol II is accomplished by the formation of a stable preinitiation complex (PIC) consisting of Pol II and the general transcription factors (GTFs), TBP, TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH on the core promoter (Hahn et al. 2011). The core promoter is an ~60 base pair DNA sequence surrounding the transcription start site, which is defined as the minimal promoter element required for correct initiation of transcription *in vitro* (Smale et al. 2003). Assembly of the PIC takes place in a sequential manner. The first event in core promoter recognition is the binding of TATA-binding protein (TBP) to the highly conserved TATA box sequence (Hahn et al. 2011, Thomas et al. 2006). The sequence-specific interactions between TBP and the minor groove of the TATA box introduce a sharp 90 degree bend in the promoter DNA, which is critical for the subsequent recognition of this complex by other GTFs (Nikolov et al. 1994, Nikolov et al. 1996). Other factors, namely TFIIA and TFIIB, recognize the platform formed by the TBP-TATA complex and enter the PIC. TFIIB is crucial for the recruitment of Pol

II and the correct positioning of the DNA strands in proximity with the Pol II active center. TFIIF, a heterodimer of Tfg1 and Tfg2, joins the PIC along with Pol II, and has important functions in proper transcription start site determination and transcription elongation. TFIIE, also a heterodimer, consists of the Tfa1 and Tfa2 subunits, and is recruited later (Hahn et al. 2011). Its complete array of functions is not fully understood yet, but TFIIE is thought to play a role in stabilizing the open complex, and it also makes important contacts with TFIIH (Grunberg et al. 2012) and Mediator (Esnault et al. 2008). TFIIH is a multi-subunit complex composed of 3 modules with structural and enzymatic activities (Gibbons et al. 2012). It is the last GTF recruited to the PIC and it plays critical roles in transcription by 1) unwinding the DNA strands surrounding the transcription start site to form the open complex, and 2) phosphorylating the C-terminal tail of the largest Pol II subunit, Rpb1, thus facilitating the transition to promoter escape (Hahn et al. 2011).

TBP-TATA Box Interactions

Crystal structures of TBP from different species bound to the TATA box DNA have revealed the fine molecular details of how this factor interacts with and shapes the promoter DNA (Nikolov et al. 1994, Nikolov et al. 1996, Geiger et al. 1996). TBP is a saddle-shaped molecule with a concave surface lined with hydrophobic residues. This hydrophobic surface interacts with the minor groove of the 8-bp TATA box sequence, altering the shape of the traditional B-form DNA. Importantly, two pairs of conserved phenylalanine residues positioned on either side of the saddle insert within the DNA strands flanking the TATA box and deform it into the 90° bend characteristic of the TBP-TATA complex (Pardo et al. 2000). Crystal structures of TBP with ten naturally existing TATA elements of various transcriptional strengths demonstrate that the TATA-TBP complex adopts the same conformation in all cases examined. While TBP can recognize a variety of TATA sequences that differ from the optimal recognition element, a C or a G in specific positions of the TATA box would result in a steric clash with residues in the TBP concave surface (Patikoglou et al. 1999). The significance of the TATA-TBP interaction in transcription activation is underscored by the results of a screen for TBP mutants deficient in response to activators. Surprisingly, the majority of TBP mutations mapped to the DNA binding surface and interfered with its ability to recognize the TATA box (Lee et al. 1995).

Diversity of Core Promoter Elements

Aside from the TATA box, the core promoter in metazoan cells contains a number of other elements that are involved in directing PIC assembly (Smale et al. 2003, Juven-Gershon et al. 2008b, Juven-Gershon et al. 2006b, Butler et al. 2002). The first indication for this came from the observation that the TBP-Taf complex TFIID gives an extended DNase I footprint downstream from the transcription start site of certain promoters, suggesting that there are elements besides the TATA box that can interact with the transcription machinery in a sequence-specific manner (Sypes et al. 1994, Nakatani et al. 1990). Further investigation led to the discovery of the Inr (Initiator), DPE (Downstream Promoter Element), DCE (Downstream Core Element), and the MTE (Motif 10 element), all of which are conserved from Drosophila to humans and directly interact with various TFIID subunits (Burke et al. 1996, Burke et al. 1997, Burke et al. 1998, Kadonaga 2002, Kutach et al. 2000, Lim et al. 2004). Additionally, the BRE elements, located upstream and downstream of the TATA box (independently or together) are recognized by the general transcription factor TFIIB (Deng et al. 2005, Deng et al. 2006, Deng et al. 2007). These and other studies have shown that despite the evolutionary conservation of the

TATA box, it is utilized by a limited number of promoters (approximately 20%), and other elements are equally important for PIC assembly (Basehoar et al. 2004, Yang et al. 2007).

Core promoters are important regulators of transcription

There is ample evidence that the core promoter is more than a static platform to which the transcription machinery is recruited; it can contribute to the combinatorial regulation of gene expression. One of the first indications for this comes from a screen in Drosophila for enhancer elements that exhibit specificity toward either a TATA- or a DPE-containing promoter. The results indicated that while the majority of regulatory elements can work in either promoter context, some enhancers can only activate DPE-containing promoters or TATA-containing promoters (Butler et al. 2001). A different example using the murine terminal deoxynucleotityltransferase (TdT) gene, which is essential for early lymphocyte development, demonstrates that its Inr-containing promoter is strengthened by the insertion of a canonical TATA element. However, the inserted TATA element is unable to substitute for a nonfunctional Inr, indicating that these two elements can work together but are not functionally redundant (Martinez et al. 1995). Another study of the Drosophila Hox cluster, which is a key determinant of the segment identity of the embryo, provides evidence for promoter selectivity by the ANT-C enhancer, which preferentially activates a TATA-containing promoter over an equidistant DPE-containing gene. The insertion of a TATA element in the DPE promoter resulted in its inappropriate activation by the enhancer, thus showing that core promoter elements are important players in developmental gene regulation (Ohtsuki et al. 1998). Finally, it was shown that Caudal, a master regulator of the predominantly DPE-containing Hox genes, has a high specificity for DPE versus TATA-containing promoters. Caudal’s inability to activate TATA-containing promoters is determined by the presence of BRE elements flanking the TATA motif (Juven-Gershon et al. 2008a). It is thus likely that diverse core promoter elements are utilized differentially during complex developmental pathways to fine-tune gene regulation.

Targeting of TBP to the core promoter by SAGA and TFIID

In addition to its association with the TATA box, TBP is targeted to promoters via its interaction with coactivator complexes. Coactivators are protein complexes that bridge the interaction between activators and the basal transcription machinery, and they are essential for gene expression *in vivo* and *in vitro*. In the eukaryote *Saccharomyces cerevisiae*, two multi-subunit protein assemblies, TFIID and SAGA, both conserved in higher eukaryotes, are responsible for delivering TBP to the promoter (Hahn et al. 2011). SAGA stands for Spt–Ada–Gcn5–acetyltransferase, referring to its multi-modular composition, consisting of activator binding, enzymatic, Taf-containing, and TBP-binding subdomains (Samara et al. 2011). Importantly, the SAGA subunits involved in TBP interaction are Spt3 and Spt8, although TBP does not associate tightly with SAGA (Mohibullah et al. 2008). Yeast TFIID is composed of 14 Tafs and TBP and the overall three dimensional structure and subunit organization of the complex is conserved from yeast to humans (Hahn et al. 2011). It consists of three major lobes arranged in a horseshoe shape, with TBP nestled in the inner surface formed by the different lobes (Leurent et al. 2004). Recent cryo-EM structures of human TFIID have revealed the structural association between the Tafs which form the core of TFIID (Bieniossek et al. 2013), as well as the flexibility of the lobes and the interaction of TFIID with promoter DNA (Cianfrocco et al. 2013).

While both TFIID and SAGA interact with TBP and they share a number of common Taf subunits, their roles in transcription are not interchangeable. Individual depletion of Tafs (Moqtaderi et al. 1996), followed by later genome-wide assessment of the effects of Taf depletion and bioinformatics studies (Shen et al. 2003) have demonstrated that yeast genes have a preference for TFIID or SAGA: many genes can use either TFIID or SAGA, some can use both, and a class of genes can use exclusively TFIID or SAGA (Huisinga et al. 2004, Lee et al. 2000). Genes requiring SAGA function are usually classified as strong, inducible promoters, rapidly activated during conditions of stress. On the other end of the spectrum, a large number of yeast genes with steady expression during exponential growth are TFIID-dependent (Huisinga et al. 2004). A similar classification has been made for human genes (Yang et al. 2007)

TFIID is a coactivator and a core promoter recognition factor

TFIID contains the TATA box-binding subunit TBP and a number of Tafs which bind to other core promoter elements (Hahn et al. 2011, Thomas et al. 2006). Studies with Drosophila and human TFIID have shown that the binding of Tafs to the Initiator and DPE can occur independently of binding of TBP to the TATA box, and that all three interactions can synergize to result in higher than expected levels of transcription when all three elements are combined in one promoter (Cianfrocco et al. 2013, Parry et al. 2010, Juven-Gershon et al. 2006a). Specific TFIID-recognition motifs have not yet been defined in yeast (Smale et al. 2003), although this does not exclude their potential existence. Complementing the *in vitro* studies defining TFIID-core promoter interactions, a recent EM structure showing the path of promoter DNA along human TFIID has solidified the view of TFIID as a complex with many core promoter recognition modules (Cianfrocco et al. 2013).

In addition to the promoter-binding role of TFIID, its function as a coactivator has been well documented in studies with different organisms. *In vitro* studies in mammalian systems have demonstrated the requirement for TFIID during attempts to obtain activated transcription using purified GTFs (Dynlacht et al. 1991). Further support for the notion that TFIID interacts with activators comes from EM structures of yeast and human TFIID in complex with different activator proteins (Papai et al. 2010, Liu et al. 2009). However, the details of how activators interact with and bring TFIID to the promoter are not well understood.

Finally, TFIID makes important contacts with chromatin via the interactions of some of its subunits with modified histone tails. The human Taf3 has a Phd finger domain, which binds to trimethylated H3K4, and this association is important for recruitment of TFIID during transcription activation of specific promoters (Lauberth et al. 2013). Along the same lines, human Taf1 has a double bromodomain capable of recognizing acetylated lysines (Jacobson et al. 2000). These two chromatin binding domains are absent from yeast TFIID (van Ingen et al. 2008, Vermeulen et al. 2007). However, the TFIID-associated protein Bdf1 is homologous to the double bromodomain of higher eukaryote Taf1 and binds to acetylated nucleosomes, thus offering a potential link between TFIID and chromatin (Matangkasombut et al. 2000). A recent detailed map of yeast transcription factor occupancies places Taf1 and Bdf1 in overlapping regions of the promoter, thus strengthening this claim (Rhee et al. 2012).

The TATA box is not a feature of most core promoter elements

Despite the prevalence of the TATA box as a feature of most promoters studied in isolation, several high throughput sequencing and bioinformatics studies in yeast and higher eukaryotes have shown that it is not a widely used DNA element. Extensive mapping of transcription start sites and promoter elements in humans and mice revealed that the majority of promoters of genes with constitutive levels of expression lack a well-defined transcription start site or a TATA box (Yang et al. 2007, Carninci et al. 2006). The class of promoters which initiate from a single well-defined nucleotide is much smaller and is enriched in highly expressed tissue-specific genes regulated by a TATA box (Carninci et al. 2006). Studies in yeast have uncovered a similar trend, with only 10-20 % of promoters classified as TATA-containing. Interestingly, these promoters are also characterized by high levels of expression induced by stress and a dependence on the SAGA coactivator. On the contrary, the majority of genes fall into the TATA-less category and overlap extensively with the class of genes that are TFIID-dependent and have steady levels of expression (Basehoar et al. 2004). A more recent study of general transcription factor binding in yeast has questioned the validity of this classification by determining that in nearly all yeast genes, the location of TBP with reference to the transcription start site is identical and the sequence utilized has one or two nucleotide differences with the consensus TATA element (Rhee et al. 2012). However, these results do not speculate whether the so called “TATA-like” elements found in TATA-less genes function equivalently to a classical TATA box.

Importance of the UAS for Specifying Coactivator Preferences

There are different views regarding the specification of coactivator preferences of yeast promoters. One possibility is that the upstream activating sequence (UAS), where transcription factors bind, is dominant because the activators that bind to it favor the recruitment of one coactivator complex over another. Support for this model comes from both *in vivo* and *in vitro* studies showing that the yeast acidic activators Gal4 and Gcn4 preferentially interact with SAGA

(Bhaumik et al. 2001, Bhaumik et al. 2004, Li et al. 2002, Fishburn et al. 2005, Reeves et al. 2005). In yeast, the ribosomal protein gene (RPG) cluster has served as a useful system for studying coactivator requirements. These genes are highly expressed in exponentially growing cells and they are coordinately regulated by a Rap1-containing activator (Mencia et al. 2002). The majority of RPGs are TATA-less and depend on TFIID for expression (Basehoar et al. 2004). It has been proposed that the UAS, and the Rap1 binding sites in particular, are essential for the recruitment of TFIID to RPG promoters. Several studies showed that the fusion of a minimal element from the *RPS8A* UAS containing just the Rap1 sites upstream of a TAF-independent core promoter is sufficient to confer TFIID-dependence (Mencia et al. 2002, Garbett et al. 2007). Further evidence for the activator-dependent recruitment of TFIID comes from *in vitro* experiments demonstrating the direct interaction between Rap1 and several TFIID-specific Tafs (Garbett et al. 2007, Papai et al. 2010, Layer et al. 2010).The converse has also been demonstrated by one group who reported that upstream sequences from Taf-independent genes such as *ADH1* and *GAL1* cannot recruit Tafs when fused to the TFIID-dependent *RPS5* core promoter, resulting in failure to transcribe the chimeric construct (Li et al. 2002). Finally, several groups have demonstrated that Tafs are recruited to RPG promoters *in vivo* in the absence of functional TBP, suggesting that the interaction between activators and Tafs is direct (Mencia et al. 2002, Shen et al. 2003).

Role of the Core Promoter in Determining TFIID or SAGA Dependence

However, a different interpretation of the TBP-independent recruitment of Tafs is that they can interact directly with core promoter elements other than the TATA box, thus allowing for their independent recruitment. Evidence that the core promoter determines coactivator preferences in yeast comes from studies performed in the Ptashne and Green labs. Cheng and coworkers used *in vivo* reporter assays to demonstrate that fusions of a Taf-independent UAS and the TFIID-dependent *RPS5* core promoter are TFIID-dependent (Cheng et al. 2002). Shen and Green came to the same conclusion (Shen et al. 1997). One group found that insertion of a canonical TATA box in the *RPS5* promoter did not alter coactivator requirements when the endogenous UAS was used, but that a combination of a Taf-independent UAS and a canonical TATA box converted the *RPS5* core promoter to TFIID-independence. Most likely, both the UAS and the core promoter, including sequences other than the TATA box are important for specifying the optimal coactivator.

Structure models of the transcription preinitiation complex gleaned from TATA-containing promoters

Biochemical and structural studies of PICs formed on TATA-containing promoter DNA, as well as structural studies of GTFs, have generated exquisite models of the PIC structure and the path of promoter DNA (Miller et al. 2006, Grunberg et al. 2012, Sainsbury et al. 2013, Lariviere et al. 2012, Kostrewa et al. 2009, He et al. 2013). The positions of the GTFs TBP, TFIIB, TFIIF, and TFIIE relative to Pol II have been revealed for the yeast PIC assembled on *HIS4* DNA. Some of the most recent findings are that TFIIE spans the Pol II cleft and its interaction with the Ssl2 subunit of TFIIH is important for the positioning of the TFIIH complex close to the transcription start site in the PIC (Grunberg et al. 2012). Combined with EM structures outlining the organization of yeast TFIIH (Gibbons et al. 2012), there results pose intriguing questions about the placement and orientation of the entire TFIIH complex within the PIC. Other key recent discoveries include the crystallization of the mediator head module

(Robinson et al. 2012, Lariviere et al. 2012) and the creation of models for the interaction of this complex with Pol II and the rest of the PIC (Cai et al. 2012, Bernecky et al. 2011). These studies have opened the field for probing the interactions of GTFs with Mediator and ultimately creating a PIC model that includes the structures of Mediator and TFIIH.

Description of Dissertation

The focus of this work is predominantly on the study of yeast TATA-less promoters. In chapter one, I describe my initial attempts to use different yeast TATA-less promoters to develop an *in vitro* transcription system. I narrow down my efforts to the core sequence of the Ribosomal Protein gene *RPS5* and show that it can be used successfully to isolate active PICs with transcription start sites corresponding to the *in vivo* mapped initiating sequences. I further demonstrate that this promoter requires TFIID for transcription *in vitro* and it also needs TBP, despite lacking a TATA box. I show that the requirements for active transcription from this promoter are different from those of the TATA-containing *HIS4* promoter, as *RPS5* requires more than the general transcription factors. The *in vitro* assembled immobilized transcriptional templates formed with *RPS5* also have unique requirements and the search for these additional proteins is still ongoing. I also describe results suggesting that the position of the N-terminal portion of TFIIB might be identical between the *RPS5* and *HIS4* PICs.

Chapter two explores the position and function of TBP on the TATA-less promoter *RPS5*, taking advantage of the developed *in vitro* transcription system as well as techniques for *in vivo* analysis of gene expression. I describe the phenotypes of a panel of TBP mutants along its hydrophobic DNA-binding surface. I found that the majority of these mutations were able to support viability of yeast, but show a range of sensitivities to the drug SM (sulfometuron methyl), implicating a defect in transcription of genes activated by Gcn4. SM is a drug often used to mimic amino acid starvation and ability to grow in the presence of SM requires a functional Gcn4 pathway. I explore the observed SM phenotypes further by introducing the DNA binding mutants into a cell that has a temperature sensitive copy of TBP. Upon heat shock, many of the TBP mutants are able to rescue transcription defects specifically on the TATA-less Ribosomal Protein genes, but not on the TATA-containing Gcn4-dependent genes. These results correlate with the SM sensitivity phenotypes of the TBP mutants, with SM-sensitive mutants generally failing to activate transcription of Gcn4-dependent genes.

I further characterize the TBP-TATA-less promoter DNA interactions by demonstrating that TBP can bind to *RPS5* and possibly other TATA-less promoters. Mapping the TBP binding sites on the *RPS5* promoter revealed two distinct sites. Surprisingly, deletion of both of these sequences had no effect on *in vitro* transcription from the *RPS5* promoter, while mutation of both sites with GC-rich DNA lead to a measurable decrease in transcription.

Chapter three differs from the previous two chapters, as it looks at PIC interactions formed on the TATA-containing *HIS4* promoter. My goal was to find proteins that are in close proximity to the Rpb5 subunit of Pol II, testing the hypothesis that the lower jaw of Pol II, close to the transcription start site, is near the TFIIH complex. I created amber mutations covering the surface of Rpb5 and used the amber codon suppression system (Chen et al. 2007) to incorporate the photocrosslinker Bpa at each of these positions. Crosslinking experiments revealed two positions in the N-terminal domain of Rpb5, Q101 and K122, to crosslink to multiple polypeptides in the PIC. I also showed that Rpb5 Q101-bpa crosslinks to the small subunit of TFIIE in a PIC formed with purified factors and I demonstrated that Q101 crosslinks to multiple polypeptides *in vivo*. I developed a method to purify Rpb5-Q101 from *in vivo* crosslinked cells and from PICs. While multiple tests failed to detect crosslinking of Rpb5 to TFIIH subunits, studies are ongoing to identify the crosslinked proteins by mass spectrometry.

CHAPTER 1: *IN VITRO* CHARACTERIZATION OF TATA-LESS PROMOTERS

INTRODUCTION

Transcription from most Pol II promoters rarely involves the use of a classical TATA box or other well-defined TFIID binding sites (Basehoar et al. 2004, Carninci et al. 2006, Yang et al. 2007), yet *in vitro* transcription systems tend represent this relatively small fraction of all promoters. The most convincing examples of *in vitro* transcription from TATA-less promoters come from work in Drosophila, where promoters with different elements known to bind TFIID, such as the Initiator (Inr), Downstream Promoter Element (DPE), and an element which does not bind TFIID such as the TCT element, have been used successfully in cell-free transcription systems (Parry et al. 2010). Some human promoters containing an Inr (Martinez et al. 1995), or a more recently described element, the XCPE2 (Anish et al. 2009), have also been demonstrated to work *in vitro*. However, in the yeast *Saccharomyces cerevisiae*, where the discovery and characterization of core promoter elements other than the TATA box have proven elusive, no *in vitro* transcription system based on a TATA-less promoter has been developed yet. This is a problem that needs to be addressed, because while yeast is a useful model organism for studying eukaryotic transcription, most of the current structural and biochemical data has been generated from TATA-containing promoters, which represent a small class of all yeast genes (20%) (Basehoar et al. 2004).

The ribosomal protein (RP) gene cluster in yeast is the group of genes which ranks highest in a genome-wide classification of TATA-less promoters (Basehoar et al. 2004). These genes are transcribed at very high levels in exponentially growing cells and their levels of expression are synchronously regulated by the upstream activity of the TOR nutrient sensor pathway (Klein et al. 1994). RP gene transcription is rapidly down regulated as a result of different types of stress. Most RP promoters have upstream binding sites for the transcriptional activator-repressor Rap1, and Rap1 binding has been shown to be essential, but insufficient for activation (Wade et al. 2004, Zhao et al. 2006). Two other transcription factors, Fhl1 and Ifh1 are recruited to RP gene upstream sequences in a Rap1-dependent manner and function to activate transcription from these promoters (Wade et al. 2004, Schawalder et al. 2004, Zhao et al. 2006). In addition to Rap1, Fhl1, and Ifh1, a number of other factors have demonstrated roles in RP gene transcription (Hall et al. 2006, Marion et al. 2004, Jorgensen et al. 2004) and a consensus for the minimal number of activators needed for transcription from this major class of genes is yet to be reached.

An important feature of RP genes is that they are among the most dependent on the general transcription factor TFIID (Huisinga et al. 2004, Shen et al. 2003, Lee et al. 2000). Work *in vivo* and *in vitro* has demonstrated that Rap1 binding sites are critical for the recruitment of TFIID to these promoters and that Rap1 directly binds to several TFIID subunits (Mencia et al. 2002, Papai et al. 2010, Layer et al. 2010, Garbett et al. 2007). While an *in vitro* transcription system using an RP upstream activating sequence (UAS) containing Rap1 binding sites has been successfully implemented (Garbett et al. 2007), an *in vitro* transcription system using a TATA-less core promoter has not been developed in yeast. In this chapter, I describe my attempts to use a number of yeast RP and other TATA-less promoters to establish such a system and I show that the *RPS5* core promoter can recapitulate many of the features of TATA-less RP promoters observed *in vivo*. I further demonstrate that the requirements for TATA-less transcription differ significantly from those of TATA-containing promoters.

MATERIALS AND METHODS

Transcriptionally-competent whole cell extract preparation

Twelve liters of cells were grown in YPAD (1% w/v Difco Yeast Extract, 2% w/v Difco Peptone, and 3% w/v Dextrose) supplemented with 0.004% w/v Adenine if needed to an OD_{600} of ~2-3. After harvesting by centrifugation, cells were washed once with distilled water and the combined pellet was transferred to 50 ml conical tubes, frozen in liquid nitrogen and kept at -80°C until further processing. Cell pellets were thawed in a room temperature water bath, transferred to ice and resuspended in ~1ml/1g cell pellet cold lysis buffer supplemented with DTT and protease and phosphatase inhibitors. The lysis buffer composition was as follows: 200 mM Tris-acetate, pH 7.9, 390 mM ammonium sulfate, 20% glycerol, 1 mM EDTA. 1mM DTT, and inhibitors (0.1 mM PMSF, 2 mM Benzamidine, 2.5 µg/ml leupeptin, 1.5 µg/ml pepstatin, 10 µg/ml TPCK, 10 µg/ml TLCK, 0.2 mM NaF, 0.2 mM β-glycerophosphate, 0.2 mM activated sodium orthovanadate, 300nM Trichostatin A) were added immediately before use. The resuspended cell pellets were placed in the chilled chamber of a bead-beater (BioSpec products) filled half-way with chilled 0.5 mm glass beads (approximately 175 ml beads per 350 ml chamber) and the slurry was topped off with additional buffer to minimize air bubbles. Cells were lysed with twelve 30 second pulses and 2 min breaks and kept cold throughout the procedure. Crude extracts were collected and additional proteins extracted by washing the beads with 50 ml cold lysis buffer. The combined crude extract was clarified by centrifugation (30 min 16,000×*g*) to remove cell debris, then spun at 4°C for 3 hours at 200,000×*g* in a Beckman 50.2 Ti rotor. The clarified extract was transferred to a glass beaker, solid ammonium sulfate ground to a fine powder was added to 0.337 g/ml and the mixture was stirred 30 min at 4°C. The precipitated proteins were harvested by ultracentrifugation (15 min at 200,000×*g* in a Beckman

50.2Ti rotor at 4°C) and resuspended slowly in a minimal amount of cold dialysis buffer (20mM Hepes, 7.9, 20% glycerol, 10 mM magnesium sulfate, 10 mM EGTA, 5 mM DTT, inhibitors), approximately 1 ml/L culture. The whole cell extract was dialyzed against three 1L changes of dialysis buffer until the conductivity of the extract diluted at 1:200 was less than 100 µS/cm. The resulting whole cell extract was clarified at 27,000×*g* and stored at -80°C. For the preparation of TFIID-depleted whole cell extracts, yeast strain Shy343 (Taf11ts) and the isogenic wild type strain Shy 342 were grown at 25°C until reaching the correct OD. Upon harvesting, the cells were resuspended in the equivalent amount of 37°C YPAD media and incubated in a 37°C shaker for 1 hour. The remainder of the whole cell extract procedure was unmodified.

Immobilized template assay

DNA templates were generated by PCR from pNM2 or pSH515 (containing the *RPS5* core promoter and *HIS4* core promoters respectively, cloned between the XhoI and BamHI restriction sites of pSH515) with primers pBio965 (biotin-TAATGCAGCTGGCACGACAGG) and pNOT (GGCCGCTCTAGCTGCATTAATG). The 651 and 629-bp products were used to generate immobilized templates as described previously by (Ranish et al. 1999) For transcription from the immobilized *RPS5* promoter, the protocol was modified slightly: PICs were formed using 240 µg whole cell extract supplemented with 30 – 120 µg nuclear extract per 2x PIC (50 µl reaction). The rest of the protocol was identical for *HIS4* and *RPS5*.

TFIID purification via TAP tag

TFIID complex was purified from yeast strain Shy626 expressing a TAP-tagged Taf13. 12 L of culture was grown in YPD (3% Dextrose) to OD600 ~ 2.0 and harvested by centrifugation. The cell pellet was washed in 200 ml cold TAP extraction buffer (40 mM HEPES 7.5, 10% glycerol, 150 mM NaCl, 0.1% Tween 20) without protease inhibitors or DTT, and resuspended in 150 ml cold TAP extraction buffer with DTT and protease inhibitors. Cells were lysed in a bead beater (BioSpec Products) as described for WCE preparation and the supernatant clarified by centrifugation in FL14 rotor (~ 15,000 x g 30 min at 4°C). The clarified supernatant was subjected to ultracentrifugation in a fixed angle 50.2 Ti rotor (Beckman) at 150,000 x g for 90 min at 4°C. The supernatant was either quick frozen in liquid nitrogen and stored at -70°C until further processing, or directly used for IgG purification. 2.5 ml IgG Sepharose beads (GE Healthcare) were washed 3 times with 14 ml TAP extraction buffer and resuspended in TAP extraction buffer to a total volume of 8 ml. The IgG bead slurry was added to the whole cell extract and incubated on a roller at 4°C overnight. The beads were collected by centrifugation and the volume was brought down to ~10 ml by dumping the flow through whole cell extract. The slurry was transferred at 4°C to a 5/8” diameter by 5” length econocolumn connected to PE160 tubing attached to the bottom of the column. After draining the depleted extract, the IgG beads were washed with 5 x 10 ml TAP extraction buffer, followed by 4 x 10 ml TEV cleavage buffer (10 mM Tris pH 8, 150 mM NaCl, 0.1% NP-40, 0.5 mM EDTA, 10% glycerol). The IgG beads were resuspended in 4 ml TEV cleavage buffer and transferred to a 15 ml screwcap tube. 25 µg TEV protease were added and the slurry was incubated overnight on a roller at 4°C. 1 ml Calmodulin agarose beads (Stratagene) were equilibrated by washing 3 times with 14 ml Calmodulin binding buffer (15 mM HEPES pH8, 1 mM MgOAc, 1 mM Imidazole (high purity), 2 mM $CaCl_2$, 0.1% NP-40, 10% glycerol, 150 mM NaCl). The IgG resin suspension cleaved overnight was transferred to the econocolumn described above and the supernatant collected in a screwcap tube. The column was washed with 3 column volumes (~ 7.5 ml total) of calmodulin binding buffer and the wash was combined with the supernatant obtained from TEV cleavage. 8.6 µl 1M calcium chloride was added to the combined flow through and wash fractions and the pooled TEV eluate was added to 1 ml equilibrated calmodulin agarose beads. After a 4 hour incubation at 4°C (or overnight), the beads were washed twice with 14 ml calmodulin binding buffer, centrifuging each time to pellet the beads and incubating each wash ~5 min at 4°C. The beads were then washed 3 times with 14 ml each calmodulin wash buffer and resuspended in 3 ml calmodulin wash buffer (identical to calmodulin binding buffer, but detergent and protease inhibitors can be omitted if needed). The resultant slurry was transferred to a 10 ml disposable column (BioRad) with the flow rate adjusted to ~0.5 ml/min. The bound TFIID complex was eluted by sequential addition of 11 0.5 ml fractions of room temperature calmodulin elution buffer (15 mM HEPES pH8, 1 mM MgOAc, 1 mM Imidazole (high purity), 2 mM EGTA, 10% glycerol, 150 mM NaCl). The elution procedure was carried out at room temperature. Fractions were analyzed by silver staining and the desired peak fractions pooled. Unless otherwise specified, DTT (1 mM final concentration) and protease inhibitors (same as used in nuclear extract preparation) were added to all buffers immediately before use.

TFIID purification by Flag affinity resin

TFIID was purified from yeast strain Shy763 with a 3 x Flag epitope on Taf1 and Taf3. 4 L of culture grown to OD600 3-5 were used and the method for “Large-scale minichromosome preparation for MS analysis” described in (Akiyoshi et al. 2009) was followed. TFIID was eluted from the Flag Dynabeads competitively in 3 elutions (15 min at room temperature with nutation) using 150 µg/ml 3 x Flag peptide (Sigma) diluted in BH150 buffer.

Crude TFIID purification by Flag-agarose affinity resin

~ 500 µl of whole cell extract prepared from yeast strain Shy763 expressing 3 x Flag-tagged Taf1 and Taf3 was dialyzed for 1 hour at 4°C into WCE dialysis buffer (described in WCE preparation) supplemented with protease inhibitors to remove DTT. 120 µl of Flag M2 agarose bead slurry (Sigma) equilibrated in dialysis buffer were added and the extracts incubated for 1.5 hours at 4°C. The depleted WCE was added to a fresh batch of M2 beads (120 µl) and the 1.5 hour incubation was repeated. The beads from the two consecutive depletions were combined and washed twice in 1 ml dialysis buffer supplemented with protease inhibitors. Bound TFIID was eluted competitively by incubation with 120 µl 3 x Flag peptide (Sigma) at 0.25 µg/µl for 20 min at room temperature with nutation. The elution was repeated once and the eluates were frozen on dry ice and stored at – 70 °C.

TBP purification

TBP core (Δ aa 2-60) or full length TBP were expressed as N-terminal SUMO-fusion proteins from pIK8 (6 x HIS-SUMO-TBP in pET 21a) and pIK12 (6 x HIS-SUMO- Δ2-60 TBP in pET 21a) in *E. coli* strain BL21(DE3)RIL. 2 L of cells were grown at 37 °C in 2xYT media supplemented with 50 µg/mL Ampicillin until reaching OD600 ~ 0.7. Expression was induced with 1 mM IPTG for 3 hours at 37 °C. Cells were collected by centrifugation, washed in 50 ml lysis buffer (30 mM Tris-HCl, pH 7.5, 500 mM KCl, 40 mM imidazole, 10% glycerol), resuspended in 30 mL lysis buffer, quick frozen in liquid nitrogen and stored at -70°C until further use. DTT and PMSF were added to the thawed pellet to a final concentration of 1 mM each and cells were lysed by sonication. The resulting lysates were clarified by centrifugation and incubated with 2 ml equilibrated nickel-sepharose beads (GE Healthcare) for 1 hour at 4°C.

After collection, the beads were washed 4 times with lysis buffer (10 minutes at 4°C per wash) and the bound protein eluted in 4 x 4 ml fractions (elution buffer: 30 mM Tris 7.5, 500 mM KCl, 500 mM imidazole, 10% glycerol). After pooling the appropriate fractions, the protein was dialyzed into SUMO cleavage buffer (30 mM Tris-HCl 7.5, 300 mM KCl, 10% glycerol) over the course of 3 hours at 4°C with 3 x 1L buffer changes. The dialyzed protein was spun to remove precipitated material, recombinant SUMO protease was added to a final concentration of ~2.2 µg/ml, and digestion was carried out for 3 hours at 4 °C. The SUMO tag and protease were removed by incubating the cleaved protein with nickel-Sepharose beads (reserved from nickel-Sepharose purification step) for 1 hour at 4°C. The flow through and a 4 ml wash in lysis buffer were pooled. The salt concentration of the digested protein sample was reduced to 100 mM and the protein was purified on a Source 15S column equilibrated in 5% Buffer B (20 mM Tris-HCl pH 7.8, 100 mM KCl, 10% glycerol) over a gradient from 100 mM to 500 mM KCl. Fractions were analyzed by Coomassie staining, pooled and concentrated in Amicon Ultra or YM-10 concentrators with 10K molecular weight cutoff (Millipore). 1 mM DTT was added to all buffers immediately before use.

In vitro transcription and primer extension

In vitro transcription reactions were performed in a 25 µl reaction containing 20 mM HEPES (pH 7.6), 100 mM potassium acetate, 1 mM EDTA, 5 mM magnesium acetate, 3 mM dithiothreitol (DTT), 1 µl nucleoside triphosphates (NTPs) (10 mM each) 38 mM creatine phosphate, 0.03 units creatine phosphokinase, and 4 units RNase Out (Invitrogen). One microliter of supercoiled plasmid DNA template (150 ng for pSH515, 300 ng for all TATA-less promoter constructs) was added to each reaction mixture, along with 24 ng recombinant Gal4-

VP16 activator, and the activator and DNA were allowed to interact for 10 min at room temperatures. Transcription was initiated by addition of 150 -250 µg whole cell extract or purified factors (68 ng TBP, 26 ng TFIIB, 54 ng TFIIF, 12 ng TFIIE, 85 ng TFIIH, and 280 ng Pol II) and allowed to proceed for 40 min at room temperature. 180 µl stop mix consisting of 100 mM sodium acetate, 10 mM EDTA, 0.5% SDS, and 17 µg/ml tRNA (Sigma) was added to stop the reaction. Reaction mixtures were phenol-chloroform (2:1) extracted once, and the RNA was ethanol precipitated, washed in 80% ethanol, and dried in a speed vac. The pellets were resuspended in 10 µl primer annealing mix containing 5 mM Tris (pH 8.3), 75 mM potassium chloride, 1mM EDTA, and either 32P-labeled lacI primer (~5×10^5 cpm) or 65 µM 700IR fluorescently labeled lacI (Li-Cor Biosciences). Reaction mixtures were incubated for 45 min at 48°C (32P-labeled lacI) or 55°C (700IR-lacI). Next, 20µl cDNA synthesis mix (25 mM Tris pH 8.3, 75 mM potassium chloride, 4.5 mM magnesium chloride, 15 mM DTT, 150 µM deoxynucleoside triphosphates (dNTPs), and 100 units MMLV-RT (Invitrogen)) was added to each reaction mixture, and the mixture was incubated for 30 min at 37°C. Reactions were stopped by ethanol precipitation. The pellets were washed with 80% ethanol, dried, resuspended in 3 µl RNase A (40 µg/ml), and incubated for 3 min at room temperature before the addition of 3 µl formamide loading dye containing bromophenyl blue. Just before electrophoresis, samples were heated for 1 min at 90°C, transferred to ice, and run on denaturing 8% acrylamide gels. Gels were visualized with a PhosphorImager (32P; Molecular Dynamics) or with an Odyssey scanner (700-IR; Li-Cor).

FeBABE cleavage assays

TFIIB-FeBABE derivatives were assembled into PICs by using immobilized template DNAs derived from pSH515 (*HIS4*) and pNM2 (*RPS5*) essentially as described in (Chen et al. 2004). A mixture of whole cell extract and nuclear extract was used in the formation of PICs on the *RPS5* immobilized template. Cleavage assays and analysis of cleavage products were performed as described previously (Eichner et al. 2010).

Sequencing analysis of transcription start sites

In vitro transcription / primer extension was carried out using the pNM2 plasmid template as described in *In vitro transcription and primer extension.* The resulting single-stranded DNA was resolved on a denaturing 8% acrylamide urea gel. A sequencing reaction using the pNM2 plasmid template and 32-P-labelled lac I primer was carried out using a Sequenase 2.0 DNA sequencing kit (USB) following the manufacturer's protocol and run adjacent to the primer extension reactions.

RESULTS

Initial attempts for in vitro transcription with different TATA-less promoters

The biochemical characterization on TATA-less promoters has been hindered by a lack of a model promoter with strong activity *in vitro*. My initial work focused on establishing a robust *in vitro* transcription assay with a TATA-less promoter that could serve as a model for this promoter type. I tested a number of different TATA-less promoters, designated as such by computational methods (Basehoar et al. 2004) (Table 1). The majority are ribosomal protein genes, which are characterized by high levels of expression *in vivo* (Warner 1999), thus making them good potential candidates for strong promoter activity. The upstream regulators necessary for the expression of these genes are poorly understood, which is why we tested two types of promoter constructs: core (-200 bp to ATG translation start site) and UAS+core (-500 bp to ATG translation start site). All the constructs summarized in Table 1 were inserted into a vector backbone with a single Gal4 binding site 37 bp upstream of the point of insertion. This allowed us to assay the activity of the different promoters in the presence or absence of recombinant Gal4-VP16 activator. Of all the promoters tested, only three showed promising *in vitro* transcription activity: *RPS5*, *RPL25*, and *HIS3*.

The *HIS3* promoter has been characterized previously as having a canonical TATA box, which drives transcription from a downstream transcription start site, +13, and a non-canonical TATA-like element which drives transcription independently from an upstream +1 start site (Iyer et al. 1995). While I was able to observe these *in vivo* characterized start sites in an *in vitro* transcription with the *HIS3* promoter, when I mutated the canonical *HIS3* TATA box, activity from both sites was abolished completely (data not shown). My result suggested that the *HIS3* transcription that I observed *in vitro* was driven by a canonical TATA element, thus yielding this promoter inappropriate for studying TATA-less transcription. I chose to focus on the *RPS5* core promoter, because the results obtained with it were the most consistent.

The yeast TATA-less promoter RPS5 functions in vitro

The *RPS5* promoter construct used throughout this work is shown as a schematic in Figure 1A. The majority of the *in vitro* transcription assays were performed with a supercoiled plasmid template containing this promoter and a single upstream Gal4 binding site and assayed by primer extension. Transcription using this promoter and a standard nuclear extract resulted in poor activity (Neeman Mohibullah, Steve Hahn, unpublished results). However, I was able to adapt a published procedure for yeast whole cell extract preparation (Garbett et al. 2007) that resulted in high activity from the *RPS5* promoter when recombinant Gal4-VP16 activator was included (Figure 1B, lanes 2-4). The basal transcription levels from this promoter were very low or undetectable (Figure 1B, lane 1). Addition of α-amanitin inhibited transcription, demonstrating that the observed transcripts were due to Pol II activity.

In order to verify that the observed primer extension products are indeed initiating from the correct transcription start sites, primer extension reactions of *RPS5* plasmid transcription were resolved on a high-resolution sequencing gel and the start sites were mapped to base pair accuracy (Figure 1C). These start sites matched the start sites from this promoter mapped *in vivo* using high throughput sequencing technologies (Zhang et al. 2005). In addition, I carried out *in vivo* primer extension on the *RPS5* promoter and found that the primer extension products from this reaction matched the *in vitro* detected start sites (not shown). There results suggest that *in vitro* transcription from the *RPS5* core promoter models the *in vivo* activity.

Biochemical assays to characterize the structure and assembly mechanism of a PIC on a TATA-less promoter require the isolation of a stable complex. I was able to show that PICs isolated on immobilized *RPS5* promoter DNA retain their activity after washing away unbound proteins (Figure 1D). The immobilized *HIS4* promoter serves as a control for a stable TATA-driven promoter. For either promoter, washing immobilized PICs before initiating transcription resulted in minimal loss of activity (compare lanes 1, 2 to 3 ,4 for *HIS4* and 5,6 to 7, 8 for *RPS5*) and the intensity of the bands suggest comparable levels of transcription from both promoters. Thus, the *RPS5* core promoter can serve as a useful biochemical tool for *in vitro* studies of TATA-less promoter organization.

The RPS5 core promoter is TFIID-dependent and SAGA-independent in vitro

Coactivator requirements for the *RPS5* promoter have been studied extensively *in vivo*, providing convincing evidence that TFIID-specific Tafs are enriched on this promoter and required for transcription, while SAGA inactivation has modest to no effect (Shen et al. 1997, Huisinga et al. 2004). To test whether these observations held through in the *in vitro* transcription system, I used a yeast strain harboring a Taf11 temperature-sensitive mutation (Komarnitsky et al. 1999). Shifting this mutant to the restrictive temperature results in severe defects in the expression of TFIID-dependent genes. Whole cell extracts prepared from the Taf11ts strain grown at 37°C were deficient in transcription from the *RPS5* promoter (Figure 2A, lane 3). Addition of TFIID, highly purified via TAP tag, restored transcription to wild type levels, (Figure 2A, lane4), while addition of purified TBP or SAGA did not (not shown), suggesting that the transcription defect is due to depletion of TFIID-specific Tafs. Interestingly, when a crudely purified TFIID, immunoprecipitated from a whole cell extract via Flag tag, was used to rescue transcription, it resulted in transcription levels much higher than observed in the wild type extract. A mock-purified TFIID from an untagged strain had no effect. This strong stimulation of transcription by a crude TFIID fraction was observed repeatedly with different TFIID preparations, implying that a TFIID-interacting factor may be responsible for enhancing transcription from the *RPS5* promoter.

To test the effect of SAGA depletion on *RPS5* transcription *in vitro*, I prepared a whole cell extract from an *Spt7* deletion strain. Deletion of this SAGA subunit has been shown to disrupt the integrity of the complex (Wu et al. 2002). Therefore, we thought that extracts prepared from an *Spt7*Δ strain would be deficient in SAGA activity. We observed that the Spt7Δ extract was functional on the *RPS5* promoter and that addition of purified SAGA did not promote transcription and was inhibitory at high levels (Figure 2C). It is, however, likely that the inhibition by SAGA is due to high amounts of salt in the buffer, as transcription from *RPS5* appears to be more sensitive to salt concentration. In addition, experiments done at a later time demonstrated that Spt7 deletion does not lead to loss of SAGA activity on naked DNA templates *in vitro* (Yan Han, unpublished). Therefore, this approach to testing the requirement of *RPS5* for SAGA is inconclusive. Deletion of another SAGA subunit and assessing the effects on SAGA depletion in the context of chromatin might be more appropriate.

TBP is required for transcription from RPS5

Because the *RPS5* promoter lacks a well-defined TATA box, it was important to investigate whether TBP is required for transcription. To answer this question, I used a whole cell extract from a TBP mutant strain, I143N, which has been demonstrated previously to be deficient for *in vitro* transcription on the *HIS4* promoter (Ranish et al. 1999). This extract was unable to transcribe the *RPS5* promoter either (Figure 3, lane 2). Supplementing reactions with recombinant purified TBP restored transcription (Figure 3, lanes 3-6), as did the purified TBP core (Figure 3, lanes 7-10), which has a 60 amino acid deletion in the N-terminal portion of the protein. The truncated version of TBP was used, because it had been shown previously to have comparable activity to the full length protein (Zhou et al. 1995) and TBP mutants used in subsequent experiments were found to be more stable in TBP core. These *in vitro* transcription results demonstrate that even though *RPS5* lacks a TATA box, it still requires TBP to initiate transcription.

General transcription factors are insufficient for RPS5 transcription

In vitro transcription from promoters with a canonical TATA box, such as the *HIS4* promoter, can be accomplished by a minimal set of highly purified general transcription factors (TFs): TBP, TFIIB, TFIIF, TFIIE, TFIIH, and Pol II (Fishburn et al. 2012). Therefore, we asked whether this was true for the *RPS5* promoter. Attempts to use purified factors on the *RPS5* promoter demonstrated that the general transcription factors alone are insufficient to generate transcripts from *RPS5* (Figure 4A lanes 3 and 7). Supplementing the purified factors with recombinant TFIIA or purified TFIID did not promote transcription either. Only upon addition of whole cell extract (WCE) and Gal4-VP16 was active transcription achieved. These results imply that (a) critical factor(s) required for *RPS5* transcription is(are) missing from the set of purified factors, or that an important post-translational modification is lost during the recombinant expression.

Additional attempts were made to uncover the requirements for basal transcription from the *RPS5* promoter. When TFIID and TFIIA were added to a TFIID-depleted extract, transcription was restored (Figure 4B top, lanes 2-4). However, a combination of the general transcription factors, TFIID, and another coactivator, Mediator, did not produce transcripts from *RPS5* (Figure 4B top, lane 7). It is interesting to note that the extracts and purified transcription factors tested on *RPS5* show very different activities when used on the *HIS4* promoter. Purified TFIID is active on the *HIS4* promoter, but TBP alone is even more efficient (Figure 4B, bottom, lanes 7 and 8), while TBP alone is not sufficient for transcription from *RPS5*. These results suggest that there is an additional factor other than the general transcription factors, TFIID, and Mediator, which is needed for *RPS5* transcription. The above experiments highlight the fact that TATA-less and TATA-dependent promoters have different transcription requirements.

Characterizing in vitro transcription from RPS5 immobilized promoter DNA

Exploring the requirements for transcription from PICs formed on immobilized *RPS5* promoter DNA provided further evidence that TATA-less promoters have distinct biochemical features. When a *HIS4* immobilized template is used to form PICs with a nuclear extract, these complexes are highly stable to washes and retain their ability to function as active transcription complexes. In contrast, my first attempts to form PICs on the *RPS5* promoter using a whole cell extract, determined to be more active on this promoter, resulted in loss of activity upon washing (Figure 5A). I also found that the addition of actinomycin c1 during the primer extension step of the reaction resulted in lower activity on either promoter, thus discontinuing its use.

To find conditions that supported stable PIC formation on *RPS5*, I tested a variety of factors for the ability to stabilize PICs on this promoter. I found that addition of purified factors or TFIID before washing the PICs had no effect on their stability, and neither did supplementing PICs with those factors after washing (Figure 5B and not shown). This eliminated the possibility that loss during washing steps or insufficient amounts of these proteins are the reasons why *RPS5* PICs are unstable to washes. Surprisingly, adding a small amount of nuclear extract to *RPS5* PICs, in addition to the whole cell extract, resulted in transcription activity that was stable to washes (Figure 5B, lanes 3, 6, and 7). It is also interesting to point out that when a nuclear extract was used to form PICs on *RPS5*, these complexes had no transcription activity unless they were washed, which rendered them somewhat active (Figure 5B, lanes 1 and 2). It is unclear why this effect was seen repeatedly, but one possibility is that the washing steps remove an activity present in nuclear extracts that is inhibitory to *RPS5* transcription. Finally, I observed that including Mediator in the PIC formation reactions had a small, but reproducible positive effect on PIC stability (Figure 5B, lane 9).

Comparing the immobilized template results obtained with *RPS5* to the *HIS4* promoter further supports the claim that these two promoters have different requirements for formation of stable active PICs. Like *RPS5* PICs, complexes formed on *HIS4* with a whole cell extract were not stable. However, in contrast to *RPS5*, addition of whole cell extract to *HIS4* PICs did not enhance their stability (Figure 5C). Another observed difference was that the activity of *HIS4* complexes formed with a nuclear extract was not enhanced by washing the PICs, as was observed with *RPS5*.

I used the immobilized template assay to ask whether the general transcription factors were stably associated with the *RPS5* promoter. Western blots of washed PICs eluted from *RPS5* demonstrate that all factors tested, namely Rpb1 (Pol II), Rpb2 (Pol II), Tfg2 (IIF), Tfa2 (IIE), Kin28 (IIH), TBP, and Toa2 (TFIIA), were associated with *RPS5* (Figure 6). The recruitment of these proteins to the promoter was enhanced by the presence of VP16 activator. Since multiple *in vivo* reports have demonstrated that TFIID is found in higher levels at the *RPS5* promoter than at the promoters of TATA-dependent genes (Mencia et al. 2002, Huisinga et al. 2004), I was hoping to observe this difference *in vitro* as well. However, results were obscured by any background binding to the DNA and by the fact that high levels of cofactors are recruited by the VP16 activator.

With these experiments, I demonstrated that the *RPS5* promoter can be used to isolate active transcription complexes *in vitro* and that there are significant differences in the biochemistry of PICs formed on *HIS4* and *RPS5* that need to be explored.

Position of the TFIIB N-terminal region is maintained in the RPS5 PIC

After developing a method for the isolation of active transcription complexes on the *RPS5* promoter, it became possible to ask questions about the structural organization of the PIC. I used a previously established hydroxyradical cleavage (FeBABE) assay (Chen et al. 2004) to compare the position of TFIIB in the *HIS4* and *RPS5* PIC. Cleavage induced by probes that were attached to unique residues in TFIIB was detected by western blotting of specific PIC components. Residues positioned in the zinc ribbon (R37 and S53) and B-finger (S61 and S67), the two N-terminal regions of TFIIB, cleaved Rpb1 and Rpb2 at identical positions whether the PIC was formed on the *HIS4* or *RPS5* promoter (Figure 7). However, residues in the core domain of TFIIB (M135 and I145) only cleaved Rpb1 and 2 in the *HIS4* PIC ((Chen et al. 2004), Table 2). These results raise the possibility that the interaction between Pol II and the TFIIB N-terminal domains, which is very well conserved among species and other RNA polymerases which use a B-type factor (Chen et al. 2003), has remained unchanged in the *RPS5* PIC. The position of the core domain of TFIIB may still be different, because it is unclear what the path of the promoter DNA is in this TATA-less gene. However, due to unresolved difficulties with the immobilized template assay on the *RPS5* promoter, it is not possible to draw such a conclusion without additional experiments.

CONCLUSIONS

The results presented in this chapter demonstrate that the *RPS5* core promoter can be used successfully to study TATA-less transcription *in vitro*. Transcription observed from this promoter is dependent on TFIID and TBP, while the general transcription factors, TFIID, and Mediator are insufficient for transcription. The *RPS5* promoter can be used to purify stable transcription complexes formed with a combination of nuclear and whole cell extracts. Initial probing of PICs formed on the *RPS5* core promoter using hydroxyradical cleavage probes in the N-terminal regions of TFIIB suggests that at least part of the Pol II-IIB interface is identical to what was observed in TATA-containing promoters.

Table 1 List of TATA-less promoters used for *in vitro* transcription studies

Promoter	*Type*	*Gal4 binding site*	*In vitro transcription*
RPS5	Core	Yes	Yes
	Core+UAS	Yes	No
RPL5	core	Yes	Very weak
RPL9a	Core	Yes	No
	Core+UAS	Yes	No
RPL25	Core	Yes	Yes
RPL37a	Core+UAS	Yes	No
RPL40a	Core	Yes	No
	Core+UAS	Yes	No
RPS11b	Core	Yes	No
	Core+UAS	Yes	No
TUB2	Core	Yes	No
TRP3	Core	Yes	No
HIS3	Core	Yes	Yes (TATA-driven)

Table 2 Summary of cleavage experiments performed with different TFIIB FeBABE derivatives in PICs formed on the *HIS4* and *RPS5* promoters.

Cleavage products were detected in the Rpb1 and Rpb2 subunits of Pol II. NR* indicates an interesting cleavage pattern which was observed, but could not be reproduced. It may be interesting to repeat with an optimized immobilized template on the *RPS5* promoter.

Cleavage detected:	*IIB-FeBABE:*	*HIS4*	*RPS5*
Rpb1 C-terminal	R37C	+	+
	S53C	+	+
	S61C	+	+
	I145C	-	-
Rpb1 N-terminal	R37C	+	+
	S53C	+	+
	S61C	-/+	-/+
	I145C	-	-
Rpb2 N-terminal	R37C	+	+
	S67C	+	-/+
Rpb2 C-terminal	R37C	+	-/+
	S53C	+	-
	D88C	+	-
	D103C	+	NR*
	M135C	+	NR*
	I145C	+	NR*

CHAPTER 2: POSITION AND FUNCTION OF TBP ON THE YEAST TATA-LESS PROMOTER *RPS5*

INTRODUCTION

Some of the major unanswered questions regarding transcription from TATA-less promoters concern the role of TBP. Does TBP bind to TATA-less promoters? If so, where does it bind? Is the binding as important for transcription as it is in TATA-containing promoters? Is TBP even near the DNA of TATA-less promoters? There is some experimental evidence which suggests that the DNA binding role of TBP may not be critical for transcription from TATA-less promoters and that other factors, namely the Taf subunits of TFIID are involved in the initial recognition of TATA-less promoters. Early work in the human system showed that a TBP mutant defective in TATA box binding fails to transcribe a TATA-containing promoter, yet is active on a TATA-less Inr-containing promoter in the context of the TFIID complex (Martinez et al. 1995). In addition, chromatin immunoprecipitation experiments in yeast have made the case that the TATA-less TFIID-dependent RP promoters have a high occupancy ratio of Tafs to TBP, while TATA-containing TFIID-independent genes have a low Taf to TBP ratio (Mencia et al. 2002). These results lead us to hypothesize that the DNA-binding function of TBP is not required for transcription from TATA-less promoters. I tested this idea by generating a panel of mutations in the DNA binding surface of TBP and assessing their ability to function in transcription on TATA-less and TATA-containing promoters. In addition, I made the surprising discovery that while TBP can bind to the TATA-less *RPS5* promoter at two distinct sites *in vitro,* deletion of these binding sites had no effect on transcription.

MATERIALS AND METHODS

Yeast strains

All strains used for viability assays and qPCR analysis are derivatives of the TBP shuffle strain SHY737. Strains were constructed by transformation by lithium acetate followed by plasmid shuffling by 5-FOA selection.

Yeast culture and manipulations

Yeast were grown under standard conditions in rich media (YPD, 3% Dextrose) or Glucose Complete (GC) media with 2% dextrose and lacking the appropriate amino acids. For inductions of Gcn4-dependent genes, cells were grown in GC media lacking isoleucine and valine and induced with 0.5 μg/ml sulfometuron methyl (SM) in dimethyl sulfoxide (DMSO) for one hour. Experiments with TBP^{ts} mutants were carried out by growing cells to OD_{600} ~ 0.8 at 25°C and shifting cells to 37°C for 30 minutes, followed by a one hour SM induction at 37°C. Viability assays were carried out on GC plates lacking isoleucine and valine with or without 3 μg/ml SM. Five-fold serial dilutions were spotted and plates were incubated at 30°C for 2 days before being photographed, with the exception of SM plates which were photographed after 5 days.

Design and generation of mutations on the DNA-binding surface of TBP

Mutations in the DNA-binding surface of TBP were based on previously published reports of conserved residues involved in DNA binding (Kim et al. 1993, Tan et al. 1996, Lee et al. 1995, Reddy et al. 1991). We selected specific residues based on their position along the TBP-TATA interface, which was analyzed using the visualization tools PyMol and Unix 3D. The selected residues were changed to alanine, under the assumption that more drastic changes such as insertion of bulky aromatic residues or radical mutations might disrupt the overall structure of the protein. Mutants were generated by Quickchange mutagenesis in plasmids pIK12 (6xHIS-SUMO-TBPcore (Δ2-60)) for recombinant expression, pIK25 (3 x Flag-TBP in PRS314) for *in vivo* expression, and pSH227 (TBP ORF under control of SP65 promoter) for *in vitro* transcription/translation.

Analysis of mRNA levels by RT-qPCR

The total mRNA extraction and RT-qPCR procedures used are described in (Knutson et al. 2011). Due to decrease in the levels of many genes typically used as normalization controls in TBP mutants ts-1 and ts-2, relative levels were not normalized to a reference gene. RNA amounts used for cDNA synthesis served as a normalization control instead.

Expression and purification of TBP mutants

Recombinant expression and purification of TBP mutants followed the procedure outlined in Chapter 1 *"TBP purification."* All mutant TBP proteins were in the context of TBP core.

In vitro translation of TBP mutants

In vitro transcription and translation of TBP mutants was carried out using 2μg AccI-linearized plasmid template. The procedure outlined in (Hahn Cell 1989) was followed, the only difference being the method of visualization of translated proteins: instead of S35-labelled methionine, western blotting with polyclonal TBP antisera was used.

Electrophoretic Mobility Shift Assays (EMSA)

Binding assays were performed in 20 µl reactions containing binding buffer (final concentration: 20 mM Tris-HCl pH 8.0, 60 mM KCl, 5 mM MgCl2, 4% glycerol), ~6000 cpm double stranded promoter DNA probe (5’ end-labelled by PCR and gel purified using Qiagen gel extraction kit), 0.5% Brij58, 0.3 µg/ml BSA (NEB), 0.125 µg poly[dG-dC], and proteins diluted in protein dilution buffer (20 mM Tris, pH 7.9, 150 mM KCl, 1 mM DTT, 10 % glycerol, 50 µg/ml BSA (NEB)). The proteins were added last and the reactions incubated for 30 min at room temperature before loading to native acrylamide gels (6 % acrylamide; 0.2 mM MgOAc was included in the gel and in the 1xTGOE running buffer). Electrophoresis was carried out for 1 hour at 4°C. Gels were vacuum dried on Whatman paper and analyzed by PhosphorImager.

DNase I footprint assays

DNase I footprint analysis of protein-DNA interactions followed a standard protocol (Brenowitz, Current Protocols in Molecular Biology, 1989) with several modifications. The conditions for protein-DNA binding were identical to the ones described in *Electrophoretic Mobility Shift Assays (EMSA)*. ~20,000 cpm of end-labelled double-stranded DNA probe was used (same probe used for EMSA). After 30 minute incubation of DNA and proteins, 2 µl of DNase I (USB, diluted to the appropriate concentration) were added and the reaction was stopped after incubating exactly one minute at room temperature by the addition of 160 µl stop mix (20 mM EDTA, 31 µg/ml tRNA (Sigma) final concentration). Reactions were ethanol precipitated, the pellets resuspended in formamide loading buffer and loaded on a denaturing acrylamide-urea sequencing gel. Gels were analyzed by PhorphorImager. A range of DNase I concentrations and digestion times were assayed to determine the optimal conditions for the particular protein-DNA complex.

In vitro transcription assays

In vitro transcription assays were performed as described in Chapter 1.

RESULTS

Mutants in the DNA-binding surface of TBP support viability of yeast

One of the important questions regarding TATA-less promoters is what the function of TBP is on these promoters. It is especially interesting to investigate whether the DNA-binding function of TBP is utilized on promoters lacking a distinct binding site for TBP and whether TBP is even in contact with the DNA on these genes. In order to address these questions experimentally, I created a series of mutations along the DNA-binding surface of TBP, which were intended to disrupt its binding to the TATA box, and assessed the effects of these mutations *in vivo* and *in vitro*.

I transformed a panel of eleven TBP DNA-binding mutants (Figure 8A) into a TBP shuffle strain and removed the wild type copy of TBP by 5-FOA selection. Surprisingly the majority of mutants were able to support viability of yeast, with the exception of V71A, T112K, and V161A, which were inviable. The rest of the mutant TBPs displayed a range of growth phenotypes: R105A had growth defects, L205A and F207A had a mild phenotype, and the rest were indistinguishable from the wild type. These result support my hypothesis that the DNA binding function of TBP is not required for the expression of most genes that support the growth of yeast under optimal conditions, as these genes tend to lack a TATA box (Basehoar et al. 2004).

On the contrary, I hypothesized that the ability of TBP to bind tightly to a TATA box sequence will be critical for the expression of genes turned on during conditions of stress. The drug SM is often used to mimic amino acid starvation and induce a cluster of TATA-driven genes regulated by the activator Gcn4 (Knutson et al. 2011). Inability to grow in the presence of SM can be an indicator of failure to express sufficient levels of genes in this pathway. I tested the same panel of TBP mutants for their ability to grow on SM plates and found that most of them were sensitive to the drug: R105A and L205 were inviable, R196 and F207 had severe growth defects, and F116A and S118L had an intermediate phenotype. The inability of most TBP DNA-binding mutants to grow on SM provides indirect support for the hypothesis that the ability of TBP to bind to a TATA sequence is much more important for the expression of genes which are dependent on this DNA element.

TATA-less and TATA-dependent genes require TBP for expression in vivo

It was my goal to compare the extent to which TBP DNA-binding mutants can function on TATA-less and TATA-dependent genes. However, TBP is an essential gene and some of the DNA-binding mutants that I created were inviable. To circumvent this problem and any potential complications arising from indirect effects of TBP mutants with compromised growth, I designed a system in which two copies of TBP are present in the cell: a previously published temperature sensitive TBP (Cormack et al. 1992) and a TBP DNA-binding mutant. By growing cells at the restrictive temperature, TBP^{ts} should cause gene expression levels to decrease dramatically, allowing for the extent to which each TBP DNA-binding mutant rescues transcription to be measured (Figure 9).

I tested the expression levels of many different classes of genes in two TBPts mutants, T111I and S136N, during a time course of growth at the restrictive temperature. Both mutants caused very low levels of expression of both TATA-less and TATA-containing genes after as early as 30 minutes of heat shock treatment (Figure 10). I chose the 90 minute time point for subsequent experiments, since the levels of Ribosomal Protein genes decrease as a result of heat treatment in a wild type TBP context, as well as in TBP mutants. In the wild type, however, levels start to go back up after 30 minutes (Figure 10A and not shown). The levels of *HIS4* were dependent on TBP for expression under basal conditions (Figure 10B), and also for SM-dependent induction (not shown).

Mutants in the DNA-binding surface of TBP preferentially support transcription on TATA-less promoters

I used RT-qPCR to assess the ability of the panel of eleven TBP DNA-binding mutants to rescue transcription of both TATA-less and TATA-containing promoters using the scheme described in Figure 9. Upon treatment with SM, the expression levels of the Gcn4-dependent genes *HIS4* and *SNZ1* were induced five- and ten-fold respectively in wild type cells. A wild type copy of TBP expressed from a plasmid was able to rescue this level of induction fully, while the majority of DNA-binding mutants failed to do so (Figure 11A). Two exceptions were V203A and K218A, which rescued transcription levels to 60% and 50% of the wild type, respectively.

In contrast to the observations made at TATA-containing genes, the expressions levels of all TATA-less genes tested (*RPS5, RPL9a, RPL5*) were rescued by the majority of TBP DNA-binding mutants. Mutants V71A, F116A, S118L, R196A, and L205A all supported mRNA levels approximately 80% of wild type (Figure 11B). T112K and V203A were the only mutants that couldn't rescue transcription to at least 50% of wild type levels. The inability of T112K to rescue transcription on both TATA-less and TATA-containing genes could be due to its low protein expression, while all other mutant TBPs were expressed normally (Figure 12). It is interesting that mutations which caused SM-sensitivity, such as F116A, S118L, R196A, were efficient at rescuing transcription on the RP, but not on the Gcn4-dependent promoters. The qPCR data obtained with TBP DNA-binding mutants provides further support to the idea that the ability of TBP to bind TATA box DNA is less critical for transcription from TATA-less promoters.

Mutants on the DNA-binding surface of TBP function in transcription of the RPS5 promoter in vitro

The *in vivo* analysis of the effects of TBP mutants on transcription is one way to address the question whether TBP's DNA-binding role is required for the transcription of different classes of genes. However it has the disadvantage of potential nonspecific effects. To gain a more complete picture, I measured at the ability of purified TBP DNA binding mutants to function in transcription on the *RPS5* and *HIS4* promoters. Interestingly, the mutant T112K, which was purified as a recombinant protein (Figure 13C), was able to rescue transcription on the *RPS5* promoter with a TBP-deficient whole cell extract. The same TBP mutant was inactive on the *HIS4* promoter (Figure 13A). The results obtained with TBP mutant V161A were surprising, as it was very active on both *RPS5* and *HIS4* (Figure 13A), but unable to support cell grown *in vivo*. These results are, however, consistent with the fact that V161A appears to bind DNA according to a gel shift assay performed with a fragment of the *HIS4* promoter spanning the TATA box (Figure 13B). T112K, on the other hand, did not bind to the TATA box of *HIS4*, thus confirming it as a good choice for a DNA-binding mutant of TBP.

Mutants on the DNA binding surface of TBP fail to bind both the HIS4 and RPS5 promoter sequences in vitro

To supplement the *in vivo* transcription analysis of TBP DNA-binding mutants with information about their ability to bind TATA box DNA, I performed gel shift assays using *in vitro* translated TBP mutants and the *HIS4* promoter probe described above. This method was chosen over recombinant purification, as most of the mutants were insoluble. Ten of the eleven TBP mutants were tested for their ability to bind to the *HIS4* promoter and only the *in vitro* translated wild type TBP and mutant V203A were able to form a complex with DNA (Figure 14). Surprisingly, V161A, which was found to bind DNA when expressed as a recombinant protein, did not bind DNA as an *in vitro* translated full-length TBP. It is possible that some of the mutants' DNA-binding activity may be sensitive to deletion of the N-terminal portion of TBP. The observation that V203A binds to the *HIS4* promoter is notable, as this mutant had the highest level of transcription rescue on *HIS4* and *SNZ1 in vivo*.

Finally, I was interested in testing the ability of the TBP mutants to bind to the *RPS5* promoter, because it was important to show that they fail to interact with this promoter. I performed gel shift assays with an *RPS5* promoter probe and purified TBP DNA-binding mutants. Surprisingly, my results indicate that TBP core is able to bind the *RPS5* promoter, although much less strongly than it binds to *HIS4*, and addition of purified TFIIA shifts the complex further and stabilizes it (Figure 15). Similarly to my observations on the *HIS4* promoter, mutants T112K and F116A failed to bind to *RPS5*, while V161A did. The presence of TFIIA caused a weak shift with the F116A mutant, which may be important for the stabilization of TBP on TATA-less promoters *in vivo*.

TATA-less promoter sequences may bind TBP

It has been shown previously that TBP can bind to a wide variety of TATA sequences with high affinity, and it binds to DNA nonspecifically, but the binding is weaker by a few orders of magnitude (Hahn et al. 1989). However, the ability of TATA-less promoter DNA to bind TBP had not been addressed. Therefore, I tested the ability of different TATA-less promoter sequences to compete with the *HIS4* TATA box for binding to TBP. Using a gel shift assay with a labelled *HIS4* promoter probe, I showed that addition of cold unlabeled *HIS4* or *RPS5* competitor, but not *HIS4* competitor with a mutated TATA box was able to reduce the detectable levels of TBP binding. Importantly, the molar amounts of wild type *HIS4* and *RPS5* probes needed to accomplish this competition were comparable, suggesting that the affinity of TBP core to these sequences might be similar (Figure 16). Experiments with the TATA-less promoters *RPL9a* and *RPL25* demonstrated similar results (Figure 17). These findings collectively lead to the surprising conclusion that TBP is capable of binding to TATA-less promoters as strongly as to TATA-containing DNA.

TBP forms a complex with TFIIA on the RPS5 promoter and binds to RPS5 at two distinct sites

To extend the findings that wild type TBP can bind to TATA-less promoter DNA, I characterized its binding sites on the *RPS5* promoter. Side-by-side gel shift assay comparing the binding of TBP to *RPS5* and *HIS4* promoters revealed that TBP core, but not full length TBP, binds to *RPS5* in the absence of TFIIA (Figure 18). In contrast, full length TBP was able to bind to the *HIS4* promoter when TFIIA was omitted from the reaction, albeit more weakly than the core of TBP. When TFIIA was included, both the core and full length versions of TBP bound to the *RPS5* promoter, suggesting that TFIIA might have an important role in stabilizing TBP on this TATA-less promoter.

A DNase I footprint assay performed with the same *RPS5* sequence used for the gel shift studies described above revealed that TBP binds to this promoter at two distinct sites (Figure 19). The upstream binding site was positioned approximately 60 bp away from the transcription start site, and the downstream site was approximately 30 bp away. This binding was observed both in the presence and in the absence of TFIIA and the amount of TBP used was equivalent to the amount needed to footprint the *HIS4* TATA box (Figure 19 right). Unfortunately, attempts to obtain a footprint of TFIID over the *RPS5* promoter were unsuccessful. A representative attempt is shown here (Figure 19 middle) and the preparation of TFIID used was demonstrated to be active in transcription both on the *RPS5* and *HIS4* promoters (Figure 4B).

The TBP binding sites on the RPS5 promoter are redundant

The discovery of two distinct binding sites on the *RPS5* promoter prompted the question whether they make a functional contribution to transcription. I created a series of mutant *RPS5* transcription plasmids designed to answer this question. I made single 8 base pair deletions in the middle of each binding site, as well as a construct with both sites deleted (Figure 20). I also replaced the middle 8 base pairs of each binding site with a sequence from the *HIS4* promoter which does not bind TBP, also in single and double combination (Figure 20). Surprisingly, all of these mutant promoters, with the exception of the double replacement mutant, had wild type levels of activity in an *in vitro* transcription assay (Figure 21). The positions of the starts sites were not altered in any of the mutant promoters. The findings that mutating significant portions of the *RPS5* promoter had a minimal effect on transcription are consistent with *in vivo* results, which showed a similar phenomenon (Sugihara et al. 2011). DNase I footprints with the double deletion and double replacement *RPS5* mutant promoters demonstrated that TBP is still bound despite removal of its binding sites, although the location and number of sites where TBP binds is altered (Figure 22). It appears that removal of both TBP binding sites prompts the protein to use another site on the *RPS5* promoter, which can still lead to productive transcription, but somehow replacement of both sites with GC-rich *HIS4* sequences leads to TBP binding that is not conducive to transcription.

Insertion of a canonical TATA box sequence in the RPS5 promoter enhances in vitro transcription

It is also interesting to explore the effect of introducing a canonical TBP binding site into the TATA-less *RPS5* promoter. Changing a single base pair in the sequence from TCTAAAAT to TATAAAAT creates a TATA box in *RPS5,* just upstream of the FP1 binding site defined in the footprinting analysis. The *in vitro* transcription activity of this construct was stronger than the activity of the wild type *RPS5* promoter. This *RPS5*-TATA sequence supported levels of transcription comparable to those observed with the *HIS4* promoter (Figure 23). It would be interesting to explore the reason for this increase in activity and to test whether this new *RPS5* promoter is dependent on TFIID for its activity.

CONCLUSIONS

In this chapter I have presented *in vivo* evidence that the DNA-binding function of TBP may not be required for transcription from the TATA-less ribosomal protein genes. I showed that several TBP mutants which fail to bind TATA-box DNA *in vitro* can rescue defective transcription on the RP promoters but not on the TATA-containing Gcn4-dependent genes. I further demonstrated that at least one DNA-binding mutant, TBP T112K, was able to rescue *in vitro* transcription on the *RPS5* promoter when a TBP-deficient extract was used. This same mutant did not support transcription on the TATA-containing *HIS4* promoter. Gel shift experiments showed that TBP can form a complex on the *RPS5* promoter in the presence of TFIIA and competition experiments with other TATA-less promoters suggested that different TATA-less promoters interact with TBP. DNase I footprint mapping of the TBP binding sites on the *RPS5* promoter revealed two distinct positions. However, deleting both sites had no effect on transcription from the *RPS5* promoter, and footprinting analysis of mutant promoters showed that TBP is still capable of binding. These experiments suggest that the DNA binding requirements of TBP on TATA-less promoters are flexible, contrary to the case at TATA-containing genes.

CHAPTER 3: MAPPING THE POSITION OF RNA POL II SUBUNIT RPB5 IN THE TRANSCRIPTION PRE-INITIATION COMPLEX

INTRODUCTION

The eukaryotic Pol II enzyme is composed of 12 subunits termed Rpb 1-12, ten of which are essential for viability. These subunits are arranged in a globular core, which contains the DNA cleft, the nucleotide triphosphate entry pore and the active center (Cramer et al. 2008). Two subunits, Rpb4 and Rpb7 form a heterodimer located at the base of Pol II. This region is termed the stalk and can dissociate from Pol II under certain conditions (Jasiak et al. 2008). The Rpb5 subunit is also part of the Pol I and pol III enzymes and is located on the lower jaw of Pol II, the bulk of which is formed by the largest Pol II subunit, Rpb1. The jaw is a mobile region of the polymerase which changes conformations during transcriptional elongation (Cramer et al. 2008). The position of Rpb5 close to the transcription start site suggests a function in stabilizing the transcription bubble or interactions with factors responsible for holding the open DNA strands (Zaros et al. 2007). Experiments with archaeal transcription factors have indeed demonstrated that the C-terminal domain of Rpb5 functions in open complex formation and this role can be performed interchangeably between the yeast and archaeal proteins (Grunberg et al. 2010). Yeast Rpb5 contains an additional N-terminal domain, absent in archaea, but conserved among eukaryotes, which has an unknown function and a more peripheral position on the surface of Pol II (Zaros et al. 2007).

Recent findings have defined the position of the Ssl2 subunit of TFIIH in the PIC, placing it close to Rpb5 (Grunberg et al. 2012). This model has led to speculations of what other TFIIH subunits might be located in proximity to Rpb5. TFIIH is a ten subunit complex required to generate the transcription bubble. It has three distinct domains that show great flexibility and contain three different enzymatic activities: Rad 3 helicase activity, required for the transcription-coupled DNA repair function of TFIIH, Kin28 (Cdk7 in humans) kinase activity necessary for Pol II CTD phosphorylation at Serine 5, and the putative DNA translocase activity of Ssl2 essential for open complex formation (Gibbons et al. 2012, Hahn et al. 2011). In addition to the described interactions of TFIIH with TFIIE and promoter DNA (Kim et al. 2000, Miller et al. 2006, Grunberg et al. 2012), some studies have also suggested that it interacts with the Mediator complex (Imasaki et al. 2011). However, the Mediator-TFIIH interface is not well defined.

Given the likelihood of finding an interaction between Rpb5 and TFIIH, we introduced the photoreactive crosslinking amino acid Bpa at surface-exposed residues in Rpb5 and asked whether Rpb5 crosslinks to TFIIH in PICs assembled *in vitro* or in living cells. The results presented here show that Bpa substitutions at residues Q101 and K122 of Rpb5 crosslink to multiple polypeptides in PICs, and experiments to identify these proteins are underway.

MATERIALS AND METHODS

Yeast strain construction and growth conditions

All Rpb5 amber strains generated for this study are derived from IKY30 (Rpb5 shuffle strain with 6xHis-linker-RPB3) by plasmid shuffling/5-FOA selection. The amber mutations in Rpb5 were generated in pRS425/Adh1p-13xMyc-Rpb5 and this plasmid was co-transformed with pLH157 (Chen et al. 2007) by Lithium acetate transformation. Transformants were grown on GC plates lacking appropriate amino acids and supplemented with 1 mM Bpa before 5-FOA selection and 1 mM Bpa was included in the media from this point. Amber strains were grown as starter cultures in GC media lacking tryptophan and leucine and inoculated into YPD (3 % Dextrose) for nuclear extract preparation.

Generation of Rpb5 amber mutations

Surface-exposed residues in Rpb5 were selected using Naccess software (S. Hubbard and J. Thornton 1992-6) and chosen to have low conservation among species based on T-coffee protein sequence alignment. The low degree of conservation criterion was used in order to minimize the chance of introducing Bpa at residues with important function. Mutagenesis was carried out using Quickchange.

Nuclear extract preparation

Yeast nuclear extract preparation from amber mutant strains was as described in (Chen et al. 2007). Due to poor growth of the amber strains used in these experiments, 150 ml of saturated overnight culture per liter YPD was used to inoculate 4 L of YPD (3 % glucose). Cells were incubated at 30 °C for 2 days until reaching an OD600 ~3-4.

Pol II purification

Purification of Pol II was carried out as described in (Fishburn et al. 2012). Pol II was purified from yeast strain IKY39 (derived from IKY30 by plasmid shuffle; pRS425/Adh1p-13x-Myc-Rpb5 [Q101amber]; pLH157) inoculated into 12 L of YPD (3 % Dextrose) supplemented with 1 mM Bpa.

Bpa crosslinking and analysis of crosslinking results

UV crosslinking of PICs formed with a nuclear extract was performed as described in (Chen et al. 2007). For crosslinking of PICs using purified factors, proteins were added in the amounts specified in Chapter 1 *In vitro transcription and primer extension* during the PIC formation step and the rest of the procedure was identical to the one for PICs formed with nuclear extracts. *In vivo* crosslinking was carried out by harvesting 200 ml of OD_{600} ~ 1 culture grown in GC media supplemented with 1mM Bpa. Cells were resuspended in 30 ml GC media and half were reserved for a –UV control. The other half was plated in three wells of a 6-well cell culture dish and irradiated using the same instrument and setting used for PICs. Both the control and the irradiated culture were washed in deionized water and the cell pellets frozen on dry ice and stored at -70 °C. Whole cell extracts were prepared using a mini bead beater in RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10 % glycerol, 1 % NP-40 0.5 % Sodium deoxycholate, 1 % SDS). Approximately 1 mg of WCE diluted 3 fold in RIPA buffer lacking SDS was used to immunoprecipitate Rpb5 with 10 µl Myc-agarose slurry (SIGMA). The IP reaction was allowed to proceed for at least 2 hours (or overnight) at 4 °C and the beads were washed 3 times with RIPA buffer containing 0.1 % SDS. The Myc-tagged protein was eluted by boiling the beads in 2xNuPAGE sample buffer (Life Sciences) for 5 min at 95 °C, or in 0.1 M ammonium hydroxide, following the Myc-agarose beads manufacturer’s protocol (Sigma). Protease inhibitors (same as for nuclear extract preparation) were included in the WCE and IP buffers, but DTT was only added before gel electrophoresis.

RESULTS

Incorporation of Bpa on the surface of Rpb5

The peripheral position of Rpb5 in the lower jaw of Pol II makes it an attractive candidate for testing interactions with proteins whose location in the PIC has not been determined yet. We used the amber codon suppression system, described in detail in (Chen et al. 2007), to incorporate the photoactive crosslinking amino acid Bpa into surface exposed residues of the Rpb5 protein (Figure 24). Since a large portion of the C-terminal domain of Rpb5 is buried in the interface with Rpb1, we placed the majority of amber mutations in the eukaryote-specific N-terminal domain of Rpb5. After transforming the Rpb5 mutants into an Rpb5 yeast shuffle strain and eliminating the wild type copy of Rpb5 by 5-FOA selection, I found that most of the amber mutants, with the exception of Rpb5 M121, were viable in the presence of the suppressor tRNA and tRNA synthetase (Table 3). All of the amber mutants grew better in the presence of Bpa in the media and the protein levels of the full-length protein product was higher in cells grown in Bpa-supplemented media (not shown).

Rpb5 residues Q101 and K122 crosslink to a number of proteins in the PIC

To determine which proteins are positioned near Rpb5 in the Pol II PIC, I prepared nuclear extracts from the Rpb5 amber cells amenable to such large scale growth (Table 3). These extracts were used to form PICs on the *HIS4* promoter. After washing PICs to remove nonspecifically bound proteins, half of each PIC reaction was treated with UV to induce crosslinking. The PIC components were resolved on SDS PAGE gels, followed by western blotting against an N-terminal Myc tag on Rpb5 revealed distinct high molecular weight products which were present only in the +UV lanes.

Attachment of Bpa to two Rpb5 residues situated in the back portion of Pol II, F29 and Q32, generated weak a crosslinking band of high molecular weight (Figure 25). Given the proximity of these amino acids to Rpb2, it is possible that the crosslinking product is an Rpb5-Rpb2 fusion protein. Alternatively, it could represent an interaction between Rpb5 and a Pol I or Pol III subunit, as Rpb5 is a component of all three RNA Pols (Cramer et al. 2008). Inserting Bpa at positions H99 and F135, both somewhat buried into the surface of Rpb5, did not result in any detectable crosslinks (Figure 25). Most notably, Bpa at residues Q101 and K122, which are located close to each other in an alpha helix region of Rpb5 (Zaros et al. 2007), generates a number of strong crosslinking bands, ranging in estimated molecular weights from 30 to140 kDa (Figure 26 and Table 3). The fact that these two amino acids reside close to each other and produce a similar crosslinking pattern suggests that they share interactions with the same target(s).

Attempts to identify proteins crosslinking to Rpb5 Q101 by candidate approach

Identifying the proteins crosslinked to a Bpa-labelled protein requires estimation of the molecular weight of the crosslinked protein, followed by predicting its identity and testing potential candidates by epitope tagging and western blotting. Given our hypothesis that inserting Bpa into Rpb5 will help us define the position of additional subunits of TFIIH and that several of the crosslinked proteins were of the correct size to be TFIIH components, we explored the idea that Rpb5 crosslinks to TFIIH. I tested this prediction using multiple methods: mixing nuclear extracts with epitope tags on TFIIH subunits along with Q101-bpa extract, probing western blots from crosslinking against TFIIH proteins, immunodepleting TFIIH to see if the crosslinked bands would disappear or decrease in intensity, and ultimately, Flag-tagging individual TFIIH subunits in the context of Q101-amber and preparing nuclear extracts from these new strains. None of these approaches suggested crosslinking of TFIIH to Rpb5. I also predicted that the largest crosslinking product may be Rpb1, because Q101 and K122 are both very close to Rpb1. However, probing with Rpb1 anti-sera did not reveal crosslinking to the largest subunit of Pol II.

Rpb5: Q101-bpa crosslinks to the small subunit of TFIIE in minimally assembled PICs on the HIS4 promoter

Due to the difficulties of identifying a crosslinking target among the hundreds of proteins in a nuclear extract, I decided to try purifying Pol II with the crosslinker attached to position Q101, which had yielded the strongest crosslinks, and use this purified Pol II to crosslink PICs assembled with only the minimal set of required transcription factors. I was able to successfully purify RNA Pol II from the Rpb5: Q101 amber strain and to demonstrate its ability to stimulate transcription in a minimal transcription system consisting of the purified general transcription factors and the *HIS4* promoter (Figure 27).

Crosslinking the purified polymerase separately from other proteins yielded a novel crosslink corresponding to a 16 KDa protein (not shown). Based on the crystal structure of the Pol II enzyme and the known molecular weights of the individual subunits (Cramer et al. 2008), I predict this protein to be Rpb6, although this crosslink could be attributed to a contaminant in the Pol II prep.

I tested the crosslinking of Pol II to various GTFs by excluding them from the PIC reaction when PICs were assembled on the *HIS4* promoter. If one of the GTFs is responsible for the crosslink, the high molecular band should disappear from the UV-treated reaction when the protein is omitted. TFIIH is the last of the GTFs to join the PIC during assembly (Hahn et al. 2011, Thomas et al. 2006). When I left it out of the PIC reaction, the recruitment of the other PIC components was unaffected and surprisingly, the crosslinking pattern was identical to the one observed when all the GTFs were present (Figure 28). TFIIE enters the PIC before TFIIH and helps to recruit and stabilize TFIIH (Hahn et al. 2011). Excluding TFIIE from the crosslinking reaction led to unexpected loss of crosslinking (Figure 28, left). When I probed the PICs with an antibody against Tfa2, the small subunit of TFIIE, there was a signal that overlapped with the crosslinking band detected with α-Myc antibody and this signal was only present in the UV-treated samples that had TFIIE (Figure 28 right).

In vivo crosslinking in the Rpb5: Q101-amber strain

In addition to crosslinking transcription factors in PICs, which provides a way to isolate transcriptionally active Pol II complexes, Bpa crosslinking can also be utilized to capture protein-protein interactions in living cells. I was interested in probing the surroundings of Rpb5: Q101 *in vivo* and comparing my observations with those made at isolated PICs. It is important to note that Ppb5 is also a subunit of Pol I and Pol III, therefore making it likely that any *in vivo* crosslinks could be due to interactions made at the Pol I and III PICs. There are also many polymerases engaged in active elongation, and *in vivo* crosslinking would also capture interactions made by elongating polymerase. Finally, UV treatment most likely results in activation of the DNA repair machinery, which could also contribute to the observed crosslinks.

Crosslinking cells carrying amber mutation at Rpb5 Q101 resulted in a very similar crosslinking pattern to the one observed *in vitro*. However, the two smaller crosslinking products were of slightly lower molecular weight (Figure 29). While I have not been able to reveal the identity of these crosslinking proteins, I was able to show that they can be purified by immunoprecipitation via a Myc epitope tag on Rpb5 (Figure 29). The ability to observe these crosslinking bands by silver stain make the case that this method can be used for an unbiased identification of crosslinking products by mass spectrometry (Majmudar et al. 2009, Guerrero et al. 2006).

CONCLUSIONS

Introducing Bpa at positions Q101 and K122 of Rpb5 resulted in crosslinks to a number of proteins ranging from ~30to 140 kDa when crosslinking was performed in the Pol II PIC using a nuclear extract. Additionally, Q101-bpa crosslinked to the small subunit of TFIIE, Tfa2, in PICs formed using purified general transcription factors. Crosslinking cells carrying Q101-Bpa *in vivo* resulted in a slightly altered but similar crosslinking pattern to the one observed *in vivo*. While multiple attempts to test TFIIH subunits as candidate crosslinking proteins did not succeed, experiments to identify the crosslinking targets of Rpb5 using mass spectrometry are under way.

Rpb5 amber mutant	*GC-LW*	*GC-LW+1mM Bpa*	*NE*	*cross-linking*	*Size of cross-linked protein*
pRS315	-	-	-	-	-
Adh1p-13xMYC-RPB5 wt	+++	+++	✓	✗	-
N8	-	++	ND	ND	
F29	+/-	++	✓	✓	> 140 KDa
Q32	+/-	+++	✓	✓	> 140 KDa
E40	+/-	+++	ND	ND	
E85	ND	ND	ND	ND	
H99	+/-	++	✓	✗	
Q101	-	++½	✓	✓	130-140 KDa 55 Kda <55 Kda (Tfa2)
Y112	-	+++	ND	ND	
M121	-	-	ND	ND	
K122	-/+	++	✓	✓	130-140 KDa 55 KDa
F135	-/+	+++	✓	✗	
I154	-	+½	ND	ND	
R180	-/+	++	ND	✓	130-140 KDa 55 KDa
R192	-/+	++	ND	✓	

Table 3 Summary amber mutations introduced in Rpb5 and crosslinking products that were detected.

ND stands for Not Determined, due to extremely slow growth of these strains. NE stands for nuclear extract.

DISCUSSION

Why are the majority of TATA-less promoters inactive in vitro?

I have successfully developed an *in vitro* transcription system using the *Saccharomyces cerevisiae RPS5* TATA-less core promoter. However, most of the TATA-less promoters tested, the majority of which are also RP promotes, were inactive *in vitro* and it is important to understand why. The inability to obtain active transcription from many RP promoters *in vitro*, even though these genes are some of the most active *in vivo* (Warner 1999), points to the fact that there is still a great deal of mystery surrounding the transcription regulation of RP genes. The full array of transcription activators required for RP promoters is yet to be determined. While it has been demonstrated convincingly that RP genes depend on Rap1 for their expression (Wade et al. 2004, Mencia et al. 2002), and that Rap1 is important for the recruitment of TFIID (Mencia et al. 2002, Garbett et al. 2007), it is unclear what specifies Rap1 as an activator, as it can act as a transcriptional repressor at other promoters . It is also well known that Rap1 activates the glycolytic genes which are not TFIID-dependent (Hu et al. 2007). The need for the transcriptional activators Fhl1 and Ifh1, whose recruitment to RP promoters depends on Rap1, has been established *in vivo* (Wade et al. 2004, Schawalder et al. 2004), but their role *in vitro* has been difficult to dissect. One study demonstrated that the requirement for Fhl1 and Ifh1 can be bypassed *in vitro* when the UAS from the *RPS8A* promoter was used to drive transcription (Garbett et al. 2007), and this finding agrees with the result that a double deletion of Fhl1 and Ifh1 is viable (Wade et al. 2004, Schawalder et al. 2004). Additional complications arise from the identification of two other factors, Hmo1 and Sfp1, which bind to the upstream sequences of RP promoters and are important regulators of transcription from this gene cluster (Hall et al. 2006, Jorgensen et al. 2004, Marion et al. 2004), but the mechanisms through which they modulate transcription are unknown. Therefore, the regulation of RP transcription is complex and *in vitro* studies with these promoters will be greatly enhanced by a more thorough understanding of the mechanisms by which transcription is activated from these genes.

It is possible to explain the inactivity of RP promoters by the presence of repressive elements in the promoter sequences used in this study. There is at least one report which demonstrates that the co-repressor Crf1 can modulate transcription from RP promoters *in vivo* (Martin et al. 2004), although these results appear to be strain-dependent (Zhao et al. 2006). Even though the details of RP gene repression are not well understood, it is certainly possible that this could be an explanation for the low activity of RP promoter constructs which include the full UAS. It could be useful to deplete the levels of Crf1 in an extract to see whether this results in a higher activity.

In the case of constructs where the core region was used in combination with the Gal4-VP16 activator, lack of activity could be due to incompatibility between the UAS and the core promoter sequence. *In vivo* experiments with chimeric promoter constructs combining TFIID-dependent and independent UAS and promoter constructs have resulted in somewhat inconsistent results, with examples of both compatibility and incompatibility between core promoter and UAS (Shen et al. 1997, Cheng et al. 2002, Li et al. 2002). Therefore, it seems plausible that UAS sequences which recruit TFIID are best suited for transcription from TFIID-dependent, TATA-less promoters.

Lastly, our construct using the *RPS5* promoter included sequences downstream from the ATG start of translation, while all the other promoter constructs were designed to end at the ATG. It is possible that these downstream sequences make an important contribution to transcription. However, *in vivo* deletion analysis of the *RPS5* promoter did not indicate that the

DNA sequence past the ATG is critical for transcription (Sugihara et al. 2011), and it is unlikely, although still possible, that lack of these downstream DNA sequences is the reason why the other RP promoters were inactive *in vitro*.

Utility of the RPS5 promoter as a model TATA-less TFIID-dependent gene

The transcription system described here is the first example of a yeast TFIID-dependent TATA-less core promoter shown to be active *in vitro*. Several other examples of similar, though not identical systems have been described by others. The upstream region from the TATA-less *RPS8A* core promoter, including two Rap1 binding sites, has been used in conjunction with the TATA-containing *PGK1* core promoter. This construct requires TFIID, Rap1, and TFIIA for transcription with a whole cell extract (Garbett et al. 2007, Layer et al. 2013), and as such is a model of a TFIID-dependent TATA-containing promoter. More recently, an *in vitro* transcription system using the yeast TATA-less *TMT1* promoter, which is regulated by the Gcn4 activator, has been developed (Seizl et al. 2011). This core promoter forms a complex with TBP and TFIIB *in vitro*, but it is unclear whether it requires TBP, TFIID, or either.

Our *in vitro* transcription system is dependent on TFIID, consistent with the *in vivo* evidence (Tsukihashi et al. 2001, Huisinga et al. 2004). The use of VP16, rather than the natural activators, to drive transcription from this promoter is a drawback that we have not been able to overcome. However, *in vivo* experiments have shown that an acidic activator such as Gal4 can be used to activate transcription from the *RPS5* core promoter and the coactivator requirements of the native promoter were preserved (Cheng et al. 2002). As an additional attempt to use the native activators of the *RPS5* promoter, we purified a minimal activation domain of Ifh1 fused to the Gal4 DNA-binding domain. This construct was able to activate transcription both from the

TATA-containing Gal genes and from the TATA-less *RPS5* promoter *in vivo* (Zhong et al. 2010). In our *in vitro* experiments, however, this recombinant activator had poor activity on *RPS5*, while it was very active on the TATA-containing *HIS4* promoter. This result supports the notion that a complementary combination of activator and DNA elements is needed for the full activity of the RP promoters and these combinations need to be dissected in greater detail.

Transcription requirements of TATA-less promoters

There are no examples of *in vitro* transcription systems using a minimal set of purified transcription factors and a TATA-less promoters such as the ones described for TATA-containing promoters (Fishburn et al. 2012, Knuesel et al. 2009). Our attempts to develop such a system using the GTFs that are sufficient for high levels of basal transcription from *HIS4* did not yield active transcripts. Addition of TFIID and TFIIA, which are particularly important for transcription from TATA-less RP promoters (Layer et al. 2013), did not result in enhancement of transcription, and neither did addition of Mediator. Some reports have suggested that the NC2 complex is important for transcription from TATA-less promoters (Hsu et al. 2008), but we have not tested this hypothesis.

Attempts to identify novel or unsuspected proteins involved in transcription of TATA-less promoters have included mass spectrometry analysis of PICs isolated on the *RPS5* promoter. These experiments did not reveal an obvious candidate. However, it is certainly reasonable to suspect that the use of the Gal4-VP16 activator during PIC formation has obscured the results. Collectively, these findings highlight the fact that TATA-less promoters have additional transcription factor requirements and there is a need for more sophisticated ways of analyzing the components of a TATA-less PIC. One such very promising technique is the isolation of native

DNA elements along with their associated protein components, followed by subsequent analysis by mass spectrometry (Dejardin et al. 2009).

Importance of DNA-binding function of TBP on TATA-less promoters

The transcription analyses of mutants in the DNA-binding surface of TBP extend on previous findings which have proposed that TBP's DNA binding role is important for transcription from TATA-containing, but not TATA-less promoters (Martinez et al. 1995). My findings support this hypothesis, by demonstrating that DNA-binding mutants are able to rescue transcription from TATA-less, but not TATA-containing promoters, both *in vivo* and *in vitro*. It would be interesting to compare the genomic localization of TBP deficient in DNA-binding on TATA-less and TATA-containing promoters. According to our hypothesis, TBP can still be recruited to TATA-less promoters via interactions of the Tafs with promoter DNA or the nucleosomes surrounding the promoter region. If the interactions between TBP and Tafs are solely responsible for bringing TBP to TATA-less promoters, disrupting the TBP-Taf interface should result in decreased transcription levels from TATA-less promoters. At least one such mutant is available (Ranallo et al. 1999) and can be used to test this idea experimentally.

Assuming that TBP's DNA binding function is less critical for transcription from TATA-less promoters, it is interesting to consider the possibility that TBP may be located farther away from DNA in TATA-less promoters. Single molecule analysis of the TBP-TATA box interactions and degree of bending caused by TBP has made the surprising discovery that binding TBP to a mutant TATA box results in same degree of bend as observed with a canonical TATA sequence, but lower percentage of bent complexes (Blair et al. 2012). This finding fits with the idea that TBP would still be located on the TATA-less promoter DNA as opposed to away from it. Unfortunately, our experiments do not address this question directly, and neither do *in vivo* chromatin immunoprecipitation experiments. Positioning of a crosslinking or protein/DNA cleavage probe either on the TATA-less promoter DNA or TBP itself (Theisen et al. 2010, Miller et al. 2006) would be one way to gain understanding on this intriguing question.

Proposed models for PIC formation on TATA-less promoters

One of the outstanding questions regarding TATA-less promoters is whether the PIC is assembled at a well-defined location or it has a more flexible position. Results to support both models exist. On one hand, most mammalian promoters of housekeeping genes are of the ‘broad’ type, meaning that transcripts originate from a region spanning 50-100 base pairs (Carninci et al. 2006). Because mammalian Pol II initiates transcription at a specified distance downstream from the site of PIC formation (Thomas et al. 2006), this observation suggests that multiple sites of PIC formation are present on these promoters. In addition, many broad promoters lack any recognizable core promoter elements and are often located in GC rich regions (Kadonaga 2012). It is thought that the anchoring of TFIID via its interactions with modified nucleosomes is important for placement of the PIC.

Yeast TATA-less promoters also lack any characterized DNA elements which could serve as the site for PIC nucleation. However, TATA-less promoters in yeast are predominantly AT-rich, which distinguishes them from mammalian promoters. One proposed model for PIC formation involves the strategic positioning of nucleosomes. Fine mapping of nucleosomes surrounding the transcription start sites of all yeast genes suggests that the +1 nucleosome is at the border of the transcription start sites specifically of RP promoters, indicating that it may somehow function to limit the site of PIC formation (Rhee et al. 2012). In further support of this idea, the same study found that the Taf1 subunit of TFIID, which interacts with the double bromodomain protein Bdf1 (Matangkasombut et al. 2000), has a binding site that overlaps with the +1 nucleosome and Bdf1. It is plausible that the binding of Bdf1 to acetylated histones serves to recruit TFIID to a specific position of TATA-less promoters, thus defining the site of PIC formation.

My findings support the idea that TATA-less promoters may allow for greater flexibility of the positioning of the PIC. I found two distinct sites on the *RPS5* promoter where TBP resides *in vitro*. However, deletion of both sites had no effect on transcription from this promoter. These results are in agreement with the *in vivo* findings reported by (Sugihara et al. 2011), demonstrating that large portions of the *RPS5* promoter can be mutated with no loss of activity. Interestingly, the upstream TBP binding site that I identified overlaps perfectly with the high resolution TBP binding site found *in vivo* for this promoter (Rhee et al. 2012). These findings suggest that TBP is positioned *in vivo* at a site that it has high affinity for *in vitro,* but there is significant flexibility in the DNA sequences which can be used to form productive transcription complexes.

Insights from Bpa crosslinking with Rpb5

I have identified a surface on the Rpb5 protein defined by the residues Q101 and K122 which is in close proximity to a number of proteins in the PIC. Q101 was found to crosslink to the small subunit of TFIIE in PICs formed with purified transcription factors. According to the most recent models of the yeast PIC structure (Grunberg et al. 2012), TFIIE is positioned far away from the site of Bpa incorporation. However, only the positions of the structured domains of TFIIE have been modeled and their interactions in the PIC mapped. Tfa2 has an N-terminal and a C-terminal unstructured region with undetermined locations in the PIC. Deletion analysis has shown that the entire N-terminal unstructured domain of Tfa2 can be removed without any phenotype in yeast, while deletion of the C-terminal unstructured region results in slow growth (Grunberg et al. 2012). Making recombinant TFIIE mutants with deletions of each of these two regions and testing them in crosslinking will help narrow down the segment of Tfa2 which is close to Rpb5. Additional support for positioning parts of Tfa2 in this region of the PIC comes from previous crosslinking results with Bpa incorporated in Rpb1, which crosslinked to Tfa2 (Chen et al. 2007).

My attempts to demonstrate crosslinking between Rpb5 and TFIIH did not show such an interaction, despite many indications that TFIIH is positioned close to the transcription start site (Kim et al. 2000, Miller et al. 2006, Grunberg et al. 2012). It is possible that subunits of other complexes, such as Mediator or TFIIE are wedged in between Rpb5 and TFIIH. It will be important to uncover the identity of the crosslinks generated by Rpb5, especially since the Q101 residue crosslinked to multiple polypeptides *in vivo*, highlighting the large number of proteins positioned in this region of the PIC in all conditions tested. In this work, I have demonstrated that crosslinked Rpb5 can be purified successfully and this strategy is the basis for developing an unbiased method for crosslinked protein identification via mass spectrometry (Guerrero et al. 2006, Majmudar et al. 2009, Krishnamurthy et al. 2011)

FIGURES

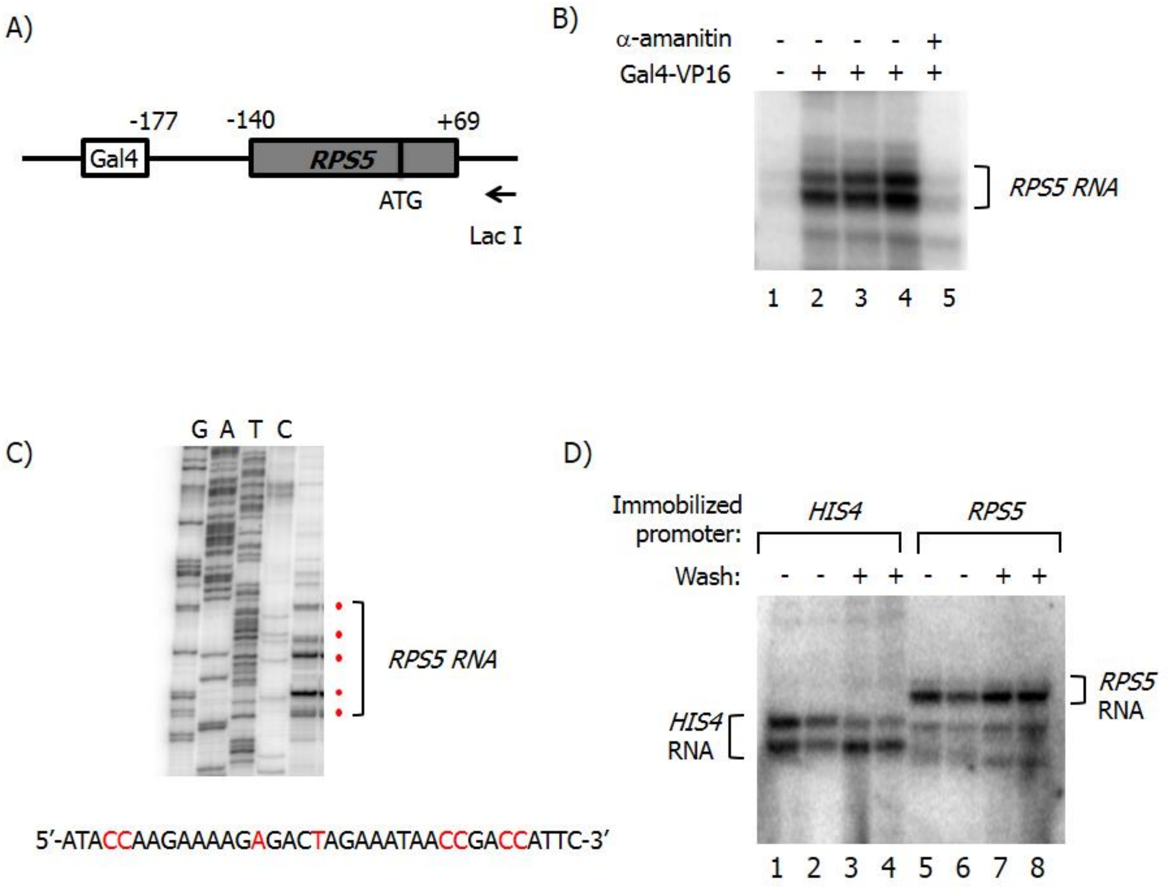


Figure 1 The yeast TATA-less promoter *RPS5* is active in an *in vitro* transcription system A) Diagram of the *RPS5* core promoter construct used for *in vitro* transcription assays. The core promoter of *RPS5* (-140 to +69 bp relative to the ATG) was inserted into pSH515 downstream from a single Gal4-binding site. Primer extension was assayed using a lacI oligo. B) *In vitro* transcription reactions with the *RPS5* core promoter using a whole cell extract were visualized by primer extension with 32P-labeled lacI oligo. Recombinant Gal4-VP16 activator and α-amanitin were included where indicated. C) Primer extension of *RPS5* plasmid transcription was resolved on a sequencing 8% Acrylamide gel and *in vitro* transcription start sites are marked by red dots. The red bases in the *RPS5* promoter sequence correspond to *in vitro* mapped transcription start sites. The coding strand is shown D) Immobilized template assay showing stability of transcripts from immobilized *HIS4* (lanes 1-4) or *RPS5* promoters (lanes 5-8). A mixture of nuclear and whole cell extracts was used to form the *RPS5* PICs.

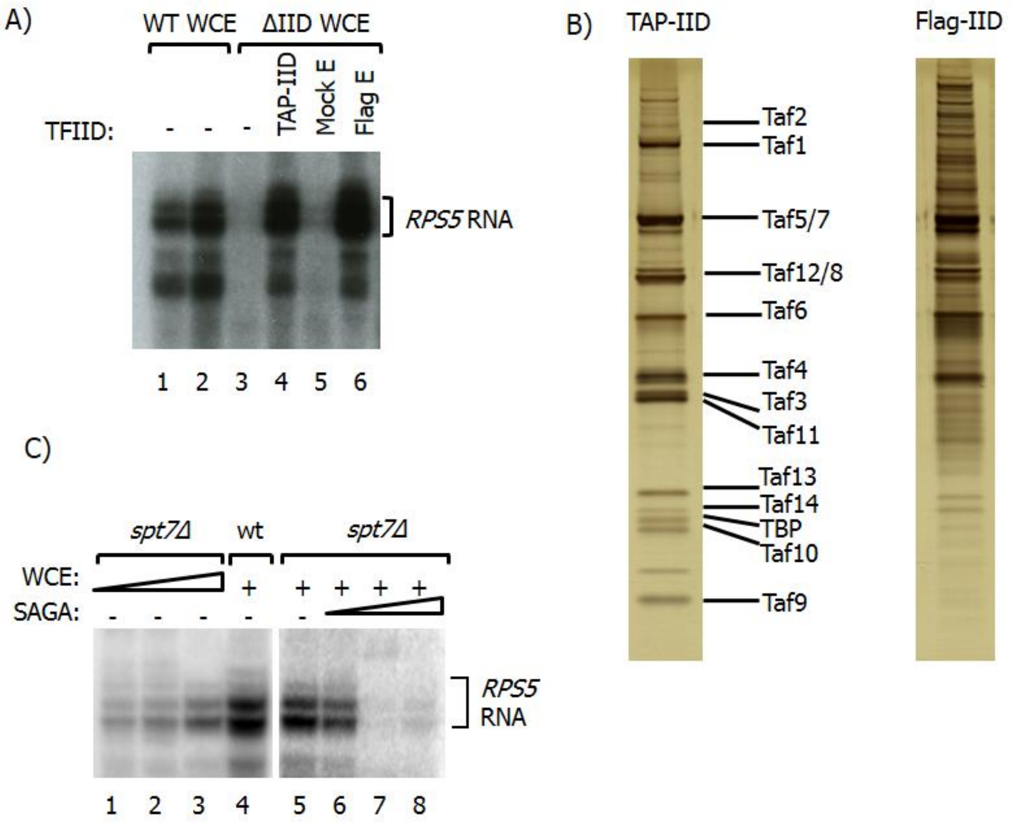


Figure 2 The *RPS5* core promoter is TFIID-dependent and SAGA-independent *in vitro*
A) *In vitro* transcription on an *RPS5* plasmid template using a Taf11ts whole cell extract (lanes 3-6) supplemented with Tap-purified TFIID (lane 4), mock-immunoprecipitated TFIID from an untagged whole cell extract (lane 5), or immunoprecipitated TFIID via triple Flag tag on Taf1 and Taf3. Lanes 1 and 2 show transcription using a whole cell extract from an isogenic strain. B) TFIID used to supplement transcription reactions in A) purified via Tap-tag on Taf13 (left) or crudely purified via Flag tag on Taf1 and 3 (right). C) *In vitro* transcription on an *RPS5* plasmid template with a whole cell extract prepared from an *Spt7* deletion strain. Tap-purified SAGA was added where indicated.

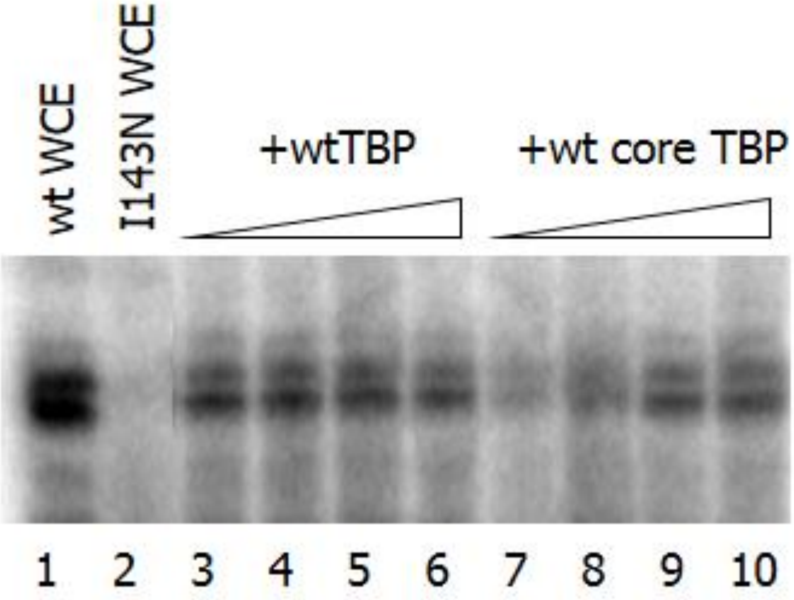


Figure 3 TBP is required for transcription from the *RPS5* promoter
In vitro transcription from *RPS5* plasmid template using a whole cell extract from a TBPts strain harboring mutation I143N. Purified recombinant full length TBP or wild type TBP core (Δ2-60) were added where indicated.

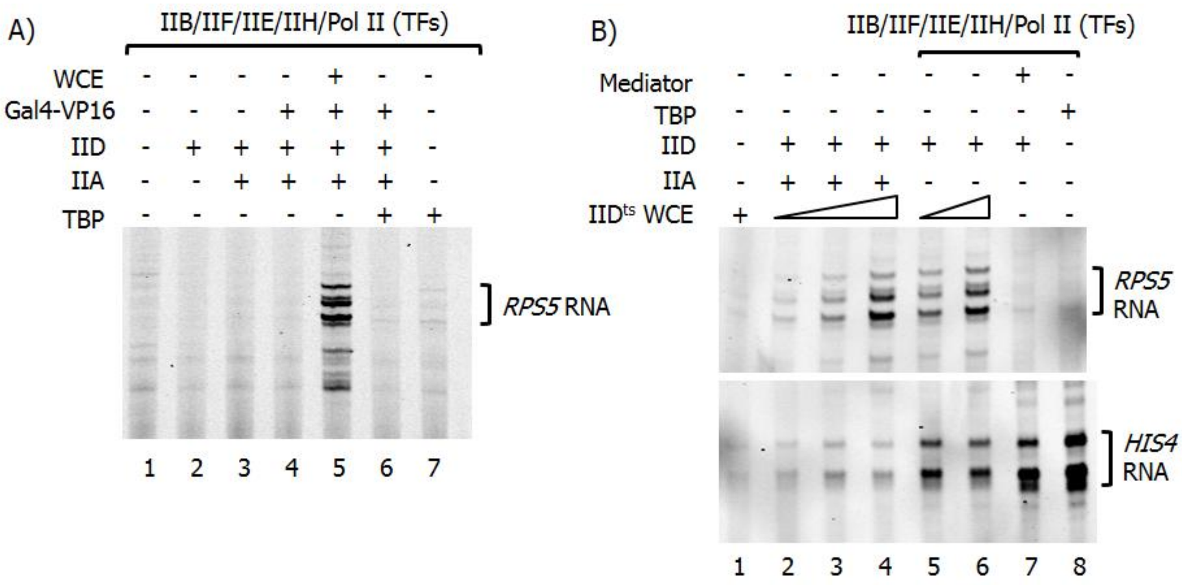


Figure 4 General transcription factors are insufficient for *RPS5* transcription
A) *In vitro* transcription from *RPS5* plasmid template using purified general transcription factors is visualized by primer extension with a fluorescently labelled lacI oligo. Wild type whole cell extract and Gal4-VP16 activator were added where indicated. B) *In vitro* transcription from *RPS5* and *HIS4* plasmid templates performed as described in A). The whole cell extract added to reactions 1 through 6 was from a Taf11ts strain. These experiments were performed by James Fishburn.

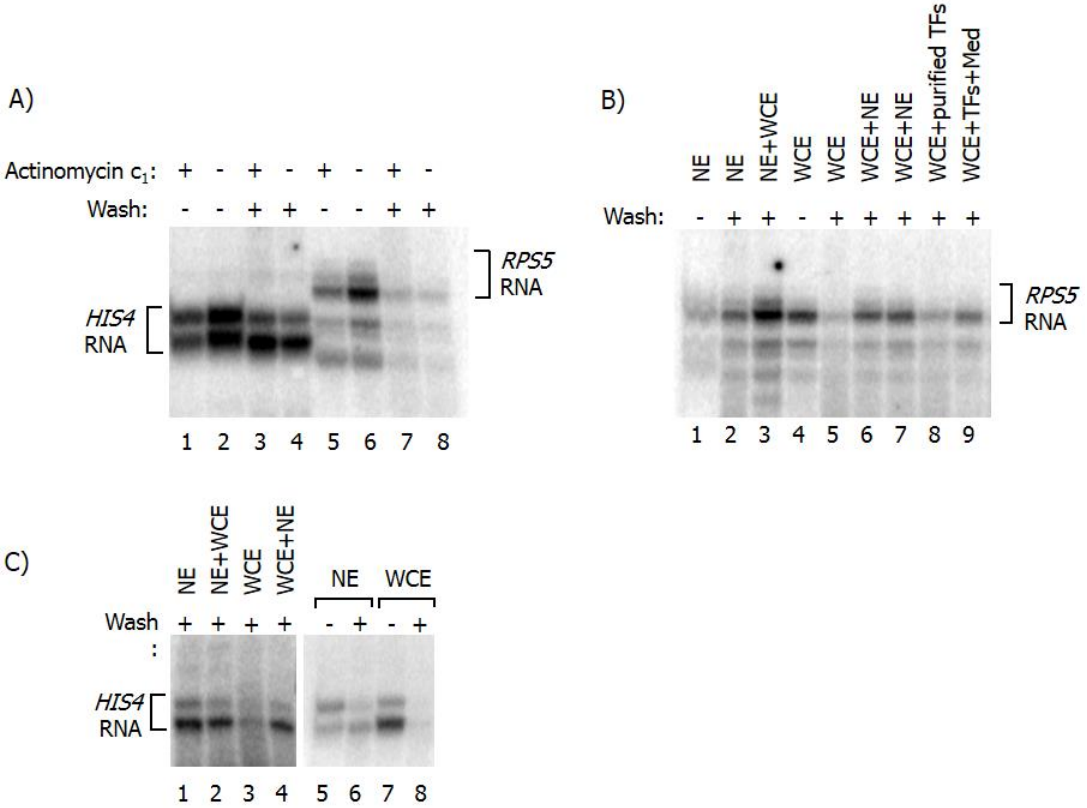


Figure 5 Characterizing *in vitro* transcription from *RPS5* immobilized promoter DNA

A) Immobilized template transcription from *HIS4* promoter (lanes 1-4) and *RPS5* promoter (lanes 5-8) performed as described in Figure 1. *HIS4* PICs were formed with a nuclear extract and *RPS5* PICs were formed with a whole cell extract. Gal4-VP16 was included in all reactions. Actinomycin c_1 was added to indicated reactions during the primer extension steps. B) Immobilized template transcription reactions on *RPS5* template performed as in A). Different combinations of nuclear extract (NE) and whole cell extract (WCE) were used to form PICs. Purified transcription factors (TBP, TFIIB, TFIIF, TFIIE, TFIIH, TFs) were added to reaction 8 during PIC formation (before washing PICs). Purified TFs and Tap-purified Mediator (Med) were included in the PIC formation mix for reaction 9. C) Immobilized template transcription reactions on *HIS4* template performed as in A) and B), using a combination of nuclear extract (NE) and whole cell extract (WCE).

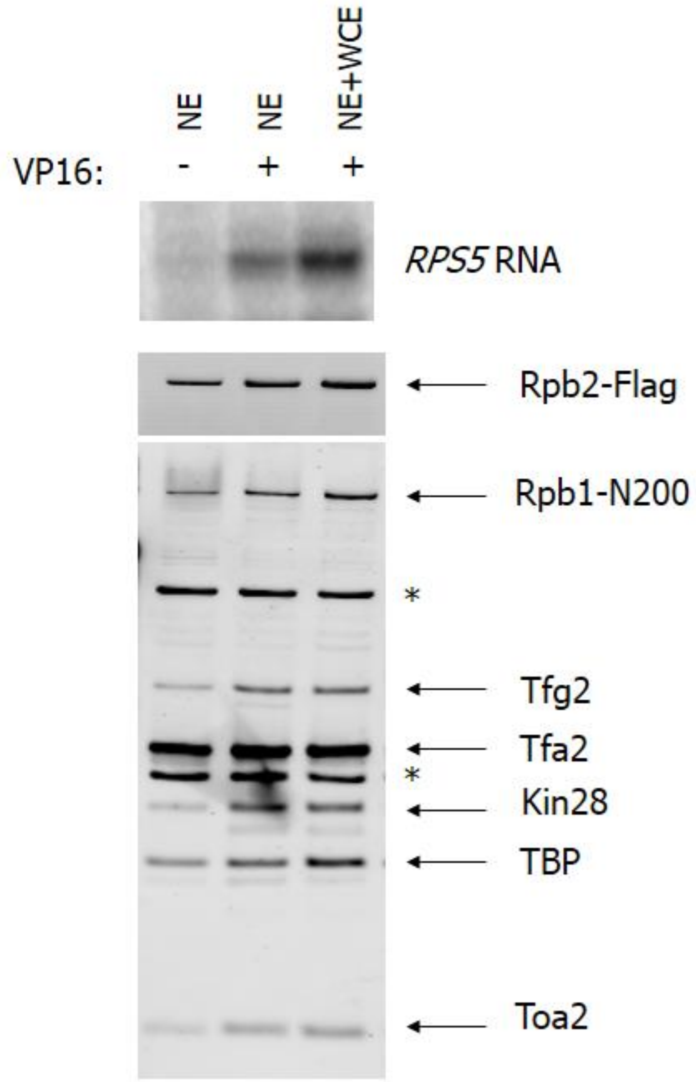


Figure 6 General transcription factors associate with the immobilized *RPS5* promoter

Immobilized PICs were formed on *RPS5* template as described in Figures 1 and 5 and after washing, PICs were split into two equal reactions. One was used to initiate transcription and assayed by primer extension (top panel). The other reaction was digested with PstI to release PICs from the dynabeads and the elutions were analyzed by Western blotting. General transcription factors were probed using polyclonal antibodies, as indicated in the bottom two panels.

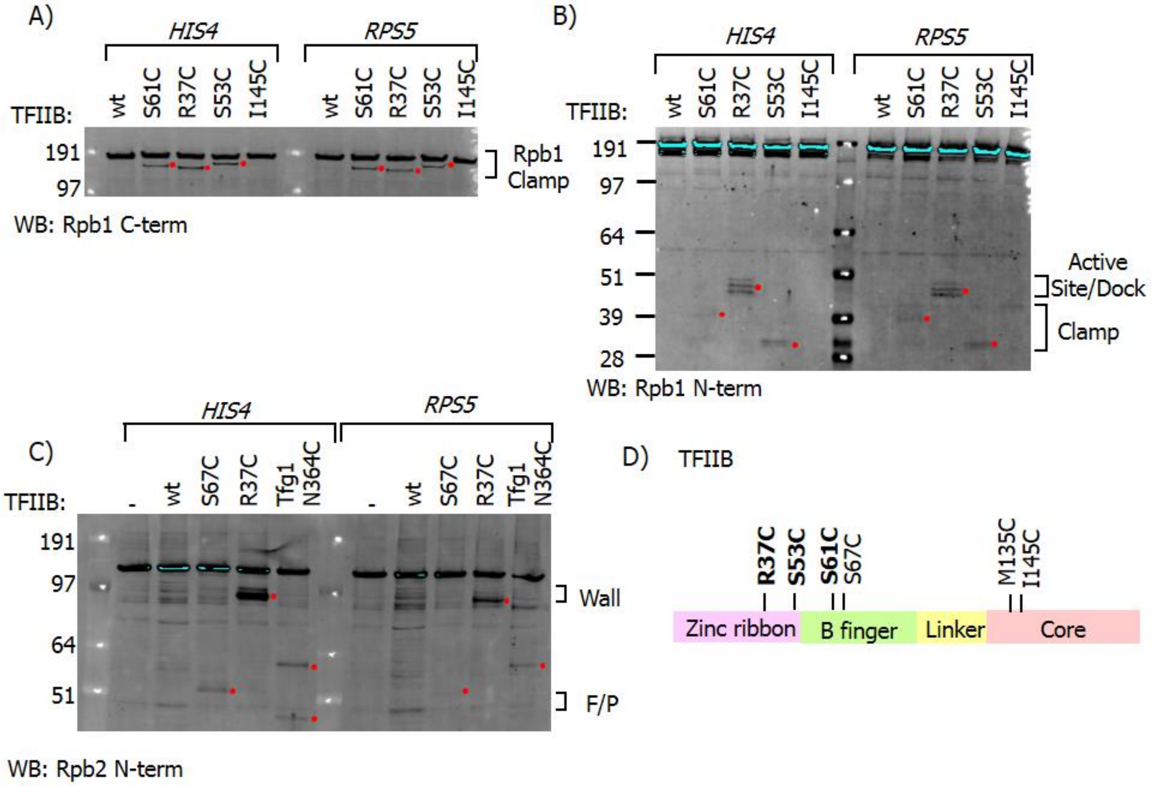
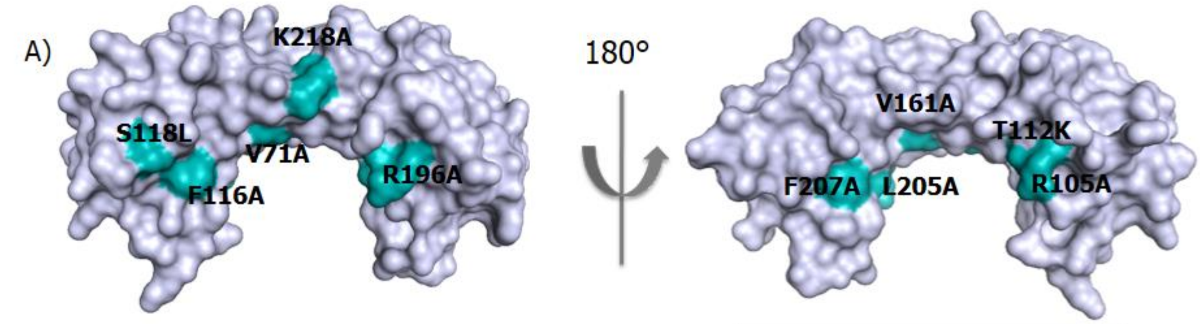
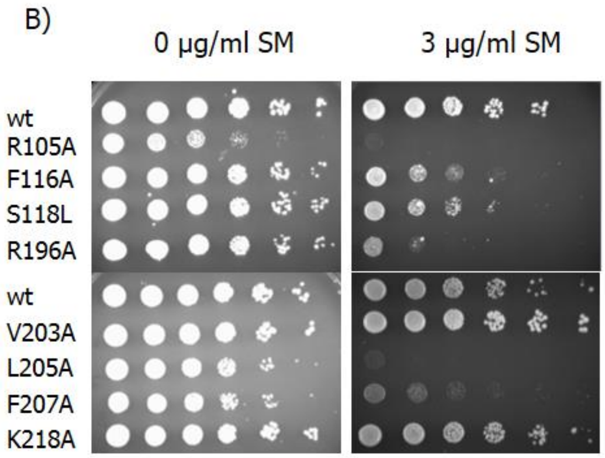


Figure 7 Position of the TFIIB N-terminal region is maintained in the *RPS5* PIC Immobilized PICs were formed on *HIS4* or *RPS5* template as described in previous figures and recombinant TFIIB labelled with FeBABE at the indicated position was included in the PIC reaction. After washing PICs, FeBABE cleavage was performed and the resulting products were eluted from the beads in NuPAGE sample buffer and analyzed by western blotting. A) Cleavage products on Rpb1 detected by probing with monoclonal antibody 8WG16 against the C-terminus of Rpb1. B) Cleavage products on Rpb1 detected by polyclonal antibody against the Rpb1 N-terminus. C) Cleavage on the N-terminus of Rpb2, detected by monoclonal Flag antibody. For *HIS4* PICs, a nuclear extract from an N-terminally Flag-tagged Rpb2 strain was used. For *RPS5* PICs, a combination of nuclear and whole cell extracts from N-terminally Flag-tagged Rpb2 strain was used.





TBP mutant	0 μg/ml SM	3 μg/ml SM
wt	++++	++++
V71A	-	-
R105A	++	-
T112K	-	-
F116A	++++	++
S118L	++++	++
V161A	-	-
R196A	++++	+
V203A	++++	++++
L205A	+++	-
F207A	+++	+
K218A	++++	++++

Figure 8 Phenotypes of TBP DNA-binding mutants
A) Mutations made along the TBP DNA-binding surface are shown in cyan on a 3D model generated from the x-ray structure of *S. cerevisiae* TBP (1YTB) using PyMol. B) Strains expressing the indicated TBP mutants as the only source of TBP were grown to saturation and spotted as five-fold serial dilutions on glucose complete plates lacking the appropriate amino acids, with or without 3 μg/ml SM.

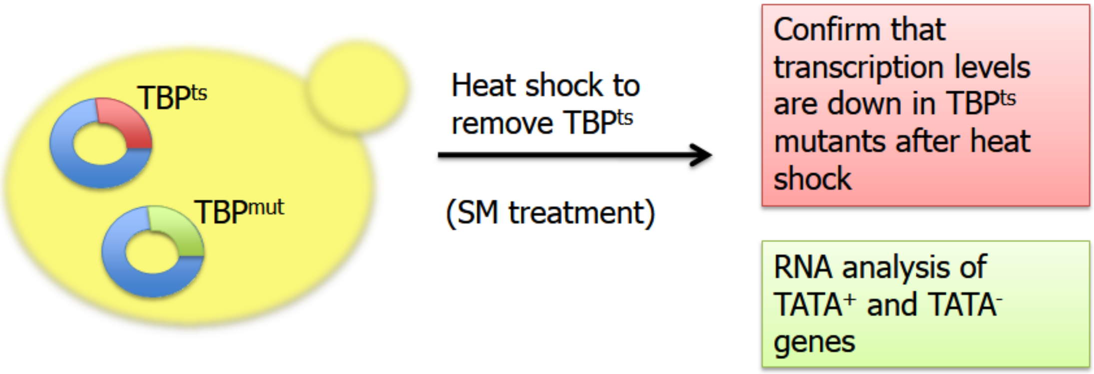


Figure 9 Strategy for *in vivo* analysis of the transcription effects of TBP mutants
TBP DNA-binding mutants are transformed in a cell expressing a temperature-sensitive mutant of TBP (T111I). Yeast are shifted to 37°C to remove TBP^{ts}, while simultaneously treated with SM to induce expression of Gcn4-dependent genes. At the 90 min time point, cells are harvested for RNA and protein analysis.

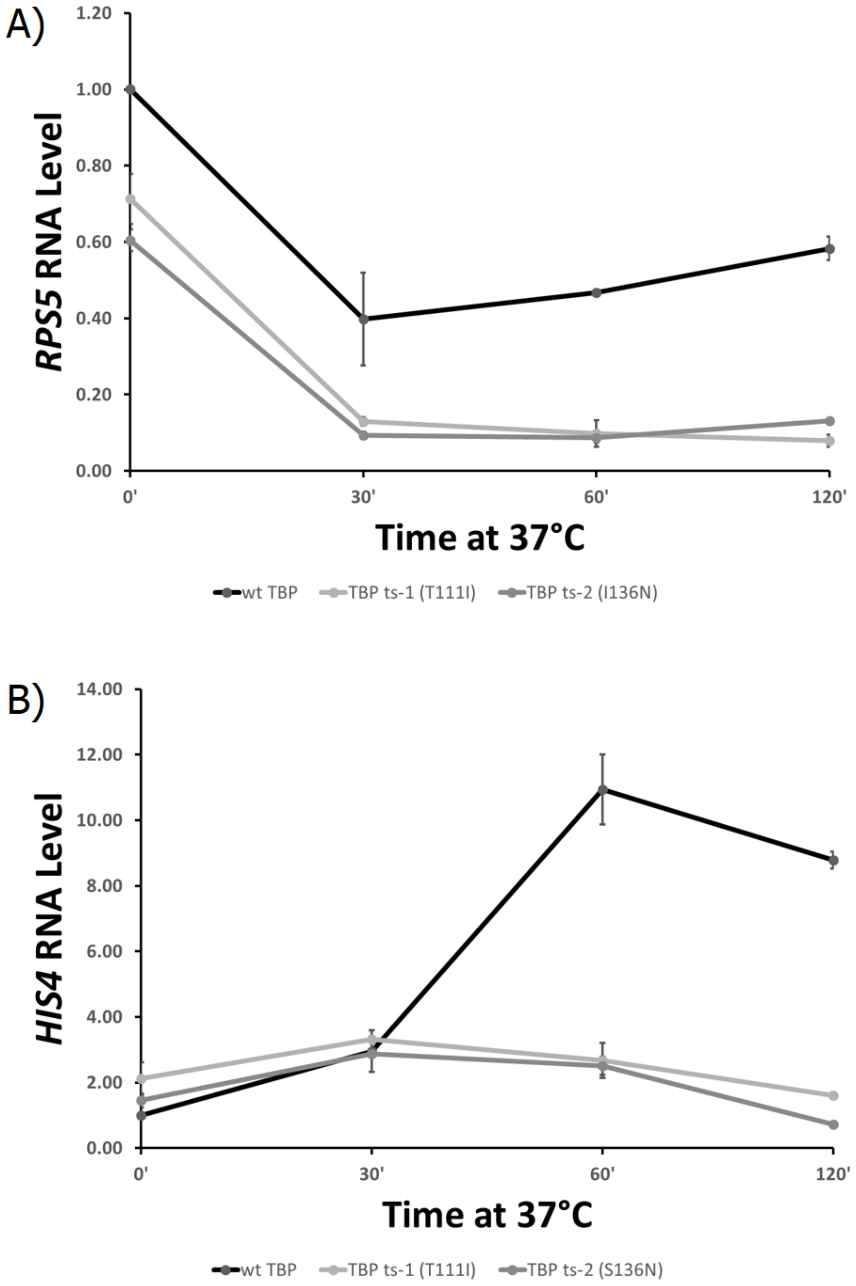


Figure 10 TBP is required for transcription from TATA-less and TATA-containing genes *in vivo*

Time course expression analysis of cells incubated at 37°C. RNA levels were analyzed by RT-qPCR and normalized to the level of the wild type TBP strain at time zero. Results are from biological duplicates and error bars represent standard deviation from the mean. A) *RPS5* levels from wild type and TBP^{ts} mutant T111I (ts-1) and TBP^{ts} mutant TS136N (ts-2). B) *HIS4* levels of uninduced cells.

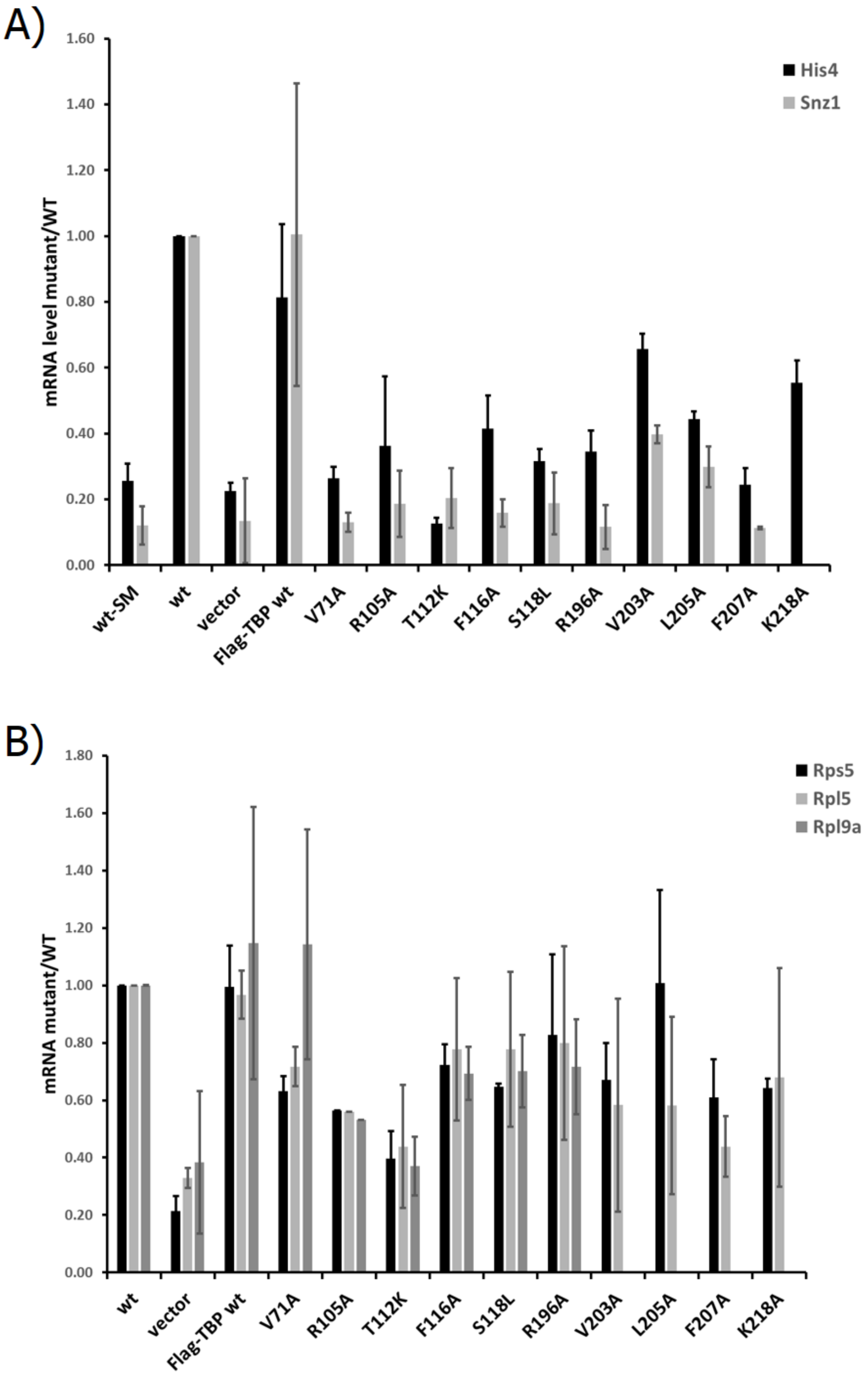

A)
mRNA level mutant/WT
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
His4
Snz1
wt-SM
wt
vector
Flag-TBP wt
V71A
R105A
T112K
F116A
S118L
R196A
V203A
L205A
F207A
K218A
B)
mRNA mutant/WT
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
Rps5
Rpl5
Rpl9a
wt
vector
Flag-TBP wt
V71A
R105A
T112K
F116A
S118L
R196A
V203A
L205A
F207A
K218A

Figure 11 *In vivo* analysis of the ability of TBP DNA-binding mutants to rescue transcription in TBPts yeast
RNA levels of TATA-less and TATA-containing genes following heat shock to inactivate the temperature-sensitive allele of TBP (TBP T111I)). Yeast co-expressing TBP T111I and a TBP DNA-binding mutant were grown to log phase at the permissive temperature of 25°C. Upon reaching the desired OD, SM was added to 3 µg/ml SM and cells were shifted to 37°C for 90 min. Cells were harvested for RNA and protein expression analysis. Expression was quantified by RT-qPCR and normalized to the wild type control. A) TATA-driven Gcn4-dependent genes *HIS4* and *SNZ1*. B) TATA-less ribosomal protein genes *RPS5*, *RPL5*, and *RPL9a*. All experiments were performed in biological duplicates and error bars represent standard deviation from the mean.

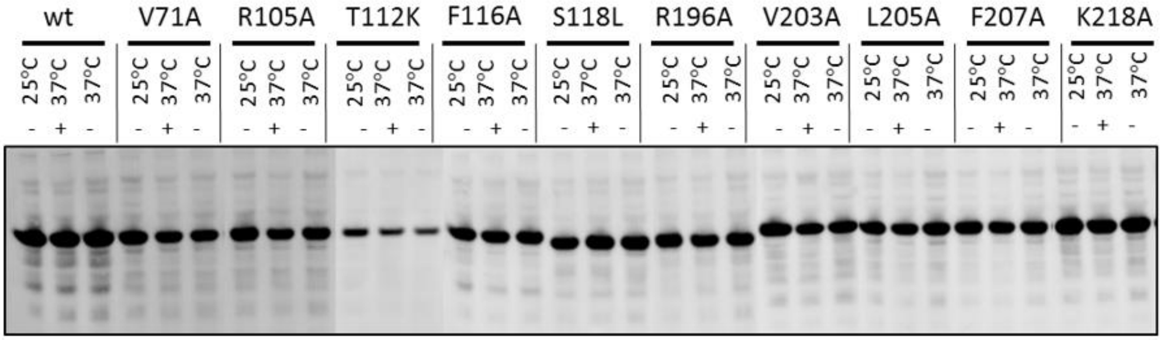


Figure 12 Expression levels of TBP mutants at different temperatures
Protein expression levels of TBP DNA-binding mutants grown as described in Figure 9 before and after 90 min of heat shock. Mutant TBP levels were detected by probing with Flag antibody against N-terminally tagged TBP.

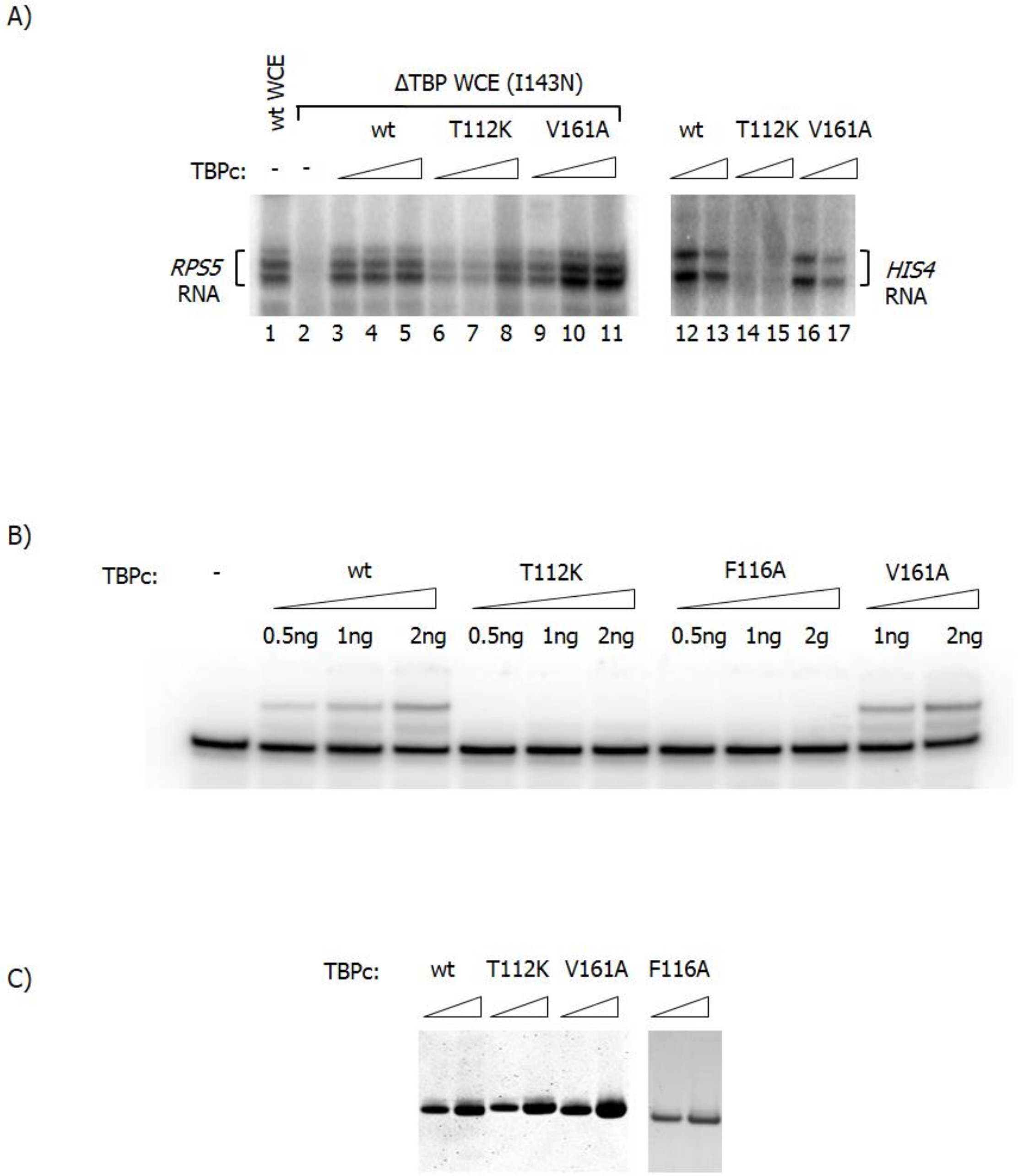


Figure 13 TBP DNA-binding mutants rescue transcription on the *RPS5* promoter
In vitro transcription assay with plasmid *RPS5* and *HIS4* templates. For *RPS5*, a whole cell extract from a temperature sensitive strain harboring the mutation I143N in TBP was supplemented with the indicated TBP mutants, which were purified as recombinant proteins. B) Gel shift assay with a radiolabelled double stranded probe containing the *HIS4* TATA box and the indicated recombinant TBP mutants, all in the context of core TBP (TBPc, Δaa 2-60). C) Coomassie-stained SDS PAGE gel shows purified recombinant TBP mutants were used in transcription and gel shift experiments.

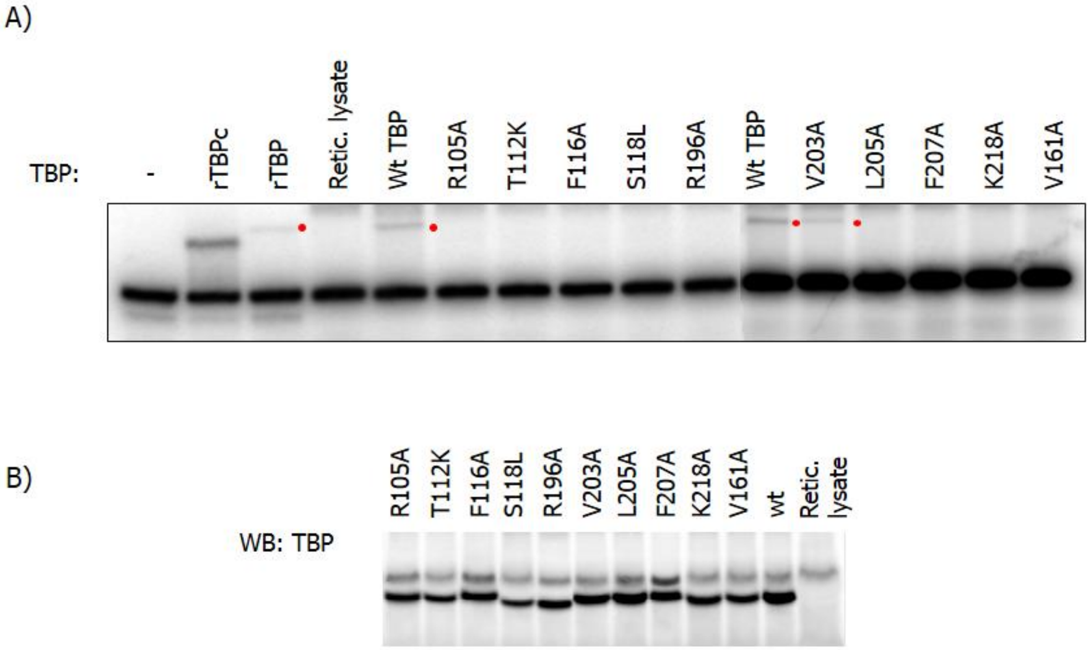


Figure 14 *In vitro* translated TBP DNA-binding mutants fail to bind the *HIS4* TATA box
A) Gel shift assay with a double stranded *HIS4* TATA box probe and TBP DNA-binding mutants generated by *in vitro* translation. Recombinant core wild type TBP (TBPc) and wild type full length TBP (rTBP) serve as positive controls. All TBP mutants are in the context of the full length protein. The position of the TBP-DNA complex is marked by a red dot. B) Western blot probed with polyclonal TBP antibody shows *in vitro* translated TBP mutants.

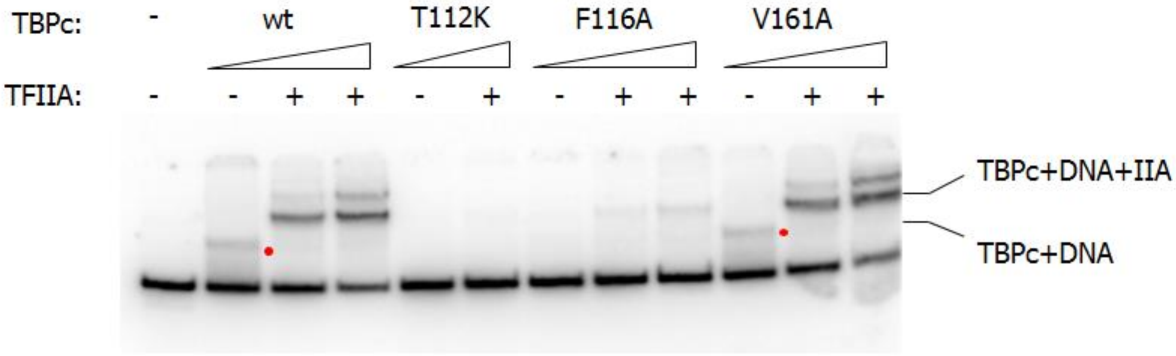


Figure 15 TBP DNA-binding mutants fail to bind the *RPS5* promoter
Gel shift assay with a double stranded *RPS5* promoter probe and indicated recombinant TBP mutants (described in Figure 13). TBPc-DNA complex is marked by a red dot.

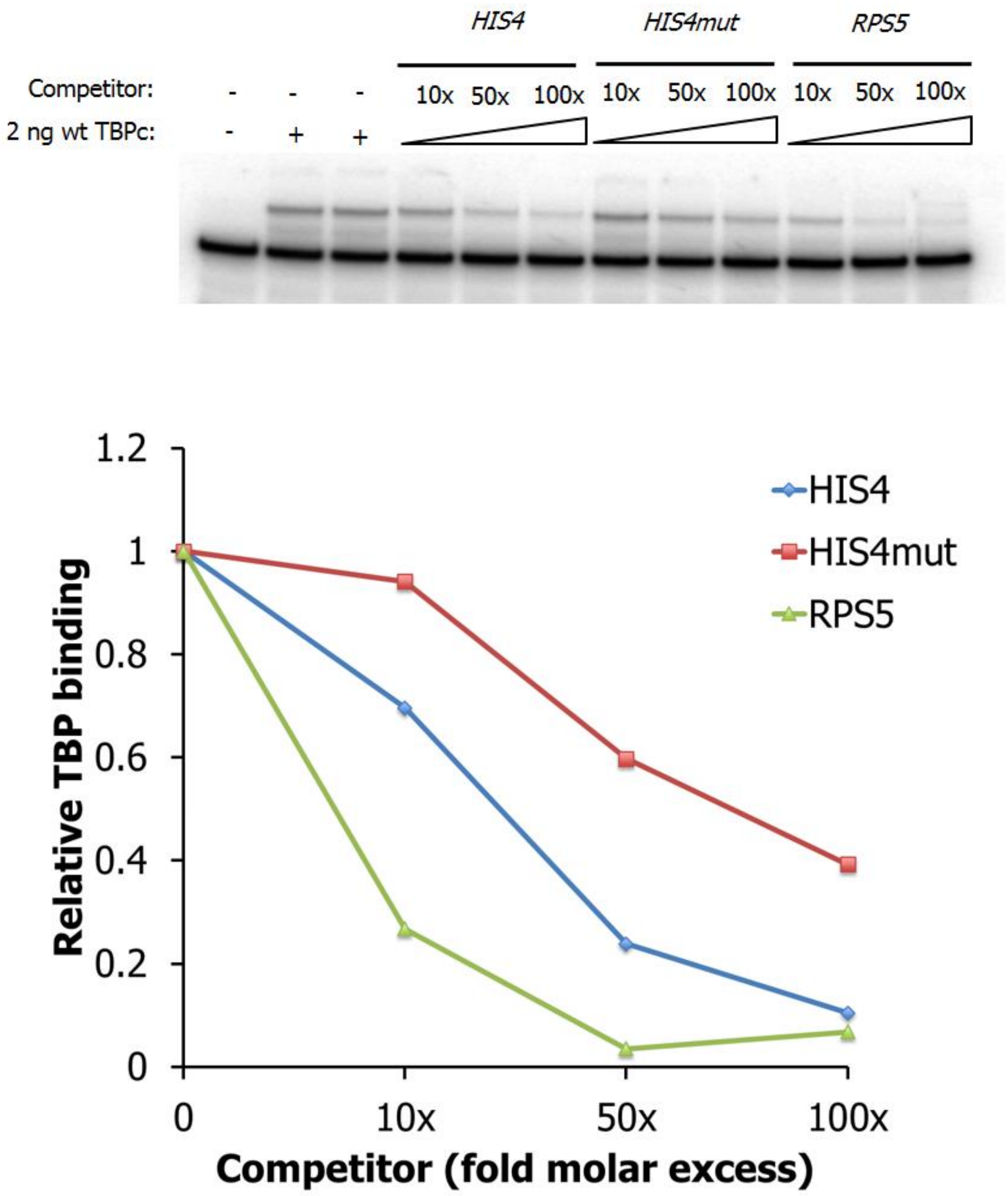


Figure 16 *RPS5* promoter probe competes with *HIS4* for binding to TBP
Gel shift assay using radio-labelled *HIS4* TATA box probe and wild type TBP core. Incremental amounts of competitor unlabeled promoter probe, ranging from a 10 to 100 fold molar excess, were included in the reactions as shown. Competitor probes were respectively wild type *HIS4* TATA box DNA (*HIS4*), *HIS4* with a mutated TATA box (*HIS4*mut), and wild type *RPS5* promoter. The results were quantified using ImageQuant software and normalized to the level of TBP binding in the absence of competitor.

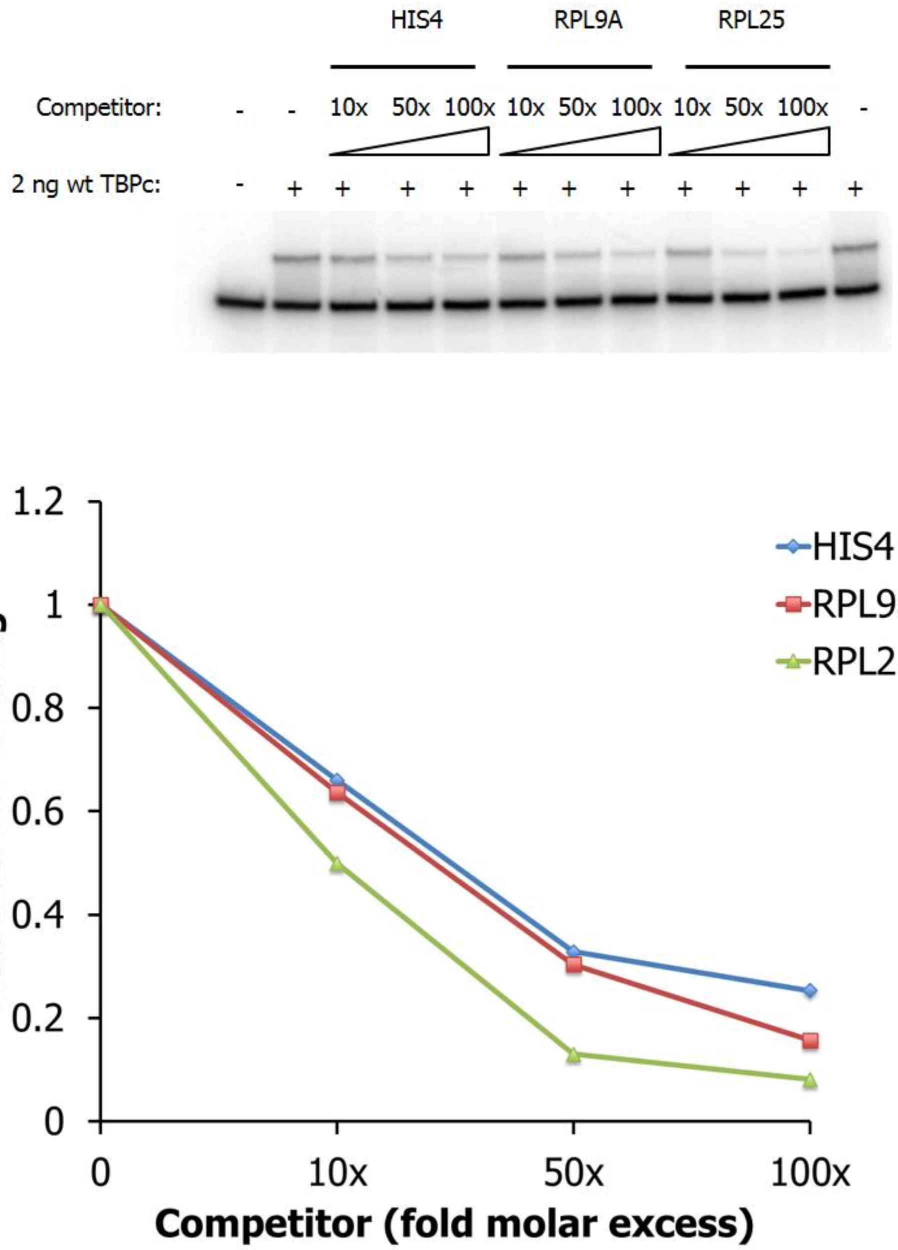


Figure 17 Different TATA-less promoters compete with *HIS4* TATA box for TBP binding Gel shift assay using radio-labelled *HIS4* TATA box probe and wild type TBP core and unlabeled competitor probes. The *RPL9a* and *RPL25* promoters were used as unlabeled competitors, along with the wild type *HIS4* TATA box fragment.

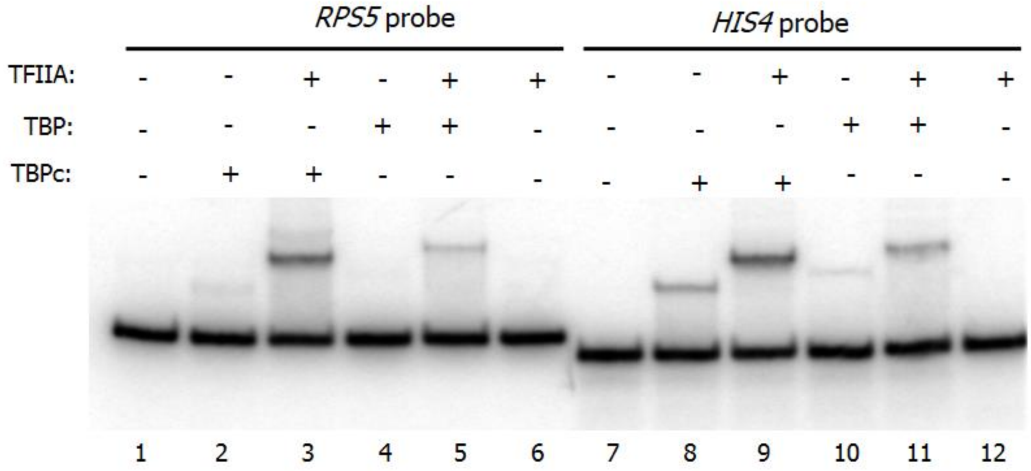


Figure 18 TBP binds to the *RPS5* promoter and forms a complex with TFIIA
Gel shift assay using a radio-labelled *RPS5* or *HIS4* promoter probe and recombinant full length TBP or TBP core (TBPc). Recombinant TFIIA was included as indicated.

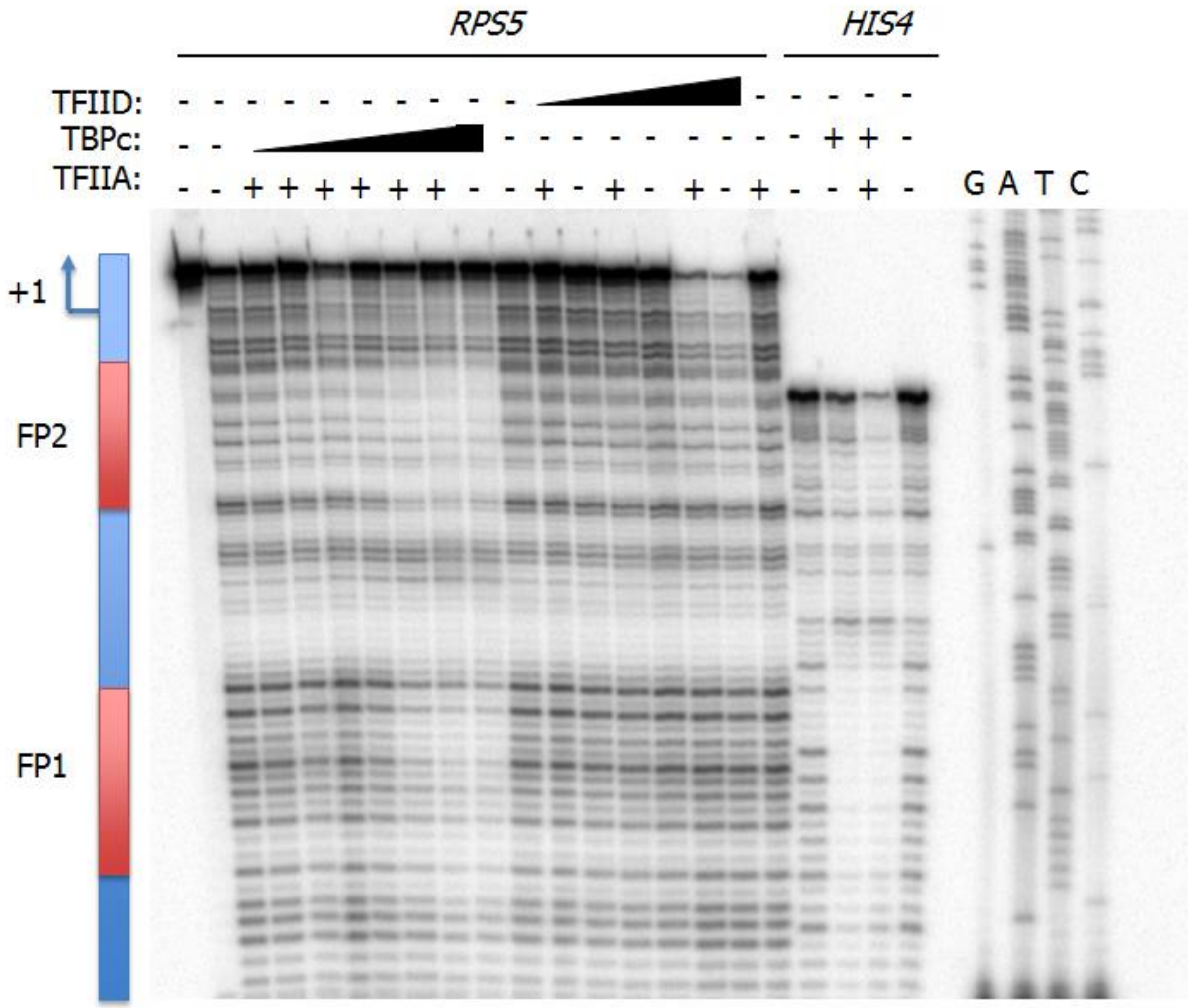

RPS5
HIS4
TFIID: - - - - - - - - - - - - - - -
TBPc: - - - - - - - - - - - - + + -
TFIIA: - - + + + + + + - - + - + - + - + - - + -
G A T C
+1
FP2
FP1

Figure 19 TBP binds to two sites on the *RPS5* promoter

DNase I footpinting assay performed with a 5’ radiolabeled *RPS5* promoter DNA (same as used in gel shift assays) and recombinant core TBP with or without TFIIA. Flag-purified TFIID with or without TFIIA was used in place of TBP where indicated. The *HIS4* TATA box-containing promoter probe, also labeled on the 5’ end, was included as a positive control for TBP binding. A sequencing ladder generated from the *RPS5* promoter is included on the right. A schematic diagram if the *RPS5* promoter with TBP binding sites shown as red boxes is displayed to the left of the DNase I footprint. FP1 refers to the upstream TBP binding site, FP2 to the downstream TBP binding site, and +1 to the transcription start site.

RPS5 promoter mutants

wt	TCTAAAATATTTTGTACTTTTTATCAATACTTAAAATTTCATTTTGAATTAAAATTACTTTTTATTTTAATTACACTCCTTT
ΔFP1	Δ8bp TCTAAAATATTTTGTACTTTTTATCAATACTTAAAATTTCATTTTGAATTAAAATTACTTTTTATTTTAATTACACTCCTTT
ΔFP2	Δ8bp TCTAAAATATTTTGTACTTTTTATCAATACTTAAAATTTCATTTTGAATTAAAATTACTTTTTATTTTAATTACACTCCTTT
ΔFP1ΔFP2	Δ8bp Δ8bp TCTAAAATATTTTGTACTTTTTATCAATACTTAAAATTTCATTTTGAATTAAAATTACTTTTTATTTTAATTACACTCCTTT
FP1mut	TCTAAAATATTTTGTACTTCCTCGAGATACTTAAAATTTCATTTTGAATTAAAATTACTTTTTATTTTAATTACACTCCTTT
FP2mut	TCTAAAATATTTTGTACTTTTTATCAATACTTAAAATTTCATTTTGAATTAAAATTACCTCCGATGTTAATTACACTCCTTT
FP1mutFP2mut	TCTAAAATATTTTGTACTTCCTCGAGATACTTAAAATTTCATTTTGAATTAAAATTACCTCCGATGTTAATTACACTCCTTT

Figure 20 ***RPS5*** **mutant promoter sequences**
Sequence of wild type *RPS5* promoter with TBP binding sites highlighted in red. Below the wild type sequence, a series of deletion and substitution mutants are displayed, with the mutated sequence boxed. FP1 and FP2 refer to the upstream and downstream TBP binding sites, respectively, as defined by the DNase I footprinting experiment in Figure 19.

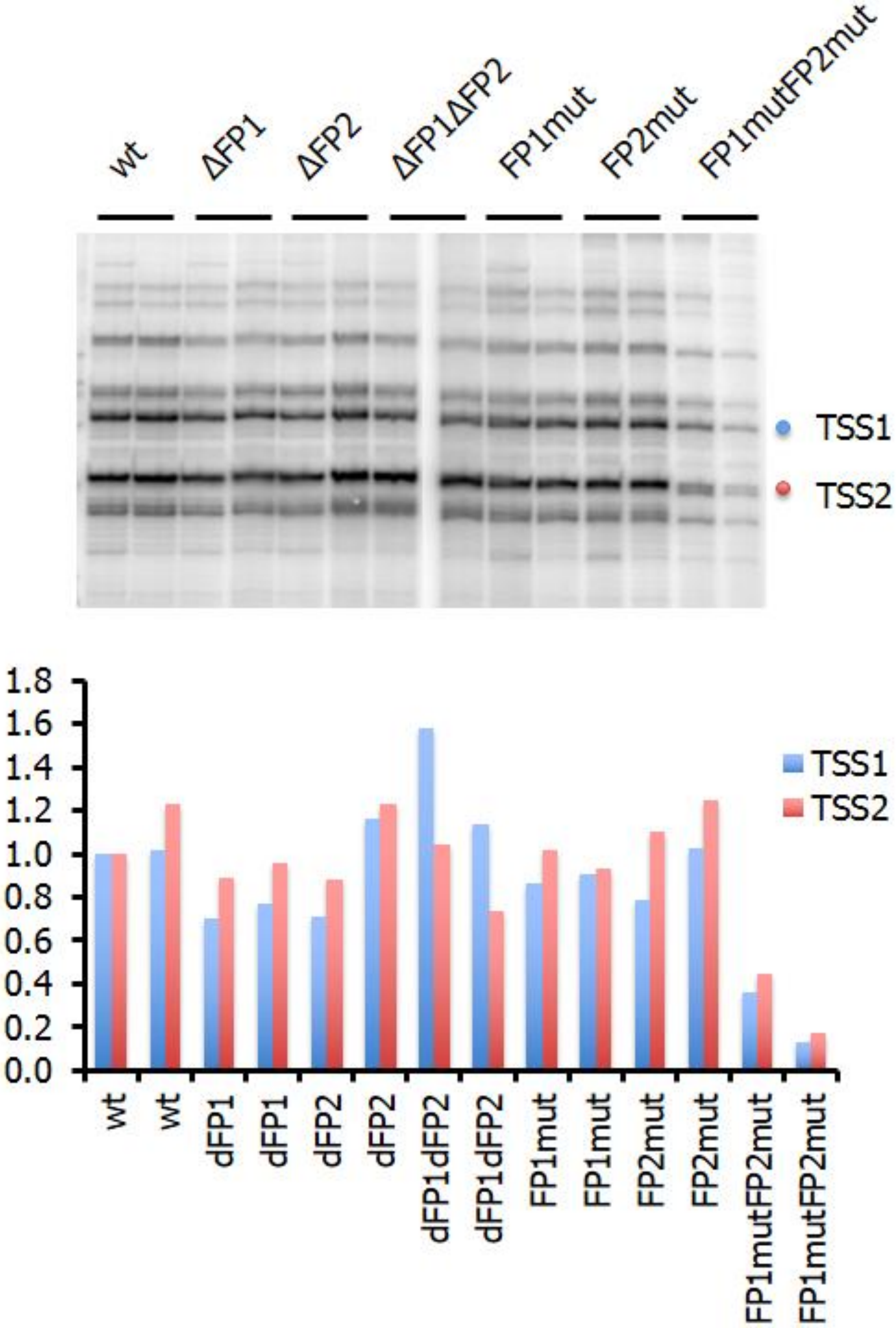


Figure 21 Deletion of TBP bindings sites from *RPS5* promoter does not affect *in vitro* transcription

RPS5 mutant promoter constructs (described in Figure 18) were used for *in vitro* plasmid transcription, the products assayed by primer extension and resolved on a high resolution sequencing gel. The major start sites are marked by a red dot (TSS2, downstream site) and a blue dot (TSS1, upstream site) and quantified using ImageQuant software. Each mutant construct was tested in duplicate and the results are expressed as a fraction of the wild type transcription levels.

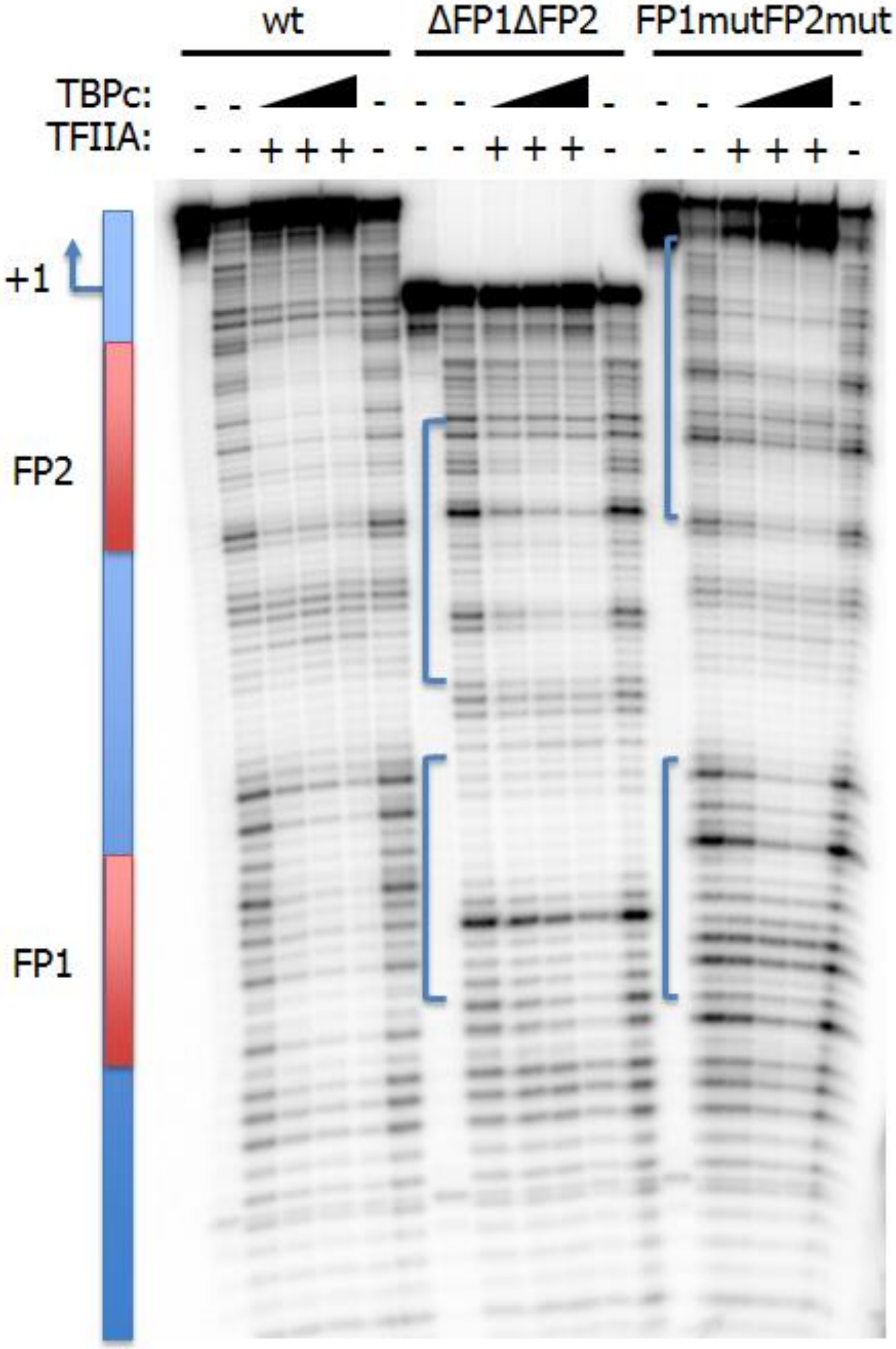


Figure 22 TBP binding to mutant *RPS5* promoter probes is altered
DNase I footprinting assay as described in Figure 19. The *RPS5* probes used are wild type (wt), both TBP binding sites deleted (ΔFP1ΔFP2), and both TBP binding sites mutated (FP1mutFP2mut). TBP binding to mutant *RPS5* probes in marked by blue brackets.

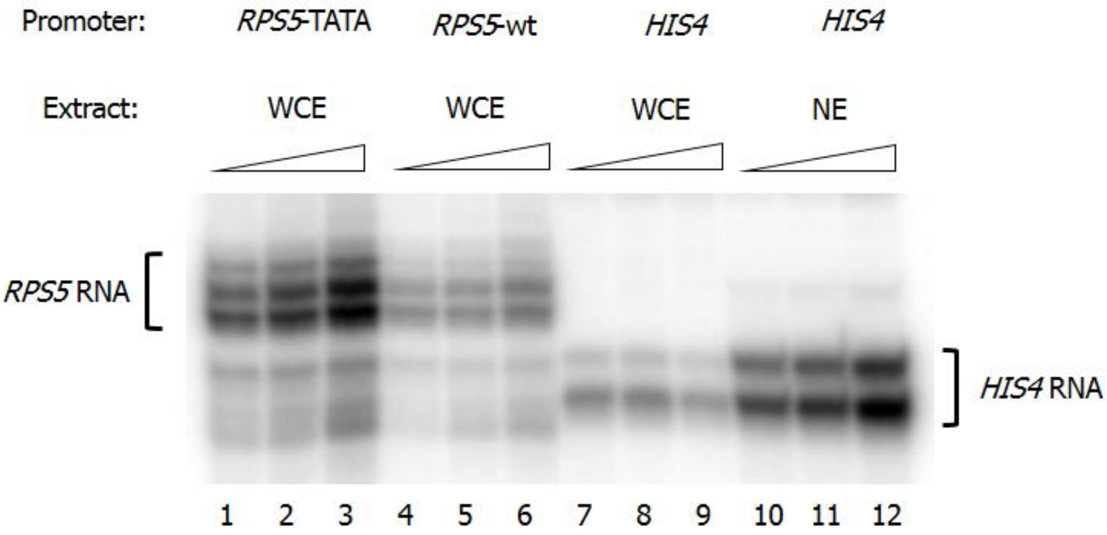


Figure 23 Insertion of a canonical TATA box in the *RPS5* promoter results in higher transcription levels

In vitro transcription assay with *RPS5* promoter variants and *HIS4* promoter, performed as described in previous figures. The *RPS5*-TATA mutant has a single C→A mutation, resulting in a canonical TATA box TATAAAAT. Lanes 7-12 explore the ability of whole cell extract (WCE) or nuclear extract (NE) to transcribe the *HIS4* promoter.

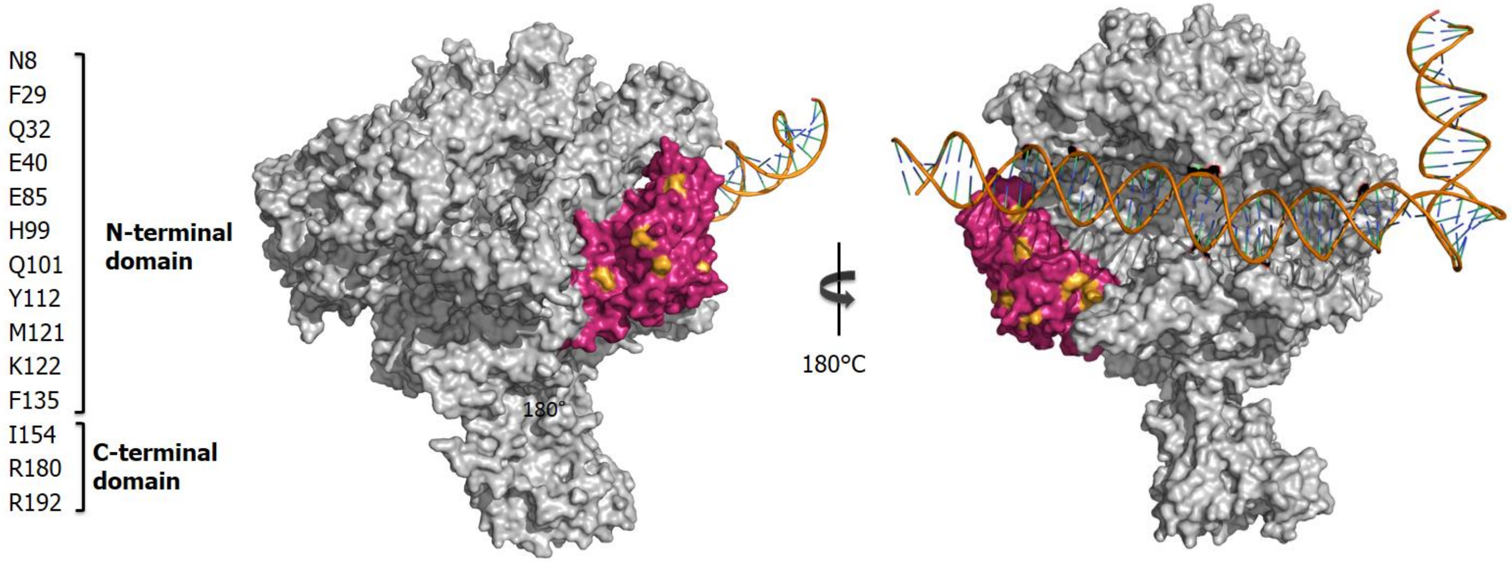


Figure 24 Sites of Bpa incorporation on the Rpb5 surface
The surface-exposed residues shown on the left in the N- and C-terminal domains of Rpb5 were mutated to amber stop codons. Yeast strains were generated in which each mutated residue was suppressed with Bpa, resulting in site-specific incorporation of the photocrosslinker. Sites of Bpa attachment are displayed in orange and Rpb5 is pink, shown in the context of yeast RNA Polymerase II.

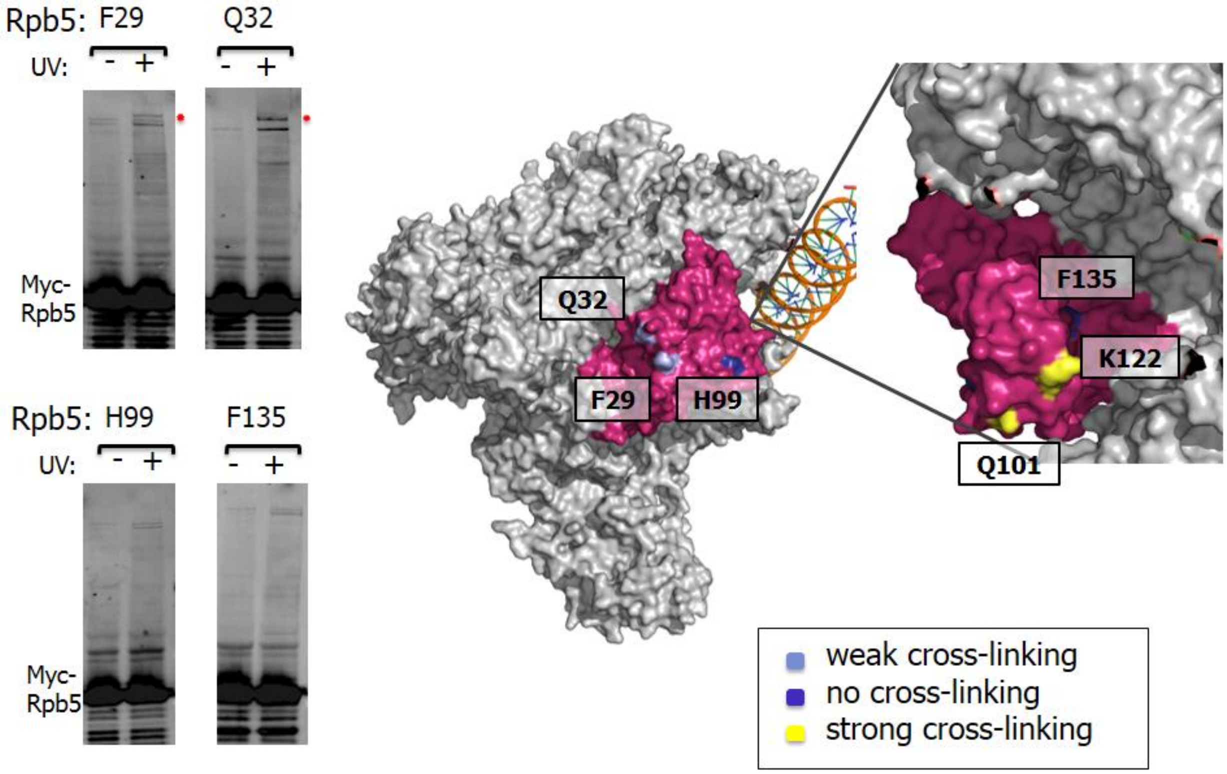


Figure 25 Rpb5 crosslinking results in the PIC
Western blots of PICs formed using Rpb5-Bpa nuclear extracts. The residue of Bpa attachment is indicated above each panel. After washing unbound proteins, each PIC reaction was split in two and one half was treated with UV, and the other used as a negative control. Blots were probed with α-Myc antibody, which recognizes a 13xMyc tag on the N-terminus of Rpb5. The Rpb5-crosslinked fusion proteins are marked with a red dot. The right panel displays the sites of Bpa incorporation in the context of Pol II (grey) and Rpb5 in pink.

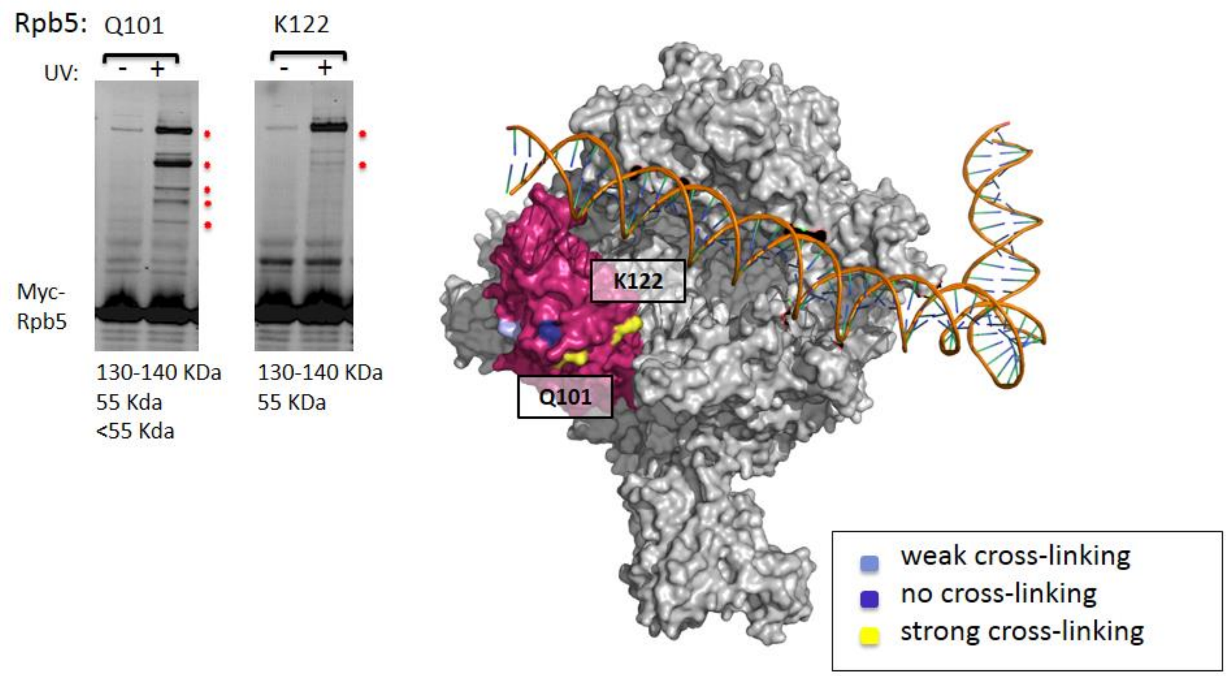


Figure 26 Rpb5 crosslinking results in the PIC with residues Q101-bpa and K122-bpa
Western blots of PICs formed using Rpb5-Bpa nuclear extracts. The residue of Bpa attachment is indicated above each panel. After washing unbound proteins, each PIC reaction was split in two and one half was treated with UV, and the other used as a negative control. Blots were probed with α-Myc antibody, which recognizes a 13xMyc tag on the N-terminus of Rpb5. The Rpb5-crosslinked protein fusion proteins are marked with a red dot. The right panel displays the sites of Bpa incorporation in the context of Pol II (grey) and Rpb5 in pink.

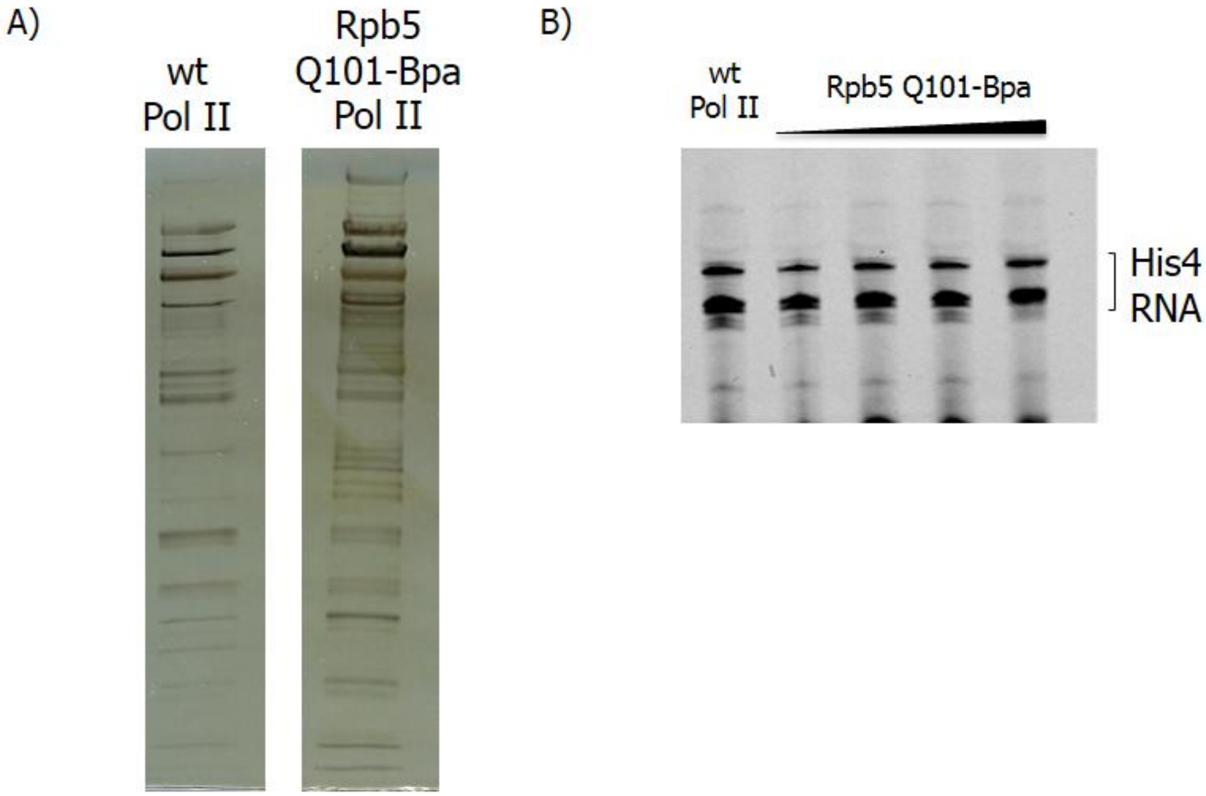


Figure 27 Pol II: Rpb5 Q101-Bpa purification and transcription activity
A) Silver stained gel of RNA Poll II purified from yeast via 6xHIS tag on Rpb3. Wild type Pol II is shown on the left, and Pol II from Rpb5 Q101-amber strain grown in the presence of Bpa is shown on the right. B) *In vitro* transcription reaction using *HIS4* promoter plasmid template and purified transcription factors. Wild type Pol II is used as a control and the activity of a titration of Pol II: Rpb5 Q101bpa is tested.

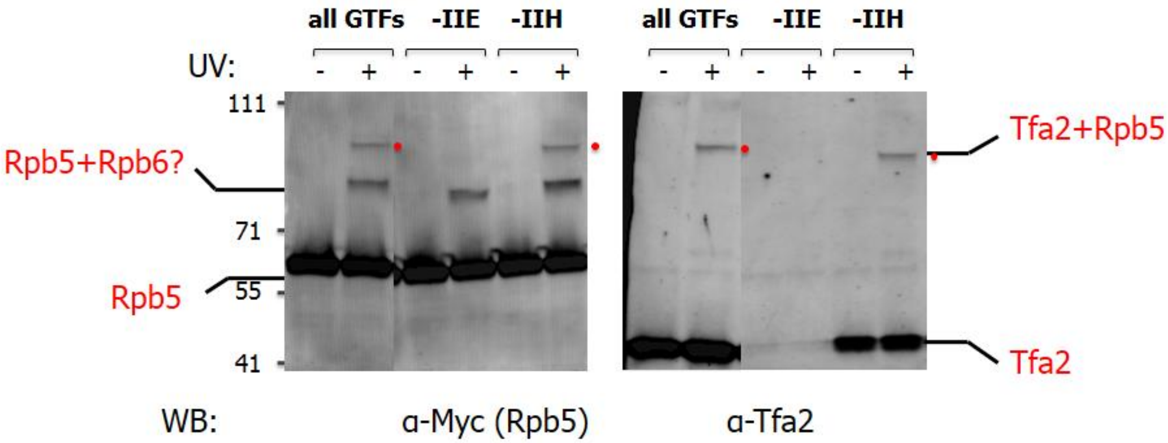


Figure 28 Rpb5 Q101 cross-links to the small subunit of TFIIE
Western blots of PICs formed using purified general transcription factors (GTFs: TBP, TFIIB, TFIIF, TFIIE, TFIIH, Pol II). The Pol II used in these reactions contains Rpb5 Q101bpa (shown on Figure 27). After washing unbound proteins, each PIC reaction was split in two and one half was treated with UV, and the other used as a negative control. Some GTFs are omitted from the PIC reactions as indicated. The blot on the left was probed with mouse α-Myc antibody, which recognizes a 13xMyc tag on the N-terminus of Rpb5. The left right panel represents the same membrane probed with rabbit α-Tfa2 polyclonal antibody. The Tfa2-Rpb5 crosslinked product is marked with a red dot.

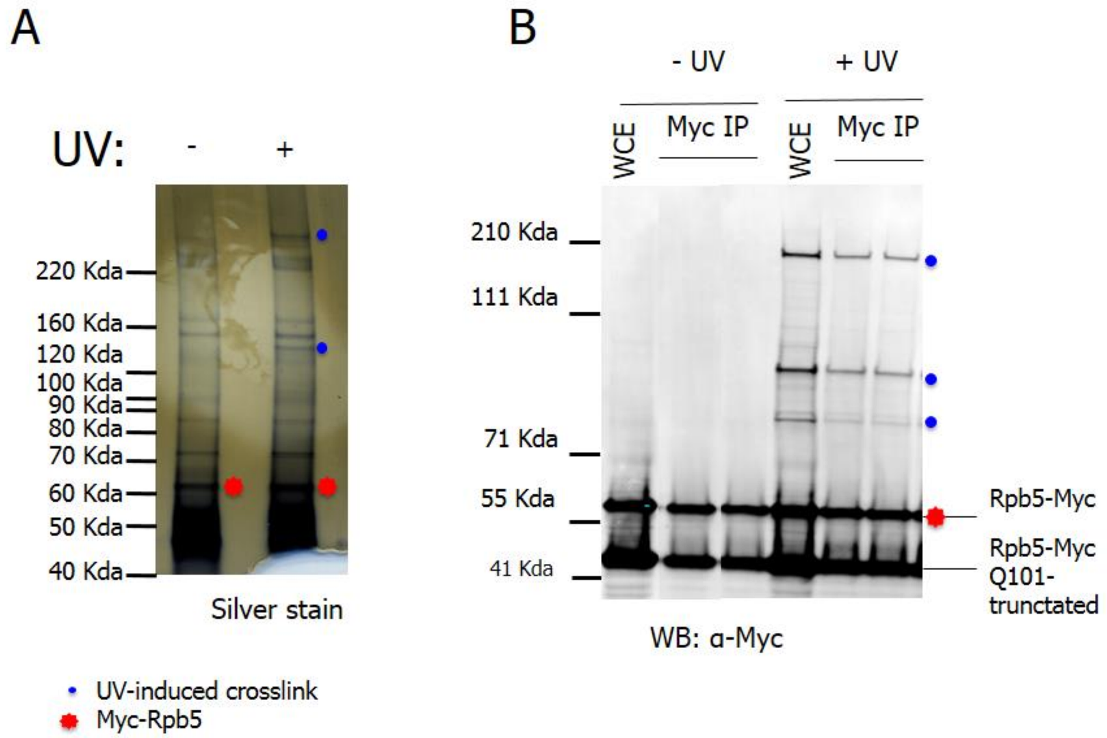


Figure 29 IP of Myc-tagged Rpb5 from whole cell extracts of *in vivo* crosslinked yeast containing Rpb5 Q101 Bpa
Yeast strain expressing Rpb5:Q101-amber was grown in Bpa containing media until log phase. Upon harvesting, half of the cells were treated with UV and whole cell extracts were prepared from these cells and the –UV control. Rpb5 along with any crosslinked polypeptides was immunoprecipitated using Myc-agarose beads and eluted in NuPAGE sample buffer. A) Silver stained gel of Rpb5 immunoprecipitated from yeast whole cell extracts via its N-terminal Myc tag. 20 % of the elutions was run on 3-8% Tris-acetate gel. B) 10% of the elutions was run on 3-8% Tris-acetate gel and probed with α-Myc antibody. Myc-Rpb5 is marked with a red dot, crosslinked peptides are marked with blue dots.

REFERENCES

Akiyoshi, B., Nelson, C. R., Ranish, J. A. & Biggins, S. 2009. Quantitative proteomic analysis of purified yeast kinetochores identifies a PP1 regulatory subunit. *Genes Dev,* 23**,** 2887-99.

Anish, R., Hossain, M. B., Jacobson, R. H. & Takada, S. 2009. Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. *PLoS One,* 4**,** e5103.

Basehoar, A. D., Zanton, S. J. & Pugh, B. F. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell,* 116**,** 699-709.

Bernecky, C., Grob, P., Ebmeier, C. C., Nogales, E. & Taatjes, D. J. 2011. Molecular architecture of the human Mediator-RNA polymerase II-TFIIF assembly. *PLoS Biol,* 9**,** e1000603.

Bhaumik, S. R. & Green, M. R. 2001. SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes Dev,* 15**,** 1935-45.

Bhaumik, S. R., Raha, T., Aiello, D. P. & Green, M. R. 2004. In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev,* 18**,** 333-43.

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.

Blair, R. H., Goodrich, J. A. & Kugel, J. F. 2012. Single-molecule fluorescence resonance energy transfer shows uniformity in TATA binding protein-induced DNA bending and heterogeneity in bending kinetics. *Biochemistry,* 51**,** 7444-55.

Burke, T. W. & Kadonaga, J. T. 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev,* 10**,** 711-24.

Burke, T. W. & Kadonaga, J. T. 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. *Genes Dev,* 11**,** 3020-31.

Burke, T. W., Willy, P. J., Kutach, A. K., Butler, J. E. & Kadonaga, J. T. 1998. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harb Symp Quant Biol,* 63**,** 75-82.

Butler, J. E. & Kadonaga, J. T. 2001. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev,* 15**,** 2515-9.

Butler, J. E. & Kadonaga, J. T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev,* 16**,** 2583-92.

Cai, G., Chaban, Y. L., Imasaki, T., Kovacs, J. A., Calero, G., Penczek, P. A., Takagi, Y. & Asturias, F. J. 2012. Interaction of the mediator head module with RNA polymerase II. *Structure,* 20**,** 899-910.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A., Taylor, M. S., Engstrom, P. G., Frith, M. C., Forrest, A. R., Alkema, W. B., Tan, S. L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S. M., Wells, C. A., Orlando, V., Wahlestedt, C., Liu, E. T., Harbers, M., Kawai, J., Bajic, V. B., Hume, D. A. & Hayashizaki, Y. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet,* 38**,** 626-35.

Chen, H. T. & Hahn, S. 2003. Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol Cell,* 12**,** 437-47.

Chen, H. T. & Hahn, S. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell,* 119**,** 169-80.

Chen, H. T., Warfield, L. & Hahn, S. 2007. The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol,* 14**,** 696-703.

Cheng, J. X., Floer, M., Ononaji, P., Bryant, G. & Ptashne, M. 2002. Responses of four yeast genes to changes in the transcriptional machinery are determined by their promoters. *Curr Biol,* 12**,** 1828-32.

Cianfrocco, M. A., Kassavetis, G. A., Grob, P., Fang, J., Juven-Gershon, T., Kadonaga, J. T. & Nogales, E. 2013. Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell,* 152**,** 120-31.

Cormack, B. P. & Struhl, K. 1992. The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell,* 69**,** 685-96.

Cramer, P., Armache, K. J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G. E., Dengl, S., Geiger, S. R., Jasiak, A. J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C. D., Lehmann, E., Leike, K., Sydow, J. F. & Vannini, A. 2008. Structure of eukaryotic RNA polymerases. *Annu Rev Biophys,* 37**,** 337-52.

Dejardin, J. & Kingston, R. E. 2009. Purification of proteins associated with specific genomic Loci. *Cell,* 136**,** 175-86.

Deng, W. & Roberts, S. G. 2005. A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev,* 19**,** 2418-23.

Deng, W. & Roberts, S. G. 2006. Core promoter elements recognized by transcription factor IIB. *Biochem Soc Trans,* 34**,** 1051-3.

Deng, W. & Roberts, S. G. 2007. TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma,* 116**,** 417-29.

Dynlacht, B. D., Hoey, T. & Tjian, R. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell,* 66**,** 563-76.

Eichner, J., Chen, H. T., Warfield, L. & Hahn, S. 2010. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J,* 29**,** 706-16.

Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van Berkum, N., Boschiero, C., Holstege, F. & Werner, M. 2008. Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Mol Cell,* 31**,** 337-46.

Fishburn, J., Mohibullah, N. & Hahn, S. 2005. Function of a eukaryotic transcription activator during the transcription cycle. *Mol Cell,* 18**,** 369-78.

Fishburn, J. & Hahn, S. 2012. Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol Cell Biol,* 32**,** 12-25.

Fuda, N. J., Ardehali, M. B. & Lis, J. T. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature,* 461**,** 186-92.

Garbett, K. A., Tripathi, M. K., Cencki, B., Layer, J. H. & Weil, P. A. 2007. Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. *Mol Cell Biol,* 27**,** 297-311.

Geiger, J. H., Hahn, S., Lee, S. & Sigler, P. B. 1996. Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science,* 272**,** 830-6.

Gibbons, B. J., Brignole, E. J., Azubel, M., Murakami, K., Voss, N. R., Bushnell, D. A., Asturias, F. J. & Kornberg, R. D. 2012. Subunit architecture of general transcription factor TFIIH. *Proc Natl Acad Sci U S A*, 109**,** 1949-54.

Grunberg, S., Reich, C., Zeller, M. E., Bartlett, M. S. & Thomm, M. 2010. Rearrangement of the RNA polymerase subunit H and the lower jaw in archaeal elongation complexes. *Nucleic Acids Res,* 38**,** 1950-63.

Grunberg, S., Warfield, L. & Hahn, S. 2012. Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nat Struct Mol Biol,* 19**,** 788-96.

Guerrero, C., Tagwerker, C., Kaiser, P. & Huang, L. 2006. An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol Cell Proteomics,* 5**,** 366-78.

Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. 1989. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc Natl Acad Sci U S A,* 86**,** 5718-22.

Hahn, S. & Young, E. T. 2011. Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics,* 189**,** 705-36.

Hall, D. B., Wade, J. T. & Struhl, K. 2006. An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in Saccharomyces cerevisiae. *Mol Cell Biol,* 26**,** 3672-9.

He, Y., Fang, J., Taatjes, D. J. & Nogales, E. 2013. Structural visualization of key steps in human transcription initiation. *Nature,* 495**,** 481-6.

Hsu, J. Y., Juven-Gershon, T., Marr, M. T., 2nd, Wright, K. J., Tjian, R. & Kadonaga, J. T. 2008. TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. *Genes Dev,* 22**,** 2353-8.

Hu, H. & Li, X. 2007. Transcriptional regulation in eukaryotic ribosomal protein genes. *Genomics,* 90**,** 421-3.

Huisinga, K. L. & Pugh, B. F. 2004. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. *Mol Cell,* 13**,** 573-85.

Imasaki, T., Calero, G., Cai, G., Tsai, K. L., Yamada, K., Cardelli, F., Erdjument-Bromage, H., Tempst, P., Berger, I., Kornberg, G. L., Asturias, F. J., Kornberg, R. D. & Takagi, Y. 2011. Architecture of the Mediator head module. *Nature,* 475**,** 240-3.

Iyer, V. & Struhl, K. 1995. Mechanism of differential utilization of the his3 TR and TC TATA elements. *Mol Cell Biol,* 15**,** 7059-66.

Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science*, 288**,** 1422-5.

Jasiak, A. J., Hartmann, H., Karakasili, E., Kalocsay, M., Flatley, A., Kremmer, E., Strasser, K., Martin, D. E., Soding, J. & Cramer, P. 2008. Genome-associated RNA polymerase II includes the dissociable Rpb4/7 subcomplex. *J Biol Chem,* 283**,** 26423-7.

Jorgensen, P., Rupes, I., Sharom, J. R., Schneper, L., Broach, J. R. & Tyers, M. 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev,* 18**,** 2491-505.

Juven-Gershon, T., Cheng, S. & Kadonaga, J. T. 2006a. Rational design of a super core promoter that enhances gene expression. *Nat Methods,* 3**,** 917-22.

Juven-Gershon, T., Hsu, J. Y. & Kadonaga, J. T. 2006b. Perspectives on the RNA polymerase II core promoter. *Biochem Soc Trans,* 34**,** 1047-50.

Juven-Gershon, T., Hsu, J. Y. & Kadonaga, J. T. 2008a. Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. *Genes Dev,* 22**,** 2823-30.

Juven-Gershon, T., Hsu, J. Y., Theisen, J. W. & Kadonaga, J. T. 2008b. The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol,* 20**,** 253-9.

Kadonaga, J. T. 2002. The DPE, a core promoter element for transcription by RNA polymerase II. *Exp Mol Med,* 34**,** 259-64.

Kadonaga, J. T. 2012. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol,* 1**,** 40-51.

Kim, J. L., Nikolov, D. B. & Burley, S. K. 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature,* 365**,** 520-7.

Kim, T. K., Ebright, R. H. & Reinberg, D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science,* 288**,** 1418-22.

Klein, C. & Struhl, K. 1994. Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. *Mol Cell Biol,* 14**,** 1920-8.

Knuesel, M. T., Meyer, K. D., Bernecky, C. & Taatjes, D. J. 2009. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev,* 23**,** 439-51.

Knutson, B. A. & Hahn, S. 2011. Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Mol Cell Biol,* 31**,** 818-31.

Komarnitsky, P. B., Michel, B. & Buratowski, S. 1999. TFIID-specific yeast TAF40 is essential for the majority of RNA polymerase II-mediated transcription in vivo. *Genes Dev,* 13**,** 2484-9.

Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M. & Cramer, P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature,* 462**,** 323-30.

Krishnamurthy, M., Dugan, A., Nwokoye, A., Fung, Y. H., Lancia, J. K., Majmudar, C. Y. & Mapp, A. K. 2011. Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo. *ACS Chem Biol,* 6**,** 1321-6.

Kutach, A. K. & Kadonaga, J. T. 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol,* 20**,** 4754-64.

Lariviere, L., Plaschka, C., Seizl, M., Wenzeck, L., Kurth, F. & Cramer, P. 2012. Structure of the Mediator head module. *Nature,* 492**,** 448-51.

Lauberth, S. M., Nakayama, T., Wu, X., Ferris, A. L., Tang, Z., Hughes, S. H. & Roeder, R. G. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell,* 152**,** 1021-36.

Layer, J. H., Miller, S. G. & Weil, P. A. 2010. Direct transactivator-transcription factor IID (TFIID) contacts drive yeast ribosomal protein gene transcription. *J Biol Chem,* 285**,** 15489-99.

Layer, J. H. & Weil, P. A. 2013. Direct TFIIA-TFIID Contacts Drive Budding Yeast Ribosomal Protein Gene Transcription. *J Biol Chem*.

Lee, M. & Struhl, K. 1995. Mutations on the DNA-binding surface of TATA-binding protein can specifically impair the response to acidic activators in vivo. *Mol Cell Biol,* 15**,** 5461-9.

Lee, T. I., Causton, H. C., Holstege, F. C., Shen, W. C., Hannett, N., Jennings, E. G., Winston, F., Green, M. R. & Young, R. A. 2000. Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature,* 405**,** 701-4.

Leurent, C., Sanders, S. L., Demeny, M. A., Garbett, K. A., Ruhlmann, C., Weil, P. A., Tora, L. & Schultz, P. 2004. Mapping key functional sites within yeast TFIID. *EMBO J,* 23**,** 719-27.

Li, X. Y., Bhaumik, S. R., Zhu, X., Li, L., Shen, W. C., Dixit, B. L. & Green, M. R. 2002. Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. *Curr Biol,* 12**,** 1240-4.

Lim, C. Y., Santoso, B., Boulay, T., Dong, E., Ohler, U. & Kadonaga, J. T. 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev,* 18**,** 1606-17.

Liu, W. L., Coleman, R. A., Ma, E., Grob, P., Yang, J. L., Zhang, Y., Dailey, G., Nogales, E. & Tjian, R. 2009. Structures of three distinct activator-TFIID complexes. *Genes Dev,* 23**,** 1510-21.

Majmudar, C. Y., Wang, B., Lum, J. K., Hakansson, K. & Mapp, A. K. 2009. A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew Chem Int Ed Engl,* 48**,** 7021-4.

Marion, R. M., Regev, A., Segal, E., Barash, Y., Koller, D., Friedman, N. & O'shea, E. K. 2004. Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc Natl Acad Sci U S A,* 101**,** 14315-22.

Martin, D. E., Soulard, A. & Hall, M. N. 2004. TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell,* 119**,** 969-79.

Martinez, E., Zhou, Q., L'etoile, N. D., Oelgeschlager, T., Berk, A. J. & Roeder, R. G. 1995. Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. *Proc Natl Acad Sci U S A,* 92**,** 11864-8.

Matangkasombut, O., Buratowski, R. M., Swilling, N. W. & Buratowski, S. 2000. Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev,* 14**,** 951-62.

Mencia, M., Moqtaderi, Z., Geisberg, J. V., Kuras, L. & Struhl, K. 2002. Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast. *Mol Cell,* 9**,** 823-33.

Miller, G. & Hahn, S. 2006. A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat Struct Mol Biol,* 13**,** 603-10.

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 Dev,* 22**,** 2994-3006.

Moqtaderi, Z., Bai, Y., Poon, D., Weil, P. A. & Struhl, K. 1996. TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature,* 383**,** 188-91.

Nakatani, Y., Horikoshi, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R. G. & Freese, E. 1990. A downstream initiation element required for efficient TATA box binding and in vitro function of TFIID. *Nature,* 348**,** 86-8.

Nikolov, D. B. & Burley, S. K. 1994. 2.1 A resolution refined structure of a TATA box-binding protein (TBP). *Nat Struct Biol,* 1**,** 621-37.

Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G. & Burley, S. K. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A*, 93**,** 4862-7.

Ohtsuki, S., Levine, M. & Cai, H. N. 1998. Different core promoters possess distinct regulatory activities in the Drosophila embryo. *Genes Dev,* 12**,** 547-56.

Papai, G., Tripathi, M. K., Ruhlmann, C., Layer, J. H., Weil, P. A. & Schultz, P. 2010. TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature,* 465**,** 956-60.

Pardo, L., Campillo, M., Bosch, D., Pastor, N. & Weinstein, H. 2000. Binding mechanisms of TATA box-binding proteins: DNA kinking is stabilized by specific hydrogen bonds. *Biophys J,* 78**,** 1988-96.

Parry, T. J., Theisen, J. W., Hsu, J. Y., Wang, Y. L., Corcoran, D. L., Eustice, M., Ohler, U. & Kadonaga, J. T. 2010. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev,* 24**,** 2013-8.

Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodadek, T. & Burley, S. K. 1999. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev,* 13**,** 3217-30.

Ranallo, R. T., Struhl, K. & Stargell, L. A. 1999. A TATA-binding protein mutant defective for TFIID complex formation in vivo. *Mol Cell Biol,* 19**,** 3951-7.

Ranish, J. A., Yudkovsky, N. & Hahn, S. 1999. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev,* 13**,** 49-63.

Reddy, P. & Hahn, S. 1991. Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region. *Cell,* 65**,** 349-57.

Reeves, W. M. & Hahn, S. 2005. Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol,* 25**,** 9092-102.

Rhee, H. S. & Pugh, B. F. 2012. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature,* 483**,** 295-301.

Robinson, P. J., Bushnell, D. A., Trnka, M. J., Burlingame, A. L. & Kornberg, R. D. 2012. Structure of the mediator head module bound to the carboxy-terminal domain of RNA polymerase II. *Proc Natl Acad Sci U S A*, 109**,** 17931-5.

Sainsbury, S., Niesser, J. & Cramer, P. 2013. Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature,* 493**,** 437-40.

Samara, N. L. & Wolberger, C. 2011. A new chapter in the transcription SAGA. *Curr Opin Struct Biol,* 21**,** 767-74.

Schawalder, S. B., Kabani, M., Howald, I., Choudhury, U., Werner, M. & Shore, D. 2004. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature,* 432**,** 1058-61.

Seizl, M., Hartmann, H., Hoeg, F., Kurth, F., Martin, D. E., Soding, J. & Cramer, P. 2011. A conserved GA element in TATA-less RNA polymerase II promoters. *PLoS One,* 6**,** e27595.

Shen, W. C. & Green, M. R. 1997. Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator. *Cell,* 90**,** 615-24.

Shen, W. C., Bhaumik, S. R., Causton, H. C., Simon, I., Zhu, X., Jennings, E. G., Wang, T. H., Young, R. A. & Green, M. R. 2003. Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *EMBO J,* 22**,** 3395-402.

Smale, S. T. & Kadonaga, J. T. 2003. The RNA polymerase II core promoter. *Annu Rev Biochem,* 72**,** 449-79.

Sugihara, F., Kasahara, K. & Kokubo, T. 2011. Highly redundant function of multiple AT-rich sequences as core promoter elements in the TATA-less RPS5 promoter of Saccharomyces cerevisiae. *Nucleic Acids Res,* 39**,** 59-75.

Sypes, M. A. & Gilmour, D. S. 1994. Protein/DNA crosslinking of a TFIID complex reveals novel interactions downstream of the transcription start. *Nucleic Acids Res,* 22**,** 807-14.

Tan, S., Hunziker, Y., Sargent, D. F. & Richmond, T. J. 1996. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature,* 381**,** 127-51.

Theisen, J. W., Lim, C. Y. & Kadonaga, J. T. 2010. Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. *Mol Cell Biol,* 30**,** 3471-9.

Thomas, M. C. & Chiang, C. M. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol,* 41**,** 105-78.

Tsukihashi, Y., Kawaichi, M. & Kokubo, T. 2001. Requirement for yeast TAF145 function in transcriptional activation of the RPS5 promoter that depends on both core promoter structure and upstream activating sequences. *J Biol Chem,* 276**,** 25715-26.

Van Ingen, H., Van Schaik, F. M., Wienk, H., Ballering, J., Rehmann, H., Dechesne, A. C., Kruijzer, J. A., Liskamp, R. M., Timmers, H. T. & Boelens, R. 2008. Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure,* 16**,** 1245-56.

Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., Van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M. & Timmers, H. T. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell,* 131**,** 58-69.

Wade, J. T., Hall, D. B. & Struhl, K. 2004. The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature,* 432**,** 1054-8.

Warner, J. R. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci,* 24**,** 437-40.

Wu, P. Y. & Winston, F. 2002. Analysis of Spt7 function in the Saccharomyces cerevisiae SAGA coactivator complex. *Mol Cell Biol,* 22**,** 5367-79.

Yang, C., Bolotin, E., Jiang, T., Sladek, F. M. & Martinez, E. 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene,* 389**,** 52-65.

Zaros, C., Briand, J. F., Boulard, Y., Labarre-Mariotte, S., Garcia-Lopez, M. C., Thuriaux, P. & Navarro, F. 2007. Functional organization of the Rpb5 subunit shared by the three yeast RNA polymerases. *Nucleic Acids Res,* 35**,** 634-47.

Zhang, Z. & Dietrich, F. S. 2005. Mapping of transcription start sites in Saccharomyces cerevisiae using 5' SAGE. *Nucleic Acids Res,* 33**,** 2838-51.

Zhao, Y., Mcintosh, K. B., Rudra, D., Schawalder, S., Shore, D. & Warner, J. R. 2006. Fine-structure analysis of ribosomal protein gene transcription. *Mol Cell Biol,* 26**,** 4853-62.

Zhong, P. & Melcher, K. 2010. Identification and characterization of the activation domain of Ifh1, an activator of model TATA-less genes. *Biochem Biophys Res Commun,* 392**,** 77-82.

Zhou, Q. & Berk, A. J. 1995. The yeast TATA-binding protein (TBP) core domain assembles with human TBP-associated factors into a functional TFIID complex. *Mol Cell Biol*, 15**,** 534-9.