BIOCHEMICAL AND GENETIC ANALYSES OF INTERACTIONS BETWEEN TRANSACTIVATORS AND TBP ASSOCIATED FACTORS IN *SACCHAROMYCES CEREVISIAE*

By

Justin H. Layer

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

In

Molecular Physiology and Biophysics

December 2010

Nashville, Tennessee

Approved:

Roland W. Stein

Raymond L. Mernaugh

Roger J. Colbran

Douglas P. Mortlock

Linda J. Sealy

ACKNOWLEDGEMENTS

I must first make clear that my family and close friends enjoy ownership in this document that signifies an experience and achievement of a lifetime. As this affair has been such a very long time coming, I unfortunately cannot keep all of those friends in order. Reason being that I feel as if many lives have passed since I began here nearly eight years ago. However I claim confidently that my parents Jane and Hugh Layer, and Dan and Tonya Snyder (my "Nashville Parents!") helped me maintain my sanity during this trying period. Even if they couldn't understand the basis of my personal crisis of any given moment, they always lent an ear and words of encouragement, or repeated advice that I had been given but was still too stubborn to follow. I will never forget all that Dan has taught me about the joys of craftsmanship, and in so doing at times gave me a needed reprieve from the depression originating from countless hours spent in a particularly frustrating workplace environment. These friends watched over me with love and helped to keep me from self-destructive actions, the irreparable kind at least.

In science, I also met some very special one-of-a-kind people. Drs. Madhu V. Singh and Manish K. Tripathi are prominent individuals I will always remember from my training days. Dr. Mourad Bendjennat made a significant impression on my scientific training. One of the biggest regrets of my life is my series of mistakes that led to the dissolution of our friendship. I was just too young and stupid to make the most of this opportunity. Dr. Xiaohui Wang showed me friendship early on, and I remiss on my failure to reciprocate for her kindness. Again, more incompetence in personal relationships; here perhaps more than any other place I would wish for second chances.

The person I owe the most beside myself for reaching this point is none other than Tony Weil. He was the only educator who could shepherd me through this process. After I said or did callous or stupid things when my professional composure eroded, he gave me second chance upon second chance, again and again. I will always hold Tony as a personal role model, as a self-made man and as a thinking man that stands out in many aspects amongst other highly intelligent individuals. I like to think the completion of my training here is a crowning achievement for him too, and to more of an extent perhaps than it has been for other of his trainees. This is because he and I went through such a turbulent series of events, due to the goings on both outside and within the lab. We endured through some serious trials and tribulations, and we confronted some of these challenges in common, like true colleagues. I really do believe that I had an extremely difficult PhD. process, and not strictly due to the particular rigor of Tony's scientific philosophy. Ultimately, I will benefit from going through this hardship while starting out in science, and hope that it will prove to be a source of strength. I very much hope that I can use my experience obtained here to obtain a more spiritually prosperous way of living and participating in scientific discovery.

TABLE OF CONTENTS

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Overview

Model organisms are used to study biological processes that share commonalities throughout the phylogenetic tree of life, with the ultimate objective of a comprehensive understanding of the molecular dynamics of even the most highly evolved creature. Study of model organisms represents a reductionist approach to our investigations, because one simplification dictates that the "least complex" organism will likely have the fewest and most easily dissected molecular processes that it can undertake. Indeed, scientific understanding proceeds through incremental gains in knowledge, and one strategy is to bracket the constituents of an organism into all of those "systems" or "circuits" functional at the level of a single cell; these include sets of macromolecular assemblies or "machines," which can be further subdivided as typically consisting of several individual protein or RNA subunits (Alberts, 1998; Alberts and Miake-Lye, 1992). The ability to define a complete collection of "protein machines" is greatly facilitated by the availability of genomic DNA sequences, since this information contains the blueprint for the identity of the individual RNA and protein molecules needed to construct a cell and hence the whole organism. However, to a cell the blueprints encoded by the genomic DNA are non-functional without molecular machineries to faithfully interpret, maintain, and propagate distribution of the informational content through

subsequent generations. It so happens that the "circuits," "machines," and "parts lists" behind these genome-directed processes are very similar whether one is examining a simple unicellular model eukaryote or a highly evolved human cell. It is through an understanding and appreciation of this basic concept that the life sciences have enjoyed such dramatic gains of knowledge in the last half-century.

This document describes the molecular process of eukaryotic transcription, the molecular mechanism by which the informational content of the genomic DNA is converted into RNA, either messenger RNA (mRNA), tranfer (tRNA), ribosomal RNA (rRNA), or other low molecular weight non-coding RNA (ncRNA). tRNA, rRNA, and the small ncRNAs such as small nucleolar (SnoRNA) or micro RNA (miRNA) represent entities that are direct participants in biological regulatory processes; these are structural and functional components of non-transcriptional gene regulation. On the other hand, mRNA is mainly thought of as the critical biological intermediate that must be further interpreted during the process of translation to yield a protein product. Thus, the function of an mRNA is as an obligatory, though indirect, contributor to cellular processes. The different kinds of RNA are synthesized by enzymes known as RNA polymerases that use DNA sequences as templates to produce specific transcripts.

In eukaryotic cells there are multiple RNA polymerases, each of which is responsible for transcription of a subset of the RNAs mentioned just above (Roeder and Rutter, 1969). All eukaryotes contain 3 nuclear DNA-dependent RNA polymerases. RNA Polymerase I (Pol I) synthesizes rRNAs that are constituents of the ribosome. RNA Polymerase III (Pol III) transcribes 5S rRNA, also a ribosomal component, as well as all of the tRNAs and some of the SnoRNAs. The overwhelming majority of the diversity in

RNA species comes from mRNA, which is synthesized by RNA Polymerase II (Pol II). Plants contain two additional nuclear RNA polymerases, Pol IV and Pol V, and there is recent evidence of a nuclear-encoded enzyme that is differentially spliced to generate an RNA polymerase that targets either to the nucleus or the mitochondrion (Kravchenko et al., 2005; Matsunaga et al., 2004; Ream et al., 2009). Other non-nuclear enzymes include the mitochondrial RNA polymerase and chloroplast RNA Polymerases in plants, both of which are considerably more simple molecules and only functional for transcription of the organelle genomes (Matsunaga et al., 2004).

Despite the diversity of function in the different final transcription products, all of these enzymes share similar overall characteristics with respect to the structure and function of their catalytic components. The three-dimensional structure of yeast Pol II was solved by Roger Kornberg's lab, and comparison of this data with that of the comparable archael RNAP structure supports the concept of conserved catalytic mechanisms (Asturias and Kornberg, 1999; Boeger et al., 2005; Bushnell et al., 2002; Bushnell and Kornberg, 2003; Bushnell et al., 2004; Chung et al., 2003; Cramer et al., 2001; Darst et al., 1991; Edwards et al., 1990; Fu et al., 1999; Gnatt et al., 1997; Gnatt et al., 2001; Liu et al.; Spahr et al., 2009; Wang et al., 2009; Westover et al., 2004a, b; Zhang et al., 1999). Further emphasizing the features shared between the eukaryotic enzymes, all three are composed of anywhere from twelve to sixteen different proteins, some of which are commonly found within several of the three RNA Polymerases, an observation that again hints at both complexity and similarities in mechanism of function (Sklar et al., 1975; Woychik et al., 1990). Even given the noted similarity between eukaryotic polymerases and the lone bacterial polymerase, there are two major

differences. First, the eukaryotic enzymes have more than twice the number of subunits as the bacterial enzyme (β, β', α^2, ϖ, and σ subunits). Secondly and in spite of this seemingly increased overall complexity, when the purified enzymes are tested in vitro none of the nuclear polymerases are capable of synthesizing a biologically relevant RNA species, a feat readily accomplished by the bacterial RNA Polymerase Holoenzyme ($\beta\beta'\alpha^2\varpi\sigma$, referring to the isolated species of enzyme that possesses all biologically necessary activities, as a cohesive assembly of proteins) (Burgess, 1969a, b; Burgess et al., 1969; Weil et al., 1979).

The differences between the bacterial enzyme and the eukaryotic RNA Polymerases indicate an increased complexity in the eukaryotic transcription program. It is not unexpected that over the course of evolution, nature created additional means of regulatory inputs to allow higher flexibility and fine-tuning of homeostatic and developmental processes. The final products of translation of different mRNA transcripts, the proteins, are basic functional units found in cells, and the physiological status or even identity of an individual cell is intimately linked to which subset of all the possible genome-encoded mRNAs and proteins are actually present (DeRisi et al., 1997). Thus cellular homeostasis and identity rely on differential gene expression, which typically leads to changes in protein abundance. As it happens, one of the major means of controlling protein abundance occurs at the level of mRNA gene transcription by RNA Pol II. There are four phases of polymerase function that can be tightly regulated: transcript initiation, promoter clearance, elongation, and termination. Inhibition or stimulation of any of these four phases of transcription can affect the quantity of

functional transcript produced. The steady state cellular abundance of a transcript has a major influence on the ability of cells to synthesize the corresponding protein.

Structure of RNA Polymerase II-Transcribed Genes

The most obvious component of a simplified gene (as illustrated in **Figure 1.1**), excluding the presence of intronic non-protein coding sequence, is the DNA sequence encoding all of the mRNA triplet codons that in turn correspond to the anticodons specifying amino acids of the protein product. This portion of a gene is referred to as the open reading frame (ORF). However, an mRNA will possess additional sequences located at both the 5' and 3' ends that also correspond to the DNA sequence, but that precede the initiating methionine codon and extend beyond the stop codon, respectively. The upstream end of this so-called 5' untranslated region (UTR) of a message always corresponds to the position of the transcription start site, the very first nucleotide transcribed by Pol II. It is not uncommon for a gene to have several unique transcription start sites that map to within 5-10 nucleotides. The 5' UTR sequence can have important impacts on post-transcriptional processing, translation, and even on subcellular localization of the mRNA. The 3' UTR also affects various properties of mRNA. Notably, there will be extra features found at the 5' and 3' ends of a typical mRNA in addition to the 5' and 3' UTR. These extra parts of the mRNA do not directly correspond to the DNA encoding the transcript. Instead, the 5' methyl guanosine cap and the 3' poly-adenosine (A) tail are added co- or post-transcriptionally. These mRNA modifications affect translation and mRNA turnover rates.

The position of the transcription start site is part of the so-called promoter region of a gene (Burke et al., 1998; Butler and Kadonaga, 2002; Juven-Gershon et al., 2006b).

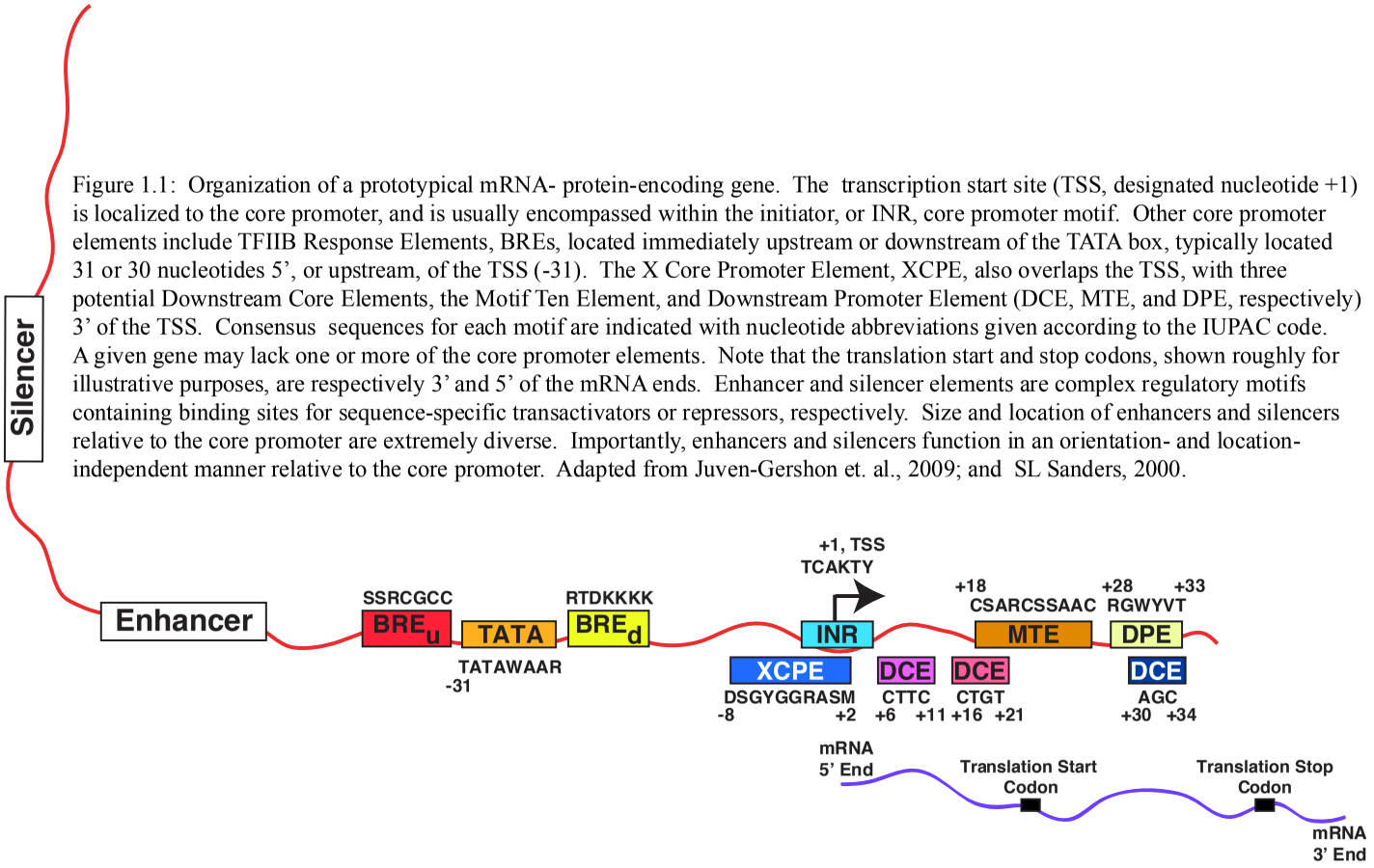


Figure 1.1: Organization of a prototypical mRNA- protein-encoding gene. The transcription start site (TSS, designated nucleotide +1) is localized to the core promoter, and is usually encompassed within the initiator, or INR, core promoter motif. Other core promoter elements include TFIIB Response Elements, BREs, located immediately upstream or downstream of the TATA box, typically located 31 or 30 nucleotides 5', or upstream, of the TSS (-31). The X Core Promoter Element, XCPE, also overlaps the TSS, with three potential Downstream Core Elements, the Motif Ten Element, and Downstream Promoter Element (DCE, MTE, and DPE, respectively) 3' of the TSS. Consensus sequences for each motif are indicated with nucleotide abbreviations given according to the IUPAC code. A given gene may lack one or more of the core promoter elements. Note that the translation start and stop codons, shown roughly for illustrative purposes, are respectively 3' and 5' of the mRNA ends. Enhancer and silencer elements are complex regulatory motifs containing binding sites for sequence-specific transactivators or repressors, respectively. Size and location of enhancers and silencers relative to the core promoter are extremely diverse. Importantly, enhancers and silencers function in an orientation- and location-independent manner relative to the core promoter. Adapted from Juven-Gershon et. al., 2009; and SL Sanders, 2000.

As shown in **Figure 1.1**, there are several DNA sequence elements that constitute a promoter; these include TATA, INR, DPE, MTE, BRE, and XCPE elements (Anish et al., 2009; Juven-Gershon et al., 2008; Kadonaga, 2002; Lim et al., 2004; Smale and Baltimore, 1989; Tokusumi et al., 2007). Importantly, the TATA box is the only identified promoter element in yeasts. Whether or not functional equivalents of the other promoter cis-elements exist in yeast remains to be determined. It is likely that additional unique yeast promoter elements remain to be discovered. In metazoan promoters, all of the known elements have a characteristic order and composition of nucleotides that vary to an extent from a consensus sequence, depending on the particular gene. This variation may affect the capacity of the promoter to stimulate transcription. The probable absence of one or more elements from some promoters also contributes to unique enhancer-promoter transcriptional properties of genes (Juven-Gershon et al., 2006). The overall function of these DNA motifs at promoters is to help localize the RNA Polymerase at the transcription start site since Pol II has no potential for sequence specific DNA binding, unlike its bacterial counterpart. Thus the promoter sequence motifs act indirectly for accurate positioning of Pol II.

Although the composition of the promoter affects the transcriptional output, these elements are not sufficient to stimulate physiological levels of transcription, either in vitro or in vivo. Additional sequence elements known as enhancers are required to regulate the amount of transcript produced, and when fused to the promoter can stimulate transcription as much as 1000 fold (simplified representation in **Figure 1.1**). Enhancers are complicated regulatory elements, typically composed of multiple individual sequence elements that act to modulate transcription rates (Szutorisz et al., 2005). In yeasts these

DNA elements are known as upstream activating (UAS) or repressing (URS) sequences, named in part because they are typically located upstream of the promoter, within 1-2 kilobases of the transcription start site (Bram et al., 1986; Lorch and Kornberg, 1985). Regardless of any other general characteristics, an overarching feature of an enhancer is the ability to function in both a distance- and orientation-independent manner. For example, vertebrate enhancers can influence transcription of their associated transcription unit when placed hundreds of kilobases upstream or downstream of the transcription start site. These properties are attributed to the ability of proteins to bind specifically to small six to twenty nucleotide cis-elements within the enhancer. There are typically many such small sequence motifs within an enhancer. The identity, number, spacing, order of arrangement, and variation of cis-elements from a consensus sequence make a huge contribution to the transcriptional properties of a gene. Thus enhancers modulate maximum output potential as well as the ability to integrate cell-type specific and/or environmental signals into a gene's transcriptional activity at any point in time. In fact, it is generally accepted that the majority of regulatory decisions are integrated into a genes function, at least initially, by virtue of the dynamic properties of the enhancer. The cellular decision making process exploits the properties of the factors that bind to the enhancer. These DNA-binding proteins are known as transfactors.

The properties of the transfactor-bound enhancer and those of the promoter, and the communication between enhancer and promoter are the drivers of transcriptional control. The outcome of this communication can affect all four stages of RNA Polymerase activity from initiation to termination. Importantly, the accessibility and hence function of enhancers and promoters can be affected by nucleosomes, protein-

DNA assemblies that consist of roughly 150 base pairs of DNA wrapped around an octamer of four histone proteins, two each of H2A, H2B, H3, and H4 (Kornberg and Lorch, 1999). The octamer of histones is tightly bound to the encompassed DNA, and this feature means that a particular sequence of DNA may be inaccessible to factors even if they could normally bind with high affinity to an element within the nucleosome (Luger et al., 1997). Thus nucleosomal architecture influences the activity of both enhancers and promoters by limiting access of critical regulatory factors to the DNA. However, to say that chromatin is simply repressive is inaccurate (Durrin et al., 1992; Han and Grunstein, 1988; Han et al., 1988; Kim et al., 1988). While the physical presence and precise location of nucleosomes may be relatively static over time (as determined by population measurements, and hence an average of position), their exact biochemical characteristics may vary widely, with resulting alterations in transcriptional activity. Histone proteins within the nucleosome are subject to a vast array of post-translational modifications (PTMs) (Allfrey et al., 1964; Allfrey and Mirsky, 1964; Jenuwein and Allis, 2001). PTM status of the histone proteins affects nucleosomal properties in two critical ways. First, PTMs affect the nucleosome-DNA interaction to strengthen or weaken it thereby influencing the ability of the nucleosome to compete with other DNA-binding factors for occupancy of DNA sequence elements. Note that modification on the N-terminal histone tails, where the majority of PTMs occur, probably does not influence stability of nucleosome-DNA interactions since the histone tails are unstructured and not necessarily near the DNA. Instead nucleosome-DNA interaction is probably modulated through histone residues located in the nucleosomal interior such as H3 lysine 56 (Shahbazian and Grunstein, 2007; Xu et al., 2005). Second, histone

modifications can stimulate or repress association of the nucleosome with histone-binding factors that may ultimately stimulate or repress transcription (Hecht et al., 1995; Hecht et al., 1996). Thus PTM status of histones can stimulate or inhibit protein-protein interactions between histones and other regulatory factors, and therefore affect DNA-independent factor localization around regulatory regions in the DNA. Histone residues in the amino terminal tails are believed to be modified to influence protein-protein interactions between the nucleosome and other coregulators. Histones are subjected to phosphorylation, acetylation, methylation, ubiquitination, sumoylation, glycosylation, ADP-ribosylation, and other PTMs on many residues of all four proteins, and there is evidence of combinatorial variety in the final outcome of such modifications. It has been hypothesized that these PTMs comprise a "code" that dictates how combinations of modifications alter the accessibility and interaction with other transcription proteins and hence function of the underlying DNA (Strahl and Allis, 2000; Sun and Allis, 2002). In addition, there are histone variants that have different properties than the canonical histones and these variants can be inserted into nucleosomes at specific locations in genes (Krogan et al., 2003; Mizuguchi et al., 2004; Santisteban et al., 2000). Although it is difficult to state any immutable rules regarding whether chromatin architecture will be stimulatory or repressive to transcription in a give context, it is clear that gene expression can be heavily dependent upon chromatin-directed activities.

To summarize which gene characteristics influence RNA polymerase activity, we can then say that the promoter DNA elements help localize the enzyme and variations in promoter DNA sequence composition can affect the efficiency of transcription. Function and identity of enhancer DNA elements have a more profound affect on transcription

rates. Finally, chromatin can be highly regulated and dynamic over the course of the cell cycle and in response to environmental changes. Chromatin structure affects the behavior of both enhancer and promoter and even the ability of the polymerase to proceed once transcription is initiated. The decision to make the correct transcript at the right time and in the appropriate amount is tied to all three components of gene structure: promoter, enhancer, and chromatin structure. The combination ultimately affects the ability of RNA polymerase to do its job. We must now begin to consider how regulatory inputs are made into both the promoter and enhancer through associated regulatory proteins, how chromatin affects and is affected by these transactions, and how regulation of the promoter, the enhancer, and the overlying chromatin structure collectively regulate transcription of a given gene.

Basal Transcription Factors and Coregulators

As mentioned above RNA polymerase II cannot accurately initiate synthesis of a transcript alone, unlike the bacterial enzyme whose σ and α subunits contain promoter cis-element DNA binding domains (Browning and Busby, 2004). σ factor binds to specific sequence elements in all bacterial core promoters and thus contributes to correct polymerase localization. Pol II lacks comparable specific DNA binding capacity, and in fact it took over a decade between the initial identification of the enzyme and the discovery of Polymerase accessory factors that would allow accurate 5' mRNA end production in vitro (Weil et al., 1979). This in turn led to the discovery of the protein factors that allow accurate Pol II transcription (Matsui et al., 1980). Extensive purification efforts identified six distinct chromatographic fractions required to

recapitulate accurate in vitro transcription using purified Pol II (Conaway et al., 1996; Feaver et al., 1991; Flores et al., 1988; Gileadi et al., 1992; Hahn et al., 1989a; Henry et al., 1992; Reinberg and Roeder, 1987a, b; Sayre et al., 1992; Sumimoto et al., 1990). These were subsequently more extensively purified and termed TFIIA, B, D, E, F, and H, where the acronyms specify Transcription Factor, RNA Pol II transcription, and the letter suffix indicates order in which components were identified. Continuing investigation using additional promoters as templates in transcription assays confirmed that different target promoters exhibit differences in the degree of dependency upon individual factors. After genome sequencing revealed the composition of many gene core promoters, it became clear that no single DNA element is present in every described promoter, an observation likely indicative of mixed requirements or differences in mechanism of function for these transcription factors from gene to gene. As such we can consider it less accurate to refer to these factors as 'General Transcription Factors', but should describe them as 'Basal Transcription Factors' because it is fair to say that they are often all required for low-level transcription, at least in vitro (Sikorski and Buratowski, 2009). Each factor performs a distinct role in formation of a Pol II-containing pre-initiation complex (PIC) on model genes in vitro (see **Figure 1.2**).

TFIIA facilitates TFIID loading onto the promoter, in part by acting as an anti-inhibitor to factors that restrict TFIID-promoter binding (Ma et al., 1996; Weideman et al., 1997). TFIIB stabilizes the IIA/IID complex on the promoter to augment association of TFIIF and Pol II in the PIC and it also helps in choice of the transcription start site (Nikolov et al., 1995). TFIIB can reportedly auto-acetylate and this seems mandatory for transcription to occur but the exact mechanism of this PTM remains obscure

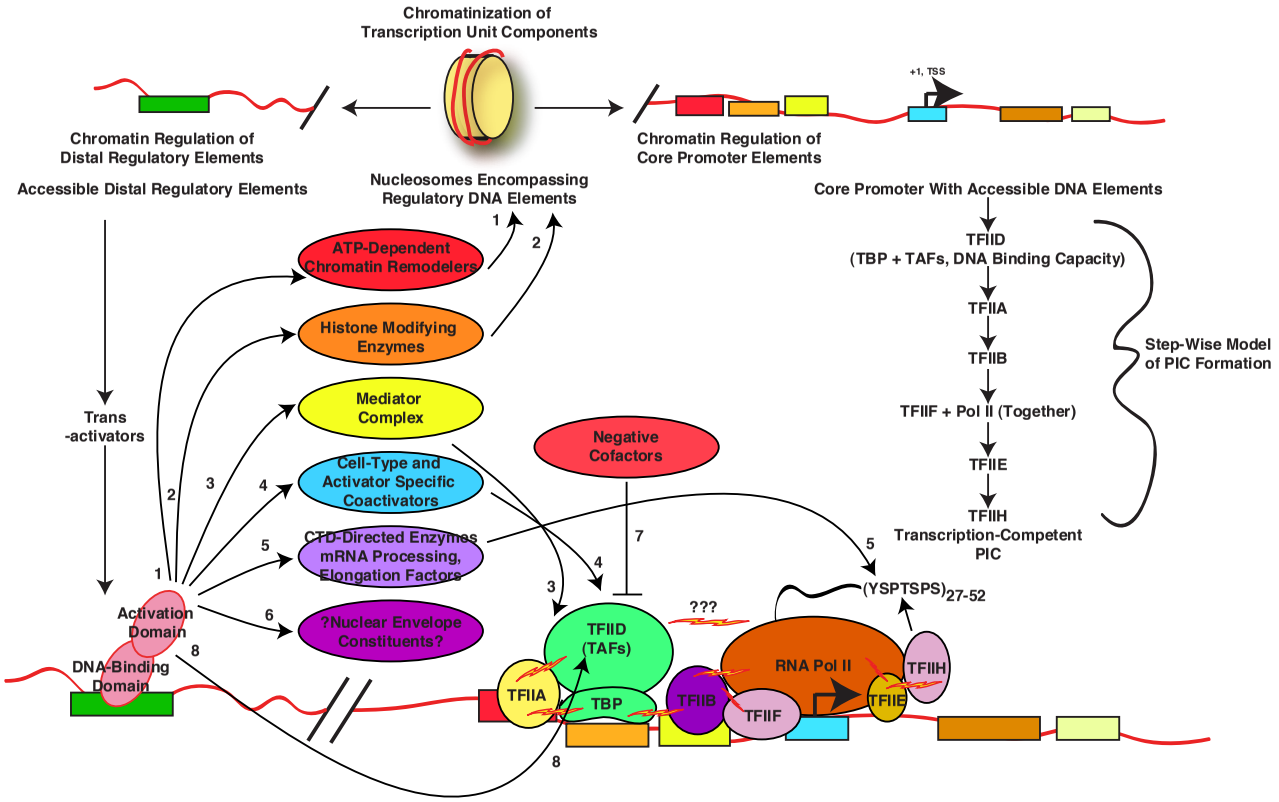


Figure 1.2: Protein-DNA and Protein-Protein Interactions in mRNA gene regulation. Enhancer or Core Promoter Sequence Elements may be free or encompassed in chromatin that restricts access of factors that bind to DNA. Functional Trans-activators will bind to their cis elements in absence of a chromatin barrier, and use their activation domains to engage other factors by protein-protein interaction. Targets include Coregulatory protein complexes with activity directed towards alleviation of repressive chromatin architecture, the Mediator complex which bridges enhancer-bound activators to the promoter bound Pre-Initiaion complex (PIC) or other regulatory complexes. Cell-type specific coactivators collaborate with activators and Mediator to stimulate PIC function. Activator contact with CTD-directed activities can influence the protein-protein interactions of the Pol II CTD to include elongation factors and enzymes involved in 5' and 3' mRNA processing and splicing. Activator contact may extend to Nuclear Pore Components to affect the sub-nuclear locale of transcribed genes. All of these factors ultimately affect PIC function or formation that can occur through the step-wise program shown at the right. Protein-Protein interactions between PIC components are indicated by the jagged arrows. Coregulators can counteract the action of negative cofactors that inhibit PIC formation or function. Finally, activators may directly contact the PIC; this has most often been attributed to direct contact with TAF components of TFIID, which also recognizes most of the core promoter sequence elements. Adapted from R.G. Roeder, 2005.

(Choi et al., 2003). TFIID is the most complex of all the factors, with 15 subunits, and contains the major DNA binding activity; IID association at promoters was classically described as rate-limiting for PIC formation (Buratowski et al., 1989; Van Dyke et al., 1988). TFIID is the focal point of this dissertation. TFIIE function is tied to that of TFIIH as IIE helps bring IIH to the PIC and also stimulates IIH-intrinsic enzymatic activities described below (Ohkuma et al., 1995; Ohkuma and Roeder, 1994). TFIIF directly associates with Pol II and acts as an adaptor between it and the IIB/IID assembly, which in turn allows clearance and transcription. Finally, TFIIH provides several enzymatic activities including those needed for promoter melting and modification of Pol II by phosphorylation, both of which help Pol II transition out of the PIC and into elongation complexes (Lu et al., 1992).

The enormous amount of investigation that led to the understanding presented in the last paragraph relied upon sophisticated biochemical experiments. An 'order of action' model was presented based upon detailed biochemical complementation experiments (see **Figure 1.3**, (Buratowski et al., 1989; Hawley and Roeder, 1985, 1987; Van Dyke et al., 1988)), although as stated the 'rules' for factor responsiveness probably vary somewhat from gene to gene. The complete purification of each factor led to the identification, isolation, and cDNA cloning for each subunit of each complex; 2 for TFIIA, 1 for TFIIB, 15 or more for TFIID, 2 for TFIIE, 3 for TFIIF, 10 for TFIIH, and 12 Pol II subunits (DeJong and Roeder, 1993; Dynlacht et al., 1991; Horikoshi et al., 1989a; Moqtaderi et al., 1996b; Pinto et al., 1992; Poon et al., 1995; Reese et al., 1994; Sun and Hampsey, 1995; Young and Davis, 1983). All of the factors have been purified from budding yeast and human sources, and some have been isolated from fission yeast, fruit

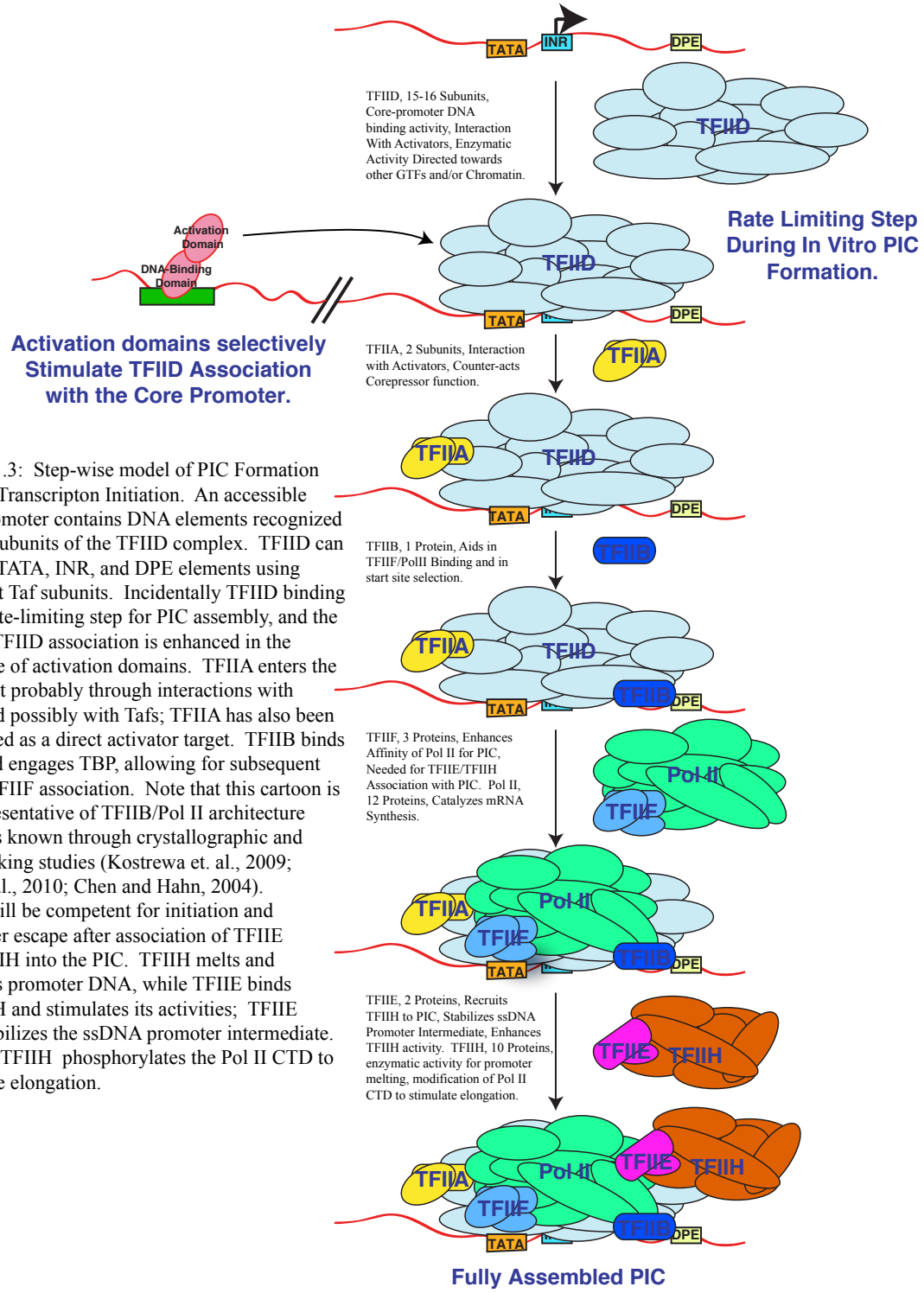


Figure 1.3: Step-wise model of PIC Formation During Transcripton Initiation. An accessible core promoter contains DNA elements recognized by Taf subunits of the TFIID complex. TFIID can bind to TATA, INR, and DPE elements using different Taf subunits. Incidentally TFIID binding is the rate-limiting step for PIC assembly, and the rate of TFIID association is enhanced in the presence of activation domains. TFIIA enters the PIC next probably through interactions with TBP and possibly with Tafs; TFIIA has also been suggested as a direct activator target. TFIIB binds next and engages TBP, allowing for subsequent Pol II/TFIIF association. Note that this cartoon is notrepresentative of TFIIB/Pol II architecture which is known through crystallographic and crosslinking studies (Kostrewa et. al., 2009; Liu et. al., 2010; Chen and Hahn, 2004). Pol II will be competent for initiation and promoter escape after association of TFIIE and TFIIH into the PIC. TFIIH melts and unwinds promoter DNA, while TFIIE binds to TFIIH and stimulates its activities; TFIIE also stabilizes the ssDNA promoter intermediate. Finally, TFIIH phosphorylates the Pol II CTD to facilitate elongation.

fly, and mouse. The first major conclusion is that there is a high level of phylogenetic amino acid sequence conservation in eukaryotes for every factor subunit. The second conclusion is that in a limit case, 46 polypeptides are required to fulfill a set of biochemical functions satisfied by merely five proteins of the bacterial Holoenzyme, $\beta\beta'\alpha^2\sigma$. Actually, initial studies using TATA-containing promoters showed all of these factors are required for activator-responsive in vitro transcription, although activator-independent transcription can utilize just the TBP subunit of TFIID and under certain circumstances become independent of TFIIE and TFIIH as well, whereas on TATA-less, INR-containing promoters all of the factors are absolutely required for transcription (Conaway and Conaway, 1990b; Pugh and Tjian, 1990; Smale et al., 1990). Contrast this situation again with the *E. coli* RNA polymerase, where both promoter recognition and activator responsiveness of polymerase can be entirely mediated by just one polypeptide, sigma, associating with the 'core' $\beta\beta'\alpha^2$ assembly. The apparent complexity in eukaryotes suggests that there are many polypeptide candidates for mediating activator responsiveness available in eukaryotic Pol II transcription. This situation is attractive for allowing entry of diverse regulatory inputs with equally diverse functional consequences, all of which could contribute to fine-tuning of transcriptional output. Because it was initially suggested that multiple proteins within the different PIC- forming complexes may directly respond to activators, they could qualify as coregulators, which are defined as factors that contribute to enhancer-mediated regulation of Pol II activity. However as time has passed, dozens of additional factors qualifying as coregulators have been identified (Roeder, 2005). For a summary of some of the different classifications of coactivators, see **Figure 1.2**. Some of the first studies hinting of regulatory entry points,

exclusive from basal transcription factors, identified a so-called eukaryotic RNA Polymerase II 'holoenzyme' that contains Pol II, some of the basal factors, and an additional thirty-plus novel proteins that stimulate a significantly higher amount of transcription than Pol II alone when combined with TFIIB, IIF, IIE, IIH and the TBP subunit of TFIID (Kelleher et al., 1990; Kim et al., 1994; Koleske and Young, 1994). This holoenzyme was also required for activator responsiveness in the same biochemical settings that did not include holo-TFIID (Koleske and Young, 1994). The novel holoenzyme components were subsequently referred to as the Mediator of transcription, because it appears to directly modulate the effects of both positive and negative regulatory inputs to Pol II, both from enhancer bound activators and repressors and directly from signal transduction pathways. It is currently believed that a freestanding form of a Mediator complex, distinct from Pol II holoenzyme, is mainly responsible for the additional coregulatory function of this multisubunit complex in vivo (Bhoite et al., 2001; Bryant and Ptashne, 2003; Kuras et al., 2003).

It has become increasingly clear that the cellular machinery regulates many steps of transcription besides PIC formation and function. There are other aspects of Pol II function that are modulated besides its localization to the PIC. The C-terminus of the largest subunit, Rpb1, contains an array of seven amino acid repeats called the C-terminal domain, or CTD (Corden et al., 1985; Dahmus, 1983). There are 27 repeats in yeasts and 52 in humans. The consensus peptide repeat sequence is YSPTSPS (Nonet et al., 1987a; Nonet et al., 1987b). Each repeat can be post-translationally modified on the serine, threonine, and tyrosine residues (Feaver et al., 1991). Differential modification status distinguishes free Pol II (unphosphorylated) from actively transcribing Pol II (elongating

transcripts) and the enzyme present in PICs (Payne et al., 1989). The elongating polymerase is extensively phosphorylated on serine2 and serine5 residues within each repeat, a partially active form is modified on the second serine alone, and the inactive, PIC restricted, polymerase is hypophosphorylated. Ser5 phosphorylation is added by kinase activity in TFIIH while Ser2 is modified by another kinase, pTEFb in mammals and CTKs 1 and 2 in yeasts (Lu et al., 1992; Marshall et al., 1996; Marshall and Price, 1995). The combination of modifications and the extent of overlap throughout the 27 to 52 repeats has major affects on the repertoire of factors that will associate with the CTD, and a 'CTD code' hypothesis exists which predated the 'histone code' model described earlier (Buratowski, 2003). These models of PTM function specify combinatorial affects of individual modifications on different residues, with different combinations resulting in potentially diverse functional outcomes. Thus potential exists for widely different species of Pol II isoforms, each with different combinations of modifications on individual Pol II molecules. Also, as in the case of histone modifications, a major outcome of differential modification is altered capacity for protein-protein interactions between the CTD and additional regulatory factors (Brookes and Pombo, 2009). For example Ser5 modification is important for association of the 5' mRNA capping machinery with template-engaged, promoter-proximal Pol II, whereas Ser2 status influences the association of pre-mRNA splicing factors and 3' PolyA tailing machinery during transcription elongation (Cho et al., 1997). Phospho-Ser2 stimulates association of histone methyltransferases that put chromatin into a state compatible with Pol II transcription elongation (Stock et al., 2007). The CTD is also regulated by the prolyl isomerase Pin1 in mammals and this affects and is affected by the extent of

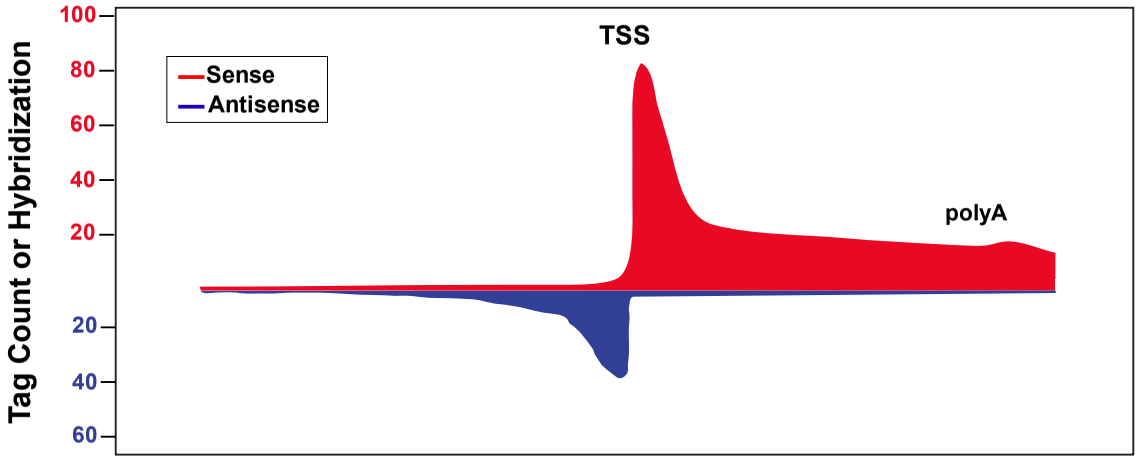
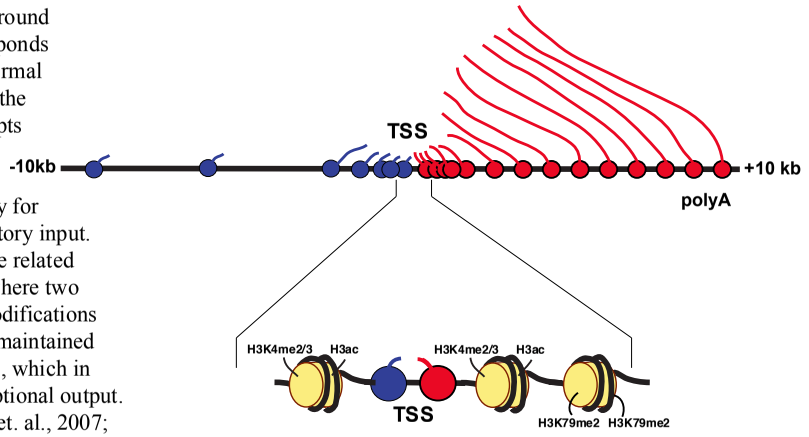


Figure 1.4: An alternative to the step-wise model of PIC Formation. Core et. al. performed deep sequencing analyses of nascent cellular RNAs and found that peaks of transcriptional activity localize around the promoters of 30% of human genes. Top, peak height corresponds to abundance of nascent transcripts with red peaks indicating normal genic transcripts and blue indicates transcripts originating from the non-genic, or antisense, direction. Observance of these transcripts indicates that many genes have transcriptionally engaged polymerase at all times, and hence functional PICs. However, the template engaged polymerases are deficient in their capacity for elongation; this characteristic may provide for additional regulatory input. One purpose of the bidirectionally active PICs and Pol II may be related to a characteristic chromatin structure around core promoters, where two transcriptionally permissive nucleosomes with characteristic modifications flank core promoters. The status of these nucleosomes may be maintained by the partially active, bi-directionally transcribing polymerases, which in turn keeps genes in a state poised for rapid increases in transcriptional output. Results illustrated here are in accord with findings of Guenther et. al., 2007; Nechaev et. al., 2010; Preker et. al., 2008; Seila et. al., 2008.



phosphorylation, which collectively regulate CTD association with pre-mRNA processing factors (Albert et al., 1999; Kops et al., 2002; Xu et al., 2003; Xu and Manley, 2007). The regulation of CTD-directed factor activity, including Pin1 and the mentioned kinases and equivalent phosphatases, as well as all of the phosphorylated/prolyl isomerised-CTD associated factors, are potential points of regulatory input after PIC formation.

Very recent studies indicate that regulation of Pol II at a post-recruitment/post-PIC formation step is very prevalent in regulating individual gene expression, and that a low level of transcription is common at all times on many genes (see **Figure 1.4**). These findings have their basis in genome-wide localization of Pol II and phospho-isoforms, and more importantly by localization of Pol II activity, also in a genome-wide manner (Core and Lis, 2008; Core et al., 2008; Guenther et al., 2007; Kininis et al., 2009; Muse et al., 2007; Nechaev and Adelman, 2008; Nechaev et al.; Preker et al., 2008; Preker et al., 2009; Seila et al., 2008; Seila et al., 2009; Zeitlinger et al., 2007). The implication is that most of the Pol II is localized around gene promoters along with some or all of the basal factors and possibly the mediator, and that these Pol II molecules are actually transcriptionally active because they generate high levels of very short transcripts and a very low level of intermediate to full length transcripts. Template-engaged promoter-proximal Pol II is co-localized with negative elongation factors such as NELF and DSIF, which can restrict the ability of Pol II to completely escape or clear the promoter. Complementary biochemical studies indicate that the rate limiting step for conversion of Pol II to the fully active form in vitro includes pTEFb mediated phosphorylation of Ser2 in CTD repeats, which promotes dissociation of NELF, association of numerous other

positive chromatin-directed factors with the CTD as mentioned above, and the phosphorylation-mediated conversion of DSIF into a positive elongation factor (Chen et al., 2009; Hartzog et al., 1998; Kim et al., 2003; Kim et al., 2001; Narita et al., 2003; Renner et al., 2001; Wada et al., 2000; Wada et al., 1998a; Wada et al., 1998b; Wu et al., 2003; Yamada et al., 2006; Yamaguchi et al., 1999; Zhu et al., 2007). The reversal of CTD phosphorylation events, by phosphatases Fcp1 in yeasts and SCP in mammals, is associated with transcriptional termination and recycling of the resulting hypophosphorylated enzyme into additional rounds of initiation, promoter escape, elongation, and termination (Lin and Dahmus, 2003).

The discussion of biochemical rate-limiting steps has therefore expanded from the initial model of stepwise PIC formation to include Mediator involvement and the modification of the Pol II CTD that in turn facilitates transcriptional elongation. In other words, both factor association with promoters and modification status of those factors can be important. The concept of factor association assumes a chromatin context where nucleosomes present in and around gene regulatory elements do not prohibit downstream binding events. For a generalized example of how chromatin both influences and is influenced by gene regulation and transcriptional activation, see **Figure 1.5**. The finding that dozens of unique coregulatory complexes contain chromatin-directed enzymatic activity underscores the dramatic influence that the nucleosome can have on gene regulation (Tsukiyama and Wu, 1997; Wu, 1997). Many such coregulators are molecular motors that hydrolyze ATP and use the liberated energy to physically relocate nucleosomes to different locations on enhancers or promoters or even completely evict

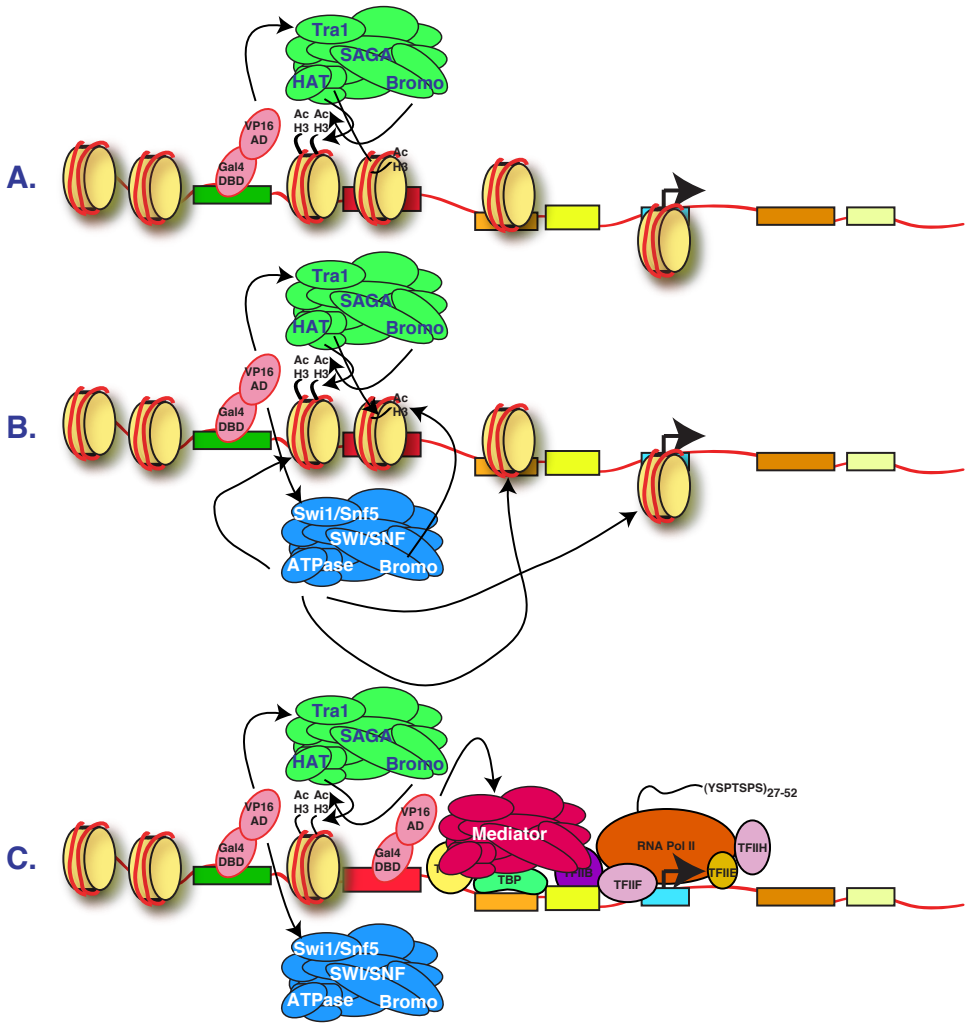


Figure 1.5: Chromatin regulation of enhancer and core promoter accessibility. **A**. Nucleosomes restrict factor access to enhancer and core promoter DNA elements. Transfactor activation domains, such as the chimeric Gal4-VP16 fusion, directly interact with chromatin modifying enzymes like SAGA where Tra1 is one specific subunit contacted by activators in vivo (Bhaumik et. al., 2005). Activator mediated recruitment of SAGA results in localized histone H3 hyperacetylation mediated by the Gcn5 acetyltransferase and SAGA is further stabilized by binding to acetyl-lysines via its bromodomain, also contained in Gcn5. **B**. Activation domains also directly interact with SWI/SNF complex subunits through the Swi1 and Snf5 subunits. (Prochasson et. al., 2003). Like SAGA, SWI/SNF is stabilized both by contacts with activator but also through contacts of intrinsic bromodomains with acetyl-lysines in nucleosomes. (Hassan et. al., 2002). Localization of SWI/SNF activity at enhancers results in remodelling of enhancer and promoter nucleosomes by the SWI/SNF Snf2 ATPase. **C**. The remodeled enhancer and promoter is now more accessible to binding factors. Additional transactivator binding can result in recruitment of additional coregulators such as Mediator, also by direct protein-protein interaction. In turn PIC components will assemble either as a holoenzyme or through stepwise association with the promoter DNA. The removal of repressive chromatin allows transcription to occur. Note that in vivo correlates to this amalgamated example exist, for example on the yeast *GAL1-10, PHO5*, and *HO* genes, (Bhaumik et. al., 2002; Reinke et. al., 2003; Kuras et. al., 2003; Cosma et. al., 1999); and the human PS2, p21, and PEPCK genes (Métivier et. al., 2003; Espinosa et. al., 2003; Li et. al., 2010).

nucleosomes from DNA (Tsukiyama et al., 1994). The archetypes of these factors include the SWI/SNF and ISWI families of ATP-dependent chromatin remodelers (Cairns et al., 1994; Cairns et al., 1996; Peterson et al., 1994; Tsukiyama et al., 1995; Vary et al., 2003). Chromatin remodelers can be both coactivators and corepressors, depending on the exact role of the nucleosome being mobilized on an individual gene.

Another class of coregulators includes the so-called 'writers of the histone code' that modify histones by acetylation, methylation, phosphorylation, ubiquitination, sumoylation, etcetera. The first described example is the Gcn5 acetyltransferase, which acetylates lysine residues on histone H3 (Brownell and Allis, 1995; Brownell et al., 1996). Gcn5 is also a 'reader' of the histone code, as it contains another domain called a bromodomain that has increased affinity for binding to acetylated H3 relative to unmodified H3 (Horn and Peterson, 2001; Marmorstein, 2001; Marmorstein and Berger, 2001). The equivalent 'code reader' in the methyltransferases is the chromo, chromoshadow, and/or plant homeodomain (PHD) in these complexes. These methyltransferases domains bind to methylated histones (Chen et al., 1999). Thus capacity within chromatin-directed coregulators for physical alterations of the chromatin is often combined with the ability to recognize introduced changes. It seems that the modification status can be self-stabilizing or self-propagating, since an initial modification event can promote the continued presence of the responsible enzyme, or even distinct enzymes that introduce other modifications but recognize the modification introduced by the first coregulatory enzyme. This may be important for long-term regulation of genomic loci or 'memory', by marking regions of the genome with a unique physical architecture (Kundu et al., 2007; Kundu and Peterson, 2009), also keeping in

mind that steady state abundance of a particular modification at a given genomic region is dependent upon the balance of activity between the 'writers,' 'readers,' and 'erasers' of histone modifications such as phosphatases, deubiquitinases, deacetylases and demethylases (Rundlett et al., 1996; Tsukada et al., 2006; Vidal and Gaber, 1991) (Daniel et al., 2004). The presence of particular histone modifications in enhancer or promoter nucleosomes can now be used as a surrogate marker of gene transcriptional status, active or inactive or somewhere in between, depending on the particular combinations of modifications enriched in the enhancer, promoter, and/or ORF (Karlic et al.). These observations are consistent with the long-established, genetically-defined concept of regions of the genome being described as transcriptionally permissive 'euchromatin' and repressed chromosomal regions as 'heterochromatin' In spite of this apparent targeting of coregulatory enzymatic activity to specific genome regions, it is important to remember that with the exception of a few of the basal transcription factors, most coregulators lack sequence specific DNA binding activity. Therefore it can be imagined that one possible mechanism of coregulator targeting to specific genomic regions relies upon their capacity for modified histone binding. While this is probably true to a limited extent, it is important to point out that much of the histone deposition in the genome occurs concomitant with DNA replication in S phase of the cell cycle. As it happens, most of the newly deposited histones are initially found as posttranslational modification-deficient isoforms. Therefore a 'pioneer' round of modification is undoubtedly necessary to facilitate later stabilization of coregulatory binding. Another fact is that ectopic placement of defined enhancer elements in genomic locales is sufficient to confer a heterologous pattern of histone modification (Bhaumik and Green, 2003). This is also

genetically dependent upon the activator proteins that bind to the enhancers. Altogether these facts suggest that there is a functional link between activator proteins and chromatin-directed coregulatory complexes. Like a huge portion of regulatory biology, these phenomena are directed by protein-protein interactions, in this case interactions between transactivators and transcriptional coregulators. An attempted summary of some of these interactions is illustrated in **Figure 1.5**. Before discussing the details on what is known about activator-coregulator interactions, it is appropriate to discuss biochemical characteristics of activator proteins.

Transfactors

All transfactors influence transcription by association with enhancers, most commonly through direct DNA binding in cis-elements. Another common feature of this diverse class of proteins is a modular structure and presence of several functionally distinct domains (Keegan et al., 1986; Ma and Ptashne, 1987a). At a minimum, a transfactor will usually have a DNA binding domain (DBD), responsible for sequence-specific association with cis-elements contained in enhancers, and an activation or repression domain, depending on whether the factor is an activator or repressor (Kadonaga et al., 1988). For clarity, activation domains are operationally defined by fusion to a heterologous DBD, resulting in a chimeric protein that is typically necessary and sufficient to stimulate expression of a reporter gene containing cis-elements recognized by the specified DBD.

The observation that many so-called ‘master regulatory genes’ of development and differentiation encode transfactors underscores both the importance and high level of

regulation of these proteins (Affolter et al., 1990; Muller et al., 1988; Otting et al., 1988; Qian et al., 1989). In general, transcription factors are among the most rare of all types of proteins in vivo, often present at just hundreds of molecules per cell, quite low considering that there are an estimated 30,000 Pol II enzymes per cell (Borggrefe et al., 2001; Kimura et al., 1999). A consequence of the low cellular abundance is that there is a large excess of potential binding sites present throughout the genome, meaning that all available sites will likely never reach saturation. Low transfactor steady state levels, a specific binding site composed of just a few base pairs, and a genome that can potentially provide an excess of millions of base pairs of DNA to compete with the specific binding site for limiting transfactor molecules are all indications that the affinity of a transfactor for its binding site is generally extremely high. However, all transfactors have some intrinsic non-specific binding affinity for DNA, but this affinity is typically three to four orders of magnitude lower than that for consensus binding sites. This general biochemical characteristic has been exploited to purify transfactors by site-specific DNA affinity chromatography techniques, which allowed the initial cloning of many transfactors prior to availability of genome sequences (Kadonaga and Tjian, 1986). It is estimated that there are more than 1500 genes in the mammalian genome that encode transfactors.

The necessity for an intrinsically high affinity for binding sites means that transfactor-enhancer interaction can be extremely stable. However, occupancy of enhancers by transactivators is a common point of target gene regulation. It is easy to imagine that presence of activation domains at enhancers could stimulate promoter activity and inappropriate gene expression, and thus binding is a logical event to set as

rate limiting for gene activation. Not surprisingly, nature has evolved myriad mechanisms for restricting transfactor activity. Transcriptional control of transfactor gene expression is often used in developmental pathways; translational control of transfactor mRNA is another obvious cellular mechanism since enhancers won't be bound by proteins that are not present (Hinnebusch and Natarajan, 2002). Transfactor mRNA can be kept inactive by sequestration away from the translational machinery via control of nuclear export or trafficking into so-called P-bodies, or bound by cytoplasmic factors such as microRNAs that inhibit translation. Likewise, decreased protein stability via post-translational modification such as ubiquitination can keep transfactor abundance low, but reversal of negative modifications will allow protein to accumulate in response to appropriate signaling events (Honda et al., 1997; Honda and Yasuda, 1999). Protein localization is another regulated aspect; dynamic phosphorylation status can provide a means to shift the balance between nuclear import and export, leading to nuclear depletion and cytoplasmic accumulation whenever gene transcription is inappropriate (Kaffman et al., 1994; O'Neill et al., 1996). Membrane tethering and regulated proteolysis can be used to direct subcellular localization (Hua et al., 1993; Wang et al., 1994; Yokoyama et al., 1993). Characteristics of the DNA binding domain can be directly regulated by modification, most commonly phosphorylation but also methylation or acetylation, which can increase or decrease sequence-specific binding affinity (Gu and Roeder, 1997; Huang et al.; Huang et al., 2006; Huang et al., 2007; Luo et al., 2004). Many transactivators are only functional in a multimerized form, often as homo- or heterodimers (Abate et al., 1990). Therefore, oligomerization status can be tightly controlled, by mechanisms including but not limited to regulation of heterodimer partner

abundance/stability, localization, or posttranslational modification, providing yet another means to limit specific DNA binding activity.

In instances where DNA binding is constitutive, activation domain potency can be modulated by PTM. Other proteins may engage the activator, blocking or competing with the ability of the activation domain to contact other responsive factors that would stimulate transcription (Ma and Ptashne, 1987b). The nuclear receptor family of transfactors are activated by appearance and binding of unique small molecule ligands, and there are typically dramatically different biochemical and cell biological differences between the apo- and ligand bound state of these proteins (Means et al., 1975). This can be affected by mechanisms such as those described above and also by many others. It has become increasingly evident that these proteins can even be converted from activator to repressor or vice versa in a ligand-dependent manner (Fondell et al., 1996a; Fondell et al., 1996b; Fondell et al., 1993). Along these lines transfactors are probably not strictly activators or repressors in a general sense; instead they function in widely different manners depending on environmental conditions and on which particular enhancer/gene they are binding (Kurtz and Shore, 1991). The complexity of regulation of a lone transfactor is generally extensive, since the variety mentioned above are only a small representation of those mechanisms probably in existence. Added to that, combined utility of different mechanisms is likely the order of the day.

Moreover, one must take into account the complexity of enhancers found in natural genes. Although there are only a few model genes for which the enhancer is perhaps completely characterized with respect to the proteins involved, such as the Interferon-β 'enhanceosome', we know that vertebrate enhancers can include

contributions from many distinct transfactors (Agalioti et al., 2000; Ford and Thanos; Munshi et al., 1999; Sheppard et al., 1999; Thanos, 1996; Yie et al., 1999a; Yie et al., 1999b). The quantity of transcriptional stimulation that an enhancer can provide is generally greater than the sum of that provided by the individual components. This indicates that separate enhancer-bound transfactors may be responsible for directing association with distinct coregulatory complexes. In other words different enzymatic activities can be directed into stimulating transcription through a variety of mechanisms (note chromatin as one possibility, see **Figure 1.2**). There may be different stages of gene activation, each of which is required in a specific order of occurrence, with distinct cohorts of coregulators functioning within each stage, and with the involvement of individual coregulators specified by one or a few of the enhancer-bound transfactors. Evidence for the existence of such phenomena is available from studies on several model genes in yeast and human cells. Chromatin immunoprecipitation (ChIP) studies looking at occupancy of numerous factors during a time-course of gene activation confirm a stepwise transcriptional program. The yeast *GAL*, *HO,* and GCN-pathway genes and the human PS2 and p21 genes provide just a few examples (Bryant and Ptashne, 2003; Cosma et al., 1999; Espinosa et al., 2003; Govind et al., 2005; Metivier et al., 2003). There appears to be both an order and a periodicity of coregulator association with target gene enhancers and promoters. In the case of the PS2 gene, this periodicity is correlated with a dynamic pattern of transfactor-enhancer association (Metivier et al., 2003). The complexity of the enhancer, coupled with the plethora of coregulators involved and added kinetic components during the activation process do not lend themselves to easily understanding the mechanism of coregulator targeting. Nevertheless, enhancer activity is

undoubtedly tied to protein-protein interactions between activators and coactivators. Simplified biochemical and genetic systems have provided many of the clues about these interactions (Swanson et al., 2003). A minimalist approach to investigation is necessary to confront the problem of coregulator targeting, since even though a limit case is to consider the interaction between one transfactor and one coregulator, there remains the fact that a typical coregulator is structurally and functionally complicated.

Transfactor-Coregulator Interactions

Having re-introduced the concept of a defined order of events in the pathway to gene activation, we will consider here the earliest possible rate-limiting steps. Transactivator occupancy of the enhancer is the most obvious. Even if they are present in the nucleus and competent for both DNA binding and transactivation, transfactors may not be able to access their cis-elements within the enhancer if it is occluded by nucleosomes. Partial nucleosomal occlusion of the enhancer provides a context where dynamic regulation and mechanism of co-regulatory function can be productively studied. The yeast *PHO5* gene is a physiological model for this situation (Almer et al., 1986). There are two high-affinity binding sites for the Pho4 activator, but the more promoter-proximal site is occluded by a nucleosome that prohibits occupancy by Pho4 (Venter et al., 1994). Full enhancer occupancy and gene activation requires intact binding sites, nuclear-localized Pho4, and histone acetyltransferase activity provided by Gcn5 and Esa1 (Barbaric et al., 2001) (Svaren et al., 1994). Kinetic studies indicate that a rise in histone H3 acetylation precedes full de-repression of the gene; this is presumably due to Pho4-mediated recruitment of acetylases (Reinke and Horz, 2003). Pho4

occupancy, as scored by genomic footprinting or ChIP, becomes apparent first on the accessible binding site, then on the previously inaccessible binding site (Barbaric et al., 2007). Pho4 binding of the second enhancer occurs coincident with the apparent removal of the blocking nucleosome, as scored by ChIP or nuclease accessibility. Similarly, the core promoter region becomes depleted of nucleosomes around this time. Removal of the histone acetylase by deletion of the *GCN5* gene lengthens the time necessary to de-repress the gene (Barbaric et al., 2001; Gregory et al., 1998). This is an example of an activator, chromatin, and histone acetylase-dependent gene activation (Svaren and Horz, 1997).

Detailed biochemical experiments with defined components substantiate these descriptive findings. Using a chromatin-reconstituted and immobilized DNA template containing Gal4 activator binding sites, a promoter, and nucleosome positioning sequences around the enhancer and promoter, Workman and colleagues were able to demonstrate activator dependent targeting and stabilization of the Gcn5-containing SAGA complex and the SWI/SNF chromatin remodeler (Hassan et al., 2001a; Hassan et al., 2001b; Hassan et al., 2002; Neely et al., 1999). Specifically, association of SAGA and SWI/SNF with the chromatinized template required pre-incubation with Gal4-VP16 activator. Subsequent addition of SAGA and SWI/SNF only resulted in binding of the complexes to the template if both template-bound activator and acetyl-CoA were present, indicating coregulator binding was stabilized by both activator and coregulator-mediated acetylation of histones on the template. Moreover, competition of Gal4-VP16 from the template using a molar excess of oligonucleotide binding sites destabilized SAGA and SWI/SNF binding. Not surprisingly, several groups were able to demonstrate direct

protein-protein interaction between Gal4 as well as other yeast activators and subunits of SAGA and SWI/SNF (Bhaumik and Green, 2001; Brown et al., 2001; Neely et al., 2002; Prochasson et al., 2003). Genetic disruption of the binding surfaces provided by SAGA and/or SWI/SNF subunits resulted in reduced transcription of target genes in vivo.

Perhaps the most convincing experiment demonstrating the interaction between an activator and a chromatin-directed coregulator again involved the yeast Gal4 protein and the SAGA complex. Bhaumik and colleagues used GFP fusion proteins and live cell imaging to demonstrate fluorescence resonance energy transfer, or FRET, between Gal4 and a single SAGA subunit (Bhaumik et al., 2004). When combined with the other genetic and biochemical experiments, this paints a clear picture of direct protein-protein interaction between activators and chromatin-directed coregulators. A variety of additional studies including other activators and coregulators suggest this is not uncommon; for example the related p300 and CBP coregulators that directly associate with the cAMP response element binding protein, CREB, are acetyltransferases whose activity is highly important for transcription in vitro and in vivo (Eckner et al., 1994; Kundu et al., 2000; Kwok et al., 1994; Scolnick et al., 1997). Connections between activators and chromatin-directed coregulators are illustrated in **Figure 1.5**.

The Mediator complex has been the subject of many studies seeking to identify direct activator-coregulator interactions. Using affinity purification of the thyroid hormone receptor from human cell lines, Roeder and colleagues identified a set of co-fractionating proteins with homology to yeast RNA Pol II holoenzyme subunits (Fondell et al., 1996b). The copurification of these proteins depended on the presence of thyroid hormone in the growth media of the cells used as the initial source of material, indicating

that only transcriptionally active thyroid hormone receptor engages these proteins, a property consistent with their identity as coregulators. Additional studies aimed at directly purifying these TRAPs, or thyroid receptor activator proteins, proved their coregulatory role and demonstrated its function in the form of a discrete high molecular weight complex. Studies by Tjian's lab characterized coactivators of the Vitamin D Receptor and Sp1 activators and identified a similar complex, termed DRIP or CRSP, respectively. In parallel, the Activator Responsive Complex, or ARC, was also identified by Tjian as providing direct responsiveness to several activators in vitro (Naar et al., 1998a; Naar et al., 1999; Naar et al., 1998b; Ryu and Tjian, 1999; Ryu et al., 1999; Taatjes et al., 2002). Further investigation has shown that multiple Mediator subunits provide interaction surfaces for activators, that these interaction partners can vary somewhat between different activators, and certain activators can likely use multiple domains to interact with more than one mediator subunit at a time (Ge et al., 2008; Grontved et al.; Malik et al., 2004). This observation is probably highly relevant in the broader relationships between various activators and coactivators.

In vivo studies using yeast demonstrated a set of very important observations about Mediator function. Ptashne and coworkers demonstrated that a fusion protein consisting of the Pho4 DNA binding domain and the Gal11 mediator subunit provided both chromatin remodeling and transcriptional activation at the *PHO5* gene in vivo (Gaudreau et al., 1997). There are several implications of this work. First, the fusion protein directed function from the enhancer and not from the core promoter. Second, chromatin remodeling was stimulated even though there was no indication that Mediator subunits could provide this function. Thus, Mediator was suggested to regulate the

function of the enhancer, chromatin remodeling, and the core promoter. It also seemed unlikely that Mediator would do this in the form of a holoenzyme complex, since no aberrant upstream transcripts were detectable as would be expected if Pol II colocalized with the Pho4-Gal11 fusion at the enhancer. Chromatin immunoprecipitation studies performed over a time course of gene activation later showed that Mediator subunits preferentially localize to the enhancers of the *HO* and *GAL* genes in yeast, and do so at time points preceding the association of Pol II with the promoter and ORF (Bhoite et al., 2001; Bryant and Ptashne, 2003). Collectively, these studies indicate a coregulatory role for Mediator, in many cases implicating direct interaction with activators, meaning that the Mediator functions to physically bridge or link enhancer with promoter, perhaps by interacting with both activators and basal transcription factors.

Interactions between activators and either chromatin-directed complexes or the Mediator complex were not the first postulated transfactor-coregulator protein-protein interactions facilitating communication from enhancer to promoter. At a point in history when only activators and basal transcription factors had been identified, it was logical to look at interactions between these components. As has already been mentioned here, the best example of a coactivator available at the time were the bacterial σ factors, which can specifically recognize the promoter, associate with the Polymerase, and directly interact with activator proteins. TFIID activity had been shown to act early in stepwise PIC formation, and was the only basal factor able to directly bind TATA box DNA found in promoters (Buratowski et al., 1989; Van Dyke et al., 1988). This led to comparisons between TFIID and σ. For these reasons the possibility of TFIID coactivator function was pursued early on (Abmayr et al., 1988).

TFIID: Requirement of Taf Proteins for Activator-Stimulated Transcription

Perhaps the longest recognized hallmark of TFIID is its ability to bind to the TATA box that is located in model promoters that were used as DNA templates in biochemical assays. Therefore specific DNA binding assays were used to guide fractionation of TFIID activity from cellular extracts. This proved very difficult in animal systems, probably owing to the fairly low abundance and instability of the factor. However, using soluble yeast extract as a starting material, a single polypeptide possessing specific TATA box binding activity was purified (Cavallini et al., 1988). Protein sequencing and cloning of the single-copy, essential yeast TBP gene quickly followed, and the recombinant protein was able to functionally substitute for the native protein in biochemical complementation of DNA binding and in vitro transcription (Hahn et al., 1989b; Horikoshi et al., 1989a, b). Importantly, the recombinant yeast protein supported these functions whereas all other reaction components were derived from human sources, demonstrating the high evolutionary conservation of protein activity between yeasts and human. Consistent with this idea, the sequence of yeast TBP protein was useful for cloning of TBP-encoding genes from several organisms including flies and humans (Gasch et al., 1990; Hoey et al., 1990; Hoffmann et al., 1990; Lichtsteiner and Tjian, 1993; Muhich et al., 1990; Peterson et al., 1990; Peterson and Tjian, 1993).

Interestingly, in no case could any of the recombinant TBPs support activator-stimulated transcription in vitro, whereas less pure fractions containing native, high molecular weight form of TBP could do so. This observation, with its origins in a

'failed' experiment, led to some of the first speculation of the identity of coactivators in the eukaryotic transcription system (Pugh and Tjian, 1990). Properties of TBP again became useful in pursuit of these elusive coactivators, as it was possible to generate high affinity, high specificity antibodies using the recombinant protein as an antigen. Use of these antibodies allowed the immuno-affinity purification of a protein fraction containing TBP and 10 to 15 TBP associated factors, or TAFs, that could provide both core promoter recognition and response to activators in an in vitro transcription system (Dynlacht et al., 1991; Meisterernst and Roeder, 1991; Pugh and Tjian, 1991; Tanese et al., 1991). As with TBP, Taf-encoding cDNAs were obtained from different organisms, first from flies, followed by human and lastly from yeast a few years later; again this information indicated a high overall degree of Taf sequence conservation (Chiang and Roeder, 1995; Dynlacht et al., 1993; Gill and Tjian, 1992; Goodrich et al., 1993; Hisatake et al., 1993; Hisatake et al., 1995; Hoey et al., 1993; Hoffmann and Roeder, 1996; Klemm et al., 1995; Kokubo et al., 1993a; Kokubo et al., 1994a; Kokubo et al., 1993b; Kokubo et al., 1993c; Kokubo et al., 1993d; Moqtaderi et al., 1996b; Poon et al., 1995; Reese et al., 1994; Ruppert et al., 1993; Takada et al., 1992; Verrijzer et al., 1994; Weinzierl et al., 1993b; Yokomori et al., 1993). Characterization of purified recombinant Tafs revealed that some could engage in binary interactions with activators in vitro, results that at least partially explained the basis of activator responsiveness (Goodrich et al., 1993; Hoey et al., 1993; Kashanchi et al., 1994). The existence of a complex functionally analogous to bacterial σ, conferring both activator binding and sequence-specific promoter binding within the PIC, was the first example of a factor present in

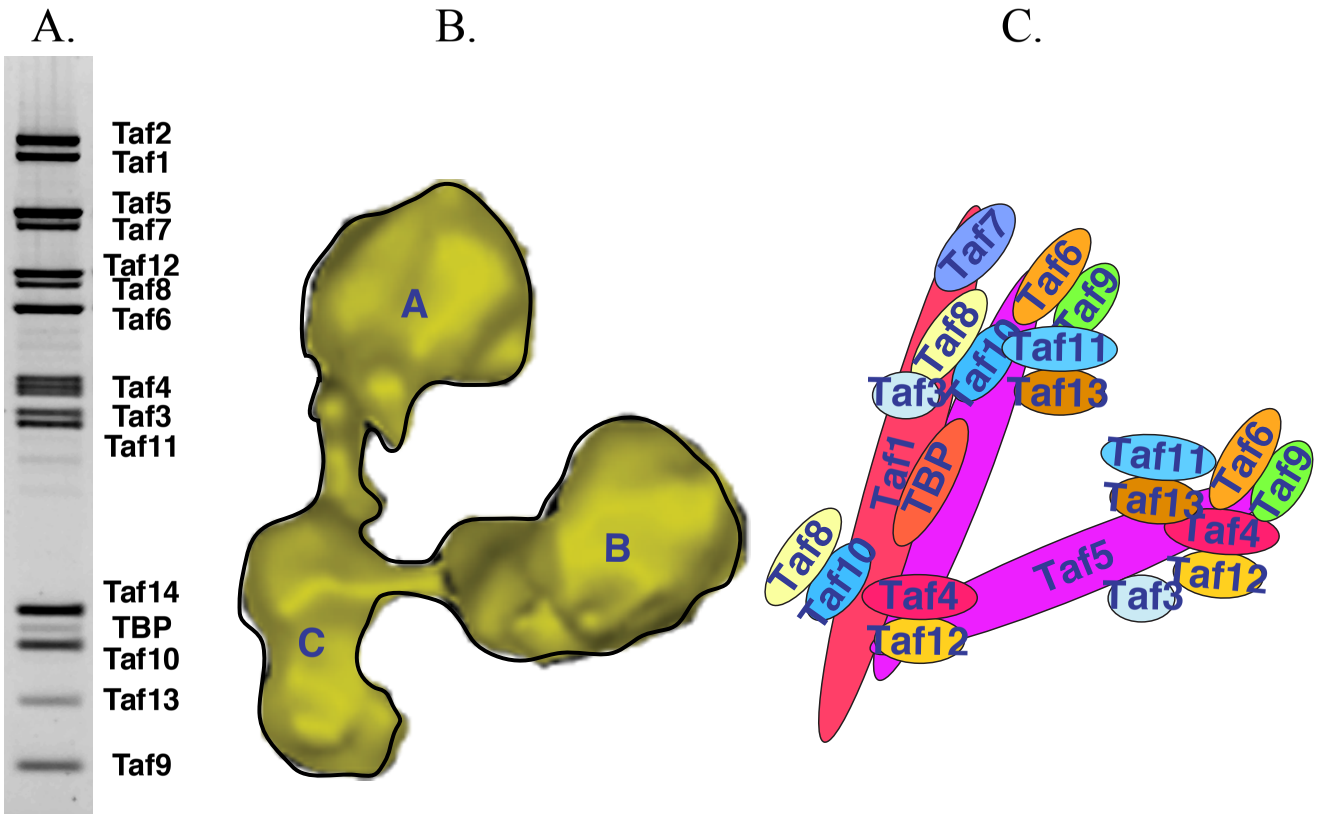
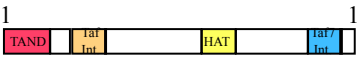
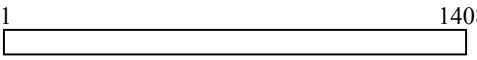
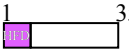
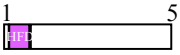
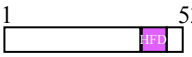


Figure 1.6: Connection between TFIID-Taf composition and structure. **A**. Sypro Ruby-stained SDS-PAGE analysis of purified yeast TFIID showing the mobility of individual subunits. (Sanders et. al., 2003). **B**. Refined image of fixed and negatively stained yeast TFIID as obtained by electron microscopy. (Leurent et. al., 2002). The characteristic tri-lobed, horseshoe shape indicates three primary densities of Taf location within the complex, termed lobes A, C, and B starting from the top and in anti-clockwise direction. Note that human and yeast TFIID share a similar global architecture. (not shown, Grob et. al., 1999). **C.** Yeast TFIID is the only example in which systematic localization of individual Tafs has been carried out. Taf7 is exclusively located in the C lobe, while the single molecule of Taf1 extends from the N- to C-terminus in both C and A lobes. Other Tafs are typically resident in two or more lobes, consistent with their increased stoichiometry in the complex and with the propensity of histone fold-containing Tafs to engage in selective heterodimerization with other Tafs. Importantly, note that Taf4, Taf12, and the Taf5 N-terminus colocalize in the C-lobe; this feature is likely important for interaction of the Rap1 transactivator with TFIID in budding yeast. Illustration is representative of data from Leurent et. al., 2002, 2004.

	Evolutionarily Conserved and/or Known Functional Domains are Shaded	TFIID Specific?	Moles Taf/ Moles TFIID	Functional Characteristics
Taf1		Yes	1	TBP Interaction/ Negative Regulation, Scaffold for TFIID Assembly, Putative Acetyltransferase
Taf2		Yes	<1	Very little structural/functional information. Sensitive to degradation during isolation of TFIID
Taf3		Yes	1	Histone Fold Domain Mediates Dimerization with Taf8 and Taf10
Taf4		Yes	3	Atypical bifurcated Histone Fold Domain Mediates Dimerization with Taf12
Taf5		No	2	N-terminal domains possibly allow dimerization, C-terminal WD40 Repeats Interact with Taf6/9 heterodimers
Taf6		No	2	Histone Fold Domain Mediates Dimerization with Taf9
Taf7		Yes	1	N-terminal Domain interacts with sub-stoichiometric Bdf1 component. Interacts with Taf1 C-terminus
Taf8		Yes	1	Non-essential Histone Fold Domain Mediates Dimerization with Taf3
Taf9		No	>3	Histone Fold Domain Mediates Dimerization with Taf6
Taf10		No	2	Histone Fold Domain Mediates Dimerization with Taf3
Taf11		Yes	1	Histone Fold Domain Mediates Dimerization with Taf13
Taf12		No	2	Histone Fold Domain Mediates Dimerization with Taf4
Taf13		Yes	1	Histone Fold Domain Mediates Dimerization with Taf11
Taf14		No	>3	Only non-essential Taf in TFIID, present in many other complexes
TBP		No	1	TATA-binding protein, major DNA binding activity in yeast TFIID. Present in many other complexes regulating transcription by RNA Pol I, II, and III

seemingly all eukaryotes (Conaway and Conaway, 1990a). For an overview of TFIID composition, structure, and function; see **Figure 1.6** and **Table 1.1**.

Transfactor-TAF/TFIID Interaction

Upon consideration of TFIID complex characteristics, a paradox became quickly apparent, although it once again had its basis in comparisons to the bacterial system. The TFIID complex containing TBP and Tafs clearly provides promoter recognition and facilitates direct interaction between activators and the PIC. But why are as many as 15 distinct proteins required for TFIID function in eukaryotes when just one or two proteins (i.e. σ, α) can communicate with a plethora of activators in bacteria? Even in light of current information, such as the ability of activators to contact multiple subunits of coregulator complexes like Mediator and TFIID, the high complexity is not well understood, but the basis of this complexity surely lies in the diverse functional capabilities among the different Tafs. TFIID can do quite a lot more than bacterial σ, and its additional features are likely put to use in the more complicated mechanisms of eukaryotic gene regulation.

In vitro TFIID-Taf interaction with activators is long recognized and biochemical studies showed that transfactors lacking an activation domain could not utilize TFIID to stimulate in vitro transcription (Gill et al., 1994). Partial reconstitution of TFIID complexes from a subset of recombinant Taf proteins provided the first evidence for specific targets of activation domains, since omission of certain Tafs also resulted in lack of activator responsiveness during in vitro transcription assays (Chen et al., 1994; Weinzierl et al., 1993a). Many studies have since shown specific in vitro interactions between Tafs and a variety of activators (Kashanchi et al., 1994; Thut et al., 1995).

These interactions were first described as important for mediating 'recruitment' of TFIID to promoters in vitro (Dynlacht et al., 1991). TBP binding affinity for the TATA box is strong, but the strength of TBP interaction with non-TATA sequence is also significant, as compared to the binding affinities of typical activators for non-binding site DNA. As such, activator-TFIID interaction was proposed to increase the affinity of TFIID for TATA box, perhaps by affecting TBP conformation or location within the complex (Coleman and Pugh, 1995; Horikoshi et al., 1992; Lee et al., 1991; Nikolov et al., 1995). Additionally, the N-terminus of Taf1 from various organisms contains a domain that engages TBP through the same surfaces important for TATA binding, thereby inhibiting TBP-DNA interaction (Kokubo et al., 1994b). It has been postulated that activator-Taf interaction promotes the rearrangement of this configuration such that TBP in TFIID is made available for DNA binding (Albright and Tjian, 2000; Nishikawa et al., 1997; Oelgeschlager et al., 1996). It should be pointed out that any models involving modulation of TFIID-TATA DNA binding activity would involve TFIIA. The binding of the TFIID complex to TATA is not readily observed in vitro unless TFIIA is also present or very specific biochemical conditions are provided, specifically the inclusion of super-physiological levels of magnesium ions. (Geiger et al., 1996; Yokomori et al., 1994; Zerby and Lieberman, 1997).

The limitation of proposals involving activator-mediated TFIID-TATA recognition is that the majority of promoters in eukaryotic genomes do not contain a consensus TATA box. As such, models solely involving activator-mediated TBP-TATA interaction are deficient owing to the lack of general applicability. But how then can TFIID engage TATA-less promoters? Recombinant Taf2 and Taf6/9 dimers have been

shown to directly interact with INR and DPE core promoter elements, respectively (Burke and Kadonaga, 1996, 1997; Chalkley and Verrijzer, 1999; Emami et al., 1997; Kaufmann et al., 1998; Kaufmann and Smale, 1994; Kaufmann et al., 1996; Martinez et al., 1994; Verrijzer et al., 1995; Verrijzer et al., 1994). Taf-INR or DPE interaction could possibly be influenced by activator-TFIID interaction. But again, many promoters lack these cis-elements. Furthermore, there is no evidence that non-TATA core promoter cis-elements exist in lower eukaryotes, although existence of a functional equivalent cannot presently be ruled out.

It is possible that TFIID could localize to chromatinized promoters in the absence of Taf-DNA binding. The Taf1, Taf3, and Taf5 subunits possess respective bromo-, PHD finger, and WD40 domains. The Taf1 bromodomain potentiates high affinity binding to acetylated histone H4 peptides in vitro; likewise the Taf3 PHD finger recognizes H3 triply methylated on lysine four (Jacobson et al., 2000; van Ingen et al., 2008; Vermeulen et al., 2007). Taf5 WD repeats could also bind methylated histones or nucleosomes. The chromatin marks recognized by Taf1 and Taf3 are generally associated with transcriptionally permissive or active genomic regions (Guenther et al., 2007). In theory, activators, assuming the requisite histone modifications are also present, could modulate TFIID association with certain nucleosomal gene promoters.

Furthermore, it appears that the Taf1 bromodomain can bind non-histone acetyllysine-containing substrates, the example being acetylated p53. This interaction appears to facilitate Taf1 interaction with the p21 enhancer, potentially resulting in the formation of a DNA loop between p53-localized, enhancer bound Taf1 and other TFIID Tafs already present at the core promoter (Li et al., 2007). Taf1 also appears to possess

histone acetyltransferase activity, thus TFIID is potentially both a 'writer' and a 'reader' of histone/chromatin modifications (Mizzen et al., 1996). As in the case of the Taf1 bromodomain, Taf1-mediated acetylase activity may not be limited to histone substrates; for example p53 or other GTFs might be acetylated by Taf1 (Imhof et al., 1997).

There are many caveats to the observations suggesting connections to chromatin. First, these putative TFIID-chromatin interactions remain primarily untested in the context of TFIID, since all of the experiments described above were conducted with isolated Tafs. Second, these TFIID-histone interactions were not tested in the context of intact nucleosomes. Third and finally, the responsible domains in Taf1 and Taf3 are not found in Tafs from all species, so it is unlikely that chromatin targeting of TFIID is absolutely conserved, at least not through Taf1 and/or Taf3 dependent mechanisms. Nevertheless, models where TFIID affinity for promoter DNA elements or promoter chromatin is increased by interaction with activators remain attractive for a number of reasons, even if as yet unproven in vivo. Gene activation could occur only when activator is present, when appropriate histone modifications are present, or both. As always, a combination of activator status, promoter element accessibility, and histone status could all contribute to establishing TFIID-promoter association as rate-limiting. The development of more powerful genetic reagents in higher eukaryotes is the roadblock that prevents the proof of activator-mediated TFIID recruitment. Very specific mutations in activators and their Taf binding partners are needed that selectively disrupt function of the protein surfaces involved. The prediction would be that PIC formation and transcription would be inhibited due to reduced TFIID association and/or function on promoters directly regulated by the activator in question.

High-resolution structural information would be quite useful to direct mutational analysis of activator-Taf interaction surfaces, but this has remained elusive in the case of TFIID, primarily because of its large size and subunit complexity. There are a few x-ray structures of individual Taf domains, primarily the histone-fold domain pairs, and several high-resolution structures have been solved for TBP (Chasman et al., 1993; Geiger et al., 1996; Kim et al., 1993; Nikolov et al., 1995; Nikolov et al., 1992). Structures of the Taf1 bromodomain, the heterodimerized histone-fold domains of Taf4 with Taf12 and Taf6 with Taf9, and portions of the Taf4 and Taf5 amino termini have all been solved (Bhattacharya et al., 2007; Birck et al., 1998; Jacobson et al., 2000; Romier et al., 2007; Wang et al., 2007; Werten et al., 2002). Most of these studies suffer the same limitations of any experiments that examine Taf characteristics outside the context of both the intact proteins and TFIID holocomplex; if framed within TFIID those findings may be dramatically different on the basis of three-dimensional conformation. However, the Taf1 structure was actually crucial for identifying residues in Taf1 responsible for binding acetylated p53, and unmodifiable p53 variants and bromodomain-substituted Taf1 forms were put to use in ChIP experiments to look at changes in Taf and/or activator occupancy on the p21 enhancer and promoter during a time course of gene activation (Li et al., 2007).

As discussed above, p53-Taf1 interaction is arguably the only example of activator-mediated Taf recruitment resulting from direct protein-protein interaction that is supported *both* by solid in vitro biochemical and in vivo molecular genetic data. Unfortunately, there are numerous limitations to the general applicability of the p53-Taf1 example. First and most obvious, this again concerns an activator-Taf interaction and not

necessarily one of an activator-TFIID complex interaction. By necessity, this must involve subcomplexes of TFIID and de novo assembly of the complex in situ on promoters. There is some evidence for these TFIID subcomplexes but the physiological significance is uncharacterized and unknown, except in the in vivo context of p21 transcription (Demeny et al., 2007). Secondly, this mechanism is probably not generally utilized, even on the p21 gene, since different p53-controlled transcriptional programs are apparently employed depending on the induction method. Specifically, UV irradiation promotes Taf1 association with the p21 enhancer and promoter, but other genotoxic agents invoke a transcriptional response that does not involve Taf1 (Donner et al., 2007; Espinosa et al., 2003). Third, a p53- and Taf1-dependent gene-looping event remains speculative, because this type of event is very difficult to document in vivo, and is dependent on the use of a technically challenging, gene-specific protocol called the chromosomal conformational capture (3C) assay (Ansari and Hampsey, 2005; Laine et al., 2009; Singh and Hampsey, 2007); such experiments have not yet been reported for the p53-Taf1 interaction on p21. Thus the simultaneous involvement of Taf1 at both enhancer and core promoter in this context has not truly been confirmed. It should also be noted that in the context of TFIID, p53 does not interact with just Taf1, but also Taf5, Taf6, and TBP (Liu et al., 2009). However, DNA was not included in these particular biochemical experiments, and this could alter the spectrum of p53-Taf interactions.

An understanding of overall three-dimensional TFIID structure has been made available by electron microscopy and digital reconstruction. There are now low-resolution structures available for TFIID derived from human, budding yeast, and fission yeast (Andel et al., 1999; Brand et al., 1999; Elmlund et al., 2009). Overall similarity of

the three-lobed, horseshoe shape between species is observed, not unexpected given the conserved polypeptide sequences of the subunits. Structural methodology involving difference mapping has been used to roughly localize TBP and all 13 integral Tafs within budding yeast TFIID (Leurent et al., 2002; Leurent et al., 2004; Papai et al., 2009). This data is consistent with a modular TFIID organization suggested by earlier biochemical experiments, with histone-fold dimers residing in all three lobes and nucleating around two Taf5 densities, with a single TBP-Taf1-Taf7 module on the outside of the half of the complex containing the so-called A and C lobes (Chen et al., 1994; Guermah et al., 2001; Hoffmann et al., 1996; Oelgeschlager et al., 1996). This defined structure, and the observed presence of slightly different context-dependent conformations of TFIID, suggest the existence of intrinsic flexibility within the complex that is consistent with speculation of a 'hinge' function underlying the three-lobed configuration (Elmlund et al., 2009; Grob et al., 2006; Liu et al., 2008). Indeed, TFIID appears capable of adopting varying degrees of 'open' to 'closed' configurations. Consequently, models of activator-TFIID interaction in TFIID recruitment could mechanistically include alteration of TFIID conformation as a result of interaction with activators, and these conformational changes could then alter the binding to core promoter elements such as TATA, DPE, or INR or to PTMs on chromatin or other promoter-bound proteins, leading to stabilized TFIID-promoter association. Alternatively, dynamic conformational changes might affect many other characteristics of TFIID, like intrinsic enzymatic activities, or recruitment of other GTFs, Pol II, and/or additional coregulators (Dikstein et al., 1996; Imhof et al., 1997; Mizzen et al., 1996; O'Brien and Tjian, 1998; Pham and Sauer, 2000).

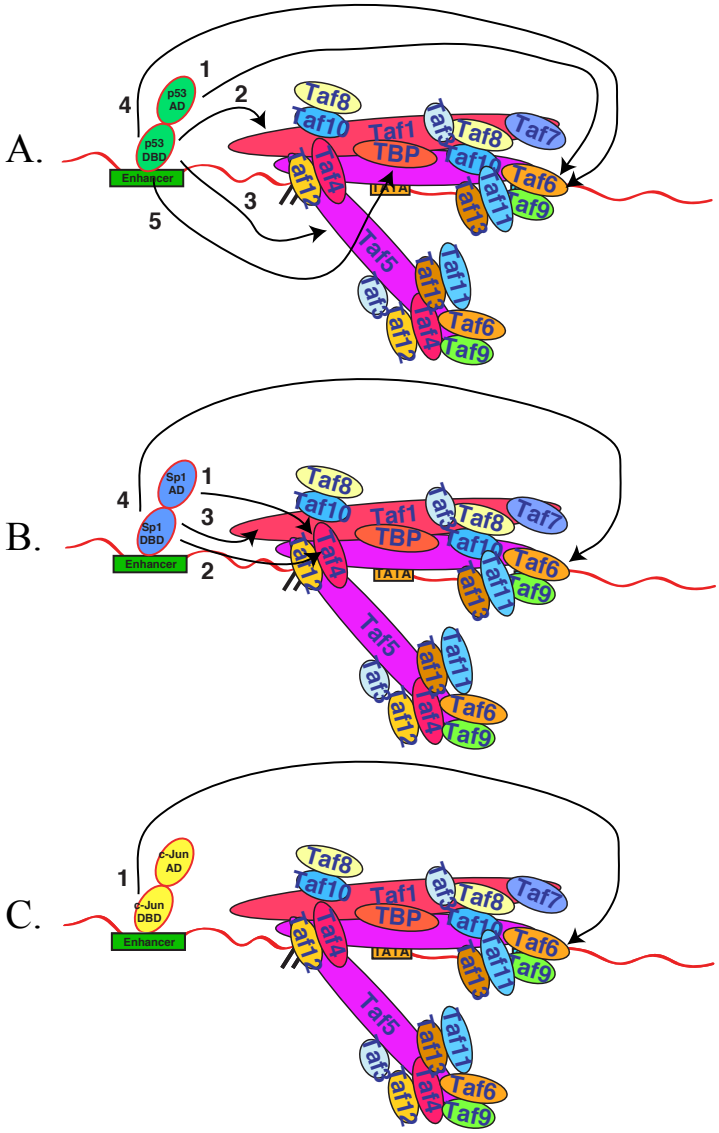


Figure 1.7: Activator-Taf Interaction in the context of TFIID. Purifed human TFIID was saturated with p53, Sp1, or c-Jun activators. Crosslinking and EM studies indicated the Taf components bound by each activator, and the activator domains involved. Representation of data from Liu et. al., 2009. **A.** p53 activator contacts Taf6 using its activation domain, while Taf1, Taf5, Taf6, and TBP are contacted by the p53 DBD. Importantly, it is not clear in any example whether each of the Taf molecules super-stoichiometric in the complex is engaged with activator or if there is some specificity of activator-Taf interaction based on location in the complex. **B.** The Sp1 activation domain contacts Taf4, while the DBD again contacts Taf4 but also Taf1 and Taf6. **C.** the c-Jun DBD selectively contacts Taf6. Note that DNA was actually not included in these binding assays, so the existence of activator DBD contacts with TFIID may be altered in that physiological context. Any conformational change happening in TFIID as a result of activator contact may occur only in the presence of core promoter DNA, a possibility that so far remains untested. Importantly, this study did not address the individual Taf domains involved in TFIID-activator interaction. Nevertheless this study is the only one besides that of Garbett et. al., 2007 that addresses activator-Taf interaction within the context of TFIID.

There is already a strong basis for activator-mediated conformational changes in coactivator structure. Perhaps the best example comes from low-resolution analysis of the Mediator complex in the absence and presence of several different activators and the RNA Pol II CTD (Naar et al., 2002; Taatjes et al., 2002; Taatjes et al., 2004). Activator-specific conformations were observed, consistent not only with interaction being achieved via different Mediator subunits, but also with potentially diverse functional outcomes within activator- and gene-specific transcription. Accordingly, EM difference mapping showed that c-Jun, Sp1, and p53 activators, which engage in binary interactions with only partially overlapping subsets of Tafs, localized to distinct sites in TFIID (Liu et al., 2009). In contrast to the Mediator studies, none of these activators induced dramatic conformational changes in TFIID comparable to those observed in activator-Mediator interaction. As summarized in **Figure 1.7**, these results regarding activator-TFIID interaction are surprising in some ways, but there are also technical limitations that could be influencing the interpretation of results. First, images were obtained from fixed and negative stained samples in low resolution. Harsh sample preparation could have disrupted conformational characteristics of TFIID, and subtle alterations could have escaped detection at the resolution achieved. Second, it is conceivable that TFIID interaction with core promoter elements or chromatin could contribute to conformational differences while in presence of activators. In other words, activators or other binding surfaces could synergize to obtain the final allosteric effects on the complex. Obviously this is the more physiological context, but neither DNA nor chromatin was used for analysis of TFIID binding with c-Jun, Sp1, or p53.

So while one must be careful not to trivialize the existing studies and findings, much of the current body of knowledge of activator-TFIID interaction suffers from a common weakness; absence of one or more components that would be found at genes in cells. Investigation of interactions using isolated Tafs instead of the complex is just as perilous as studying the complex in the absence of DNA or chromatin, or even additional factors like TFIIA. One might say that what has been missing is a combined approach not reliant upon one particular methodology and/or derivatives thereof, but instead the use of techniques of genetics, biochemistry, cell biology, and structural biology in a complementary way. This is a formidable undertaking but it actually has been demonstrated in other contexts; namely in defining mechanism(s) of bacterial activator-RNA Polymerase interaction. Once again I will use bacterial transcription to illustrate the diversity in how these interactions can be achieved, and suggest how this information might be applied to direct a cohesive strategy for characterizing activator-TFIID cooperation in eukaryotes.

Additional Paradigms from Bacterial Models

In *E. coli*, roughly 300 transcription factors can directly regulate some of the 2000 different promoters in the genome (Browning and Busby, 2004). This regulation is accomplished by direct interaction of the activators with ~8 potential isoforms of theholo-RNA polymerase, each differing on the basis of the σ factor they contain. Some genes use just one transcription factor to up-regulate polymerase activity on the promoter. This regulatory event is referred to as 'simple activation' and can be further categorized into three basic mechanisms (diagrammed in **Figure 1.8**). In class I activation, the binding

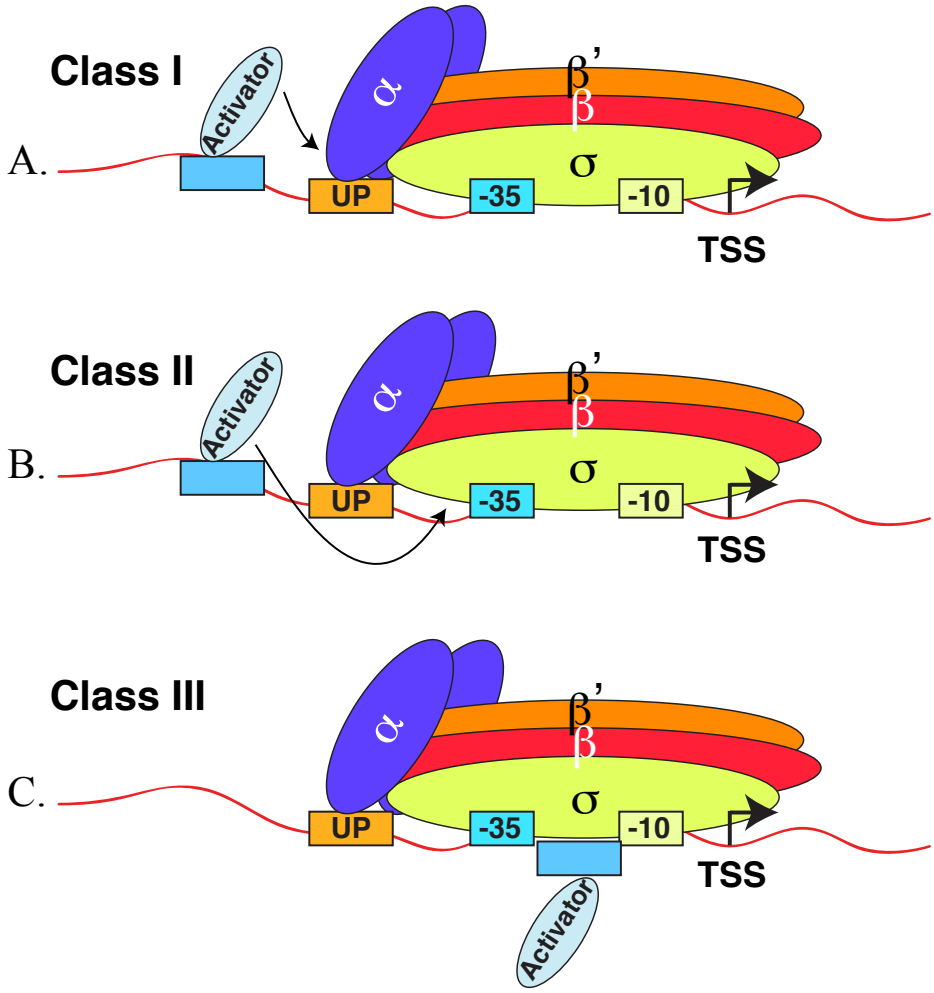


Figure 1.8 Simple Activator-Coregulator Interactions as exemplified by bacterial activator-RNA Polymerase contacts. Most situations involve a stabilization or 'recruitment' of the five subunit polymerase by activator to the typical core promoter elements UP, -35, and -10. **A.** Class I activation: direct protein-protein interaction between activator and the α polymerase subunit. **B.** Class II activation: as in A except the σ polymerase subunit is directly contacted. **C.** Class III activation: The event of DNA binding by activator induces a rearrangment in the core promoter DNA, which is more favorable for the association of polymerase. Note that in contrast to the situation for activator-TFIID contact, here it is more common for a single activator to contact just one subunit of the coactivator, in this case the RNA polymerase itself. Adapted from Browning and Busby, 2004. Compare these simplified examples with either of the three activator-TFIID interactions described in Figure 1.7. The possibile mechanisms of activator-TFIID interactions are significantly expanded given the current state of knowledge.

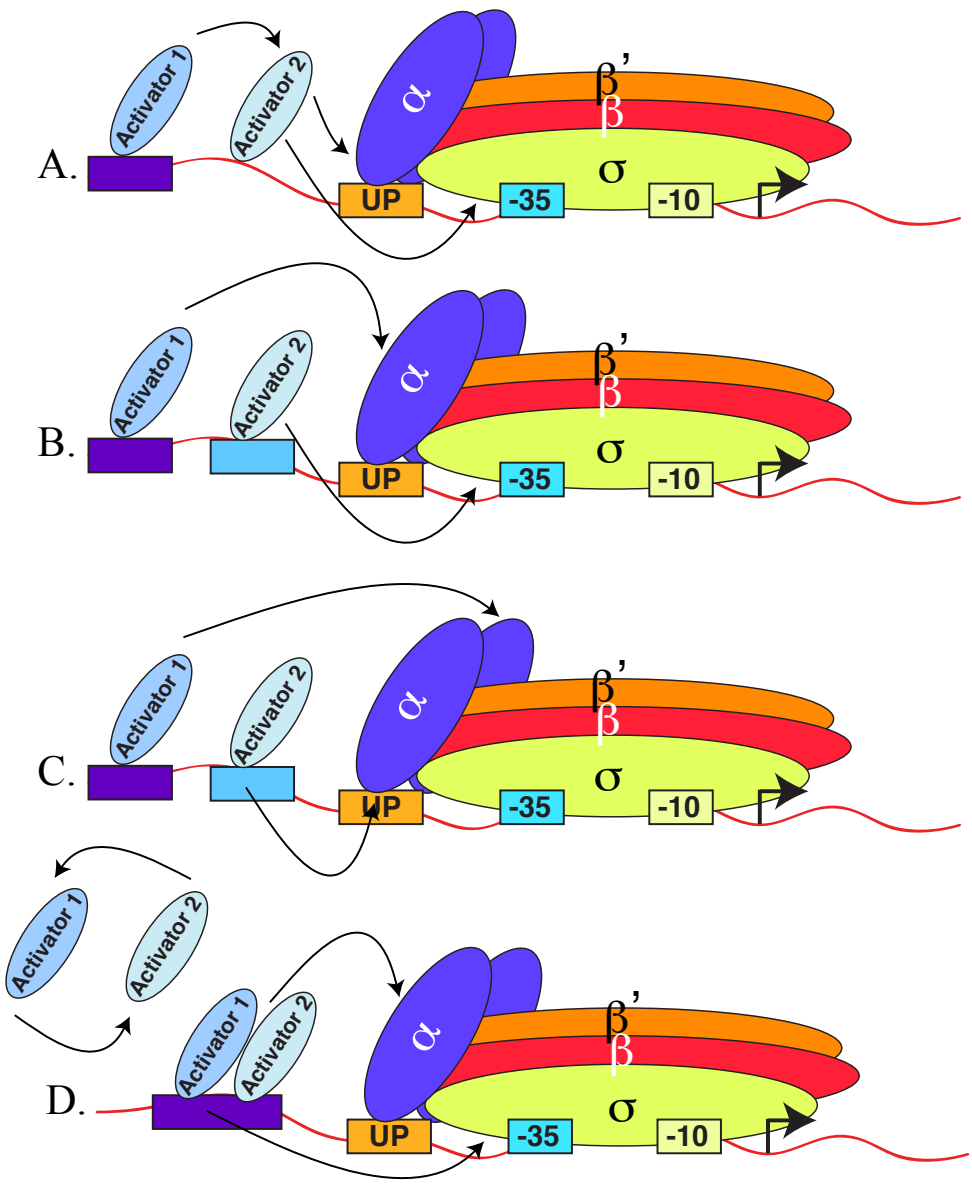


Figure 1.9 Models for independent inputs from two transactivators in bacteria. **A.** One activator influences the position of a second, which is then better positioned to engage Polymerase, for example by either class I or class II mechanisms. **B.** Combined Class I and Class II mechanisms. **C.** Independent Class I contacts. **D.** Cooperative binding of DNA by activators, which are subsequently able to stabilize Polymerase on the promoter by various mechanisms. Many more mechanisms exist than those described in A-D, and combinations of these and other mechanisms help to diversify transcriptional regulation in prokaryotes. Since the variety of both transactivators and combinatorial control is greater in eukaryotes, it is to be expected that many more mechanisms of regulation of RNA Pol II function will be in place, even within the realm of activator-TFIID interaction. For example, p53, Sp1, and c-Jun, which bind distinct but overlapping sets of Tafs, often collaborate to regulate common target genes. Again consider the sampled mechanisms shown here, the possibility of combinatorial regulation of/by eukaryotic activators, and the mechanisms detailed in Figure 1.7 in considering the potential complexity of activator-TFIID interaction, within the context of a target gene in vivo; adapted from Browning and Busby, 2004.

site for the activator is located further upstream from the binding sites for σ in the promoter and activator will engage polymerase by contacting the α subunit(s). Class II is similar except that the activator-binding site will be closer to the promoter, perhaps even overlapping with a σ binding site, and instead of α the activator will directly engage σ. In contrast to classes I and II, class III does not necessarily involve protein-protein interaction, but instead the DNA binding of activator within the promoter serves to change the DNA topology, allowing polymerase to associate more readily; this mechanism illustrates the important point that promoters are not merely simple linear double helices, since even in bacteria, chromosomal conformation can be used to regulate genome function.

As is the case in eukaryotes, it is much more common for multiple bacterial activators to influence polymerase function, and composite 'enhancers' are frequently observed in prokaryotes. To describe increased complexity a bit further, I should mention some of the known mechanisms by which two activators work together to influence promoter/polymerase function (**Figure 1.9**). In 'repositioning' mechanisms, using cooperative protein-protein interactions, an initial activator can help adjust the movement of a second activator from inefficient stimulatory locations on DNA to more efficient ones. A second type of 'repositioning' would involve a change in DNA conformation mediated by one activator, which serves to allow direct contact of the second activator with the polymerase. As such the binding sites on the DNA do not actually change, but the position of the activator-DNA complex does. Activation by two different proteins can strictly involve protein-protein interactions with the polymerase. For example, polymerase could be simultaneously contacted on the α subunit and on the

σ subunit, each interacting with different activators. Or, since the α subunit is present as a homodimer within the enzyme, distinct activators could contact each monomer. These possibilities represent combinations of simple Class I and II activation. Of course amalgamation with Class III mechanisms can diversify these schemes, even with three or more functioning activators.

Multiple activators can function even when only one is directly engaging the polymerase. Protein-protein interactions between activators can increase the binding (i.e. cooperativity) of both to DNA, allowing them to stably associate with the enhancer and recruit the polymerase through a simple Class I or Class II mechanism (see (Griffith et al., 1986; Hochschild and Ptashne, 1986). When one considers that repressors will often co-occupy regulatory sequences along with activators, a single activator could serve only to counteract a repressor, allowing net activation by a second activator through simple mechanisms. These are just a few of the known mechanisms of activator function in bacteria and there are probably simple but undiscovered pathways awaiting discovery. By including schemes with additional activators or repressors, the variety can increase a great deal.

In bacteria where polymerase subunits themselves provide all cellular coregulatory function, there are still multiple proteins that can interact with activators. In addition, multiple domains within those two subunits can function in an activator-dependent fashion. One must reconsider that in eukaryotes there are dozens if not hundreds of coactivators whose contact with activators can be just as necessary for transcription as that of polymerase with activators in prokaryotic cells. Many of those complexes include a dozen or more subunits that can provide interaction surfaces, as is

the case with Mediator (34 subunits) and TFIID (15 subunits). Explicitly, consider again that TFIID has fifteen subunits and this is three times as many proteins as bacterial polymerase, any of which could bridge the complex to activators and most likely a higher effective number when one considers TFIID size and subunit domain architecture and stoichiometry within the complex (compare activator-TFIID interactions in **Figure 1.7** with bacterial activator-polymerase interactions described in **Figures 1.8** and **1.9**). Also one must consider compartmentalization in eukaryotes. There is emerging evidence that the nuclear rim and nuclear pore complexes can have important topological affects on gene location and transcription, above and beyond what DNA conformation can achieve. It is becoming apparent that coregulation by the nuclear pore in space and time can add a whole new level of complexity as opposed to just considering the chromosomal DNA as 'free-floating' within the nucleus (Casolari et al., 2005; Casolari et al., 2004). Precise nuclear locales may contain conformationally constrained 'transcription factories,' whose formation, composition, and location may be guided by activator-coregulator interactions (Ahmed et al.; Ahmed and Brickner, 2007; Brickner et al., 2007; Brickner, 2009; Brickner and Walter, 2004; Luthra et al., 2007; Martin and Pombo, 2003; Menon et al., 2005).

The prokaryotic system is greatly simplified but this in no way suggests the gains in knowledge were easy. In fact, it is the use of facile, high-throughput microbial genetic techniques coupled with powerful biochemistry and structural biology that allowed the elaboration of the models for activator-polymerase interaction presented above. While the use of biochemistry and cell fractionation in systems such as *Drosophila*, *Xenopus*, and mammalian cell culture can rival *E. coli* or other microbes, the genetics of these

metazoan systems are more cumbersome and time-consuming. Therefore it is more challenging to apply lessons learned from exercises in biochemistry and re-test them to observe changes in transcription in vivo. The advent of RNAi and transgenic animals has helped this to a degree but these techniques are neither foolproof nor trivial to implement effectively.

The key to understanding activator-coregulator interactions in any model system is to have identified a candidate gene whose regulation can be easily studied and is dependent upon both genetically defined activators, and also on specific coregulators. It is important to confirm direct activator and coregulator association with the candidate gene regulatory sequences, both in vitro and in vivo.

In vivo investigation of Taf function in human cell lines and flies has been restricted to a very small sampling of mutant alleles (Georgieva et al., 2001; Hisatake et al., 1993; Kaufmann et al., 1996; Metzger et al., 1999; Pham et al., 1999; Sauer et al., 1996; Suzuki-Yagawa et al., 1997; Zhou et al., 1998). Unfortunately, targeted deletion of Tafs in mice repeatedly led to embryonic lethality, restricting the potential of this system for detailed investigation (Guermah et al., 2003; Indra et al., 2005; Metzger et al., 1999; Mohan et al., 2003). Mutations in fly Taf4, Taf6, and Taf9 indicated gene specific transcriptional defects in vivo, with important contributions to development and evidence for direct affects (Pham et al., 1999; Zhou et al., 1998). However, large-scale effects on transcription genome-wide have not been reported. Cell lines containing temperature sensitive mutations in hamster Taf1 or human Taf2 indicated a very important function for the proteins in cell cycle progression from G1 to S or from G2 to M, respectively (Kaufmann et al., 1998; Suzuki-Yagawa et al., 1997). These cell-cycle defects were

attributed to direct participation of Tafs in transcription of cyclin genes, with some indications that an affect on activator-Taf interactions could be involved. Somewhat surprisingly, the Taf1 mutant cell line demonstrated mis-regulation of 18% of expressed genes at the non-permissive temperature, considerably less than what would be expected given the results of various in vitro transcription experiments (O'Brien and Tjian, 2000). Regarding genetic dependencies on Tafs in TFIID, as mentioned previously there were already indications that certain activators could function in a Taf-independent manner in vitro, even when driving transcription from the same core promoter as Taf-dependent activators (Oelgeschlager et al., 1998). Collectively these results were somewhat difficult to expand upon, given the relatively limited number of mutant forms of Tafs under examination.

Yeast has been used most extensively for in vivo investigation of Taf function. The earliest studies indicated that all Taf encoding genes are essential in yeast (Poon et al., 1995; Reese et al., 1994), but this was used as an asset to direct construction of temperature conditional mutants in several different Tafs (Apone et al., 1998; Apone et al., 1996; Shen and Green, 1997; Walker et al., 1996; Walker et al., 1997). These studies, along with others based on different strategies to deplete cellular Taf protein levels, again showed that Tafs are not generally required for mRNA gene transcription (Moqtaderi et al., 1996a; Walker et al., 1996). However, the use of other distinct mutants subsequently showed that the bulk of mRNA gene transcription in fact does depend on Tafs (Holstege et al., 1998; Moqtaderi et al., 1998; Shen et al., 2003). The seeming inconsistencies between the results of experiments utilizing different Taf mutant alleles were difficult to reconcile, especially given the paucity of corresponding biochemical data.

Mechanistically, all that had been made clear was that certain genes require Taf function for transcription and other genes do not. Yeast is the eukaryotic model system with the most technical similarities to bacteria, including facile genetics, so it is ironic that TFIID biochemistry in yeast lagged behind the genetic understanding, especially in comparison to the animal systems where genetic information about TFIID function is greatly exceeded by the biochemistry.

TAF Dependent and Independent Transcription

The studies in yeast that showed broad dependencies on individual Taf function focused on those Tafs (5, 6, 9, 10, and 12) that were subsequently shown to be shared with the SAGA histone acetylase/deubiquitinase complex (Grant et al., 1998), which is also a direct target of activators (Bhaumik et al., 2004; Brown et al., 2001). These facts are inconsistent with genetic analyses of SAGA, which have shown that the genes encoding nearly all non-Taf SAGA subunits are non-essential, and deletion often has minimal impact on overall mRNA gene transcription (Bhaumik and Green, 2002). Therefore it has been speculated that essential Taf function must reside in TFIID, unless there are undiscovered Taf-containing complexes in yeast besides TFIID and SAGA. Interestingly, many genes that exhibit genetic dependencies on one complex do not have dependencies on the other, although the transcriptional deficit in the absence of SAGA is typically less profound (Huisinga and Pugh, 2004; Lee et al., 2000; van Werven et al., 2009). Genes dependent on both SAGA and TFIID do exist though (van Oevelen et al., 2005; Zhang et al., 2008). In any case, the focus on SAGA and other chromatin-directed coregulators and mediator has shifted attention away from the concept of TFIID as a

direct activator target. Nevertheless, current thinking suggests that TFIID-Taf function is required for transcription of the majority of genes, most of which are somewhat arbitrarily described as the 'housekeeping' variety. On the other hand SAGA is generally thought to provide coregulatory function on inducible genes that are transcribed under specific environmental conditions.

The basis of TFIID dependence is defined by exhibition of reduced transcription upon Taf inactivation and also by physical occupancy of target promoters by Tafs in ChIP experiments. Taf–independent promoters will exhibit a much higher occupancy by TBP relative to the occupancy by Tafs as scored by ChIP, while Taf-dependent promoters will have similar quantities of TBP and Tafs (Kuras et al., 2000; Li et al., 2000). These results are thought to indicate the presence of TBP as a component in TFIID (Taf-dependent genes), whereas TBP can exist on promoters as a component of other complexes (on Taf-independent genes). TBP regulation cannot be the sole basis on whether or not a gene is Taf-dependent or independent. Higher organisms also exhibit TBP-independent transcription; in fact a set of so-called TBP-related factors, or TRFs, exist in cell type and developmental specific patterns and are critical for expression of functionally related subsets of genes (Crowley et al., 1993; Hansen et al., 1997; Kopytova and Krasnov, 2007; Rabenstein et al., 1999). TRFs exist in complexes distinct from classical TBP-containing TFIID. The most dramatic example involves mouse TRF3, which is important for differentiation of myoblasts into mature myotubes, or muscle fibers. TRF3 is directly involved in the myotube-specific transcriptional program. Coincident with differentiation, TRF3 becomes associated in a stable complex with Taf3, while surprisingly at the same time cellular levels of most other Tafs drop to near

undetectable levels. The TRF3-Taf3 complex appears to serve as a coactivator for the muscle-specific MyoD activator through direct MyoD-Taf3 interactions (Deato et al., 2008; Deato and Tjian, 2007, 2008; Hu et al., 2008). A similar phenomenon of declining TFIID-Taf levels appears to occur in another differentiation event, the transition from hepatoblast to hepatocyte. Furthermore, Mediator complex levels also decline during this tissue differentiation step (D'Alessio et al., 2009). Collectively, these data have led to the hypothesis that TFIID is generally involved in mRNA gene transcription within cell types containing greater proliferative potential, and then becomes insignificant in differentiated, post-mitotic cells. It is notable that TFIID function in vivo has mostly been studied using transformed cell lines, the *Drosophila* embryo, and logarithmically dividing yeast cultures, contexts where the overall dependency on TFIID would be very high according to this hypothesis. However, the continual presence of one particular Taf, Taf3, is critical for myotube function, still consistent with the central importance of Tafs in transcription within all eukaryotic cells. Therefore, although we have enough examples to place universality of each Taf's function in doubt, their physiological importance remains central to life. From one perspective, the importance of Tafs has only been solidified, since their regulated and timely removal can be as important to cellular identity as their ongoing function.

It seems that the yeast work that did identify examples of Taf-dependent transcription would have indicated the genetic systems useful for an in vivo dissection of activator-TFIID interaction. Unfortunately, many of the best-studied and potentially useful genetic systems, those with characterized activators, cis-elements, and core promoters, fall into the Taf-independent category (Kuras et al., 2000; Li et al., 2000;

Moqtaderi et al., 1996a). Early differential display experiments performed by Michael Green's lab indicated that transcription of G1/S and G2/M cyclin genes were TFIID-Taf dependent (Shen and Green, 1997; Walker et al., 1997). This was not substantiated by ChIP, a technique quite new at that time, thus indirect affects were not excludable. Those genetic systems are also technically challenging to study since synchronous cell populations are needed to observe cyclical transcription and factor-promoter association, a characteristic that slowed subsequent analyses of Taf involvement therein (Spellman et al., 1998). Ribosomal Protein Genes, or RPGs, also showed up in Green's studies (Shen and Green, 1997). Study of microarray data indicated that RPG transcripts were over-represented amongst those mRNAs genetically dependent on TFIID Tafs. About sixty percent of RPGs required Taf1 function for wild-type transcript levels, as compared to twenty percent of all other transcripts (Holstege et al., 1998). Subsequent ChIP studies by the Green and Struhl labs showed that RPGs have approximately equal TFIID-Taf and TBP occupancy, indicating that TBP is present there as a component of the TFIID complex (Kuras et al., 2000; Li et al., 2000). This work established RPGs as a candidate genetic system to study the possible direct roles of TFIID on specific mRNA gene transcription.

Yeast Ribosomal Protein Genes as a Context for Studying TFIID Function

The reason for a strong dependence on TFIID function in RPG transcription was initially considered to revolve around TBP function, because a majority of the 137 RPG promoters lack a 'consensus' TATA box (Mencia et al., 2002). Several lines of evidence pointed to a linkage between TFIID function and absence of a TATA box. The

aforementioned in vitro transcription studies showed that a TATA-less core promoter template requires Tafs for in vitro transcription, even in the absence of activator (Pugh and Tjian, 1991; Smale et al., 1990). This is probably because Tafs possess the ability to directly bind other promoter DNA elements (Chalkley and Verrijzer, 1999; Verrijzer et al., 1994).

This is more difficult to test in yeast, because to date no DPE/MTE/XCPE/INR elements (see **Figure 1.1**) have been defined in this organism. Perhaps such elements exist but there is no available biochemical data such as chemical or enzymatic footprints to suggest how TFIID physically associates with yeast promoters in the absence of TATA. Informatics approaches that make use of genome sequence and microarray or deep-sequencing data have not shed any light on this matter. Nevertheless genetic experiments showed that addition of a 'consensus' TATA box to a gene with a TFIID dependent, TATA-less promoter converted that gene to TFIID independence. Other studies confirmed this correlation. Mechanistically, the reason(s) behind this remained obscure owing to the poor definition of yeast core promoter sequences.

It was perhaps experimentally less challenging to pursue the relationship between activators and TFIID dependence. The basis for a yeast gene's TFIID dependence must in some ways be linked to the activator for several reasons. Regulons activated by well-studied proteins such as Gal4 and Gcn4 are transcribed independent of Taf function (Kuras et al., 2000; Li et al., 2002; Moqtaderi et al., 1996a). Strikingly, these fairly strong activators cannot always function when used in conjunction with a Taf-dependent core promoter (Shen and Green, 1997). These results unfortunately indicate that the benefits of studying these exceptionally well-characterized activators/genes cannot be

extended to TFIID. However, besides the TATA-less nature of the majority of RPG core promoters, the Rap1 activator also directly binds most of those genes (Klein and Struhl, 1994; Lascaris et al., 1999; Lieb et al., 2001). The Struhl lab showed that artificial placement of Rap1 binding sites from an RPG directly upstream of a Taf-independent core promoter resulted in Taf-dependent transcription and increased Taf occupancy on these chimeric genes (Mencia et al., 2002). Struhl expanded on these observations considerably and proposed that Rap1-mediated increases in TFIID occupancy could occur through several mechanisms. Rap1 might directly recruit TFIID to RPG promoters or recruit chromatin-directed coactivators that modify nucleosomes, leading to their removal, thereby allowing passive association of TFIID with the cis-linked promoter. Owing to the partial core promoter independence of this phenomenon, it seemed logical that Rap1 and TFIID function together by way of direct protein-protein interactions with Tafs. The functional link between enhancer and core promoter would occur by interaction between enhancer-bound Rap1 and promoter-bound TFIID. This actually became the first reported case of activator-specified yeast TFIID function in vivo. Description of TFIID function in the yeast Ribosomal protein gene system offered the promise of getting the most information from a combined, thorough genetic and biochemical investigation.

Direct Rap1-TFIID Interaction

Once again, though a more informative series of genetic observations had been made, description of any biochemical mechanism was lacking. The findings of Struhl actually supposed a direct genetic involvement of Rap1 through manipulation of the Rap1

binding sites. Gain of function was achieved using a consensus RPG Rap1 binding site because its presence created a TFIID-dependent promoter; loss of function was observed by point mutations in the Rap1 binding site. The direct role of Rap1 in TFIID interaction was not confirmed using any loss of function mutants directly affecting Rap1 protein function. Similarly, the methodology used to test and confirm TFIID-Taf dependence was not appropriate for indicating the specific reason(s) for TFIID involvement. Taf1 protein was depleted from the cells and effects on transcript levels were subsequently measured (Mencia et al., 2002). Elimination of Taf1 protein most likely causes a destabilization of the entire complex, although Struhl's study did not address the affect of Taf1 depletion on levels of any cellular proteins, including Taf1.

To get at the mechanism behind these genetic interactions, the next step was to identify if physical interaction occurs between the proteins, and to define the Rap1 surfaces and the Taf proteins and domains involved. Such information would in turn allow an additional genetic characterization of the physical interaction. Demonstration of physiological relevance via directed genetic studies would warrant additional biochemical characterization. The situation could be examined much in the way a similar problem would be pursued with *E. coli* activator-polymerase interactions.

I should make it explicitly clear that there are other proteins besides Rap1, TFIID, and the other basal transcription factors that contribute to RPG transcription. First, Rap1 doesn't bind to all 137 RPG enhancers. Instead Abf1, or in a few cases Reb1, appear to carry out functions similar to Rap1 (Diffley and Stillman, 1989; Ju et al., 1990; Yarragudi et al., 2004; Yarragudi et al., 2007). Second, other transactivators contribute broadly to RPG transcription including the Fhl1/Ifh1 dimer, Sfp1, and Hmo1 (Hall et al., 2006;

Lempiainen et al., 2009; Marion et al., 2004; Martin et al., 2004; Rudra et al., 2005; Schawalder et al., 2004; Wade et al., 2004). RPG transcription is very tightly and coordinately controlled according to growth rate, which is influenced by environmental conditions such as nutrient quality/availability (Rudra and Warner, 2004). RPG transcription can approach 50% of total Pol II output during rapid growth or almost completely cease when cultures approach stationary phase or conditions of starvation (Warner, 1999). Rap1 occupancy appears fairly consistent regardless of transcription rate, while the other transfactors appear to dissociate from the enhancer during low transcription in response to signal transduction pathways (Berger et al., 2007; Lempiainen et al., 2009; Marion et al., 2004; Martin et al., 2004; Rudra et al., 2005; Schawalder et al., 2004; Wade et al., 2004). This has been taken to suggest that these multiple alternative transcription factors, and not Rap1, are the principal drivers of RPG transcription activation. Another piece of circumstantial evidence supporting this hypothesis is that Rap1 itself is considered a weak activator, since when fused to a heterologous DNA-binding domain such as LexA or Gal4 DBD, it does not behave like Gcn4 or Gal4 activation domains, for example (Zhao et al., 2006). The continued presence of Rap1, its apparent weak activation potential, and the eviction of the other transfactors concomitant with decreased transcription could be taken to indicate their heightened importance relative to Rap1. Unfortunately, there is little information about the amount of stimulation of RPG transcription contributed by these proteins. Moreover, there exists the possibility that Rap1 functionality is not independent of RPG transcription rate; for example Rap1 could be altered by PTMs in such a way that as yet unidentified properties, but not DNA-binding, are stimulated or inhibited in response to

changing environmental conditions. Indeed Rap1 is known to be phosphorylated (Klein and Struhl, 1994; Tsang et al., 1990). It could be that Rap1 directly influences the properties of the other transfactors, for example by stabilizing their binding, but to date there is no evidence of direct protein-protein interaction between Rap1 and Ifh1/Fhl1, Sfp1, or Hmo1 (Rudra et al., 2007). There is also little compelling evidence to suggest functional interactions between the alternate RPG transfactors and TFIID, in contrast to the connection between Rap1 and TFIID. However, there are other coregulators present and apparently functional on RPG enhancers and promoters, and their function could be tied to that of Fhl1/Ifh1, Sfp1, and Hmo1. An Rpd3 deacetylase-containing complex can repress RPG transcription rates as part of the response to environmental stress (Humphrey et al., 2004). In contrast, the Esa1 acetylase-containing NuA4 complex positively regulates transcription, is required for full Taf and TBP occupancy, and also dissociates from the enhancer during conditions of reduced transcription (Mencia et al., 2002; Reid et al., 2000; Rohde and Cardenas, 2003). Genome-wide ChIP on chip studies indicated that proteasomal components also occupy RPG enhancers, consistent with involvement of the proteasome complex in transcription (Auld et al., 2006). The proteasome could conceivably mediate transfactor and coregulator turnover on RPGs, perhaps with an accelerated turnover rate contributing to the decreased occupancy of the factors mentioned. Inhibition of the TOR signaling pathway is commonly associated with downregulation of RPG transcription and loss of Ifh1/Fhl1, Sfp1, Hmo1, and NuA4 (Powers and Walter, 1999). A summary of the inputs into RPG transcription from activators and coregulators is shown in **Figure 1.10**.

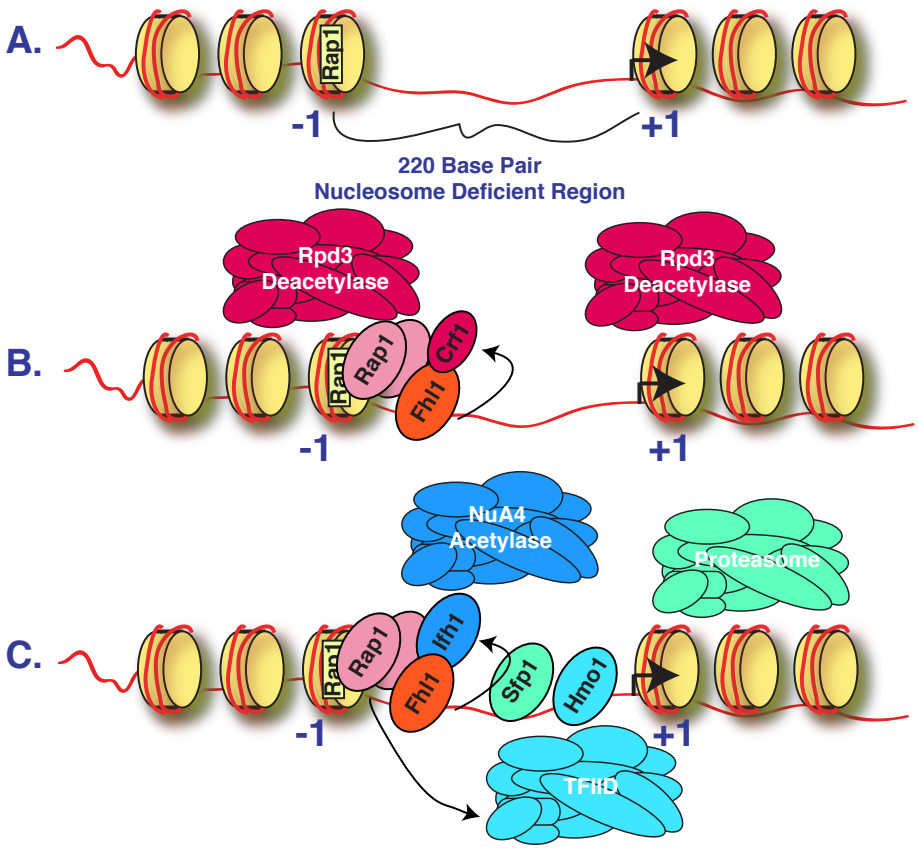


Figure 1.10: Structure of Yeast Ribosomal Protein Genes and factors regulating their transcription . **A.** Notable characteristics of Ribosomal Protein Gene DNA include an absence of obvious core promoter elements. Instead, a well positioned nucleosome overlaps slightly with DNA containing the transcription start site, and an ~220 base pair nucleosome-deficient region extends 5' of the TSS, where a dyad Rap1 binding site is contained within another well-positioned nucleosome. Importantly, Rap1 efficiently binds to this nucleosomal Rap1 enhancer element (Koerber et. al., 2009). **B.** Factor association during conditions of nutrient deficiency and low level or absent RPG transcription. Two molecules of Rap1 are constitutively bound to its nucleosomal enhancer element found in many RPGs, and Fhl1 is also commonly found in or around the nucleosome free region, although its exact mechanism of association is unknown. The Crf1 repressor is engaged with Fhl1. The Rpd3 histone deacetylase complex is selectively present under these conditions although the substrate is unknown. **C.** Conditions of high level RPG transcription, where the 137 genes account for ~50 % of Pol II initiation events genome-wide. In addition to Rap1 and Fhl1, the Ifh1 protein has replaced Crf1 for binding to Fhl1, and Sfp1 and Hmo1 associate in or around the nucleosome-free region. Like Fhl1, their exact mechanism of binding remains unclear. The NuA4 Histone H4 acetylase and the TFIID complex are physically associated and both are genetically required for high-level RPG transcription. There also evidence for physical association of the Proteasome complex or subcomplexes and it is genetically required for transcription as well. Note that the only known protein-protein interactions are between Fhl1 and its binding partners, and Rap1 and the TFIID complex. Indeed, Rap1-TFIID interaction is the only known direct activator-TFIID interaction documented in budding yeast. Note that the potentially flexible configuration of the nucleosome-free region could be important for facilitating the interaction between upstream bound Rap1 and TFIID localized around the transcription start site.

The overall relationship between Rap1, the other activators, Rpd3, NuA4, the proteasome, and TFIID is not at all clear since the exact role that each of these factors play in RPG transcription is unknown; the connection between Rap1 and TFIID remains the best described. However, because there are clearly other important transfactors and coregulators, it is unlikely that overall stimulation of transcription occurs in a way strictly analogous to simple bacterial class I, II, or class III activation mechanisms, though with respect to the contribution of Rap1 and TFIID, these mechanisms may still be relevant. Our lab began testing this possibility by looking at genetic dependencies on specific Rap1 domains for TFIID-driven RPG transcription. We also began to characterize the possibility and mechanism of physical interaction between Rap1 and TFIID (Garbett et al., 2007).

Importantly, in our studies the possibility of contributions from factors other than Rap1 and TFIID was minimized through the use of an artificial transgene composed of a minimal RPG enhancer fused to a TFIID-independent core promoter (Garbett et al., 2007; Mencia et al., 2002). Although the mechanism of Fhl1/Ifh1/Sfp1/Hmo1 association with RPGs is unclear, if they indeed bind via their own cis-elements then their contribution to transcription of our chimeric reporter gene is theoretically zero. Furthermore, eliminating the contribution of the other activators might also minimize contribution from the NuA4 complex and the Proteasome. Our lab confirmed Struhl's observation that the minimal 41 base pair Rap1 enhancer from *RPS8A,* which contains just two tandem 17 base pair Rap1 binding sites*,* is sufficient to confer TFIID occupancy on the otherwise TFIID-independent *PGK1* core promoter. Moreover, Rap1 itself physically occupied this gene, while specific Rap1 domains within the C-terminus were needed for full transcription of

the chimeric gene and authentic RPGs. Using immunodepletion and partial reconstitution of in vitro transcription, we confirmed the importance of the Rap1 protein and the TFIID complex in transcription of the chimeric reporter gene. Simultaneous depletion of Rap1 and TFIID from the transcription-competent extract abolished activity. Strikingly, addition of purifed Rap1 and TBP back to the extract did not rescue transcription. However, in the TFIID- and Rap1-depleted extracts, addition of Rap1 and TFIID restored activity to full levels. This experiment essentially recapitulated the very first results that indicated coactivator function of Tafs, but shown here in the context of yeast RPG transcription. The most important conclusion of this experiment is its indication that Rap1 uses Tafs to coactivate Rap1-driven RPG transcription, since add back of TFIID but not TBP restored transcription in the depleted extract. Rap1 interacted with the TFIID complex in binding assays. This result indicated direct interaction between Rap1 and the TFIID complex. At the time, this was the first report of an activator interaction with the isolated TFIID holocomplex, since all prior studies had instead measured activator interaction with isolated Tafs or subcomplexes. Consistent with the importance of Rap1 C-terminal domains for transcription noted above, the Rap1 C-terminus was also important for in vitro binding to TFIID. Despite Struhl's prediction that Rap1 invoked recruitment of TFIID to RPG promoters, in the absence of the Rap1 C-terminus TFIID-Taf occupancy on the chimeric Rap1-driven gene was not markedly decreased in vivo as scored by ChIP. But, consistent with reduced in vivo transcription in the absence of the Rap1 C-terminus, Pol II occupancy was low. These results suggest that Rap1 may affect TFIID function subsequent to its association with RPG promoters, perhaps by modulating the ability of TFIID to stimulate pre-initiation complex formation or function. This result

therefore indicates a heretofore-unrecognized activator-dependent mechanism of TFIID function, because transcription is reduced while TFIID occupancy remains consistent in the absence of Rap1 function, arguing against Rap1-mediated TFIID recruitment to RPG promoters. Finally, we confirmed that Rap1 contacts three specific Tafs in vitro; Taf4, Taf5, and Taf12. As such, genetic analyses of these Tafs promised to shed more light on the mechanism of Rap1-TFIID interaction and TFIID function in RPG transcription. I characterized interaction of Rap1 with Taf12 in considerable detail (Garbett et al., 2007).

Initial Study of Rap1-TFIID Genetic and Physical Interaction

The initial work of Struhl provided a nice series of observations using purely genetic methodology, which as a sole strategy is always at risk of over-interpretation due to potentially misleading indirect affects. It would have been difficult to extend the initial characterization of Rap1-TFIID genetic interactions without some supporting biochemical data, which would in turn allow additional iterative rounds of enlightening genetic and biochemical analyses. This held promise to us because of our interest in TFIID, and our ability to use the yeast system for both genetics and biochemistry. Other researchers were interested as well, recognizing these studies as the first detailed characterization of activator-TFIID interaction in yeast. Fortunately we were largely able to recapitulate the major observations of Struhl's work, confirm that there is a direct biochemical basis behind his genetic observations, use genetics along with biochemistry to more precisely define the Rap1 domain(s) mediating interaction with TFIID, provide clues about the potential contribution of three specific TFIID-Taf subunit's involvement,

and allow some speculation about the mechanisms of Rap1-TFIID binding on RPGs and the initiation steps affected by the interaction in vivo (Garbett et al., 2007).

If there was a weakness of our work prior to initiation of the experiments described in this dissertation, it was the failure to provide genetic evidence about the specific Taf functionalities mediating the interaction. While removal of specific Rap1 protein regions had affects on in vitro interaction with TFIID, affects on transcription in vivo, and defective post-TFIID binding PIC formation/function in vivo, there was no equivalent information about individual Taf involvement other than the in vitro interactions. I did perform a detailed characterization of Taf12 and identified specific N-terminal residues required for in vitro interaction with Rap1. However, this was somewhat ambiguous given existing genetic information about Taf12, because sequence encoding this protein region is neither required for viability nor is its removal associated with any loss of growth phenotype (Moqtaderi et al., 1996b). If the Taf12 domain were truly required for a Rap1-TFIID interaction that drives high level RPG transcription, which is in turn essential for rapid cellular growth, we would expect that abolishing Rap1-Taf12 interaction would have an immediate and obvious growth phenotype. Clearly, that wasn't the case so we had to consider some additional possibilities. First, the in vitro mapping of interaction domains might not reflect the true nature of Rap1-Taf12 interaction. Second, a specific domain or domain(s) in Taf12 might be redundant with parts of other Tafs. Taf4 and Taf5 were good candidates since these TFIID subunits were identified as Rap1 binding partners at the same time as Taf12. Because the possibility of redundancy was easier to test, it was chosen as a means to address our

deficient understanding of genetic and biochemical interactions between Tafs and Rap1 in Rap1-TFIID mediated RPG transcription.

A series of assumption and predictions had to be considered before initiating studies characterizing putative Rap1 Binding Domains, or RBDs, in Taf4 and Taf5. First, we expected that one or more unique domains in each protein allow the interaction with Rap1. Proteins are modular with multiple domains, and individual domains often provide distinct components of overall protein function. As an example, within Taf5 there are the C-terminal WD repeats which are known to mediate direct interaction with the Taf6/Taf9 heterodimer and also possibly contribute to methyl-lysine recognition (Dynlacht et al., 1993; Fitzgerald et al., 2006). At the Taf5 N-terminus there are the so-called N-terminal domains, or NTDs (1 and 2), that are speculated to be involved in Taf5 dimerization (Bhattacharya et al., 2007; Romier et al., 2007). Thus we have to think of potential Rap1 interaction surfaces in several ways; certain Taf domains are clearly involved in TFIID integrity via Taf-Taf interactions, such as Histone fold domains, and these are important for TFIID formation and stability, while other Taf domains are responsible for allowing interactions with proteins that are not integral TFIID components, such as Rap1 (Lawit et al., 2007; Yatherajam et al., 2003). Individual domains could also function for TFIID DNA binding or enzymatic activity (Chalkley and Verrijzer, 1999; Mizzen et al., 1996). Of course there could be the complication of shared overall functionality between multiple distinct domains.

My objective was to identify the smallest portions of each Taf responsible for in vitro interaction with Rap1. The logic was that this would then allow me to target a very specific locale of the overall coding sequence for targeted mutagenesis, in an attempt to

test the physiological significance in vivo. Unfortunately at the time my project was initiated there was very little structural data that could be used to guide site-directed mutagenesis studies. As such I had to begin with a semi-random strategy to analyze structure and function of Taf4 and Taf5, with respect to the interaction with Rap1. One speculation coming from my studies on Taf12, our previous studies on Taf4, and structural information about the Taf4/12 heterodimer was that since the Taf12 histone fold domain did not appear to engage Rap1 but did engage Taf4, then perhaps the Taf4 domains responsible for engaging Taf12 would not be involved with Rap1 interaction (Garbett et al., 2007; Thuault et al., 2002). But again except for Taf4/12 histone fold domains, there was no structural information available (Werten et al., 2002). However, Taf sequences are fairly well conserved in evolution at the amino acid-level, and because conservation of sequence often indicates conservation of function, I could use deletion and point mutation of conserved coding residues to guide my analyses of Taf domain interaction with Rap1. By focusing on a minimal contiguous section of amino acid sequence, I hoped to only affect Taf function pertinent to interaction with Rap1 and not other processes such as TFIID integrity.

A second prediction relates to the importance of RPG transcription to cellular growth capacity. To grow and divide at their maximum rate yeast cells must synthesize 2000 ribosomes per minute. RPG transcription rates are limiting for ribosome abundance in yeasts, even though one of every three mRNAs is from an RPG. RPG mRNA stabilities are also considerably lower than the average non-RPG transcript. Thus, meeting demand for ribosomes requires an enormous amount of transcription of the 137 genes that encode the 79 ribosomal proteins (Gorenstein and Warner, 1976; Warner,

1999). Disruption of Rap1-TFIID interaction through targeted mutagenesis of Rap1 binding domain-coding sequence in Tafs should have an obvious negative effect on cellular growth. If a particular Taf RBD-targeted mutation does cause a strong negative growth phenotype, then the degree of growth defect should be commensurate with a reduction in RPG transcription relative to wild type. Moreover since Rap1 and TFIID directly bind most RPGs, and all 137 genes are coordinately regulated, I expected that mutations affecting Taf RBD function should reduce transcription of the whole gene family.

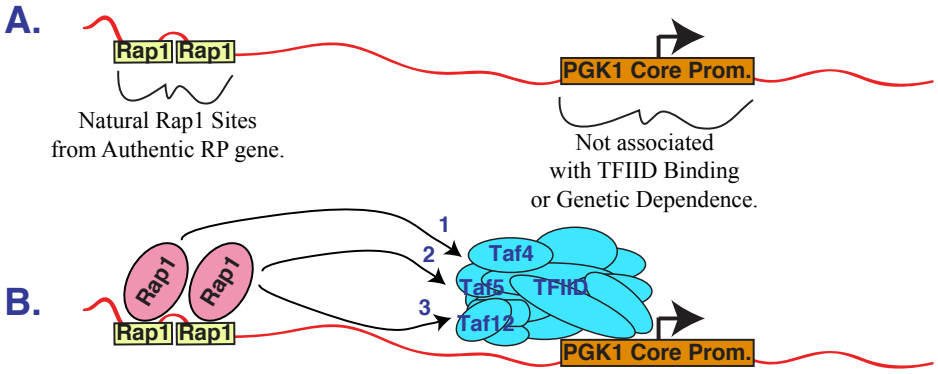
I planned a back and forth application of genetic and biochemical approaches. Our hypothesis predicted that mutations putatively affecting Taf RBD function should be associated biochemically with proteins possessing altered affinity of binding to Rap1, in vitro and in vivo. Test of Rap1-Taf binding in vitro would utilize conventional protein-protein interaction assays with purified recombinant Rap1 and Tafs; Rap1-TFIID binding would use TFIID purified from Taf-RBD mutant yeast strains. Affect on Rap1-TFIID binding might also be measured using immunoprecipitation and western blotting. Importantly, I planned to test for off-target affects on TFIID stability, potentially introduced by targeted RBD mutation, by examining Taf-Taf interaction and/or TFIID complex integrity within mutant strains.

I should explain why a parallel dissection of both Taf4 and Taf5 was planned. This is due to the possibility of redundancy within TFIID for Rap1 binding, maybe due to a mechanism involving fractional contribution(s) from Taf4, Taf5, and Taf12. Given the absence of phenotype in Taf12 mutants, it seemed that simultaneous alterations in each Taf-encoding gene might be needed to manifest a growth and transcriptional defect. To

test this redundancy hypothesis, I planned to create yeast strains bearing mutations in multiple Taf-encoding genes and look for intensified reductions in growth rate, RPG transcription, and Rap1-TFIID interaction.

The possible roles of three different Tafs raised another hypothesis, that Rap1 uses multiple independent domains as surfaces for interaction with unique Taf domain surfaces. Although our previous genetic and biochemical analyses of Rap1-TFIID/Taf interactions indicated that both the DNA binding domain and C-terminus of Rap1 are involved, this work did not permit identification of a specific C-terminal domain. Actually, there are no less than 4 independent domains in the Rap1 C-terminus, only one of which has been purported to contain transactivation potential (Freeman et al., 1995; Hardy et al., 1992a). Moreover, the most C terminal domain has a known role in chromosomal telomere length homeostasis through transcription-independent functions (Feeser and Wolberger, 2008; Lustig et al., 1990). Telomere homeostasis is believed to have important relationships with cellular lifespan and growth properties (Bitterman et al., 2003). The involvement of this so-called silencing domain in multiple cellular processes might serve as a relay between them. In any case I wanted to describe the involvement of Rap1 domains in TFIID interaction in greater detail than we had previously done.

To summarize, the objectives of my project were to map segments of amino acid sequence in Taf4 and Taf5 that are bound by Rap1 in vitro. I then wanted to subject the DNA sequence encoding these domains to truncation and deletion mutagenesis to determine if they are important to Taf function in cellular growth. I next intended to attempt identification of amino acid substitutions within the RBDs associated with



C. Which Taf4, 5, and 12 domains are involved?

Do those Taf domains contribute to cell growth?

Do amino acid substitutions in those Taf domains reduce cell growth?

Do amino acid substitutions in those Taf domains reduce RPG transcription?

What is the relationship between Taf domains and Rap1 domains?

Figure 1.11: Rap1-TFIID interaction in target gene regulation. **A**. To study the possibility of direct interaction between Rap1 and TFIID, as originally suggested by the widespread usage of both factors on *RP* genes, the minimal Rap1 binding sites of the *RPS8A* gene were fused to a promoter segment of the *PGK1* gene. This chimeric gene was introduced into yeast and scored for TFIID binding and genetic dependence on *TAF1*. TFIID binding was strictly dependent upon integrity of the Rap1 binding sites, as seen by Struhl and colleagues previously (Mencia et. al., 2002). Note that genetic dependence on Rap1 was independently verified as removal of the Rap1 C-terminus caused defective TFIID association with the reporter gene and defective transcription; authentic *RPG*s were also affected (Garbett et. al., 2007). Importantly, the use of the minimal reporter gene in all likelihood minimized, or more likely abolished, the contribution of other *RPG* transfactors, since the only *RPG*-related DNA sequence were the Rap1 binding sites, and use of the normally TFIID-independent *PGK1* promoter minimized any potential relationship between TFIID and core promoter sequence motifs. Furthermore, use of a naked DNA template in vitro transcription system recapitulated Rap1 and TFIID dependencies seen in vivo, suggesting that chromatin and chromatin-directed factors have a negligible impact on Rap1-TFIID interaction. **B.** Basic model for Rap1-TFIID interaction. Rap1 bound to the enhancer directly contacts the TFIID complex. This probably involves multiple domains in Rap1 since separate direct contacts are made with in the so-called Rap1 DBD, but also with the C-terminus, which contains no less than four separate domains. The Taf4, 5, and 12 subunits were implicated but the relevance of their contribution to interaction remained largely untested. Based upon gene architecture, it seems likely that the mechanism of Rap1-TFIID interaction involves significant rearrangement of DNA sequences in between the Rap1 binding sites and the TFIID-bound promoter. **C.** Questions to be addressed by this dissertation project regarding the Rap1-TFIID interaction.

conditional loss of growth phenotypes in yeast. Carefully selected conditional mutant strains would be examined for levels of RPG transcripts to test if loss of growth corresponded to defective RPG transcription. Transcript analyses would be extended to whole-genome analysis by microarray, with the objective of testing if introduced mutations in Tafs affected all the RP genes, as predicted given the RPGs widespread association with Rap1 and TFIID. Given that Taf4, Taf5, and Taf12 may provide overlapping and/or redundant Rap1 binding domains, I wanted to test whether synthetic phenotypes occur when *taf4* and *taf5* mutants are combined. Next I wanted to test if RBD-substituted Tafs have a reduced ability to interact with Rap1 in vitro. To look at the status of Rap1-TFIID interaction, I wanted to isolate TFIID complexes from wild type and mutant strains and test their capacity for binding recombinant Rap1. I thought that creation of double *taf4 taf5* mutants would be more useful than single mutants for looking at Rap1 interaction with the TFIID complex, considering the possibility of functional overlap. Finally, I wanted to carry out a more precise analysis of individual Rap1 domain contribution to interaction with Tafs. These objectives are also discussed in **Figure 1.11**.

Hypothesis and Specific Aims

My research was designed to test the hypothesis that the interaction of Rap1 with TFIID is mediated through binding of Rap1 to distinct and unique domains (Rap1 Binding Domains, RBDs) present in the Taf4, Taf5, and Taf12 subunits of TFIID. My studies are predicated on the fact that transcription factor Rap1 is the key activator driving transcription of the RP genes, and that the growth rate of yeast cells is limited by RP gene transcription and subsequent ribosome production. Consequently I predicted

that removal of or amino acid substitution(s) within the RBDs would decrease RP gene transcription concomitant with negative effects upon yeast cell growth.

To test this hypothesis I developed the following **Specific Aims.**

1. Map the RBDs of Taf4, 5, and 12 using deletion mutagenesis and in vitro Rap1-Taf binding studies.

2. Determine the effects of deletion and targeted mutations within the *TAF4* and *TAF5* RBD-encoding regions upon yeast cell growth and RP gene transcription.

3. Determine whether mutation of Taf4, 5, and 12 RBDs affected their affinity of binding for Rap1.

4. Test whether mutations in *taf4* and *taf5* alleles exhibit synthetic genetic interactions with each other and mutant variants of *rap1*.

CHAPTER II

STRUCTURAL AND FUNCTIONAL ANALYSES OF RAP1-TAF INTERACTION

Considerations in Dissecting Rap1-Taf Interactions

In planning a strategy involving biochemical and genetic investigation of the molecular structure-function relationships of Tafs (4, 5, and 12) and Rap1, the single most important consideration is that all four proteins are encoded by single copy genes in the haploid genome which are essential for viability. This fact indicates that one or more domains in each protein is/are responsible for that vital characteristic. In the case of Tafs the mechanisms behind domain requirement can be simplified into two categories: protein domain functions intrinsic to TFIID or protein domain functions extrinsic to TFIID. The former can be subdivided into domain homo-oligomerization or domain interaction with other Taf(s) within TFIID. Extrinsic functions might include intrinsic enzymatic activities, DNA-binding, or interaction with proteins that are not TFIID subunits. Rap1 is a candidate for this last possibility therefore my idea was, if at all possible, to identify domains within each Taf that interact with Rap1 but do not engage in Taf-Taf interaction or possess enzymatic or DNA-binding activity. I considered aspects of protein structure to the extent that I could, given the lack of high-resolution structural data mentioned in Chapter 1. A potential confounding issue in my analyses was the fact that protein domains can sometimes be composed of residues that are non-contiguous in protein sequence. Taf4 is a case in point, since the histone fold domain (HFD) that mediates heterodimerization with Taf12 is bisected by a large stretch of amino acids that

are dispensable for interaction with Taf12 (Thuault et al., 2002). Hence the Taf4 HFD is bipartite, and composed of an amino-terminal portion and the so-called conserved C-terminal domain, or CCTD. With this caveat in mind I chose to begin mapping studies by using as unbiased an approach as possible, by first implementing ordered truncations to define the amino and carboxy termini of any Rap1-binding domain (RBD)-containing protein fragments in Taf4, Taf5, and Taf12. It follows that if Rap1 interaction with that particular protein fragment is truly important, then both the serial truncation and the selective deletion of that domain should compromise cellular growth. To supplement and extend the deletion mutagenesis studies I planned to extend manipulation of putative Taf domains to the level of single amino acid residues, in an attempt to reduce the chance of gross disruption of overall protein function, while simultaneously generating a detailed structure-function map of the respective domains. But once again, I needed to keep my objective in accordance with the level of resolution of Taf domains that I expected to obtain in a reasonable amount of time. Site-directed mutagenesis of individual codons might be overly optimistic in pursuit of associated mutant phenotypes in vivo, especially given the lack of structural information about Taf domains. A semi-randomized strategy of amino acid substitution had proven useful in the lab's previous experiences and could be expected to allow reliable, predictable and accurate testing of my hypothesis that the Taf RBDs are important for cell growth and RPG transcription. The implementation of yeast genetics also promised to speed things up. Again considering the predicted essential nature of Taf4, 5, and 12 RBDs, my best-case scenario was isolation of conditionally lethal mutants with amino acid substitutions confined to the area deemed responsible for Rap1 interaction. In this chapter I will first describe the methodology

behind my mutational studies and then discuss the results of the actual experiments where I mapped and characterized the Rap1 binding domains of Taf4, Taf5, and Taf12.

METHODS

Bacterial and Yeast Strains and DNA Manipulations

A large number of plasmid constructs were generated during the course of dissertation research; I estimate that I made in excess of two thousand unique constructs. In addition many recombinant proteins were prepared; more than three hundred were expressed and purified. Therefore every effort was made to generate these reagents as economically as possible, both in terms of monetary and time investment.

Standard *E. coli* manipulations for construct generation were performed using strain BW23474 (Liu et al., 1998; Liu et al., 2000). This strain has two advantages. First, it is compatible with replication origins found in the pUNI series of plasmids, which are created using conventional restriction enzyme-mediated subcloning and subsequently utilized for recombination-based subcloning via LoxP sites and bacteriophage P22 Cre Recombinase. The second advantage is that BW23474 can easily be made chemically competent to very high transformation efficiencies ($\sim 10^8$ - $\sim 10^9$ colony forming units per microgram DNA), which is useful for difficult or challenging subcloning experiments but also for library construction that mandates high complexity and representation. All bacteria used in this study were made chemically competent using room temperature cultures, growing logarithmically in S.O.C. media (following exact protocol in (Inoue et al., 1990). The Univector system was utilized to create yeast

expression vectors in certain instances; reactions between pUNI and pHOST vectors included Cre Recombinase, expressed and purified in-house. This methodology is superior to other recombination-based cloning strategies such as the Gateway system, because it relies only upon easily-produced Cre Recombinase instead of multiple proprietary (and costly) enzymatic components. Univector recombination reactions were transformed into *E. coli* strain DH5 (*F'* δ80Δ*lacZ* ΔM15 Δ(*lacZYA* ⁻*argF*)U169 *deoR recA1 endA1 hsdR17*(ρK° mK⁺)*phoA supE44 lambda*⁻ *thi*⁻*1*). Irrespective of whether traditional subcloning or recombinase-based cloning was used, the desired constructs were usually recovered from subcloning reactions at high efficiences (typically exceeding 80% correct, insert-bearing clones). While fully expected when using the Univector system, the success in the traditional subcloning experiments can be attributed to strict use of sequential restriction enzyme digestions, dephosphorylation of vector recipient ends, and gel purification of both vector and insert preparations.

Two *E. coli* strains were used for protein expression. Both are based upon BL21 Gold DE3 ((*B F*⁻ *ompT hsdS*(ρB⁻ μB⁻) *dcm*⁺ *Tet gal* λ(DE3 [*lacI lacUV5*⁻T7 gene 1 *ind1 sam7 nin5*]) *endA Hte*)). The 8 tRNA encoding plasmid pRARE was isolated from Rosetta DE3 (Novagen, *ompT hsdSB*(ρB⁻ μB⁻) *gal dcm* λ(DE3 [*lacI lacUV5*⁻T7 gene 1 *ind1 sam7 nin5*]) pRARE (CamR). This low-copy pACYC184-based plasmid encodes tRNAs corresponding to rare codons for arginine, isoleucine, glycine, leucine, proline, methionine, threonine, and tyrosine. pRARE was used to transform BL21 Gold DE3 to chloramphenicol resistance and the resulting strain, BL21 Gold DE3 pRARE, was made competent as above and used for all subsequent protein expression experiments. To make the second expression strain, the RIL plasmid from Arctic Express DE3 RIL

(Stratagene, *B F⁻ ompT hsdS* (ρB^- μB^-) *dcm*$^+$ *Tet gal* λ(DE3 [*lacI lacUV5*-T7 gene 1 *ind1 sam7 nin5*]) *endA Hte* [*cpn10 cpn60 Gent*] [*argU ileY leuW Str*] was isolated and used to transform BL21 Gold DE3 to streptomycin resistance. The resulting strain, BL21 Gold DE3 RIL Strept, including four tRNAs for arginine, isoleucine, and leucine on a pSC101-based plasmid, was made competent as above and used for all experiments involving coexpression of multiple proteins. This strain is compatible with other plasmids containing pBR322 and/or p15A replication origins, such as the popular T7 expression vectors based on pET or pACYC184 backbones. These two strains share several indispensable attributes: strains grow more robustly than other rare codon-containing strains such as Rosetta DE3, simplifying expression experiments; these strains exhibit high transformation efficiency when made chemically competent as described above allowing ligation or ligation-independent subcloning reactions to be directly transformed into the expression strain; finally as codon bias is probably the most common obstruction in successful bacterial expression of eukaryotic proteins, both *E. coli* strains largely overcome this problem due to their overexpression of multiple rare tRNAs.

Several yeast strains were prepared. Because strain-specific phenomena have been reported in the regulation of Ribosomal Protein Gene transcription (Zhao et al., 2006), and because my pilot experiments used a mixed assortment of different strain backgrounds, I decided to base my experiments on several common laboratory background strains. Initial experiments used SEY6211 (*MATa leu2-3,112 ura3-52 his3-Δ200 trp1-Δ901 ade2-101 suc2-Δ9; GAL*) because a previous lab strain contained a chromosomal *taf5* deletion in this background. An advantage of this strain is that it exhibits a high sprorulation efficiency that can be put to use when advanced genetic

experiments involving strain crosses are needed. SEY6211 is also good for experiments using cellular protein extracts because it does not appear to contain the proteolytic activity associated with some strains. BY4741 (*MATa his3Δ1 leu2Δ0 LYS2 met15Δ0 ura3Δ0*) was used because existing lab strains with either chromosomal *taf4* or *taf12* deletions were on this background. BY4741 was used for the genome sequencing project and several systematic strain collections are designed on this background. BY4741 is Trp^+, which reduces the nutritional markers available for genome manipulation, but Trp^+ strains are not cold sensitive or Gal^-, characteristics that can be useful for certain experiments. W303a (*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) is my strain of choice for genome manipulations, because it exhibits high transformation efficiency in conjunction with the Lithium Acetate/Polyethylene Glycol-based protocol I use for all yeast transformations (Gietz and Schiestl, 2007). W303a has the broadest range of nutritional markers of the three strains further enhancing its usefulness when planning complex plasmid-based genetic experiments. Unfortunately, unlike the other strains there is no difference in nutritional markers between W303 mating types unlike SEY6211 and BY4741, so genetic experiments involving mating, sporulation, and tetrad dissection are less convenient and more time consuming in W303a. Other disadvantages are that W303a is Gal^- and cold sensitive owing to the *trp1-1* allele, and appears less suitable than SEY6211 or BY4741 for experiments involving preparation of cellular extracts (JHL, personal observation).

"Shuffling" strains were created in each of these strain backgrounds to perform genetic experiments with *TAF4, TAF5, TAF12*, and *RAP1*. There are two modifications made to the parental genetic background to create a shuffling strain. Firstly, a "covering"

plasmid (typically a WT version of the gene of interest) is needed that contains a nutritional gene that can be selected for by omitting a nutrient such as a particular amino acid or nucleotide. However this nutritional marker should also be counterselectable such that addition of a specific compound selects against growth of cells that contain the covering plasmid. In this case the selectable/counterselectable marker was the *URA3* gene, which was selected for by use of media lacking uracil, and selected against by inclusion of the nucleotide analog 5-fluoro-orotic acid (FOA); FOA is toxic when metabolized to pyrimidine (Sikorski and Boeke, 1991). The *URA3*-marked covering plasmid contained the protein-coding sequence of either *TAF4, TAF5, TAF12*, or *RAP1*. Expression of the protein must be directed from 5' and 3' regulatory sequences unrelated to those directing expression of *TAF4, TAF5, TAF12*, or *RAP1* from their normal respective chromosomal loci, in order to avoid direction of homologous recombination events into the plasmid-borne sequence. Second, a DNA fragment designed to replace or knockout the chromosomal *TAF* or *RAP1* open reading frame is needed and this DNA fragment will include an expression cassette for either KAN^{MX} or HPH^{MX}, genes that encode resistance to either G418 or Hygromycin B, respectively. In addition to an antibiotic resistance gene this DNA fragment will contain homology arms that correspond to the 5' and 3' sequences immediately upstream or downstream of the *TAF* or *RAP1* initiator methionine codon or stop codon, respectively. The length of homology used varied from ~0.4 to 1.2 kb, although homology as short as 45 bp can direct homologous recombination of a cassette to the intended locus. Longer homology arms increase both the likelihood and efficiency of a precisely targeted recombination/gene replacement event, and were employed for this reason. Covering plasmids were created

by ligating *TAF4* or *TAF12* ORFs as 5' BamHI and 3' SalI DNA fragments into equivalently digested p416ADH (Funk et al., 2002), while *TAF5* and *RAP1* covering plasmids were created by inserting 5' NdeI and 3' XhoI DNA fragments into p416ADH. The 5' NdeI sites of inserts and 5' BamHI site of p416ADH were made blunt using the Klenow fragment of *E. coli* DNA Polymerase I; the 3' XhoI site of inserts and vector allowed directional ligation into vector. *TAF4, TAF5, TAF12* and *RAP1* fragments originated from bacterial expression vectors that had been sequenced in their entirety (described below). Targeting fragments were created in pAG32 (Goldstein and McCusker, 1999). 5' homology arms were typically created as PCR-generated 5' SpeI-3'EcoRV fragments, while 3' homology arms were typically designed as 5'BglII-3'HindIII fragments. Targeting constructs were created via four-component ligations including SpeI-HindIII-digested vector, EcoRV-BglII-digested hygromycin-resistance cassette, a SpeI-EcoRV-digested 5' homology arm, and BglII-HindIII-digested 3' homology arm. To convert these targeting constructs from hygromycin resistance to G418 resistance, an EcoRV to BglII fragment derived from pFA6 KAN^{MX} (Wach et al., 1994) was liberated and ligated to equivalently digested pAG32-based targeting vectors. Hygromycin resistance targeting vectors were digested with SpeI and HindIII, G418 resistance targeting vectors were digested with SpeI and NotI; enzymes were heat-inactivated and DNA ethanol precipitated and resuspended in 10 mM Tris pH 7.9, 0.1 mM EDTA. 25-50 μg of cut targeting fragment and 5-10 μg of corresponding covering plasmid were mixed and used to transform SEY6211, BY4741, or W303 (from 2 to 3 individual 250 ml YPD-media cultures grown to mid-log phase (A^{600} 0.4-1.7)); the high efficiency TRAFO yeast transformation method was used exclusively (Gietz and

Schiestl, 2007). Yeast cells from the large-scale one-tube transformation reaction were distributed equivalently across 20-40 100 mm agar plates containing SC-Ura media; plates were incubated at 30° C. Upon evidence of colony formation, typically requiring 2-3 days incubation, each individual plate (~1000 colonies) was replica-plated using sterile velvet cloths to YPD agar plates containing hygromycin or G418 at the appropriate concentration (Goldstein and McCusker, 1999; Wach et al., 1994). These were further incubated for 3 to 7 days, at which point antibiotic-resistant colonies were picked to 0.5 ml YPD cultures. These cultures were grown an additional 3-7 days, and then plated onto 150 mm agar plates containing synthetic complete media plus or minus 1 mg/ml 5-FOA. Use of 96 well plates and a pinning device resulted in 96 clones being analyzed per 150 mm plate. Clones exhibiting normal growth rates on the plate lacking 5-FOA, but exhibiting complete absence of growth on the plate containing 5-FOA were repicked and tested for growth at various temperatures. The most robust strains, (although little variation in phenotype was observed) were transformed with *HIS3*-marked plasmids either lacking or containing *TAF4, TAF5, TAF12*, or *RAP1,* and tested for growth at 23°, 30°, and 37° C on SC-His plates either containing or lacking 5-FOA. Strains that grew equivalently at each temperature in the presence of the respective wild type *TAF/RAP1* gene and in the presence of 5-FOA, but not when *TAF/RAP1* was absent, were designated for use in future experiments as the shuffling strain, because these strain's growth properties indicated dependence on a *URA3*-marked plasmid containing the *TAF* or *RAP1* gene. To the best of my knowledge, this procedure I have designed for generating shuffling strains is the most expedient method available (JHL, unpublished observation). Note that the reason this works is due to the use of heterologous regulatory sequences

(*ADH1*) to drive *TAF* or *RAP1* expression on the *URA3*-marked covering plasmid, therefore disruption events are restricted to the chromosomal *TAF* or *RAP1* locus. Because *TAF* and *RAP1* genes are essential for viability, yeast clones that are Ura^{+}, $Hygro^{R}$ or $G418^{R}$, and FOA sensitive by necessity contain a chromosomal insertion/disruption at the target gene locus that is complemented by the *URA3* marked covering plasmid. The absence of any other overt growth phenotypes indicates that there are likely not additional/random/unwanted genomic insertions conferring antibiotic resistance in strains exhibiting the growth phenotypes I described above.

Expression Vectors

TAF4, TAF5, TAF12, and *RAP1* fragments including the complete open reading frames were PCR amplified from yeast genomic DNA and inserted into pBG101 (http://structbio.vanderbilt.edu/wetlab/vectors.php), pET28A, or pET33B (latter two from Novagen). PCR was performed using a mixture of in-house expressed and purified Pfu and Taq DNA polymerases (expression and purification discussed below). Pfu polymerase has one of the lowest misincorporation rates of commonly available hyperthermophilic polymerases due to its intrinsic proofreading activity; its low processivity is counteracted by including Taq polymerase (Taq has better processivity but no proofreading activity). This blend results in a highly economical method to produce useful quantities of gene fragments essentially free of unwanted PCR-introduced point mutations. The proofreading capability of Pfu is due to its 3'-5' exonuclease activity, which is further exploited in a Ligation Independent Cloning (LIC) strategy (Aslanidis and de Jong, 1990; Aslanidis et al., 1994; Haun et al., 1992; Vaillancourt et al., 2000).

The pBG series of vectors can be digested with BamHI and HindIII, gel purified, and incubated with Pfu in the presence of dTTP, which leaves LIC compatible overhangs (JHL, unpublished observation). PCR primers for *TAF* or *RAP1* amplification include 5' overhangs C CAG GGG CCC GGA **T**CC NNN NNN *ATG* for the 5' primer and GC CGC CGC AAG C**T** NNN NNN *TTA* for the 3' primer, where the bold **T** nucleotide represents the endpoint for a single stranded LIC overhang resulting from incubation of gel purified, double stranded PCR products with Pfu in the presence of dATP. The optional NNN NNN sequence represents the appropriate location for nesting a unique restriction enzyme site, which can be used to recover DNA fragments from pBG series vectors for conventional ligation-mediated subcloning into other bacterial or yeast expression vectors. Note that this included restriction enzyme site is on the 5' flank of an *initiator methionine* or *stop codon*, so as to give inserts translational initiation and termination signals that are independent of destination vector sequences. 5' primer restriction sites were BamHI for *TAF4* and *TAF12* (NNN NNN sequence not required, it is already included in the generic 5' LIC primer) and NdeI (CAT ATG) for *TAF5* and *RAP1* (*ATG* not required); 3' primer restriction sites were SalI for *TAF4* and *TAF12* and XhoI for *TAF5* and *RAP1*. An additional 20 to 26 nucleotides are added to the 3' end of each primer, which are complementary to *TAF* or *RAP1* DNA sequence. The LIC-compatible PCR primers were a maximum of 45 nucleotides in length. Primers were designed as such to allow serial truncation of each *TAF* gene from the amino and carboxy terminal coding sequences; the location of truncation points was chosen based upon alignment of Taf protein sequences from *S. cerevisiae*, *C. albicans*, *S. pombe*, *C. elegans*, and *D. melanogaster*. Using genomic DNA as template, full length *TAF* and *RAP1*

coding sequences were amplified as described above. Products were precipitated and gel purified then incubated with Pfu in 1X Pfu PCR buffer (Lu and Erickson, 1997) and 1 mM dATP for one hour at 65°. A portion of the reaction products were mixed with LIC-treated pBG101 (dTTP) and allowed to cool to room temperature over the course of 30 minutes, then an appropriate amount of the LIC mixture was used to transform BW23474 with selection for kanamycin resistance. Insert bearing clones were sequenced in their entirety to confirm the appropriate reading frame and absence of any mutations. Full length and 100% correct clones (typically 80-100% of those examined) were used as templates for the PCR amplification of *TAF* truncations; Taq polymerase was omitted from these reactions to minimize the chance of introducing unwanted mutations. Pfu alone is normally sufficient for any amplification reaction insofar as the template is a plasmid bearing the target sequence. LIC was used to insert the systematic families of *TAF* truncations into pBG101. The *TAF5* family was excised from the pBG101 vectors by digestion with NdeI and XhoI and subcloned into pET28A. Likewise, *RAP1* was excised from pBG101 and transferred to pET28A and pET33B. Families of *TAF5* and *TAF12* truncations in pET28A or pBG101 and pET33B *RAP1* were directly transformed to BL21 Gold DE3 pRARE from the LIC reaction. The *TAF4* family in pBG101 (LIC reaction) was mixed with pACYC11B *TAF12* and transformed directly to BL21 Gold DE3 RIL Strept. Two to three individual clones of each construct were picked to liquid LB cultures with kanamycin and chloramphenicol, a portion of each liquid culture was expanded for protein expression while the remainder was used for plasmid miniprep. Insert-positive clones were sequenced from 5' ends using either GST oligo primers (pBG101) or the T7 primer (pET28A and pET33B); sequence from 3' end was

determined using the T7 terminator primer. Inserts were excised from those vectors found to express protein of the expected molecular weight and lacking point mutations; these inserts were subcloned to an appropriate yeast expression vector.

Yeast vectors were based upon the "Funk" series (Funk et al., 2002); these vectors contain the polylinker of pBluescript SK$^-$, the minimal centromeric portion of chromosome IV, and a replication origin from chromosome I. Several steps were used to create the yeast expression constructs for *TAFs* and *RAP1*. First, the *ADH* promoter fragment of p413 *ADH* was removed by digestion with KpnI and SpeI, and replaced with PCR-generated KpnI-SpeI fragments containing the 5' genomic regulatory sequences of *TAF4, TAF5, TAF12*, or *RAP1*. Next a synthetic oligonucleotide duplex encoding three tandem copies of the HA epitope and the nuclear localization signal of the SV40 large T-antigen was inserted between the SpeI and BamHI sites of each p413 Promoter Vector. The triple HA tag/NLS (HA^3/NLS) was designed using optimized codon usage for yeast (www.kazusa.or.jp/**codon**/) and dual glycine codons were encoded between each HA repeat and the NLS; this epitope tags ensures unbiased detection of different Taf and Rap1 protein variant levels and nuclear localization in the event that natural Taf or Rap1 NLS are disrupted. Using this design, the reading frame of the p413 Promoter HA^3 /NLS vectors is the same as that of the pBG series of bacterial expression vectors that *TAF* and *RAP1* variants were initially subcloned into. Inserts from *TAF4* bacterial expression vectors were inserted into p413 *TAF4* HA^3/NLS as BamHI/SalI fragments, the p413 *TAF12* HA^3/NLS family of vectors was created similarly. p413 *TAF5* HA^3/NLS vectors were created by cutting pET28A *TAF5* vectors with NdeI, filling in the end with the Klenow fragment of *E. coli* DNA polymerase, and liberating the *TAF5* fragment with

XhoI, followed by inserting these fragments into p413 *TAF5* HA^3/NLS cut with BamHI and XhoI where the BamHI site was cut and filled in with Klenow fragment. *RAP1* vectors were generated as with *TAF5* except that the *RAP1* promoter- HA^3/NLS fragment was first inserted into p415 rather than p413. p415 confers ability to grow on media lacking leucine, whereas p413 confers the ability to grow on media lacking histidine. Because all inserts had already been characterized within the bacterial expression vectors, all yeast expression vectors were deemed ready to use upon confirmation of DNA insert presence and length.

Site directed mutagenesis by gene SOEing (Ho et al., 1989) was used to create internal deletions in *TAF4*, *TAF5*, and *RAP1*. SOEing mutagenesis involves a three step PCR procedure using Pfu and a kanamycin-marked *TAF4, TAF5*, or *RAP1*-containing vector as template for the first two PCR steps. These first two steps involve creation of a product that slightly overlaps the 5' end of the gene region being mutagenized, likewise a product is created that slightly overlaps the 3' end of the targeted gene region. Each product is gel purified, and an equimolar mixture of the two products is used as the "template" for a third PCR reaction, using the outside primers employed in the initial two PCR reactions. The slight overlap between the two initial PCR products effectively creates a "megaprimer" allowing the synthesis of a product equivalent in length to the sum of the two initial PCR products that contains the desired site-directed alteration. Restriction enzyme sites encoded in the outside PCR primers allow the final product to replace a gene fragment within the appropriate yeast expression vector. For *TAF4* the most convenient sites are an internal EcoRV site and the 3' SalI site. *TAF5* site-directed mutagenesis used internal SmaI and EcoRI sites. *RAP1* site-directed mutagenesis used an

internal BlpI sites and a 3' XhoI site. Highly efficient site-directed mutagenesis occurs only when the third PCR product is less than 1 KB in length, and the recipient vector is digested to completion and treated with phosphatase to prevent unwanted vector closures during ligation (JHL, personal observation). Note that this method of site-directed mutagenesis allows essentially unlimited manipulation of DNA sequence, allowing seamless fusion of any possible desired DNA fragments. One need only have the cloned DNA in hand, and be able to identify unique restriction sites near the region(s) being manipulated.

Protein Expression and Purification

Transformations to BL21 Gold DE3 pRARE or BL21 Gold DE3 RIL Strept were seeded on agar plates containing all antibiotics (kanamycin, chloramphenicol, and streptomycin, or just the first two) and incubated overnight at 30° C. A single colony was picked the next morning or early afternoon into a 5 ml LB culture containing all antibiotics and placed into a shaking incubator at 30° C until early or late evening. At this point the cultures were just visibly turbid (A^{600} 0.2-0.8). A 1/1000 dilution was made into fresh auto-induction media ((Sreenath et al., 2005; Studier, 2005; Tyler et al., 2005) containing all antibiotics and cultures were grown with aeration at room temperature. Using this procedure, protein expression is typically evident at 16 hours post-inoculation and is maximal at 24 to 30 hours post-inoculum, depending on the individual protein. Cells were harvested at 4000 rpm for 5-10 minutes in a Beckman J6-HC swinging bucket centrifuge. Expression by auto-induction facilitates the parallel expression of multiple proteins because it is not necessary to simultaneously monitor culture densities in order to

determine the appropriate time to add inducing agent; instead all cultures grow to near saturation and automatically induce protein expression at that time. Consequently cell pellet mass is almost always nearly identical irrespective of the protein being expressed, a feature that greatly helps simplify extract preparation and the overall lack of variation in protein yield from protein to protein. As an added advantage, because of the additional harvested cell mass, the amount of recombinant protein per unit culture volume typically exceeds the yield from equivalent conventionally-induced cultures by ten to twenty fold. Most importantly, the only limitation to the number of different recombinant proteins that can be expressed using this method is the amount of shaking incubator space available.

Unless otherwise stated, all protein isolation procedures were performed at 4° C in conjunction with ice-cold buffers and pre-chilled equipment such as ultracentrifuge bottles and rotors. Pfu DNA Polymerase was expressed from vector pET11A Pfu and purified as follows (Lu and Erickson, 1997). Frozen cell pellets from 6 liters auto-induced culture were dispersed in 150 ml of ice-cold buffer containing 50 mM Tris pH 7.6, 150 mM potassium acetate 0.1% Triton X100, 10% Glycerol, supplemented with protease inhibitors TPCK, TLCK, Pepstatin, Leupeptin, Aprotinin, Benzamidine, and PMSF (Sanders et al., 2002a). To disperse the frozen cell pellet the buffer was added and containers with buffer and frozen cell pellet was allowed to rotate gently, ~150 rpm, on a New Brunswick floor model shaker at room temperature. The cell suspension was poured into a 250 ml plastic beaker and kept on ice. A Model 250 Branson Sonifier outfitted with the large probe was used to reduce the viscosity of the cell suspension by 5 rounds of sonication at 90% duty, setting five, with 6 bursts per round of sonication. The sonicated extract was cleared of debris in a Beckman 45 Ti preparative rotor in the LE-

80K ultracentrifuge, 43,000 rpm for one hour. The soluble extract was poured into a 300 ml ehrlenmeyer flask with a screw-on lid and sealed tightly, and placed in a 75° C water bath for 30-45 minutes. Heat-denatured materials were pelleted using a Sorvall SS34 rotor and a 15 minute spin at 14,000 rpm. The supernatant, primarily containing recombinant Pfu DNA Polymerase, was applied to an equilibrated 75 ml bed of Phosphocellulose P11 cation exchange resin (Whatman) packed in a glass column. The column was washed with resuspension buffer until UV absorbance of the effluent returned to baseline levels, (2-3 column volumes). Pfu was eluted using a 4 column volume salt gradient of 150-1000 mM potassium acetate with buffer components as above. Pfu eluted at 600-700 mM potassium acetate (determined using a conductivity meter and a salt standard curve). Peak fractions were pooled and dialyzed against Pfu storage buffer (Lu and Erickson, 1997) containing 50% Glycerol. At this stage Pfu was 95-99% pure as determined by staining of SDS-PAGE gels and was highly active for PCR. A working stock was prepared by diluting this material 40 fold. This enzyme preparation was subsequently used to generate every recombinant DNA construct made in the lab during my tenure.

Taq DNA Polymerase was expressed from vector pTTQ18 TAQ and purified using a combination of methods (Engelke et al., 1990; Pluthero, 1993). The standard procedure is very similar to that followed for Pfu except the heat-cleared extract was dialyzed against buffer containing ammonium sulfate at 50% saturation (.291 grams ammonium sulfate per ml of extract) at 4° C, overnight. Insoluble material was pelleted in the SS34 rotor as above and resuspended in Engelke's starting buffer containing 50 mM potassium acetate, and applied to a packed column of Bio-Rex 70 cation exchange

resin. After washing with buffer containing 50 mM potassium acetate, the column was step eluted with two column volumes of buffer containing 400 mM potassium acetate. Peak Taq-containing fractions were pooled and dialyzed against Taq storage buffer containing 50% glycerol (Engelke et al., 1990). As with Pfu, Taq prepared this way is highly concentrated, pure, and active. This Taq preparation, which was used for analytical PCR such as colony PCR and ChIP, was essential for creation of mutagenized *TAF* libraries, and also was blended with Pfu when amplifying targets from genomic DNA. Recombinant dUTPase from *Pyrococcus woesei* (*Pwo*) was prepared and used to supplement difficult PCR amplification reactions ((Dabrowski and Kiaer Ahring, 2003). Pwo dUTPase is expressed from pET30 Pwo and purified by successive heat treatment of extract and immobilized metal affinity chromatography (IMAC, Ni-NTA agarose, Qiagen). dUTPase helps to maintain high activity of Pfu polymerase during thermal cycling reactions by limiting the concentration of uracil created by heat-stimulated cytosine deamination; because uracil is an allosteric inhibitor of Pfu its removal by hyperthermophillic dUTPase during thermal cycling facilitates successful high-fidelity PCR amplification of targets longer than 2 kb.

Rap1 was expressed from pET33B. Cloning *RAP1* into pET33B using the NdeI and XhoI sites resulted in inclusion of an N-terminal hexahistidine tag followed by a "kemptide" sequence allowing high-efficiency phosphorylation by the catalytic subunit of bovine heart cyclic AMP-dependent protein kinase. Frozen cell pellet from 6 liters auto-induced culture was homogenized and lysed in buffer containing 20 mM HEPES pH 7.9, 200 mM potassium acetate, 20 mM imidazole, 0.1% Triton X100, 10% glycerol and protease inhibitors as above. Insoluble debris and unbroken cells were removed using a

Beckman LE80K ultracentrifuge and 45Ti rotor by spinning for 1 hour at 43,000 rpm. Nucleic acids and Rap1 were precipitated from this cleared extract with polyethyleneimine (Burgess, 1991). Rap1 was recovered using 400 mM potassium acetate-containing buffer washes of the pellet and then precipitated using ammonium sulfate (35% of saturation, .194 grams ammonium sulfate per ml of soluble material). This pellet was dissolved in buffer as above except with ionic strength of 50 mM. Soluble material was directly applied to a phosphocellulose column and eluted using a gradient of 150-1000 mM potassium acetate in starting buffer. Peak fractions identified by SDS-PAGE were pooled and directly applied to an IMAC column (Ni-NTA Agarose, Qiagen). After step-eluting this column with 300 mM imidazole in starting buffer, peak fractions identified by A280 were pooled and dialyzed over 12 hours against two 1 liter changes of 20 mM HEPES pH 7.9, 150 mM Potassium Acetate, 0.1% Triton X100, and 30% glycerol. At this stage Rap1 was highly concentrated and electrophoretically homogenous. Rap1 prepared by this method was subsequently shown to be active in a variety of protein-protein interaction assays, highly active for protein-DNA interaction as scored by EMSA, and useful for affinity purification of specific anti-Rap1 IgGs.

The Taf12 family of proteins was expressed using pBG101. This vector encodes N-terminal His^6- and GST-tags, and a cleavage site for human rhinovirus (HRV) 3C protease. Cell pellets from 50 ml auto-induced cultures were resuspended in 3 ml buffer containing 25 mM HEPES pH 7.9, 500 mM NaCl, 30 mM Imidazole, 0.1% Triton X100, 10% Glycerol, and protease inhibitors. The cell suspension was sonicated using the Model 250 Branson Sonifier as described above. Insoluble debris was cleared using the Sorvall SS34 rotor as above, and the supernatant was applied to a 0.2 ml bed of Ni-NTA

agarose and mixed in batch at 4° C for 1 hour. Resin was pelleted by gentle centrifugation and the supernatant gently poured off and saved at -80° C. The column was washed twice with the resuspension buffer, twice with the same buffer except containing 200 mM NaCl, and eluted using the last buffer supplemented with imidazole at 300 mM. All eluted material was applied to a 0.5 ml bed of Q sepharose fast flow resin (Pharmacia) and mixed in batch at 4° for 1 hour. The mixture was poured into a small spin column (BioRad) and the eluate was recovered by centrifugation. A large portion of the major contaminants bound to the resin while the breakthrough fraction, containing most of the Taf12 forms that were input, was applied to a 0.5 ml bed of SP sepharose fast flow resin (Pharmacia) and mixed in batch at 4° C for 1 hour. After pelleting beads and discarding the supernatant, the column was rapidly washed with buffer containing 25 mM HEPES pH 7.9, 200 mM NaCl, 0.1% Triton X100, 10% Glycerol, and protease inhibitors, then step eluted using the same buffer except containing 1000 mM NaCl. Most of the remaining contaminants were found in the SP breakthrough fraction. The bulk eluent was diluted three fold with buffer lacking NaCl, and used for Far Western protein-protein binding studies, as described below. The Taf12 protein family purified in this manner was essentially free of higher molecular weight contaminants as indicated by SDS-PAGE, with remaining contaminants most likely corresponding to proteolytic fragments of Taf12 forms.

The Taf4 family was expressed initially from pBG101 but was found to be highly degraded when purified from soluble bacterial extracts. The solution to the degradation problem was to coexpress the Taf4 family along with Taf12, which was expressed from pACYC11B (Fribourg et al., 2001). To include extra tRNAs for rare codons found in

eukaryotic genes, the plasmid pSC101 RIL was included in design of a bacterial expression strain appropriate for carrying pBG101 *TAF4* along with pACYC11B *TAF12*, which requires selection with chloramphenicol. This was the primary rationale for creation of bacterial strain BL21 DE3 Gold RIL Strept. Cell pellets from two liters of autoinduced culture corresponding to each Taf4 variant were resuspended in buffer containing 25 mM HEPES pH 7.9, 150 mM NaCl, 30 mM Imidazole, 0.1% Triton X100, 10% Glycerol, and protease inhibitors and sonicated as described for Pfu. Insoluble debris was removed by a 30 minute spin in the 45 Ti rotor at 43,000 rpm, and supernatant was applied to a 5 ml bed of Ni-NTA agarose and bound in batch with mixing for 1 hour at 4° C. After washing with the starting buffer, the column was eluted with starting buffer supplemented with 300 mM imidazole. The resulting material was incubated with Q sepharose as for Taf12, which bound a substantial portion of contaminants but not Taf4-Taf12 complexes. Unbound material was subjected to chromatography on SP Sepharose as was described for Taf12. The 1 M NaCl eluant from SP sepharose was subjected to preparative SDS-PAGE. This served to remove any remaining contaminants, degraded Taf4 fragments, and Taf12, allowing recovery of essentially pure Taf4 variants of the correct molecular weight. 4%/10% SDS-PAGE gels were cast with preparative combs, and after running the predominant protein species (Taf4 variants and Taf12) were identified by staining briefly with 4 M sodium acetate (Higgins and Dahmus, 1979). Excised bands were diced, and each band, corresponding to one Taf4 variant, was placed into one fixture in a Model 422 electoelution apparatus (BioRad) and recovered by being left in the electroeluter overnight at 20 volts. Two Model 422 apparati were operated in parallel thus two runs were necessary to recover the eluted material corresponding to all

Taf4 variants. Protein preparations were essentially homogenous at this stage and used for Far Western Analyses as described below.

Taf5 variants were initially expressed from pBG101 but subsequently expressed from pET28A (see explanation below). Any Taf5 variants containing the C-terminal portion of the protein are found in the insoluble fraction of bacterial extracts, irrespective of temperature during induction of expression. Cellular extracts were prepared from frozen cell pellets corresponding to two liters of induced culture. Homogenates were prepared in 25 mM HEPES pH 7.9, 500 mM NaCl, 30 mM imidazole, 0.1% Triton X100, 10% glycerol, and protease inhibitors using the Branson Model 250 Sonifier as described for Pfu and Taf4/Taf12 variants. Insoluble material was pelleted using the Sorvall SS34 rotor, the supernatant was discarded, and the insoluble pellet was washed using sonication in starting buffer where 0.5% sodium deoxycholate was substituted for Triton X100. The remaining insoluble material was recovered by centrifugation, the supernatant discarded, and the pellets resuspended in buffer containing 25 mM HEPES pH 7.9, 150 mM NaCl, 30 mM imidazole, 6 M guanidine hydrochloride 0.1% Triton X100, 10% glycerol, and protease inhibitors. After mixing, including dispersion with wooden toothpicks and rotation of sealed tubes, non-proteinaceous debris was again removed using the SS34 rotor. Solubilized material was applied to a 5 ml bed of Ni-NTA agarose and incubated in batch with mixing at room temperature. After two hours binding, resin was recovered by gentle centrifugation, supernatant was discarded and two washes using guanidine buffer were performed. Afterwards two additional washes were performed using buffer containing 4 M urea as a substitute for guanidine hydrochloride. Note that this modification was important in lieu of preparative SDS-PAGE and/or Far Western

Analyses, as guanidine hydrochloride-containing buffers are incompatible with SDS-PAGE. Denatured protein was eluted from Ni-NTA using buffer containing 25 mM HEPES pH 7.9, 150 mM NaCl, 300 mM imidazole, 4 M urea, .1% Triton X100, 10% glycerol, and protease inhibitors. Taf5 variants were subjected to preparative SDS-PAGE and recovered as specified for Taf4 variants.

Steady-state levels of Taf4, Taf5, Taf12 and Rap1 variants in yeast cells were determined using SDS-PAGE and immunoblotting. Because the serial truncations and internal deletions of Tafs and Rap1 might remove epitopes recognized by our polyclonal antibody preparations, which were raised using full-length Tafs or Rap1 as antigens, it was important to include a suitable epitope tag in the expression vectors encoding Tafs and Rap1. The HA epitope from the Influenza virus haemaglutinin protein was selected because commercially available antibodies against this epitope tag are highly sensitive and selective when using yeast extracts (Kolodziej and Young, 1991). Incidentally, detection of Taf and Rap1 variants was performed using our polyclonal antibodies, and this indicated the distribution of protein-specific epitopes recognized by the polyclonal antibodies, when compared to a parallel blot that was incubated with anti-HA antibody (not shown). Cells were often grown at various temperatures corresponding to permissive, semi-permissive, or non-permissive temperatures according to the Taf allele in question. Because yeast cells have a dense cell wall enriched in complex carbohydrates, cell pellets were harvested from logarithmically growing cultures containing plasmids encoding Taf or Rap1 variants, and suspended in 100 mM NaOH. This caustic treatment acts to partially hydrolyze the cell wall material, rendering cells sensitive to detergent-mediated lysis (Kushnirov, 2000). After a five-minute NaOH

treatment, cells were pelleted by centrifugation, the supernatant discarded, and cells resuspended in 1X NuPAGE sample buffer at 75° C. 5-10% of the resuspended volume was directly loaded on BioRad or Invitrogen precast 4-12% gradient SDS-PAGE gels and resolved at 150-180 volts for one to two hours. Gels were then electroblotted to Immobilon P (Millipore) for 2-2.5 hours at 12 volts, and membranes were blocked in 5% non-fat dry milk (NFDM) in 10 mM Tris pH 7.5, 150 mM NaCl (TBS) for 30 minutes at room temperature. After three five minute washes in TBS, blocked membranes were incubated with a 10 ml solution containing 1% NFDM in TBS that also contained antibodies. Antibody incubations were performed in seal-a-meal bags at room temperature for no less than 3 hours to overnight at 4° C. All Taf and Rap1 variants were detected using anti-HA antibody at 1:5,000 dilution (monoclonal 3F10 horseradish peroxidase conjugate, Roche Diagnostics) while anti-actin mouse monoclonal antibody (Abcam) was used to provide an internal loading control. After the primary incubation, three five minute washes were performed followed by a 15 minute incubation at room temperature in a seal-a-meal bag containing secondary antibody solution (rabbit anti-mouse IgG at 1:10,000 dilution in 1% NFDM/ TBS). This served for detection of the actin signal. The membrane was washed once more and bound antibody-HRP conjugates were detected using ECL (Pierce Super Signal Pico West) and a time course of film exposures.

Far Western Interaction Assay

Because the achievable resolution of defining Rap1 binding domains in Tafs was significant for the success of my experimental agenda, it was clear that examination of a

collection of Taf N- and C-terminal truncations was prudent. Several major considerations went into the choice of methodology for defining Rap1 binding domains. Firstly, it was important to measure relative binding affinities of Rap1 with Taf truncation variants in a parallel manner, although practically speaking, the greater the complexity of an experimental design, the lower the likelihood of obtaining meaningful data. Second, both and N- and C-terminal truncation variants of Tafs should inform about the location of Rap1 in a consistent way, but solution-based assays were ruled out for at least some of the Taf variants, particularly N-terminal truncations of Taf5, since those proteins were expressed in an insoluble form. Coexpression of Taf5 with Taf6 and Taf9 was pursued as this basic approach has resulted in production of recombinant Taf5/Taf6/Taf9 complexes in the Baculovirus system (Berger et al., 2004), but this did not provide satisfactory results using *E. coli* coexpression. For these reasons the 'Far Western' or 'Overlay' assay was employed as it allows parallel comparison of many protein variant's binding activity but with minimal handling, and also does not mandate availability of soluble protein variants to measure binding events. This is a widely accepted approach that has been used to make many important biological observations, such as the identification of targeting proteins that localize enzymes of cell signaling (Carr et al., 1992) and characterization of *E. coli* activator-polymerase interactions (Burgess et al., 2000).

Taf12, Taf4 or Taf5 forms were loaded in 1X NuPAGE sample buffer without heating on pre-cast 26 well 4-12% SDS PAGE gels (BioRad 345-0125) and run at 150 volts for 90 minutes in 1X MOPS NuPAGE running buffer. Gels were immediately electro-blotted onto PVDF membranes (Millipore Immobilon P) for 120 minutes at 12 volts in an Idea Scientific Genie Blotter apparatus using NuPAGE transfer buffer with

10% methanol. Membranes were placed on a rotary incubator at 4° C in renaturation buffer containing 20 mM HEPES pH 7.6, 75 mM KCl, 2.5 mM magnesium chloride, 0.1 mM EDTA, 0.05% Triton X100, 1 mM DTT, and 1% w/v BSA (Roche Ultra Pure, cat no. 100377, lot no. 93097226). The large amount of BSA in the renaturation buffer serves to compete away SDS bound to proteins embedded in the membrane, theoretically facilitating their partial renaturation. However, BSA is not as effective at blocking non-specific sites of interaction for Rap1 and/or antibodies that are subsequently incubated with the membrane. To improve signal-to-noise ratio and overall data quality, after 90 minutes the initial renaturation buffer was discarded and replaced with buffer containing 5% w/v NFDM instead of BSA. Following 30 minutes incubation at room temperature on a rotary incubator, membranes were washed at room temperature with five three-minute washes using renaturation buffer without BSA or milk. Next membranes were incubated in a heat-sealed bag on a tiltboard at 4° C for 120 minutes with renaturation buffer plus BSA at 1% W/V and Rap1 at 3.5 nM (0.12 ml buffer per cm^2 membrane). This and all subsequent steps were performed at 4° C using chilled buffer. Excess Rap1 was washed away using three one-minute washes with renaturation buffer. Antigen-purified anti-Rap1 rabbit IgG was incubated with membranes at 1:20,000 dilution in renaturation buffer plus 1% w/v milk for 60 minutes. Excess anti-Rap1 IgG was washed away then HRP conjugated goat anti-rabbit IgG (Roche) was applied to membranes at a 1:20,000 dilution for 10 minutes. After one more washing step, Taf-Rap1-IgG Complexes were detected via ECL (Pierce Super Signal Pico West) and film exposure (2-5 minute exposure times). Parallel controls included membranes treated identically except that either Rap1 or primary anti-Rap1 antibody were omitted from binding

solutions; these membranes presented no detectable signal in lanes containing Taf forms (not shown, (Colbran et al., 2003)). Notably, these experiments were initially performed using Rap1 that was labeled with γ^{32}P-ATP using the catalytic subunit of Protein Kinase A (gift of Drs. Sharron Francis and Jackie Corbin) using home-made components and the procedure detailed in the PKAce kit manual (Novagen). Note that Rap1 is a good substrate for phosphorylation by PKA irrespective of whether the N-terminal "Kemptide" is present on the recombinant protein. The significance of this has not been pursued. Radiolabeled Rap1 was desalted from free ATP using a 5 ml column of G25 Sephadex equilibrated in renaturation buffer plus 1% BSA, and used in binding reactions as detailed above. After removing the Rap1 binding reaction from the membrane, it was washed briefly, placed in an acetate envelope and subjected to autoradiography without an intensifying screen for ~36 hours at room temperature.

Plasmid Shuffle Assay

In the study of both *TAF* genes/proteins and *RAP1*, it is advantageous that in every case the protein-coding genes are necessary for cellular growth. Thus simple plate-based growth assays allow for rapid complementation tests of gene and protein function(s). The design and construction of yeast "shuffling strains" and sets of *TAF* and *RAP1* variants encoded within streamlined plasmid vectors allows rapid complementation tests using the plasmid shuffle assay (Sikorski and Boeke, 1991). As described above, these yeast strains were created along with yeast expression vectors containing systematic variants of *TAFs* or *RAP1,* and these vectors were transformed into the corresponding *TAF* or *RAP1* yeast shuffling strain. For *TAFs* these vectors conferred the ability to grow

on media lacking histidine while *RAP1* vectors allowed growth on media without leucine. Transformants were plated onto the appropriate selective media and grown at 30° C for 2-3 days. A single colony was then picked to a 5 ml liquid culture with the appropriate selective media and grown at room temperature without aeration for several days after the culture had appeared to reach saturation (typically 5-14 days). Note that the inclusion of uracil in the media used for both plating and liquid cultures expedites the loss of the *URA3*-marked plasmid that carries the WT copy of the *TAF* or *RAP1* gene found in the parental shuffling strain. This means that if a particular *TAF* or *RAP1* variant possesses all the characteristics needed to support growth, it is possible for the cells to lose the wild type *TAF* or *RAP1* gene found on the *URA3*-marked plasmid and become resistant to the drug 5-FOA, which is a poison to any cells containing the *URA3* gene product. Within a single cell, loss of the *URA3* plasmid occurs through a chromosomal mis-segregation event that only becomes evident within a mixed population over many generations, and several generations are also needed to allow turnover of the *URA3* gene product that confers sensitivity to 5-FOA. This is the reason for including uracil in all growth media, beginning at the transformation step, when one is seeking to look at phenotypes attributed to a particular *TAF* or *RAP1* variant, uncovered by loss of the wild-type *TAF* or *RAP1* allele. To test if a particular variant can act in a dominant manner, it is necessary to omit both histidine (or leucine) and uracil in order to force the maintenance of the wild type-containing plasmid within each cell in the population. For these reasons it is convenient to prepare two liquid cultures from each transformation plate; one lacking just histidine to be used for the complementation test and the second culture lacking both histidine and uracil for test of dominance of the *TAF* or *RAP1* variant of interest. To prepare serial

dilutions and spot them onto a plate, the 5 ml liquid cultures are mixed very well and a 250 μl portion is placed into the leftmost row of a sterile 96 well Bacti Plate (Nalgene-Nunc). Using an 8-channel pipette, 50 μl of each of these cell suspensions are diluted into 150 μl of sterile water and this is repeated six successive times working from left to right, resulting in one undiluted culture sample and 7 one-to-four serial dilutions for each *taf* or *rap1* variant. Each successive dilution is mixed by no less than 10 pumps of the 8-channel pipette before transfer to the next serial. One half of each 96 well plate is touched with a 48-prong pinning tool that lifts ~10 μl of cell suspension from each well, and the pinning tool is then carefully laid onto a 150 millimeter-diameter petri plate containing well-dried solid selective media, resulting in transfer of a uniform bead of liquid suspension onto the agar. The pinning tool is flame-sterilized, cooled briefly, and the remaining 48 wells are partially transferred to the agar plate. Several identical agar plates are prepared this way for the purpose of comparing growth at different temperatures and with different media ingredients. Thus a typical experiment examining complementation properties and also testing for dominance of twelve variants would use twelve SC-His cultures, twelve SC-His-Ura cultures, two 96 well plates, and nine agar plates. Three agar plates would contain media lacking histidine or leucine but also uracil and these would be grown at 23°, 30°, and 37° C. Six plates would contain media lacking histidine or leucine only, three of these would also contain 5-FOA at 1 mg/ml and thus two plates each would be kept at 23°, 30°, and 37° C. Plates are photographed from three to 7 seven days after cell inoculation. After representative images are obtained, a single colony corresponding to a strain containing each variant, is again transferred to liquid culture and used for immunoblotting experiments intended to compare steady-state

protein levels of different variants. This information is necessary to interpret complementation patterns in light of contribution of individual protein fragments; for example if a particular variant is of a reduced cellular concentration relative to the wild-type form, one cannot necessarily extrapolate the importance of specific regions of a protein directly to the complementation properties.

Yeast Two-Hybrid Assay

It was important to establish an alternative method for determining affects of Rap1 or Taf alteration on protein-protein interactions. Two limitations of the Far Western method are that it is potentially labor intensive and does not easily give quantitative information about protein-protein interactions. By comparison, the yeast two-hybrid method is an accepted means of measuring protein-protein interaction albeit indirectly through a genetic readout. Moreover this method does not require any purified recombinant proteins and can provide both quantitative and qualitative information about interactions (Bai and Elledge, 1996, 1997; James et al., 1996). Wild-type *RAP1* was cloned into pGBDU C1 using a blunted NdeI end on the *RAP1* fragment and a blunted BamHI end on the vector while a XhoI site was used on the *RAP1* 3' end that was ligated to a SalI site in the vector. The resulting vector directed expression of a fusion of the Gal4 DBD to the amino terminus of full-length Rap1, using the *ADH1* promoter and a high copy yeast 2μ replication origin; selection for the plasmid was maintained with SC-Ura. *TAF5* and derivatives were cloned into pGBDU C1 using the same strategy. *TAF4* and derivatives were cloned into pGBDU C1 as BamHI/SalI fragments. These constructs or empty pGBDU C1 vector were transformed into yeast strain PJ69-4A (James et al.,

1996). This strain is superb for testing two-hybrid interactions because it has multiple reporter genes including a *GAL7* enhancer/promoter-*ADE2* fusion that allows growth on media lacking adenine, but only when strong two-hybrid interactions occur. This *ADE2* reporter gene is therefore much more selective for authentic interactions, eliminating many false-positives that can occur when testing for expression of other reporter genes such as *HIS3*. Other reporter genes in PJ69-4A encode *E. coli* β-galactosidase and yeast α-galactosidase. These enzymes are readily scored by colorimetric assays and provide quantitative information about positive interactions. PJ69-4A derivatives containing pGBDU C1 with or without *RAP1, TAF4, TAF5* and variants of the latter two were grown in liquid SC-Ura media and transformed with a systematic collection of vectors encoding each TFIID subunit, subcloned in pACT2.2 (Yatherajam et al., 2003) while empty pACT2.2 and a vector encoding full-length Gal4 served as negative and positive controls for two-hybrid interactions, respectively. pACT2.2 expresses N-terminal fusions of the Gal4 AD and a single HA epitope tag to the protein of interest, using the *ADH1* promoter and a high-copy yeast 2μ replication origin (Bai and Elledge, 1997). pGBDU C1 and pACT2.2 derivatives were simultaneously maintained in PJ69-4A using selection with SC-Ura-Leu. Single colonies corresponding to each pGDDU C1 / pACT2.2 pairwise combination were picked to liquid SC-Ura-Leu media and grown to saturation at room temperature without aeration over the course of 2-3 weeks. Serial dilutions were then prepared and spotted onto plates containing either SC-Ura-Leu, SC-Ura-Leu-His plus 1 mM 3-aminotriazole (to inhibit growth permitted by leaky expression of the *HIS3* gene), or SC-Ura-Leu-Ade, where the latter two score for two-hybrid interactions that activate expression of *HIS3* or *ADE2* reporter genes, respectively. Plates were incubated

at room temperature from between 4 days to 2 weeks. A significantly longer time is often needed for strains containing true positive interactions to grow on the media that select for interactions. While a low level of growth on SC-Ura-Leu-His plus 1 mM 3-aminotriazole is typically observed even in the absence of bait-(DBD fusion) or prey-(AD fusion) encoding plasmid this phenomenon is completely suppressed when scoring for activation of the *ADE2* reporter gene on SC-Ura-Leu-Ade. This makes growth patterns significantly easier to assess as to whether or not a true two-hybrid interaction is occurring, as compared to scoring *HIS3* reporter gene activity. Moreover, different interaction partners exhibit a range of colony color on SC-Ura-Leu-Ade. Strong interactions result in rapidly dividing white colonies and weak interactions provide slow growing red colonies while intermediate interactions give different shades of pink colony color (JHL, data not shown). Differences in affinity of protein-protein interactions between different two-hybrid partners probably account for at least some of these observable colony color variations, through indirect mechanisms involving differential *ADE2* reporter gene expression could also occur. Expression levels of Gal4-DBD fusions expressed from pGBDU C1 vectors were determined using western blotting with polyclonal antibodies against Rap1, Taf5, or Taf4. Expression levels of Gal4-AD fusions expressed from pACT2.2 were made using anti-HA antibody 3F10 (Roche Diagnostics).

PCR-Mediated Random Mutagenesis

The parallel pursuit of identifying Rap1 interaction domains within Taf4, Taf5, and Taf12, while considered necessary for success of this project, placed a practical limitation on the amount of resolution that could be obtained. Only being able to identify

rather expansive protein segments, coupled with the complete absence of structural information about these segments at that time meant that a random mutagenesis strategy was the only realistic option to obtain information about mutational sensitivity of DNA encoding these Rap1 interaction domains. I employed my expertise with PCR to achieve this goal. The basic strategy employed PCR amplification of RBD coding sequences under conditions of reduced enzyme fidelity. The basic formulation and protocol are outlined in Cadwell and Joyce, 1992. There are three major modifications to the described protocol. First, a 10X PCR buffer stock was changed to 100 mM Tris-Cl pH 8.3, 500 mM KCl, and 70 mM magnesium chloride, because the above-referenced buffer composition did not work. Second, a homemade preparation of Taq DNA polymerase described above was used. Third, because Taf4 and Taf5 Rap1 interactions domain-coding sequences were flanked on both sides by other unrelated but important coding sequences, and mutagenized sequences had to be restricted to those encoding the Rap1 interaction domains and used to replace the wild type sequences in cloned *TAF4* and *TAF5* genes, long PCR primers had to be used that included appropriate restriction enzyme sites but that avoided unwanted mutagenesis by inclusion of those sequences within the primers. Therefore, PCR primers were designed to be as much as ~100 nucleotides in length and were subjected to preparative denaturing PAGE to isolate full-length primers away from truncated synthesis byproducts. *TAF4* oligonucleotides were designed to use an internal 5' EcoRV site and the artificial 3' SalI site, but were extended to include codons from amino acids 250 and 350, respectively (therefore ~38 codons were included in the 3' primer). *TAF5* oligonucleotides were designed to use an internal 5' SmaI site and a 3' EcoRI site; it was not necessary for these primers to be as great in

length as those for *TAF4*. Otherwise, the relevant ingredients to successful mutagenic PCR are an increased magnesium chloride concentration, as mentioned, the inclusion of Manganese chloride at 0.5 mM final, modified dNTP concentrations (fresh 10X stock 2 mM dGTP and dATP, 10 mM dCTP and dTTP), and increased concentration of Taq DNA polymerase. It was important to add the manganese chloride at a later time to avoid precipitation. These modifications to standard PCR increase the misincorporation frequency of Taq DNA polymerase several hundred fold. 15 pmoles of each primer and 0.5-1 ng of pBG101 *TAF4* or pET28A *TAF5* were also included in each 50 μl PCR reaction. 30 rounds of thermal cycling were used with 94°, 45°, and 72° C, with 1 minute holds per step. Products of ninety six 50 μl PCR reactions were pooled, ethanol-precipitated, and subjected to preparative electrophoresis in 1.5% agarose gels. DNA was recovered using the Qiagen gel extraction kit and 5-10 spin columns per DNA fragment. A large starting quantity of mutagenized DNA fragments was needed due to the low final recovery of insert after successive rounds of restriction digestion and gel purification.

Library Construction

To generate a high complexity plasmid library, it is necessary to have an extremely low background of parental plasmids DNAs present upon transformation of the ligations reactions to bacteria. This was achieved by extensively digesting the vector preparations, treating them with alkaline phosphatase to prevent re-closure of vectors, and gel purification. Plasmid p413 *TAF4* HA^3/NLS (250-500 μg) was digested sequentially with EcoRV, AatII, and SalI followed by alkaline phosphatase treatment and

gel purification. The use of two different restriction enzymes prior to SalI treatment increases the proportion of vector molecules able to accept an EcoRV/SalI insert, since SalI cuts circular DNA at very low efficiency. Plasmid p413 *TAF5* HA^3/NLS (250-500 μg) was digested sequentially with XmaI and EcoRI followed by alkaline phosphatase treatment and gel purification. XmaI digestion was allowed to proceed 24 hours at room temperature to maximize vector cleavage. Each successive enzymatic treatment was preceded by extraction using the Qiagen PCR clean-up kit. Pools of mutagenized *TAF4* and *TAF5* fragments were prepared by successive EcoRV/SalI/gel purification or XmaI/EcoRI/gel purification, respectively. Pilot ligation reactions performed overnight at room temperature confirmed an extremely low vector background and allowed determination of optimal vector to insert ratios giving maximal ligation efficiency. Subsequently, a large-scale ligation reaction was performed in a 200 μl final volume overnight at room temperature. The next morning the two large ligation reactions were distributed into 20 bacterial transformation reactions containing 10 μl of each ligation reaction and 50 μl of BW23474 *E. coli* that had been made chemically competent and frozen at -80 degrees C the week before (transformation efficiency ~ 10^9 CFU/microgram DNA). After incubation on ice for 30 minutes, 20 seconds heat shock at 37°, and 1 hour recovery at 37° C in .5 ml fresh SOC medium, each transformation reaction was plated onto one 150 mm petri dish containing LB Agar plus ampicillin at 100 ug/ml and incubated overnight at 30° C. The following morning/early afternoon, each plate had ~1000 large visible colonies. 5 ml of fresh LB liquid plus ampicillin was added to each plate, which were then wrapped in parafilm, stacked on an orbital shaker, and left to rotate gently for 30 minutes. Afterwards the LB liquid in each plate, corresponding to

either *TAF4* or *TAF5* mutagenized libraries, was decanted into a fernbach flask containing 900 ml LB plus ampicillin and incubated at 30° C with shaking for ~5 hours, which resulted in a significantly turbid culture. The *E. coli* were harvested by centrifugation and plasmids were extracted using the Qiagen Giga Prep kit. Assuming that bias towards certain clones was avoided during the incubation of plates and expansion of libraries in liquid culture, each library contained as many as 20,000 unique clones.

Library Screening

A log phase 250 ml YPD liquid culture of the appropriate shuffling strain was transformed with a portion of the corresponding mutagenized plasmid library. Ten 150 mm petri plates containing SC-His agar, each having 500-1000 yeast colonies after two days incubation at 30° C, were used to pick single colonies directly to wells of sterile 96 well bacti plates containing 200 µl SC-His liquid media. Forty to fifty 96 well plates were prepared per library and incubated at room temperature for from four to seven days. Each 96-well plate contained a clone transformed with empty pRS413 or p413 *TAF4/TAF5* as negative and positive controls, respectively. Thus between ~3500-5000 individual mutant clones were examined, or at maximum ~25% of total clones present in each library. Each well was mixed by ten strokes with an eight channel pipettor, and then a 96 pin device was used to transfer a sample of each well to six replicate 150 mm plates containing SC-His with or without 5-FOA. Accordingly two plates each were incubated at 23°, 30°, and 37° C and examined after 3 to 7 days. The comparison of plates without or with 5-FOA and incubated at different temperatures identified individual colonies that

were defective for growth at elevated temperature but only in the presence of 5-FOA. Several hundred such colonies were repicked from the 23° plates without 5-FOA for *TAF4* and *TAF5* strains. These cultures were grown to saturation and then subjected to serial-dilution growth analyses on plates containing or lacking 5-FOA with incubation at various temperatures. Clones that again exhibited temperature conditional growth but which grew comparably to wild type at lower temperatures were further screened by western blot analyses. Loading-adjusted immunoblots using anti-HA and anti-actin antibodies were used to select clones that exhibited steady state Taf4 or Taf5 protein abundance very similar to wild type. Strains that exhibited "tight" growth, with little visible perturbation at lower temperature but severely compromised or eliminated growth at 37° C, and consistent steady-state Taf4 or Taf5 protein abundance after heat shock were expanded from single colonies picked from 23° SC-His plates containing 5-FOA. Plasmids were recovered from these strains and transformed back into *E. coli* according to the procedure of M.V. Singh (Singh and Weil, 2002). Next, DNA sequencing was used to identify the changes in coding sequence and allow exclusion of any *taf4* or *taf5* alleles that had coding changes outside codons 250-350 of *TAF4* or 147-290 of *TAF5,* as well as alleles that coded for deletion, insertions, or frameshifts within these sequences. Finally, plasmids associated with all of the specified parameters were re-transformed into the yeast shuffling strains and confirmed to confer tight temperature conditional growth after shuffle with 5-FOA. Note that as performed, these genetic screens did not nearly approach saturation based on estimated library complexity, and can only identify recessive, loss-of-function, temperature conditional alleles. Three unique *taf4* alleles

were identified along with thirty unique *taf5* alleles. Only two of the thirty *taf5* alleles were isolated twice, consistent with the screen not having been performed to saturation.

RESULTS

Rap1 Interacts With Specific Regions of Taf12, Taf4, and Taf5

The investigation into Rap1-TFIID interaction had been ongoing within the lab for several years prior to my participation in the project. From the beginning of my involvement, my objective was characterization of Rap1 interaction with Tafs, although a former post-doc had gained some clues about this. Data indicated direct Rap1 binding by Taf4, Taf5, and Taf12 and my experimental data confirmed this observation. Another previous observation was that Rap1-Taf12 binding appeared to involve the highest affinity of interaction. For this particular reason, I first wanted to determine the site(s) within Taf12 that mediated interaction with Rap1. We decided to base our strategy on use of the Far Western interaction assay since it had proven useful up to that point for characterizing these interactions, and because it would lend itself to analysis of many Taf12 variants to help localize the Rap1 interaction domain(s). Even so, I designed a systematic family of N and C-terminal Taf12 truncations, to be expressed as fusions to GST, allowing the opportunity to examine the interactions using GST pulldown interaction assays as an experimental alternative. A second reason for preparing Taf12 variants as fusions to GST was that the tag might obviate problems with protein folding and solubility anticipated to arise with some truncation variants. All of the designed Taf12 variants, including the full-length protein and 19 truncations, were successfully

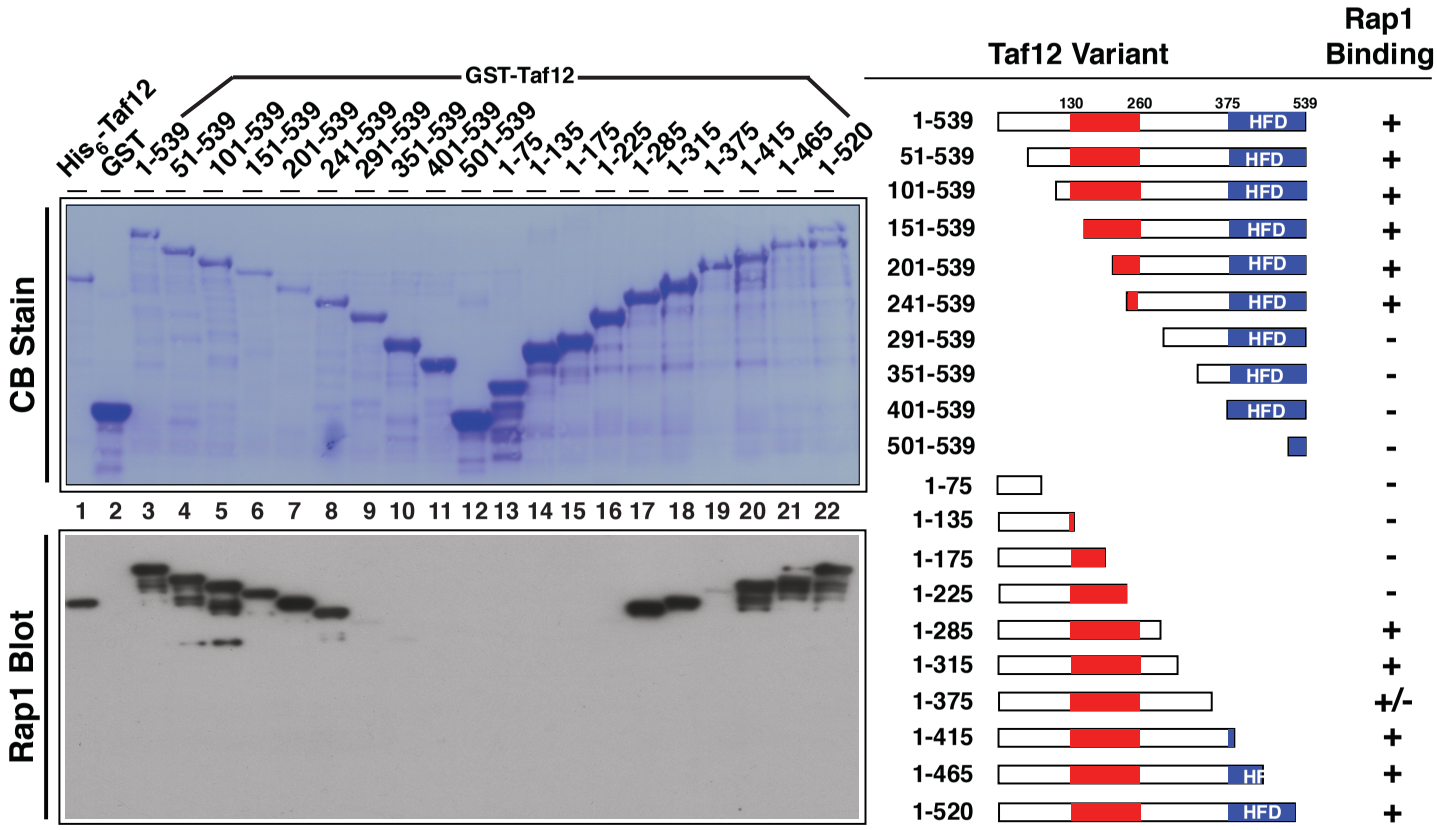


Figure 2.1: Mapping the Rap1 Binding Domain (RBD) of Taf12. Systematic serial truncations of Taf12 were designed to originate either from the N- or C-terminus, and were expressed and purified from *E. coli* as fusions to GST. Proteins were fractionated on two parallel SDS-PAGE gels, transfered, and either stained to indicate total protein content (upper panel; CB stain) or processed for Far Western as detailed in the methods section (bottom panel; Rap1 Blot). The structure and Rap1 binding properties of each Taf12 variant are summarized at the right. These data indicate that the most amino terminal portion of the Taf12 RBD is between residues 241 and 291, while the most carboxy-terminal portion of the RBD is between residues 225 and 285. Therefore the Taf12 RBD must reside between residues 241 and 285. Consistent with this an additional fusion protein including residues 201-315 appeared fully active for Rap1 binding in this assay (not shown). The Taf12 HFD is the only portion with strict sequence conservation from yeast to man, but the region indicated in red and containing at least some of the RBD is conserved in several fungal species. Data are adapted from Garbett et. al., 2007.

purified from the soluble extracts of bacteria expressing the recombinant proteins. These proteins were run on SDS-PAGE gels, alongside full-length Taf12 lacking the GST tag and recombinant GST alone, and analyzed by Far Western as described above. Data are shown in **Figure 2.1**. Consistent loading of the gel and equivalent transfer to the PVDF membrane are indicated as shown by a representative membrane stained with Coomassie Brilliant Blue. Rap1 interacted with Taf12 variants missing as many as 241 residues from the amino terminus but not with a variant missing 291 residues. Consistent with this observation, the amino terminal 285 residues of Taf12 were sufficient for interaction with Rap1, but the amino terminal 225 residues were not. Note that a variant containing the amino terminal 375 residues consistently failed to interact with Rap1 in this assay, the reason behind this lack of binding is not clear but I suspect this may be due to a limitation of protein refolding after SDS-PAGE and electroblotting, which is peculiar to that particular variant. Altogether, the data were reproducibly consistent with the existence of Rap1-Taf12 interaction occurring through the involvement of Taf12 residues between 241 and 285. Subsequent experiments showed that other Taf12 fragments containing residues ~201-315 were both necessary and sufficient for interaction with Rap1 in the Far Western assay (not shown).

After the characterization of the Taf12 Rap1 interaction domain(s), equivalent information about Taf4 RBD(s) was pursued next. In proceeding this way, I was considering that Rap1 interacts with Taf4, but also that Taf4 interacts with Taf12, and the heterodimer of the two proteins co-localize to multiple locations within the TFIID complex. Participation of both Taf4 and Taf12 in Rap1 binding was taken to indicate the possibility of redundancy of Taf contacts for Rap1 within TFIID. Thus information about

the interactions of each Taf with Rap1 might be required to adequately dissect Rap1-TFIID interactions. Having obtained information about Rap1 and Taf12 using the Far Western assay, this seemed the logical approach to extend to Rap1-Taf4 interaction. Full-length protein and truncations were again expressed as GST fusions in bacteria. Unfortunately, it was very difficult to obtain useable Taf4 variants. Expression and purification of Taf4 variants was plagued by low protein solubility and stability. Much of the difficulty was obviated by developing a new strain of bacteria compatible with coexpression of two different proteins from separate vectors, and with a third vector supplying rare tRNAs. As such Taf4 and variants were coexpressed along with Taf12 in BL21 DE3 Gold RIL Strept. Coexpression of Taf4 with Taf12 seemed to stimulate proper folding and solubility of Taf4 and also provide protection from proteolytic degradation. Preparative SDS-PAGE was included in the purification scheme to remove protein contaminants that might otherwise complicate the interpretation of the Far Western analyses. Gel purification also removed the full length Taf12 that was present in every preparation; this preparative step was intended to facilitate GST pulldown experiments to compare binding of Rap1 to full length GST Taf4 and derivatives. The results of the Far Western experiments are shown in **Figure 2.2**. Taf4 without a GST tag was used as a positive control because the family of Taf4 derivatives possessed this tag. GST was again used as a negative control and Taf3 was used as an additional negative control. Rap1 failed to interact with either recombinant Taf3 or GST. Rap1 interacted with Taf4 variants missing as many as 253 N-terminal residues, but not when 281 amino acids were removed. Amino acids 1-344 were sufficient for interaction with Rap1 but a fragment containing residues 1-316 did not interact with Rap1. Collectively, these results

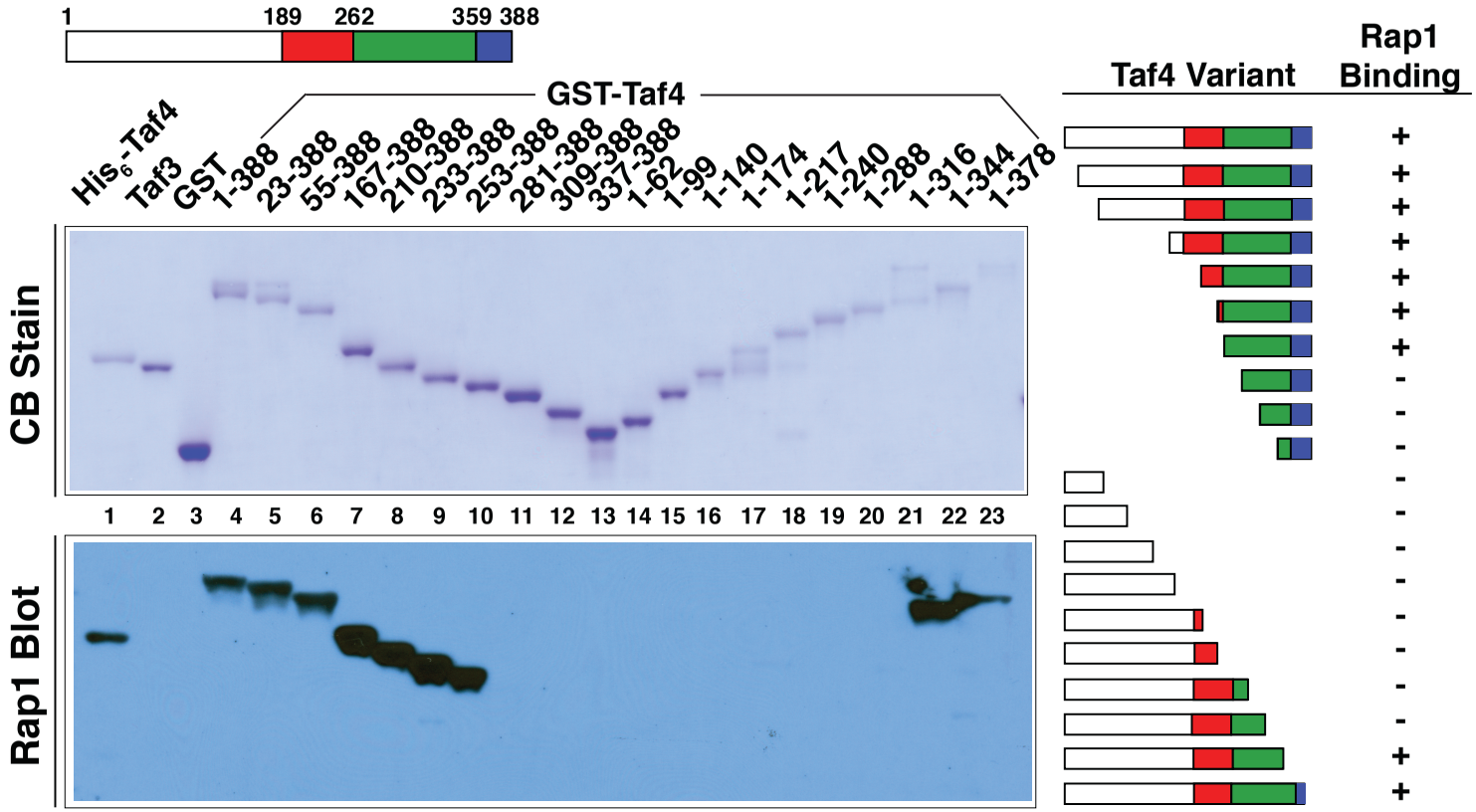


Figure 2.2: Mapping the Rap1 Binding Domain (RBD) of Taf4. Systematic serial truncations of Taf4 were designed to originate either from the N- or C-terminus, and were expressed and purified from *E. coli* as fusions to GST. Proteins were fractionated on two parallel SDS-PAGE gels, transfered, and either stained to indicate total protein content (upper panel; CB stain) or processed for Far Western as detailed in the methods section (bottom panel; Rap1 Blot). The structure and Rap1 binding properties of each Taf4 variant are summarized at the right. These data indicate that the most amino terminal portion of the Taf4 RBD is between residues 253 and 281, while the most carboxy-terminal portion of the RBD is between residues 316 and 344. Therefore the Taf4 RBD must reside between residues 253 and 344. Consistent with this an additional fusion protein including residues 253-344 appeared fully active for Rap1 binding in this assay (not shown). The Taf4 RBD (green colored region) is positioned in between the N- and C-terminal portions of the HFD (colored red and blue), and does not appear to contribute to HFD function in any way. Therefore, within the Taf4/Taf12 heterodimer, the mapped Taf4 RBD is in theory available for protein-protein interaction. Data are adapted from Layer et. al., 2010.

indicated that the Rap1 interaction domain(s) resided within residues 253 and 344 of Taf4. A fragment containing residues 253-344 fused to GST was sufficient for interaction with Rap1 in this assay (not shown).

Having successfully characterized interactions between Rap1 and both Taf12 and Taf4, it seemed logical to obtain equivalent information about Taf5. Once again, this was done in consideration of how Rap1 actually interacts with Tafs in the physiological context; that of the TFIID complex. And again, I was considering the possibility of redundancy existing in the mechanism of Rap1-Taf-TFIID interaction. Like with Taf4, Taf5 presented problems, but to a more significant extent. I found that while full length Taf5 and variants thereof are primarily insoluble when expressed in bacteria, some soluble protein could be obtained when expression was performed at lower growth temperatures and also if a solubility-enhancing tag like GST or MBP was appended to the protein. In this way I was able to obtain the systematic family of Taf5 truncations for use in the Far Western Assay. Unfortunately, when tested side by side with appropriate positive and negative controls, Rap1 failed to interact with any members of the Taf5 family. However, full length Taf5 expressed without the GST tag and purified under denaturing conditions was reproducibly capable of interacting with Rap1 within the same assays that failed to detect interaction with GST Taf5 and variants prepared from soluble proteins. I considered that it was the denaturing buffer conditions employed during purification of insoluble Taf5 that somehow rendered Taf5 competent to interact with Rap1 in the Far Western Assay. As such, denaturing preparations were made for the GST Taf5 family. Again, none of these proteins interacted with Rap1, although full length untagged Taf5 made and tested in parallel once again did interact with Rap1. This

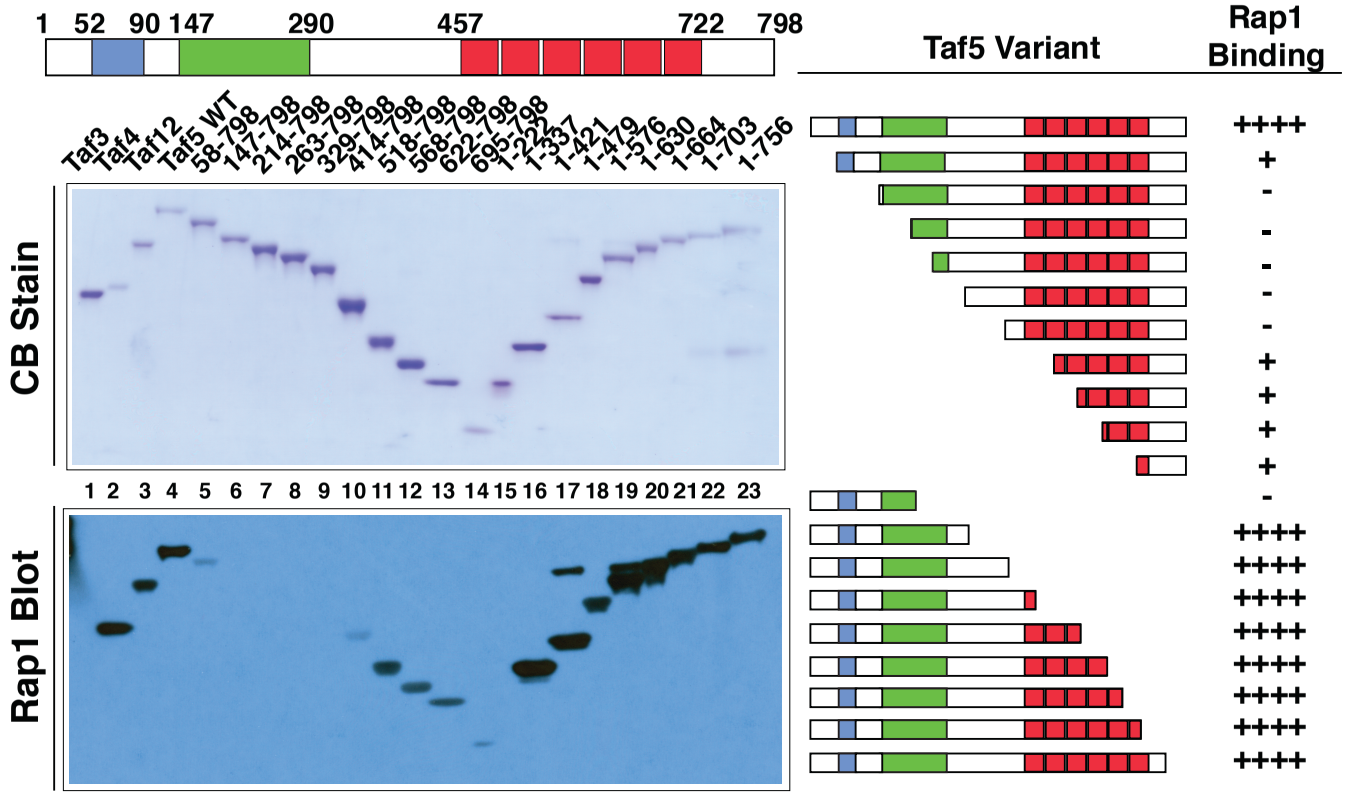


Figure 2.3: Mapping the Rap1 Binding Domain (RBD) of Taf5. Systematic serial truncations of Taf5 were designed to originate either from the N- or C-terminus, and were expressed and purified from *E. coli* as hexahistidine fusions. Proteins were fractionated on two parallel SDS-PAGE gels, transfered, and either stained to indicate total protein content (upper panel; CB Stain) or processed for Far Western as detailed in the methods section (bottom panel; Rap1 Blot). The structure and Rap1 binding properties of each Taf5 variant are summarized at the right. These data indicate that the most amino terminal portion of the Taf5RBD is between residues 1 and 58, while the most carboxy-terminal portion of the RBD is between residues 222 and 337. Therefore the Taf4 RBD must reside between residues 1 and 337. Consistent with this the variant including residues 1-337 appeared fully active for Rap1 binding in this assay (lane 16). The Taf5 RBD (blue and green colored region) contains both NTD1 and NTD2, and thus both probably contribute to interaction with Rap1. Notably, an N-terminal GST tag essentially eliminated the ability of Taf5 to interact with Rap1 in this assay, again consistent with the participation of the Taf5 N-terminus in the interaction. Note that the low-level binding of Taf5 variants in lanes 10-14 (5-20% of wild type) was reproducible, but the significance of this is doubtful and has not been pursued. Data are adapted from Layer et. al., 2010.

indicated that the amino terminal GST tag was interfering with the ability of Rap1 to interact with the protein(s). As such the entire Taf5 family was transferred from pBG101 to pET28A to allow expression without the GST tag. Rapid switching of expression vector backbone was simplified by the logical, streamlined LIC strategy used to create all Taf12, Taf4, Taf5, and Rap1 constructs, which itself would not have been possible without having had made my own preparation of Pfu polymerase. Expression and purification of Taf5 variants from pET28A under denaturing conditions resulted in the data shown in **Figure 2.3**. In these experiments Taf3 was again used as a negative control, and Taf4, Taf12, and full length Taf5 was used a positive control. Importantly, the similar quantity of Rap1 binding observed in these experiments indicated that Rap1 binds to Taf4, Taf5, and Taf12 with similar affinity. Binding of Rap1 to Taf5 was significantly reduced when as little as 58 amino terminal residues were removed, and eliminated when additional protein sequence was removed. On the other hand a Taf5 fragment containing just the amino terminal 337 residues was sufficient for interaction with Rap1 but the 222 amino terminal Taf5 residues were not. Notably there did appear to be some interaction between Rap1 and the Taf5 C-terminus, specifically within the WD repeats, but this only occurred when all residues amino-terminal of the WD repeats were removed. Therefore the results indicated overall that the amino terminus of Taf5 contains the Rap1 interaction domain(s). Recognize this is also consistent with the inability of full length Taf5 bearing a large amino terminal GST tag to interact with Rap1; it seems likely that placement of a large epitope tag at the Taf5 amino terminus reduces access for Rap1 to interact with Taf5 there. Note that the smallest Taf5 fragment that was able to interact with Rap1 contained residues 1-337; several other variants with

residues removed from either amino or carboxy terminal ends of this fragment did not interact with Rap1. Therefore the lowest amount of resolution of the Rap1 interaction domain(s) was available with Taf5.

Sequence Encoding the Rap1 Binding Domain of Taf12 is Non-Essential for Viability

Identification of Rap1 binding domain(s) contained within residues ~200-300 of the 539 residue Taf12 protein was a bit surprising. We expected that if truly critical for Rap1-TFIID interaction and thus Ribosomal Protein Gene (RPG) Transcription, then the DNA sequence encoding the responsible portion of Taf12 would be both necessary for cellular growth and encode amino acid sequence conserved in evolution. Available evidence from the literature suggested that this N-terminal Rap1-binding portion of Taf12 was dispensable for cellular growth (Moqtaderi et al., 1996b). Additionally, the only portion of yeast Taf12 with significant evolutionary conservation resides between ~400-539, which contains the histone fold domain that mediates dimerization with Taf4 (Thuault et al., 2002; Werten et al., 2002). Nevertheless, we decided to use our available systematic Taf12 truncation family and assess the complementation properties in yeast, to confirm the literature and see for ourselves if we could observe any significant growth defects associated with removal of the Rap1 interaction domain(s) in Taf12. Taf12 and derivatives were expressed from a single copy *HIS3*-marked plasmid under control of natural *TAF12* regulatory sequences, within my *TAF12* shuffling strain. As shown in **Figure 2.4**, all of these variants appeared to be recessive to wild type *TAF12*. When cells were grown in the presence of 5-FOA, which killed any cells containing *TAF12* on the *URA3*-marked plasmid, it was clear that removal of as many as 350 N-terminal amino

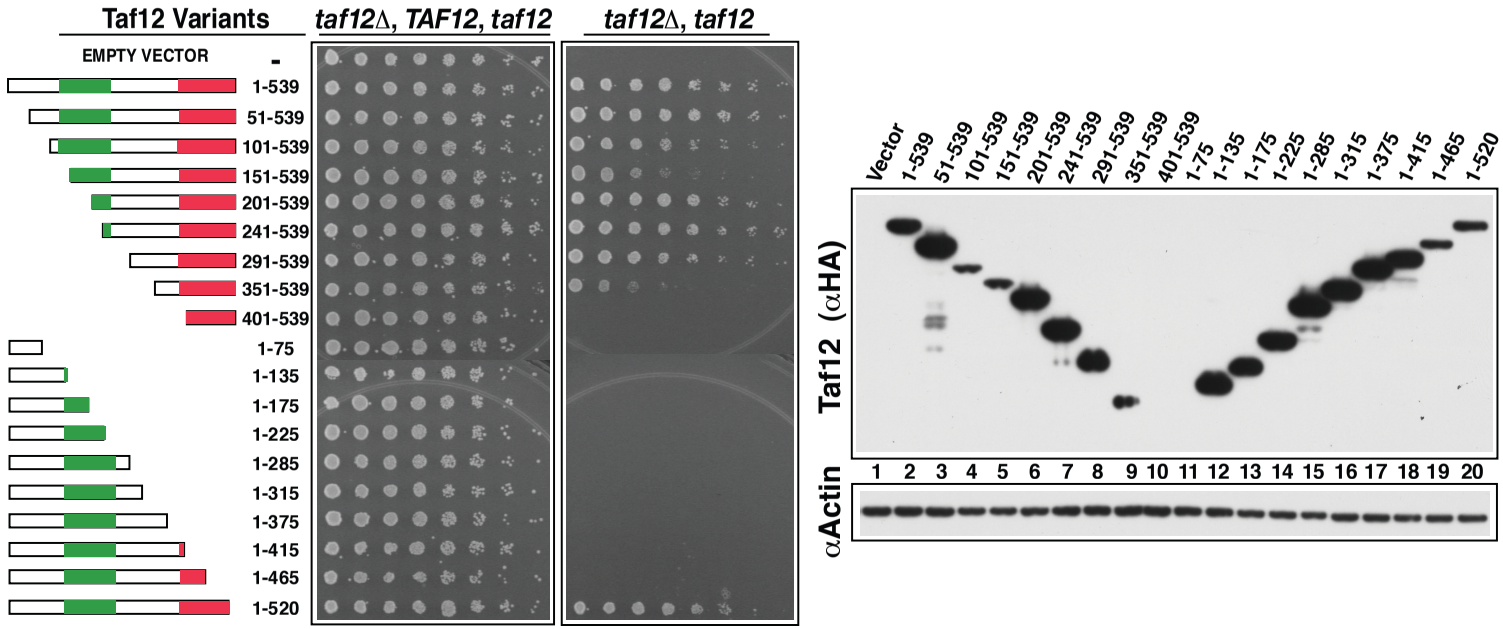


Figure 2.4: Mapping regions of Taf12 necessary for growth. The systematic serial truncations of *TAF12* were tested in a plasmid shuffle complementation assay for their ability to rescue the inviable phenotype of a chromosomal *TAF12* deletion. Serial 1:4 dilutions of each strain were spotted in replicate onto agar plates lacking or containing 1 mg/ml 5-FOA and incubated at several temperatures for several days. 30^{o} data are shown. Vector containing *TAF12* was able to complement as were most N-terminal truncation variants, with the exception of the 401-539 variant. On the other hand, only the 1-520 C-terminal truncation variant could complement, and this variant was inviable at 37^{o} (not shown). Collectively, this data indicates that the portion of Taf12 containing the RBD is completely dispensable for growth, and that the only critical portion is the C-terminal HFD colored in red. As indicated by the accompanying western blot, almost all of the variants were present at or near wild type steady state abundance, with the exception of the 401-539 and 1-75 variants (lanes 10 and 11), a result also in agreement with the idea that most of Taf12 is not needed for cell growth, with the HFD being the only critical part of the protein. The significance of the Rap1 binding domain in Taf12 to RPG transcription is brought into question by these results. For this reason, little additional effort was invested in characterizing the Taf12 RBD. Data are adapted from Layer et. al., 2010.

acids was compatible with growth. Consistent with the requirement of Taf12 for viability (see empty vector strain) and importance of the C-terminus, any significant alteration of this region resulted in a loss of viability. Notably, removing just 19 C-terminal residues, while compatible with viability, conferred temperature conditional growth (not shown). Two of the amino-terminal truncations, containing residues 151-539 or 351-539, showed a consistent slow-growth phenotype on 5-FOA. However, the loading adjusted immunoblot shown below the growth data indicated that both of these variants were of reduced steady-state abundance compared to wild type. Thus although the 351-539 variant is missing the entire Rap1 binding domain, its slow growth phenotype is not expressly due to this deficiency. I attempted to clarify this issue by subcloning the Taf12 family into another yeast vector intended to direct high-level expression. Neither the complementation patterns nor the pattern of Taf12 variant steady state abundance relative to wild type changed from the prior result, so this did not reconcile whether the growth deficiency in the 351-539 variant is strictly due to absence of the Rap1 interaction domain (not shown). Note that four Taf12 variants that did not support viability but were capable of interaction with Rap1 in the Far Western accumulated to levels greater than or equal to wild type (lanes 15-18). Therefore the available information confirmed what the literature already reported, that the RBD within Taf12 is dispensable for viability (Moqtaderi et al., 1996b).

Sequences Encoding the Rap1 Binding Domains of Taf4 and Taf5 Are Essential For Viability

Given the lack of a discernible phenotype with any of the Taf12 variants, it was necessary to proceed with in vivo analyses testing growth requirements for Taf4 and Taf5 Rap1 interaction domains. Experience had shown us that only severe loss of growth phenotypes typically preceded observance of affects on RPG transcription; it seemed unlikely we would see this phenomenon in the case of the Taf12 Rap1 interaction domain. As was the case with Taf12 though, the initial design of the Taf4 and Taf5 deletion families facilitated their movement into yeast expression vectors for plasmid shuffle complementation assays. **Figure 2.5** shows the results of complementation tests and a western blot measuring steady state Taf4 variant protein abundance. Recall that the Rap1 interaction domain in Taf4 was localized to residues 253-344. Note that removal of residues between 210 and 233 resulted in loss of viability. Removal of either the first 21 or 91 residues caused a slow growth phenotype, but the steady state levels of the corresponding proteins were barely detectable. On that note, there was a wide range of steady state protein levels compared to wild type. The 210-388 form, which supported

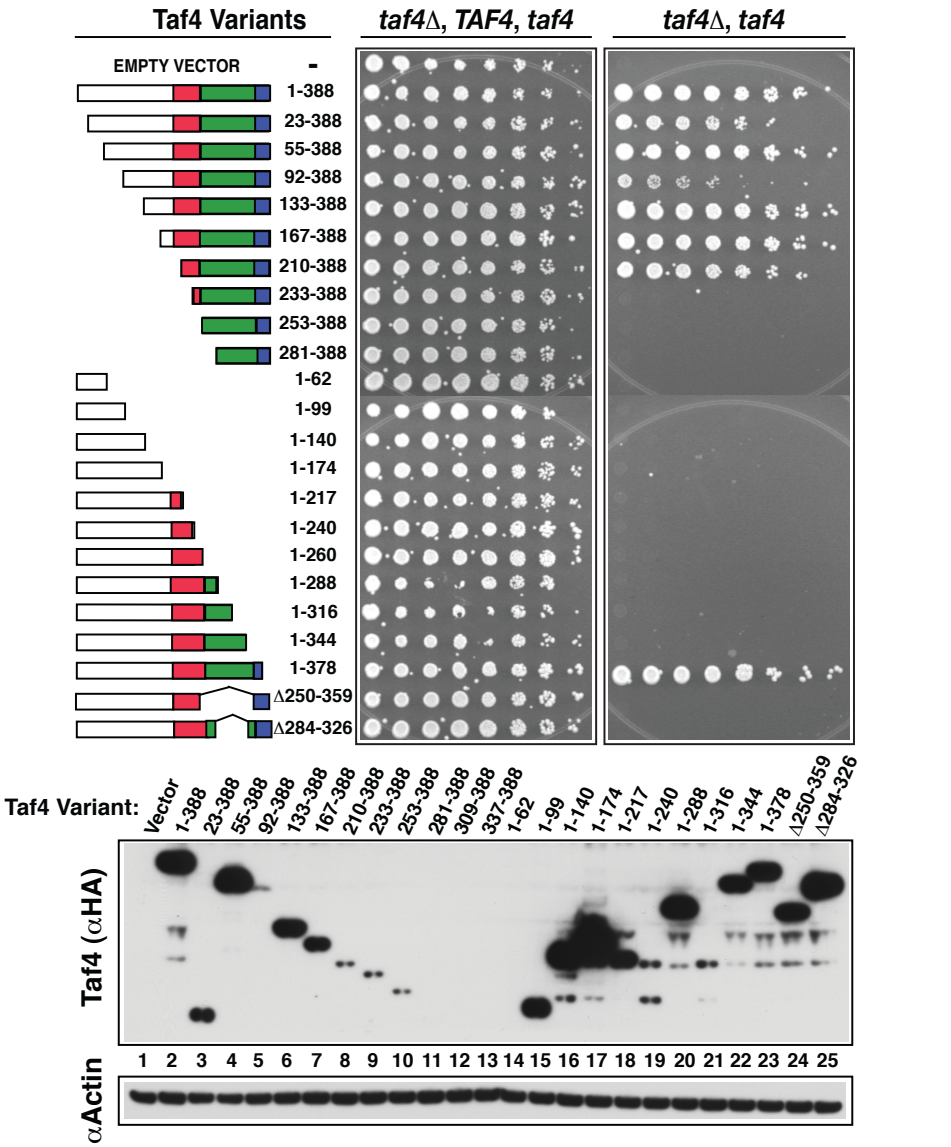


Figure 2.5: Mapping regions of Taf4 necessary for growth. The systematic serial truncations of *TAF4* were tested in a plasmid shuffle complementation assay exactly as was done for *TAF12*. Vector containing *TAF4* was able to complement as were most N-terminal truncation variants, with the exception of variants with the N-terminal 232 or more residues absent. On the other hand, only the 1-378 C-terminal truncation variant could complement, and this variant was inviable at 37^{o} (not shown). Collectively, this data indicates that all portions of the Taf4 HFD are completely needed for growth, and do not address the role of the RBD in cell growth. Creation of the two internal deletion mutants Δ250-359 and Δ284-326, which lack some or all of the mapped RBD, demonstrate the requirement of the RBD for cell growth. As indicated by the accompanying western blot, variants deviated widely from wild type steady state abundance, but abundance did not necessarily correspond to complementation profile. The Δ250-359 and Δ284-326 variants were present at or greater than wild type levels. Data are adapted from Layer et. al., 2010.

viability at the wild type rate, was also present at reduced levels. Therefore in the case of Taf4 it was somewhat difficult to reconcile steady state protein abundance with complementation patterns. At the C-terminus, removal of just 10 residues conferred a temperature conditional phenotype (1-378, data not shown); additional removal was not compatible with life (see 1-344). Importantly the 1-344 variant protein accumulated to approximately wild type levels, consistent with a critical function of residues 345-388, which are not involved in Rap1 interaction. Because amino acids flanking the Rap1interaction domain(s) at both the amino and carboxy terminus were necessary for viability, additional Taf4 variants had to be generated to specifically address the requirements of those particular Taf4 residues for viability. Site directed mutagenesis using the gene SOEing technique allowed the removal of residues 250-350 while leaving Taf4 otherwise intact. To create a Taf4 variant with a smaller deletion removing residues 284-326, *TAF4* was subcloned as a BamHI-SalI fragment into p413 *TAF4* HA^3NLS that had been cut with BamHI and XhoI thereby destroying the XhoI site in the vector backbone. This construct was then cut with XhoI and subsequently made blunt using the Klenow fragment, then digested with ZraI. The vector backbone was next gel purifed and re-ligated at low DNA concentration. This resulted in seamless deletion of residues 284-326 with preservation of the reading frame. When tested in complementation assays neither the form missing residues 250-350, nor the variant missing 284-326 could support viability, even though these proteins accumulated to levels equal to or greater than that of the wild type protein. This result indicated that like the residues flanking the Rap1 interaction domain, these amino acids are also necessary for viability. Therefore the

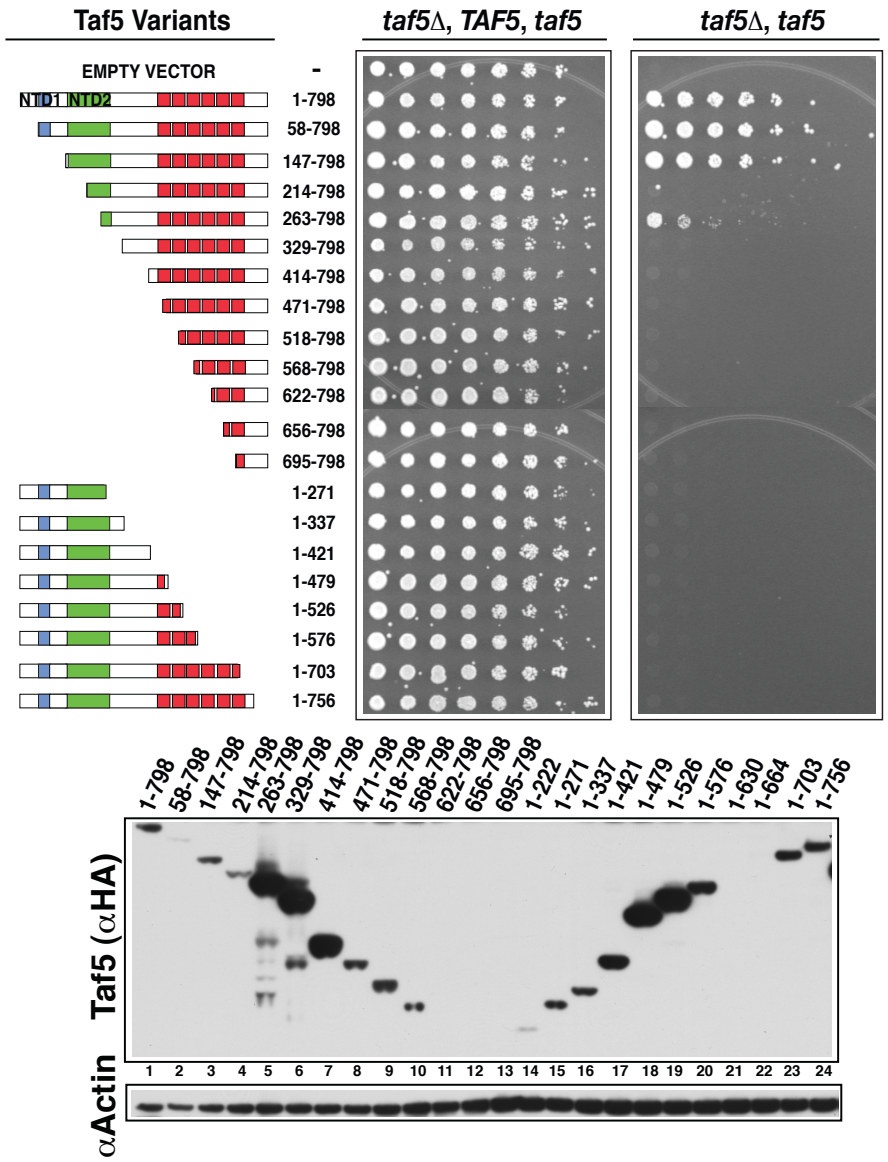


Figure 2.6: Mapping regions of Taf5 necessary for growth. The systematic serial truncations of *TAF5* were tested in a plasmid shuffle complementation assay exactly as was done for *TAF12* and *TAF4*. Vector containing *TAF5* was able to complement as were a few N-terminal truncation variants. The 263-798 variant, but not 214-798, was able to complement. Note however the large excess of the 263-798 variant demonstrated in the accompanying western blot, which may account for the complementation properties of the 263-798 variant, although it cannot support viability at 37^{o} (not shown). Note that there are three lysine residues between 214 and 263; these residues may play a role in Taf5 stability, possibly through ubiquitin- and/or SUMO-dependent pathways. On the other hand, none of the C-terminal truncation variants could complement. Collectively, this data indicates that the N-terminal NTD1 and NTD2 portions of the Taf5 RBD, colored blue and green, respectively, are important for cell growth, as are the non-RBD C-terminal WD repeats. Data are adapted from Layer et. al., 2010.

portion of Taf4 responsible for interaction with Rap1 in vitro indeed plays a critical role in cellular growth.

In parallel the complementation properties of the Taf5 truncation family were tested in yeast (**Figure 2.6**). Removal of as many as 146 amino terminal residues was compatible with normal growth. As a portion of this sequence was found to be important for interaction with Rap1, this result indicates that at least a portion of the Taf5 Rap1 interaction domain is not needed for cellular growth. However, additional removal of amino terminal Taf5 domains up to residue 213 was strictly incompatible with viability.Strangely, a variant lacking as many as 263 amino terminal residues consistently grew, albeit at a reduced rate and in a temperature sensitive manner. Note that this 263-798 variant exhibited steady state protein abundance 5-10 fold greater than wild type; this characteristic may be related to the ability to complement since this form is always effectively overexpressed, perhaps providing a measure of dosage compensation. Consistent with the important nature of amino terminal Taf5 residues, both the 329-798 and 414-798 variants, which are also present at elevated steady state levels, are completely unable to support growth. These two variants are either nearly lacking or completely missing the mapped Rap1 interaction domain(s). With respect to the Taf5 C-terminus, it appears that all or nearly all of this portion of Taf5 is required for growth. The least disturbed variant, missing residues 757-798, can only grow at low temperatures and very slowly at that (not shown).

To better characterize the Taf5 amino terminus, two additional variants were created using site directed gene SOEing. The hope was that these variants would better explain the relationship and relevance of the so-called N-terminal domains 1 and 2

(NTD1 and 2). These domains represent the most conserved residues in the Taf5 amino terminus and as such it was reasonable to consider them as of primary significance with respect to the observed complementation properties, and for interaction with Rap1. Because NTD1 appeared dispensable as shown using the systematic truncation family and NTD2 seemed to play a bigger role, I designed variants missing all of NTD2 or all of NTD2 but also NTD1. The complementation properties of these variants are shown in **Figure 2.7**. Deletion of residues 147-290 conferred a temperature conditional growth phenotype, where this strain was unable to grow at temperatures above 30^{o} C. Simultaneous removal of NTD2 and also NTD1 worsened this phenotype, causing a constitutive severe slow growth phenotype and this strain was also unable to grow at temperatures above 23-25^{o} C. Note that both variants were of higher steady state abundance than WT, particularly the double deletant, which may alleviate the deficiencies associated with these variants, by driving protein-protein interactions whose affinity was reduced by the deletions. Note that the worsened phenotype is considered a synthetic genetic interaction, often found when distinct proteins or protein domains contribute to the same molecular process. Since the entire Taf5 amino terminus was needed for effective interaction with Rap1, this new finding supported the idea that these domains in Taf5 really do contribute to interaction with Rap1 in vivo, and also suggested the importance of this interaction to cellular growth.

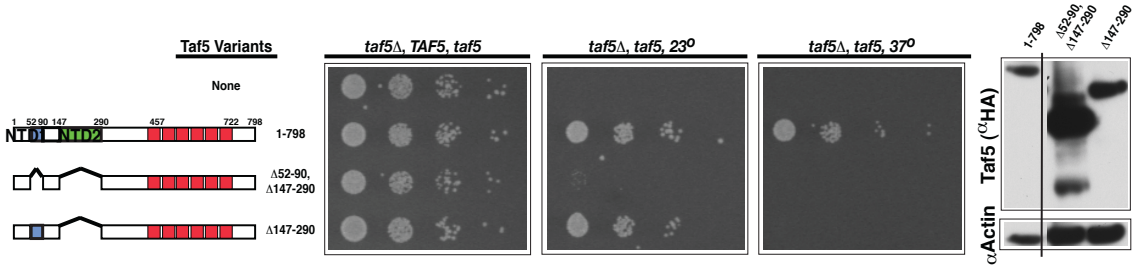


Figure 2.7: Mapping contribution of Taf5 N-terminal domains 1 and 2 to growth. The internally deleted *taf5* variants were tested in a plasmid shuffle complementation assay, as previously. Growth of the ΔNTD1 ΔNTD2 variant was barely visible, and could be seen only at room temperature. A higher amount of visible colonies was apparent after extended incubation, but this variant grows at 5-10% of the wild type rate (not shown). On the other hand, the ΔNTD2 variant could grow as wild type at room temperature, but not at all at 37°. These results are consistent with both Taf5 domains making important contributions to Rap1-TFIID interaction and RPG transcription. The accompanying western blot indicates the high expression level of the deletants relative to wild type. Data are cropped and spliced from the same image of the same experiment, as indicated by the dividing line.

Rap1 Binding Domains of Taf4 and Taf5 Do Not Contribute to Taf-Taf Interaction

An alternative methodology was sought to characterize interactions between Rap1 and Tafs, with an eye for obtaining both qualitative and quantitative information. The yeast two-hybrid technique fulfilled these criteria. Unfortunately Rap1 did not interact with any Tafs when fused to the Gal4 DBD, where Tafs were fused to the Gal4 AD. A possible explanation for the lack of interaction with Tafs was that the Gal4 DBD Rap1 fusion protein was unstable or expressed at very low levels. This undesirable feature might be further complicated by, and contribute to, a reduced availability for participation in activation of two-hybrid reporter genes, particularly since the fusion protein presumably maintained its ability to bind to the many natural Rap1 binding sites distributed throughout the genome. This is not a new observation or suggestion (Zhao et al., 2006). Nevertheless, the Gal4 DBD Rap1 fusion was active for interaction with Gal4 AD Rif2, a protein known to interact with the C-terminus of Rap1 to mediate telomeric silencing (positive control, not shown). The Rap1-Rif2 interaction might simply be of a higher affinity than Rap1-Taf interactions and thus activation of reporter genes is possible with the Rap1-Rif2 two-hybrid pair even though the fusion proteins are poorly expressed.

To attempt to address this issue the reciprocal experiment was performed with Tafs fused to the Gal4 DBD and Rap1 fused to the Gal4 AD. Unfortunately, the Gal4 AD Rap1 fusion appeared to have a toxic affect on yeast cell growth, because no transformants could be obtained that contained the pACT2.2 *RAP1* construct. Further attempts would have necessitated creation of additional constructs, none of which promised to provide useful information, so at this stage two-hybrid experiments to study Rap1-Taf interactions were abandoned.

To salvage some useful information after the investment of time made in establishing reagents for this procedure, the requirement for the Taf4 and Taf5 domains of interest for interaction with other Tafs was pursued. Gal4 DBD fusion of Taf4, Taf5, and deletion derivatives were created using pGBDU C1. These were transformed into the reporter strain along with a vector encoding TBP or one of the fourteen Tafs, so that interaction with every single stoichiometric component of the TFIID complex was examined. Using an *ADE2* reporter gene as the readout for two-hybrid interaction, I identified interactions between Taf4 with Taf9 and Taf10 (**Figure 2.8**). The pattern of interaction did not change when residues 250-350 were removed from Taf4. Gal4 DBD Taf5 interacted with Gal4 AD Taf4, Taf6, Taf9, and Taf10. Again, this pattern of interaction did not change when both NTD1 and NTD2 were deleted from Taf5.

The lack of interaction between Taf4 and Taf12 was unexpected since it is widely accepted that these proteins directly bind in a strong manner. The observed of interaction between Taf5 and Taf4, but not the reciprocal interaction, was also troubling. There are two additional reports in the literature concerning systematic two-hybrid interactions using *TAF* genes from yeast or plants (Lawit et al., 2007; Yatherajam et al., 2003). There

are considerable technical differences between these two studies and my own experiments, not the least of which include using *TAF* genes from plants instead of yeast, or use of a different (and less selective) reporter gene to indicate two-hybrid interactions. Thus published work with yeast Tafs identified interactions between Gal4 DBD-Taf4 and Gal4 AD-Taf10 and Gal4 AD-Taf12, but not Gal4 AD -Taf9; work with plant Tafs found the exact same pattern of interactions. The study using yeast Tafs identified interactions between Gal4 DBD-Taf5 with Gal4 AD-Taf4, Gal4 AD-Taf10, and Gal4 AD-Taf12, while plant Gal4 DBD-Taf5 interacted with Gal4 AD-Taf4, Gal4 AD-Taf5, Gal4 AD-Taf8, and Gal4 AD-Taf9. Thus a commonality between the different results is the observance of interactions between Gal4 DBD-Taf5 and Gal4 AD -Taf4, but not the reciprocal interaction. This and other discrepancies might be sorted out using combinations of both N- and C-terminal fusions of Tafs to the Gal4 DBD and AD, in case the exclusive N-terminal location of epitope tags is site-specifically interfering with certain interactions. Until this and systematic use of reporter genes is implemented, it is difficult to extrapolate results of two-hybrid experiments from one report to another. For my purposes, the strongest conclusion that could be made from my studies is that the regions of Taf4 and Taf5 proposed to be important for interaction with Rap1 do not appear to influence Taf-Taf interactions identified by directed two-hybrid screening. With the caveat of false negatives accompanying use of this assay, we could proceed with some extra confidence in additional mutational analyses of DNA encoding these Taf domains, without concern about gross perturbation of TFIID structure caused by disruption of Taf-Taf interactions.

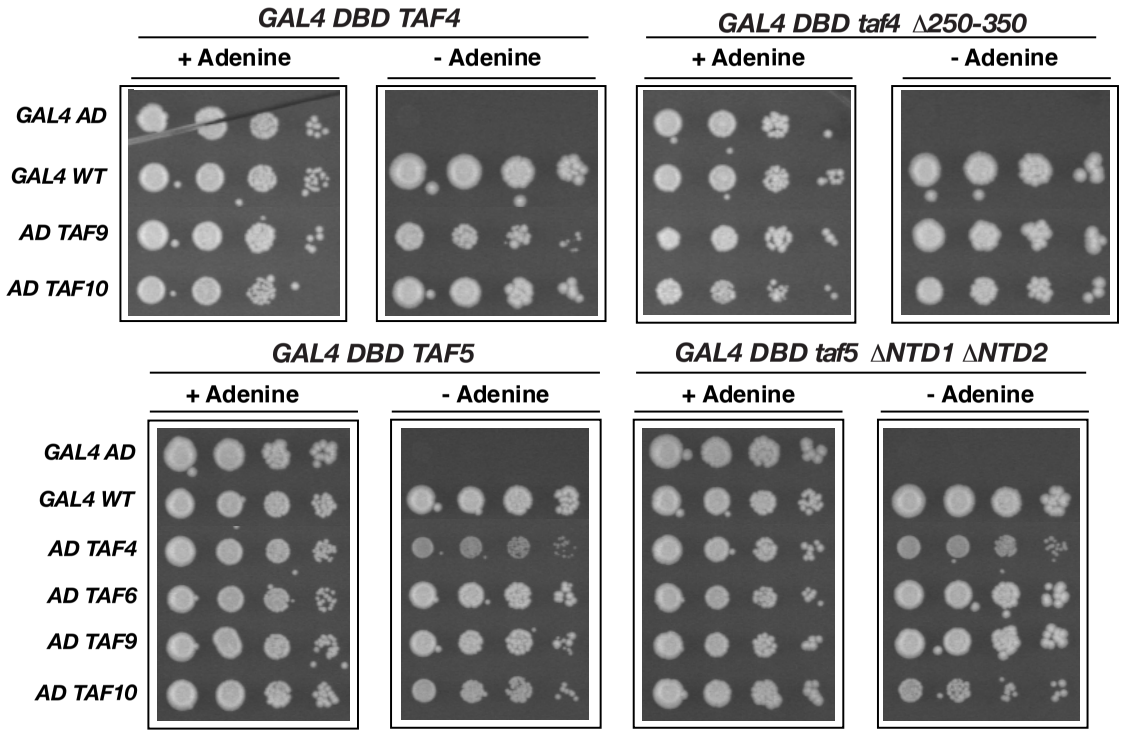


Figure 2.8: Yeast two-hybrid assays comparing interaction pattern of Gal4-DBD-Taf4 or Gal4-DBD-Taf5 fusions or equivalent RBD-deleted variants with various Gal4-AD-Taf binding partners. Activation of the chimeric *GAL7-ADE2* reporter gene in strain pJ69-4a allowed growth on media lacking adenine. Media also lacked uracil and leucine to select for plasmids bearing DBD- and AD-fusions, respectively. Gal4-DBD-Taf4 interacted with AD-Taf9 and AD-Taf10. Surprisingly, DBD-Taf4 did not interact with AD-Taf12, even though the two form a very stable heterodimer in vivo and in vitro. Interaction patterns did not change when the Taf4 RBD was omitted. Gal4 DBD-Taf5 interacted with AD-Taf4, 6, 9, and 10. Thus the Taf5-Taf4 interaction was not observable in a reciprocal manner. Again, the pattern of interaction did not change when the Taf5 RBD was omitted. These results indicated that the Taf-Taf interactions detected by two-hybrid interaction studies are unaffected by deletion of the Taf4 and Taf5 RBDs, suggesting any mutations introduced into these *TAF*'s RBD-encoding sequence should not affect Taf-Taf interaction or TFIID stability. Unfortunately, Taf-Rap1 interactions could not be scored in this assay. Data are adapted from Layer et. al., 2010.

Point Mutagenesis of Sequences Encoding Rap1 Binding Domains Confers Temperature Conditional Growth

Having mapped Rap1 binding domains in Taf4, Taf5, and Taf12, and demonstrating that the responsible portions of Taf4 and Taf5 are necessary for cellular growth, I had a good basis for attempting isolation of mutant alleles that affect TFIID function by way of an affect on the portions of interest. Because an obvious growth phenotype occurred upon removal of the Rap1 interaction domain(s) in Taf4 and Taf5 but not with Taf12, only Taf4 and Taf5 were pursued further. The objective of this part of the study was to attempt to identify mutants that show growth defects, but that more importantly display a defect in RPG transcription. I wanted to rule out indirect affects to the best of my ability and as such sought to avoid the use of large deletions in building my model of the role of Rap1-Taf-TFIID interaction in RPG transcription. Thus point mutants were most desirable. Again, there was little basis for designing mutants on the information about structural features, since nothing was known at the time I started my Taf4/Taf5 structure-function studies. Moreover, our experience has shown that multiple amino acid substitutions are often needed to provide enough disturbance of Taf function so as to allow detection of a growth phenotype. Both of these limitations could be overcome by partial randomization of DNA sequence encoding the Taf domains of interest. Successful creation of *taf* point mutants was achieved by a combination of judicious subcloning strategy, PCR-mediated mutagenesis under conditions of low-fidelity amplification, high quality library construction, and efficient growth-based library screening available through use of yeast genetics as detailed above. There were several criteria set forth prior to screening for mutants. The purpose of these criteria was

to eliminate indirect affects as much as possible in judging the suitability of particular unique mutants. Therefore I was able to focus on the contribution of individual mapped Rap1 interaction domains in Taf4 and Taf5, and truly and selectively test the contribution of those domains to growth and RPG transcription.

Mutagenesis of *TAF4* was successful in that three unique mutant alleles demonstrating temperature conditional growth were isolated (**Figure 2.9**). Alleles were named with a number designation corresponding to the order in which they were isolated and finally selected. Each allele encoded a protein exhibiting steady state abundance and mobility very similar to wild type, as expected given that this was factored into my selection of mutant alleles. All three alleles encoded multiple amino acid substitutions, from three to six amino acid substitutions coded per allele. Interestingly the allele with the strongest loss of growth phenotype, *taf4-219*, contained the fewest coding changes. The identification of multiple mutants suggested firstly that this region alone is responsible for providing a critical Taf function, and also again suggested that individual amino acids within the putative Rap1 binding region of Taf4 are important for cellular growth, although the occurrence of multiple substitutions per allele complicated more precise interpretation. Since all three alleles conferred a loss-of-growth phenotype, but were recessive to wild type, it was inferred that the encoded proteins are defective for a specific part of overall protein function.

While this goal of the project concerning Taf4 met with success, by comparison mutagenesis of *TAF5* was a smash hit. In total 30 unique recessive mutant alleles were isolated that conferred a temperature conditional loss of growth phenotype. All isolated alleles except one contained multiple amino acid substitutions. For practical reasons

further use of all identified *taf5* mutants was deemed excessive; the experiments that will be described below in detail were limited to four particular mutants (**Figure 2.9**) that contained either one, five, or eight different coding changes. Again, each Taf5 protein was of similar mobility and abundance relative to wild type, as mandated by my selection criteria. Growth characteristics were generally more desirable with "tighter" phenotypes evident than seen with *taf4* mutants; this is attributed to a larger set of *taf5* mutants to choose from that allowed the 'best' mutants to be used. The specific reasons behind the higher success rate of *TAF5* mutagenesis are discussed below. In any case the randomization experiments involving DNA encoding Taf4 and Taf5 Rap1 binding domains supported the importance of these domains to cellular growth and showed that individual amino acids contribute to RBD function. The occurrence of strong growth phenotypes is consistent with, but does not prove, a defect in RPG transcription in mutant cells since RPG transcription is a limiting factor for cellular growth capacity. The next logical step was to examine levels of RPG transcripts in RNA from wild type and mutant cultures grown at permissive and non-permissive growth temperatures.

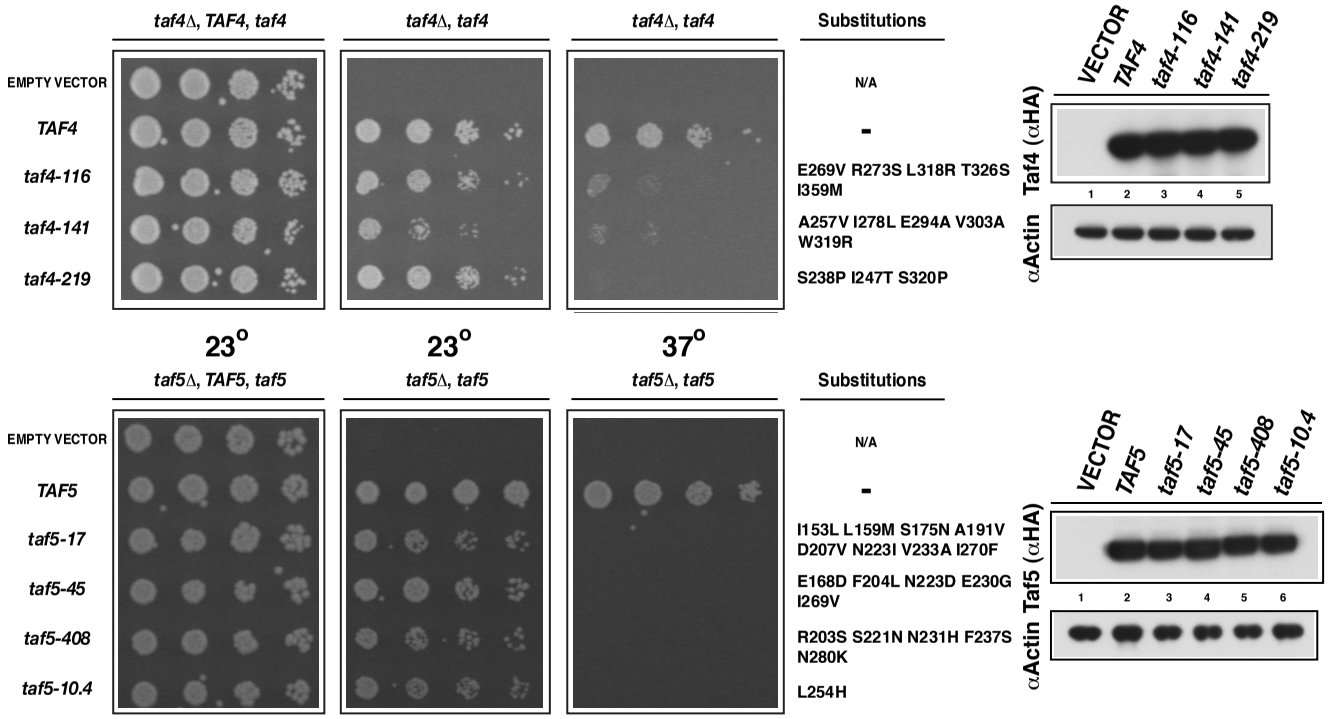


Figure 2.9: *taf4* and *taf5* temperature conditional mutants. Sequence encoding the RBDs of Taf4 and Taf5 were mutagenized by error-prone PCR, and amplified DNA was used to create libraries of mutant alleles. A genetic screen was performed using each library with the corresponding shuffling strain, with selection for temperature conditional growth. After filtering identified alleles by a number of criteria described in the Methods section, 3 *taf4* ts^+ and 30 *taf5* ts^+ mutant alleles were selected for final use. All three *taf4* and four *taf5* alleles were retransformed and subjected to plasmid shuffle at various temperatures, as shown. Almost all mutant alleles encoded multiple amino acid substitutions. *taf4* mutants grew slowly at 37^o whereas *taf5* mutants could not grow at that temperature. As mandated by the selection criteria, all alleles encoded proteins that were stable at all temperatures and present at levels approximately equal to wild type. Western blot data from 30^o -grown cells are shown. Data are adapted from Layer et. al., 2010.

Discussion

The objectives of this phase of my thesis research were to identify and characterize Rap1 binding domains in TFIID that are potentially shared or distributed between three Tafs, and were predicated on the concept that any such domains make a critical contribution to cellular growth in part via participation in RPG transcription. Only by uncovering growth phenotypes associated with alteration of these Taf domains could I expect to successfully and meaningfully characterize the function of those domains. To restate in other terms, those loss-of-growth phenotypes associated with Taf mutants should be coincident with the molecular phenotype of reduced RPG transcription, the readout that we needed to ultimately measure. It would be impractical to first measure for RPG transcriptional phenotypes, so the correlation (albeit not strict or exclusive to RPGs) of RPG transcriptional defects with growth defects was mandatory for execution of this project.

The first Rap1 binding domain to be characterized was that of Taf12. This was roughly mapped to residues ~200-315 of the 539 amino acid protein. The best characterized feature of Taf12 is its C-terminal histone fold domain that participates in direct interaction with Taf4. There is very little information about the Taf12 N-terminus, probably because in higher eukaryotes Taf12 is a much smaller protein, essentially containing just the HFD. In addition the yeast-specific N-terminus has enjoyed little investigation because it was initially shown to be dispensable for viability, (Moqtaderi et al., 1996b) in contrast to essential yeast Taf domains that have been characterized in any detail such as HFDs of Taf4 and Taf12, Taf1 domains, or WD repeats of Taf5. The lack of evolutionary conservation hinders study of the Taf12 N-terminus even further. The

only noticeable feature(s) within the Taf12 RBD are several patches of consecutive glutamine residues. Patches of contiguous glutamine residues are found in the metazoan Taf4 amino terminus, but not in the yeast Taf4 amino terminus. Incidentally, the glutamine-rich Taf4 amino terminus has been shown to participate in coregulatory protein-protein interactions with metazoan activators such as Sp1. It is reasonable to propose then that the yeast Taf12 amino terminus fulfills roles normally tasked to Taf4 in higher organisms, an idea made all the more reasonable since Taf4 and Taf12 bind directly and are hence tightly colocalized within the TFIID complex. Perhaps this concept could include Taf12-mediated coactivation for Rap1 by TFIID in yeast. It remains difficult to test at present given the lack of Taf12 mutants exhibiting strong growth phenotypes.

The Rap1 binding domain in Taf4 was found to reside in a portion of the protein that interrupts the bipartite Taf4 histone fold domain. This domain is believed to provide a function independent of the Taf4 interaction with Taf12. For example Taf4 variants missing this domain were previously shown to copurify with Taf12 when coexpressed in *E. coli* (Thuault et al., 2002). While this portion of Taf4 is conserved, the identity of amino acids is rather loosely defined, with the major characteristic being a relatively hydrophilic overall character. Consistent with these features, this 'linker' region was proposed to be unstructured and solvent exposed since it did not show up in the x-ray crystal structure of the Taf4-Taf12 heterodimer (Werten et al., 2002). If these features are of significance in vivo in the context of TFIID, then it leaves open the possibility that this domain is available for interaction with activators such as Rap1. And again noting some of the differences between yeast Tafs and metazoan Tafs, the yeast Taf4 protein is

relatively small and missing several domains found in the metazoan protein. These metazoan Taf4 domains have been implicated in direct activator interaction. Maybe yeast Taf4 is still an important coregulatory contact, but it must utilize the identified domain to carry out this role. This clue that there are different protein features present or absent depending on the organism is perhaps indicating to us that the composite and/or collective features of the Taf4-Taf12 subcomplex are the more relevant context within which to visualize overall coregulatory responsibility. Distribution of function between either one or multiple TFIID subunits is a phenomenon that was first described for the chromatin/acetyllysine binding function of TFIID, found either within just metazoan Taf1 or shared between yeast Taf1 in conjunction with yeast Bromodomain Factor 1; Bdf1 (Matangkasombut et al., 2000).

The characterization of *TAF5* described here is the most extensive in the literature to date. The function of most Taf5 domains is proposed to be related to Taf-Taf interaction and this is supported by coexpression/copurification studies and immuno-electron microscopy data. It seems likely that the WD repeats function predominantly or wholly to interact with Taf histone fold heterodimer pairs like Taf4/12 and Taf6/9. This hypothesis is supported by the yeast two-hybrid data reported here and elsewhere, although there are caveats to this type of experiment that will be discussed at greater length below. Perhaps less attention has been paid to the Taf5 amino terminus and this may be related to several reports in the literature. First, a study to characterize the highly related *S. pombe* Taf5 protein indicated that the N terminus is dispensable for viability (Yamamoto et al., 1997) although subsequently it has become clear that there is a flaw with this study. Second, our immuno-electron microscopy studies and biochemical

characterization of TFIID showed that two moles of Taf5 are present per mole of TFIID and that the Taf5 N-termini colocalize in the TFIID EM structure (Leurent et al., 2004; Sanders et al., 2002a). It was noted that the NTD1 domain of most species Taf5 proteins have significant homology to the Lis-1 protein, which forms an obligate homodimer (Kim et al., 2004), and this fact along with the other observations were used to propose that Taf5 dimerizes in TFIID and this is mediated by the N-terminus. Thus the only proposed function of the Taf5 N-terminus was, like the C-terminus, to facilitate Taf-Taf interaction and TFIID structure and/or integrity. This prediction has not been borne out by any experimental data in the literature. Firstly, neither the systematic two-hybrid analyses of yeast Taf interactions nor my equivalent studies identified interaction of DBD Taf5 with AD Taf5, as would be expected if the protein formed a homodimer (although plant Taf5 has been reported to self-interact in two hybrid assays, (Lawit et al., 2007; Yatherajam et al., 2003). Second, a form of Taf5 missing the N-terminus co-immunoprecipitated with the full length protein in yeast expressing both forms of Taf5, indicating that N-terminal to N-terminal contacts are not required for multiple molecules of Taf5 to be contained within the TFIID or SAGA complexes (Romier et al., 2007). Third, biochemical analysis of the recombinant Taf5 N-terminus indicated a lack of oligomerization in gel filtration and analytical density-gradient ultracentrifugation experiments (Bhattacharya et al., 2007; Romier et al., 2007). Fourth, during the course of my studies the x-ray structures were determined for the human, yeast, and *E. cuniculii* (a parasitic microorganism) Taf5 amino terminus by independent groups; neither of the remarkably similar x-ray structures indicated crystallographic dimerization (Bhattacharya et al., 2007; Romier et al., 2007). In addition it was shown that certain organisms

completely lack the portion of the Taf5 amino terminus with homology to Lis1 (Romier et al., 2007). Thus the weight of evidence suggests that the Taf5 amino terminus possibly does not directly contribute to Taf5 dimerization or does so in a manner dependent on either additional Tafs, or dependent on posttranslational modifications. The practical implication of these observations is that as far as is known, the Taf5 amino terminus is available for interaction with other proteins besides itself or besides other Tafs.

The initial finding that the Rap1 binding domain in Taf12 is non-essential for growth was discouraging but easily reconcilable with available data about Taf stoichiometry and location within the TFIID complex. There are three different Tafs with which Rap1 can interact in vitro thus this opens the possibility of redundancy in the context of Rap1-TFIID interaction. The lack of any observable phenotypes in Taf12 truncation mutants did not 'soften the blow' when I reconciled this observation with the facts; at that time the practical implication was that I would be dealing with four proteins rather than just two (Rap1 and Taf12). These logistics were yet another obvious prelude to my PhD experience involving quite an extended/protracted campaign. Besides the possibility of redundancy amongst different protein, I should again mention the possibility of a 'distribution of function' model. Some of the possible 'distribution of function' relates to subtle differences between Tafs from different common model organisms, where it is possible that there is differential responsibility between orthologous proteins, but that overall function within the physiological context is maintained. Thus I would propose that overall function of the Taf4/Taf12 heterodimer is conserved even though individual protein characteristics are not. Clearly the known function associated with histone fold domains, that of heterodimerization, is conserved.

A proposed critical structural role of Taf4 within TFIID may also be tied to histone fold domain function along with Taf12 (Wright et al., 2006). Thus the remaining known role for the Taf4/Taf12 heterodimer is to function in direct activator-TFIID contact, and my findings are in overall agreement with this idea. But for practical reasons, Taf12 was not investigated much further, in part because I needed to characterize Taf4 and Taf5 first and then was obtaining promising information in the study of those two proteins. Because of the non-essential nature of the Taf12 RBD, I think that future experiments would have to include combination of *taf12* mutants with *taf4* or *taf5* mutants. For example, it would be interesting to determine if chimeric proteins, that include the Taf12 RBD fused to Taf4 and/or Taf5 variants deleted for their respective RBDs, could complement growth of yeast cells expressing RBD deleted-Taf12. Another simpler experiment would involve testing for synthetic phenotypes when various Taf12 truncations were combined with the Taf4 and/or Taf5 mutants within the same strain background. I initiated such studies though technical problems prevented me from drawing any conclusions, I do consider these experiments worthy of re-investigation. A lack of time and manpower has delayed re-attempts of these experiments. The resources at hand were directed towards characterization of Taf4 and Taf5 RBDs.

The finding of a Taf4 Rap1 binding domain essential for viability and located between residues 253-344 is the first report of a protein-protein interaction mediated by that portion of Taf4 in yeast; the only other known protein-protein interaction was between Taf4 and Taf12 as mediated by the bipartite Taf4 HFD, also essential for viability. Other features of this 'linker' portion of Taf4 include a generally unstructured domain organization when observed in isolation. We have observed that this portion of

Taf4 is phosphorylated on four residues when isolated from yeast (Manish K. Tripathi, unpublished observation). There are growth phenotypes associated with genetic manipulation of certain of these phosphorylated residues (JHL, unpublished) but detailed analyses to look at important features such as stoichiometry or dynamics of the modifications or characterization of responsible kinases has not been done. DNA binding activity of this portion of Taf4 has been reported (Shao et al., 2005), but the physiological relevance of this remains obscure. Whether DNA binding is mediated by this domain in vivo is unclear. Since this portion of Taf4 does not appear to interact with Tafs, as determined by my two-hybrid analyses, it certainly remains possible that DNA binding could occur through this domain. Human Taf4 was reported to interact with the GTF TFIIA and also with a repressor protein through the equivalent protein domain (Guermah et al., 2001; Olave et al., 1998). The interaction of this Taf4 domain with TFIIA in yeast may be of significance to our analyses and will be discussed at length in an ensuing chapter. However, it remains equally possible that the seemingly uncommitted character of this Taf4 domain renders it available to interact with activators such as Rap1.

After I mapped the Taf5 RBD to the amino terminus, it seemed that this would represent another non-essential RBD like the case with Taf12. I was led to believe this since studies in *S. pombe* had indicated that the Taf5 N-terminus is dispensable for viability (Yamamoto et al., 1997) but it turns out there is a serious flaw with this study; the shuffling strain used contained an intact chromosomally encoded *TAF5* gene because there are two gene copies in the *S. pombe* genome. I did confirm that NTD1 is dispensable for viability and a study published while this work was in progress mirrored my results (Romier et al., 2007). However, my addition to the body of knowledge

showed that NTD2 is needed for growth at elevated temperatures, and indicated a collaborative relationship between NTD2 and NTD1. Together these two domains are almost entirely essential for growth; strains missing both domains are barely alive and cease to grow with even the most minor environmental insult. A Taf5 protein missing NTD2 or both NTD1 and NTD2 is present in high amounts relative to the wild type protein, which may be the main reason these proteins can support growth, although the possibility has not been tested by "under-expression" studies. It seems likely that these domains may normally play a role in de-stabilizing Taf5 steady-state abundance, perhaps through the ubiquitin-proteasome degradation system. It has not been demonstrated that Taf5 is ubiquitinated, but Taf5 is known to be modified by the similar small ubiquitin-like moiety (SUMO) (Boyer-Guittaut et al., 2005). These modifications occur through lysine residues and NTD1 and NTD2 are highly enriched in lysine. The consensus SUMO modification site contains the consensus ZKXE peptide sequence, where Z refers to any hydrophobic residue. Several candidate sites loosely matching this consensus were identified in the Taf5 amino terminus. I performed some preliminary analysis of the role of lysine residues in Taf5 steady state abundance, using substitution of arginine for individual lysines. The results were inconclusive, and it has been shown that mutation of many lysine codons to arginine can be needed to retard overall affects on protein steady state abundance, as was demonstrated with the human p21 protein (Bloom et al., 2003; Bornstein et al., 2003). I did not proceed with generation of mutants where multiple lysine codons are switched to arginine, which is the next logical step in pursuit of the importance of ubiquitination/sumoylation in Taf5 steady state abundance. Ubiquitination or sumoylation might be more directly pursued by use of overexpressed hexahistidine

tagged-ubiquitin and mass spectrometry (Mayor and Deshaies, 2005). In any case it was comforting that the Taf5 domains implicated in Rap1 binding are clearly important for cellular growth as expected if truly important for Rap1-TFIID driven RPG transcription.

Once again, the essential nature of the putative Rap1 binding domains in Taf4 and Taf5 may be unrelated to an interaction with Rap1 in vivo. The most obvious alternative reason for this is responsibility for interaction with other Tafs for proper assembly of TFIID. Other possibilities include DNA binding mediated by the Taf4 domain of interest, although that possibility is difficult to test given the current state of knowledge. Another mechanism requiring the Taf5 N-terminus is oligomerization, although as I have discussed, this seems unlikely.

Taf-Taf interaction was the most testable hypothesis and I pursued that using systematic directed two-hybrid interaction screening. Even a non-specialist with the technique might ask if it is appropriate to use this yeast-based technique to study yeast proteins. One could envision that introduced yeast proteins, even those fused to the Gal4 DBD or AD, might compartmentalize to an organelle where they are restricted from participating in the two-hybrid, or even if properly compartmentalized they could be restricted from interacting with two-hybrid partners due to sequestration within their natural yeast protein complex(es). The latter situation would be more applicable to Tafs since they are nuclear proteins and present within endogenous high molecular weight TFIID. A related 'sequestration' caveat might apply to Rap1, which is a nuclear sequence-specific DNA binding protein; if the Gal4 DBD or AD fusion protein spends a majority of its time bound to the numerous natural Rap1 binding sites distributed throughout the genome it might be unavailable for participation in two-hybrid

interactions. Nevertheless, using the two-hybrid system Rap1 was characterized with respect to protein-protein interaction partners (Hardy et al., 1992b), which resulted in the identification of Rap1 interacting factors, or Rifs, one of which I used as a positive control in my two-hybrid studies with Rap1. It is ironic then but not surprising that this technique was not useful for characterization of Rap1-Taf interactions. The reason for me to say that this is no surprise is because this is a technique subject to both false positives and false negative results. Most of the attention on improvements to the system have focused on elimination of false positives. For example, this was the reason for creation of strain PJ69-4A, with its stringent *ADE2* reporter gene. There are less obvious solutions to the problem of false negatives. In fact there are practical issues which only worsen this situation including use of different expression vector, host, and reporter gene combinations amongst different studies. Prominent examples of this in this literature are the cited attempts at systematic characterization of yeast and plant Taf-Taf interaction, where there are differences in the findings between those studies and also between each of those studies and my findings. To reduce the chances of false negatives, I think it would be judicious to confirm interactions using as many reporter genes as possible, to examine the affects of changing locations of the epitope tags, and to attempt variation of expression levels through use of additional promoters and vector replication origins. Systematic optimization of Taf two-hybrid interactions would be worth looking into, particularly since we have many of the relevant plasmids and strains in place. Sorting out the differences in literature regarding Taf two-hybrid interactions would be a pursuit worthy of a PhD. Unfortunately this could not be justified during my time here. My first addition to the available reagents would be Tafs fused at their C-termini to Gal4 DBD

and/or AD, instead of just the N-termini. For reasons that have been mentioned, I suspect that different epitope tags can introduce artifactual effects into protein characteristics, and this has been seen by me and by others in the case of Taf5.

The advent of *E. coli*-based two-hybrid systems has provided an alternative to problems exclusively associated with the yeast system. I employed one such system involving a fusion between the lambda repressor DBD and Tafs or Rap1 and another fusion of Tafs or Rap1 to a portion of the omega subunit of *E. coli* RNA polymerase (Dove and Hochschild, 2004). Unfortunately this system has its own problems, some of which are protein specific. For example many Tafs cannot be effectively expressed in *E. coli* unless measures are taken to address problems with codon bias or solubility or stability issues (as is the case with Taf4 and Taf5). Thus while appropriate positive controls functioned effectively for me in this assay, no positive interactions could be detected between the proteins I was seeking information about. Additional improvements will need to be implemented in this system before it will be suitable for analysis of Taf protein-protein interactions.

Two-hybrid interactions between Taf4 and Taf5 with others Tafs were identified and interaction patterns did not vary between full length Taf4 or 5 and variants lacking Rap1 binding domains, indicating that these domains have no apparent responsibility for Taf-Taf interaction. I will say once again that there is the caveat of false negatives, meaning that the Taf4 and 5 domains actually do contribute to Taf-Taf interactions, albeit Taf-Taf interactions that my experimental set-up will not detect. Even so, we felt justified in proceeding with mutagenesis of sequence encoding RBDs in Taf4 and Taf5. Although I recognized the possibility that I was missing the whole story, I had obtained

enough information to be able to interpret the mechanisms underlying any putative growth and/or RPG transcription phenotypes. Therefore at this stage, regional randomization of *TAF4* and *TAF5* was considered appropriate.

However, randomization of sequence encoding of RBDs was expected to be quite labor intensive (see above). To try and circumvent this challenge I performed a set of preliminary experiments, with the logic that nature might be able to provide me with the material to test the in vivo function of the Taf5 RBD. I had in hand cDNAs for the *Drosophila* and human Taf5 genes, and used these to create chimeric proteins containing the C-terminus of yeast Taf5 and the N-terminus of one of either ortholog. The expectation was that the encoded proteins might represent a form with a partial loss of function within the N-terminal portion, and normal function within the C-terminus. Unfortunately the human/yeast Taf5 chimera presented the same growth phenotype as the yeast variant with NTD1 and NTD2 deleted. This can be explained because of the evolutionary distance between the yeast N-terminus and the human version; these have more individual changes within individual amino acid residues than a closer ortholog such as that of flies. The human variant could provide none of the function provided by the yeast version. On the other hand, the fly/yeast Taf5 chimera could not be tested because of apparent toxicity of the protein; no transformants were obtained containing the sequence-verified chimeric construct thus it seemed to act as a dominant negative. It would be necessary to use a regulatable on/off promoter to drive expression of this chimera, in order to definitively test whether it is a true dominant negative. These type of experiments were shown to be of limited utility in the case of chimeras between human and yeast Taf4, so at one level my lack of success was unsurprising although it was worth

a try (Thuault et al., 2002). I chose to abandon use of the dTaf5/yTaf5 chimeric protein, since such experiments might prove overly complicated, and consequently proceeded with the more conventional generation and isolation of mutant alleles.

Dozens of *taf4* alleles conferring temperature conditional growth were identified and characterized. When the mobility of the encoded proteins was analyzed by SDS-PAGE and western blotting, it became clear that many of the corresponding mutant alleles likely encoded proteins with an abnormal number of amino acids. Consistent with the hypothesis, many of the alleles included frameshifts in sequence encoding the CCTD, which was not desired since this portion of the protein does not contribute to Rap1 interaction. Previously I have observed that one of the C-terminal Taf4 truncations was associated with a temperature conditional phenotype, and that a previously described *taf4* allele encodes a protein with a frameshift in the CCTD as well as other histone fold domain alterations (JHL, data not shown, (Shen et al., 2003). Since many of the CCTD variants contained just a single nucleotide deletion, I suspect that truncated PCR primer oligonucleotides introduced the technical artifact responsible for much of this problem, which is not surprising given that the primer was by necessity ~100 nucleotides long. Gel purification of the primer was evidently not sufficient for removal of N-1 length synthesis products. In retrospect, a better method of creating the mutagenized library would have started with introduction of silent mutations encoding unique restriction enzyme sites closer to the region encoding the RBD. This would in turn allow the use of shorter oligonucleotides and avoid creation of any alleles that affected portions of Taf4 other than the RBD. With this simple modification in creation of the mutagenized *taf4*

library, the number of mutants identified would speculatively have been similar to the situation with *taf5*.

I was fortunate enough to isolate 3 unique *taf4* temperature conditional alleles. It is always good to have more than one mutant available, to avoid the observance of artifacts due to allele-specific phenotypic changes. Moreover, if I had gone ahead and performed a saturating screen of the mutagenized library I might have identified a couple more useful *taf4* alleles, but I was discouraged from this pursuit during a PhD committee meeting. The logic behind this decision was that even though the genetic screen for *taf4* RBD mutants was not as effective as I might have hoped, afterwards I was positioned to determine the affect of these mutants on RPG transcription. The reagents needed to advance the project were thus available after discovery of these three *taf4* alleles. Their identification represented the first *taf4* point mutant temperature conditional alleles with coding substitutions in a domain other than the HFD. This result again supported the importance of the putative Taf4 RBD to cellular growth.

A few additional strategic decisions were made before undertaking the mutagenesis of Taf5 RBD-encoding sequence. First, I decided to restrict the mutagenesis to NTD2-encoding sequence even though Rap1 interacts through the entire N-terminus, not just NTD2 but likely NTD1 as well. This was done for both theoretical and practical reasons. First, NTD1 is non-essential, so it would be difficult to establish the contribution of substitutions there to growth phenotypes without performing a supplemental site-directed point mutagenesis study. Second, the convenient location of restriction enzyme sites for generation of the mutant library almost perfectly flanked NTD2-encoding sequence. Third, I had established that deletion of NTD2 alone

conferred a rather tight temperature conditional growth phenotype, so it seemed a safe bet that amino acid substitutions there would be sufficient to confer similar phenotypes. Around 30 unique temperature conditional mutants were identified and 2 more represented reisolation of the same allele, consistent with the number of colonies screened being far less than the total number of clones available in the mutagenized library. I attribute the relatively greater success of *TAF5* mutagenesis as compared to *TAF4* to two reasons. First, the natural restriction sites lying very close to NTD2 coding sequence facilitated the use of short DNA primers for PCR mutagenesis, and these primers were also subjected to preparative PAGE and as such were essentially all unit length and thus ~100% correct with a very low proportion able to cause frameshifts. Second, unlike the Taf4 C-terminus where deletion or alteration of the CCTD confers temperature conditional growth, Taf5 residues C-terminal to the Taf5 RBD are strictly required for growth and as such any unwanted alterations to that protein region would result in a mutagenized clone that would score completely negative for growth in my genetic screen.

The generation of the *taf5* alleles represents the first identification of point mutants in a portion of *TAF5* other than the WD repeats. The first characterized *taf5* temperature conditional mutant contained substitutions in the WD repeats, and a subsequent study generated a number of mutants all of which also encoded WD substitutions (Apone et al., 1996; Durso et al., 2001). Notably, several of the mutants identified in the second study also contained substitutions in the N-terminus. I obtained these mutants with the idea of combining the 5' mutant gene fragments from these alleles with a 3' wild type gene fragment to test the affect of the N-terminal substitutions

encoded in these alleles. The success of my *taf5* mutant screen reduced the priority of performing this experiment. My results again indicate that temperature conditional *taf5* mutants can be isolated that do not disturb the function of the WD repeats. Because I placed a priority on selecting *taf4* and *taf5* mutants that did not appear to affect bulk protein stability, and because the identified domains did not appear to influence Taf-Taf interaction, we could anticipate that the growth and putative transcription phenotypes would be independent of gross disruption of TFIID structure. Consistent with this, the growth characteristics of the mutant strains indicate that the cells can recover from a period of heat shock lasting as long as 12 hours. The recovery from heat shock indicated two things about possible molecular phenotypes. First, the altered proteins are probably not completely disrupted, misfolded, and degraded from an unfolded protein response. There are reports of *taf* mutant alleles that encode unstable proteins and this can lead to degradation of the TFIID complex, and cells containing these are less able to recover from heat shock (Moqtaderi et al., 1996b; Reese and Green, 2003). Perhaps a complete removal/degradation of TFIID from such heat-treated cells renders them susceptible to terminal cell cycle arrest. This idea is related to the second indication about the molecular phenotype of my novel *taf* mutants. TFIID is important for transcription of genes required for cell cycle progression such as G1 and G2 cyclins needed for the G1/S and G2/M transitions; TFIID dependency involves Taf1 but also the WD repeats of Taf5 (Apone et al., 1996; Walker et al., 1997). Because my mutants were able to resume growth quickly after removal from the restrictive temperature, this phenotype indicates that there was probably little affect on the transcription of such cell cycle-specific genes, which are TFIID-dependent but Rap1-independent. By extension, the growth defect

could logically be expected as an affect, at least in part, on Rap1 and TFIID dependent ribosomal protein gene transcription. A test of my hypothesis was appropriate at this time, and the results of these experiments are presented and discussed in the next chapter.

CHAPTER III

ANALYSES OF THE ROLE OF RAP1-TAF INTERACTION IN RNA POLYMERASE II TRANSCRIPTION IN VIVO

Considerations in Testing for Molecular Phenotypes Associated with Novel *taf4* and *taf5* Temperature Conditional Mutants

The identification of Rap1 binding domains in Taf4 and Taf5, the observance that these domains are required for cellular growth without apparent roles in TFIID integrity, and the sensitivity of these domains to amino acid substitutions conferring conditional growth all supported the notion that Rap1-Taf4 and Rap1-Taf5 interaction are physiologically relevant. There are several complications to interpretation of phenotypes associated with *taf4* and *taf5* mutants not the least of which is that TFIID, and thus Taf4 and Taf5, are involved in the transcription of many genes (Huisinga and Pugh, 2004; Shen et al., 2003). Since TFIID and Rap1 commonly regulate RPGs these seem like great candidates for genes experiencing reduced transcription rates in *taf* mutants at the non-permissive temperature, as compared to wild type. However, it is theoretically possible that defective transcription of just one gene in the entire genome could cause a dramatic growth deficiency phenotype, even in the absence of an affect on RPG transcription. For example, abolishing transcription of a single *TAF4* and/or *TAF5* dependent cyclin-encoding gene could cause such an affect. However, as mentioned, comparison of several studies in the literature would argue against this scenario in the case of these novel *taf* mutants, as do the growth cessation/recovery phenotypes

mentioned in the last chapter (Apone et al., 1996; Reese and Green, 2003; Walker et al., 1997).

There were several key considerations in testing the affect of *taf* mutants on RPG transcription. Proper experimental design, technique, and choice of controls were especially paramount to allow unbiased and meaningful interpretation of results. RPG transcription rate is exquisitely linked to the growth stage and nutrient availability in the culture environment; subjecting cells to heat shock causes a momentary disruption of RPG transcription even in wild type strains (Sanders et al., 1999). Common laboratory manipulations such as centrifugation can also disrupt RPG transcription rates (JHL, personal observation). Of course it is essential to prepare high-quality RNA from all cultures of interest, which in and of itself is a non-trivial undertaking. These and other technical considerations had to be accounted for.

As with the immunoblotting data already shown, it was critical to provide internal controls for accurate comparison of RPG transcript abundance from different strains. A related control relates to the possibility of off-target affects. Since incubating yeast cultures above a threshold temperature causes a transient cessation of RPG transcription, it is necessary to sample cultures after a more chronic heat shock to differentiate the natural physiological response to the treatment versus defective regulation occurring via the altered temperature sensitive Taf4 or Taf5 proteins. This experimental characteristic increases the chance of observing off-target affects on gene transcription that is normally TFIID-independent. As such it was important to confirm that the abundance of TFIID-*independent* transcripts remained unperturbed. Ideally, I hoped to observe very little reduction in transcript abundance of genes that are independent of TFIID and Rap1, but a

reduction in transcription of coordinately regulated, Rap1- and TFIID-dependent genes; in other words the majority of Ribosomal Protein Genes. A caveat to this is that other activators could theoretically function in a manner analogous to Rap1, by using the identified surfaces in Taf4, Taf5, and/or Taf12 to utilize TFIID as a coactivator. Genes utilizing such activators would also be affected in *taf4* and *taf5* mutants grown at the non-permissive temperature.

There is an additional caveat in that because Taf5 is also a constituent of the SAGA complex, the *taf5* mutants could be scoring SAGA function rather than TFIID function, or in a more complicated scenario the *taf5* mutants could be scoring the requirement for both SAGA and TFIID at the same target genes (Grant et al., 1998; van Oevelen et al., 2006; Zhang et al., 2008). The bulk of data suggests that SAGA does not directly regulate RPG transcription, but I felt it important to confirm this using my own yeast strains and methods (Huisinga and Pugh, 2004). Even still, affects on other genes might occur through a SAGA-dependent pathway. There is no functional data about the mechanistic role of Tafs in SAGA function, so this would be a difficult issue to contend with. Aside from SAGA, I thought it important to consider the possibility of additional contributing coregulatory complexes. The NuA4 complex is proposed to stimulate RPG transcription but the mechanism of targeting is completely unknown although like SAGA, NuA4 contains the Tra1 subunit that is a direct target of the Gal4 activator (Bhaumik et al., 2004; Reid et al., 2000). The Mediator complex is another activator-targeted coregulatory complex in Pol II transcriptional regulation but its role in RPG expression had not been characterized at the time these studies began and as such was worthy of investigation.

Most of the commonly used assays that measure specific mRNA abundance such as northern blot, RNase protection, S1 nuclease, primer extension, or RT-PCR can be used to monitor at maximum a few dozen different transcripts within a reasonable amount of time. Therefore, while one or more of these techniques might be useful for testing for an RPG transcriptional phenotype in the *taf* ts mutants, none would be suited to look at more than a few TFIID- and Rap1-independent transcripts; this is important to ascertain to what extent non-RPG transcripts are affected in the mutants. Likewise, given that there are well over one hundred RPGs possibly affected in the *taf* ts mutants, it would not be possible to fully authenticate the expected broad phenotype for the coordinately regulated gene family. No such limitations exist when applying whole transcriptome analyses such as serial analysis of gene expression (SAGE), microarray, or RNA-seq because these methods can in theory simultaneously quantify every transcript encoded by the genome (Holstege et al., 1998; Nagalakshmi et al., 2008; Velculescu et al., 1997). However, being that quantification of transcript abundance typically involves a population measurement, information about abundance of cell-cycle regulated transcripts is not easily extracted unless the population of cells being studied is first synchronized at a specific stage of the cell cycle (Spellman et al., 1998). Thus information about this particular class of genes, some of which are Taf dependent, would not be obtained unless extra steps were taken to ensure their inclusion in population-based studies of transcript abundance. I considered these steps to be beyond the scope of my study due to the technical difficulty involved. I did plan on testing whether the entire RPG family was affected in my novel *taf* mutants and a microarray experiment seemed the most practical method to achieve this. Due to the caveats mentioned I did not

anticipate being able to conclude that an RPG-exclusive transcription phenotype was responsible for the growth phenotypes associated with *taf* mutants.

In the likely event of an RPG transcription phenotype occurring in the *taf4* and *taf5* mutants, I would also need to test/confirm the mechanistic basis of that phenomenon. Since the objective of the domain-directed random mutagenesis was to create and identify mutant alleles encoding proteins compromised in their ability to interact with Rap1, it seemed logical that Rap1-TFIID interaction would be thusly affected in vivo at the non-permissive or semi-permissive temperature. Deficient Rap1-TFIID interaction would compromise RPG transcription rates, which would in turn reduce cellular growth capacity. There are several methods that could be used to score Rap1-TFIID interaction and these range from fully in vivo approaches that do not require any cell disruption, to in vitro interaction assays just using purified TFIID and Rap1, to extract-based methods that fall somewhere in between with respect to physiological context. Our lab has a lot of experience using co-immunoprecipitation (Co-IP) and this is an example of an extract-based method to measure protein-protein interactions (Sanders et al., 2002b). Besides being a non-in vivo method, there are additional drawbacks to the technique like the need for good antibodies or epitope tags to be available. Also the technique relies on western blotting, which does not readily lend itself to precise quantification of interactions. Most notable, Co-IP does not demonstrate direct interactions per se, but must be supplemented with other techniques such as in vitro interaction assays using pure components for confirmation. However, because of our experience with the method, Co-IP was our first choice for characterizing Rap1-TFIID interactions and comparing them between wild type and *taf* mutants. My second alternative was the use of isolated TFIID and Rap1 to

measure interactions in vitro. Use of isolated proteins to assay interaction suffers from the main limitation of being able to obtain TFIID from wild type and *taf* mutant yeast cells in sufficient yield and purity to do the experiments. It is quite formidable to isolate TFIID from a specially engineered strain, even though that strains grows as well as wild type, because a large quantity of starting material is needed and the procedure is labor intensive, technically challenging, and expensive (Sanders et al., 2002a). In order to isolate TFIID containing altered Taf proteins, the mutant yeast would first have to be engineered to include an N-terminal epitope tag on Taf1 to allow efficient affinity purification of the complex. Second, large-scale cultures would need to be grown to serve as the starting material for TFIID purification. Third, intact TFIID would need to be successfully isolated from the mutant cells and proven to possess some of the same apparent biochemical properties as TFIID isolated from wild type cells. A not-inconsiderable amount of technical skill would be required to succeed in each of these steps. Successful isolation of TFIID from mutant cells has never been reported, so the technical hurdles might be insurmountable. A purely in vivo approach would involve the use of fluorescent protein epitope tags that could act as sensors of fusion protein proximity (binding) in Fluorescence resonance energy transfer (FRET) assays. By introducing enhanced cyan fluorescent protein (ECFP) and enhanced yellow fluorescent protein (EYFP) epitope tags at the C-termini of every SAGA complex component and to the Gal4 activator, Bhaumik and colleagues were able to identify subunit specific Gal4-SAGA protein-protein interactions in yeast cells using live cell imaging with confocal microscopy. The Bhaumik study represents the only FRET-based characterization of eukaryotic activator-coregulator interactions reported in the literature to date (Bhaumik et

al., 2004). While no equivalent analyses of mutants has been reported, in theory the technique could provide quantitative information describing altered/weakened interactions between Tafs and Rap1 such as what may be occurring in the *taf* mutants. So while on paper FRET looks like the best approach, it suffers from our lack of experience with it, and the problems of artifacts that can come with the use of epitope tags, particularly the larger/bulkier variety such as the EXFPs. Epitope tags were documented by me and others to cause problems with Taf4, Taf5, and Rap1 depending on which tag was used and where the tag was appended. Such problems might eliminate the possibility of using FRET to characterize Rap1-Taf interactions in wild type and mutant cells.

The limitation of all of the above approaches is that they are describing characterization of Rap1-TFIID interactions in the context of a single *taf* mutant. If the assumption is that a single Rap1-Taf-TFIID binding event is the driver of RPG transcription, then it is reasonable to expect a deficiency in Rap1-TFIID interaction in the mutants. However, this seems unlikely given that three different Tafs are potentially contacted by Rap1 in vivo, each of these three Tafs are present in multiple molecules per TFIID molecule, and are present in different locations in the complex. Even if a simple binary binding event is the purpose and final outcome of Rap1-TFIID interaction, elimination/reduction of interaction between Rap1 and a particular Taf within the complex would at most eliminate 33% of overall Rap1-TFIID interaction, assuming independent binding sites are provided by each Taf molecule (see **Figure 1.6**). Moreover, it is possible that the mutants could experience a deficit in interaction between Rap1 and TFIID immeasurable by the techniques employed, but still present in vivo and

severe enough to cause a reduction in RPG transcription incompatible with rapid cellular growth. In fact, I considered it worth determining first if multiple *taf* mutants could be combined into the same strain. If such strains proved to be viable and able to grow at wild type rates, even at 10% of wild type rate, such multiply mutant strains could provide a useful avenue to apply one of the above-mentioned techniques of quantifying Rap1-Taf-TFIID interaction in these strains. Creation of these strains would involve traditional yeast genetic crosses along with phenotypic and genotypic confirmation of presence of multiple mutant *taf* alleles.

The problem of examining the mutational sensitivity of Rap1-TFIID interaction stems from potential differences between the mode of Rap1-Taf interaction and the mode of Rap1-TFIID interaction. The project initiated with the mapping of RBDs in each Taf, so it would be logical to expect the proteins encoded by the loss of function *taf* alleles to exhibit reduced affinity for Rap1. As a supplement to defining Rap1-TFIID interaction in the mutants, characterization of in vitro interactions between Rap1 and the altered Taf proteins would be of value. The problem with doing this experiment relates to the relatively poor quantitative capabilities of the assays already in place such as the Far Western. Other related techniques such as the in vitro pulldown also do not lend themselves to precise quantification of binding affinities. The majority of the time, workers use these approaches only for qualitative purposes rather than both qualitative and quantitative measurement of interaction. The most quantitative technique used for examining intermolecular interactions that we had used in the lab was fluorescence anisotropy, and in the format in which we had used anisotropy it could not be expected to be useful due to its limited capacity for throughput, because I could expect a need to

perform many experiments to accurately and precisely define the interactions of interest (Gumbs et al., 2003). The need for a high-throughput, quantitative method was my prime motivation for investing so much effort into setting up the yeast and bacterial two-hybrid techniques, but unfortunately neither was applicable for detecting the Rap1-Taf interactions. Therefore, I could expect to introduce novel techniques into the lab in order to quantify binding interactions between Rap1 and the altered Taf4 and Taf5 proteins.

Of course the physiological relevance of simple binary interactions between Rap1 and Tafs are of dubious significance in the absence of in vivo interaction data documenting loss of protein function, or at least observance of an affect on binding in the context of TFIID. The stated problem of redundancy or overlapping function between Tafs in mediating Rap1 interaction with TFIID might allow observance of a deficit in Rap1-TFIID binding only if multiple Taf RBDs were altered. Because the available *taf4* and *taf5* mutants have rather dramatic growth phenotypes, and more importantly because the affected protein domains are ostensibly contributing to the same molecular pathway, it is easily conceivable that it would be impossible to obtain strains containing *taf* RBD mutants corresponding to both proteins. It is common for mutations in different genes whose protein products contribute to the same process to result in loss of viability (known as synthetic lethality), even when the single mutants do not display overt phenotypes (Davierwala et al., 2005). If that turned out to be the case, then I would not be able to obtain more precise genetic reagents to confirm loss of Rap1-TFIID interaction. An inviable strain obviously could not be used to quantify Rap1-TFIID interaction using any available methods.

However, the use of advanced genetic approaches, meaning combination of different mutations and looking for enhanced growth defects, holds additional value. This relates to the lack of severe growth defects in certain *rap1* mutants. We found that Rap1 utilizes both its centrally located DBD and unidentified portion of the C-terminus to interact with TFIID in vitro (Garbett et al., 2007). While the Rap1 DBD is required for cell growth, cells completely lacking the C terminus are able to grow, albeit at a reduced rate. Deletion of individual C terminal domains, including Tox, AD, and Silencing domains results in much more subtle phenotypes (as seen in **Figure 3.8,** (Yu et al., 2001). If one or more specific C-terminal Rap1 domains were involved in TFIID binding, strains missing both this domain might exhibit a growth phenotype if a mutation in one the Taf RBDs were introduced. Such an approach promised to reveal several features of Rap1-Taf/TFIID interaction; it would indicate which specific C-terminal Rap1 domain(s) were important for cellular growth, suggest which Tafs were contacted by the Rap1 C-terminus, and in doing both of the above provide further indication of the physiological relevance of Rap1-Taf interaction.

To summarize, the objectives of the experiments described in this chapter were to determine if an RPG transcription defect was occurring in the *taf ts* mutants at the non-permissive temperature, determine the extent to which other genes in the genome experienced transcriptional defects, identify if RPGs were actually broadly affected as would be expected given their coordinately regulated nature, and test whether Rap1-TFIID or Rap1-Taf interactions were actually compromised in the novel *taf4* and *taf5* mutants. The latter objective would be approached using a variety of techniques, including test of synthetic genetic interactions, which also promised to characterize Taf

binding domains in Rap1. The methods used to complete these objectives, a description of experimental results and a discussion of the significance of the findings are described in the ensuing sections of this chapter.

METHODS

Temperature Shift, RNA Extraction, and Transcript Analyses by 5' Primer Extension

A single colony from each shuffled *TAF4*, *TAF5*, *taf4 ts*, or *taf5 ts* strain was picked from a fresh 23° C 5-FOA-containing plate used for the serial dilution growth analyses. Each colony was grown to saturation in YPAD at 23° C (3-5 days) then re-inoculated into two separate 50 ml YPAD cultures and shaken overnight at the permissive temperature of 23° C. The quantity of each pre-culture used as an inoculum to the 50 ml cultures was estimated based upon preliminary growth curve analyses, and were adjusted so that cells would reach log-phase at approximately the same time the following morning. For example, certain mutant strains grow at ~10% the rate of the equivalent wild type strain and would therefore need ~10 fold more cells as inoculum to reach the desired cell density in the same time frame as wild type. The next day cultures were grown until they each reached an optical density at 600 nanometers (A^{600})of ~1. At that point 50 ml of fresh pre-warmed 50° C YPAD was added to one of the two cultures and placed in a 37° C incubator. The other culture was spiked with 50 ml of 23° C YPAD and retained at room temperature. After two hours, cultures were rapidly harvested by filtration and total RNA immediately isolated by hot acidic phenol extraction and ethanol

precipitation (Belinchon et al., 2004). The two hour time point was selected because at that stage wild type cells have mostly recovered from the transient cessation of RPG transcription that occurs upon heat shock but this time is short enough to avoid some of the off-target affects associated with more chronic heat shock (Reese and Green, 2003). Cultures were harvested by filtration because this is up to ten times more rapid than centrifugation and consequently results in better yield of RNA, with the transcriptome representing the fully heat shocked state, rather than the partially heat shocked state resulting from culture cooling during the centrifugation run (Belinchon et al., 2004). Filters were immediately submerged in hot acidic phenol and 50 mM acetate, 10 mM EDTA, 0.5% SDS buffer maintained in a 65^{o} C water bath (Schmitt et al., 1990). This method captured RNA representative of the cellular pool because degradation and synthesis were rapidly halted by the harsh conditions. The acidic phenol removed much of the liberated DNA because it is acid-labile, moreover nuclei were left mostly intact by this procedure and large intact chromosomal DNA partitioned into the organic phase along with the nuclei, while cytoplasmic RNA and some of the nuclear RNA was captured in the aqueous phase. The aqueous and organic phases were split by centrifugation and the aqueous phase subjected to one more 10 minute incubation in fresh hot acidic phenol. Subsequently the RNA was extracted two times more with a room temperature mixture of 1:1 acid phenol/chloroform and twice thereafter with chloroform. The RNA was ethanol precipitated after incubation of tubes on dry ice, and resuspended in 10 mM Tris pH 7.9, 0.1 mM EDTA, 0.1% SDS. After incubating the resuspended RNA at 65^{o} C to fully dissolve it, one more round of ethanol precipitation was performed. This last step served to remove any residual phenol which can interfere with accurate

quantification of RNA concentration by spectrophotometry. Concentration of final resuspended RNA was measured in quadruplicate and normalized to 0.5-1 mg/ml. A^{260}, A^{280}, and A^{270} measurements were taken to estimate nucleic acid content, protein content, and organic contamination. A^{260}/ A^{280} ratios were ~2, and the A^{270} values fell in between the value of A^{260} and A^{280} if organic contamination was absent. Equivalent sample-to-sample concentration and integrity of both high and low molecular weight RNA was confirmed respectively by electrophoresis in both denaturing 1.2% agarose and 10% polyacrylamide gels, followed by staining with Sybr Gold (Invitrogen) and imaging using a BioRad Pharos Fx machine.

Analysis of specific messages was by 5' primer extension. Complementary gel-purified (unit length) oligos were labelled with T4 Polynucleotide Kinase and γ -^{32}P-ATP (MP Biochemicals, 7000 Ci/mmole) then de-salted by Sephadex G-50 chromatography. Specific activities were greater than ~6000 cpm/fmole as determined by liquid scintillation counting, except that *RDN1* and *U3* probes were adjusted to ~600 and ~200 cpm/fmole, respectively, by dilution with unlabeled probe. For best results, I found it was preferable to label oligonucleotide probes as soon as possible upon receipt of freshly prepared radioisotope and to use the labeled probes for primer extension quickly. A mixture containing ~35 fmoles of each labeled probe (*RDN1, PGK1, U3, RPS2, RPS3*, and *RPS5*) were simultaneously mixed with 5 μg total RNA and quantitatively precipitated along with a labeled 330 bp DNA fragment recovery control. The pellet was resuspended well in annealing buffer (1X AMV RT buffer plus 0.5 mM dNTPs), heated briefly at 100° then placed at 65° for 30 minutes. Samples were spun briefly after which Avian Myeloblastosis Virus Reverse Transcriptase (AMV RT, Promega) was added and

incubation continued at 42° for 45 minutes. Samples were then precipitated, resuspended in formamide sample buffer, heat denatured, and resolved on 8% polyacrylamide sequencing gels (IBI Model 45 STS) made in 1.5X Tris/Borate/EDTA buffer. Sequencing gels were run at constant temperature, 50°-55° C, as monitored using an IBI laboratory surface thermometer. Dried gels were exposed to a phosphoimager (Kodak K-Screen) overnight and imaged using a BioRad Pharos Fx Imager. Image analysis and quantification was done exclusively using BioRad Quantity One software. Dried gels were also subjected to film autoradiography at room temperature without an intensifying screen for 2-5 days. Additional primer extension was done using a variety of probes to Taf-dependent genes including *RPS12, RPS13, RPS15, RPS20, RPL3, RPL5, RPL28, RPL30,* and Taf-independent genes *HHT3, ADH1, ADH3, TPI1*, and *SCR1*.

Strains used for the analyses of SAGA and Mediator contribution to RPG transcription were procured from the Open Biosystems Systematic Gene Deletion Collection. These strains contained a G418 resistance-marked disruption of genes encoding single SAGA or Mediator complex subunits and were created on the BY4741 background. 100 ml log phase cultures of these strains were grown at room temperature and processed for RNA which was subjected to 5' primer extension as described above.

RNA Analyses by Slot Blot and Microarray

Measurement of total polyA$^+$ RNA content is a useful indicator of RPG transcript abundance, since ~1 out of 3 mRNAs originates from the 137 RPGs in yeast (DeRisi et al., 1997; Gorenstein and Warner, 1976; Hereford and Rosbash, 1977; Warner, 1999). The most convenient method to do this is the Oligo dT slot blot. Using a slotted vacuum

manifold and Zeta Probe GT membrane material (BioRad), 1.8 μg of RNA sample in 300 μl of 1 M ammonium acetate was applied with slow vacuum to one slot of the vacuum manifold, this was done in triplicate and the vacuum set so that the full volume of one 300 μl sample was fully applied at the time that the third sample was just being introduced into the apparatus. After all samples were applied, membrane-bound RNA was UV crosslinked using a Stratalinker machine. Next the membrane was blocked using 25 ml of pre-hybridization solution (5X SSPE, 0.5X Denhardt's solution, 0.5% SDS, 0.1 mg/ml salmon sperm DNA) in a glass tube with rotation at 37° C for three hours. Afterwards fresh pre-hybridization solution replaced the solution in contact with the membrane; it contained ~160 pmoles of γ–^{32}P-labelled oligo dT 20 mer probe at specific activity of ~250 cpm/fmole. The specific activity of this probe was kept low by providing a large excess of cold ATP to the labeling reaction. Hybridization proceeded overnight at 37° C and was followed by three 15 minutes washes in 2X SSPE/0.1% SDS then three 15 minutes washes in 0.2X SSPE/0.1% SDS, all at room temperature. The damp membrane was placed into an acetate envelope and exposed to a phosphoimager at room temperature for at least 2-3 hours.

Preparation for microarray analyses was initiated after carrying out several primer extension and slot blot experiments. RNA was available from two independent biological replicates, representing unique mutant clones, temperature shift growth experiments, and RNA preparations. These were the same samples that had been tested by primer extension and slot blot, and originated from strains *TAF5*, *taf5-17*, *taf5-45*, *taf5-408, taf5-10.4,* and *taf1 ts2*. RNAs were again subjected to two rounds of ethanol precipitation because it was found that 0.1% SDS in the resuspension buffer interfered

with downstream procedures. At that point all samples were submitted to the Vanderbilt Microarray Shared Resource for quality control. RNA concentration was checked using a nanodrop instrument (values were essentially the same as my spectrophotometer data) and RNA integrity was checked by Bioanalysis (precast agarose gels to compare sample to sample variation in large ribosomal RNA content). cDNA was prepared from 10 μg total RNA using Moloney Murine Leukemia Virus Reverse Transcriptase (MMLV RT) and oligo dT (specifically prepared for conjugation of Cy Dyes through imido ester linkages). After successful cDNA synthesis and cleanup/desalting of reactions, Cy3 or Cy5 dye was conjugated to 23° or 37° samples, respectively. An equivalent mass of each Cy3 and Cy5/23° and 37° pair were combined, dried down, and resuspended in 12 μl of hybridization mix, each of which contained a unique sample tracking control ("bar code"). 6 μl of this mix was applied to one slide in a Nimblegen 12-plex oligonucleotide array (Roche-Nimblegen, Indianapolis, Indiana) and incubated with the array at 37° for 17 hours. Therefore two-color hybridizations were carried out for six different strains, with biological replicates for each strain, resulting in 12 different hybridizations per array. Washing was performed according to Nimblegen instructions and the array was dried using an array centrifuge. The array was scanned using an AXON 4000B instrument according to the Nimblegen instructions for two-color hybridizations. Raw data were imported into the Arraystar software program and decrypted. Files were submitted to the NCBI GEO database (GSE20444 record). HEAT map clustering analysis of transcripts was performed using Arraystar.

Immunological Techniques

To measure Rap1-TFIID interaction and TFIID integrity in extracts of wild type and mutant strains, it was necessary to perform co-immunoprecipitation experiments. To avoid the generation of multiple epitope-tagged strains, it was most practical to use polyclonal antibodies to Rap1 and several Tafs. The use of such antibodies has the additional advantage that epitope accessibility is less of a factor than can potentially occur when using epitope tags, because with polyclonal antibodies epitopes are potentially distributed throughout the protein rather than localized to one protein area. Fortunately we had antiserum from rabbits immunized with full length Rap1 or each of the Tafs. I thought it best to affinity purify the IgGs from these antisera, in order to increase the specific antigen binding activity of these antibodies. A 5-10 ml portion of the latest bleed from each of two rabbits was mixed and a saturated solution of ammonium sulfate was slowly added to 33% of saturation (.182 grams ammonium sulfate per ml of antisera) with end-over-end mixing at 4° C. A portion of 1/3 the volume of the final necessary amount of saturated ammonium sulfate was added twice at five minute intervals, followed by two successive additions of 1/6 the volume needed. Precipitated material including total IgG was collected by centrifugation in the SS34 rotor. The pellet was dried using kimwipes and resuspended in 25 mM HEPES pH 7.6, 150 mM NaCl, 0.1% Triton X100, and 10% glycerol and applied to a 1 ml packed bed of Protein A sepharose (Sigma) and bound in batch with mixing at 4° C for 3 hours. Beads were collected and the supernatant decanted and saved. The column was washed three times with the above buffer and IgG was eluted by addition of 1 ml of 100 mM Sodium Citrate pH 3.5, 0.1% Triton X100, and 10% glycerol. This eluate was adjusted to near neutral

pH by addition of 10 μl of 1 M Hepes pH 7.6. In some cases these fractions, containing essentially pure total IgG, were subjected to antigen affinity chromatography. Full length soluble antigens were expressed in *E. coli.* Rap1 was expressed and purified as described above. Taf1 was expressed from pET28A *TAF1* (full length NheI/XhoI fragment) and purified by successive Ni-NTA and SP sepharose columns. Taf3 was expressed from pET28A *TAF3* (full length NdeI/BamHI fragment) and purified by successive Ni –NTA and Q sepharose columns. Taf4 was expressed from pBG100 *TAF4* (full length BamHI/SalI fragment) coexpressed along with Taf12 from pACYC11b *TAF12* (full length NdeI/BamHI fragment) and purified as described above. Taf6 was expressed from pBG101 *TAF6* (full length BamHI/HindIII fragment), coexpressed with Taf9 expressed from pACYC11b *TAF9* (full length NdeI/BamHI fragment) and purified by successive Ni-NTA and Glutathione sepharose columns. Taf8 was expressed from pET28A *TAF8* (full length NdeI/BamHI fragment) and purified by successive Ni-NTA and Q sepharose columns. Taf10 was expressed from pBG101 *TAF10* (full length NdeI/BamHI fragment) and purified by successive Ni-NTA and Glutathione sepharose columns. Taf11 was expressed from pET28A *TAF11* (full length NdeI/XhoI fragment), coexpressed with Taf13 expressed from pACYC11b *TAF13* (full length NdeI/XhoI fragment) and purified by successive Ni-NTA and SP sepharose columns. TBP was expressed from pBG101 *TBP* (full length BamHI/XhoI fragment) and purified by successive Ni-NTA and Glutathione sepharose columns. Purified proteins were again bound to a 1 ml bed of Ni-NTA sepharose, and unbound protein was removed by washing. Whole IgG preparations supplemented with 10 mM imidazole were incubated with the Ni-NTA-immobilized antigen, unbound and loosely bound IgG was removed by washing, and antigen-specific

antibodies were eluted from the immobilized antigen with 25 mM PIPES pH 6.8, 150 mM Sodium Acetate, 2.5 M Magnesium Acetate, 0.1% Triton X100, and 10% Glycerol. The excess magnesium was subsequently removed by extensive dialysis against the same buffer lacking the magnesium acetate. Taf4/12, Taf6/9, and Taf11/13 heterodimer preparations were used to affinity purify both binding partner-specific IgGs. Full length Taf2 was not successfully expressed in bacteria, and Taf7 could also not be obtained from bacteria due to apparent toxicity of plasmids encoding the full length protein. As such Taf2 and Taf7 were expressed and purified from Sf9 cells infected with recombinant baculoviruses were used to affinity purify anti-Taf2 and anti-Taf7 IgGs. Previously prepared aliquots of antigen affinity purified anti-Taf5 IgGs were used.

30° C represents a "semi-permissive" temperature for growth of *taf5* conditional mutants. Higher growth temperatures (heat shock) could not be used because this resulted in extracts containing excessive proteolytic activity. Importantly, the strain background also ended up being important for the success of these experiments and the SEY6211 *TAF5* shuffling strain was selected because it appeared to consistently yield the best whole cell extracts with respect to Taf content and integrity. *taf5* mutants were the focus of these experiments because they exhibited much stronger RPG transcription phenotypes than the *taf4* mutants. 250 ml YPAD cultures were grown overnight at 30° so that they reached an OD^{600} of 0.5-1 the next morning. Cells were harvested by centrifugation and resuspended in 0.6 ml buffer for every 50 OD^{600} units of cells. The initial buffer contained 20 mM HEPES pH 7.6, 200 mM Sodium Acetate, 0.1% Triton X100, 10% Glycerol, 1 mM DTT, and a panel of protease inhibitors. 0.6 ml portions of each cell suspension were pipetted into screw cap tubes containing ~0.5 ml acid washed glass

beads and disrupted in a 96 chamber bead beater with one 30 second pulse. The bottom of the screw cap tube was punctured with a 20 gauge needle and placed into a 2 ml eppendorf tube which was subsequently spun for 30 seconds at 2000 rpm in an eppendorf 5417C centrifuge to separate unbroken cells and extract from the glass beads. The punctured screw cap tube was removed and the 2 ml tube containing the lysate was centrifuged for 10 minutes at 14,000 rpm. Each 1 ml of supernatant was mixed with 0.5 ml of a 1:1 slurry of Protein A sepharose in the extraction buffer and incubated for 30 minutes at 4° in order to preclear the extract. The beads were pelleted by centrifugation and a portion of the precleared extract was set aside as an input sample. Immunoprecipitations were prepared using 100 μl of precleared extract, 1-5 μg of antibodies, 0.5 μl of 10 mg/ml ethidium bromide (BioRad), 10 μl of 1:1 Protein A Sepharose slurry, and buffer to a final volume of 200 μl then incubated on a tiltboard overnight at 4°. Immunocomplexes on the Protein A beads were recovered by centrifugation at 3000 rpm for 30 seconds and the supernatant aspirated with a 20 gauge needle attached to a vacuum trap. Beads were rapidly washed three times with starting buffer with each wash aspirated with the needle/vacuum trap. The last wash was aspirated extensively, and the beads were eluted with 20 μl 1X NuPAGE sample buffer at 75°. 2.5% of input and 50% of immunoprecipitation were loaded in equal 10 μl volumes onto 26 well precast 4-12% NuPAGE gels and run at 180 volts until the dye front reached the bottom of the gel, then electroblotted to PVDF membranes. Western blotting was done with antibody dilutions from 1:500 to 1:10,000, with all antibodies mixed and bound to membranes simultaneously.

Attempts were made to increase the stringency of IP conditions by altering salt and detergent concentrations. Salt concentrations of 100, 200, 300, 400, and 500 mM sodium acetate were systematically combined with Triton X100 at 0.01, 0.05, 0.1, 0.25, and 0.5 %. Salt concentrations greater than 300 mM negatively affected Rap1 stability, while those significantly less than 200 mM negatively affected TFIID stability and IP stringency. Detergent concentration greater than 0.5% Triton X100 disrupted TFIID integrity. Therefore a second titration of salt concentration was performed with 200, 225, 250, 275, and 300 mM sodium acetate and 0.2, 0.25, 0.30, 0.35, and 0.5 % Triton X100. Maximal stringency with acceptable Rap1 stability and TFIID integrity was achieved with 20 mM HEPES pH 7.6, 225 mM Sodium Acetate, 0.25% Triton X100, and 10% Glycerol, and this buffer condition was used for high stringency Rap1-TFIID coimmunoprecipitation using anti-Taf7 IgG.

Quantitative Protein-Protein Interaction Assays

The Far Western assay was initially used in attempts to characterize the binding of Rap1 to altered Taf5 forms. The yeast vectors containing all of the different *taf5* alleles were digested with XbaI/XhoI to recover full-length *taf5* genes. These fragments were ligated to pET28A digested with NheI/XhoI to yield bacterial expression vectors encoding full-length Taf5 forms containing amino acid substitutions. Full length Taf5 proteins were expressed and purified from inclusion bodies as described previously and tested in parallel for binding to Rap1 in the Far Western assay. A series of reproducible experiments using different Rap1 protein concentrations in the binding reactions failed to detect any quantifiable difference in Rap1 binding between the different Taf5 forms.

The main weakness of the Far Western in a quantitative role is its reliance upon film exposures to obtain final binding data, which can introduce the non-linearity associated with autoradiography and/or enhanced chemiluminescence. The conventional in vitro pulldown methods do not suffer this limitation if relatively pure proteins are used and bound proteins are visualized by staining of SDS-PAGE gels rather than western blotting. To this end Rap1 was expressed with an N-terminal GST tag (111 kDa protein) from pBG101, purified as described in Chapter 2, and immobilized on glutathione sepharose. To obtain soluble Taf5 forms, sequences encoding the amino terminal 337 residues were PCR amplified from yeast vectors containing *taf5* mutant alleles and inserted into pET28A as NdeI/XhoI fragments. Taf5, Taf5-17, Taf5-45, Taf5-408, or Taf5-10.4 proteins were prepared from 6 liters of autoinduced cultures per protein. Two 5 ml Ni-NTA columns were prepared for each protein. Soluble extract was incubated with one of the columns by batch binding for 1-2 hours at 4° C, the beads recovered by centrifugation, and the supernatant applied to the second 5 ml Ni-NTA column in the same manner. The first column was washed and recombinant Taf5 forms eluted by imidazole competition. Eluted material was kept at 4° C, and the first column was regenerated by treatment with 1 M hydrochloric acid for five minutes and subsequently re-equilibrated in the appropriate buffer. The supernatant of the binding reaction with the second column was then applied to the regenerated column. The second column was washed, eluted, and regenerated as well. This process was repeated iteratively until each column had been used three times for a total of six elutions, which effectively depleted the large majority of recombinant Taf5 forms from the soluble extract prepared from six liters of culture. The next day, the six elutions of each Taf5 form were pooled. Using a

Pharmacia FPLC system with a 50 ml superloop loading apparatus attached, pools of Ni-NTA-purified proteins were applied to a 5 ml Hi-Trap SP sepharose column, which was subsequently washed and subjected to gradient elution over five column volumes with buffer containing from 150 to 600 mM sodium acetate. Peak fractions were pooled and dialyzed against 25 mM HEPES pH 7.6, 150 mM sodium acetate, 0.1% Triton X100, and 30 % glycerol. The column was washed extensively and succeeding Taf5 variants were fractionated the same way. Taf4 variants were expressed from pBG100 (inserts subcloned as BamHI/SalI fragments from yeast expression vectors) alongside pACYC11b *TAF12* in BL21 DE3 RIL Strept, essentially as described previously except that each preparation began with cell pellets from 6 liters induced culture. A 3 to 5 fold molar excess of Taf5 or variants or Taf4/12 or variants were incubated with immobilized GST or GST Rap1 (10-20 μl of 1:1 resin slurry) in 100 μl binding reactions performed in Far Western binding buffer. Incubation was carried out at room temperature for one hour after which three 250 μl washes were performed and GST or GST Rap1 and bound proteins were eluted using 30 μl of binding buffer supplemented with 10 mM reduced glutathione. 33% of the eluted material from each binding reaction and 2% of the inputs were run on 4-12% NuPAGE gels and stained with Sypro Ruby, then visualized with the PharosFx imager (BioRad). As with the Far Western assay, this approach did not detect any difference in Rap1 binding between wild type forms and the variants. Systematic titration of buffer components and incubation conditions including time courses and incubation at different temperatures yielded either equivalent negative or uninterpretable results.

The failure of the Far Western and GST pulldown assays to yield results indicative of compromised Rap1-Taf interaction can be attributed to artifacts unique to each assay. For example, neither approach is carried out at the non-permissive temperature where growth and transcriptional phenotypes were observed in vivo; these assays require rather strict conditions to give interpretable results and those do not appear to be compatible with the 'heat shock condition.' For example, renaturation and Rap1 binding in the Far Western must be performed at 4° C to observe binding, and binding reactions in the GST pulldown could not be performed at temperatures above room temperature without compromising the interaction between GST and immobilized glutathione. Apparently what was needed was a technique with considerably increased throughput to identify optimized binding parameters and allow statistical power through use of replicate reactions; such a technique could be used along with the large quantities of pure Taf proteins and variants that I had already made for attempts at GST pulldowns. Dr. Scott Miller, a postdoctoral fellow in the lab, used the protein preparations I generated to set up the biolayer interferometry technique to study Rap1-Taf binding. The basis of this technique is the immobilization of a recombinant protein onto derivatized sensor probes that are present in 96 replicates mounted on a biorobot ('Octet' instrument, ForteBio Inc.). Each sensor probe has optical detectors that can emit fixed wavelength light and detect incidence of reflection. The composition and quantity of protein bound to the derivatized probes affects the angle of reflection of the detected light. Moreover, light reflecting properties of the derivatized probes can be monitored in real time, for example at time points after immersing the probes into buffer solutions containing different protein binding partners of the probe-immobilized test protein, or with different

concentrations of binding partners, different binding buffer composition, or binding conducted with varying temperatures. Systematic variation of binding partner protein concentration can allow the determination of rates of binding and dissociation and facilitate the construction of binding isotherms that allow extrapolation of rate constants for association and dissociation. The ratio of dissociation rate constant to association rate constant gives an indication of the strength of protein-protein interaction (i.e. K^D). Because sensor probes were coated with recombinant streptavidin, chemical modification of Rap1 by biotinylation allowed immobilization of Rap1 through the extremely high affinity interaction between streptavidin and biotin. Rap1 was biotinylated at one of several cysteine residues using a thiol reactive biotin derivative (Sulfo-link, Pierce). Note that Rap1 contains four cysteine residues including a dyad in the N-terminus and two separate residues in the C-terminus. Thus we did not anticipate that modification of these residues would have any affect on Rap1 function with respect to Rap1-Taf interaction, because the Rap1 N-terminus is neither required for growth nor for interaction with the TFIID complex, while modification of the C-terminal residues were not expected to affect Rap1-Taf interaction because these cysteines are located in the silencing domain, which also did not appear important for Rap1-Taf interaction. Nevertheless, modified Rap1 was tested in DNA-binding gel shift assays, and DNA binding activity did not appear any different from the unmodified protein. Therefore constant amounts of Rap1 were used in titration binding reactions where the concentration of non-binding Taf3, Taf4/Taf12 heterodimers, the Taf5 amino terminus, or the variants of these Tafs were monitored in real time for rate of association with and dissociation from Rap1. This allowed comparison of binding properties of Tafs and altered derivatives, and in addition

to allowing more rapid optimization of binding conditions, the 96 well format allowed statistical analyses because many replicate reactions could be done in parallel. Data were analyzed using the software package included with the Octet instrument. All of this work was performed by Dr. Miller.

Synthetic Genetic Interaction Studies

To create strains containing multiple deletions in different Tafs or Rap1, different chromosomal knockout markers and covering plasmids had to be created. These reagents were used to create haploid strains with different mating types along with the different knockout markers. W303 was chosen for these strain constructions because it contains the most nutritional selection options, and because the strain has superb transformation efficiency which helps both in creating the initial knockout strains, and subsequent introduction of plasmids. W303a was used to make a chromosomal *taf5* knockout plasmid shuffle strain as described previously, the knockout was marked by G418 resistance. This strain was transformed with *HIS3*-marked plasmids containing *TAF5* or *taf5* conditional alleles and subjected to plasmid shuffle. In parallel W303α was used to make a chromosomal *taf4* or *rap1* knockout plasmid shuffle strain where the knockout was marked by Hygromycin resistance. The W303α *taf4* shuffling strain was transformed with *TRP1*-marked plasmids containing *TAF4* or *taf4* conditional alleles, while the W303α *rap1* shuffling strain was transformed with *LEU2*-marked plasmids containing *RAP1* or *rap1* mutant alleles. Each of the shuffled, kanamycin-resistant W303a *taf5* strains was then transferred into liquid culture (SC-His) as were each of the hygromycin resistant W303α *taf4* or *rap1* strains (these were grown in SC-Trp-Ura or

SC-Leu-Ura, respectively). Crosses were set up at room temperature by mixing a portion of each *taf5* strain culture with each *taf4* or *rap1* strain culture and incubating overnight. The next morning an appropriate serial dilution was plated onto YPD containing hygromycin and G418 at the appropriate concentrations and the plates were incubated at room temperature. Upon appearance of defined single colonies, each plate was replica-plated onto another fresh plate containing hygromycin and G418 to ensure that all colonies consisted entirely of diploid cells. A single colony of each diploid strain was picked and mixed into 0.5 ml distilled water. 100 μl portions were then spread across each of four plates containing SC-Trp-Ura or SC-Leu-Ura, for *taf5 taf4* strains or *taf5 rap1* strains, respectively. Upon appearance of large, defined single colonies, each plate was replica-plated to sporulation media plates supplemented with 0.5% w/v potassium acetate and 0.25 M sorbitol (included both to select against non-sporulated cells but prevent excessive osmotic shock of conditional mutants). After several days of incubation, each of these plates was overlaid with a sterile buffered solution containing 0.5 M sorbitol and 0.1X lyticase. The overlay solutions for each strain's four plates, containing partially digested spores, were pooled and subjected to mild sonication. The sonifier probe was rinsed thoroughly with ethanol before sonication of each strain suspension. Appropriate dilutions of the solutions containing disrupted asci were plated onto YPD media containing hygromycin and G418 and incubated at room temperature. These represented master plates used for a series of replica plating phenotypic analysis that tested for temperature sensitivity (indicative of recessive *taf5* allele presence and absence of *TAF5*), His^+ phenotype (*taf5* plasmid), and Trp^+ Ura^+ (*taf4* and *TAF4* plasmids) or Leu^+ Ura^+ (*rap1* and *RAP1* plasmids). Temperature sensitive strains that

were His^+, Trp^+, and Ura^+ (*taf5 taf4*) or His^+, Leu^+, and Ura^+ (*taf5 rap1*) were genotyped for ploidy and mating type by PCR. Because the proteins encoded by the plasmid-borne *taf5*, *taf4*, and *rap1* alleles were epitope tagged, resulting in a mobility shift on SDS-PAGE, western blotting with polyclonal antibodies against each protein served as another indirect indicator of *taf5, taf4*, or *rap1* genotype and strain ploidy. To test for synthetic growth phenotypes, strains containing the above growth and nutritional phenotypes and that were haploid/*Mat a* were expanded into room temperature liquid cultures containing SC-trp or SC-leu media, grown for 5-7 days to allow for loss of the *URA3*-marked plasmids containing either *TAF4* or *RAP1*, and subjected to serial dilution growth analysis on 150 mm plates with appropriate media containing or lacking 5-FOA; identical plates were prepared and incubated in parallel at 23°, 30°, and 37° C. At least two independent clonal isolates were tested to confirm the evidence of synthetic growth phenotypes. In total more than 52 unique strains were created; isolation of multiple clones containing all of the desired attributes mandated examination of 100-200 individual colonies on average thus more than 10,000 colonies were screened to complete these experiments.

RESULTS

Neither the SAGA Nor Mediator Complexes Contribute to RPG Transcription

The purpose of identifying specific domains in Tafs responsible for interaction with Rap1 was to inform a directed mutagenesis/genetic test of Taf function that would hopefully not affect other Taf/TFIID functions. However, in addition to having several

distinct roles in the context of TFIID function, both Taf5 and Taf12 contribute to gene regulation as subunits of the SAGA complex (Grant et al., 1998). Thus it is possible that Taf5 and Taf12 function as coregulators for Rap1 not only as members of the TFIID complex, but potentially as members of SAGA, though this has not been explicitly tested and reported in the literature. Clearly, if SAGA were a contributor to RPG regulation, it would be difficult to differentiate the molecular mechanism of any transcriptional defect found in *taf5* mutants, with respect to whether TFIID or rather SAGA function were compromised. Moreover, there are many multisubunit complexes that could theoretically fulfill a coregulatory role on RPGs; examples include the Mediator complex and the NuA4 histone acetylase, a complex that has been documented to directly regulate RPGs (Reid et al., 2000). The participation of these other complexes in RPG transcription would complicate interpretation of TFIID function, particularly the role of TFIID in direct activator interaction, since Mediator and SAGA have been shown to perform their function at least in part via direct contact with activators. Fortunately, the majority of genes encoding subunits of Mediator and SAGA are present in single-copy and are non-essential for cellular growth; a characteristic that greatly simplified genetic tests of their contribution to RPG transcription. Furthermore, strains containing deletions of most of these subunit-encoding genes are commercially available. These strains and their cognate wild type control were grown to mid-log phase, harvested by filtration and RNA was extracted. The method used for measurement of specific transcript abundance was the 5' primer extension method that is well suited to my purposes for several reasons (Ghosh et al., 1978; McKnight et al., 1981). First, although the technique is somewhat less sensitive than other radioactive detection methods (northern blot, RNase protection, S1

nuclease), that is less of a concern when studying RPG transcripts because these are among the most abundant mRNAs present in a rapidly growing yeast cell and thus easily scored. Second, like the other radioactive methods, primer extension lends itself to processing numerous samples in parallel, because cross contamination of reaction samples is much less of a concern than with RT-PCR because there is no exponential amplification of target signal. Third, the method gives rather crisp raw data due to the use of DNA sequencing gels so bands are quite sharp, similar to results obtained with RNase protection or S1 nuclease assays. Moreover, probe preparation is somewhat more straightforward than RNase protection or RT-PCR, because a synthetic DNA oligo needs only to be labeled by 5' phosphorylation using T4 polynucleotide kinase. One more advantage is that by definition, primer extension provides not only quantitative information about transcript abundance, but also qualitative data as the method scores the location of the transcription start site. It can be envisioned that the mechanism of Rap1-TFIID interaction/TFIID function might involve an affect on start site selection. Primer extension provides this information by default, although RNase protection and S1 nuclease assays can also map start sites but must be deliberately set up to do so. Note that all choices of methodology do not measure transcription directly since they score only steady state abundance rather than synthesis rates per se. However, the initiation rate of RPG transcription is quite high and RNA half-life is quite short, thus these steady state measurement methods are a reasonable proxy of initiation rates for RPG mRNAs.

As scored by primer extension, the abundance of several transcripts within Mediator and SAGA deletion strains is shown in **Figure 3.1**. *RPS2, RPS3, RPS5*, and *RPS13* are from RPGs while *PGK1* was analyzed as a control that is TFIID-independent

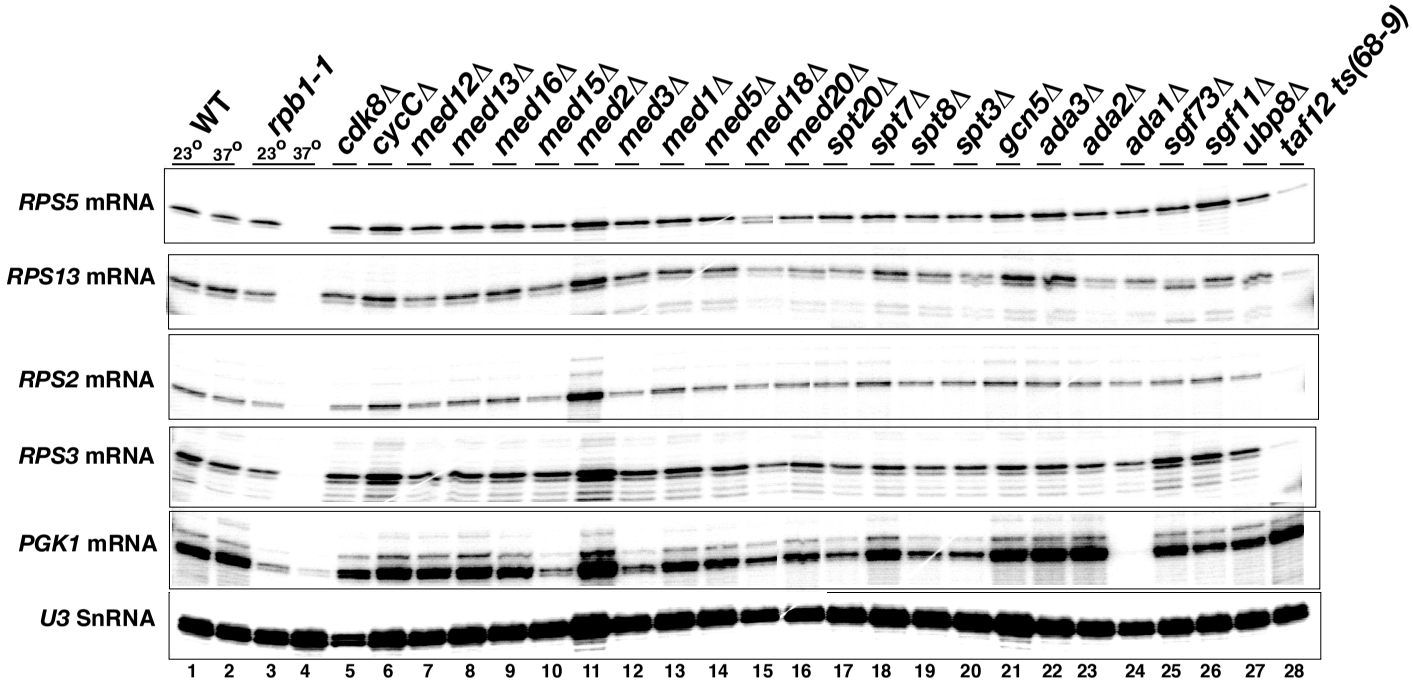


Figure 3.1: Genetic dependence of RPG transcription on on Mediator and SAGA subunits. The strains indicated, that contained deletions of individual subunit-encoding genes, were grown at room temperature in rich media (YPAD) until mid-log phase, cells were collected by rapid filtration and RNA was prepared immediately by organic extraction. Primer extension with individual probes indicated was used to score transcript abundance in an equivalent amount of RNA from each strain. Control strains wild type, Pol II mutant *rpb1-1*, and the *taf12 ts* mutant behaved appropriately. In contrast to the control mutant strains none of the Mediator (lanes 5-16) or SAGA strains (lanes 17-27) exhibited deficiencies in *RPS5, RPS13, RPS2*, or *RPS3* mRNAs, while *U3* was also fairly consistent, as expected. On the other hand, *med15*Δ, *med3*Δ, and *med18*Δ strains had reductions in *PGK1* mRNA, as did the *spt20*Δ, *spt8*Δ, *spt3*Δ, and *ada1*Δ strains. Thus these data collectively indicate that Taf-dependent RPG transcription is Mediator- and SAGA-independent, while Taf-independent *PGK1* transcription is Mediator- and SAGA-dependent. In light of the sharing of certain Tafs between SAGA and TFIID, RPG transcription can be productively studied through the lens of genetic *TAF* contribution occuring exclusively through TFIID-dependent pathways. Data are adapted from Layer et. al., 2010.

(but possibly SAGA and/or Mediator dependent). The very abundant U3 transcript is TFIID- and SAGA-independent and provided a useful loading control. Wild type cells were grown at low and high temperatures to provide a control for the *rpb1-1* strain; all mRNA transcription is compromised in this temperature sensitive strain at high but not low temperature, with a mutation affecting the largest Pol II subunit (Nonet et al., 1987a). The *taf68-9* temperature sensitive strain which encodes an altered form of TFIID- and SAGA-shared Taf12 was also analyzed as a control since this mutation could possibly affect both SAGA- and TFIID-dependent transcription (Reese and Green, 2003). Twelve different viable Mediator subunit deletion strains were analyzed and are labeled according to the submodule of the holocomplex in which each subunit is believed to reside (Bourbon et al., 2004). Eleven different viable SAGA deletion strains were analyzed that collectively represent the majority of non-Taf SAGA subunits. As expected U3 transcript levels were fairly even from strain to strain. Each RPG transcript was also fairly even irrespective of strain background, except for deficiencies in the *rpb1-1* mutant and the *taf68-9* mutant, which were grown at both permissive and non-permissive temperature. Use of these two mutants confirms the involvement of Pol II and Taf12 in RPG transcription, although since little to no effect is seen in any of the SAGA mutants, it seems likely that Taf12 contribution is occurring through its participation as a member of the TFIID complex and not SAGA. Likewise, there is at best a minor contribution from Mediator subunits to RPG transcription. On the other hand, Mediator seems to function in *PGK1* transcription, since mRNA levels are notably reduced in the *med15, med3*, and *med18* strains (lanes 10, 12, and 15, respectively). SAGA subunits Spt20, Spt8, Spt3, and Ada1 (lanes 17, 19, 20, and 24) all appear to be

important for *PGK1* transcript levels with the *ada1* strain showing the most profound affect. Collectively these data clearly argue against any significant direct contribution from either Mediator or SAGA to RPG transcription. These results set the stage for testing the affect of *taf4* and *taf5* mutants in RPG transcription.

Reduced Ribosomal Protein Gene Transcripts in *taf4* and *taf5* Mutants

Before embarking on a detailed gene-by-gene analysis of transcript abundance in *taf4* and *taf5* mutants, I decided to grow the strains, perform the temperature shifts, prepare the RNA from unshifted and shifted cultures, and carry out a bulk analysis of changes in total cellular RNA abundance. The logic behind this decision was that because RPG transcripts represent ~33% of cellular mRNA, any profound affect on RPG transcripts would cause a noticeable reduction in the cellular mRNA pool of polyA$^+$ (Warner, 1999). Total cellular mRNA is easily detectable through the 3' polyA tail by use of hybridization with end-labelled oligo dT (Reese and Green, 2003). The *rpb1-1* strain was again used as a positive control since the mutation causes a cessation of all transcription at the non-permissive temperature. Tony Weil helped with these experiments and found that there was a noticeable reduction in total polyA$^+$ mRNA content in each of the *taf4* and *taf5* temperature conditional mutants. RNA was prepared from 3 different *taf4* mutants, 30 different *taf5* mutants, and several positive and negative controls in addition to wild type and *rpb1-1* strains; characterization of each RNA preparation was achieved using the convenient oligo dT slot-blot method and a range of loss-of-function phenotypes was noted (not shown). Having seen this affect, I decided to check an RPG transcript in these same RNA preparations. Using one-tube multiplex

primer extension with probes for *U3*, *PGK1*, and again for *RPS5*, I analyzed all 45 strains (two wild type, three *taf4* mutants, thirty *taf5* mutants, and several control strains). *U3* signal was constant as expected and so was *PGK1*, while there was a mild reduction in *RPS5* transcripts in *taf4* mutants. On the other, every single *taf5* mutant showed a dramatic reduction in *RPS5* but not *PGK1* transcripts, indicating an essential role of the Taf5 amino terminus in RPG transcription (not shown). To extend this analysis, and to make these experiments more manageable, I decided to carry forward using only a subset of the *taf5* mutants, since all of them appeared to confer a loss of RPG transcripts at the non-permissive temperature. It was convenient to analyze a couple positive control strains, the negative control wild type strains, the three available *taf4* mutants, and four of the *taf5* mutants. The multiplex primer extension assay was extended to include probes for *RPS2, RPS3*, and *RDN1*, in addition to *U3, PGK1*, and *RPS5*. *RDN1* promised to provide useful information about off-target affects, since this RNA is synthesized by Pol I and thus is TFIID-independent, however, there is crosstalk between RPG transcription and *RDN1* transcription as they all encode constituents of the ribosome; the probe used for *RDN1* also scored the 5' end of the immature transcript which is rapidly removed co- and post-transcriptionally therefore as monitored *RDN1* RNA abundance is a very close approximate of transcription rates. Representative results are shown in **Figure 3.2**. Note that it is important to compare transcript levels from 37^{o} samples (even numbered lanes) to the equivalent analyses for the wild type strain (lanes 2 and 10). Elevated temperature caused a mild reduction in *RPS5* transcripts in *taf4-116* and *taf4-219*, but not in *taf4-141*. Quantification of signals were normalized to *U3* abundance and represented as a percentage of the wild type at 23^{o}, indicating a reduction from 40 to 60% in the two

affected mutants (bar graph data). Note that these strains had elevated *RPS5* mRNA levels compared to wild type when grown at low temperature, so the relative reduction is actually considerably greater than wild type. The same basic pattern was evident when examining *RPS2* or *RPS3* mRNA levels (quantification not shown). *RPS2, RPS3*, and *RPS5* are all much more affected in the each of the *taf5* mutants, although not quite as much in the *taf5-408* mutant. Quantification indicates a 50 to 80 percent reduction in *RPS5* transcripts. Importantly, in no case were there reductions in *RDN1, U3*, or *PGK1* levels, consistent with the phenotypic affect having selectivity as to which genes are affected by *taf* mutants. Note that these experiments were performed a number of times with the use of independently grown cultures and different RNA preparations. The data were qualitatively remarkably similar from experiment to experiment, but differences in probe labeling and gel handling rendered some differences in exact quantitative similarity, which obviated the inclusion of error bars with the graphical data. Nevertheless, the fact that different strains recapitulated the same phenomenon amongst the same group of transcripts again and again strongly supports the existence of a bona fide *RPG* transcriptional defect in the *taf* mutant strains. Also consistent with these observations, an extension of these experiments revealed reductions in transcript levels of *RPS12, RPS13, RPS15, RPS20, RPL3, RPL5, RPL28, and RPL30.* Non-RPG transcripts were also analyzed, specifically *HHT3, ADH1, ADH3, TPI1*, and *SCR1*. Of these only *TPI1* exhibited a reduction in transcript abundance, consistent with the lack of an RPG-exclusive dependence on the Taf4 and Taf5 domains affected by the mutations.

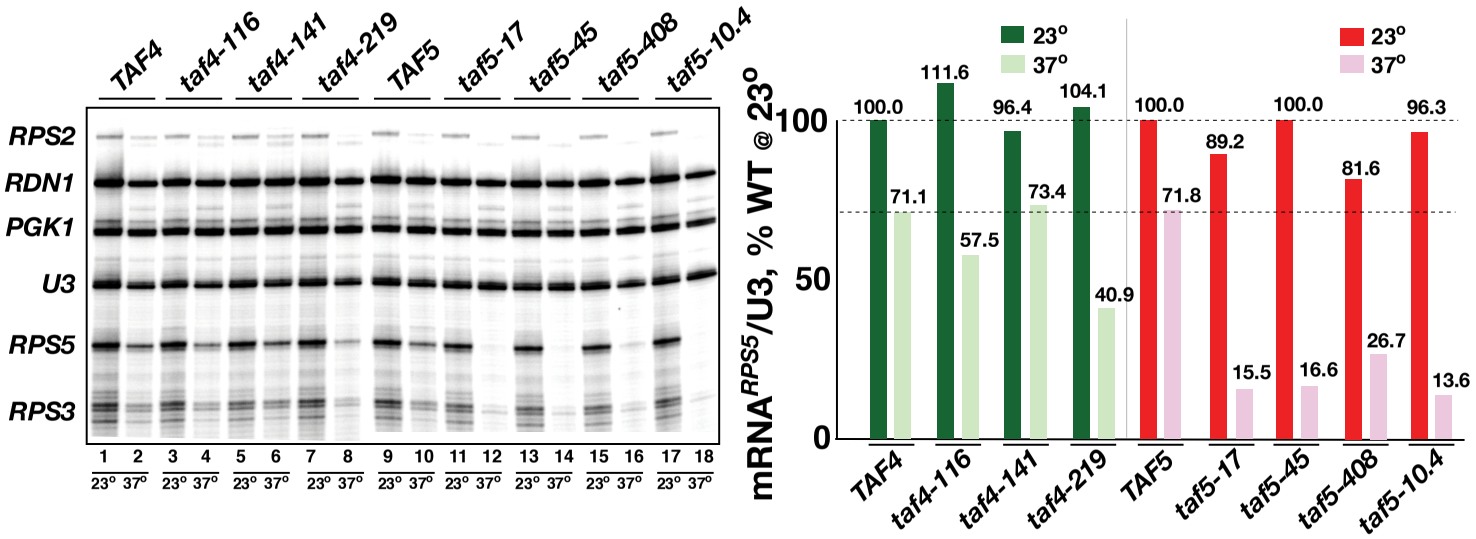


Figure 3.2: RPG transcript levels in *taf4* and *taf5* ts^+ mutants. Duplicate cultures of each strain were grown in parallel to mid log phase in YPAD at room temperature. One culture was rapidly shifted to 37° while the other was retained at room temperature and incubation was continued for two hours. Cells were harvested by rapid filtration and immediately subjected to organic extraction. An equivalent amount of total RNA from each strain and condition was analyzed by multiplex 5' primer extension including probes to six different transcripts, 3 RPGs and 3 controls (*U3, PGK1, RDN1*). The reduction in RPG transcript abundance as a result of heat shock is evident in the wild type strains. A more significant reduction between room temp and 37° occurred in the *taf4-116* and *taf4-219* mutants as compared to wild type (autoradiograph at left, quantification at right). As indicated by the quantification graph, *RPS5* mRNA dropped 54.1% and 63.2% respectively, as compared to 38.9% in the cognate wild type. *taf4-141* exhibited no apparent defect in RPG transcript abundance. Every *taf5* ts^+ mutant exhibited a much more severe phenotype, with reductions from 83.4% to 54.9% as compared to 38.2% in the cognate wild type, with *taf5-45* being the most affected and *taf5-408* the least. All three RPG transcripts behaved in a quantitatively similar manner, and *PGK1* abundance actually increased in every strain at 37° (not shown). Results are representative of three or more completely independent experiments. The RPG transcription defect in the *taf* ts^+ strains is consistent with defective Rap1-Taf interactions occuring as a result of the introduced mutations and amino acid substitutions.
Data are adapted from Layer et. al., 2010.

The limitation of these experiments is that this approach is not readily amenable to looking at all 137 RPG transcripts, or to monitor the large number of non-RPG transcripts. Also, these assays do not lend themselves to precisely quantifying phenotypic affects with statistical significance. Microarray analyses can do all of the above with relative ease if the experimental format is chosen properly. Oligonucleotide arrays are especially useful if they are designed to provide good specificity and sensitivity; both can be achieved by providing multiple probes per transcript (Gasch, 2002). We therefore chose to perform such experiments and opted to use Nimblegen 12-plex arrays because these arrays utilize from 4-6 probes per transcript for ~5700 of the ~6200 yeast genes. Moreover these arrays have 12 individual microarrays per slide so up to 12 independent samples can be monitored in parallel. We included biological replicates for all strains and conditions to strengthen any potential statistical analyses beyond what is automatically performed with the software package obtained with purchase of the microarray. One set of biological replicates was generated by Tony Weil and one by myself. We first analyzed the *taf* mutants along with the equivalent wild type strain. The ArrayStar program (DNA Star, Inc.) allowed us to create a hierarchical clusterogram which is shown in **Figure 3.3**. Presentation of data was achieved by comparing the 23° wild type data with the 37° data of the *taf5-17* mutant and selecting those transcripts that exhibited a two-fold up or down change at the 95% confidence interval. These 1347 gene transcripts were ranked from highest to lowest abundance from left to right, and their abundance was compared amongst each of the *taf5* mutants grown at 23° and 37°. The RPG transcripts, the most abundant class in the cell, grouped into a tight cluster at the left of the panel. Signal intensity in this region of the cluster,

which included 110 of the 137 RP genes, dropped significantly at 37° as compared to 23° in each of the mutants as compared to the effect seen in wild type. Note that the overall change in signal intensity was less dramatic in the *taf5-408* mutant as compared to the other mutants, in agreement with the gene-by-gene analysis presented in **Figure 3.2**. It is clear that many transcripts were elevated with heat shock in the mutants just as well as wild type, indicating that these genes do not depend on the affected Taf5 domain for their transcription. However, many more genes other than RPGs were negatively affected by heat shock, indicating a more widespread role for this Taf5 domain in stimulation of transcription. Additional data analyses have not been performed to characterize these phenomena in more detail. This set of microarray experiments fulfilled their main objective, which was to test if there was an affect on the entire RPG regulon, as would be expected if Rap1-Taf5-TFIID interaction is important for transcription of RPGs. It remained to be seen if the biochemical explanation for the affect seen in the *taf5* RBD mutants was truly due to defective Rap1-TFIID interaction.

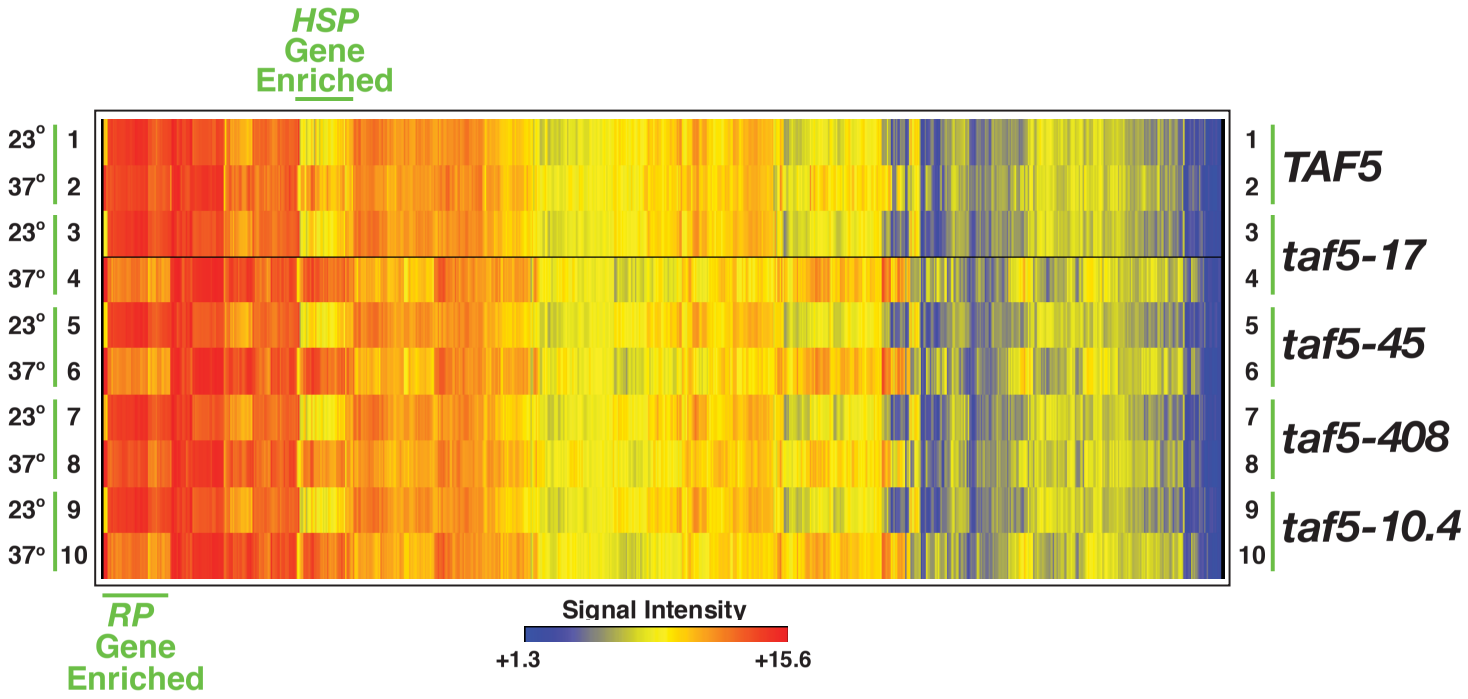


Figure 3.3: Transcriptome analyses in wild type and *taf5 ts*$^{+}$ mutants at permissive and non-permissive temperatures. Microarrays were performed using RNA from the experiment shown in Figure 3.2, labeled with Cy3 or Cy5 and hybridized to Nimblegen 12-plex oligonucleotide arrays containing an average of six unique probes for ~5700 of the ~6200 yeast genes. Computational analysis selected the transcripts that showed 2-fold or more abundance change between the permissive and non-permissive temperature in the *taf5-17* mutant. The abundance of this set of transcripts was then compared between all the strains and conditions and used to form the HEAT map shown above. 1347 transcripts met these criteria and vary in abundance over a fifteen-fold range, with abundant transcripts shown in red, intermediately abundant transcripts in yellow, and less abundant transcripts in blue. Although this cluster analysis was essentially unsupervised, 110 of the 137 RPG transcripts clustered together and showed a consistent drop in abundance at the non-permissive temperature as compared to wild type. Many other transcripts were affected negatively, and many inducible transcripts were not mis-expressed in the mutants. Collectively these results are consistent with a negative affect on RPG transcript abundance, across the entire regulon, in *taf5 ts*$^{+}$ mutants, with affects also occuring on non-RPG transcripts in the *taf5 ts*$^{+}$ mutants, and other genes being completely *TAF5* independent. Most importantly, these data are consistent with a direct Rap1-TFIID interaction occurring on many of the coordinately-regulated Ribosomal Protein Genes, and with a loss of productive interaction occurring in the *taf5 ts*$^{+}$ mutants at the non-permissive temperature. Note that further analyses of this data are pending. Data are adapted from Layer et. al., 2010.

Rap1 Interaction with TFIID Containing Altered Taf5 Appears Unaffected

Several methods were considered for testing the interaction of Rap1 with TFIID. The ideal technique would monitor interactions in vivo without disruption of cells. The FRET technique has this property. Unfortunately, neither I nor any lab colleague had experience with FRET, and moreover I knew that large epitope tags could potentially cause problems when appended to the Rap1 C-terminus, the Taf4 C-terminus, or the Taf5 N-terminus, which are the preferred locations in each protein where a ECFP/EYFP tag would have the best chance of documenting FRET interactions. We had previously used an in vitro pulldown technique using immobilized TFIID complex purified from yeast to document interaction with purified recombinant Rap1 (Garbett et al., 2007). Using this method, it was possible to get rather clean protein preparations to use for Rap1-TFIID interaction assays and even provide some quantitative data about the Rap1-TFIID interaction. The logistics of performing these experiments are intimidating with respect to the skill and effort needed to get this working. Nevertheless I took the previous success to indicate that it might be possible to adapt the in vitro Rap1-TFIID pulldown assay to include the use of TFIID complexes containing altered Taf4 and Taf5, purifed from the mutant strains. First the *TAF1* locus had to be genetically modified to encode an N-terminal HA tag to allow affinity purification of TFIID by anti-HA immunoaffinity chromatography. I created a construct to allow this genome modification and transformed it into BY4741 *taf4* and SEY6211 *taf5* shuffling strains. The tagging construct was marked by hygromycin resistance while the *taf4* and *taf5* knockout markers were kanamycin resistance. Correctly tagged strains were chosen from hygromycin-resistant clones by immunoblotting with anti-HA antibody and looking for a single band

at the mobility of Taf1. The *taf4* and *taf5* temperature conditional alleles were reconfigured to remove the epitope tags and were then shuffled into the modified shuffling strains. After confirming temperature sensitivity of the resulting strains, single colonies were progressively expanded into larger and larger liquid cultures. Finally, a cell pellet from 36 liters of 23°-grown cultures was obtained for each of six temperature conditional strains (*taf4-116, taf4-219, taf5-17, taf5-45, taf5-408, taf5-10.4*) and each cognate wild type (total of 192 separate 1.5 liter cultures). Transcriptionally active whole cell extracts prepared from each cell pellet and a portion of each of these extracts was subjected to large scale cation-exchange chromatography on BioRex70 resin (Woontner et al., 1991). The next step of the purification scheme was to perform the immunoaffinity anti-HA step, but I was unsuccessful in this part of the purification and failed to obtain any Tafs/TFIID from the BioRex70 fractions. Therefore, I did not get the TFIID preparations needed to perform in vitro pulldown assays between Rap1 and immobilized TFIID. However, I did document the transcriptional activity of the whole cell extracts, which were consistently active for transcription by RNA Pol I and RNA Pol III, but showed deficiencies in Pol II transcription in extracts prepared from *taf4* or *taf5* mutant strains (not shown) (Klekamp and Weil, 1982; Schultz et al., 1991). These experiments indicated that the phenotypes seen in vivo were likely caused by a true defect in transcription initiation, as would be expected if TFIID/Taf function was the compromised activity.

Thus I turned to co-immunoprecipitation of Rap1 with TFIID using extracts prepared from wild type and mutant strains. I decided to focus on the *taf5* strains because they exhibited a more dramatic transcriptional phenotype, and because of this I suspected

that Rap1-TFIID interaction was potentially compromised to a greater extent in these strains. But first I prepared antigen-affinity purifed IgG to many of the TFIID subunits and Rap1 by using existing stocks of whole anti-serum and expressing and purifying full-length antigens along with ammonium sulfate precipitation and Protein A affinity chromatography prior to antigen-affinity chromatography on Ni-NTA-immobilized antigens. This gave me an indispensable asset of polyclonal antibody preparations with excellent sensitivity and selectivity in both immunoprecipitation and immunoblot assays, allowing avoidance of the artifacts that can be introduced by epitope tags. More importantly, I hoped to systematically monitor both TFIID integrity and association with Rap1 using this comprehensive panel of reagents. Initial IP experiments with extracts from heat-shocked cultures revealed consistent proteolytic degradation of Tafs in extracts of wild type cells (not shown), so a semi-permissive growth temperature of 30^{o} was chosen because the cells grew, albeit slowly, at this temperature and thus it was possible that Rap1-TFIID interaction was appreciably reduced in a manner that could be successfully documented. **Figure 3.4** shows results of these IP reactions. Rap1-Taf Co-IP was documented in extracts from wild type and three *taf5* mutants. Immunoprecipitation was performed with negative control IgG to show specificity and 14 antibodies to account for almost all TFIID subunits. These immunoprecipitation reactions included ethidium bromide to exclude the possibility of protein-DNA interactions contributing to Co-IP, as such any co-precipitation is due to protein-protein interaction (either direct or indirect, (Lai and Herr, 1992)). In extracts prepared from wild type cells, Taf1, Taf5, Taf4, and Taf10 reproducibly co-precipitated, consistent with each protein existing as a member of the TFIID complex in these extracts. Rap1 also co-

precipitated with every single Taf, again consistent with Rap1 engaging the TFIID complex in vivo. Note that there are separate bands with different mobility detected with polyclonal anti-Rap1 antibodies, indicating that multiple Rap1 isoforms are actively engaged with TFIID. An additional Rap1 species of further increased molecular weight was also present in some but not all of the immunoprecipitation reactions (not shown) although I suspect it is present in all reactions but is outside of the detection limits of the assay in current configuration. These high molecular weight Rap1 forms suggested that a portion of the TFIID-associated Rap1 contained one or more post-translational modifications, such as phosphorylation or ubiquitination. Note that enrichment of Rap1 varied depending on the IP antibody used; the strongest enrichment occurred when using antibodies to Taf5, Taf6, Taf10, Taf12, and TBP; all of which are constituents of the SAGA coregulatory complex. Note that the reciprocal interaction did not occur; IP with anti-Rap1 antibodies enriched Rap1 but not any detectable Tafs (not shown). This is likely due to the relatively low proportion of total cellular Rap1 that is involved in gene transcription; the majority of Rap1 is bound to repetitive elements near each of the 32 telomeres, which are transcriptionally silent and depleted of coregulators (Negrini et al., 2007). Related to this observation of a lack of reciprocal immunoprecipitation, the altered mobility Rap1 isoforms did not appear as enriched when immunopurifying Rap1 using anti-Rap1 IgG, as compared to isolation of Rap1 indirectly via Co-IP with anti-Taf IgG, again suggesting that these Rap1 isoforms are preferentially associated with TFIID. Immunoprecipitations from wild type extracts were done in parallel with reactions from mutant strain extracts. There was essentially no change in the pattern of Taf co-precipitation, indicating that TFIID integrity is maintained in the mutant cells.

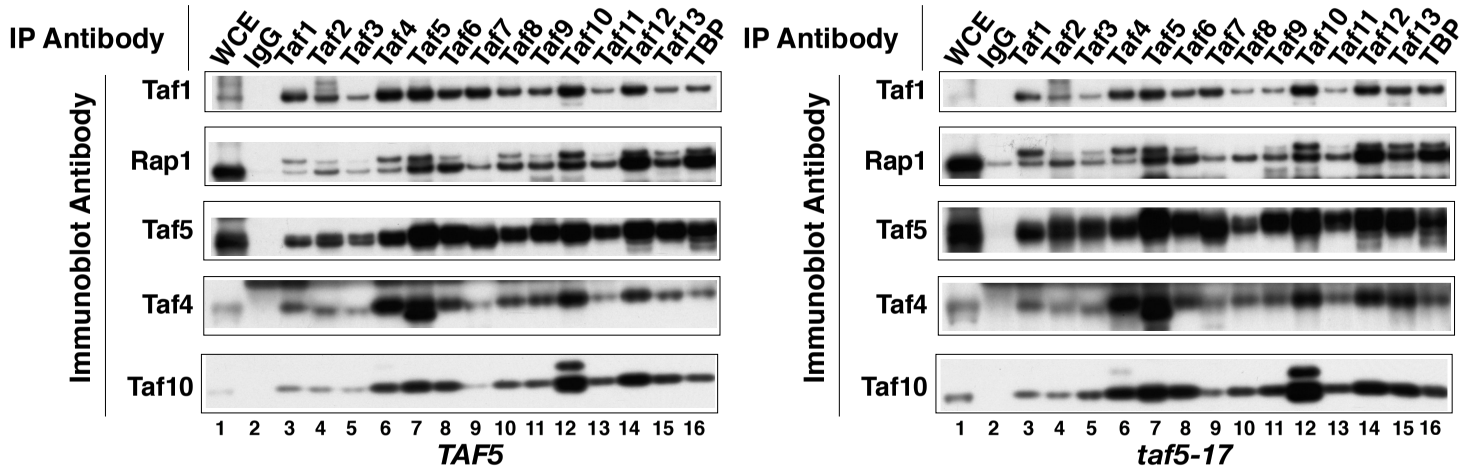


Figure 3.4: Co-immunoprecipitation of Rap1 with TFIID subunits in extracts from wild type and a *taf5 ts*$^{+}$ mutant. The indicated strains were grown in YPAD until mid-log phase at 30^{o}, a semi-permissive temperature for the conditional mutants. Soluble extracts were prepared immediately by glass bead lysis in buffer containing 20 mM HEPES pH 7.6, 200 mM Sodium Acetate, 0.1% Triton X100, 10% Glycerol, 1 mM DTT, protease inhibitors, and 25 ng/μl ethidium bromide (to block DNA-dependent co-IP). Soluble extract, antibodies, and protein A-sepharose beads were used to prepare immunoprecipitations, which were then analyzed by immunoblotting. As expected, all Tafs co-precipitated in extracts from both wild type and the mutant, consistent with conservation of TFIID integrity in the mutant strain extract. Rap1 also consistently co-precipitated with every Taf, and several bands of varying mobility and all attributable to the Rap1 antibody were observed. This is consistent with Rap1-TFIID interaction in yeast extracts, and the participation of Rap1 isoforms in the interaction. The Rap1 isoforms probably correspond to both multiply phosphorylated and/or ubiquitinated or sumoylated derivatives. Most notably, the pattern of Rap1 interaction with Tafs was essentially identical in extracts of both strains; a result consistent with a multivalent mode of interaction between Rap1 and TFIID, involving multiple Tafs. Results did not vary when different *taf5* mutants were used as source of extract (not shown). All data are derived from the exact same film exposure and are 100% representative of five complete, independent experiments. Data are adapted from Layer et. al., 2010.

However, there was also no apparent difference in the pattern and amount of Rap1 co-precipitation with Tafs. This result indicated that Rap1 interaction with TFIID and Tafs was not significantly reduced in vivo in any of these mutant strains under the experimental conditions used. To test the possibility that the extraction and IP conditions were not sufficiently stringent to discriminate between full-strength and weakened interactions between Rap1 and TFIID, buffers constituents were systematically titrated in conjunction with IPs using the anti-Taf7 antibody. Maximal salt and detergent concentrations that allowed both preservation of Rap1-TFIID interaction and Rap1/TFIID integrity were identified. Using these buffer conditions, the temperature of IP incubation was varied from 4° to 30° in a further attempt to reduce weakened Rap1-TFIID interactions. The results are shown in **Figure 3.5**. The amount of Rap1 co-precipitating with anti-Taf7 antibodies did not appreciably change in any strain extract or condition except in extracts of the *taf5-408* mutant where decreased Rap1 co-precipitation occurred at all temperatures tested. However, there was also a reduction in the quantity of Taf1 co-precipitating with anti-Taf7, which indicate a reduction in Taf1 stability and/or TFIID complex integrity. Thus it is not possible to state that Rap1 association with the intact TFIID complex was truly affected in this particular mutant. In conclusion Rap1 interaction with TFIID did not appear to be affected in the *taf5* mutants examined using a very thorough analysis by co-immunoprecipitation assays. The possibility of not having the correct experimental conditions remained as an explanation although this was pursued to the best of my ability given the time constraints. The problem of redundancy between different Rap1-binding Tafs present within TFIID also offered a very legitimate explanation for the failure to document defective Rap1-TFIID interaction.

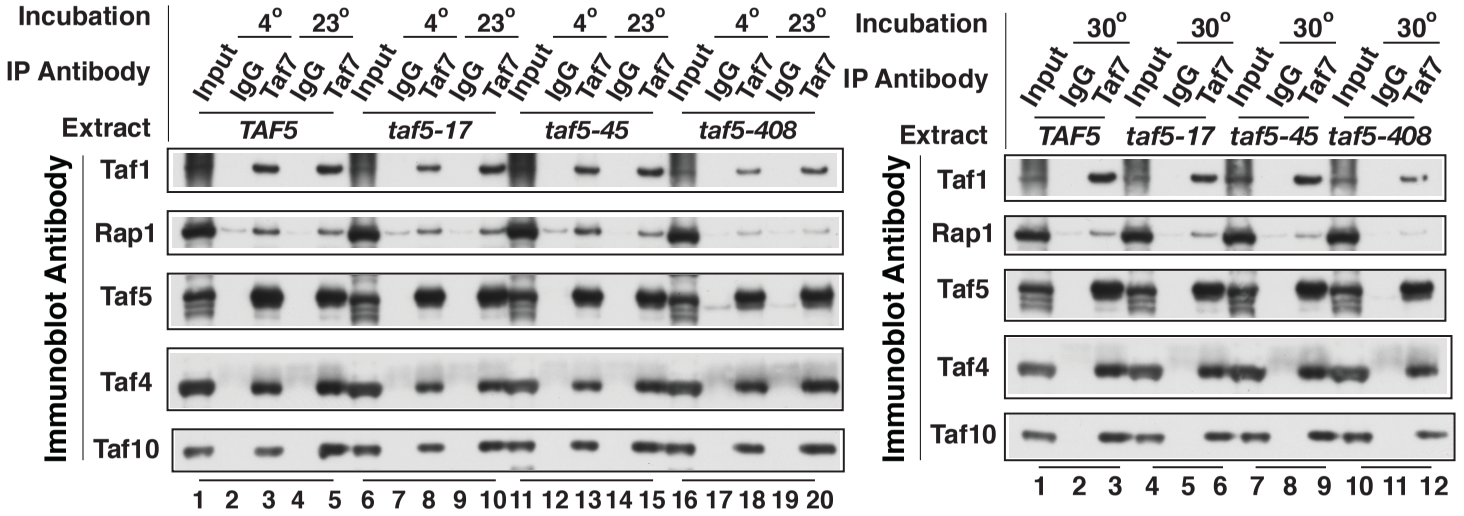


Figure 3.5: Co-immunoprecipitation of Rap1 with TFIID subunits in extracts from wild type and *taf5 ts*$^{+}$ mutants, under higher stringency buffer conditions. Soluble extracts were prepared exactly as in Figure 3.4 except buffer contained 20 mM HEPES pH 7.6, 225 mM sodium acetate, 0.25% Triton X100, 10% Glycerol, 1 mM DTT, protease inhibitors, and 25 ng/µl ethidium bromide. Taf7 antibody was used exclusively for IP, and IP reactions were performed at 4^{O} as previously, and also at 23^{O} and 30^{O} (shown at right). As before, all Tafs strictly co-precipitated in extracts from both wild type and the mutants, consistent with conservation of TFIID integrity in the mutant strain extracts. Rap1 also consistently co-precipitated with every Taf. The pattern of Rap1 interaction with Tafs was essentially identical in extracts of each strain; a result consistent with a very strong interaction between Rap1 and TFIID involving multiple Tafs. All data are derived from the exact same film exposure. Note that observance of the Rap1 isoforms of altered molecular weight are only detectable with very long film exposures when using the Taf7 antibody for IP, a peculiarity of that antibody. Data are adapted from Layer et. al., 2010.

Combinatorial Mutation of *TAF4* and *TAF5* Results in Lethality

Since Taf4, Taf5, and Taf12 all contain Rap1 binding domains, it is possible that each protein acts either independently or collaboratively to promote Rap1 interaction with TFIID in vivo. In support of this notion, distinct domains of Rap1 contribute to interaction with TFIID in vitro; thus one could hypothesize that each Rap1 domain is responsible for interacting with distinct subsets of Tafs (Garbett et al., 2007). Furthermore, nearly every RPG enhancer contains two binding sites for Rap1, hence the multiple enhancer-bound Rap1 molecules may be capable of direct interaction with TFIID (Lieb et al., 2001). Perhaps each molecule of Rap1 uses a unique mode of binding to TFIID. In any case, the existence of three distinct binding sites, each of which is present in multiple copies per TFIID molecule, and localized to the same general regions of holo-TFIID argues that there may be redundancy involved in the exact mechanism of interaction. Regardless of the exact mode of Rap1-TFIID/Taf interaction, introduction of both mutant *taf4* and *taf5* genes into yeast cells should weaken or reduce the Rap1-TFIID interaction to a greater extent than in either single mutant, and perhaps allow the observation of a reduced or defective Rap1-TFIID interaction using the co-immunoprecipitation technique. In other words, the specific objective of this experiment was not to examine a growth phenotype per se, but rather to generate a reagent that would be more useful for biochemical tests of our model of Rap1-TFIID-Taf interaction. The creation of *taf4 taf5* strains required a substantial amount of yeast genetics including independent haploid strain construction, crosses to generate diploids, induction of cells to undergo meiosis, re-isolation of haploid strains, and then several phenotypic and genotypic tests to confirm strain identity, with actual experiments being performed using

the standard plasmid shuffle technique. Results representative of these plasmid shuffles are shown in **Figure 3.6**. It was possible to obtain strains containing three *taf5 ts* alleles combined with each of the three *taf4 ts* alleles. However, upon shuffling no viable double mutant strains were recovered, while the single *taf* mutants grew as previously documented. The incubation temperature for the plates shown was 30°, however reduced temperature did not allow any growth of the double mutant strains (not shown). This is an example of synthetic lethality. Therefore no additional tests of Rap1-TFIID interaction in compound mutants could be made. However, this provided further evidence of the importance of these domains to cellular growth, and argues that these two Taf domains contribute to the same molecular process. Perhaps that process, for which these two protein domains act in common, is to allow Rap1 interaction with the TFIID complex. Unfortunately, additional headway could not be made to investigate Rap1 interaction with TFIID in *taf* mutant strains.

Rap1 Interaction With Altered Taf4 and Taf5 is Compromised In Vitro

To attempt to provide information about the molecular mechanism behind the reason for loss of growth in the *taf ts* mutants, it was necessary to return to investigation of binary Rap1-Taf interactions using in vitro techniques, since semi-in vivo approaches like the yeast two-hybrid and co-immunoprecipitation had only yielded negative results. The most obvious method was the Far Western, since it had been used initially to define Rap1 binding domains in all three Tafs (4, 5, and 12). To this end, full-length *taf5* mutant alleles were transferred from yeast to *E. coli* expression vectors; the proteins were

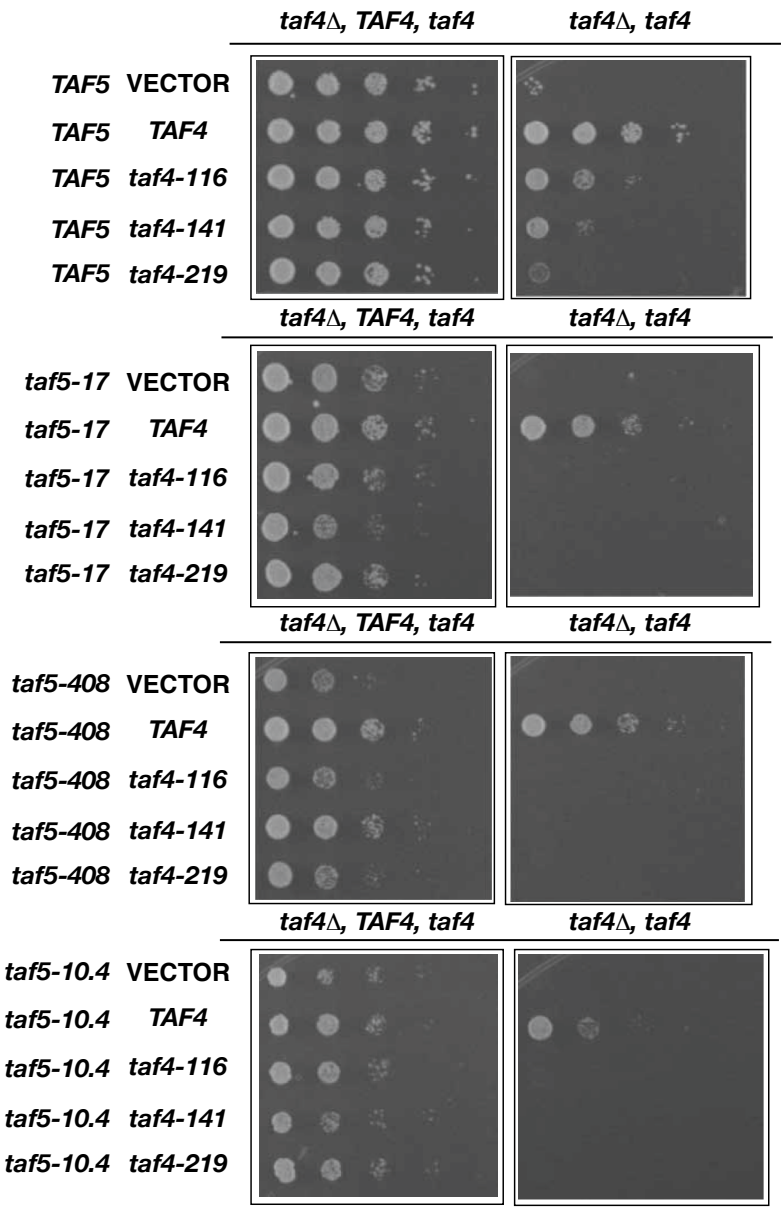
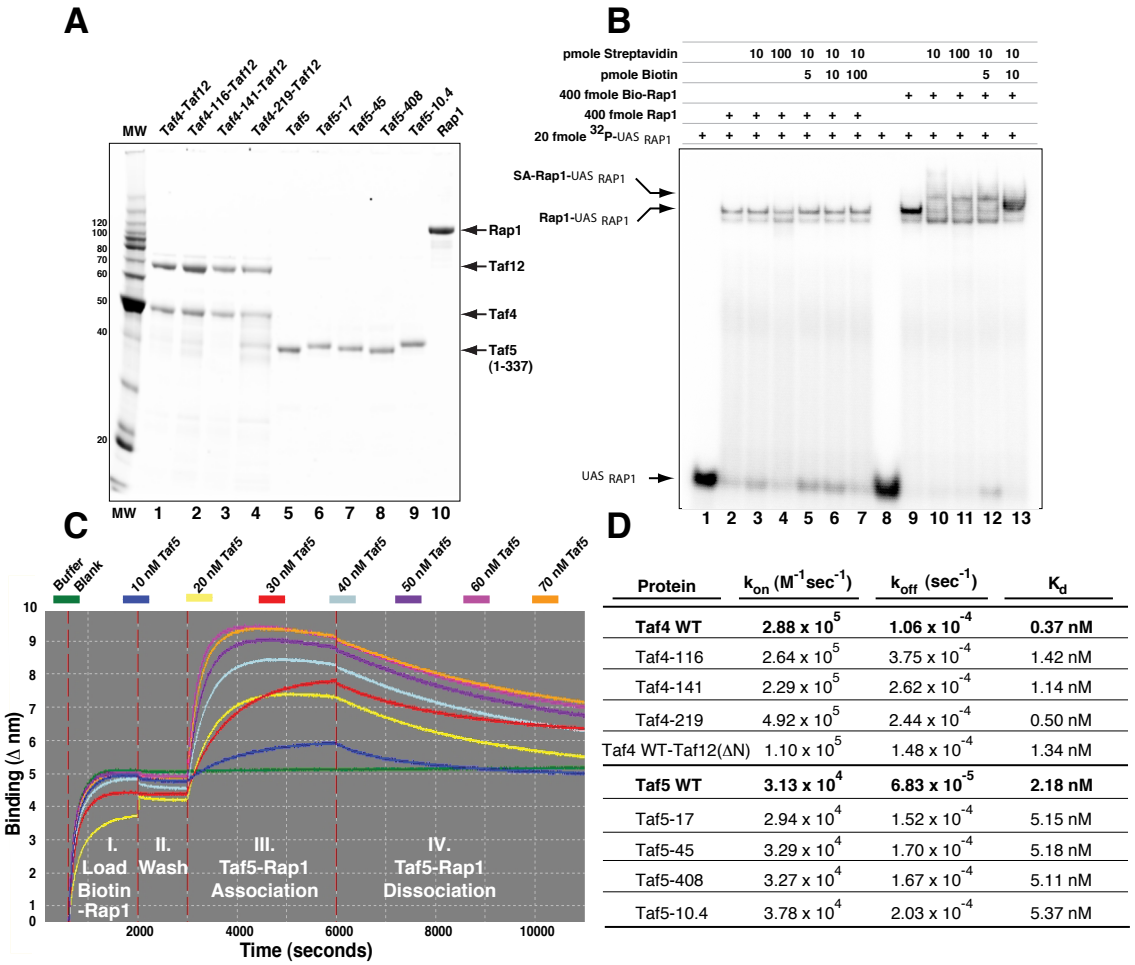


Figure 3.6: Synthetic lethality of *taf5* and *taf4* mutants. Shuffled, haploid strains, pseudo-diploid for *taf5*, containing either shuffled-in *TAF5* or *taf5* ts^+ alleles, were mated with unshuffled, haploid strains of the opposite mating type, pseudo-triploid for *taf4* and containing both *TAF4* and one of three *taf4* ts^+ alleles. Haploid, ts^+ strains pseudo-diploid for *taf5* and pseudo-triploid for *taf4* were isolated and subjected to plasmid shuffle at 30^o. Whereas the presence of the *TAF5* allele supported the isolation of viable *taf4* strains, presence of *taf5* ts^+ alleles could not, indicating a synthetic genetic interaction, possibly indicative of participation in the same molecular process. Data are adapted from Layer et. al., 2010.

expressed and purified from inclusion bodies, then used in the standard Far Western assay. There was no apparent difference in Rap1 binding to wild type full length Taf5 or the equivalent altered proteins (not shown). Alteration of the incubation parameters and buffer constitution to increase the stringency of binding reactions did not reveal any differences. Thinking that the sensitivity of Far Westerns might be a limitation in observance of subtle differences in binding affinity, I next established a more conventional GST pulldown assay, with GST Rap1 acting as 'bait' and soluble Taf5 fragments (residues 1-337, containing NTD1 and NTD2) or Taf4/Taf12 heterodimers as 'prey.'Again, there appeared to be no difference in binding of Taf5 variants to GST-Rap1 as compared to the wild type Taf5 amino terminus (not shown). The relatively cumbersome nature of Far Westerns and GST-pulldowns did not facilitate rapid or comprehensive optimization of binding reaction conditions, neither did these methods lend themselves to a quantitative description of binding events. Given these constraints and the availability of large amounts of suitable recombinant proteins, we opted to establish a new assay possessing both high-throughput and quantitative capabilities. The protein preparations used are shown in **Figure 3.7**. Note that Taf4 was co-expressed and purified as a heterodimer with Taf12, whose association was not affected by the amino acid substitutions in the Taf4 RBD, at least not within the fairly harsh conditions under which these proteins were prepared. The practical consequence of this coexpression/copurification strategy is that at least two binding sites are present in the wild type Taf4/Taf12 heterodimer preparation, and one intact RBD is theoretically always provided by wild type Taf12. Also note the mobility variation of the wild type Taf5 N-terminal fragment as compared to equivalent proteins expressed from mutant

alleles. Rap1 cysteines were covalently modified with biotin; the biotinylated Rap1 retained full DNA binding activity, as shown in the upper right panel of **Figure 3.7** that displays the results of EMSAs. This figure shows that addition of both biotin and streptavidin did not preclude DNA binding of modified or unmodified Rap1, as most clearly indicated by the enhanced mobility shift of biotin-streptavidin-Rap-DNA quaternary complexes. Moreover, the interaction of biotin-Rap1 with immobilized streptavidin biosensors allowed real-time monitoring of light scattering properties of streptavidin-biotin-Rap1. A representative real-time kinetic trace of streptavidin-biotin-Rap1-Taf5 complex binding and dissociation using the 96-well 'Octet' instrument with optical biosensors is shown in **Figure 3.7**. Biotin-Rap1 was first loaded onto the streptavidin biosensors (step I), and in a second step excess Rap1 was washed away (step II). Next the Rap1-containing biosensors were immersed in individual wells containing different concentrations of the wild type Taf5 N-terminus (step III); Taf5-Rap1 binding was monitored as increased signal (Δnm; y-axis) until after binding reached equilibrium. The biosensors containing Rap1-Taf5 complexes were then transferred to wells containing buffer alone and dissociation of the protein complexes commenced (Step IV). Taf5-Rap1 association occurred in a rapid, concentration dependent manner, and dissociation proceeded more slowly. Given the real-time data, rate constants describing both association and dissociation rates were derived, with the ratio representing the equilibrium dissociation constant, or K^D, which indicates the relative strength of interaction between different Taf5 variants or Taf4/12 heterodimer preparations and Rap1. K^D values indicated that altered Taf4 variants bound slightly less tightly than the



Protein	k_{on} ($M^{-1}sec^{-1}$)	k_{off} (sec^{-1})	K_d
Taf4 WT	**2.88×10^{5}**	**1.06×10^{-4}**	**0.37 nM**
Taf4-116	2.64×10^{5}	3.75×10^{-4}	1.42 nM
Taf4-141	2.29×10^{5}	2.62×10^{-4}	1.14 nM
Taf4-219	4.92×10^{5}	2.44×10^{-4}	0.50 nM
Taf4 WT-Taf12(ΔN)	1.10×10^{5}	1.48×10^{-4}	1.34 nM
Taf5 WT	**3.13×10^{4}**	**6.83×10^{-5}**	**2.18 nM**
Taf5-17	2.94×10^{4}	1.52×10^{-4}	5.15 nM
Taf5-45	3.29×10^{4}	1.70×10^{-4}	5.18 nM
Taf5-408	3.27×10^{4}	1.67×10^{-4}	5.11 nM
Taf5-10.4	3.78×10^{4}	2.03×10^{-4}	5.37 nM

Figure 3.7: Interaction of Rap1 with Taf4 and Taf5 proteins bearing amino acid substitutions in the RBDs. Proteins used are shown in the upper left hand panel, **A**. It was necessary to prepare Taf4 and variants as heterodimers with Taf12 by co-expression in *E. coli*, so at least two RBDs are present per heterodimer, with the Taf12 RBD being wild type in each case. The N-terminal 337 Taf5 residues, containing all of the RBD, was expressed in order to obtain soluble protein in *E. coli*. Rap1 was modified by biotinylation of cysteine residues (three per molecule and none in portion of Rap1 relevant to Rap1-Taf interaction). Retention of DNA binding activity by biotin-Rap1 is shown in the upper right hand panel, **B**. Biotin and/or streptavidin had no affect on native recombinant Rap1 binding to a radio-labeled RPG binding element in EMSA (lanes 2-7), whereas streptavidin caused a supershift of biotinylated Rap1-DNA complexes, which could be reversed by inclusion of excess free biotin (lanes 9-13). Therefore Rap1 DNA-binding, the only measureable biological activity, was unaffected by site-specific modification of Rap1 and thus modified Rap1 was deemed suitable for Rap1-Taf interaction studies. Biotin-Rap1 was used along with Taf4/12, Taf5, and derivatives from mutant alleles in biolayer-interferometry assays, an example of which is shown in the lower left hand panel. Interferometry experiments depend upon biotin-Rap1 binding to immobilized streptavidin, and occur in four steps which are monitored in real-time, **C**. Kinetic data allows calculation of rates of association and dissociation, and therefore determination of the associated rate constants, which in turn allow calculation of an apparent affinity of binding. A summary of data for wild type proteins and derivatives is shown in the table at the lower right, **D**. Preparation of vectors and recombinant proteins by J.H.L.; all protein-DNA binding and protein-protein interaction by Scott G. Miller. Data are adapted from Layer et. al., 2010.

wild type protein, and the Taf12 amino terminus also contributed to the interaction with Rap1 in this assay. While the wild type Taf5 amino terminus did not bind as tightly as the Taf4/Taf12 heterodimer, each Taf5 variant bound ~2 fold less tightly than the wild type protein. Altogether, these results were consistent with *taf* temperature conditional alleles encoding proteins that were defective for interaction with Rap1. The exact relationship between affects on Rap1 binding and the previously documented transcriptional phenotypes amongst the different mutant alleles remains somewhat less clear. This disparity will be discussed in some detail below, but the thrust of any argument will probably relate to the connection between observations made here, in vitro, and the situation actually occurring in vivo, with all the additional complexities inherent to that context. For now, note that all the experimental data shown was obtained using room temperature kinetic binding experiments, and different results may be anticipated if the temperature of reactions were raised. At 37^{o}, binding affinities may be much more significantly reduced than we have observed here.

Synthetic Genetic Interaction Between *TAF5* and Specific Mutant *RAP1* Variants

While we, and others, had previously documented functional relationships between Rap1 and TFIID in vivo, my study was still lacking in a demonstration of Rap1-Taf relationship, outside of purely biochemical experiments. Moreover, the level of understanding of particular protein domains within Tafs and their contribution to interaction with Rap1 had surpassed the known information available regarding Rap1 structure-function relationships. To address both these deficiencies, I decided to try and use a genetic approach to define connections between Rap1 domain functionality and Taf

domain functionality, in vivo. The basic strategy followed that used in creation of *taf4 taf5* strains. But since I did not have any *rap1* conditional mutants in hand, and because my goal was to obtain information about contribution of individual Rap1 domains, I generated the series of mutants shown in Figure 3.8. The focus of this study was on the Rap1 C-terminus, since deletion of the C-terminus is compatible with viability, unlike the DBD, which is strictly required for growth. Therefore I hypothesized that one or more specific domain(s) in the Rap1 C-terminus mediates interaction with one or more of the TFIID Tafs. While I could expect that removal of one Rap1 C-terminal domain at a time could cause a mild growth defect (at most), a more pronounced growth defect could possibly occur when *taf4* or *taf5* mutations were also present. A synthetic genetic interaction such as this would be indicative of biochemical crosstalk between specific Rap1 C-terminal domains and specific Taf domains. As a first step, several *taf5* mutants were combined with the systematic *rap1* mutants; the *TAF5* strain was combined with *rap1* mutants as control. Results of growth assays where *RAP1* was shuffled out to expose phenotypes associated with *rap1* mutants are shown in **Figure 3.8**. *RAP1* was essential as indicated by the lack of growth in the presence of vector missing the gene. The DBD, encompassed by residues 361 and 596, was also necessary for viability. The so-called Tox domain, which negatively affects growth when included in overexpressed Rap1, was not required for growth, and neither was the putative AD (Freeman et al., 1995). Residues of unknown function from 678 to 695 appeared to make a contribution to growth since strains missing this region grew at a reduced rate, as did a strain missing a large portion of the Silencing Domain (residues 763-827); slow growth of this strain may relate to compromised regulation of telomere homeostasis (Hardy et al., 1992a). On

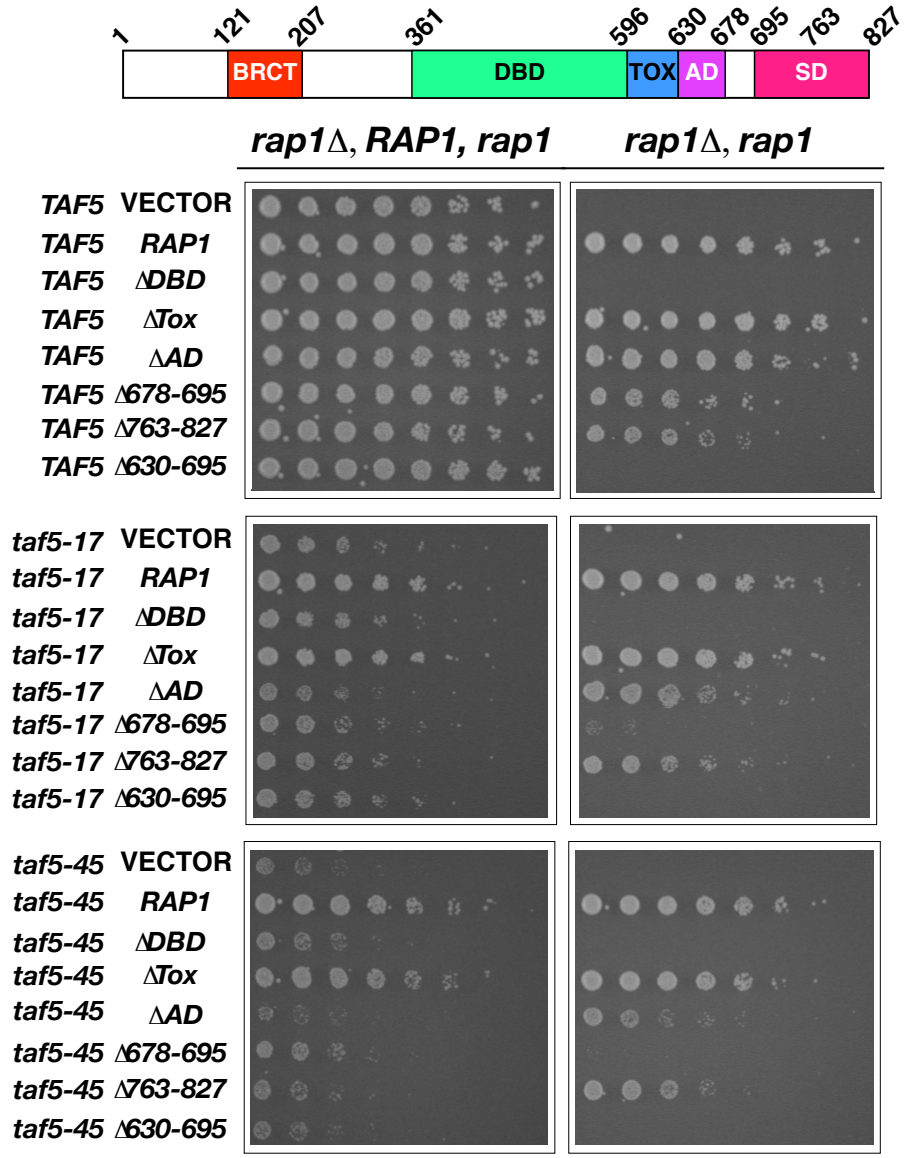


Figure 3.8: Synthetic lethality of *taf5* and *rap1* mutants. Haploid, ts^+ strains pseudo-diploid for *taf5* and pseudo-triploid for *rap1* were isolated and subjected to plasmid shuffle at 30^{o}. *rap1* ΔDBD and Δ630-695 could not support viability even in the presence of *TAF5*. Surprisingly, the Δ678-695 Rap1 variant, able to grow in the presence of *TAF5*, was extremely slow-growing and nearly inviable in the presence of *taf5* mutants. Besides this, another notable synthetic genetic interaction is the enhanced growth of unshuffled *taf5* strains containing extra *RAP1* or Δ*Tox* alleles (left hand panels). These were the most technically and logistically demanding experiments described in this dissertation. Data are adapted from Layer et. al., 2010.

the other hand, removal of the AD along with residues 678-695 was not compatible with growth (Δ630-695). This is consistent with collaboration between the Rap1 AD and the residues of unknown function, perhaps these collectively contribute to transactivation potential (possibly by interaction with Tafs). Note that in the *TAF5* background, all *rap1* alleles were recessive to the wild type *RAP1* allele, since the unshuffled strains grew equivalently.

On the other hand, in each of the *taf5* mutants, unshuffled strains containing *RAP1* or the *ΔTox* mutant exhibited an elevated growth rate relative to those strains lacking extra full-length or nearly full-length *RAP1*, indicative of partial suppression of the slow-growth phenotype of the *taf5* strains that occurred at the 30° incubation temperature used for these experiments. This was the first indication of a positive genetic interaction between *TAF5* and *RAP1* revealed by these experiments. A more obvious indicator of the relationship between the genes occurred when the Δ678-695 Rap variant was the sole source of Rap1 in the *taf5* mutant strains; the failure of this *rap1* allele to support growth when combined with *taf5* mutants represents a negative synthetic genetic interaction between DNA encoding this Rap1 domain and the *taf5* alleles. Such synthetic genetic interactions were taken to suggest the possibility of biochemical interaction between the Rap1 C-terminus and the Taf5 N-terminus; this biochemical interaction is likely to be direct given the variety of other observations I had made regarding relationships between these proteins.

DISCUSSION

Whereas the experiments detailed in Chapter 2 were intended to define protein domains within Tafs that mediate interaction with Rap1, the series of experiments described in this Chapter were intended to put precedence on testing the physiological relevance of the Rap1-Taf-TFIID interaction(s). The molecular physiological process in question was transcription, particularly that of Ribosomal Protein Genes, since these are the common target of the protein factors in question (Lieb et al., 2001; Mencia et al., 2002). However, it was by no means reasonable to hypothesize that these genes would be the only ones negatively affected within the novel *taf4* and *taf5* temperature conditional yeast strains that I generated for my thesis project. While it seemed reasonable to expect that all or most RPGs would be negatively affected, this characteristic, along with the identity of some of the affected non-RPG transcripts, could all be tested for using unbiased measurements of transcript abundance, in this case microarray analyses. The second major impetus of this series of experiments was to monitor possible molecular defects underlying loss of growth and RPG transcription phenotypes. Since the mutant *taf* alleles were rationally designed to minimize affects on Taf function other than Rap1 interaction, it seemed that a loss of function in Rap1 interaction would be the best candidate to test as being causative for growth phenotypes. With the caveat that there are likely to be several Rap1 binding sites within TFIID, provided by three different Taf proteins each present in multiple copies, I was prepared for negative results in assay of Rap1-TFIID interaction status in mutant strains (Garbett et al., 2007; Leurent et al., 2002; Leurent et al., 2004). Since assay of Rap1-TFIID

interaction presents technical as well as theoretical challenges, I also planned measurements of binary Rap1-Taf interaction in the context of altered Taf4 and Taf5 proteins. While this type of assay is readily performed in vitro, the physiological significance of binary interactions between activators and subunits of high molecular weight complexes is much more difficult to document in vivo. Assay of synthetic genetic interactions is a powerful, albeit indirect, strategy to test for functional interactions between multiple factors in the context of the whole organism. Given that I had some prior knowledge of both Rap1 and Taf protein domains involved in physical interaction, I could design a rational domain-directed synthetic genetic interaction study.

The complexity of the in vivo situation is always the challenge when attempting extrapolation from in vitro observations obtained using a minimal set of components. Transcription is certainly no exception to this, as stated before there are a huge number of individual proteins and many multisubunit complexes potentially involved at every step of transcription regulation. I am studying the connection between enhancer bound Rap1 and promoter bound TFIID but there are many candidates for bridging adaptors, or multisubunit coactivators that could possibly fulfill the same responsibilities more traditionally associated with TFIID. SAGA and Mediator are two complexes that could in theory coactivate RPG transcription by direct interaction with Rap1. Moreover, SAGA also contains Taf5 and Taf12, so it would be possible for Taf5 or Taf12 biochemical function to occur solely or partially through SAGA (Grant et al., 1998). A SAGA-dependent pathway of biochemical contribution would be difficult to decipher using genetic approaches such as generation of *taf* mutant alleles. Therefore it was reassuring to see that neither SAGA nor Mediator make any significant contribution to RPG

transcription, as evidenced by consistent steady state RPG transcript abundance within a broad collection of strains lacking subunits of each complex. A concern of mine was the authenticity of these strains, since none were prepared by me, although I verified each strain to the best of my ability and perhaps more tellingly, I observed that some did display overt growth phenotypes. However, I was most reassured because there were *PGK1* transcription defects within some of these strains. Thus the deletion of SAGA and Mediator subunits was effective in reduction of gene transcription (i.e. *PGK1*), just not that of RPGs. There was a recent report of the direct involvement of Mediator in RPG transcription, which used ChIP of epitope tagged proteins to show Mediator association at RPG promoters while an *med17* temperature conditional mutant strain was used to show genetic dependence on Mediator (Ansari et al., 2009). Personally, I feel that there are serious flaws with this report, since previous studies suggested that ChIP of tagged Mediator subunits is problematic, and further this report included only one Mediator mutant, leaving open the possibility of allele-specific affects (artifacts). The study I carried out used a large panel of isogenic genetic reagents, and allows a much stronger conclusion arguing against the involvement of either SAGA or Mediator. This fact simplifies the interpretation of genetic experiments involving TFIID Taf-encoding genes, where I endeavored to implicate the importance of Taf-RBDs in RPG transcription.

The observance of deficits in bulk polyA$^+$ RNA levels in temperature shifted *taf4* and *taf5* ts cultures, as scored by oligo dT-probed slot blots, was a strong indication that RPG transcription was compromised in these mutants at the non-permissive temperature. RPGs account for ~50% of all Pol II transcription initiation events in the genome and thus a negative affect on these 137 genes should confer a significant alteration in overall

gene expression. It was surprising that only mild RPG transcriptional defects occurred in the *taf4 ts* mutants when individual mRNA abundance was scored by primer extension. I feel that this is a limitation of the particular mutants that I generated; the three *taf4* strains used were the only *taf4* mutants that I isolated because there was a flaw in my plasmid library construction. This mistake generated many alleles with mutations outside the targeted area. Thus while the experiment was performed using the best reagents available at that time, but certainly not the ideal. Since then I have generated additional *taf4* mutants that almost certainly will be more appropriate for testing the contribution of the Taf4 RBD to RPG transcription. Even though the *taf4* alleles used were quite possibly suboptimal for testing my hypothesis, two of the three nevertheless did manifest defects in RPG transcript abundance. This result was easily overshadowed since the novel *taf5* mutants displayed a more dramatic RPG transcript reduction. Personally, I do not think this result reflects an elevated importance of the Taf5 RBD relative to the Taf4 RBD but rather simply reflects the relative quality of the genetic reagents used in the study. There was a much larger pool of *taf5* mutants to choose from whereas the *taf4* mutants represent the sum total of those identified, and consequently the *taf5* alleles were of higher quality as scored by several criteria. Still, the availability and application of more than one allele per *taf* strengthened this study, since it is widely acknowledged that the use of multiple unique genetic reagents will allow stronger conclusions (Durso et al., 2001). Moreover, a consistent reduction in the levels of distinct RPG transcripts occurred in the *taf* mutants, which was expected given that RPGs are coordinately regulated through mechanisms involving both Rap1 and TFIID. However, we did not know if all 137 transcripts were affected. This question and also knowledge about affects on other

non-RPG transcripts were of interest to us. The microarray experiments were executed to address these unknowns. A broad RPG transcription phenotype indeed occurred in the *taf5* mutants, as 110 of the 137 transcripts consistently clustered together as a group that showed decreased abundance at elevated growth temperature. Furthermore, the cluster analysis indicates that many other genes were affected by loss of Taf5 functionality (~600 downregulated greater than two fold, but not mRNA of *RAP1*, *TAFs* or any other important RPG regulators). The loss of some of these genes mRNA may be due to a secondary affect, occurring as a result of many perturbations on protein abundance or function, not exclusive to Tafs. A more interesting and testable hypothesis for the mechanism behind these deficiencies is that many other transfactors might use TFIID coregulatory function in a manner similar to Rap1. To test this hypothesis, it would be necessary to identify putative Taf5-interacting transfactors. The identity of such transfactors would be facilitated by a gene ontology (GEO) clustering analyses of the array data, which groups genes based on the biological process(es) in which the encoded protein participates. Like RPGs, other coordinately regulated clusters of genes often utilize one or more common transfactors to activate their transcription in parallel. GEO-grouped genes could be compared with ChIP-ChIP data that is available for several hundred yeast transfactors (Harbison et al., 2004; Lee et al., 2002). The comparison of this ChIP-ChIP data with equivalent data for TFIID occupancy, and also genetic dependence on Tafs, might indicate those genes that are commonly directly dependent both on a particular transfactor and also on TFIID (Lee et al., 2000; Shen et al., 2003). I analyzed the *taf1-ts2* mutant strain that has been used as a benchmark for establishing TFIID dependence of individual gene transcription, so we have our own dataset to

compare genes negatively affected in *taf5 ts* mutants to those affected in *taf1 ts2* (Huisinga and Pugh, 2004). This comparison might help to differentiate between the requirement of a gene for Taf5 function in SAGA versus Taf5 function within TFIID. With respect to those transfactors affecting genes other than RPGs, I must point out that Rap1 directly regulates ~450 additional genes, so some of those might share TFIID dependence and include direct Rap1-TFIID interaction in their regulation (Lieb et al., 2001). At one level, this seems a bit unlikely or alternatively, more complicated, since genes that I tested such as *PGK1*, *ADH1*, and *ADH3* are direct Rap1 target genes but are clearly Taf and TFIID-independent (Chambers et al., 1989). With respect to non-RPG activators, we do have some biochemical evidence that the SBF transfactor, which is a direct activator of a collection of cell cycle-regulated genes, can directly interact with Tafs (Sanders et al., 2002b). Unfortunately, my microarray data will probably not shed light on any affects on SBF target genes, since those transcripts are cell cycle regulated and need to be quantified within synchronized cultures to determine if a defect does occur in the *taf5 ts* mutants (Spellman et al., 1998). An experiment with synchronously growing cultures would be interesting to perform in the *taf5 ts* mutants, since it seems reasonable to propose that SBF might interact with TFIID, and Taf5, in a manner similar to Rap1. To summarize this set of experiments, there are many facets of the microarray data that have not been visited in detail and the dataset promises to be a source of novel information, and subsequent publications for the lab. My specific question was answered by completing these experiments; essentially the entire RPG regulon was negatively affected by incubating *taf5 ts* strains at the non-permissive temperature. This is consistent with the Taf5 RBD functioning critically in RPG transcription. The effect

was clear, it remained to be seen if the cause could be legitimately ascribed to a defective Rap1-TFIID interaction.

Having had established a bona fide RPG transcription phenotype in *taf* mutants designed to interfere with Rap1-Taf4/Taf5 interaction, it was next logical to examine Rap1-TFIID interaction in these strains. The expectation that it would be difficult to observe such an affect, owing to possible inherent redundancy, did not alleviate the disappointment when this prediction was fulfilled by negative experimental results. One may ask just why there is such a strong transcriptional phenotype in the *taf5* mutants without a detectable affect on Rap1-TFIID interaction. Such strong phenotypes have been associated with *taf* mutants that compromise TFIID integrity and/or stability (Komarnitsky et al., 1999). Clearly this is not the case with these particular mutants, since there is no change in quantity or quality of Tafs that co-immunoprecipitate in extracts of mutants as compared to wild type. The use of a comprehensive panel of antibodies for IP solidifies the argument that TFIID integrity remains unaffected in the mutants generated by this study. The immunoprecipitation data on Taf-Taf association in extracts of the mutant strains is also consistent with my yeast two-hybrid data, which indicated a lack of involvement of Taf4 and Taf5 RBDs in Taf-Taf interaction. The lack of a noticeable reduction in Rap1-TFIID association in the immunoprecipitation experiments might be explained by the existence of some form of compensation with respect to conservation of Rap1-TFIID interaction in the mutant strains. This seems a bit unlikely given the strength of the transcriptional phenotype; were there effective compensation one would not anticipate the growth and transcription phenotypes being so severe. The more likely reason for the failure to observe defective interaction, in my

opinion, is the limitation of the technique used. While I made an attempt at systematic buffer optimization in order to identify optimal salt and detergent concentrations, these were not combined with other parameters such as pH of extraction and IP. Neither were identity of different salts nor detergents pursued although these could well contribute to the nature of the interaction(s) and affect of *taf* mutations. Perhaps the other techniques I have mentioned, such as FRET or in vitro pulldowns with TFIID isolated from mutant strains would readily identify differences between wild type and mutants. Unfortunately I think this is a long way from fruition, since these methods are not without their own difficulties and caveats, which will also need to be sorted out through additional hard work, much like what was necessary to get the IP assays working. With respect to FRET, I think the judicious use of epitope tag locations will be of exceptional importance for these experiment to have any chance of success. The apparent integrity of TFIID in extracts of mutants can be taken as a further indication that it should be possible to purify the complex from mutant strains using our standard isolation procedure. But this will take the right person with the right skill-set and an eye for detail to do this and re-establish in vitro Rap1-TFIID interaction assays. So it is fair to say the Co-IP experiments did provide some insights such as the observance of conserved TFIID integrity and an apparent association with specific Rap1 isoforms. The identification of any post-translational modification in this Rap1 population and their locations within the protein will be interesting to investigate and could very well provide some insight into the nature of the Rap1-Taf/TFIID interaction and possible dynamic regulation of the interaction. Another positive note about this round of Co-IP experiments has to do with the evident possibility of yet another type of experiment that might be put to use to

characterize Rap1-TFIID interaction. Nearly every individual Taf protein was prepared in recombinant form for use in affinity purification of specific anti-Taf IgG. These proteins were generally prepared in soluble form with good purity and yield. In theory it should be possible to attempt reconstitution of full or partial TFIID complexes using recombinant proteins prepared using the methods I have worked out. Indeed, partial reconstitution of multisubunit complexes using purified subunits, originally pioneered using *E. coli* RNA Polymerase, has been applied to human and fly Tafs and TFIID in several previous studies (Chen et al., 1994; Guermah et al., 2001; Ishihama and Ito, 1972). One can easily envision the incorporation of recombinant Tafs encoded by mutant alleles into reconstitution experiments (Tang et al., 1996; Tang et al., 1995). Thus full or partially reconstituted TFIID complexes containing altered proteins might be prepared for use in Rap1-TFIID interaction studies. However promising and/or technically challenging this particular endeavor may be, like all the other possible approaches, TFIID reconstitution still suffers the limitation that the mutant alleles described by my studies may be too phenotypically strong (or exhibiting too dramatic a reduction in RPG transcription as a result of heat shock) to be useful for the full panel of experiments I would like to use in testing the model of Rap1-TFIID interaction.

Having put much effort into establishing the Co-IP procedures, and obtaining at least some useful observations, I wanted to try and address the potential caveat of inherent redundancy of Rap1-Taf-TFIID interaction. Thus several of the available *taf5* mutants were systematically combined with *taf4* mutants. The objective was to make strains with multiply disrupted surfaces within Tafs used by Rap1 to interact with TFIID. If such strains could be identified and isolated, it might be possible to observe a deficit in

interaction as compared to single mutants or wild type using Co-IP assays. Unfortunately the possible allelic combinations produced only strains that were inviable. The molecular mechanism behind this phenotype can only be speculated on; it may be due to the loss of Rap1-TFIID interaction or could be a consequence of compromised TFIID integrity or stability, or any other myriad reasons. Again, this question of whether or not I used the 'right' mutants in these experiments is one that must be asked (Ebright, 1991). The best mutants would be strong enough to observe phenotypes, but not so strong as to chronically manifest a portion of the overall phenotype. Perhaps the 'right' mutants for this type of experiment are those that do not display as dramatic a phenotype on their own, relative to the mutant *taf4* and *taf5* alleles that were used. While I feel the overall qualities of the *taf5* alleles are quite good by comparison with other mutant strains I have worked with, nearly every one I identified encodes multiple amino acid substitutions, and this appears to be a general feature of most every useful allele present in the libraries I constructed. Perhaps a novel but more systematic targeted mutagenesis could identify *taf5* alleles with even 'tighter' phenotypes, which could subsequently be combined with existing or to-be-identified *taf4* alleles, in order to generate viable strains useful for testing the model. In fact I have made progress in identifying novel targeted *taf4* alleles, in a study that will be discussed in the next chapter. The generation of better reagents will be the holdup for fulfilling the promise of *taf* synthetic interaction studies. The study I did perform served to further indicate the importance of the identified Taf domains, but the molecular mechanism can only be inferred and not proven. Unfortunately, inviable strains cannot provide additional evidence of molecular mechanism. Indeed, dead cells will tell no tales. So while we got an additional clue from the synthetic genetic

interaction study between *taf4* and *taf5* alleles, we were left wanting for means to carry forward with analysis of Rap1-TFIID interaction. Since compound *taf* mutant strains were likely necessary to justify carrying forward with extract-based Co-IP approaches, and these strains could not be obtained, the next best available option at this juncture was to express recombinant Taf4 and Taf5 proteins corresponding to the mutant alleles and measure the capacity of each of these proteins for interaction with Rap1.

Far Western and GST pulldown assays did not identify any difference in the interaction between altered Taf5 proteins and Rap1. Thinking that these methods lacked the sensitivity or capacity for rapid optimization to determine if any real deficiencies were occurring, we turned to a fairly new method that provides kinetic protein binding data, termed biolayer interferometry. Thus we have to date utilized five different methods to observe interactions between Rap1 and Tafs or TFIID: in vitro TFIID pulldown, Far Western, GST pulldown, co-immunoprecipitation with Tafs from yeast extracts, and biolayer interferometry. This last technique allowed the determination of binding constants describing the affinity of Rap1-Taf interaction. The consequence for interaction of Taf4 with Rap1 was different depending upon the Taf4 allele used. Somewhat paradoxically, the Taf4 variant corresponding to the allele with the most severe transcriptional defect, *taf4-219*, demonstrated the least perturbed Rap1 binding affinity. The protein corresponding to the *taf4-141* allele exhibited a reduced Rap1 binding affinity, even though this strain did not have a profound RPG transcription defect. It was also somewhat surprising that Taf4 bound to Rap1 more tightly than Taf5, even though much more severe transcription phenotypes were observed in the *taf5* mutants. However, recall that most of these measurements were made using recombinant

Taf4-Taf12 heterodimers thus two Rap1 binding sites may be available; consistent with this hypothesis removal of the Taf12 N-terminus results in a protein that should be compromised for interaction with Rap1 but fully capable of dimerizing with Taf4, and Taf4/12 dimers containing truncated Taf12 bind Rap1 with affinity similar to Taf5. However, the affinity of such an altered heterodimer for Rap1 is still slightly greater than Taf5 affinity for Rap1, which could partially explain why none of the *taf4* mutants have as strong a transcriptional phenotype as *taf5* mutants, and why truncation of the Taf12 N-terminus causes no obvious in vivo phenotype at all. In other words the Taf4-Taf12 heterodimer may provide a localized, high affinity 'platform' within TFIID for Rap1 binding. With respect to Taf5 variants, although each binds Rap1 with reduced affinity compared to the wild type protein, there is a consistent quantitative reduction, which is surprising given that the *taf5-408* allele has a less severe transcriptional phenotype than the other mutants; one might expect that this particular variant would bind better than the others but not as well as wild type. Altogether, it is very difficult to draw connections between the transcriptional phenotypes associated with particular mutant alleles and the Rap1-binding properties of proteins corresponding to those alleles. I do not feel it is a particularly worthwhile exercise for several reasons, not the least of which is that these experiments, while apparently reproducible at the quantitative level, were actually not performed with multiple independent preparations of Taf4 and Taf5 variant proteins. Thus binding characteristics could be artifacts of peculiarities unique to the preparations studied. Most fundamentally, these binding measurements are performed outside the context of the TFIID complex where these proteins are naturally found and which is naturally bound by Rap1. Thus protein conformation could be radically different here in

these experiments with isolated Tafs than what is found in the context of TFIID, where many other Taf-Taf, Taf-DNA, and Taf activator interactions are occurring, and where any combination of these different types of interactions could converge to influence the properties of those Taf domains responsible for interaction with Rap1. Also, the recombinant proteins were produced in *E. coli* and thus are devoid of any post-translational modifications that could occur in yeast. Indeed, the Taf4 and Taf5 domains of interest are in all likelihood subject to PTM in vivo. We have evidence that the Taf4 RBD is subject to phosphorylation in vivo, and there is some likelihood that the Taf5 RBD is subject to ubiquitination and/or sumoylation. Both types of modification could be important for RPG transcription; Rap1 along with Tafs are probably conduits for such regulatory information stream since it is known that multiple signaling pathways converge upon RPG transcriptional regulation. In light of the discussed caveats, the physiological relevance of these in vitro Rap1-Taf binding experiments will await further investigation, and will benefit greatly from corresponding binding data obtained using Rap1 and TFIID preparations containing altered Tafs (whether obtained in vitro or in vivo).

Yeast genetic approaches afforded the opportunity to perform in vivo tests of the relationship between the Taf5 RBD and Rap1. Clearly the choices of *TAF5* genetic reagents were my novel mutant alleles, but the equivalent choice for *RAP1* presented more of a dilemma. We have had some previous trouble with the authenticity of published *rap1* temperature conditional mutants (Freeman et al., 1995; Kurtz and Shore, 1991; Lustig et al., 1990). I attempted to reconstitute several *rap1 ts* alleles using site-directed mutagenesis but the resulting sequence-correct clones did not manifest the

growth phenotype predicted by the literature. Thus although the use of point mutant and/or temperature conditional mutants is always preferred to alleles containing large deletions, I needed to turn to deletion mutants to pursue analyses of genetic interaction between *TAF5* and *RAP1,* since generation of novel point mutant *rap1* alleles was not practical given the time constraints. To move ahead I systematically deleted sequence encoding individual Rap1 domains beginning with the Rap1 DBD, since the amino terminus is completely dispensable for cell growth, and because we had seen TFIID interaction with the DBD and C-terminus, but not the N-terminus. I expected that certain particular C-terminal domains would be important for cellular growth on their own, since deletion of the whole C-terminus results in a severe slow growth phenotype. Consistent with the literature, Rap1 forms lacking the DBD failed to support viability (Freeman et al., 1995). Therefore I could not examine genetic interactions between *rap1* DBD mutants and *taf5* mutants. My suspicion was that any observable genetic interactions would occur between *rap1* activation domain mutants and the *taf5* mutants, since the affected Rap1 domain would predictably function to stimulate/activate RPG transcription. However, I wondered if the Rap1 AD had been mapped in its totality (Hardy et al., 1992a). After all there was a stretch of residues C-terminal to the AD that had no known function, but which by inference could associate with AD function, owing to these domains' proximity. Simultaneous deletion of sequence encoding the AD and this region of unknown function resulted in lethality in otherwise wild type strains, suggesting that these domains indeed functionally overlap. This was a novel observation. Perhaps the contiguous stretch of residues participates in the same process, namely transactivation and/or physical interaction with Tafs. Consistent with this possibility, the *rap1* mutant

missing the region C-terminal to the AD exhibited strong synthetic negative interactions with each *taf5* mutant, suggesting a functional interaction between specific portions of the Rap1 C-terminus and the Taf5 N-terminus.

Many potentially informative experiments could follow in a logical progression to extend these observations, using very similar techniques. First, I would expect to be able to identify point mutants in this region of *RAP1* that display similar synthetic interactions with *taf5* mutants, because if this protein domain is truly functional in contribution to Rap1-Taf5 physical interaction, then individual Rap1 amino acid residues there should contribute to contact between the two proteins. Similarly one would expect that certain specific changes in this *RAP1* sequence might suppress the effect of the *taf5* mutants. If the conformation of the Taf5 N-terminus is altered within the mutants such that the domain(s) cannot effectively interact with the surfaces provided by the Rap1 C-terminus, then perhaps compensatory conformational changes caused by mutation of *RAP1* could counteract the changes in *TAF5* to reintroduce a productive interaction. This is one theoretical basis for the past success of suppressor genetics. Thus one can propose many experiments to test for both negative and positive genetic interactions. If a weakened protein-protein interaction between Rap1 and Tafs occurring within either *taf5* or *rap1* mutants is truly responsible for the phenotypes, then it should be possible to rescue growth of the *taf5* mutants at the non-permissive temperature by overexpressing Rap1. Weakened interactions and the accompanying hastened turnover of interaction might be compensated by a mass action affect of increased cellular concentration of an individual binding partner. This is another commonly practiced method of suppressor genetics (i.e. high copy suppression). Unfortunately, overexpression of Rap1 is itself associated with

significant cytotoxicity and slow growth, although removal of the so-called Tox domain alleviates this phenomenon to an extent (Freeman et al., 1995). Since the Tox domain itself doesn't contribute to genetic interaction with *TAF5*, I decided to try and rescue the growth phenotype of *taf5* mutants by overexpression of the tox-domain deleted Rap1 variant. Indeed, such overexpression alleviated the lack of growth of *taf5* mutants at elevated temperature; *taf5* strains containing excess Rap1 could grow, albeit very slowly, at 37^{o} (not shown). All of these genetic experiments could serve to shed light on the overall Rap1-TFIID interaction; perhaps Taf4 also interacts with the Rap1 C-terminus. Alternatively the Rap1 DBD might be restricted to interaction with Taf4 and the Rap1 C-terminus reserved for interaction with Taf5. Making things more complicated within the relevant context of these interations, which occurs on RPG enhancers and promoters, the presence of two molecules of Rap1 per RPG enhancer could either allow some overlap and/or redundancy in the mode of Rap1-Taf4-Taf5-Taf12-TFIID interaction, but it is just as possible that this enhancer characteristic could facilitate distinct modes of Rap1-Taf interaction. These various possibilities might be supported or ruled out by directly testing the affect of Rap1 domain alteration on the biochemical interaction with Tafs. Use of the Rap1 proteins corresponding to the deletion variants along with Taf4/12 heterodimers and the Taf5 N-terminus in the biolayer interferometry assay would be the most straightforward way to begin these experiments. I expressed the appropriate Rap1 proteins in *E. coli* and isolated them with good yield and purity but these variants have not yet been analyzed with respect to interaction with Tafs.

It is easy for me or anyone else to criticize the limitations of the experiments I did, being that we now have the benefit of hindsight. Indeed, one must start somewhere,

and in this respect it does seem clear that I made a good start into the dissection of Rap1-TFIID interaction. Thematically, I think that future studies must focus on two particular areas of deficiency. First is to address the question of whether or not the 'right' *taf* and *rap1* mutants are in hand at this time to allow the progression of our studies. Since we are talking about a set of protein-protein interactions, and since many well-known examples have been mapped to peptide levels, I would argue that we do not yet have the most ideal mutant alleles in hand since there are scattered amino acid substitutions encoded in each allele. As an example of a situation where peptide-level resolution of interaction is the norm, the protein regions responsible for recognition by monoclonal antibodies are commonly mapped to within 8 to 10 amino acids, and specific changes of just one residue within that peptide can essentially eliminate antibody-antigen interaction. If we could achieve that type of resolution of protein sequence mediating Rap1-Taf interaction, within each protein, then we could much more easily obtain precise and rationally designed mutant alleles with minimal chance for cause of off-target affects. It is quite possible that cells regulate Rap1-Taf interactions in a dynamic manner, for example by post-translational modification, in order to allow the highly environmentally sensitive regulation of RPG transcription. Identification of the locations of such putative post-translational modifications, for example in the state of high level RPG transcription versus physiological conditions wherein expression is halted, could tell us a lot about the specific residues directly involved in Rap1-Taf-TFIID interaction. Even in the absence of a PTM-directed mutagenesis study, a systematic residue-by-residue site-directed mutagenesis study of the identified Taf and Rap1 regions is an attractive approach that would surely allow us to obtain the 'right' mutants for our future studies.

The second theme I would propose to address deficiencies of my study is to look at things in the context of a 'bigger picture' of RPG transcription. I have discussed the need to more effectively examine these interactions in the context of a Rap1-TFIID interaction, but that also leaves out the fact that the proteins are almost certainly bound to DNA while they are engaged in protein-protein interaction in vivo. Perhaps this characteristic represents an upstream and downstream influence on the mechanistic outcome(s) of Rap1-TFIID interaction. Since the enhancer DNA bound by Rap1 is approximately 400 bases upstream of promoter-bound TFIID, there must be significant conformational complexity of the DNA involved in the process, otherwise it seems unlikely that the proteins could effectively interact while still bound to their respective regulatory DNA sequences. In the absence of our ability to examine these protein interactions in the context of the cellular mileu, for example by using FRET, we must extend our planned biochemical experiments to not only include Rap1 and the TFIID complex instead of recombinant Tafs. But we also must consider the affect of DNA in these interactions. Since TFIID does not bind to DNA in vitro in the absence of TFIIA, we would have to include this GTF in our future studies as well. The regulatory interplay of Rap1, TFIID, TFIIA, and DNA have actually been examined extensively by our collaborators in France, and these studies and several others will be detailed in the final Chapter of my dissertation. Structural studies combining Rap1, TFIID, DNA, and TFIIA have placed my observations within a more mechanistic context, and this has led to ideas for many more potentially informative experiments to understand the precise function of TFIID-directed transcriptional coregulation.

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

Summary

My thesis project was designed to pursue an understanding of the mechanistic outcomes of activator-TFIID interaction. I initiated my work by identifying the individual subunits of the TFIID complex mediating interaction with the yeast Rap1 activator (Garbett et al., 2007). To allow a successful genetic analysis of these interactions, the regions of each TFIID-Taf responsible for the Rap1 binding were mapped at low resolution to within several dozen to several hundred amino acid residues (Garbett et al., 2007; Layer et al., 2010). The expectation was that these regions of the Taf proteins would be important for cellular growth and RPG transcription, if the mapped RBDs were truly involved in physiologically relevant Rap1-TFIID binding events. Conversely, the recessive mutants I obtained that affected the function of these Taf domains and which exhibited growth and/or RPG transcription defects should also compromise interactions of Tafs or TFIID with Rap1. While all of these expectations were borne out at the level of Rap1-Taf interaction there are still, not surprisingly, caveats to the interpretation of my experimental results and our models as presented in (Layer et al., 2010). Thus, an extended set of future experiments that utilize both genetic and biochemical techniques will be required to more firmly establish the mechanistic outcome of Rap1-TFIID interactions. So as with any good research project my work has been informative but it remains incomplete. However, I am satisfied that the relevant

future experiments have also been hinted at by my studies. I have systematically investigated and described the function of not just one but instead four different proteins. Here I will attempt to discuss theoretical and practical ways to advance the study of Rap1-TFIID interaction, given the observations and limitations that I have revealed and/or encountered during my time studying these factors. All of the genetic investigation that was done was guided by low-resolution analyses of residues needed for in vitro protein-protein interaction and this in vitro data let me focus my genetic experiments. Since only low-level resolution was obtained, a randomization strategy (Cadwell and Joyce, 1992) for generating mutant *taf* alleles was most appropriate at the time my studies were initiated. This choice of technique was made based on my level of expertise at that time and also simply because that approach had proven successful in our laboratory (Singh et al., 2004). This mutagenesis method suffers the drawback that frequently many residues within the targeted region are altered at the same time, therefore the exact amino acid residues responsible for the cognate protein-protein interaction are usually not identified. Indeed this was the case with most all of the mutants that I described in the previous chapters, consequently these mutants are of limited utility as a basis for carrying forward a more detailed mechanistic description of Rap1-Taf interaction.

It is tempting to speculate that a fairly small number of residues directly participate in the binding events between each Taf and Rap1, with other residues in the vicinity either being unimportant or contributing to the interaction indirectly, for example by contributing to the overall conformation of that protein region or domain. Ideally one would hope to define the individual direct-binding residues, for example by biophysical

studies coupled with mutagenesis (Wakula et al., 2003). I have mentioned several times about the need to obtain the 'right' mutants, that is, those mutants that only influence the process in question and do so by the most direct mechanism possible. Presumably, such mutants would possess the most 'tight' conditional growth phenotype with little manifestation of that phenotype at the permissive condition(s), but with an obvious defect at the non-permissive environmental condition(s). Such mutants would be potentially more useful for extended genetic and biochemical experiments, such as synthetic genetic interaction analyses or isolation of TFIID from mutant strains. Multiple distinct peptide residues within each Taf and also Rap1 probably contribute to the interaction(s); the distinct N-terminal domains 1 and 2 in Taf5 both contribute biochemically to interaction with Rap1, the Rap1 DBD and domains in the C-terminus make either distinct or overlapping contributions to interaction with Tafs and/or TFIID, and the RBD in Taf4 appears to contain at maximum ~92 residues, an expanse that could easily include several distinct motifs or domains (Layer et al., 2010). To obtain the best possible mutants, it would be necessary to identify the most critical individual residues within each domain of interest. It is now within my ability to perform site-directed mutagenesis to alter every individual amino acid in each domain. The most practical way to implement this is systematic alanine block scanning mutagenesis, where one to several contiguous residues are simultaneously changed to alanine. Growth and/or transcription analyses would reveal phenotypes associated with changing particular residues to alanine. This overall strategy would serve to answer a fundamental question: what is minimally required to make a Rap1 binding domain in Tafs, or a Taf binding domain in Rap1?

What Makes a Rap1 Binding Domain? Taf4 as a Test Case.

Since the least success was obtained in generation of *taf4* mutant alleles, and because there is no structural information about the Taf4 RBD, I performed the alanine scanning mutagenesis of the Taf4 RBD-encoding sequence to try and identify better *taf4* mutants (Werten et al., 2002). These reagents would allow me to more rigorously test the importance of the Taf4 RBD in vivo, and map those residues most critical to the mechanistic function of that domain. I generated sixteen 6 amino acid Ala block mutants within *TAF4* sequences encoding the RBD of the protein. These mutants were transformed into yeast and tested for growth phenotypes using the plasmid shuffle assay. The results of this experiment are shown in **Figure 4.1**. The usual negative and positive controls of empty vector and vector expressing wild type Taf4 were performed. Two internal deletion mutants again showed that the entire Taf4 RBD, as mapped to residues 253-344, was required for growth and also that the smaller deletant lacking residues 284-326 also failed to support viability. When present as the sole source of Taf4 protein, each of the Ala block-mutated alleles were able to support growth at 30° although two alleles conferred a slow-growth phenotype; these two variants have Taf4 residues 311-316 or 317-322 changed to alanine. Moreover, the slow-growth phenotype of these two variants was more severe at 37°, with the 317-322 Ala variant unable to support viability. It is noteworthy that a subtle slow growth phenotype was evident for the 251-256, 264-269, and 269-274 Ala-block mutant variants when these strains were grown at 37°. None of the alanine substitution strains exhibited reduced steady state abundance (**Figure 4.1**), thus the growth defect phenotypes observed above are not explained by reduced protein stability and/or defective synthesis.

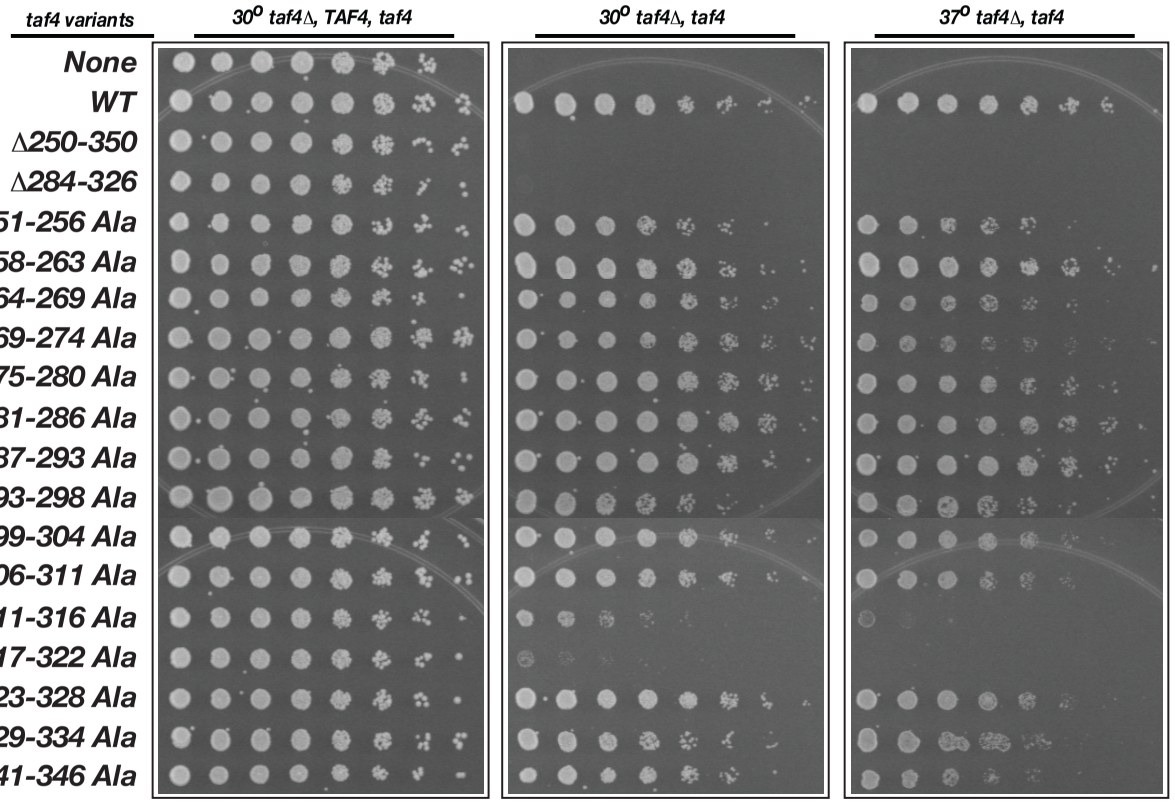
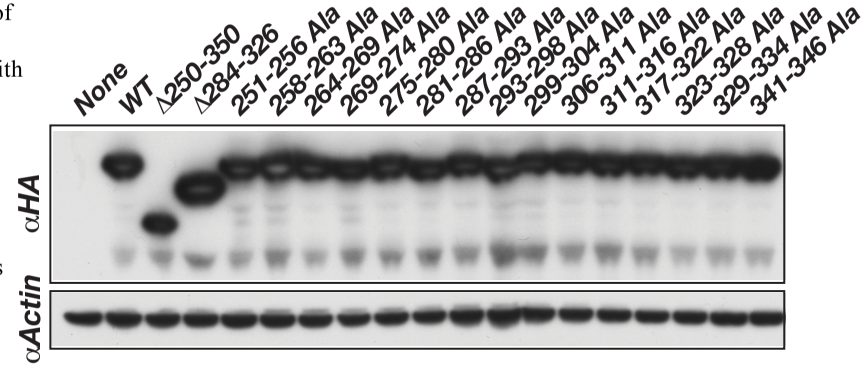


Figure 4.1: Systematic alanine-scanning mutagenesis of sequences encoding the Taf4 RBD. Taf4 variants were created where six contiguous residues were replaced with alanine and complementation properties were scored by a standard plasmid shuffle assay at 30° and 37°. Two variants affecting residues 311-322 had obvious growth defects, and were inviable at 37°. However, note that mild growth phenotypes at 37° were consistently evident with variants affecting residues 264-274, and this phenomenon is seen to a less obvious extent at 30° as well. This experiments defines the most critical residues within the Taf4 RBD. Companion immunoblot at right.



Since the *taf4* alanine scan family appeared to exhibit phenotypes ranging from no visible growth reduction, mild growth reduction at elevated temperature, to inviability, I reasoned that a synthetic genetic interaction study might expose subtle growth defects associated with some of these *taf4* mutants. Thus I sought to systematically combine these *taf4* variants with one or more of my *taf5* temperature conditional mutants. As shown in **Figure 4.2**, 14 of the 16 *taf4* Ala scan mutants were combined with strains bearing either *TAF5* or two of the *taf5* temperature conditional mutants, *taf5-17* and *taf5-10.4,* then subjected to plasmid shuffle at 30°. As controls, the two *taf4* deletion mutants and the three previous *taf4* temperature conditional mutants were combined with these *taf5* strains; as demonstrated before, these variants could either not support viability or exhibited synthetic lethality with *taf5* mutants. Additionally, since residues 311-322 had a clearly important role in Taf4 function and because we had some evidence that serine 311 was phosphorylated (Manish K. Tripathi, J.H.L., Scott G. Miller, unpublished observations) S311 was changed to either alanine or a phosphomimetic aspartate residue. When combined with the *taf5-17* or *taf5-10.4* strains, the *taf4* S311D mutant, but not the *taf4* S311A mutant, showed a slow growth phenotype indicating a functional interaction between this Taf4 residue and the Taf5 NTD2. Consistent with the importance of the portion of Taf4 containing S311, the other two *taf4* alanine substitution mutants affecting residues 311 to 322 displayed synthetic lethality when combined with either *taf5-17* or *taf5-10.4*. These results further support the importance of the identified Taf4 residues to the function of the RBD contained within that part of the protein, and indicates a

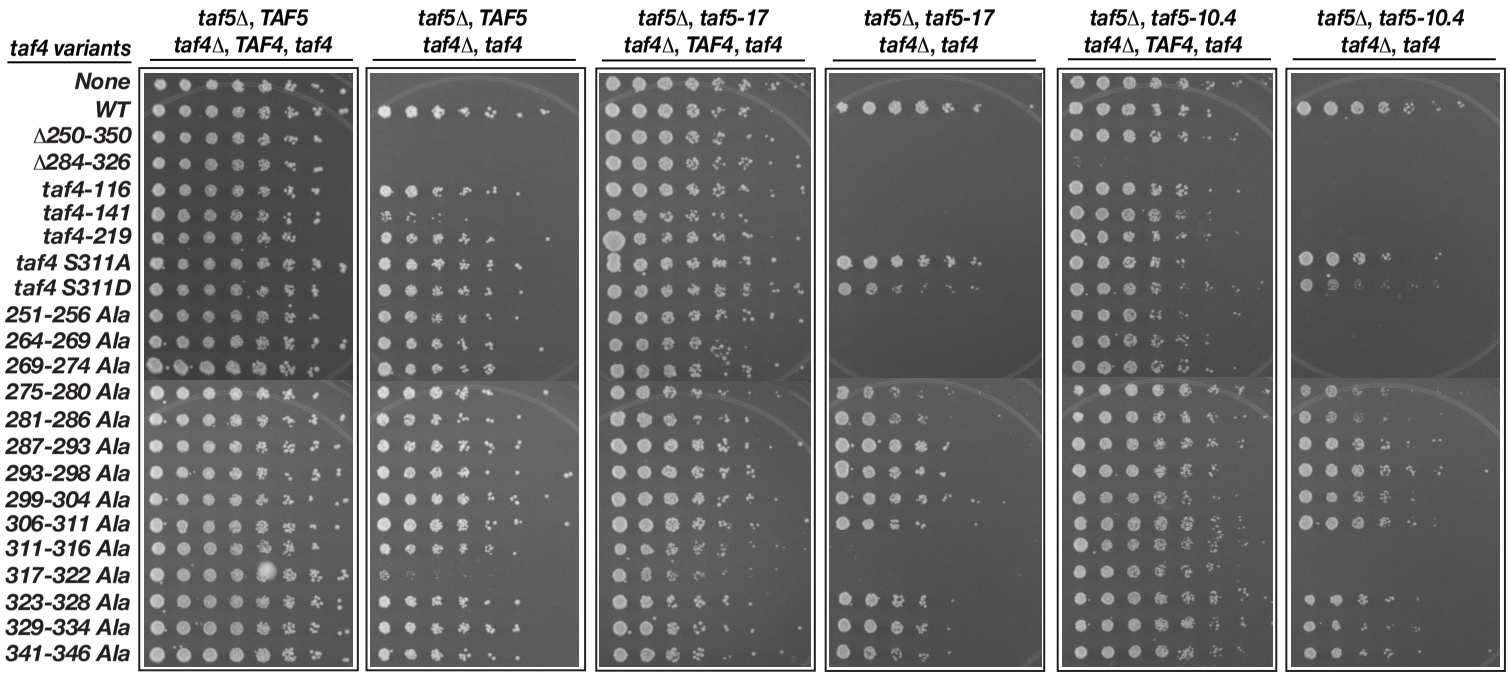


Figure 4.2: Synthetic lethality of *taf5* and *taf4* ala scan mutants. Haploid, ts^+ strains pseudo-diploid for *taf5* and pseudo-triploid for *taf4* were isolated and subjected to plasmid shuffle at 30^o. Whereas the presence of the *TAF5* allele supported the isolation of viable *taf4* strains corresponding to every ala scan mutant, presence of *taf5* ts^+ alleles was not compatible with isolation of *taf4* strains containing ala substitutions in residues 251-274 or 311-322, indicating a synthetic genetic interaction, likely indicative of participation in the same molecular process. This may indicate that the Taf5 RBD-containing N-terminus and two separate contiguous portions of the Taf4 RBD may be participating in the same molecular process, which may include interaction with Rap1.

functional overlap with the mapped Taf5 RBD. Importantly, 9 of the 14 tested Ala scan variants did not exhibit synthetic lethality with *taf5* mutants, indicating a high degree of allele specificity in these analyses. Interestingly, three Ala-block variants that affected residues 251-274 also showed synthetic lethality with both *taf5* variants, even though these *taf4* mutants only showed very mild temperature conditional slow growth phenotypes on their own **Figure 4.1**. This result suggests that two separate stretches of residues in the Taf4 RBD contribute critically to the function of that domain and these include residues 251-274 and 311-322. Therefore, through these Ala-block mutagenesis studies, the most critical residues within the Taf4 RBD were much more narrowly defined as a result of this experiment, going from 90 amino acids to less than 34 residues of interest.

Collectively, these data suggest functional interaction(s) between two separate stretches of Taf4 residues with Taf5 NTD2 residues. By inference, this result might indicate that the Taf4 RBD utilizes two different portions/domains to participate in the same cellular process as the Taf5 NTD2, which I presume to include functional interaction with Rap1. If those two Taf4 RBD components are indeed contributing to the same process, I predicted that combining mutations of each of these Taf4 RBD 'domains' would result in a synthetic growth deficiency. To test this hypothesis I constructed a series of *taf4* alleles where two distinct stretches of residues were each converted to alanines within the same protein, and tested the ability of these mutants to support growth (**Figure 4.3**). The previous negative and positive controls were included and an additional Taf4 variant was created that lacked residues 274-310, since this deletion was

predicted to have little to no effect on Taf4 function if those residues are truly inconsequential. Six different single alanine scan alleles are shown first; the two mutants affecting residues 311-322, and four additional mutants with two of those flanking the amino terminal side of 311-322 and two flanking the carboxy terminal side. Each of these six mutants was combined with one of four different alanine block substitutions; three of these affected residues 251-274 while the fourth altered residues 275-280 and was expected to be less likely to interact genetically with the six mutants affecting residues 299-334. Addition of the 251-256 alanine block to the mutants with alanine blocks in residues 299-334 had a clear and selective affect on the 311-316 and 317-322 ala mutants, 251-256 caused synthetic lethality when combined with the 317-322 block and synthetic temperature sensitivity at 30° and 37° when combined with the 311-316 block. A similar overall pattern of synthetic genetic interactions was observed when the 275-280 block mutant was combined with the 299-334 block mutants; the 311-316 or 317-322 block mutants interacted genetically with residues 275-280 while the others did not. Combination of the 264-269 ala block or the 269-274 ala block with the 299-334 block mutants had a more complex and severe pattern of synthetic genetic interaction than 251-256 or 275-280. The 311-316 and 317-322 blocks were both synthetic lethal with either of the 264-269 or 269-274 blocks. Most interestingly, the 264-269 ala block resulted in synthetic temperature conditional growth at 37° when combined with either 299-304 or 306-311 ala blocks; similarly the 269-274 block was also synthetically sick with the 299-304 block at 37°. Therefore in addition to revealing overlapping function of several distinct regions of the Taf4 RBD, this study generated several novel temperature conditional *taf4* alleles that very precisely affect specific groups of amino acids, and

could eventually prove useful in testing models of Rap1-TFIID interaction. Collectively these results are in agreement that the most critical regions of the Taf4 RBD reside amongst residues 251-374 and 311-322, although it is possible that residues 299-310 provide some function as well.

There are many implications and future directions hinted at by this systematic alanine scanning mutagenesis study of Taf4. It is worthwhile to review the finding that residues within amino acids 251-374 and also amino acids 311-322 are important, by revisiting the amino acid substitutions encoded by the three previously discussed *taf4* temperature conditional alleles, *taf4-116, taf4-141*, and *taf4-219*. Each of these three mutants contains coding changes of residues within these intervals; E269V, R273S, and L318R in *taf4-116*, A257V and W319R in *taf4-141*, while the *taf4-219* allele includes the S320P substitution. However, all three alleles also include other substitutions and thus it would be interesting to know if the changes in the intervals of interest are sufficient to confer the conditional growth phenotypes. This may be less interesting to pursue since the transcription phenotypes of those mutants were mild. In trying to get the 'right' mutants I am after, it would probably be more expedient to make substitutions at smaller intervals within residues 311-322, perhaps changing just one or two consecutive residues to alanine per mutant. If this study identified the exact residue(s) with the most critical contribution to growth, a codon randomization experiment could be used to look at how the identity of the amino acid used for substitution contributes to growth phenotype. Hopefully altering the substituent from alanine to the other 18 possible residues would identify a range of phenotypes from mild to severe loss of growth. This systematic strategy could also be extended to the 251-274 interval to comb through those residues.

Finally, to look for synthetic interactions, interesting mutants with one or a few substitutions in residues 311-322 could be combined with more precise substitutions in 251-274. Given the observation that every single or double ala block scan mutant encoded an apparently stable protein (**see figure 4.3, part two**), it seems likely that this detailed mutagenesis strategy would prove successful in generating 'better' mutants. At that point, I think the goal of finding the best mutants possible will have been reached. Before pursuing more extensive site-directed mutagenesis, I think it would be wise to examine the transcription and Rap1-binding characteristics of the newly available *taf4* ala scan mutants. Given the more severe growth phenotypes, there should be a more profound affect on RPG transcription and Rap1 binding than with those mutants already characterized. Alternatively, it is possible that the properties of these latest mutants have nothing to do with Rap1 binding. In fact, residues including 311-322 are reported to participate in direct Taf4 DNA binding in vitro (Gazit et al., 2009; Shao et al., 2005). The in vivo significance of this observation is much more obscure, since the exact DNA binding properties remain rather ill defined. These findings illustrate other possible roles for the functionality of these Taf4 residues, which may or may not be relevant to Rap1 binding. Additionally, the Ala-scan studies of the Taf4 RBD did reveal the continuation of a trend with respect to the Rap1-Taf-TFIID interaction. I already showed that several of the proteins involved, including Rap1 and Taf5, appear to utilize multiple domains for their interactions. Both the DBD and C-terminal domains in Rap1 and the NTD1 and NTD2 of Taf5 are relevant; here in the case of Taf4 we once again see that two separate stretches of amino acids appear genetically important for the biological function of the

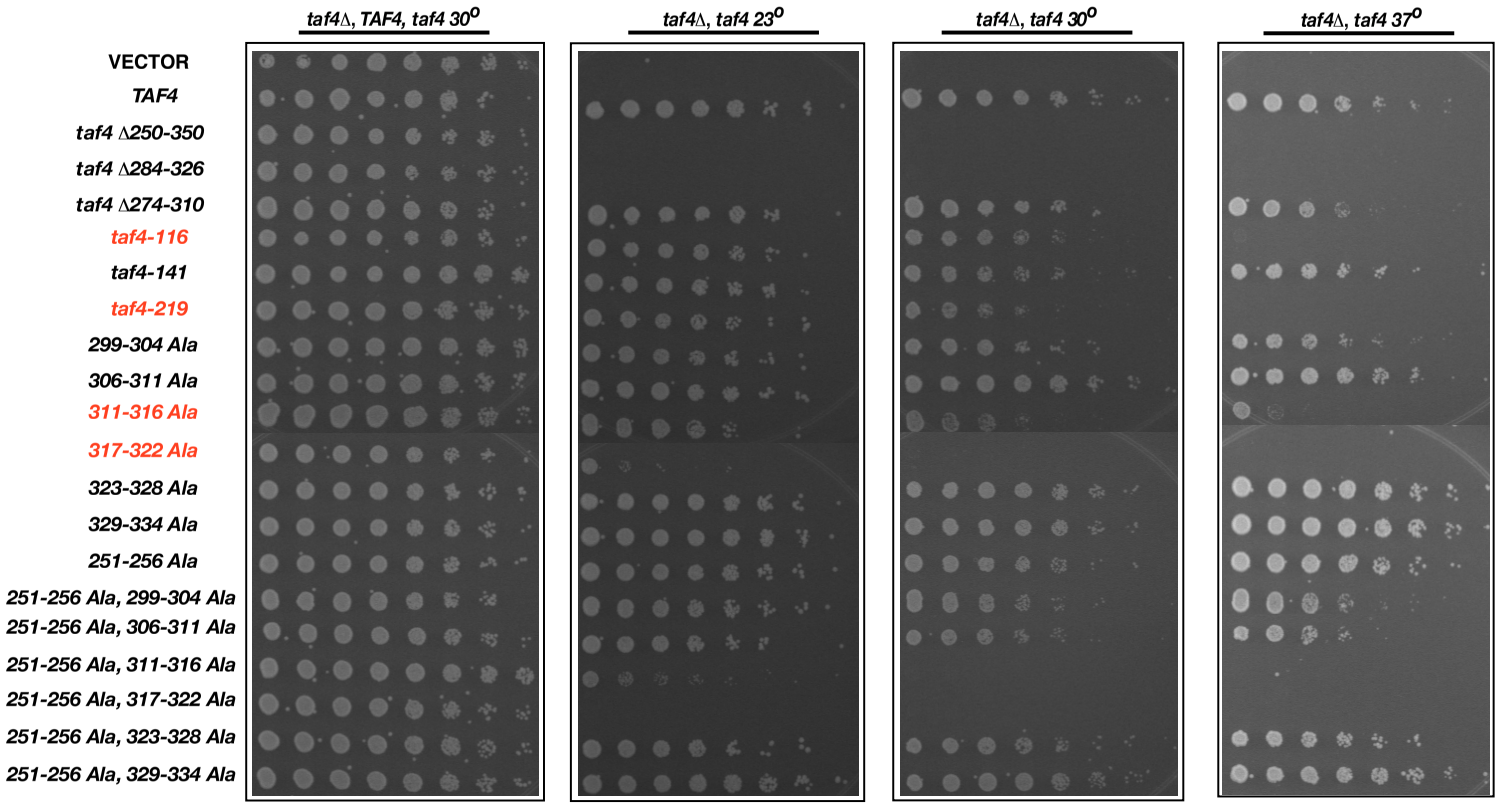
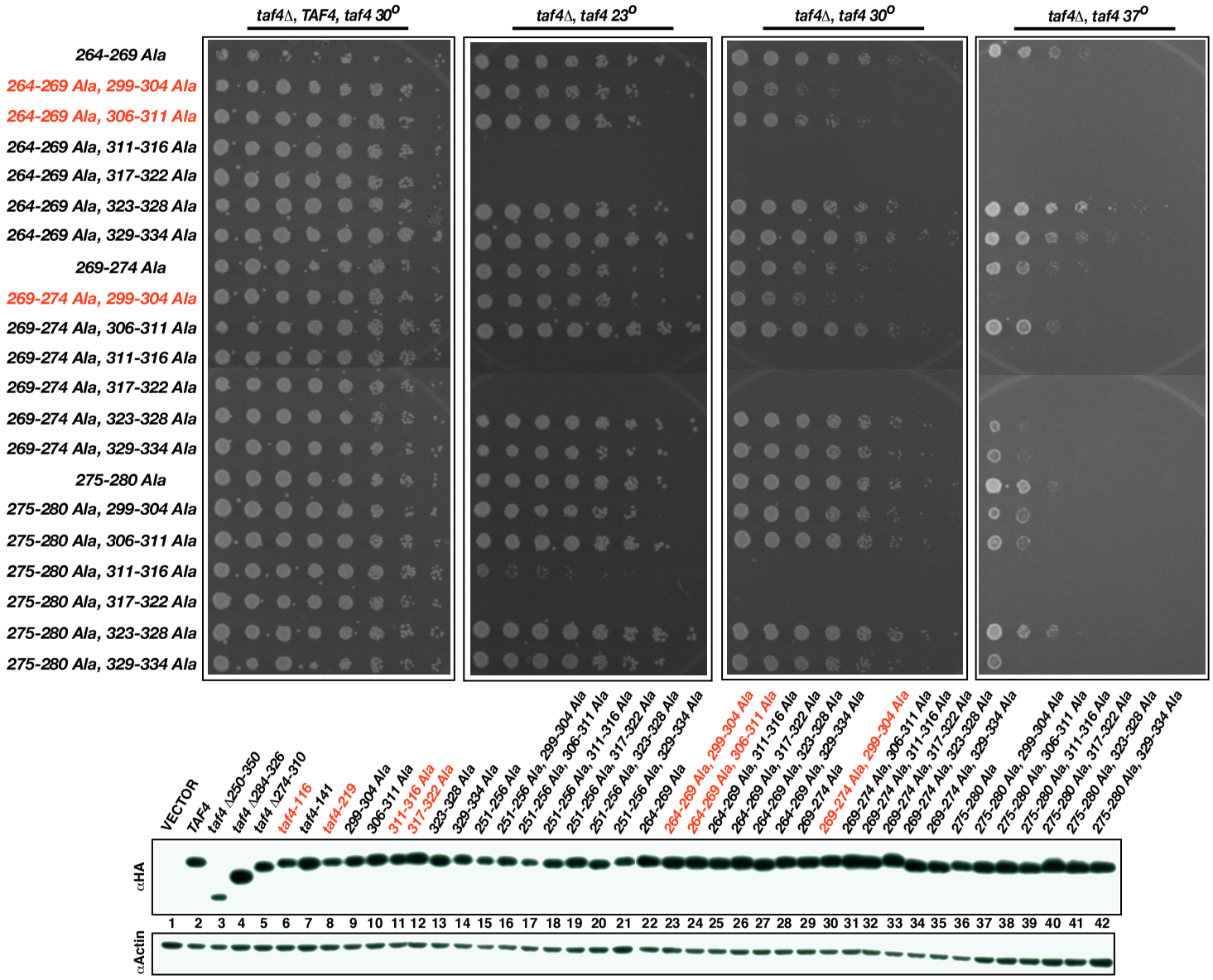


Figure 4.3: Synthetic genetic interaction of two separate portions of the Taf4 RBD-encoding sequence. The *taf4* ala scan mutants containing a single six residue alanine block conversion and *taf4* ala scan mutants containing two separate alanine block conversions were subjected to standard plasmid shuffle assays at various temperatures. Certain combinations of ala blocks cause synthetic phenotypes, whereas other combinations have no affect. The most profound affects occur when blocks affecting residues 263-274 are combined with blocks affecting residues 311-322. However, there are numerous complex patterns of genetic interactions. Variants labeled in red exhibit tight temperature conditional growth phenotypes and may have further use as reagents to look at affects on RPG transcription. Collectively, these data support the idea that two separate contiguous portions of the Taf4 RBD may be participating in the same molecular process, which may include interaction with Rap1. It remains to be seen how Rap1-Taf4 physical interaction/binding is affected by the various combinations of amino acid substitutions. Additonal growth data and the companion immunoblot are shown on the next page.


taf4Δ, TAF4, taf4 30°
taf4Δ, taf4 23°
taf4Δ, taf4 30°
taf4Δ, taf4 37°
264-269 Ala
264-269 Ala, 299-304 Ala
264-269 Ala, 306-311 Ala
264-269 Ala, 311-316 Ala
264-269 Ala, 317-322 Ala
264-269 Ala, 323-328 Ala
264-269 Ala, 329-334 Ala
269-274 Ala
269-274 Ala, 299-304 Ala
269-274 Ala, 306-311 Ala
269-274 Ala, 311-316 Ala
269-274 Ala, 317-322 Ala
269-274 Ala, 323-328 Ala
269-274 Ala, 329-334 Ala
275-280 Ala
275-280 Ala, 299-304 Ala
275-280 Ala, 306-311 Ala
275-280 Ala, 311-316 Ala
275-280 Ala, 317-322 Ala
275-280 Ala, 323-328 Ala
275-280 Ala, 329-334 Ala
VECTOR
TAF4
taf4 Δ250-350
taf4 Δ284-326
taf4 Δ274-310
taf4-116
taf4-141
taf4-219
299-304 Ala
306-311 Ala
311-316 Ala
317-322 Ala
323-328 Ala
329-334 Ala
251-256 Ala
251-256 Ala, 299-304 Ala
251-256 Ala, 306-311 Ala
251-256 Ala, 311-316 Ala
251-256 Ala, 317-322 Ala
251-256 Ala, 323-328 Ala
251-256 Ala, 329-334 Ala
264-269 Ala
264-269 Ala, 299-304 Ala
264-269 Ala, 306-311 Ala
264-269 Ala, 311-316 Ala
264-269 Ala, 317-322 Ala
264-269 Ala, 323-328 Ala
264-269 Ala, 329-334 Ala
269-274 Ala
269-274 Ala, 299-304 Ala
269-274 Ala, 306-311 Ala
269-274 Ala, 311-316 Ala
269-274 Ala, 317-322 Ala
269-274 Ala, 323-328 Ala
269-274 Ala, 329-334 Ala
275-280 Ala
275-280 Ala, 299-304 Ala
275-280 Ala, 306-311 Ala
275-280 Ala, 311-316 Ala
275-280 Ala, 317-322 Ala
275-280 Ala, 323-328 Ala
275-280 Ala, 329-334 Ala
αHA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42
αActin

RBD therein (Garbett et al., 2007; Layer et al.). It remains to be seen how much of this participation of multiple distinct domains within individual proteins is a reflection of intrinsic biological redundancy. The involvement of multiple Tafs in TFIID, along with multiple non-contiguous stretches of amino acids in both Rap1 and the Tafs, are each indicative of the importance of multiple distinct mechanisms of contact.

Even though my latest mapping experiments have shed light on the functionality of the Taf4 RBD, they are not an end in themselves, instead pointing to yet more additional experiments. It is to be expected that interaction between a large transactivator and a multisubunit coregulatory factor like TFIID will be complicated, so it is also no surprise that these mapping experiments are equally difficult. The detailed site-directed mutagenesis strategy described in this chapter should be broadly applicable, not just to Taf4 but also to Taf5, Taf12, and Rap1, or any and every additional factor we eventually define as contributing to Rap1 and TFIID-dependent RPG transcription. All that is required is the limited information that I already have about specific protein domain involvement, and basic techniques of molecular biology and yeast cell biology. Most importantly, this strategy has the potential to generate extremely precise genetic reagents, which will be all the more useful in downstream genetic and biochemical experiments. The systematic residue-by-residue mutagenesis strategy represents a rigorous approach to defining the relevant contiguous amino acid residues in each protein, but there is perhaps a more precise and possibly more rapid approach to defining the minimal residues needed for Rap1-Taf interactions. Although most of the biochemical interaction studies used recombinant proteins from *E. coli*, hence lacking yeast-specific and possibly regulatory post-translational modifications, it could be envisioned that cells utilize alterations of

specific residues to modulate the affinity and stability of Rap1-Taf interactions. More profoundly, there could be situations where Rap1-TFIID interaction is not appropriate, and cells might exploit reversible PTM(s) to disrupt these interactions. One would presume such modifications would be located at residues directly involved in the interaction, or at least proximal to the relevant protein domains. Description of such a phenomenon could help us define Rap1- and Taf-binding domains with precision, without having to first generate and sort through large quantities of site-directed mutants.

How Is Rap1-TFIID Interaction Dynamically Regulated In Vivo?

Before discussing identification PTM of Tafs and/or Rap1, it would be instructive to consider physiological situations where interaction of Rap1 with TFIID would be inappropriate and thus negatively modulated to reduce RPG transcription rates. Within that situation, simple mechanisms might include disruption of Rap1-TFIID interaction on RPGs in situ, without disruption of protein-DNA interaction, or disruption of factor-DNA binding with consequent relocation of one or both factors, or degradation of enhancer/promoter-bound proteins in situ. In all likelihood, combinations of these mechanisms with each other or any other unlisted possibilities are operative. Given that the proteins used for in vitro characterization of Rap1-Taf interaction, in particular Rap1, were prepared in *E. coli* and thus lacking yeast-specific phosphorylation, ubiquitylation, acetylation, et cetera, it seems likely that addition of modification(s) to either TFIID or Rap1 is the best bet for inhibition of the protein-protein interactions and/or RPG transcription, when appropriate. This proposal does not mean we can rule out upregulation of factors mediating removal of modifications normally stimulatory to the

interactions and RPG transcription. In any case, we might ask if any effect on these proteins is observable as a result of an environmental condition associated with down-regulation of RPG transcription. It is well documented that inhibitors of the TOR signaling pathway, such as with rapamycin, reduce bulk protein translation through inhibition of the ribosomes, but an affect also occurs through reduction in RP synthesis via reduced RPG transcription (Cardenas et al., 1999). In fact, several activators such as Fhl1/Ifh1, Sfp1, Hmo1, and the NuA4 coactivator each become dissociated from RPG enhancers when rapamycin is added to yeast cells (Berger et al., 2007; Lempiainen et al., 2009; Rohde and Cardenas, 2003; Schawalder et al., 2004). While Rap1 remains tightly bound to RPG enhancers in rapamycin-treated cells, an affect on TFIID occupancy has not been reported (Schawalder et al., 2004). Thinking that modifications of Rap1 and/or TFIID might take place during these conditions, I sought to examine the occupancy of RPGs by Rap1, Tafs, and other GTFs during rapamycin treatment in vivo. To do this I used ChIP (chromatin immunoprecipitation) with polyclonal antibodies to Pol II, TBP, Tafs, and Rap1 after exposure of cells to either rapamycin or DMSO vehicle (Hecht and Grunstein, 1999). DNA recovered by ChIP was measured by quantitative multiplex PCR and data are shown in **Figure 4.4**.

The multiplex PCR assay scored DNA content from the natural *RPS5* gene; several primer pairs were used to document relative enrichment of several distinct *RPS5* segments including the open reading frame (ORF), the enhancer where Rap1 binds, and the promoter where TBP, Tafs and Pol II bind. A primer pair recognizing the telomere-proximal region of Chromosome VI (*TELVIR)* was used to provide an internal positive and negative control; this genomic region which is bound and 'silenced' by Rap1 is

transcriptionally inactive and therefore expected to provide little enrichment of TBP, Tafs, and Pol II above background (Adkins et al., 2004). The mobility of each of the four PCR products is demonstrated in the first lane, as amplified from control unsheared, uncrosslinked genomic DNA. The next lane demonstrates a lack of PCR contamination that could confound interpretation of results, as evidenced by the absence of any PCR products in a reaction that received no template DNA. Next are the samples representing the input chromatin used for each immunoprecipitation, one each for cells treated with either vehicle DMSO or rapamycin. A small fraction of the DNA contained within each of the total inputs was serially two-fold diluted and analyzed by PCR; the approximate two-fold increase in signal abundance for each PCR product per serial dilution demonstrates that the assay was within the linear range of amplification and detection, and that each genomic region was indeed approximately equally present in each input sample. The next two lanes represent the DNA content of ChIPs for the largest subunit of RNA Pol II, from cells cultured either in the absence or presence of rapamycin, respectively. In the untreated sample, the most intense bands correspond to the core promoter and ORF, respectively most likely representative of Pol II engaged in either PICs or actively transcribing the *RPS5* gene body/ORF. There is a marked reduction in Pol II occupancy of both the core promoter and ORF in chromatin prepared from cells treated with rapamycin, consistent with the well-documented dramatic reduction in *RPG* transcripts that occurs as a result of this treatment (Cardenas et al., 1999). To score protein occupancy of a PIC component that should be specifically enriched to the promoter, TBP association was also examined. Like Pol II, TBP was specifically

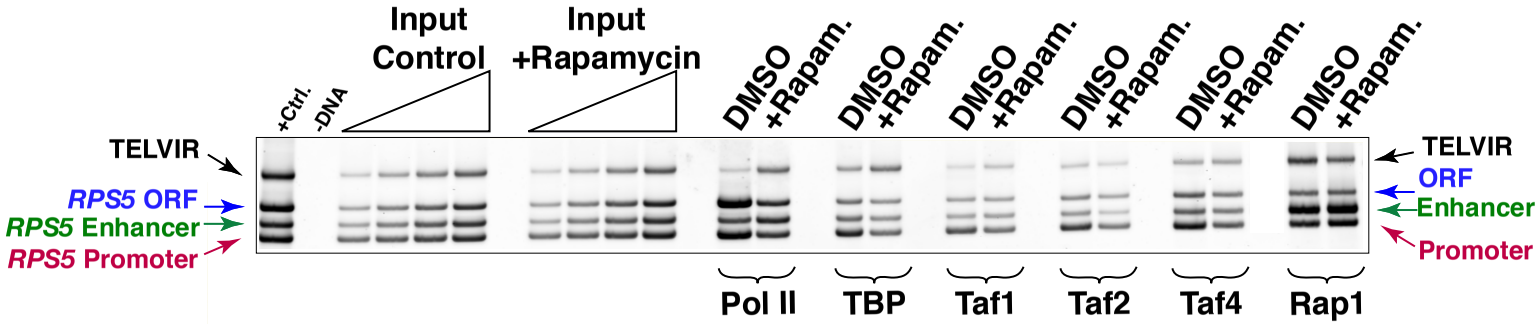


Figure 4.4: ChIP of transcription proteins present on *RPS5* in the absence and presence of rapamycin, an inhibitor of RPG transcription. Shown are the products of quantitative multiplex PCR reactions that measured the abundance of several genomic locations in total sheared chromatin and material from ChIP of those samples. Two cultures were grown and treated with either DMSO or rapamycin for two hours, then processed for ChIP. Rapamycin is a small peptide inhibitor of a prolyl isomerase important for the activity of the TOR kinase, which is a positive upstream effector of signalling cascades stimulatory to ribosome function and RPG transcription. The inputs are serial two-fold dilutions of a 1:100 dilution of the material input to one ChIP reaction. The equivalency of signal intensity of each band and the same dilution profile from the two different inputs indicates that the PCR was within the linear range of detection and that DNA concentration in each input was apparently identical. ChIP of Pol II indicates strong enrichment in the promoter and the *RPS5* ORF, consistent with its role in synthesizing *RPS5* mRNA; however these enrichments do not occur in chromatin of rapamycin-treated cells, consistent with a cessation of RPG transcription as a result of rapamycin treatment. The same phenomenon ocurred in the case of TBP and Tafs, except that proteins were only enriched at the promoter, consistent with their participation in RPG transcription as members of TFIID, functional as a promoter-bound PIC component. TBP and Tafs also disappeared upon rapamycin treatment, consistent with TFIID promoter association at *RPS5* being a highly regulated event. Given the direct inhibition of cell signalling resulting from rapamycin treatment, it is possible that TFIID Tafs are subject to PTM mediated by one or more pathways. On the other hand, Rap1 was highly enriched on the enhancer independent of rapamycin treatment. However, this does not indicate whether Rap1 maintains the same PTM profile as a result of rapamycin treatment. Perhaps TFIID Tafs and Rap1 are regulated by signalling pathways which influence their ability to interact and productively stimulate high-level RPG transcription. It will be interesting to document the spectrum of PTMs occurring on Tafs and Rap1 in the absence and presence of rapamycin.

enriched at the promoter in the absence but not in the presence of rapamycin, again indicative of a dramatic reduction in *RPS5* transcription rates consequent with drug application. Interestingly, both TBP and Pol II occupancy at *TELVIR* appeared to increase in the presence of rapamycin. The significance of this observation remains unclear. TBP is a TFIID component, but its presence in other cellular complexes does not necessarily indicate its participation in events at *RPS5* as a member of TFIID (Kuras and Struhl, 1999). Therefore the occupancy of three TFIID-specific Taf subunits Taf1, Taf2, and Taf4 was also scored. Each protein behaved identically to TBP in that all three were specifically enriched at the *RPS5* core promoter in the absence, but not the presence of rapamycin. These patterns of occupancy are entirely consistent with Tafs 1, 2, 4 and TBP participating in *RPS5* regulation as constituents of TFIID, but more importantly these results indicated that TFIID disappears (by relocation or by destruction) from the core promoter upon rapamycin treatment. Finally, as a control, Rap1 occupancy at *RPS5* was examined under both conditions. Rap1 was enriched at the enhancer and consistent with the literature, this pattern of occupancy did not change with rapamycin (Schawalder et al., 2004).

These ChIP results demonstrate that inhibition of the growth-responsive TOR signal transduction pathway results in marked alteration of TFIID association with an RPG promoter. That TOR itself is a protein kinase that phosphorylates many targets, including downstream effector kinases and phosphatases to modulate their activity (De Virgilio and Loewith, 2006), suggests a role for reversible protein phosphorylation in regulation of TFIID/Taf activity. In this case TOR is believed to function in concert with cyclic AMP-dependent protein kinase (PKA) to stimulate RPG transcription, although the

exact mechanisms are not clear (Lempiainen and Shore, 2009), but perhaps includes direct affects on reversible protein phosphorylation of factors on RPGs. Activated signalling kinases have been shown to physically occupy enhancers and promoters of target genes, so it is possible any relevant phosphorylation events could be dynamically occurring directly on Tafs while they are present on RPG promoters (Pokholok et al., 2006). Stimulatory kinases might affect modification of Taf and/or Rap1 residues that facilitate interaction between the proteins; these kinases could depend on active TOR signaling for their action on substrates. Alternatively, negatively acting kinases could be de-repressed in the absence of TOR signaling, ultimately leading to the modification of Taf and/or Rap1 residues that disrupt their interaction. This possibility can be connected with the observation that alteration of specific Taf4 residues within its RBD leads to growth phenotypes, as shown earlier in this chapter. In the case of Taf4 serine 311, conversion to aspartate, a phospho-mimetic, leads to synthetic growth phenotypes when combined with Taf5 mutants, while conversion of S311 to alanine has no such affect. It would be interesting to look for affects on Taf4 phosphorylation coincident with rapamycin treatment of cells, to see if S311 phosphorylation becomes upregulated. The S311A mutant could be resistant to the affect of the drug on RPG transcription if modification of that residue is truly important for the affects on that process. To ask more broadly how phosphorylation might influence RPG behavior, it would useful to conduct a proteomic analysis of PTMs occurring on Tafs and Rap1 isolated from cells cultured in the absence and presence of rapamycin (Huber et al., 2009). Differences in phosphorylation status of Taf RBD residues between the two conditions might indicate regulation of the Rap1-Taf interaction, and suggest those residues directly involved.

Subsequently, mutations could be introduced that change the identified serine or threonine residues to either alanine or aspartate. Cells containing the altered Tafs as the sole source of the proteins could then be treated with rapamycin and examined for an altered affect on RPG transcription status. A connection between individual modifications and candidate kinases could be pursued using available deletion strains missing the genes for those enzymes. Phospho-specific antibodies could be developed for use in kinetic ChIP assays that measure Taf/Rap1 isoform occupancy during a time course of rapamycin treatment. In parallel, the affect of kinases on in vitro interaction between recombinant Tafs and Rap1 could be examined using our established quantitative methods (see **Figure 3.7**); such assays could be supplemented with cognate Taf and/or Rap1 variants containing amino acid substitutions at particular residues.

While there is evidence suggesting that phosphorylation of Rap1 and/or Taf4 may be important, this is not the only candidate PTM that could modulate their activity. Sumoylation and ubiquitination are also reasonable possibilities; with published data showing the NTD regions of Taf5 are directly sumoylated in human cells (Boyer-Guittaut et al., 2005). Sumoylation and ubiquitination are modifications classically associated with alteration of protein abundance (Bloom and Pagano, 2005; Sacher et al., 2005). Regulated protein stability fits with my observation that deletion of the NTD2 is sufficient to cause a dramatic increase in Taf5 abundance. Residues between NTD2 214-263 are probably involved (see **Figure 2.6**), this domain includes three potentially modifiable lysines, although Taf5 stabilization is increased when NTD1 is also deleted. Thus it is possible that yeast also utilize sumoylation or ubiquitination of NTD lysine residues to modulate Taf5 stability and/or function. Some of my Co-IP data also

indirectly suggests that Rap1 is subjected to a modification that alters its molecular weight to an extent consistent with a large PTM like ubiquitination (see **Figure 3.4**). Interestingly, this effect occurs more prominently under buffer conditions including elevated salt concentrations or with larger amounts of ethidium bromide, both of which may suggest a relationship between Rap1 DNA binding status and this PTM pattern. Consistent with this hypothesis, the Rap1 DBD is highly enriched in lysine residues, as might be expected given its interaction with negatively charged DNA, however one or more of these residues might be ubiquitinated or sumoylated under certain conditions, which could in turn alter the conformation of the Rap1 DBD, change stability, or modulate Rap1's capacity for protein-protein interaction with TFIID Tafs. Any relationship between factor ubiquitination might be related to the function of the proteasome, which functions to degrade poly-ubiquitinated protein substrates (Verma and Deshaies, 2005). The proteasome also appears to directly participate in RPG transcription, where its exact function is completely unknown, but may be related to turnover of regulatory proteins in situ on promoters (Auld et al., 2006; Collins et al., 2009; Leung et al., 2008). Perhaps proteasomal activity on RPGs is stimulated after cells are treated with rapamycin, as a consequence of increased ubiquitination of enhancer/promoter-bound proteins such as Rap1 or Tafs. The net effect would be observed as the decrease in factor occupancy of RPG regulatory sequences in treated cells, as scored by ChIP. Proteomic approaches would be useful in detection and mapping of sumoylation or ubiquitination of Rap1 or Tafs, under different growth conditions (Kaiser and Wohlschlegel, 2005). The location of any modified residues could be correlated with sensitivity to amino acid substitutions of those residues, which

could also be correlated with any affect on Rap1-Taf interaction. Finally, lysine acetylation of Rap1 and/or Taf residues could conceivably be used to modulate interactions. The prevalence of this modification in non-histone proteins is only being appreciated, and the utilization of differential protein acetylation in altering protein-protein and protein-DNA interaction is well documented in the case of the highly studied p53 transactivator (Gu and Roeder, 1997; Kruse and Gu, 2008; Li et al., 2007; Loewer et al.; Luo et al., 2004; Tang et al., 2008). It stands to reason that acetylation could influence the behavior of our proteins of interest, particularly since they are dynamically regulated in vivo, with a ChIP occupancy pattern mimicking that of the NuA4 coregulatory acetylase complex but opposite that of the Rpd3 deacetylase complex (Humphrey et al., 2004; Rohde and Cardenas, 2003). NuA4 has also recently been shown to acetylate numerous non-histone substrates, so it is reasonable to propose that TFIID and/or Rap1 could be modified directly by NuA4 (Lin et al., 2009). A combination of proteomics approaches to map PTMs, site-directed mutagenesis of modified residue codons with genetic, cell biological, and biochemical assessment of phenotypic affects will likely tell us important information about Rap1 and Taf binding domains (Mayor and Deshaies, 2005). This could all be tied together by a pursuit of the kinases, phosphatases, acetylases, deacetylases, ubiquitin/SUMO ligases, deubiquitinases/SUMO proteases, and proteasome components directly responsible for Rap1 and Taf modification status, and suggest factors contributing to alteration of factor promoter occupancy in response to environmental changes. We have several good leads since specific kinases, acetylases, deacetylases, and the proteasome are all directly tied to RPG transcription. Clearly there is a wealth of biological complexity remaining to be

defined, and pursuit of those aspects discussed above is intended to better define the exact mechanism of Rap1-Taf interaction with respect to the protein domains involved. However, these questions still do not address the fundamental unanswered problem of my research interest: How does Rap1-TFIID interaction modulate RPG transcription?

The Molecular Mechanistic Outcome of Rap1-TFIID Interaction: Re-examination of Bacterial Paradigms and Evidence of Dramatic Conformational Alterations of Rap1-TFIIA-TFIID-DNA Quaternary Complexes

All of the experiments described in this chapter thus far, both completed and in progress, are intended to provide a more precise definition of those Taf4, 5, 12, and Rap1 residues involved in Rap1-TFIID interaction. At one level this is a continuation of my thesis work, but as an approach, this set of experiments does not promise to give us a rapid understanding of the remaining fundamental problem, which is to ask why these factors interact, and exactly how the interaction contributes to high-level RPG transcription. Rather, genetic studies will, at least initially, provide further proof of the importance of Rap1-Taf interaction, and also provide more suitable/appropriate reagents for additional biochemical, biophysical, and cell biological experiments intended to test the existing model.

A major deficiency of all my studies was the lack of investigation in the context of the TFIID holo-complex. The influence of target gene DNA binding on Rap1-TFIID interaction and also the likely participation of other TFIID-related factors such as TFIIA were only tangentially examined. This lack of investigation relates to technical issues with conventional approaches such as EMSA or DNase I footprinting: the optimal buffer

conditions for Rap1-DNA complex stability in vitro are not compatible with optimal conditions for TFIID-DNA complex stability (i.e. TFIID-DNA complexes are stabilized by magnesium while Rap1-DNA complexes are destabilized). The typical EMSA assay for TFIID-DNA binding utilizes a magnesium-containing agarose gel and observable TFIID-DNA gel shift also requires inclusion of TFIIA; unfortunately magnesium precludes detection of Rap1-DNA complexes in EMSA (Vignais et al., 1987; Zerby and Lieberman, 1997). Thus the most robust TFIID DNA binding protocol is not useful for studying Rap1-TFIIA-TFIID-DNA complexes. DNase I footprinting is much less sensitive than EMSA in terms of quantifying alterations in binding properties, such as affinity binding constants, but simultaneous Rap1 and TFIID footprints on an RPG promoter fragment were observed by this method. Classical studies on transfactor-TFIID interaction, as performed by the Tjian lab, suggested that an enhancement of TFIID binding to the promoter DNA might occur in the presence of Rap1. In other words, Rap1 might help 'recruit' TFIID to the promoter DNA, through an enhancement of TFIID-DNA binding affinity. The DNase I footprinting studies conducted in the lab failed to demonstrate stimulation of so-called TFIID 'recruitment' to promoter DNA (not shown). As is often the case, a negative result defies meaningful interpretation, and we decided to turn to other methodologies to investigate potential conformational alterations of TFIID resulting from interaction with Rap1. In theory some of the weaknesses of EMSA or footprinting, or any other bulk population measurement of molecular structure(s), could be circumvented through use of single-particle analyses. High-resolution electron microscopy (EM) coupled with image refinement and statistical modeling was already in use in the lab of our colleague and Tony's longtime collaborator, Dr. Patrick Schultz.

Since we had purified Rap1, TFIID, TFIIA, and promoter DNA fragments in hand from our biochemical studies, we sent these reagents to Dr. Schultz and his colleagues for examination of the various combinations using EM (Papai et al., 2010). This structural biology approach promised to give a picture on the conformation of individual molecules, which could recover certain information lost when using population biochemical techniques. But before discussing these findings, it would be instructive to first revisit some of the characteristics of prokaryotic activator-coregulator interactions, as that might help us interpret some of our own observations from the EM experiments.

Recall that most prokaryotic activators operate by directly contacting the RNA Polymerase, instead of using coactivators like TFIID, which can bridge between activators and Pol II. Two of the three simple mechanisms discussed in Chapter One involved direct activator-Polymerase contact while the third instead involved an activator-dependent physical rearrangement of the promoter DNA. The two instances of protein-protein interaction-mediated activation may be instructive for our purposes since they demonstrate that different subunit 'targets' exist within the repertoire of different prokaryotic activators, perhaps like different Taf subunits in TFIID being contacted by Rap1. However, the outcome of the prokaryotic protein-protein contacts most commonly involves 'recruitment' or 'stabilization' of Polymerase binding to the promoter. As mentioned just above, we did not consider 'recruitment' as a viable model in the case of Rap1 and TFIID on RPGs, since we had obtained only negative data regarding Rap1-mediated stimulation of TFIID DNA binding. Still, the existence of different surfaces being bound by activator and the conformational changes in the regulatory DNA sequences are both interesting possibilities hinted at by the simple models of prokaryotic

activator-Polymerase association. For our purposes, in putative EM studies these phenomena might manifest as distinct localization of Rap1 binding to TFIID regions containing several target Tafs, and/or observable changes in DNA conformation or direction resulting from Rap1-TFIID binding. Since we are dealing with a much more complicated system, it might be instructive to compare our situation to that of the models for independent activator inputs to *E. coli* RNA polymerase, as shown in **Figure 1.9**. In contrast to these situations, we are not studying two activators but just one, but for the sake of simplification TFIID might be thought of in the role of a second activator, since TFIID is also in contact with both an activator (Rap1) and also the promoter DNA (Griffith et al., 1986; Hochschild and Ptashne, 1986). Therefore, a prokaryotic situation of one or more activator/DNA complexes binding to a Polymerase/DNA complex might best be compared to a eukaryotic situation of an activator/DNA complex binding to a TFIID/DNA/TFIIA complex, as these represent the minimal eukaryotic components needed to observe an affect on factor or DNA conformational changes resulting from activator-coregulator contact. In total, the interactions between Rap1/DNA and TFIID/DNA complexes might cause significant alteration of the physical properties of each other at the level of both protein and DNA conformation; stabilization of DNA binding of either factor seems difficult to prove, but we are still left with the possibility that Rap1-TFIID interaction could stimulate a favorable conformation of a TFIIA/TFIID 'partial' PIC. The Rap1-engaged 'partial' PIC might be uniquely permissive to interaction with other downstream core promoter factors or the Polymerase itself. This is comparable to those prokaryotic systems where one or more activators interact with Polymerase to put it in a conformation favorable for activation of transcription.

There are three important characteristics of the EM experiments that allowed us to use this method. First, the basis is the analyses of a large number (10^4-10^5) of single molecules, which in total allows for calculation of a statistical average molecular shape. Second, the resolution allows us to see large macromolecular assemblies, but subunit-sized entities are not necessarily observable by themselves. Instead, the location of entities that bind TFIID are inferred on the basis of a 'difference map' that compares the shape of measured TFIID structures, that are analyzed simultaneously in the absence and presence of additional factors such as Rap1 and/or TFIIA. The extra volume of densities and/or domain positions, inferred as being occupied by TFIID-bound entities, can be compared to known X-ray crystal structures, available for portions of Rap1, TFIIA, and TBP, which provides a means of validating the identity of any observed domains. Third, from many previous experiments, we have determined the locations of all Tafs and TBP in the complex (Leurent et al., 2002; Leurent et al., 2004; Papai et al., 2009). This information can be compared to the locations of TFIID-bound entities to infer which Taf subunits are being bound by those entities.

As shown in **Figure 4.5**, when Rap1 was incubated with TFIID and imaged by EM, the location of Rap1 could be inferred from difference maps, and its position corresponded to the region of TFIID known to contain Taf5 and the Taf4/12 heterodimer. This is entirely in agreement with my numerous Rap1-Taf binding studies, which showed that Rap1 could bind specifically and with high affinity to those exact Tafs (Layer et al., 2010). Also importantly, there did not appear to be any dramatic conformational rearrangement of TFIID structure, a finding similar to those of recent EM studies of mammalian TFIID complexed with the activators Sp1, c-Jun, or p53 (Liu et al., 2009).

While each of these mammalian activators bound to distinct locations in TFIID, consistent with their unique repertoires of Taf binding, none of them dramatically changed TFIID structure. Collectively these results indicate that binding of activators to TFIID, in the absence of DNA and/or additional regulatory factors, does not lead to distinct TFIID conformations. Similarly, a conformational change did not occur when TFIID was bound to TFIIA and DNA. An additional electron density corresponding to the approximate size of TFIIA was located in the vicinity of the portion of TFIID known to contain TBP, in accord with the known direct interaction between TFIIA and TBP (Tan et al., 1996). Interestingly, an additional density was localized around the known position of Taf2, but not that of TBP, and this result was interpreted to correspond to the location of the promoter DNA. Thus in the visualized TFIID/TFIIA/DNA complex, TFIID was most likely binding to DNA through Taf2, but not TBP. This idea is supported by the known binding of Taf2 to INR sequences in the adenovirus major late promoter, and the observation made by many labs that binding of purified TFIID and TFIIA to DNA typically gives a DNase I footprint in the vicinity of promoter TATA boxes (Auty et al., 2004; Chalkley and Verrijzer, 1999; Sanders et al., 2002a). It is possible that TFIIA is also contacting TFIID through subunit contacts in addition to TBP, for example within Taf4 (Guermah et al., 2001; Olave et al., 1998). The third structural biology experiment tested the affect of combining Rap1, TFIID, DNA, and TFIIA. Over 10^5 individual molecules were scored to conduct image reconstruction, and three prominent classes of structures were evident. These corresponded to Rap1-TFIID-DNA particles, Rap1-TFIIA-TFIID-DNA particles, and TFIID unbound to additional

components; this last class served as a useful internal control for comparing density shifts within the other two classes of complexes.

The Rap1-TFIID-DNA structure did not exhibit much difference from the Rap1-TFIID structure, except that as expected the density of DNA could be detected in the vicinity of Rap1 and also Taf2, an observation again compatible with sequence-specific DNA binding occurring through those proteins. In this case, there were two additional densities in the region of TFIID containing Tafs 4, 5, and 12 instead of just one density, and several possibilities were given for this characteristic. Since two adjacent Rap1 binding sites were present in the DNA template used, the two densities could correspond to two distinct Rap1 molecules. Since the proposed Rap1 density closer to the core of TFIID was also near a proposed DNA density, and because the structure of the Rap1 DBD (Konig et al., 1996) could be docked into this density, this could correspond to the Rap1 DBD engaging with a subset of the Tafs. However, Rap1 utilizes not just the DBD but also C-terminal domains to interact with TFIID in vitro, so we cannot exclude a situation where this inner Rap1 density instead corresponds to those C-terminal domains (Garbett et al., 2007; Layer et al.). It is possible that, like the DBD, the structure of the relevant C-terminal domains could also be fit into this density, but unfortunately no X-ray structure exists for the candidate C-terminal domains. So at present we cannot discriminate between whether the two densities correspond to distinct Rap1 molecules or multiple Rap1 domains, and if they are indeed multiple domains, that there is a difference in the location of the Rap1 DBD and the Taf-interacting C-terminal domains. The presence of DNA density extending out of the complex in the vicinity of one of the Rap1 densities but not the other, and the fitting of the Rap1 DBD X-ray structure into this

interior Rap1 density can be taken to suggest, but not prove, that the interior density is indeed the Rap1 DBD directly interacting with both DNA and Tafs. This data is all entirely consistent with existing models of Rap1 and TFIID function established using structural, biochemical, and in vivo techniques, by several different labs.

The Rap1-TFIIA-TFIID-DNA structures were the most intriguing for many reasons. There was a greatly increased amount of density around the interior of TFIID, and this was again segregated into distinct entities based upon the fitting of X-ray structures of Rap1 and TFIIA/TBP/DNA complexes, and the known locations of Taf subunits and TBP within the TFIID structure. The innermost density was likely TFIIA, since it is in the vicinity of the location of TBP, again consistent with the TFIIA-TBP interaction and the fact that the TFIIA/TBP/DNA co-crystal structure could be concisely docked into the cryo-EM structure. The outermost density was proposed to be Rap1 since it remained in the vicinity of Tafs 4, 5, and 12, and because the Rap1 DBD structure could again be fitted. Notably, the positions of both TFIIA and Rap1 were slightly different from those found in the Rap1-TFIID-DNA or the TFIIA-TFIID-DNA structures. Furthermore, extra DNA density was observed around the location of TBP, in contrast to the Rap1-TFIID-DNA structure or the TFIIA-TFIID-DNA structure, strongly suggesting that in this form of complex, TFIID was engaged with promoter DNA not just through Taf2, but also through TBP. Since the location and orientation of DNA entering and exiting the protein complex could be more accurately described in this structure, and since Taf2 is believed to bind DNA downstream of TATA DNA, with both TATA and Inr DNA downstream of the Rap1 binding site, there was an indication of 'looping out' of DNA sequences intervening the Rap1-bound enhancer and the TFIID-bound promoter.

The density of this DNA 'loop' could not be detected using cryo-EM protocol most likely due to DNA conformational flexibility. Importantly, a platinum shadowing method confirmed the presence of a large DNA 'loop' in the Rap1-TFIIA-TFIID-DNA quaternary complex. This 'loop', the inclusion of additional DNA density within the TFIID core, and the conformational rearrangement of TFIIA and Rap1 position in the quaternary complex as compared to either of the tertiary complexes were suggestive of functional physical interactions between TFIIA and Rap1. The alteration of TFIID-DNA binding is indicative of a stabilization of a 'committed' TFIID-DNA conformation accomplished only by simultaneous interaction with both Rap1 and TFIIA. Again, the major point of emphasis is that the DNA is conformationally stabilized, or 'locked', in the Rap1-TFIIA-TFIID-DNA complex, and this feature may be key to how RPG transcription occurs so rapidly and efficiently, through an unusually stable TFIID-PIC configuration. The altered location/conformation of TFIIA in the quaternary complex may be significant for downstream events in PIC formation and/or function. Overall, the unique positioning of the regulatory proteins and the DNA in this quaternary complex suggests, at least in part, how RP genes are capable of remarkably high-level transcription rates. We propose that the stabilization of TFIID-promoter association afforded by the combination of Rap1 and TFIIA allows for more rapid rounds of Pol II initiation and/or re-initiation of RPG transcription. Collectively, our observations concerning conformational changes of factors and DNA and enhancer-promoter looping between DNA-bound Rap1 and TFIID represents the first description of such events in the eukaryotic system. Interestingly, similar behavior of factors has been documented in the prokaryotic system, and suggests a variety of possible attributes that may help explain

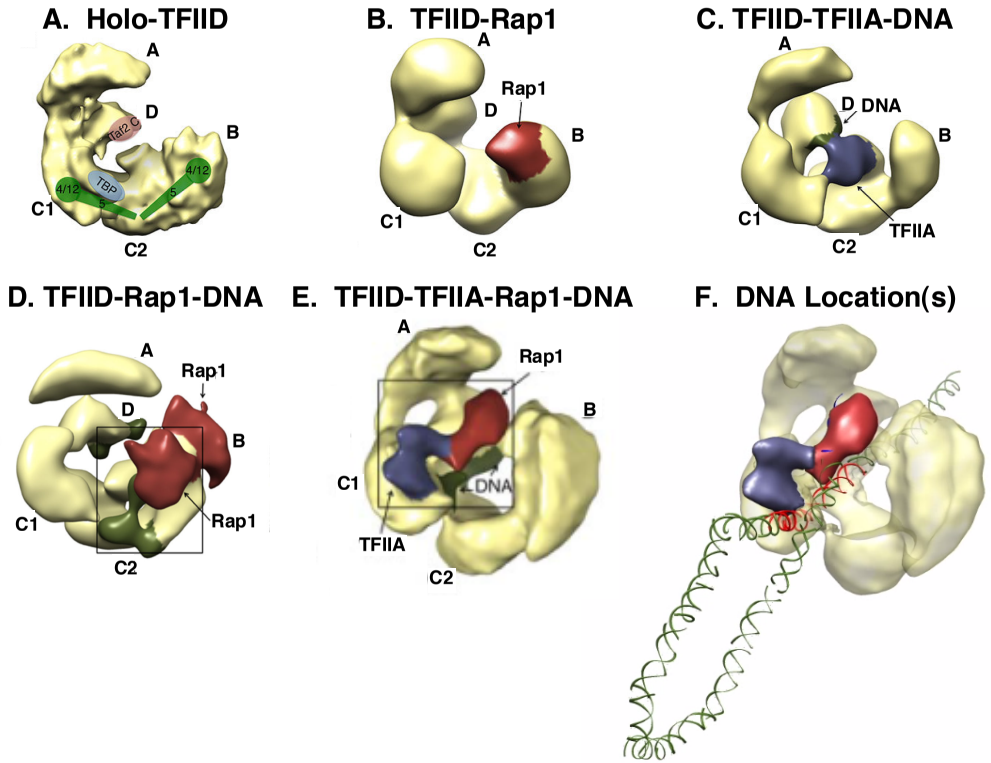


Figure 4.5: Three-dimensional EM analyses of various TFIID-containing complexes. **A.** Holo-TFIID as imaged by cryo-EM. The salient features are the multi-lobed (A, B, C1, C2, D) horseshoe-shaped configuration of the complex, in which the locations of all subunits have been previously defined (Leurent et. al., 2002, 2005, Papai et. al., 2009). Locations of DNA-binding TBP and Taf2 subunits are shown, as well as Rap1-binding Taf4, 5, and 12 subunits. **B.** The TFIID-Rap1 complex. Rap1 forms a discrete density in the vicinity of the locations of Taf4, 12, and 5. **C.** TFIID-TFIIA-DNA complex, where additional TFIIA density is located in the vicinity of TBP, and DNA is in the vicinity of the Taf2 C-terminus, but not around TBP. **D.** TFIID-Rap1-DNA complex. Rap1 is in a similar location as in panel B., but two densities are present, consistent with the presence of two Rap1 binding sites in the included DNA. Consistently, DNA density is in the vicinity of Rap1, and around the Taf2 C-terminus, as before. **E.** TFIID-TFIIA-Rap1-DNA complex. Both TFIIA and Rap1 have undergone a major conformational re-arrangement, and DNA density is now located around Taf2, but also around TBP; indicative of an alteration of mechanism of TFIID-DNA binding, possibly resulting in a stabilized, 'committed' complex. **F.** Platinum shadowing EM assays revealed the existence of a large DNA loop corresponding to sequences in between Rap1-bound elements and TFIID-bound core promoter elements. Data are adapted from Papai et. al., 2010.

the mechanism of PIC function, while hinting at numerous experiments to relieve gaps in understanding eukaryotic transcription initiation and control. In particular, these EM studies will be a great supplement to ongoing crystallographic studies of Pol II and GTFs, which to date have not included TFIID (Kostrewa et al., 2009).

There are many limitations to this EM study, most significantly the lack of correlation with the physiological relevance of the conformational rearrangements of factors and DNA observed in the Rap1-TFIIA-TFIID-DNA quaternary complex. There is also a lack of complementary biochemical validation associated with the role of particular factors in protein-protein and protein-DNA interaction. Functional studies also need to be extended to document downstream events, such as affects on further GTF association and transcription initiation.

With respect to roles of particular factors, while I was able to provide some evidence of functional interactions between Taf5 and Rap1 C-terminal domains, this has not been extended to documentation of distinct mechanisms of interaction between the Rap1 DBD and C-terminal domains with specific Tafs. Exactly which Tafs directly interact with the particular Rap1 domains is important to know. Here we are still left at the level of genetic interaction, which does not really allow an assignment of mechanism. Second, the way that TFIIA engages TFIID, with respect to TFIIA-Taf interactions, is unknown. The best candidate is the RBD of Taf4, so this will be interesting to test using competition binding reactions with Taf4/12 heterodimers, Rap1, and TFIIA. Third and related to this last point, a direct interaction between TFIIA and Rap1 was proposed based upon the EM data, but this has not been tested by direct interaction assays using just the recombinant proteins. Fourth, a constitutive DNA binding event by Taf2 was

proposed, and the previous observation that recombinant fly or human Taf2 binds to INR sequences in vitro was included in the arguments for this, but it remains untested if yeast Taf2 possesses DNA binding capability. Furthermore, to date there has been no identification of the domain(s) responsible for DNA binding by any Taf2 protein. It will be important to conduct detailed structure/function analyses of yeast Taf2 to help resolve this issue. In fact it will be important to connect the structure/function analyses of TFIIA, Taf2, Taf4, Taf5, and Rap1 with EM experiments. Ideally we would generate mutant forms of each protein that specifically affect distinct stages of the pathway to a functional PIC, and observe those defective intermediates using EM. Such data could be correlated with in vivo data from ChIP, transcript analyses, and FRET, and functional biochemical data from DNA binding/site-specific crosslinking, protein-protein interaction/site specific crosslinking, NMR, and in vitro transcription. The relevance of the large DNA 'loop' between enhancer and promoter would be confirmed in vivo using a modification of ChIP called the chromosomal conformational conformation, or 3C, assay (Laine et al., 2009; Singh et al., 2009; Tan-Wong et al., 2009). While it may be difficult to get 3C to work in this case because the relevant DNA elements are actually quite close together, it is possible that the use of 3C on RP genes in vivo, along with tools like rapamycin treatment and/or specific mutants affecting function of the relevant protein(s), could tell us a lot about the biology of enhancer-promoter physical interaction from a distance, which is becoming increasingly prominent as a highly regulated aspect of transcriptional control.

Even given the capability to generate optimal Taf2, 4, 5, and 12 mutants precisely affecting distinct components of TFIID function such as DNA binding, Rap1 interaction,

and TFIIA interaction, it would still be quite hard to place these into the relevant context of TFIID function, since it is so difficult to obtain preparative quantities of TFIID for EM and functional experiments. Isolation of fully or partially reconstituted TFIID from recombinant Tafs might be a better plan, but that also will require a lot of effort, and could turn out to be a dead end (Chen et al., 1994; Ishihama and Ito, 1972). Instead, structure/function analyses of Rap1 and TFIIA is much more attractive as a place to start, since those proteins are much easier to deal with in both genetic and preparative biochemical applications. The combination of the strategy first implemented with Taf4, that of systematic alanine scanning mutagenesis through the sequence encoding the domain of interest, could easily be applied to the Rap1 DBD and relevant C-terminal domains. Any work on the C-terminal domains will be difficult to interpret, since there is no structural information on the portions we are interested in. With respect to the DBD, it might be hard to separate DNA binding function from protein-protein interaction function. TFIIA is potentially much easier since there are X-ray structures available and the two subunits are small proteins and thus encoded by small genes, which makes site-directed mutagenesis remarkably easy. For all of these reasons and more I began a site-directed mutagenesis study of the small Toa2 subunit of TFIIA.

In spite of the available X-ray structures of the TFIIA complex, very little site-directed mutagenesis had been done on yeast TFIIA subunit-encoding genes. Three temperature conditional alleles are described in the literature, *toa1-1, toa1-25*, and *toa2-3* (Kang et al., 1995; Xie et al., 2000). The first mutant was not particularly interesting to us, since it was shown to affect interaction with another factor, the NC2 repressor complex. The molecular mechanism underlying the growth phenotypes of the other two

published yeast mutants was not clear, so these were also somewhat problematic for use in our study. Fortunately, a much more systematic analysis of individual residue function had been carried out in the case of the human small TFIIA subunit (Ozer et al., 1996). This study isolated a unique class of TFIIA mutants, those that appeared to be capable of interaction with TBP, functional in 'basal' transcription, but defective in activator-dependent transcription. Therefore we reasoned that two particular mutants, affecting residues analogous to yeast Toa2 Y10 and F71, might represent proteins defective for interaction with activators, and thus potentially defective in activator specific affects such as stimulation of TFIID conformational rearrangement and DNA binding.

When prepared in recombinant form alongside wild type TFIIA and analyzed by EM, these TFIIA variants were defective in the formation of DNA loops. DNA looping occurred at a lower frequency that that observed with wild type TFIIA (Papai et al.). Just as importantly, as shown in **Figure 4.6**, individually these *toa2* mutants displayed temperature conditional growth phenotypes when tested in plasmid shuffle complementation assays, while the double Y10 and F71 mutation resulted in lethality. The *toa2-3* allele could not support viability when tested in our assay, in contrast to the literature (Kang et al., 1995). Notably, steady state protein abundance of each mutant was very similar to wild type, with the exception of the protein from *toa2-3*, which was present at a reduced level (~50-70% of wild type protein abundance). Collectively, these results suggest that very specific residues in TFIIA participate in formation of the characteristic Rap1-TFIIA-TFIID-DNA quaternary structure observed by EM, and also show that these residues are critical to cell growth, consistent with a possible physiological role in RP gene transcription. Like each of the factors that are of particular

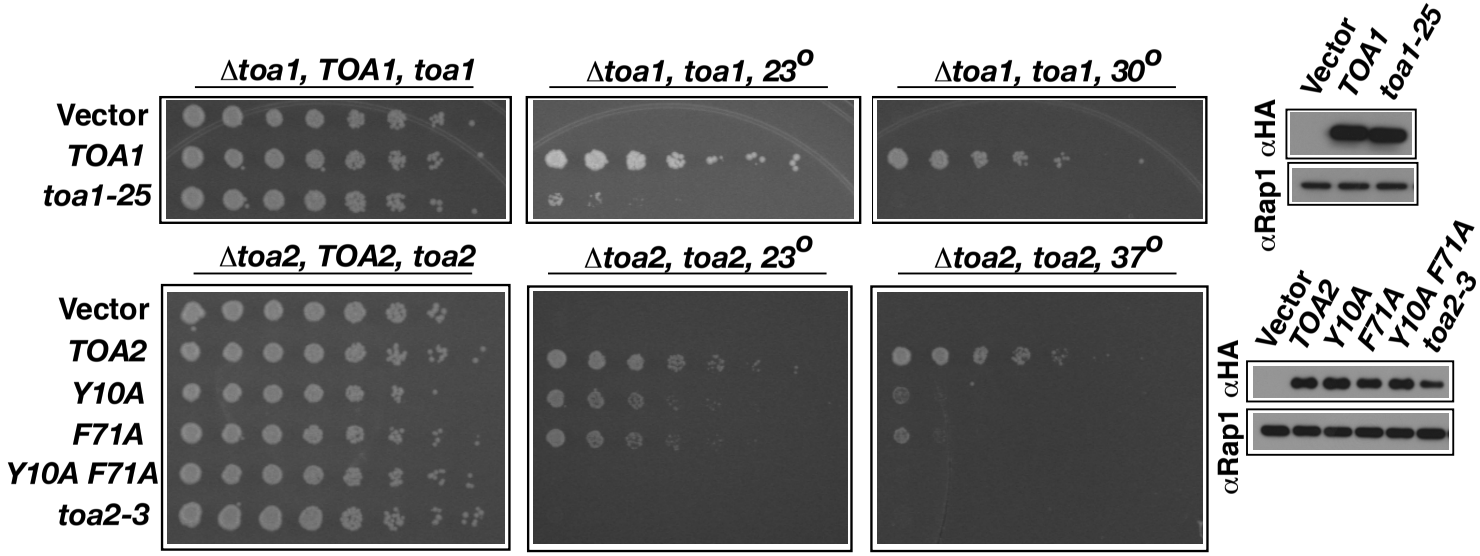


Figure 4.6: Complementation properties of various mutant TFIIA subunit-encoding alleles. TFIIA consists of two subunits that are encoded by single copy, essential genes. Top, standard plasmid shuffle assays at different temperatures were used to test the growth of strains carrying either wild type or the *toa1-25* mutant. *toa1-25* could barely complement even at 23°, and was unable to support viability at 30°, in contrast to the literature. Therefore no useful genetic reagents were in hand to test the contribution of *TOA1* to our pathways of interest. Upper right, companion immunoblot; the protein encoded by *toa1-25* was present at nearly wild type abundance. Bottom panels: complementation properties of *toa2* mutants. The only published *toa2* mutant *toa2-3* was unable to support growth at any condition tested, in contrast to the literature (Kang et. al., 1995). On the other hand, design of two additional mutants was based upon a study of the human small TFIIA subunit (Ozer et. al., 1996), and these single amino acid substitutions conferred slow growth at low temperature and were temperature sensitive. A combination of these amino acid substitutions into the same allele was incompatible with viability. As shown at the lower right, all proteins encoded by *toa2* mutant alleles were present at near wild type levels, except from *toa2-3* which was present at ~50% of wild type. Note that recombinant TFIIA containing the Y10A and F71A Toa2 variants were defective in Rap1-dependent DNA loop formation in EM-based assays (Papai et. al., 2010).

interest to us, TFIIA probably has several functions it must perform in order to allow the conformational rearrangement of the quaternary complex. It must interact with TFIID, through TBP but also probably through Tafs, and with Rap1; TFIIA also may need to undergo a physical rearrangement itself in order to accommodate the change in location stimulated by Rap1. In theory, each of these responsibilities could be dependent upon distinct residues in the TFIIA subunits. To begin to try and separate these functions, we decided to extend our mutagenesis study of Toa2. Together with a very talented and hard working student in Tony's lab, Mr. Joseph R. Cates, I set out to test the sensitivity of every Toa2 residue to alanine substitution by generation of mutants and examination of growth phenotypes. We hoped to identify those residues that make an obvious contribution to cellular growth, and in so doing, prepare genetic reagents that would be useful for functional biochemical and structural experiments. 59 mutants were planned that would change every consecutive pair of residues to alanine, so that every non-alanine residue in Toa2 would be converted. To date, 50 of these Toa2 mutants have been constructed and tested for temperature conditional growth and steady state protein abundance. As expected, a range of growth phenotypes was observed amongst different mutants, along with variable steady state protein abundance. The raw data are shown in **Figure 4.7**. In total, 18 of those 50 site-directed mutants resulted in temperature conditional phenotypes. The residues changed in those mutants were mapped onto the three-dimensional structure of TFIIA obtained by X-ray crystallography (Tan et al., 1996). **Figure 4.8** indicates that motifs distributed around the TFIIA complex are affected by those amino acid substitutions. These motifs are probably involved in several distinct aspects of TFIIA function. Alteration might be expected to disrupt TFIIA

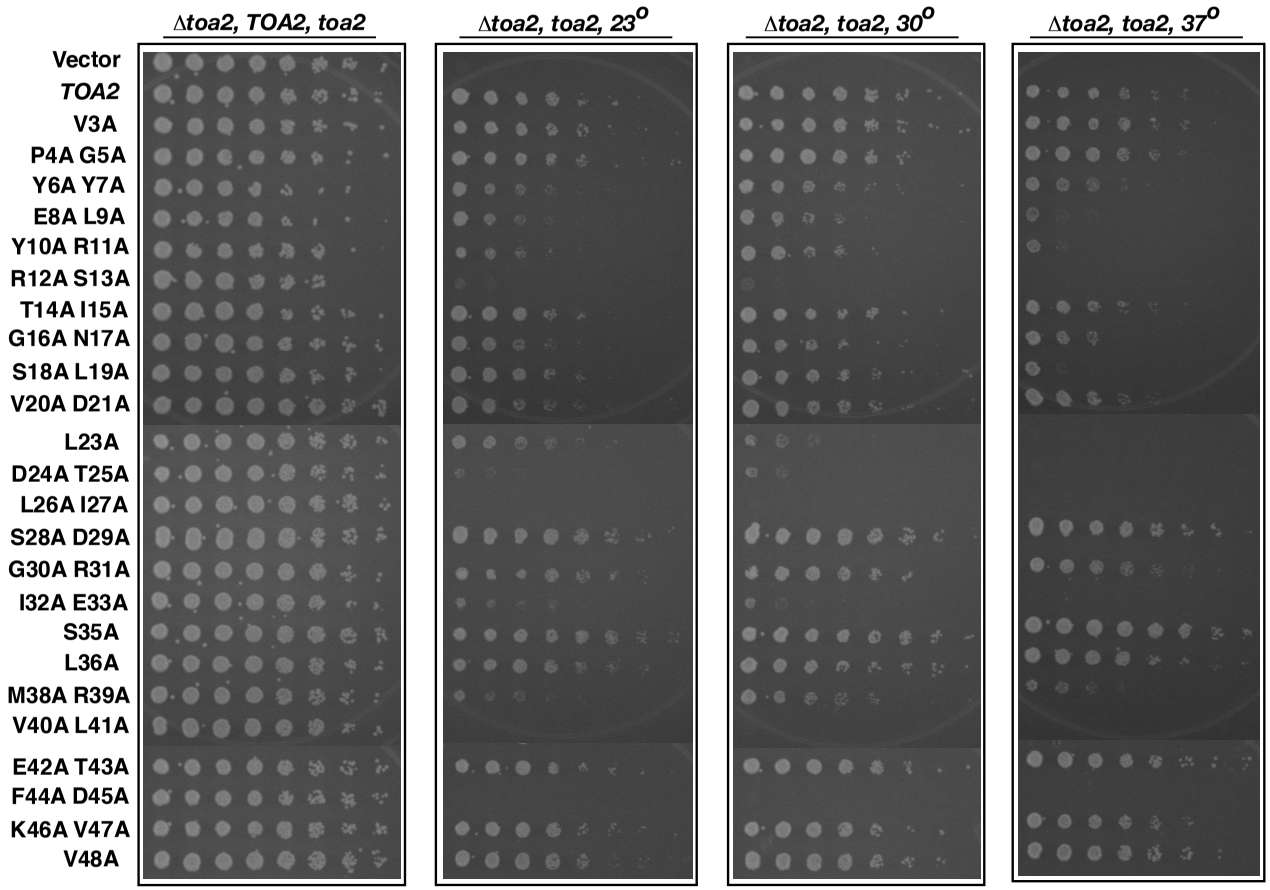
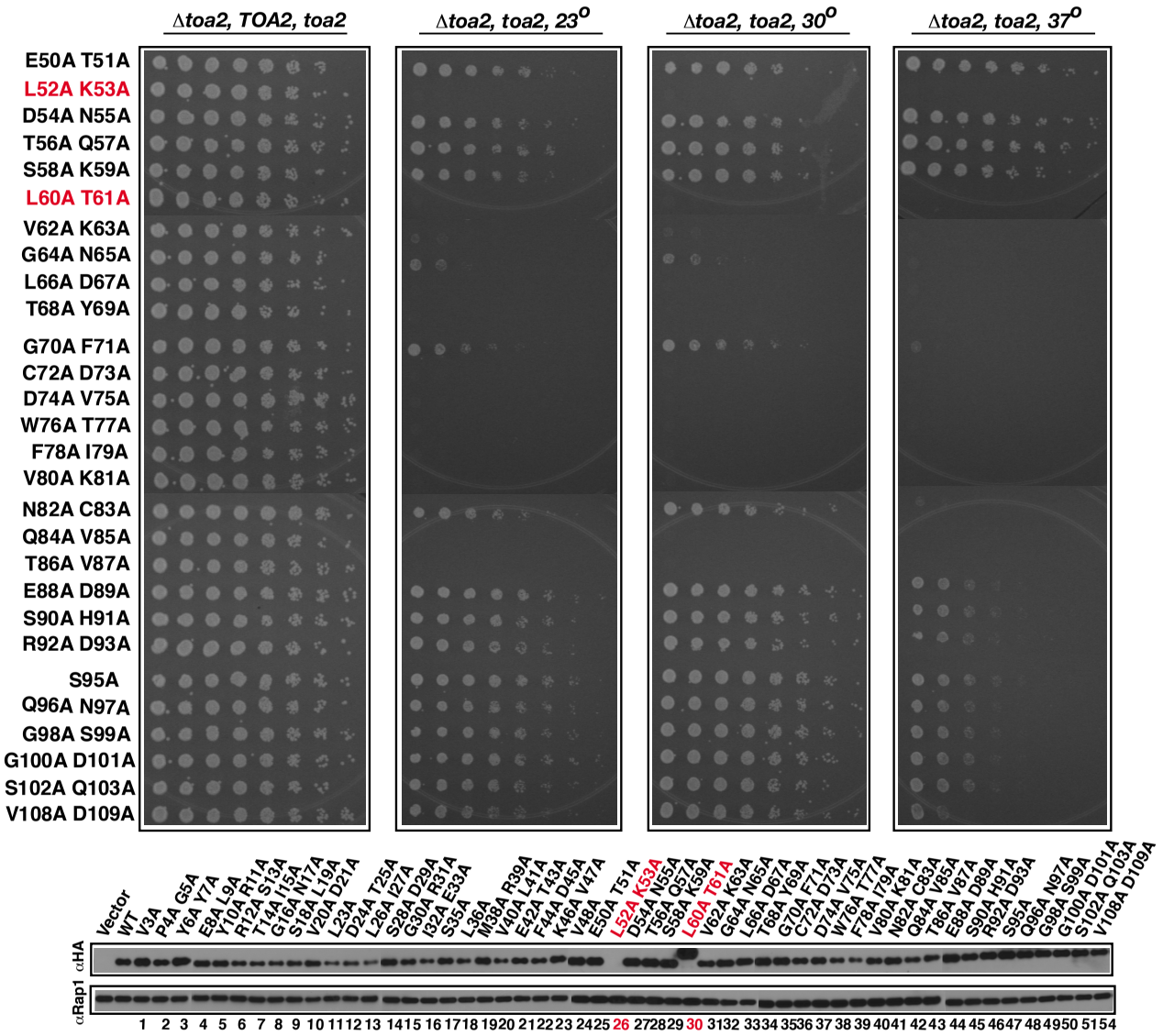


Figure 4.7: Systematic alanine scanning mutagenesis of *TOA2* and plasmid shuffle growth complementation assays. A mutagenesis strategy was designed in which every two consecutive Toa2 residues would be converted to alanine. Of 59 designed alleles 52 were prepared and tested for complementation properties and temperature sensitivity of growth. 50 of those 52 alleles were functional as indicated by complete DNA sequencing and protein expression analyses (see companion immunoblot in continuation of figure, next page; 2 non-functional alleles indicated in red). Numerous loss of growth phenotypes were observed including more than a dozen temperature conditional mutants; those will be especially useful in future genetic experiments testing the contribution to Rap1- and TFIID-dependent RPG transcriptionin vivo. Diverse Toa2 functions are probably affected. Continued on following page.


Δtoa2, TOA2, toa2
Δtoa2, toa2, 23°
Δtoa2, toa2, 30°
Δtoa2, toa2, 37°
E50A T51A
L52A K53A
D54A N55A
T56A Q57A
S58A K59A
L60A T61A
V62A K63A
G64A N65A
L66A D67A
T68A Y69A
G70A F71A
C72A D73A
D74A V75A
W76A T77A
F78A I79A
V80A K81A
N82A C83A
Q84A V85A
T86A V87A
E88A D89A
S90A H91A
R92A D93A
S95A
Q96A N97A
G98A S99A
G100A D101A
S102A Q103A
V108A D109A
Vector
WT
V3A
P4A G5A
Y6A Y7A
E8A L9A
Y10A R11A
R12A S13A
T14A I15A
G16A N17A
S18A L19A
V20A D21A
L23A
D24A T25A
L26A I27A
S28A D29A
G30A R31A
I32A E33A
S35A
L36A
M38A R39A
V40A L41A
E42A T43A
F44A D45A
K46A V47A
V48A
E50A T51A
L52A K53A
D54A N55A
T56A Q57A
S58A K59A
L60A T61A
V62A K63A
G64A N65A
L66A D67A
T68A Y69A
G70A F71A
C72A D73A
D74A V75A
W76A T77A
F78A I79A
V80A K81A
N82A C83A
Q84A V85A
T86A V87A
E88A D89A
S90A H91A
R92A D93A
S95A
Q96A N97A
G98A S99A
G100A D101A
S102A Q103A
V108A D109A
αHA
αRap1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 54

integrity via inter-subunit interactions, disrupt interaction with TBP, or inhibit TFIIA function through other means. Based upon the structural data and temperature conditional phenotypes of certain substitutions, we find it likely that amongst this last class of *toa2* temperature conditional alleles, we have generated mutants defective in several previously unexplored aspects of yeast TFIIA function. It will be interesting and informative to characterize these mutants in much more detail. These new reagents promise to help us greatly in understanding the various TFIIA-dependent aspects of TFIID- and Rap1-dependent RPG transcription.

We can clearly see that the expertise and methodology are now firmly in place to carry out a detailed, multidisciplinary characterization of TFIID and Rap1 function in RPG transcription. By using the detailed mutagenesis approaches documented in this chapter on both Taf4 and TFIIA subunits, we will much more precisely characterize each of the factors playing a direct important role in the Rap1-TFIIA-TFIID-DNA network. In so doing, our lab is now positioned to define the precise mechanism of TFIID function within the next few years.

Overview

At its core my thesis project was designed to define small segments, or domains, of Taf4, Taf5, and Taf12 that are responsible for interaction with Rap1. Subsequent elineation of the physiological relevance of Rap1-Taf interactions depended on demonstrating the in vivo significance of each RBD within the Tafs, and correlating in vitro and in vivo function of Tafs containing altered RBDs. The RBDs of all three Tafs

were initially mapped by Far Western protein blotting assays using systematic truncation variants of each Taf.

The Taf4 RBD was mapped to residues 253-344, a section of Taf4 that is largely uncharacterized, but that was known to be essential for viability in yeast. I demonstrated that the Taf4 RBD does not mediate interaction with other Tafs by using yeast two hybrid assays, an important observation since it rules out a Taf-Taf interaction-mediated explanation for the function of this Taf4 region. I confirmed the essential nature of Taf4 residues 253-344. Point mutant alleles were generated by random mutagenesis and isolated on the basis of a temperature conditional growth phenotype. These *taf4* temperature conditional mutants were defective in RPG transcription, as documented by monitoring steady-state abundance of RPG transcripts in RNA isolated from wild type and mutant cultures that had been subjected to temperature shift. Recombinant proteins corresponding to the *taf4* mutants were prepared in *E. coli* as heterodimers with the Taf4 binding partner Taf12. The amino acid substitutions encoded by the *taf4* mutant alleles resulted in reduction of in vitro binding affinity between Taf4/Taf12 heterodimers and Rap1, as scored by quantitative protein-protein interaction assays. The defect in Rap1 binding in vitro was taken to indicate a connection between the loss of growth and RPG transcription defects observed in vivo. Using a systematic alanine-scanning mutagenesis strategy, the residues within the Taf4 RBD that are essential for growth were mapped to two blocks between residues 251-274 and 311-322. Cooperativity occurs between these two amino acid sequence blocks in support of rapid cellular growth. It is likely that these two regions of Taf4 sequence correspond to structural motifs that bind Rap1 and mediate direct protein-protein interaction. Based on findings of other labs, one or both of these

two blocks may also be important for direct DNA binding and also possibly direct protein-protein interaction with TFIIA; both of these Taf4 interactions may be functionally cooperative with simultaneous or stepwise Taf4 RBD-mediated Rap1 binding.

The Taf5 segment bound by Rap1 contained two N-terminal domains, NTD1 and NTD2, both of which are evolutionarily conserved suggesting their broad importance to eukaryotic biology. While NTD1 was dispensable for growth, removal of NTD2 conferred a loss of growth phenotype that was worsened by simultaneous removal of NTD1. This result indicates collaboration between NTD1 and NTD2, possibly by collective responsibility in Rap1 interaction and RPG transcription. As with Taf4, NTD2-directed point mutant *taf5* alleles were generated and isolated by a genetic screen for temperature conditional growth. These *taf5* mutants displayed profound loss of RPG transcription across the entire regulon as scored by microarray analyses of RNA prepared from heat-shocked wild type and mutant cultures. This result is consistent with Taf5 NTD2-Rap1 binding and the importance of this interaction in driving most cellular RPG transcription. However, a defect in Rap1 association with TFIID could not be detected using co-immunoprecipitation of Rap1 with Tafs from extracts of wild type and mutant strains. This may or may not be a technical artifact originating from one or more experimental conditions. Given that recombinant proteins corresponding to *taf5* mutant alleles are defective for interaction with Rap1 in vitro, an alternative and more promising explanation for the (lack of a) result in co-immunoprecipitation assays involves redundancy in Rap1 binding mediated by the multiple Rap1-binding Tafs present in TFIID. Moreover, there appear to be multiple distinct continuous stretches of amino

acids in Taf4 (251-274 and 311-322), Taf5 (NTD1 and NTD2), and also the RBD within Taf12 (probably just one segment). Similarly, previous work in our laboratory has shown that the Rap1 DBD and one or more C-terminal domains can interact with all or a selection of the available RBDs in Tafs. Consistent with this hypothesis of RBD redundancy within Rap1-TFIID interaction, strains containing amino acid substitutions in multiple RBDs (i.e. both Taf4 and Taf5) exhibit synthetic lethality. Similarly, simultaneous removal of RBD function in Taf5 and deletion of specific Rap1 C-terminal domains confers synthetic growth phenotypes. These results strongly suggest the idea that multiple mechanisms of interaction are operative in supporting cellular growth, and hence probably involved in Rap1-TFIID binding-dependent RPG transcription.

Consistent with the functional importance of multiple Tafs (and domains) in Rap1-TFIID binding, electron microscopy difference mapping experiments revealed Rap1 binding to TFIID in regions of the complex known to contain Taf4, Taf5, and Taf12. While there are two regions of TFIID that contain all three proteins in close proximity, only one such region appeared to be bound by Rap1. The mechanistic basis of this observation remains unclear but could be related to differential PTM amongst the multiple molecules of each Taf within TFIID. The binding of Rap1 to TFIID did not appear to influence the architecture of the complex. However, the position of TFIIA and DNA binding to TFIID was dramatically altered in a Rap1-dependent fashion. This result indicates that conformational alterations in PICs may be stimulated by interactions between Rap1, TFIID, and TFIIA, and that these conformational alterations may be a hallmark of the PICs that stimulate the very high level of transcription initiation seen on the Rap1- and TFIID-bound Ribosomal Protein genes in vivo.

The connection of TFIID, Rap1, and TFIIA was pursued by a systematic alanine-scanning mutagenesis strategy of TFIIA subunit-encoding genes. Analyses of TFIIA structure and function will hopefully identify mutant variants that affect specific individual components of the interactions and conformational rearrangements. Likewise, the systematic high-resolution mutagenesis strategies developed over the course of this thesis project can be directed to any RBD, or additional Tafs or PIC components, or Rap1 itself. It is also likely that interaction and conformational rearrangements depend to some extent on dynamic PTM, and characterization of such PTMs may inform as to the mechanism and consequence of these protein-protein and protein-DNA interactions. In any case, the context of Rap1-TFIID and Rap1-Taf interaction in the stimulation of yeast RPG transcription is now the most thoroughly characterized example of activator-TFIID interaction available. This is a benefit of the systematic back-and-forth application of biochemistry, genetics, cell biology, and structural biology techniques. The philosophy of systematic design and generation of reagents, and experimentation, will move this investigation even further ahead.

The fundamental question that will be answered by the continuation of this work is to understand exactly how an activator can stimulate function of a TFIID-containing PIC, in this case where the activator does seem to merely drive the initial assembly of the PIC onto the promoter. Instead Rap1 appears to drive PIC assembly and/or function at a step subsequent to TFIIA- and TFIID-promoter association. Since the genes being studies are transcribed at a rate that is among the highest known in the literature, our work will illuminate exactly what makes highly transcribed genes unique, and this will be instructive to understanding of both normal cellular function and pathological situations.

REFERENCES

Abate, C., Luk, D., Gagne, E., Roeder, R.G., and Curran, T. (1990). Fos and jun cooperate in transcriptional regulation via heterologous activation domains. Mol Cell Biol *10*, 5532-5535.

Abmayr, S.M., Workman, J.L., and Roeder, R.G. (1988). The pseudorabies immediate early protein stimulates in vitro transcription by facilitating TFIID: promoter interactions. Genes Dev *2*, 542-553.

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

Affolter, M., Percival-Smith, A., Muller, M., Leupin, W., and Gehring, W.J. (1990). DNA binding properties of the purified Antennapedia homeodomain. Proc Natl Acad Sci U S A *87*, 4093-4097.

Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. Cell *103*, 667-678.

Ahmed, S., Brickner, D.G., Light, W.H., Cajigas, I., McDonough, M., Froyshteter, A.B., Volpe, T., and Brickner, J.H. DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. Nat Cell Biol *12*, 111-118.

Ahmed, S., and Brickner, J.H. (2007). Regulation and epigenetic control of transcription at the nuclear periphery. Trends Genet *23*, 396-402.

Albert, A., Lavoie, S., and Vincent, M. (1999). A hyperphosphorylated form of RNA polymerase II is the major interphase antigen of the phosphoprotein antibody MPM-2 and interacts with the peptidyl-prolyl isomerase Pin1. J Cell Sci *112 (Pt 15)*, 2493-2500.

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291-294.

Alberts, B., and Miake-Lye, R. (1992). Unscrambling the puzzle of biological machines: the importance of the details. Cell *68*, 415-420.

Albright, S.R., and Tjian, R. (2000). TAFs revisited: more data reveal new twists and confirm old ideas. Gene *242*, 1-13.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Allfrey, V.G., and Mirsky, A.E. (1964). Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. Science *144*, 559.

Almer, A., Rudolph, H., Hinnen, A., and Horz, W. (1986). Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. EMBO J *5*, 2689-2696.

Andel, F., 3rd, Ladurner, A.G., Inouye, C., Tjian, R., and Nogales, E. (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. Science *286*, 2153-2156.

Anish, R., Hossain, M.B., Jacobson, R.H., and 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.

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes Dev *19*, 2969-2978.

Ansari, S.A., He, Q., and Morse, R.H. (2009). Mediator complex association with constitutively transcribed genes in yeast. Proc Natl Acad Sci U S A *106*, 16734-16739.

Apone, L.M., Virbasius, C.A., Holstege, F.C., Wang, J., Young, R.A., and Green, M.R. (1998). Broad, but not universal, transcriptional requirement for yTAFII17, a histone H3-like TAFII present in TFIID and SAGA. Mol Cell *2*, 653-661.

Apone, L.M., Virbasius, C.M., Reese, J.C., and Green, M.R. (1996). Yeast TAF(II)90 is required for cell-cycle progression through G2/M but not for general transcription activation. Genes Dev *10*, 2368-2380.

Aslanidis, C., and de Jong, P.J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res *18*, 6069-6074.

Aslanidis, C., de Jong, P.J., and Schmitz, G. (1994). Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products. PCR Methods Appl *4*, 172-177.

Asturias, F.J., and Kornberg, R.D. (1999). Protein crystallization on lipid layers and structure determination of the RNA polymerase II transcription initiation complex. J Biol Chem *274*, 6813-6816.

Auld, K.L., Brown, C.R., Casolari, J.M., Komili, S., and Silver, P.A. (2006). Genomic association of the proteasome demonstrates overlapping gene regulatory activity with transcription factor substrates. Mol Cell *21*, 861-871.

Auty, R., Steen, H., Myers, L.C., Persinger, J., Bartholomew, B., Gygi, S.P., and Buratowski, S. (2004). Purification of active TFIID from Saccharomyces cerevisiae. Extensive promoter contacts and co-activator function. J Biol Chem *279*, 49973-49981.

Bai, C., and Elledge, S.J. (1996). Gene identification using the yeast two-hybrid system. Methods Enzymol *273*, 331-347.

Bai, C., and Elledge, S.J. (1997). Gene identification using the yeast two-hybrid system. Methods Enzymol *283*, 141-156.

Barbaric, S., Luckenbach, T., Schmid, A., Blaschke, D., Horz, W., and Korber, P. (2007). Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter in vivo. J Biol Chem *282*, 27610-27621.

Barbaric, S., Walker, J., Schmid, A., Svejstrup, J.Q., and Horz, W. (2001). Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. EMBO J *20*, 4944-4951.

Belinchon, M.M., Flores, C.L., and Gancedo, J.M. (2004). Sampling Saccharomyces cerevisiae cells by rapid filtration improves the yield of mRNAs. FEMS Yeast Res *4*, 751-756.

Berger, A.B., Decourty, L., Badis, G., Nehrbass, U., Jacquier, A., and Gadal, O. (2007). Hmo1 is required for TOR-dependent regulation of ribosomal protein gene transcription. Mol Cell Biol *27*, 8015-8026.

Berger, I., Fitzgerald, D.J., and Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. Nat Biotechnol *22*, 1583-1587.

Bhattacharya, S., Takada, S., and Jacobson, R.H. (2007). Structural analysis and dimerization potential of the human TAF5 subunit of TFIID. Proc Natl Acad Sci U S A *104*, 1189-1194.

Bhaumik, S.R., and Green, M.R. (2001). SAGA is an essential in vivo target of the yeast acidic activator Gal4p. Genes Dev *15*, 1935-1945.

Bhaumik, S.R., and Green, M.R. (2002). Differential requirement of SAGA components for recruitment of TATA-box-binding protein to promoters in vivo. Mol Cell Biol *22*, 7365-7371.

Bhaumik, S.R., and Green, M.R. (2003). Interaction of Gal4p with components of transcription machinery in vivo. Methods Enzymol *370*, 445-454.

Bhaumik, S.R., Raha, T., Aiello, D.P., and Green, M.R. (2004). In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. Genes Dev *18*, 333-343.

Bhoite, L.T., Yu, Y., and Stillman, D.J. (2001). The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. Genes Dev *15*, 2457-2469.

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

Bitterman, K.J., Medvedik, O., and Sinclair, D.A. (2003). Longevity regulation in Saccharomyces cerevisiae: linking metabolism, genome stability, and heterochromatin. Microbiol Mol Biol Rev *67*, 376-399, table of contents.

Bloom, J., Amador, V., Bartolini, F., DeMartino, G., and Pagano, M. (2003). Proteasome-mediated degradation of p21 via N-terminal ubiquitinylation. Cell *115*, 71-82.

Bloom, J., and Pagano, M. (2005). Experimental tests to definitively determine ubiquitylation of a substrate. Methods Enzymol *399*, 249-266.

Boeger, H., Bushnell, D.A., Davis, R., Griesenbeck, J., Lorch, Y., Strattan, J.S., Westover, K.D., and Kornberg, R.D. (2005). Structural basis of eukaryotic gene transcription. FEBS Lett *579*, 899-903.

Borggrefe, T., Davis, R., Bareket-Samish, A., and Kornberg, R.D. (2001). Quantitation of the RNA polymerase II transcription machinery in yeast. J Biol Chem *276*, 47150-47153.

Bornstein, G., Bloom, J., Sitry-Shevah, D., Nakayama, K., Pagano, M., and Hershko, A. (2003). Role of the SCFSkp2 ubiquitin ligase in the degradation of p21Cip1 in S phase. J Biol Chem *278*, 25752-25757.

Bourbon, H.M., Aguilera, A., Ansari, A.Z., Asturias, F.J., Berk, A.J., Bjorklund, S., Blackwell, T.K., Borggrefe, T., Carey, M., Carlson, M.*, et al.* (2004). A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. Mol Cell *14*, 553-557.

Boyer-Guittaut, M., Birsoy, K., Potel, C., Elliott, G., Jaffray, E., Desterro, J.M., Hay, R.T., and Oelgeschlager, T. (2005). SUMO-1 modification of human transcription factor (TF) IID complex subunits: inhibition of TFIID promoter-binding activity through SUMO-1 modification of hsTAF5. J Biol Chem *280*, 9937-9945.

Bram, R.J., Lue, N.F., and Kornberg, R.D. (1986). A GAL family of upstream activating sequences in yeast: roles in both induction and repression of transcription. EMBO J *5*, 603-608.

Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. Science *286*, 2151-2153.

Brickner, D.G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.C., Widom, J., and Brickner, J.H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol *5*, e81.

Brickner, J.H. (2009). Transcriptional memory at the nuclear periphery. Curr Opin Cell Biol *21*, 127-133.

Brickner, J.H., and Walter, P. (2004). Gene recruitment of the activated INO1 locus to the nuclear membrane. PLoS Biol *2*, e342.

Brookes, E., and Pombo, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. EMBO Rep *10*, 1213-1219.

Brown, C.E., Howe, L., Sousa, K., Alley, S.C., Carrozza, M.J., Tan, S., and Workman, J.L. (2001). Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. Science *292*, 2333-2337.

Brownell, J.E., and Allis, C.D. (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. Proc Natl Acad Sci U S A *92*, 6364-6368.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

Browning, D.F., and Busby, S.J. (2004). The regulation of bacterial transcription initiation. Nat Rev Microbiol *2*, 57-65.

Bryant, G.O., and Ptashne, M. (2003). Independent recruitment in vivo by Gal4 of two complexes required for transcription. Mol Cell *11*, 1301-1309.

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

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell *56*, 549-561.

Burgess, R.R. (1969a). A new method for the large scale purification of Escherichia coli deoxyribonucleic acid-dependent ribonucleic acid polymerase. J Biol Chem *244*, 6160-6167.

Burgess, R.R. (1969b). Separation and characterization of the subunits of ribonucleic acid polymerase. J Biol Chem *244*, 6168-6176.

Burgess, R.R. (1991). Use of polyethyleneimine in purification of DNA-binding proteins. Methods Enzymol *208*, 3-10.

Burgess, R.R., Arthur, T.M., and Pietz, B.C. (2000). Mapping protein-protein interaction domains using ordered fragment ladder far-western analysis of hexahistidine-tagged fusion proteins. Methods Enzymol *328*, 141-157.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. (1969). Factor stimulating transcription by RNA polymerase. Nature *221*, 43-46.

Burke, T.W., and 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*, 711-724.

Burke, T.W., and 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*, 3020-3031.

Bushnell, D.A., Cramer, P., and Kornberg, R.D. (2002). Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. Proc Natl Acad Sci U S A *99*, 1218-1222.

Bushnell, D.A., and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. Proc Natl Acad Sci U S A *100*, 6969-6973.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science *303*, 983-988.

Cadwell, R.C., and Joyce, G.F. (1992). Randomization of genes by PCR mutagenesis. PCR Methods Appl *2*, 28-33.

Cairns, B.R., Kim, Y.J., Sayre, M.H., Laurent, B.C., and Kornberg, R.D. (1994). A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. Proc Natl Acad Sci U S A *91*, 1950-1954.

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R.D. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell *87*, 1249-1260.

Cardenas, M.E., Cutler, N.S., Lorenz, M.C., Di Como, C.J., and Heitman, J. (1999). The TOR signaling cascade regulates gene expression in response to nutrients. Genes Dev *13*, 3271-3279.

Carr, D.W., Stofko-Hahn, R.E., Fraser, I.D., Cone, R.D., and Scott, J.D. (1992). Localization of the cAMP-dependent protein kinase to the postsynaptic densities by A-kinase anchoring proteins. Characterization of AKAP 79. J Biol Chem *267*, 16816-16823.

Casolari, J.M., Brown, C.R., Drubin, D.A., Rando, O.J., and Silver, P.A. (2005). Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. Genes Dev *19*, 1188-1198.

Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell *117*, 427-439.

Cavallini, B., Huet, J., Plassat, J.L., Sentenac, A., Egly, J.M., and Chambon, P. (1988). A yeast activity can substitute for the HeLa cell TATA box factor. Nature *334*, 77-80.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J *18*, 4835-4845.

Chambers, A., Tsang, J.S., Stanway, C., Kingsman, A.J., and Kingsman, S.M. (1989). Transcriptional control of the Saccharomyces cerevisiae PGK gene by RAP1. Mol Cell Biol *9*, 5516-5524.

Chasman, D.I., Flaherty, K.M., Sharp, P.A., and Kornberg, R.D. (1993). Crystal structure of yeast TATA-binding protein and model for interaction with DNA. Proc Natl Acad Sci U S A *90*, 8174-8178.

Chen, D., Ma, H., Hong, H., Koh, S.S., Huang, S.M., Schurter, B.T., Aswad, D.W., and Stallcup, M.R. (1999). Regulation of transcription by a protein methyltransferase. Science *284*, 2174-2177.

Chen, J.L., Attardi, L.D., Verrijzer, C.P., Yokomori, K., and Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. Cell *79*, 93-105.

Chen, Y., Yamaguchi, Y., Tsugeno, Y., Yamamoto, J., Yamada, T., Nakamura, M., Hisatake, K., and Handa, H. (2009). DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. Genes Dev *23*, 2765-2777.

Chiang, C.M., and Roeder, R.G. (1995). Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science *267*, 531-536.

Cho, E.J., Takagi, T., Moore, C.R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev *11*, 3319-3326.

Choi, C.H., Hiromura, M., and Usheva, A. (2003). Transcription factor IIB acetylates itself to regulate transcription. Nature *424*, 965-969.

Chung, W.H., Craighead, J.L., Chang, W.H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R.D., and Asturias, F.J. (2003). RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. Mol Cell *12*, 1003-1013.

Colbran, R.J., Carmody, L.C., Bauman, P.A., Wadzinski, B.E., and Bass, M.A. (2003). Analysis of specific interactions of native protein phosphatase 1 isoforms with targeting subunits. Methods Enzymol *366*, 156-175.

Coleman, R.A., and Pugh, B.F. (1995). Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. J Biol Chem *270*, 13850-13859.

Collins, G.A., Lipford, J.R., Deshaies, R.J., and Tansey, W.P. (2009). Gal4 turnover and transcription activation. Nature *461*, E7; discussion E8.

Conaway, J.W., and Conaway, R.C. (1990a). An RNA polymerase II transcription factor shares functional properties with Escherichia coli sigma 70. Science *248*, 1550-1553.

Conaway, R.C., and Conaway, J.W. (1990b). Transcription initiated by RNA polymerase II and purified transcription factors from liver. Transcription factors alpha, beta gamma,

and delta promote formation of intermediates in assembly of the functional preinitiation complex. J Biol Chem *265*, 7559-7563.

Conaway, R.C., Reines, D., Garrett, K.P., Powell, W., and Conaway, J.W. (1996). Purification of RNA polymerase II general transcription factors from rat liver. Methods Enzymol *273*, 194-207.

Corden, J.L., Cadena, D.L., Ahearn, J.M., Jr., and Dahmus, M.E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc Natl Acad Sci U S A *82*, 7934-7938.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791-1792.

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

Cosma, M.P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. Cell *97*, 299-311.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863-1876.

Crowley, T.E., Hoey, T., Liu, J.K., Jan, Y.N., Jan, L.Y., and Tjian, R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila. Nature *361*, 557-561.

D'Alessio, J.A., Wright, K.J., and Tjian, R. (2009). Shifting players and paradigms in cell-specific transcription. Mol Cell *36*, 924-931.

Dabrowski, S., and Kiaer Ahring, B. (2003). Cloning, expression, and purification of the His6-tagged hyper-thermostable dUTPase from Pyrococcus woesei in Escherichia coli: application in PCR. Protein Expr Purif *31*, 72-78.

Dahmus, M.E. (1983). Structural relationship between the large subunits of calf thymus RNA polymerase II. J Biol Chem *258*, 3956-3960.

Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R., 3rd, and Grant, P.A. (2004). Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. J Biol Chem *279*, 1867-1871.

Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 A resolution. Cell *66*, 121-128.

Davierwala, A.P., Haynes, J., Li, Z., Brost, R.L., Robinson, M.D., Yu, L., Mnaimneh, S., Ding, H., Zhu, H., Chen, Y.*, et al.* (2005). The synthetic genetic interaction spectrum of essential genes. Nat Genet *37*, 1147-1152.

De Virgilio, C., and Loewith, R. (2006). The TOR signalling network from yeast to man. Int J Biochem Cell Biol *38*, 1476-1481.

Deato, M.D., Marr, M.T., Sottero, T., Inouye, C., Hu, P., and Tjian, R. (2008). MyoD targets TAF3/TRF3 to activate myogenin transcription. Mol Cell *32*, 96-105.

Deato, M.D., and Tjian, R. (2007). Switching of the core transcription machinery during myogenesis. Genes Dev *21*, 2137-2149.

Deato, M.D., and Tjian, R. (2008). An unexpected role of TAFs and TRFs in skeletal muscle differentiation: switching core promoter complexes. Cold Spring Harb Symp Quant Biol *73*, 217-225.

DeJong, J., and Roeder, R.G. (1993). A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA. Genes Dev *7*, 2220-2234.

Demeny, M.A., Soutoglou, E., Nagy, Z., Scheer, E., Janoshazi, A., Richardot, M., Argentini, M., Kessler, P., and Tora, L. (2007). Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. PLoS One *2*, e316.

DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*, 680-686.

Diffley, J.F., and Stillman, B. (1989). Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. Science *246*, 1034-1038.

Dikstein, R., Ruppert, S., and Tjian, R. (1996). TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. Cell *84*, 781-790.

Donner, A.J., Hoover, J.M., Szostek, S.A., and Espinosa, J.M. (2007). Stimulus-specific transcriptional regulation within the p53 network. Cell Cycle *6*, 2594-2598.

Dove, S.L., and Hochschild, A. (2004). A bacterial two-hybrid system based on transcription activation. Methods Mol Biol *261*, 231-246.

Durrin, L.K., Mann, R.K., and Grunstein, M. (1992). Nucleosome loss activates CUP1 and HIS3 promoters to fully induced levels in the yeast Saccharomyces cerevisiae. Mol Cell Biol *12*, 1621-1629.

Durso, R.J., Fisher, A.K., Albright-Frey, T.J., and Reese, J.C. (2001). Analysis of TAF90 mutants displaying allele-specific and broad defects in transcription. Mol Cell Biol *21*, 7331-7344.

Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell *66*, 563-576.

Dynlacht, B.D., Weinzierl, R.O., Admon, A., and Tjian, R. (1993). The dTAFII80 subunit of Drosophila TFIID contains beta-transducin repeats. Nature *363*, 176-179.

Ebright, R.H. (1991). Identification of amino acid-base pair contacts by genetic methods. Methods Enzymol *208*, 620-640.

Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B., and Livingston, D.M. (1994). Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes Dev *8*, 869-884.

Edwards, A.M., Darst, S.A., Feaver, W.J., Thompson, N.E., Burgess, R.R., and Kornberg, R.D. (1990). Purification and lipid-layer crystallization of yeast RNA polymerase II. Proc Natl Acad Sci U S A *87*, 2122-2126.

Elmlund, H., Baraznenok, V., Linder, T., Szilagyi, Z., Rofougaran, R., Hofer, A., Hebert, H., Lindahl, M., and Gustafsson, C.M. (2009). Cryo-EM reveals promoter DNA binding and conformational flexibility of the general transcription factor TFIID. Structure *17*, 1442-1452.

Emami, K.H., Jain, A., and Smale, S.T. (1997). Mechanism of synergy between TATA and initiator: synergistic binding of TFIID following a putative TFIIA-induced isomerization. Genes Dev *11*, 3007-3019.

Engelke, D.R., Krikos, A., Bruck, M.E., and Ginsburg, D. (1990). Purification of Thermus aquaticus DNA polymerase expressed in Escherichia coli. Anal Biochem *191*, 396-400.

Espinosa, J.M., Verdun, R.E., and Emerson, B.M. (2003). p53 functions through stress- and promoter-specific recruitment of transcription initiation components before and after DNA damage. Mol Cell *12*, 1015-1027.

Feaver, W.J., Gileadi, O., Li, Y., and Kornberg, R.D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. Cell *67*, 1223-1230.

Feeser, E.A., and Wolberger, C. (2008). Structural and functional studies of the Rap1 C-terminus reveal novel separation-of-function mutants. J Mol Biol *380*, 520-531.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., and Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. Nat Methods *3*, 1021-1032.

Flores, O., Maldonado, E., Burton, Z., Greenblatt, J., and Reinberg, D. (1988). Factors involved in specific transcription by mammalian RNA polymerase II. RNA polymerase II-associating protein 30 is an essential component of transcription factor IIF. J Biol Chem *263*, 10812-10816.

Fondell, J.D., Brunel, F., Hisatake, K., and Roeder, R.G. (1996a). Unliganded thyroid hormone receptor alpha can target TATA-binding protein for transcriptional repression. Mol Cell Biol *16*, 281-287.

Fondell, J.D., Ge, H., and Roeder, R.G. (1996b). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. Proc Natl Acad Sci U S A *93*, 8329-8333.

Fondell, J.D., Roy, A.L., and Roeder, R.G. (1993). Unliganded thyroid hormone receptor inhibits formation of a functional preinitiation complex: implications for active repression. Genes Dev *7*, 1400-1410.

Ford, E., and Thanos, D. The transcriptional code of human IFN-beta gene expression. Biochim Biophys Acta *1799*, 328-336.

Freeman, K., Gwadz, M., and Shore, D. (1995). Molecular and genetic analysis of the toxic effect of RAP1 overexpression in yeast. Genetics *141*, 1253-1262.

Fribourg, S., Romier, C., Werten, S., Gangloff, Y.G., Poterszman, A., and Moras, D. (2001). Dissecting the interaction network of multiprotein complexes by pairwise coexpression of subunits in E. coli. J Mol Biol *306*, 363-373.

Fu, J., Gnatt, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R., and Kornberg, R.D. (1999). Yeast RNA polymerase II at 5 A resolution. Cell *98*, 799-810.

Funk, M., Niedenthal, R., Mumberg, D., Brinkmann, K., Ronicke, V., and Henkel, T. (2002). Vector systems for heterologous expression of proteins in Saccharomyces cerevisiae. Methods Enzymol *350*, 248-257.

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

Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R.G., and Chua, N.H. (1990). Arabidopsis thaliana contains two genes for TFIID. Nature *346*, 390-394.

Gasch, A.P. (2002). Yeast genomic expression studies using DNA microarrays. Methods Enzymol *350*, 393-414.

Gaudreau, L., Schmid, A., Blaschke, D., Ptashne, M., and Horz, W. (1997). RNA polymerase II holoenzyme recruitment is sufficient to remodel chromatin at the yeast PHO5 promoter. Cell *89*, 55-62.

Gazit, K., Moshonov, S., Elfakess, R., Sharon, M., Mengus, G., Davidson, I., and Dikstein, R. (2009). TAF4/4b x TAF12 displays a unique mode of DNA binding and is required for core promoter function of a subset of genes. J Biol Chem *284*, 26286-26296.

Ge, K., Cho, Y.W., Guo, H., Hong, T.B., Guermah, M., Ito, M., Yu, H., Kalkum, M., and Roeder, R.G. (2008). Alternative mechanisms by which mediator subunit MED1/TRAP220 regulates peroxisome proliferator-activated receptor gamma-stimulated adipogenesis and target gene expression. Mol Cell Biol *28*, 1081-1091.

Geiger, J.H., Hahn, S., Lee, S., and Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. Science *272*, 830-836.

Georgieva, S., Nabirochkina, E., Dilworth, F.J., Eickhoff, H., Becker, P., Tora, L., Georgiev, P., and Soldatov, A. (2001). The novel transcription factor e(y)2 interacts with TAF(II)40 and potentiates transcription activation on chromatin templates. Mol Cell Biol *21*, 5223-5231.

Ghosh, P.K., Reddy, V.B., Swinscoe, J., Lebowitz, P., and Weissman, S.M. (1978). Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. J Mol Biol *126*, 813-846.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc *2*, 31-34.

Gileadi, O., Feaver, W.J., and Kornberg, R.D. (1992). Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. Science *257*, 1389-1392.

Gill, G., Pascal, E., Tseng, Z.H., and Tjian, R. (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc Natl Acad Sci U S A *91*, 192-196.

Gill, G., and Tjian, R. (1992). Eukaryotic coactivators associated with the TATA box binding protein. Curr Opin Genet Dev *2*, 236-242.

Gnatt, A., Fu, J., and Kornberg, R.D. (1997). Formation and crystallization of yeast RNA polymerase II elongation complexes. J Biol Chem *272*, 30799-30805.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876-1882.

Goldstein, A.L., and McCusker, J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast *15*, 1541-1553.

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A., and Tjian, R. (1993). Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell *75*, 519-530.

Gorenstein, C., and Warner, J.R. (1976). Coordinate regulation of the synthesis of eukaryotic ribosomal proteins. Proc Natl Acad Sci U S A *73*, 1547-1551.

Govind, C.K., Yoon, S., Qiu, H., Govind, S., and Hinnebusch, A.G. (2005). Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo. Mol Cell Biol *25*, 5626-5638.

Grant, P.A., Schieltz, D., Pray-Grant, M.G., Steger, D.J., Reese, J.C., Yates, J.R., 3rd, and Workman, J.L. (1998). A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. Cell *94*, 45-53.

Gregory, P.D., Schmid, A., Zavari, M., Lui, L., Berger, S.L., and Horz, W. (1998). Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast. Mol Cell *1*, 495-505.

Griffith, J., Hochschild, A., and Ptashne, M. (1986). DNA loops induced by cooperative binding of lambda repressor. Nature *322*, 750-752.

Grob, P., Cruse, M.J., Inouye, C., Peris, M., Penczek, P.A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. Structure *14*, 511-520.

Grontved, L., Madsen, M.S., Boergesen, M., Roeder, R.G., and Mandrup, S. MED14 tethers mediator to the N-terminal domain of peroxisome proliferator-activated receptor gamma and is required for full transcriptional activity and adipogenesis. Mol Cell Biol *30*, 2155-2169.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595-606.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Guermah, M., Ge, K., Chiang, C.M., and Roeder, R.G. (2003). The TBN protein, which is essential for early embryonic mouse development, is an inducible TAFII implicated in adipogenesis. Mol Cell *12*, 991-1001.

Guermah, M., Tao, Y., and Roeder, R.G. (2001). Positive and negative TAF(II) functions that suggest a dynamic TFIID structure and elicit synergy with traps in activator-induced transcription. Mol Cell Biol *21*, 6882-6894.

Gumbs, O.H., Campbell, A.M., and Weil, P.A. (2003). High-affinity DNA binding by a Mot1p-TBP complex: implications for TAF-independent transcription. EMBO J *22*, 3131-3141.

Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989a). Identification of a yeast protein homologous in function to the mammalian general transcription factor, TFIIA. EMBO J *8*, 3379-3382.

Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989b). Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 suppressor of Ty element insertions. Cell *58*, 1173-1181.

Hall, D.B., Wade, J.T., and 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*, 3672-3679.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137-1145.

Han, M., Kim, U.J., Kayne, P., and Grunstein, M. (1988). Depletion of histone H4 and nucleosomes activates the PHO5 gene in Saccharomyces cerevisiae. EMBO J *7*, 2221-2228.

Hansen, S.K., Takada, S., Jacobson, R.H., Lis, J.T., and Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. Cell *91*, 71-83.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J.*, et al.* (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99-104.

Hardy, C.F., Balderes, D., and Shore, D. (1992a). Dissection of a carboxy-terminal region of the yeast regulatory protein RAP1 with effects on both transcriptional activation and silencing. Mol Cell Biol *12*, 1209-1217.

Hardy, C.F., Sussel, L., and Shore, D. (1992b). A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. Genes Dev *6*, 801-814.

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

Hassan, A.H., Neely, K.E., Vignali, M., Reese, J.C., and Workman, J.L. (2001a). Promoter targeting of chromatin-modifying complexes. Front Biosci *6*, D1054-1064.

Hassan, A.H., Neely, K.E., and Workman, J.L. (2001b). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. Cell *104*, 817-827.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell *111*, 369-379.

Haun, R.S., Serventi, I.M., and Moss, J. (1992). Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors. Biotechniques *13*, 515-518.

Hawley, D.K., and Roeder, R.G. (1985). Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. J Biol Chem *260*, 8163-8172.

Hawley, D.K., and Roeder, R.G. (1987). Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. J Biol Chem *262*, 3452-3461.

Hecht, A., and Grunstein, M. (1999). Mapping DNA interaction sites of chromosomal proteins using immunoprecipitation and polymerase chain reaction. Methods Enzymol *304*, 399-414.

Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M., and Grunstein, M. (1995). Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. Cell *80*, 583-592.

Hecht, A., Strahl-Bolsinger, S., and Grunstein, M. (1996). Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. Nature *383*, 92-96.

Henry, N.L., Sayre, M.H., and Kornberg, R.D. (1992). Purification and characterization of yeast RNA polymerase II general initiation factor g. J Biol Chem *267*, 23388-23392.

Hereford, L.M., and Rosbash, M. (1977). Regulation of a set of abundant mRNA sequences. Cell *10*, 463-467.

Higgins, R.C., and Dahmus, M.E. (1979). Rapid visualization of protein bands in preparative SDS-polyacrylamide gels. Anal Biochem *93*, 257-260.

Hinnebusch, A.G., and Natarajan, K. (2002). Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. Eukaryot Cell *1*, 22-32.

Hisatake, K., Hasegawa, S., Takada, R., Nakatani, Y., Horikoshi, M., and Roeder, R.G. (1993). The p250 subunit of native TATA box-binding factor TFIID is the cell-cycle regulatory protein CCG1. Nature *362*, 179-181.

Hisatake, K., Ohta, T., Takada, R., Guermah, M., Horikoshi, M., Nakatani, Y., and Roeder, R.G. (1995). Evolutionary conservation of human TATA-binding-polypeptide-associated factors TAFII31 and TAFII80 and interactions of TAFII80 with other TAFs and with general transcription factors. Proc Natl Acad Sci U S A *92*, 8195-8199.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Hochschild, A., and Ptashne, M. (1986). Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. Cell *44*, 681-687.

Hoey, T., Dynlacht, B.D., Peterson, M.G., Pugh, B.F., and Tjian, R. (1990). Isolation and characterization of the Drosophila gene encoding the TATA box binding protein, TFIID. Cell *61*, 1179-1186.

Hoey, T., Weinzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D., and Tjian, R. (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell *72*, 247-260.

Hoffmann, A., Chiang, C.M., Oelgeschlager, T., Xie, X., Burley, S.K., Nakatani, Y., and Roeder, R.G. (1996). A histone octamer-like structure within TFIID. Nature *380*, 356-359.

Hoffmann, A., Horikoshi, M., Wang, C.K., Schroeder, S., Weil, P.A., and Roeder, R.G. (1990). Cloning of the Schizosaccharomyces pombe TFIID gene reveals a strong conservation of functional domains present in Saccharomyces cerevisiae TFIID. Genes Dev *4*, 1141-1148.

Hoffmann, A., and Roeder, R.G. (1996). Cloning and characterization of human TAF20/15. Multiple interactions suggest a central role in TFIID complex formation. J Biol Chem *271*, 18194-18202.

Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell *95*, 717-728.

Honda, R., Tanaka, H., and Yasuda, H. (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. FEBS Lett *420*, 25-27.

Honda, R., and Yasuda, H. (1999). Association of p19(ARF) with Mdm2 inhibits ubiquitin ligase activity of Mdm2 for tumor suppressor p53. EMBO J *18*, 22-27.

Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T., and Roeder, R.G. (1992). Transcription factor TFIID induces DNA bending upon binding to the TATA element. Proc Natl Acad Sci U S A *89*, 1060-1064.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A., and Roeder, R.G. (1989a). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. Nature *341*, 299-303.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A., and Roeder, R.G. (1989b). Purification of a yeast TATA box-binding protein that exhibits human transcription factor IID activity. Proc Natl Acad Sci U S A *86*, 4843-4847.

Horn, P.J., and Peterson, C.L. (2001). The bromodomain: a regulator of ATP-dependent chromatin remodeling? Front Biosci *6*, D1019-1023.

Hu, P., Geles, K.G., Paik, J.H., DePinho, R.A., and Tjian, R. (2008). Codependent activators direct myoblast-specific MyoD transcription. Dev Cell *15*, 534-546.

Hua, X., Yokoyama, C., Wu, J., Briggs, M.R., Brown, M.S., Goldstein, J.L., and Wang, X. (1993). SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. Proc Natl Acad Sci U S A *90*, 11603-11607.

Huang, J., Dorsey, J., Chuikov, S., Zhang, X., Jenuwein, T., Reinberg, D., and Berger, S.L. G9a and Glp methylate lysine 373 in the tumor suppressor p53. J Biol Chem *285*, 9636-9641.

Huang, J., Perez-Burgos, L., Placek, B.J., Sengupta, R., Richter, M., Dorsey, J.A., Kubicek, S., Opravil, S., Jenuwein, T., and Berger, S.L. (2006). Repression of p53 activity by Smyd2-mediated methylation. Nature *444*, 629-632.

Huang, J., Sengupta, R., Espejo, A.B., Lee, M.G., Dorsey, J.A., Richter, M., Opravil, S., Shiekhattar, R., Bedford, M.T., Jenuwein, T.*, et al.* (2007). p53 is regulated by the lysine demethylase LSD1. Nature *449*, 105-108.

Huber, A., Bodenmiller, B., Uotila, A., Stahl, M., Wanka, S., Gerrits, B., Aebersold, R., and Loewith, R. (2009). Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. Genes Dev *23*, 1929-1943.

Huisinga, K.L., and 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*, 573-585.

Humphrey, E.L., Shamji, A.F., Bernstein, B.E., and Schreiber, S.L. (2004). Rpd3p relocation mediates a transcriptional response to rapamycin in yeast. Chem Biol *11*, 295-299.

Imhof, A., Yang, X.J., Ogryzko, V.V., Nakatani, Y., Wolffe, A.P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. Curr Biol *7*, 689-692.

Indra, A.K., Mohan, W.S., 2nd, Frontini, M., Scheer, E., Messaddeq, N., Metzger, D., and Tora, L. (2005). TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. Dev Biol *285*, 28-37.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene *96*, 23-28.

Ishihama, A., and Ito, K. (1972). Subunits of RNA polymerase in function and structure. II. Reconstitution of Escherichia coli RNA polymerase from isolated subunits. J Mol Biol *72*, 111-123.

Jacobson, R.H., Ladurner, A.G., King, D.S., and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. Science *288*, 1422-1425.

James, P., Halladay, J., and Craig, E.A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics *144*, 1425-1436.

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

Ju, Q.D., Morrow, B.E., and Warner, J.R. (1990). REB1, a yeast DNA-binding protein with many targets, is essential for growth and bears some resemblance to the oncogene myb. Mol Cell Biol *10*, 5226-5234.

Juven-Gershon, T., Cheng, S., and Kadonaga, J.T. (2006). Rational design of a super core promoter that enhances gene expression. Nat Methods *3*, 917-922.

Juven-Gershon, T., Hsu, J.Y., Theisen, J.W., and Kadonaga, J.T. (2008). The RNA polymerase II core promoter - the gateway to transcription. Curr Opin Cell Biol *20*, 253-259.

Kadonaga, J.T. (2002). The DPE, a core promoter element for transcription by RNA polymerase II. Exp Mol Med *34*, 259-264.

Kadonaga, J.T., Courey, A.J., Ladika, J., and Tjian, R. (1988). Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science *242*, 1566-1570.

Kadonaga, J.T., and Tjian, R. (1986). Affinity purification of sequence-specific DNA binding proteins. Proc Natl Acad Sci U S A *83*, 5889-5893.

Kaffman, A., Herskowitz, I., Tjian, R., and O'Shea, E.K. (1994). Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. Science *263*, 1153-1156.

Kaiser, P., and Wohlschlegel, J. (2005). Identification of ubiquitination sites and determination of ubiquitin-chain architectures by mass spectrometry. Methods Enzymol *399*, 266-277.

Kang, J.J., Auble, D.T., Ranish, J.A., and Hahn, S. (1995). Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. Mol Cell Biol *15*, 1234-1243.

Karlic, R., Chung, H.R., Lasserre, J., Vlahovicek, K., and Vingron, M. Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A *107*, 2926-2931.

Kashanchi, F., Piras, G., Radonovich, M.F., Duvall, J.F., Fattaey, A., Chiang, C.M., Roeder, R.G., and Brady, J.N. (1994). Direct interaction of human TFIID with the HIV-1 transactivator tat. Nature *367*, 295-299.

Kaufmann, J., Ahrens, K., Koop, R., Smale, S.T., and Muller, R. (1998). CIF150, a human cofactor for transcription factor IID-dependent initiator function. Mol Cell Biol *18*, 233-239.

Kaufmann, J., and Smale, S.T. (1994). Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev *8*, 821-829.

Kaufmann, J., Verrijzer, C.P., Shao, J., and Smale, S.T. (1996). CIF, an essential cofactor for TFIID-dependent initiator function. Genes Dev *10*, 873-886.

Keegan, L., Gill, G., and Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science *231*, 699-704.

Kelleher, R.J., 3rd, Flanagan, P.M., and Kornberg, R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell *61*, 1209-1215.

Kim, D.K., Inukai, N., Yamada, T., Furuya, A., Sato, H., Yamaguchi, Y., Wada, T., and Handa, H. (2003). Structure-function analysis of human Spt4: evidence that hSpt4 and hSpt5 exert their roles in transcriptional elongation as parts of the DSIF complex. Genes Cells *8*, 371-378.

Kim, D.K., Yamaguchi, Y., Wada, T., and Handa, H. (2001). The regulation of elongation by eukaryotic RNA polymerase II: a recent view. Mol Cells *11*, 267-274.

Kim, M.H., Cooper, D.R., Oleksy, A., Devedjiev, Y., Derewenda, U., Reiner, O., Otlewski, J., and Derewenda, Z.S. (2004). The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications. Structure *12*, 987-998.

Kim, U.J., Han, M., Kayne, P., and Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. EMBO J *7*, 2211-2219.

Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993). Crystal structure of a yeast TBP/TATA-box complex. Nature *365*, 512-520.

Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599-608.

Kimura, H., Tao, Y., Roeder, R.G., and Cook, P.R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol Cell Biol *19*, 5383-5392.

Kininis, M., Isaacs, G.D., Core, L.J., Hah, N., and Kraus, W.L. (2009). Postrecruitment regulation of RNA polymerase II directs rapid signaling responses at the promoters of estrogen target genes. Mol Cell Biol *29*, 1123-1133.

Klein, C., and Struhl, K. (1994). Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. Mol Cell Biol *14*, 1920-1928.

Klekamp, M.S., and Weil, P.A. (1982). Specific transcription of homologous class III genes in yeast-soluble cell-free extracts. J Biol Chem *257*, 8432-8441.

Klemm, R.D., Goodrich, J.A., Zhou, S., and Tjian, R. (1995). Molecular cloning and expression of the 32-kDa subunit of human TFIID reveals interactions with VP16 and TFIIB that mediate transcriptional activation. Proc Natl Acad Sci U S A *92*, 5788-5792.

Kokubo, T., Gong, D.W., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993a). The Drosophila 110-kDa transcription factor TFIID subunit directly interacts with the N-terminal region of the 230-kDa subunit. Proc Natl Acad Sci U S A *90*, 5896-5900.

Kokubo, T., Gong, D.W., Wootton, J.C., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1994a). Molecular cloning of Drosophila TFIID subunits. Nature *367*, 484-487.

Kokubo, T., Gong, D.W., Yamashita, S., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1993b). Drosophila 230-kD TFIID subunit, a functional homolog of the human cell cycle gene product, negatively regulates DNA binding of the TATA box-binding subunit of TFIID. Genes Dev *7*, 1033-1046.

Kokubo, T., Gong, D.W., Yamashita, S., Takada, R., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993c). Molecular cloning, expression, and characterization of the Drosophila 85-kilodalton TFIID subunit. Mol Cell Biol *13*, 7859-7863.

Kokubo, T., Takada, R., Yamashita, S., Gong, D.W., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993d). Identification of TFIID components required for transcriptional activation by upstream stimulatory factor. J Biol Chem *268*, 17554-17558.

Kokubo, T., Yamashita, S., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1994b). Interaction between the N-terminal domain of the 230-kDa subunit and the TATA box-binding subunit of TFIID negatively regulates TATA-box binding. Proc Natl Acad Sci U S A *91*, 3520-3524.

Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. Nature *368*, 466-469.

Kolodziej, P.A., and Young, R.A. (1991). Epitope tagging and protein surveillance. Methods Enzymol *194*, 508-519.

Komarnitsky, P.B., Michel, B., and Buratowski, S. (1999). TFIID-specific yeast TAF40 is essential for the majority of RNA polymerase II-mediated transcription in vivo. Genes Dev *13*, 2484-2489.

Konig, P., Giraldo, R., Chapman, L., and Rhodes, D. (1996). The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. Cell *85*, 125-136.

Kops, O., Zhou, X.Z., and Lu, K.P. (2002). Pin1 modulates the dephosphorylation of the RNA polymerase II C-terminal domain by yeast Fcp1. FEBS Lett *513*, 305-311.

Kopytova, D.V., and Krasnov, A.N. (2007). [The family of TRF (TBP-like factors) proteins]. Genetika *43*, 317-322.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature *462*, 323-330.

Kravchenko, J.E., Rogozin, I.B., Koonin, E.V., and Chumakov, P.M. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. Nature *436*, 735-739.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V.*, et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell *12*, 1565-1576.

Kruse, J.P., and Gu, W. (2008). SnapShot: p53 posttranslational modifications. Cell *133*, 930-930 e931.

Kundu, S., Horn, P.J., and Peterson, C.L. (2007). SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev *21*, 997-1004.

Kundu, S., and Peterson, C.L. Dominant role for signal transduction in the transcriptional memory of yeast GAL genes. Mol Cell Biol *30*, 2330-2340.

Kundu, S., and Peterson, C.L. (2009). Role of chromatin states in transcriptional memory. Biochim Biophys Acta *1790*, 445-455.

Kundu, T.K., Palhan, V.B., Wang, Z., An, W., Cole, P.A., and Roeder, R.G. (2000). Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300. Mol Cell *6*, 551-561.

Kuras, L., Borggrefe, T., and Kornberg, R.D. (2003). Association of the Mediator complex with enhancers of active genes. Proc Natl Acad Sci U S A *100*, 13887-13891.

Kuras, L., Kosa, P., Mencia, M., and Struhl, K. (2000). TAF-Containing and TAF-independent forms of transcriptionally active TBP in vivo. Science *288*, 1244-1248.

Kuras, L., and Struhl, K. (1999). Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. Nature *399*, 609-613.

Kurtz, S., and Shore, D. (1991). RAP1 protein activates and silences transcription of mating-type genes in yeast. Genes Dev *5*, 616-628.

Kushnirov, V.V. (2000). Rapid and reliable protein extraction from yeast. Yeast *16*, 857-860.

Kwok, R.P., Lundblad, J.R., Chrivia, J.C., Richards, J.P., Bachinger, H.P., Brennan, R.G., Roberts, S.G., Green, M.R., and Goodman, R.H. (1994). Nuclear protein CBP is a coactivator for the transcription factor CREB. Nature *370*, 223-226.

Lai, J.S., and Herr, W. (1992). Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations. Proc Natl Acad Sci U S A *89*, 6958-6962.

Laine, J.P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. Genes Dev *23*, 2604-2609.

Lascaris, R.F., Mager, W.H., and Planta, R.J. (1999). DNA-binding requirements of the yeast protein Rap1p as selected in silico from ribosomal protein gene promoter sequences. Bioinformatics *15*, 267-277.

Lawit, S.J., O'Grady, K., Gurley, W.B., and Czarnecka-Verner, E. (2007). Yeast two-hybrid map of Arabidopsis TFIID. Plant Mol Biol *64*, 73-87.

Layer, J.H., Miller, S.G., and Weil, P.A. Direct transactivator-transcription factor IID (TFIID) contacts drive yeast ribosomal protein gene transcription. J Biol Chem *285*, 15489-15499.

Lee, D.K., Horikoshi, M., and Roeder, R.G. (1991). Interaction of TFIID in the minor groove of the TATA element. Cell *67*, 1241-1250.

Lee, T.I., Causton, H.C., Holstege, F.C., Shen, W.C., Hannett, N., Jennings, E.G., Winston, F., Green, M.R., and Young, R.A. (2000). Redundant roles for the TFIID and SAGA complexes in global transcription. Nature *405*, 701-704.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I.*, et al.* (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. Science *298*, 799-804.

Lempiainen, H., and Shore, D. (2009). Growth control and ribosome biogenesis. Curr Opin Cell Biol *21*, 855-863.

Lempiainen, H., Uotila, A., Urban, J., Dohnal, I., Ammerer, G., Loewith, R., and Shore, D. (2009). Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. Mol Cell *33*, 704-716.

Leung, A., Geng, F., Daulny, A., Collins, G., Guzzardo, P., and Tansey, W.P. (2008). Transcriptional control and the ubiquitin-proteasome system. Ernst Schering Found Symp Proc, 75-97.

Leurent, C., Sanders, S., Ruhlmann, C., Mallouh, V., Weil, P.A., Kirschner, D.B., Tora, L., and Schultz, P. (2002). Mapping histone fold TAFs within yeast TFIID. EMBO J *21*, 3424-3433.

Leurent, C., Sanders, S.L., Demeny, M.A., Garbett, K.A., Ruhlmann, C., Weil, P.A., Tora, L., and Schultz, P. (2004). Mapping key functional sites within yeast TFIID. EMBO J *23*, 719-727.

Li, A.G., Piluso, L.G., Cai, X., Gadd, B.J., Ladurner, A.G., and Liu, X. (2007). An acetylation switch in p53 mediates holo-TFIID recruitment. Mol Cell *28*, 408-421.

Li, X.Y., Bhaumik, S.R., and Green, M.R. (2000). Distinct classes of yeast promoters revealed by differential TAF recruitment. Science *288*, 1242-1244.

Li, X.Y., Bhaumik, S.R., Zhu, X., Li, L., Shen, W.C., Dixit, B.L., and Green, M.R. (2002). Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. Curr Biol *12*, 1240-1244.

Lichtsteiner, S., and Tjian, R. (1993). Cloning and properties of the Caenorhabditis elegans TATA-box-binding protein. Proc Natl Acad Sci U S A *90*, 9673-9677.

Lieb, J.D., Liu, X., Botstein, D., and Brown, P.O. (2001). Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. Nat Genet *28*, 327-334.

Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., and Kadonaga, J.T. (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev *18*, 1606-1617.

Lin, P.S., and Dahmus, M.E. (2003). Dephosphorylation of the carboxyl-terminal domain of RNA polymerase II. Methods Enzymol *370*, 155-165.

Lin, Y.Y., Lu, J.Y., Zhang, J., Walter, W., Dang, W., Wan, J., Tao, S.C., Qian, J., Zhao, Y., Boeke, J.D.*, et al.* (2009). Protein acetylation microarray reveals that NuA4 controls key metabolic target regulating gluconeogenesis. Cell *136*, 1073-1084.

Liu, Q., Li, M.Z., Leibham, D., Cortez, D., and Elledge, S.J. (1998). The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. Curr Biol *8*, 1300-1309.

Liu, Q., Li, M.Z., Liu, D., and Elledge, S.J. (2000). Rapid construction of recombinant DNA by the univector plasmid-fusion system. Methods Enzymol *328*, 530-549.

Liu, W.L., Coleman, R.A., Grob, P., King, D.S., Florens, L., Washburn, M.P., Geles, K.G., Yang, J.L., Ramey, V., Nogales, E.*, et al.* (2008). Structural changes in TAF4b-TFIID correlate with promoter selectivity. Mol Cell *29*, 81-91.

Liu, W.L., Coleman, R.A., Ma, E., Grob, P., Yang, J.L., Zhang, Y., Dailey, G., Nogales, E., and Tjian, R. (2009). Structures of three distinct activator-TFIID complexes. Genes Dev *23*, 1510-1521.

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science *327*, 206-209.

Loewer, A., Batchelor, E., Gaglia, G., and Lahav, G. Basal dynamics of p53 reveal transcriptionally attenuated pulses in cycling cells. Cell *142*, 89-100.

Lorch, Y., and Kornberg, R.D. (1985). A region flanking the GAL7 gene and a binding site for GAL4 protein as upstream activating sequences in yeast. J Mol Biol *186*, 821-824.

Lu, C., and Erickson, H.P. (1997). Expression in Escherichia coli of the thermostable DNA polymerase from Pyrococcus furiosus. Protein Expr Purif *11*, 179-184.

Lu, H., Zawel, L., Fisher, L., Egly, J.M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature *358*, 641-645.

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

Luo, J., Li, M., Tang, Y., Laszkowska, M., Roeder, R.G., and Gu, W. (2004). Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo. Proc Natl Acad Sci U S A *101*, 2259-2264.

Lustig, A.J., Kurtz, S., and Shore, D. (1990). Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. Science *250*, 549-553.

Luthra, R., Kerr, S.C., Harreman, M.T., Apponi, L.H., Fasken, M.B., Ramineni, S., Chaurasia, S., Valentini, S.R., and Corbett, A.H. (2007). Actively transcribed GAL genes can be physically linked to the nuclear pore by the SAGA chromatin modifying complex. J Biol Chem *282*, 3042-3049.

Ma, D., Olave, I., Merino, A., and Reinberg, D. (1996). Separation of the transcriptional coactivator and antirepression functions of transcription factor IIA. Proc Natl Acad Sci U S A *93*, 6583-6588.

Ma, J., and Ptashne, M. (1987a). Deletion analysis of GAL4 defines two transcriptional activating segments. Cell *48*, 847-853.

Ma, J., and Ptashne, M. (1987b). The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. Cell *50*, 137-142.

Malik, S., Guermah, M., Yuan, C.X., Wu, W., Yamamura, S., and Roeder, R.G. (2004). Structural and functional organization of TRAP220, the TRAP/mediator subunit that is targeted by nuclear receptors. Mol Cell Biol *24*, 8244-8254.

Marion, R.M., Regev, A., Segal, E., Barash, Y., Koller, D., Friedman, N., and 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*, 14315-14322.

Marmorstein, R. (2001). Structure and function of histone acetyltransferases. Cell Mol Life Sci *58*, 693-703.

Marmorstein, R., and Berger, S.L. (2001). Structure and function of bromodomains in chromatin-regulating complexes. Gene *272*, 1-9.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem *271*, 27176-27183.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J Biol Chem *270*, 12335-12338.

Martin, D.E., Soulard, A., and Hall, M.N. (2004). TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell *119*, 969-979.

Martin, S., and Pombo, A. (2003). Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. Chromosome Res *11*, 461-470.

Martinez, E., Chiang, C.M., Ge, H., and Roeder, R.G. (1994). TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. EMBO J *13*, 3115-3126.

Matangkasombut, O., Buratowski, R.M., Swilling, N.W., and Buratowski, S. (2000). Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. Genes Dev *14*, 951-962.

Matsui, T., Segall, J., Weil, P.A., and Roeder, R.G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. J Biol Chem *255*, 11992-11996.

Matsunaga, M., Jang, S.H., and Jaehning, J.A. (2004). Expression and purification of wild type and mutant forms of the yeast mitochondrial core RNA polymerase, Rpo41. Protein Expr Purif *35*, 126-130.

Mayor, T., and Deshaies, R.J. (2005). Two-step affinity purification of multiubiquitylated proteins from Saccharomyces cerevisiae. Methods Enzymol *399*, 385-392.

McKnight, S.L., Gavis, E.R., Kingsbury, R., and Axel, R. (1981). Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell *25*, 385-398.

Means, A.R., Woo, S., Harris, S.E., and O'Malley, B.W. (1975). Estrogen induction of ovalbumin mRNA: evidence for transcription control. Mol Cell Biochem *7*, 33-42.

Meisterernst, M., and Roeder, R.G. (1991). Family of proteins that interact with TFIID and regulate promoter activity. Cell *67*, 557-567.

Mencia, M., Moqtaderi, Z., Geisberg, J.V., Kuras, L., and Struhl, K. (2002). Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast. Mol Cell *9*, 823-833.

Menon, B.B., Sarma, N.J., Pasula, S., Deminoff, S.J., Willis, K.A., Barbara, K.E., Andrews, B., and Santangelo, G.M. (2005). Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. Proc Natl Acad Sci U S A *102*, 5749-5754.

Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003). Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. Cell *115*, 751-763.

Metzger, D., Scheer, E., Soldatov, A., and Tora, L. (1999). Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. EMBO J *18*, 4823-4834.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T*., et al.* (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261-1270.

Mohan, W.S., Jr., Scheer, E., Wendling, O., Metzger, D., and Tora, L. (2003). TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. Mol Cell Biol *23*, 4307-4318.

Moqtaderi, Z., Bai, Y., Poon, D., Weil, P.A., and Struhl, K. (1996a). TBP-associated factors are not generally required for transcriptional activation in yeast. Nature *383*, 188-191.

Moqtaderi, Z., Keaveney, M., and Struhl, K. (1998). The histone H3-like TAF is broadly required for transcription in yeast. Mol Cell *2*, 675-682.

Moqtaderi, Z., Yale, J.D., Struhl, K., and Buratowski, S. (1996b). Yeast homologues of higher eukaryotic TFIID subunits. Proc Natl Acad Sci U S A *93*, 14654-14658.

Muhich, M.L., Iida, C.T., Horikoshi, M., Roeder, R.G., and Parker, C.S. (1990). cDNA clone encoding Drosophila transcription factor TFIID. Proc Natl Acad Sci U S A *87*, 9148-9152.

Muller, M., Affolter, M., Leupin, W., Otting, G., Wuthrich, K., and Gehring, W.J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. EMBO J *7*, 4299-4304.

Munshi, N., Yie, Y., Merika, M., Senger, K., Lomvardas, S., Agalioti, T., and Thanos, D. (1999). The IFN-beta enhancer: a paradigm for understanding activation and repression of inducible gene expression. Cold Spring Harb Symp Quant Biol *64*, 149-159.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat Genet *39*, 1507-1511.

Naar, A.M., Beaurang, P.A., Robinson, K.M., Oliner, J.D., Avizonis, D., Scheek, S., Zwicker, J., Kadonaga, J.T., and Tjian, R. (1998a). Chromatin, TAFs, and a novel multiprotein coactivator are required for synergistic activation by Sp1 and SREBP-1a in vitro. Genes Dev *12*, 3020-3031.

Naar, A.M., Beaurang, P.A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. Nature *398*, 828-832.

Naar, A.M., Ryu, S., and Tjian, R. (1998b). Cofactor requirements for transcriptional activation by Sp1. Cold Spring Harb Symp Quant Biol *63*, 189-199.

Naar, A.M., Taatjes, D.J., Zhai, W., Nogales, E., and Tjian, R. (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. Genes Dev *16*, 1339-1344.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science *320*, 1344-1349.

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D.K., Hasegawa, J., Omori, M., Inukai, N.*, et al.* (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Mol Cell Biol *23*, 1863-1873.

Nechaev, S., and Adelman, K. (2008). Promoter-proximal Pol II: when stalling speeds things up. Cell Cycle *7*, 1539-1544.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335-338.

Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. (2002). Transcription activator interactions with multiple SWI/SNF subunits. Mol Cell Biol *22*, 1615-1625.

Neely, K.E., Hassan, A.H., Wallberg, A.E., Steger, D.J., Cairns, B.R., Wright, A.P., and Workman, J.L. (1999). Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. Mol Cell *4*, 649-655.

Negrini, S., Ribaud, V., Bianchi, A., and Shore, D. (2007). DNA breaks are masked by multiple Rap1 binding in yeast: implications for telomere capping and telomerase regulation. Genes Dev *21*, 292-302.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Nature *377*, 119-128.

Nikolov, D.B., Hu, S.H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N.H., Roeder, R.G., and Burley, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature *360*, 40-46.

Nishikawa, J., Kokubo, T., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1997). Drosophila TAF(II)230 and the transcriptional activator VP16 bind competitively to the TATA box-binding domain of the TATA box-binding protein. Proc Natl Acad Sci U S A *94*, 85-90.

Nonet, M., Scafe, C., Sexton, J., and Young, R. (1987a). Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. Mol Cell Biol *7*, 1602-1611.

Nonet, M., Sweetser, D., and Young, R.A. (1987b). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. Cell *50*, 909-915.

O'Brien, T., and Tjian, R. (1998). Functional analysis of the human TAFII250 N-terminal kinase domain. Mol Cell *1*, 905-911.

O'Brien, T., and Tjian, R. (2000). Different functional domains of TAFII250 modulate expression of distinct subsets of mammalian genes. Proc Natl Acad Sci U S A *97*, 2456-2461.

O'Neill, E.M., Kaffman, A., Jolly, E.R., and O'Shea, E.K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. Science *271*, 209-212.

Oelgeschlager, T., Chiang, C.M., and Roeder, R.G. (1996). Topology and reorganization of a human TFIID-promoter complex. Nature *382*, 735-738.

Oelgeschlager, T., Tao, Y., Kang, Y.K., and Roeder, R.G. (1998). Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAFIIs. Mol Cell *1*, 925-931.

Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M., and Roeder, R.G. (1995). Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. Mol Cell Biol *15*, 4856-4866.

Ohkuma, Y., and Roeder, R.G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. Nature *368*, 160-163.

Olave, I., Reinberg, D., and Vales, L.D. (1998). The mammalian transcriptional repressor RBP (CBF1) targets TFIID and TFIIA to prevent activated transcription. Genes Dev *12*, 1621-1637.

Otting, G., Qian, Y.Q., Muller, M., Affolter, M., Gehring, W., and Wuthrich, K. (1988). Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. EMBO J *7*, 4305-4309.

Ozer, J., Bolden, A.H., and Lieberman, P.M. (1996). Transcription factor IIA mutations show activator-specific defects and reveal a IIA function distinct from stimulation of TBP-DNA binding. J Biol Chem *271*, 11182-11190.

Papai, G., Tripathi, M.K., Ruhlmann, C., Layer, J.H., Weil, P.A., and Schultz, P. TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. Nature *465*, 956-960.

Papai, G., Tripathi, M.K., Ruhlmann, C., Werten, S., Crucifix, C., Weil, P.A., and Schultz, P. (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. Structure *17*, 363-373.

Payne, J.M., Laybourn, P.J., and Dahmus, M.E. (1989). The transition of RNA polymerase II from initiation to elongation is associated with phosphorylation of the carboxyl-terminal domain of subunit IIa. J Biol Chem *264*, 19621-19629.

Peterson, C.L., Dingwall, A., and Scott, M.P. (1994). Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. Proc Natl Acad Sci U S A *91*, 2905-2908.

Peterson, M.G., Tanese, N., Pugh, B.F., and Tjian, R. (1990). Functional domains and upstream activation properties of cloned human TATA binding protein. Science *248*, 1625-1630.

Peterson, M.G., and Tjian, R. (1993). Cross-species polymerase chain reaction: cloning of TATA box-binding proteins. Methods Enzymol *218*, 493-507.

Pham, A.D., Muller, S., and Sauer, F. (1999). Mesoderm-determining transcription in Drosophila is alleviated by mutations in TAF(II)60 and TAF(II)110. Mech Dev *84*, 3-16.

Pham, A.D., and Sauer, F. (2000). Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila. Science *289*, 2357-2360.

Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell *68*, 977-988.

Pluthero, F.G. (1993). Rapid purification of high-activity Taq DNA polymerase. Nucleic Acids Res *21*, 4850-4851.

Pokholok, D.K., Zeitlinger, J., Hannett, N.M., Reynolds, D.B., and Young, R.A. (2006). Activated signal transduction kinases frequently occupy target genes. Science *313*, 533-536.

Poon, D., Bai, Y., Campbell, A.M., Bjorklund, S., Kim, Y.J., Zhou, S., Kornberg, R.D., and Weil, P.A. (1995). Identification and characterization of a TFIID-like multiprotein complex from Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *92*, 8224-8228.

Powers, T., and Walter, P. (1999). Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. Mol Biol Cell *10*, 987-1000.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851-1854.

Preker, P., Nielsen, J., Schierup, M.H., and Jensen, T.H. (2009). RNA polymerase plays both sides: vivid and bidirectional transcription around and upstream of active promoters. Cell Cycle *8*, 1106-1107.

Prochasson, P., Neely, K.E., Hassan, A.H., Li, B., and Workman, J.L. (2003). Targeting activity is required for SWI/SNF function in vivo and is accomplished through two partially redundant activator-interaction domains. Mol Cell *12*, 983-990.

Pugh, B.F., and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell *61*, 1187-1197.

Pugh, B.F., and Tjian, R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. Genes Dev *5*, 1935-1945.

Qian, Y.Q., Billeter, M., Otting, G., Muller, M., Gehring, W.J., and Wuthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell *59*, 573-580.

Rabenstein, M.D., Zhou, S., Lis, J.T., and Tjian, R. (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. Proc Natl Acad Sci U S A *96*, 4791-4796.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolic, L., and Pikaard, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol Cell *33*, 192-203.

Reese, J.C., Apone, L., Walker, S.S., Griffin, L.A., and Green, M.R. (1994). Yeast TAFIIS in a multisubunit complex required for activated transcription. Nature *371*, 523-527.

Reese, J.C., and Green, M.R. (2003). Functional analysis of TFIID components using conditional mutants. Methods Enzymol *370*, 415-430.

Reid, J.L., Iyer, V.R., Brown, P.O., and Struhl, K. (2000). Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. Mol Cell *6*, 1297-1307.

Reinberg, D., and Roeder, R.G. (1987a). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. J Biol Chem *262*, 3310-3321.

Reinberg, D., and Roeder, R.G. (1987b). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. J Biol Chem *262*, 3331-3337.

Reinke, H., and Horz, W. (2003). Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. Mol Cell *11*, 1599-1607.

Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H., and Price, D.H. (2001). A highly purified RNA polymerase II elongation control system. J Biol Chem *276*, 42601-42609.

Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett *579*, 909-915.

Roeder, R.G., and Rutter, W.J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224*, 234-237.

Rohde, J.R., and Cardenas, M.E. (2003). The tor pathway regulates gene expression by linking nutrient sensing to histone acetylation. Mol Cell Biol *23*, 629-635.

Romier, C., James, N., Birck, C., Cavarelli, J., Vivares, C., Collart, M.A., and Moras, D. (2007). Crystal structure, biochemical and genetic characterization of yeast and E. cuniculi TAF(II)5 N-terminal domain: implications for TFIID assembly. J Mol Biol *368*, 1292-1306.

Rudra, D., Mallick, J., Zhao, Y., and Warner, J.R. (2007). Potential interface between ribosomal protein production and pre-rRNA processing. Mol Cell Biol *27*, 4815-4824.

Rudra, D., and Warner, J.R. (2004). What better measure than ribosome synthesis? Genes Dev *18*, 2431-2436.

Rudra, D., Zhao, Y., and Warner, J.R. (2005). Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. EMBO J *24*, 533-542.

Rundlett, S.E., Carmen, A.A., Kobayashi, R., Bavykin, S., Turner, B.M., and Grunstein, M. (1996). HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. Proc Natl Acad Sci U S A *93*, 14503-14508.

Ruppert, S., Wang, E.H., and Tjian, R. (1993). Cloning and expression of human TAFII250: a TBP-associated factor implicated in cell-cycle regulation. Nature *362*, 175-179.

Ryu, S., and Tjian, R. (1999). Purification of transcription cofactor complex CRSP. Proc Natl Acad Sci U S A *96*, 7137-7142.

Ryu, S., Zhou, S., Ladurner, A.G., and Tjian, R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. Nature *397*, 446-450.

Sacher, M., Pfander, B., and Jentsch, S. (2005). Identification of SUMO-protein conjugates. Methods Enzymol *399*, 392-404.

Sanders, S.L., Garbett, K.A., and Weil, P.A. (2002a). Molecular characterization of Saccharomyces cerevisiae TFIID. Mol Cell Biol *22*, 6000-6013.

Sanders, S.L., Jennings, J., Canutescu, A., Link, A.J., and Weil, P.A. (2002b). Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. Mol Cell Biol *22*, 4723-4738.

Sanders, S.L., Klebanow, E.R., and Weil, P.A. (1999). TAF25p, a non-histone-like subunit of TFIID and SAGA complexes, is essential for total mRNA gene transcription in vivo. J Biol Chem *274*, 18847-18850.

Santisteban, M.S., Kalashnikova, T., and Smith, M.M. (2000). Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. Cell *103*, 411-422.

Sauer, F., Wassarman, D.A., Rubin, G.M., and Tjian, R. (1996). TAF(II)s mediate activation of transcription in the Drosophila embryo. Cell *87*, 1271-1284.

Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1992). Purification and properties of Saccharomyces cerevisiae RNA polymerase II general initiation factor a. J Biol Chem *267*, 23383-23387.

Schawalder, S.B., Kabani, M., Howald, I., Choudhury, U., Werner, M., and Shore, D. (2004). Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. Nature *432*, 1058-1061.

Schmitt, M.E., Brown, T.A., and Trumpower, B.L. (1990). A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. Nucleic Acids Res *18*, 3091-3092.

Schultz, M.C., Choe, S.Y., and Reeder, R.H. (1991). Specific initiation by RNA polymerase I in a whole-cell extract from yeast. Proc Natl Acad Sci U S A *88*, 1004-1008.

Scolnick, D.M., Chehab, N.H., Stavridi, E.S., Lien, M.C., Caruso, L., Moran, E., Berger, S.L., and Halazonetis, T.D. (1997). CREB-binding protein and p300/CBP-associated factor are transcriptional coactivators of the p53 tumor suppressor protein. Cancer Res *57*, 3693-3696.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849-1851.

Seila, A.C., Core, L.J., Lis, J.T., and Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. Cell Cycle *8*, 2557-2564.

Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem *76*, 75-100.

Shao, H., Revach, M., Moshonov, S., Tzuman, Y., Gazit, K., Albeck, S., Unger, T., and Dikstein, R. (2005). Core promoter binding by histone-like TAF complexes. Mol Cell Biol *25*, 206-219.

Shen, W.C., Bhaumik, S.R., Causton, H.C., Simon, I., Zhu, X., Jennings, E.G., Wang, T.H., Young, R.A., and Green, M.R. (2003). Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. EMBO J *22*, 3395-3402.

Shen, W.C., and Green, M.R. (1997). Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator. Cell *90*, 615-624.

Sheppard, K.A., Rose, D.W., Haque, Z.K., Kurokawa, R., McInerney, E., Westin, S., Thanos, D., Rosenfeld, M.G., Glass, C.K., and Collins, T. (1999). Transcriptional activation by NF-kappaB requires multiple coactivators. Mol Cell Biol *19*, 6367-6378.

Sikorski, R.S., and Boeke, J.D. (1991). In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. Methods Enzymol *194*, 302-318.

Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. Curr Opin Cell Biol *21*, 344-351.

Singh, B.N., Ansari, A., and Hampsey, M. (2009). Detection of gene loops by 3C in yeast. Methods *48*, 361-367.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Mol Cell *27*, 806-816.

Singh, M.V., Bland, C.E., and Weil, P.A. (2004). Molecular and genetic characterization of a Taf1p domain essential for yeast TFIID assembly. Mol Cell Biol *24*, 4929-4942.

Singh, M.V., and Weil, P.A. (2002). A method for plasmid purification directly from yeast. Anal Biochem *307*, 13-17.

Sklar, V.E., Schwartz, L.B., and Roeder, R.G. (1975). Distinct molecular structures of nuclear class I, II, and III DNA-dependent RNA polymerases. Proc Natl Acad Sci U S A *72*, 348-352.

Smale, S.T., and Baltimore, D. (1989). The "initiator" as a transcription control element. Cell *57*, 103-113.

Smale, S.T., Schmidt, M.C., Berk, A.J., and Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc Natl Acad Sci U S A *87*, 4509-4513.

Spahr, H., Calero, G., Bushnell, D.A., and Kornberg, R.D. (2009). Schizosacharomyces pombe RNA polymerase II at 3.6-A resolution. Proc Natl Acad Sci U S A *106*, 9185-9190.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell *9*, 3273-3297.

Sreenath, H.K., Bingman, C.A., Buchan, B.W., Seder, K.D., Burns, B.T., Geetha, H.V., Jeon, W.B., Vojtik, F.C., Aceti, D.J., Frederick, R.O*., et al.* (2005). Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. Protein Expr Purif *40*, 256-267.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains

poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

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

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

Sumimoto, H., Ohkuma, Y., Yamamoto, T., Horikoshi, M., and Roeder, R.G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: identification of general transcription factor TFIIG. Proc Natl Acad Sci U S A *87*, 9158-9162.

Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Sun, Z.W., and Hampsey, M. (1995). Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *92*, 3127-3131.

Suzuki-Yagawa, Y., Guermah, M., and Roeder, R.G. (1997). The ts13 mutation in the TAF(II)250 subunit (CCG1) of TFIID directly affects transcription of D-type cyclin genes in cells arrested in G1 at the nonpermissive temperature. Mol Cell Biol *17*, 3284-3294.

Svaren, J., and Horz, W. (1997). Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast. Trends Biochem Sci *22*, 93-97.

Svaren, J., Schmitz, J., and Horz, W. (1994). The transactivation domain of Pho4 is required for nucleosome disruption at the PHO5 promoter. EMBO J *13*, 4856-4862.

Swanson, M.J., Qiu, H., Sumibcay, L., Krueger, A., Kim, S.J., Natarajan, K., Yoon, S., and Hinnebusch, A.G. (2003). A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. Mol Cell Biol *23*, 2800-2820.

Szutorisz, H., Dillon, N., and Tora, L. (2005). The role of enhancers as centres for general transcription factor recruitment. Trends Biochem Sci *30*, 593-599.

Taatjes, D.J., Naar, A.M., Andel, F., 3rd, Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. Science *295*, 1058-1062.

Taatjes, D.J., Schneider-Poetsch, T., and Tjian, R. (2004). Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. Nat Struct Mol Biol *11*, 664-671.

Takada, R., Nakatani, Y., Hoffmann, A., Kokubo, T., Hasegawa, S., Roeder, R.G., and Horikoshi, M. (1992). Identification of human TFIID components and direct interaction between a 250-kDa polypeptide and the TATA box-binding protein (TFIID tau). Proc Natl Acad Sci U S A *89*, 11809-11813.

Tan, S., Hunziker, Y., Sargent, D.F., and Richmond, T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. Nature *381*, 127-151.

Tan-Wong, S.M., Wijayatilake, H.D., and Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes Dev *23*, 2610-2624.

Tanese, N., Pugh, B.F., and Tjian, R. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. Genes Dev *5*, 2212-2224.

Tang, H., Kim, Y., Severinov, K., Goldfarb, A., and Ebright, R.H. (1996). Escherichia coli RNA polymerase holoenzyme: rapid reconstitution from recombinant alpha, beta, beta', and sigma subunits. Methods Enzymol *273*, 130-134.

Tang, H., Severinov, K., Goldfarb, A., and Ebright, R.H. (1995). Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant Escherichia coli RNA polymerase. Proc Natl Acad Sci U S A *92*, 4902-4906.

Tang, Y., Zhao, W., Chen, Y., Zhao, Y., and Gu, W. (2008). Acetylation is indispensable for p53 activation. Cell *133*, 612-626.

Thanos, D. (1996). Mechanisms of transcriptional synergism of eukaryotic genes. The interferon-beta paradigm. Hypertension *27*, 1025-1029.

Thuault, S., Gangloff, Y.G., Kirchner, J., Sanders, S., Werten, S., Romier, C., Weil, P.A., and Davidson, I. (2002). Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain. J Biol Chem *277*, 45510-45517.

Thut, C.J., Chen, J.L., Klemm, R., and Tjian, R. (1995). p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. Science *267*, 100-104.

Tokusumi, Y., Ma, Y., Song, X., Jacobson, R.H., and Takada, S. (2007). The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and

TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. Mol Cell Biol *27*, 1844-1858.

Tsang, J.S., Henry, Y.A., Chambers, A., Kingsman, A.J., and Kingsman, S.M. (1990). Phosphorylation influences the binding of the yeast RAP1 protein to the upstream activating sequence of the PGK gene. Nucleic Acids Res *18*, 7331-7337.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tsukiyama, T., Becker, P.B., and Wu, C. (1994). ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. Nature *367*, 525-532.

Tsukiyama, T., Daniel, C., Tamkun, J., and Wu, C. (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell *83*, 1021-1026.

Tsukiyama, T., and Wu, C. (1997). Chromatin remodeling and transcription. Curr Opin Genet Dev *7*, 182-191.

Tyler, R.C., Sreenath, H.K., Singh, S., Aceti, D.J., Bingman, C.A., Markley, J.L., and Fox, B.G. (2005). Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination. Protein Expr Purif *40*, 268-278.

Vaillancourt, P., Zheng, C.F., Hoang, D.Q., and Breister, L. (2000). Affinity purification of recombinant proteins fused to calmodulin or to calmodulin-binding peptides. Methods Enzymol *326*, 340-362.

Van Dyke, M.W., Roeder, R.G., and Sawadogo, M. (1988). Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. Science *241*, 1335-1338.

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., and Boelens, R. (2008). Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. Structure *16*, 1245-1256.

van Oevelen, C.J., van Teeffelen, H.A., and Timmers, H.T. (2005). Differential requirement of SAGA subunits for Mot1p and Taf1p recruitment in gene activation. Mol Cell Biol *25*, 4863-4872.

van Oevelen, C.J., van Teeffelen, H.A., van Werven, F.J., and Timmers, H.T. (2006). Snf1p-dependent Spt-Ada-Gcn5-acetyltransferase (SAGA) recruitment and chromatin remodeling activities on the HXT2 and HXT4 promoters. J Biol Chem *281*, 4523-4531.

van Werven, F.J., van Teeffelen, H.A., Holstege, F.C., and Timmers, H.T. (2009). Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. Nat Struct Mol Biol *16*, 1043-1048.

Vary, J.C., Jr., Gangaraju, V.K., Qin, J., Landel, C.C., Kooperberg, C., Bartholomew, B., and Tsukiyama, T. (2003). Yeast Isw1p forms two separable complexes in vivo. Mol Cell Biol *23*, 80-91.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., Hieter, P., Vogelstein, B., and Kinzler, K.W. (1997). Characterization of the yeast transcriptome. Cell *88*, 243-251.

Venter, U., Svaren, J., Schmitz, J., Schmid, A., and Horz, W. (1994). A nucleosome precludes binding of the transcription factor Pho4 in vivo to a critical target site in the PHO5 promoter. EMBO J *13*, 4848-4855.

Verma, R., and Deshaies, R.J. (2005). Assaying degradation and deubiquitination of a ubiquitinated substrate by purified 26S proteasomes. Methods Enzymol *398*, 391-399.

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., and Timmers, H.T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58-69.

Verrijzer, C.P., Chen, J.L., Yokomori, K., and Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell *81*, 1115-1125.

Verrijzer, C.P., Yokomori, K., Chen, J.L., and Tjian, R. (1994). Drosophila TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science *264*, 933-941.

Vidal, M., and Gaber, R.F. (1991). RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in Saccharomyces cerevisiae. Mol Cell Biol *11*, 6317-6327.

Vignais, M.L., Woudt, L.P., Wassenaar, G.M., Mager, W.H., Sentenac, A., and Planta, R.J. (1987). Specific binding of TUF factor to upstream activation sites of yeast ribosomal protein genes. EMBO J *6*, 1451-1457.

Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793-1808.

Wada, T., Orphanides, G., Hasegawa, J., Kim, D.K., Shima, D., Yamaguchi, Y., Fukuda, A., Hisatake, K., Oh, S., Reinberg, D.*, et al.* (2000). FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH. Mol Cell *5*, 1067-1072.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F.*, et al.* (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev *12*, 343-356.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. EMBO J *17*, 7395-7403.

Wade, J.T., Hall, D.B., and Struhl, K. (2004). The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature *432*, 1054-1058.

Wakula, P., Beullens, M., Ceulemans, H., Stalmans, W., and Bollen, M. (2003). Degeneracy and function of the ubiquitous RVXF motif that mediates binding to protein phosphatase-1. J Biol Chem *278*, 18817-18823.

Walker, S.S., Reese, J.C., Apone, L.M., and Green, M.R. (1996). Transcription activation in cells lacking TAFIIS. Nature *383*, 185-188.

Walker, S.S., Shen, W.C., Reese, J.C., Apone, L.M., and Green, M.R. (1997). Yeast TAF(II)145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state. Cell *90*, 607-614.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203-1206.

Wang, X., Sato, R., Brown, M.S., Hua, X., and Goldstein, J.L. (1994). SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. Cell *77*, 53-62.

Wang, X., Truckses, D.M., Takada, S., Matsumura, T., Tanese, N., and Jacobson, R.H. (2007). Conserved region I of human coactivator TAF4 binds to a short hydrophobic motif present in transcriptional regulators. Proc Natl Acad Sci U S A *104*, 7839-7844.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends Biochem Sci *24*, 437-440.

Weideman, C.A., Netter, R.C., Benjamin, L.R., McAllister, J.J., Schmiedekamp, L.A., Coleman, R.A., and Pugh, B.F. (1997). Dynamic interplay of TFIIA, TBP and TATA DNA. J Mol Biol *271*, 61-75.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. Cell *18*, 469-484.

Weinzierl, R.O., Dynlacht, B.D., and Tjian, R. (1993a). Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator. Nature *362*, 511-517.

Weinzierl, R.O., Ruppert, S., Dynlacht, B.D., Tanese, N., and Tjian, R. (1993b). Cloning and expression of Drosophila TAFII60 and human TAFII70 reveal conserved interactions with other subunits of TFIID. EMBO J *12*, 5303-5309.

Werten, S., Mitschler, A., Romier, C., Gangloff, Y.G., Thuault, S., Davidson, I., and Moras, D. (2002). Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). J Biol Chem *277*, 45502-45509.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004a). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell *119*, 481-489.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004b). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science *303*, 1014-1016.

Woontner, M., Wade, P.A., Bonner, J., and Jaehning, J.A. (1991). Transcriptional activation in an improved whole-cell extract from Saccharomyces cerevisiae. Mol Cell Biol *11*, 4555-4560.

Woychik, N.A., Liao, S.M., Kolodziej, P.A., and Young, R.A. (1990). Subunits shared by eukaryotic nuclear RNA polymerases. Genes Dev *4*, 313-323.

Wright, K.J., Marr, M.T., 2nd, and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. Proc Natl Acad Sci U S A *103*, 12347-12352.

Wu, C. (1997). Chromatin remodeling and the control of gene expression. J Biol Chem *272*, 28171-28174.

Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev *17*, 1402-1414.

Xie, J., Collart, M., Lemaire, M., Stelzer, G., and Meisterernst, M. (2000). A single point mutation in TFIIA suppresses NC2 requirement in vivo. EMBO J *19*, 672-682.

Xu, F., Zhang, K., and Grunstein, M. (2005). Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell *121*, 375-385.

Xu, Y.X., Hirose, Y., Zhou, X.Z., Lu, K.P., and Manley, J.L. (2003). Pin1 modulates the structure and function of human RNA polymerase II. Genes Dev *17*, 2765-2776.

Xu, Y.X., and Manley, J.L. (2007). Pin1 modulates RNA polymerase II activity during the transcription cycle. Genes Dev *21*, 2950-2962.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol Cell *21*, 227-237.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41-51.

Yamamoto, T., Poon, D., Weil, P.A., and Horikoshi, M. (1997). Molecular genetic elucidation of the tripartite structure of the Schizosaccharomyces pombe 72 kDa TFIID subunit which contains a WD40 structural motif. Genes Cells *2*, 245-254.

Yarragudi, A., Miyake, T., Li, R., and Morse, R.H. (2004). Comparison of ABF1 and RAP1 in chromatin opening and transactivator potentiation in the budding yeast Saccharomyces cerevisiae. Mol Cell Biol *24*, 9152-9164.

Yarragudi, A., Parfrey, L.W., and Morse, R.H. (2007). Genome-wide analysis of transcriptional dependence and probable target sites for Abf1 and Rap1 in Saccharomyces cerevisiae. Nucleic Acids Res *35*, 193-202.

Yatherajam, G., Zhang, L., Kraemer, S.M., and Stargell, L.A. (2003). Protein-protein interaction map for yeast TFIID. Nucleic Acids Res *31*, 1252-1260.

Yie, J., Merika, M., Munshi, N., Chen, G., and Thanos, D. (1999a). The role of HMG I(Y) in the assembly and function of the IFN-beta enhanceosome. EMBO J *18*, 3074-3089.

Yie, J., Senger, K., and Thanos, D. (1999b). Mechanism by which the IFN-beta enhanceosome activates transcription. Proc Natl Acad Sci U S A *96*, 13108-13113.

Yokomori, K., Chen, J.L., Admon, A., Zhou, S., and Tjian, R. (1993). Molecular cloning and characterization of dTAFII30 alpha and dTAFII30 beta: two small subunits of Drosophila TFIID. Genes Dev *7*, 2587-2597.

Yokomori, K., Zeidler, M.P., Chen, J.L., Verrijzer, C.P., Mlodzik, M., and Tjian, R. (1994). Drosophila TFIIA directs cooperative DNA binding with TBP and mediates transcriptional activation. Genes Dev *8*, 2313-2323.

Yokoyama, C., Wang, X., Briggs, M.R., Admon, A., Wu, J., Hua, X., Goldstein, J.L., and Brown, M.S. (1993). SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. Cell *75*, 187-197.

Young, R.A., and Davis, R.W. (1983). Yeast RNA polymerase II genes: isolation with antibody probes. Science *222*, 778-782.

Yu, L., Sabet, N., Chambers, A., and Morse, R.H. (2001). The N-terminal and C-terminal domains of RAP1 are dispensable for chromatin opening and GCN4-mediated HIS4 activation in budding yeast. J Biol Chem *276*, 33257-33264.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet *39*, 1512-1516.

Zerby, D., and Lieberman, P.M. (1997). Functional analysis of TFIID-activator interaction by magnesium-agarose gel electrophoresis. Methods *12*, 217-223.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell *98*, 811-824.

Zhang, H., Kruk, J.A., and Reese, J.C. (2008). Dissection of coactivator requirement at RNR3 reveals unexpected contributions from TFIID and SAGA. J Biol Chem *283*, 27360-27368.

Zhao, Y., McIntosh, K.B., Rudra, D., Schawalder, S., Shore, D., and Warner, J.R. (2006). Fine-structure analysis of ribosomal protein gene transcription. Mol Cell Biol *26*, 4853-4862.

Zhou, J., Zwicker, J., Szymanski, P., Levine, M., and Tjian, R. (1998). TAFII mutations disrupt Dorsal activation in the Drosophila embryo. Proc Natl Acad Sci U S A *95*, 13483-13488.

Zhu, W., Wada, T., Okabe, S., Taneda, T., Yamaguchi, Y., and Handa, H. (2007). DSIF contributes to transcriptional activation by DNA-binding activators by preventing pausing during transcription elongation. Nucleic Acids Res *35*, 4064-4075.