UNIVERSITY OF CALIFORNIA
RIVERSIDE

Transcriptional Regulation by TAF1 Kinase

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Biochemistry and Molecular Biology

by

Thomas Charles Benedict

December 2020

Dissertation Committee:
Dr. Xuan Liu, Chairperson
Dr. Jeff Perry
Dr. Jikui Song

The Dissertation of Thomas Charles Benedict is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to acknowledge all the wonderful people who helped me achieve this dream. First, I would like to acknowledge my advisor, Xuan Liu, for all the mentorship and support throughout this long process. I would also like to thank my committee members, Dr. Jeff Perry, and Dr. Jikui Song who have supported me throughout my Ph.D. and offered invaluable guidance and advice. I would like to thank members of the Xuan Liu lab, past and current, who have contributed to the work described here. Dr. Lily Maxham trained me in the lab on many of the techniques used in my research and started the project that will be presented in Chapter 3. Lily performed the bioinformatic analyses and many of the ChIP assays that were fundamental in developing the project that my research contributed to. I would like acknowledge Dr. Joy Lin who performed the TAF1 ChIP-seq which provided critical data that contributed to the research described in Chapter 3. I would like to thank Dr. Jacob Schrier for helping with the Reactome pathway analysis and also for being a great friend and lab mate. I would like to acknowledge Dr. Carina Edmondson for the continued support both in the lab and as a friend, and for making this experience so memorable. Carina held the entire lab together during her tenor, making sure we always had everything we needed. I would like to acknowledge Qi Zhou for providing endless comedic relief during stressful times and for all the reminders that helped keep my research on track.

I would like to acknowledge my parents who supported and encouraged me throughout the Ph.D. process. My mother, Marti Benedict, for always checking in on me, providing the best advice, and for being a truly inspirational figure in my life. My father, Dr. Mark Benedict, for the invaluable support throughout my education and for being the inspiration for my career in biology. I would also like to acknowledge the rest of my family who have always been so supportive of my career in science: Megan, Jess, Cathy, Dan, Caitlin, and Mary. Lastly, I would like to acknowledge my cat, Gene, for being the best little companion.

Dedication

ABSTRACT OF THE DISSERTATION

Transcriptional Regulation by TAF1 Kinase

by

Thomas Charles Benedict

Doctor of Philosophy, Graduate Program in Biochemistry and Molecular Biology
University of California, Riverside, December 2020
Dr. Xuan Liu, Chairperson

Transcription of protein coding genes and some non-coding RNAs relies on RNA polymerase II (pol II). Positioning the polymerase at the correct location at the start of a gene is aided by a series of general transcription factors (GTFs): TFIIA, B, D, E, F, and H. Together with pol II, these factors form the preinitiation complex. TFIID is first to recognize and bind to a gene promoter and helps facilitate recruitment of the other components. TAF1 is the largest subunit of TFIID and aside from its role as a GTF, it functions to regulate gene specific transcription. TAF1 has intrinsic kinase activity and has been shown to phosphorylate other members of the general transcription machinery as well as the tumor suppressor p53. Previous studies have shown that phosphorylation of p53 by TAF1 leads to dissociation from the p21 promoter and termination of transcription. Those studies also found that TAF1 kinase responds to fluctuations in cellular ATP, having a Km_{ATP} of 1.9mM. In the early stages following DNA damage, cells undergo cellular ATP depletion which was found to effect TAF1-

mediated phosphorylation of p53 giving rise to a regulatory mechanism of p21 transcription. The studies presented here investigate the possibility that TAF1 regulates other sequence specific transcription factors, like p53, through phosphorylation. The work described in this dissertation includes the purification of many proteins for *in vitro* characterization of two novel TAF1 kinase targets, E2F1 and FOXM1. Furthermore, this research investigates the effect that TAF1-mediated phosphorylation has on E2F1 and FOXM1 promoter occupancy and transcription of target genes. My studies provide new insights into how TAF1 regulates gene-specific transcription through phosphorylation, broadening our understanding of this complex GTF.

Table of Contents	Page
Chapter 1 Introduction	
1.1 Basal Transcription Machinery of Eukaryotes	2
1.2 TAF1	8
1.3 TAF1 and p53	17
1.4 E2F1	19
1.5 FOXM1	21
1.6 Summary of Chapters	24
1.7 References	25
1.8 Tables and Figures	34
Chapter 2: Expression and Purification of TAF1, CDK8, RAP74, E2F1, FOXM1, ATF2 and SP1	
2.1 Introduction	39
2.2 Results	41
2.3 Discussion	54
2.4 Materials and Methods	57
2.5 References	63
2.6 Figures	65
Chapter 3 TAF1 regulates transcriptional activity of E2F1 and FOXM1 through phosphorylation	
3.1 Abstract	73
3.2 Introduction	74
3.3 Results	76
3.4 Discussion	83
3.5 Materials and Methods	87

3.6 References	93
3.7 Tables and Figures	98
Chapter 4 Conclusion	
4.1 Conclusion	107
4.2 References	112

List of Figures	Page
Chapter 1	
Figure 1.1 Formation of the preinitiation complex	34
Figure 1.2 Domains of TAF1	35
Figure 1.3 p53 regulation by TAF1 kinase	37
Chapter 2	
Figure 2.1 Purification of TAF1 using baculovirus	65
Figure 2.2 Purification of kinase dead A2/N7 Δ CTK TAF1	66
Figure 2.3 Purification of TAF1 NTK and CTK	67
Figure 2.4 Purification of RAP74	68
Figure 2.5 Purification of E2F1: full length, truncations, and mutants	69
Figure 2.6 Purification of bacterial and insect cell expressed FOXM1	70
Figure 2.7 Purification of ATF2	71
Figure 2.8 Purification of Sp1	71
Chapter 3	
Figure 3.1 TAF1 ChIP-seq and analysis	98
Figure 3.2 TAF1 ChIP-seq at 8 hours following UV and 8 hours plus apigenin	99
Figure 3.3 Binding patterns of FOXM1, Sp1, ATF2, and E2F1 following UV treatment	101
Figure 3.4 TAF1 phosphorylates E2F1 and FOXM1	102
Figure 3.5 TAF1 phosphorylates E2F1 at Ser375	103
Figure 3.6 Purification of RAP74, E2F1, FOXM1, ATF2, Sp1, TAF1, CDK8, Cyclin C, and MED12	104
Figure 3.7 Phosphorylation of E2F1 and FOXM1 leads to reduced promoter occupancy and target gene expression	105

List of Tables	Page
Chapter 1	
Tables 1.1 10 most commonly mutated transcription factors across 11 cancer types	36
Chapter 3	
Table 3.1 Top 100 ranked transcription factors bound to 421 promoters	100

Chapter 1: Introduction

1.1 Basal transcription machinery of Eukaryotes

Transcriptional regulation is a key component of gene expression and is vital for the proper function of cellular processes of every known organism. Dysregulation of transcription is a hallmark of developmental defects and disease in eukaryotes. Both activation and termination of transcription are highly dynamic processes involving complex interactions between the general transcription machinery and factors that associate with specific DNA sequences in response to cellular signals. Additionally, the structure of chromatin has a profound impact on transcription which can be modified via epigenetic writers and subsequently sensed by epigenetic readers in order to activate or suppress expression of genes.

RNA polymerases are the enzymes responsible for making an RNA molecule from a DNA template; the process titled transcription (Weiss and Gladstone, 1959). In eukaryotes, there are three RNA polymerases I, II, and III (it should be noted that a fourth exists in plants); RNA polymerase I transcribes 18s and 28s ribosomal RNAs and RNA polymerase III transcribes cellular 5s rRNA, tRNA, and viral associated RNAs (Herr et al., 2005; Roeder and Rutter, 1969; Roeder and Rutter, 1970). RNA polymerase II (Pol II) is responsible for b transcribing all protein coding genes to mRNA transcripts and is recruited to promoters (classified as class II promoters due to the polymerase that binds them) through interaction with general transcription factors (GTFs) (Roeder and

Rutter, 1970; Weil et al., 1979;). These transcription factors designated TFII (A-F) along with mediator complexes and RNA polymerase II form what is called the preinitiation complex (PIC) at specific locations throughout the genome which allows for controlled expression of protein coding genes (Flores et al., 1989; Flores et al., Ge at al., 1996; 1992; Matsui et al., 1980; Samuels et al., 1982; Sawadogo and Roeder, 1985). The recognition of proper genomic locations for the GTFs and RNA polymerase II to bind to relies on core promoter elements in the DNA sequence (Smale and Kadonaga, 2003). Through classical biochemical experiments and more recently with the use of modern structural analysis, such as Cryo-EM, the order of events in forming the preinitiation complex and the specific functions of each component are being elucidated.

Turning transcription on and off

Proper regulation of gene expression requires the ability to activate transcription and just as importantly, to turn off transcription at the right times. Molecular signals coming from inside and outside the cell are transduced into the nucleus through cascades of posttranslational modifications and interactions between proteins and other molecules to ultimately alter the transcription of genes. These signaling pathways often lead to alterations in the functional properties of transcription factors, affecting their ability to interact with the basal machinery in order to activate or repress transcription. A classic example of this

is mitogenic signaling which on the cellular level typically begins with the signaling molecule, EGF, binding to the receptor tyrosine kinase, EGFR, this in turn activates the MAPK cascade where activated MAPK proteins can phosphorylate transcription factors that regulate cell cycle genes (Cohen, 1965; Cohen et al., 1980; Seger and Krebs, 1995). For this pathway to successfully cause a cell to undergo mitosis, transcription of genes that halt cell cycle must be turned off, while at the same time genes that promote cell cycle must be transcriptionally activated. This means that phosphorylation, or other PTMs, by proteins activated in this pathway will enhance the ability of some transcription factors to bind onto target gene promoters while disrupting it for others. As the resulting cells exit mitosis, the transcriptional program must return to the pre-signaling state which is the case with any transcriptional response to cellular signaling (Seger and Krebs, 1995).

Formation of the preinitiation complex:

There is significant evidence suggesting the existence of two pathways leading to the formation of the preinitiation complex. One being a stepwise recruitment of the GTFs and RNA polymerase to the promoter of a gene and the other being the formation of a holoenzyme complex consisting of Pol II and some of the GTFs followed by DNA recognition and binding (Figure 1.1). In the sequential assembly pathway, TFIID recognizes and binds to a gene promoter

first, making contacts with the TATA-box as well as downstream promoter elements (DPE) and the initiator sequence at the transcription start site (TSS) (Chalkey and Verrijzer, 1999; Burke and Kadonaga, 1996). TFIIA is then recruited followed by TFIIB which binds to the B recognition element (BRE), a 7-nucleotide cis-regulatory element just downstream of the TSS (Lagrange et al., 1998). Next, RNA polymerase II in complex with TFIIF are recruited and lastly, TFIIE and TFIIH bind (Thomas and Chiang, 2006). The order of assembly in this pathway was determined through native gel electrophoresis and DNase I footprinting (Buratowski et al., 1989). This is further supported by an experiment that showed incubation of adenovirus major late promoter-containing DNA with purified TFIID before addition of the remaining components of the preinitiation complex led to a more stable complex that was not outcompeted by a second similar DNA construct (Van Dyke et al., 1989).

The second pathway leading to formation of the preinitiation complex involves a preassembled RNA polymerase II holoenzyme that includes some GTFs, chromatin remodelers and suppressers of RNA polymerase II B mutations (SRBs) (Kim et al., 1994; Koleske and Young, 1994). TFIID and TFIIA are not found in this complex but instead, recognize and bind to a promoter first and then facilitate recruitment of the RNA polymerase II holoenzyme. This pathway was discovered when it was found that RNA polymerase II not bound to DNA could be copurified with TFIIB, TFIIE, TFIIF, THIIH, GCN5, SWI/SNF and SRBs but not TFIIA or TFIID (Wu and Chiang, 1998; Wu et al., 1999). The exact components

of the holoenzyme vary with experimental procedures, though multiple research groups found a similar RNA polymerase II holoenzyme, all lacking TFIID (Thomas and Chiang, 2006). With substantial evidence supporting both preinitiation complex-forming pathways, it is likely that some combination of both are utilized inside the cell. Importantly, both the stepwise assembly pathway and the holoenzyme pathway indicate that TFIID is a key factor in promoter recognition.

TFIID

TFIID is a multiprotein 1.3MDa complex consisting of TATA box binding protein (TBP) and 13 TBP associated factors (TAFs) with TAF1 being the largest subunit and TAF13 the smallest. Cryo-electron microscopy (Cryo-EM) coupled with crystallographic studies, NMR, and chemical crosslinking mass spectrometry has allowed for relatively complete structural definition of TFIID which has been described as having a horseshoe shape (Andel et al., 1999). The complex is made of 3 lobes, A, B, and C, and exists in two major conformations, an open state which is referred to as the canonical conformation, and a rearranged state which is seen only when bound to DNA (Andel et al., 1999; Cianfrocco et al., 2013; Kolesnikova et al., 2018; Patel et al., 2018). Lobes B and C interact forming a relatively rigid structure, whereas lobe A moves up to 150Å from interacting with lobe C in the canonical state to interacting with lobe B in the

rearranged state. The geography of the DNA relative to TFIID found in Cryo-EM studies shows lobe C interacting with downstream promoter elements and lobe A (in the rearranged state, near lobe b) interacting with the TATA box (Cianfrocco et al., 2013).

Lobe C of TFIID contains a dimer of TAF6, TAF2 and a TAF1/TAF7 heterodimer, where TAF1 binds DNA at the DPE and initiator (Inr) sequences through its multiple DNA binding domains (Cianfrocco et al., 2013; Kolesnikova et al., 2018). Lobe B consists of TAF6/TAF9, TAF4/TAF12, and TAF8/TAF10 dimers each forming histone fold domains (HFD) which interact with the WD40 and helical domains of TAF5 (Patel et al., 2018). The HFD of TAF8 interacts with the TAF6 dimer and TAF2, tethering lobe B to Lobe C which form the stable TFIID core (Patel et al., 2018). Lobe A is comprised of TAF6/TAF9, TAF4/TAF12, TAF3/TAF10, and TAF11/TAF13 dimers forming an HFD octamer that resembles a nucleosome core (Patel et al., 2018). Due to the similarity in structure of lobe A's HFD octamer to a nucleosome core, it was previously predicted that this structure would be involved in DNA wrapping, however, the most recent cryo-EM structures indicate that it does not interact with DNA and does not even retain the highly positive charges necessary for binding the DNA backbone as is the case with a histone octamer (Patel et al., 2018). Here it seems the HFD octamer plays a scaffolding role (Patel et al., 2018). Similarly, TAF2 has strong sequence homology to an aminopeptidase, however, does not contain key catalytic residues and is instead utilized for scaffolding (Patel et al., 2018).

With respect to seeding the preinitiation complex, one main function of TFIID is to load TBP onto the core promoter (Patel et al., 2018; Pugh and Tjian, 1991). TBP initially interacts with lobe A and is held in an inhibitory state, upon DNA binding, TBP is released from lobe A and positioned on the promoter so that scanning can take place and finally, bending of the DNA by the fully engaged TFIID (Patel et al., 2018). TAF1's N-terminal domains, TAND1 and TAND2, block TBP from binding DNA and from associating with TFIIA; additionally, TAF11/TAF13 block the binding site on TBP for TFIIB (Kokubo et al., 1993; Liu et al., 1998; Patel et al., 2018). Once lobe C binds stably to downstream promoter elements, conformational changes in the complex allow these inhibitory interactions to be overcome and TBP can engage the promoter and recruit TFIIA and TFIIB which further stabilize its position and allow for recruitment of RNA polymerase II and subsequent transcription initiation (Patel et al., 2018).

1.2 TAF1

TAF1, the largest subunit of TFIID, is a 250kDa protein containing multiple functional domains and plays an important role in genome-wide transcriptional regulation (Figure. 1.2). TAF1 consists of a histone acetyltransferase domain (HAT), a ubiquitin activating/conjugating domain (designated UBAC or E1/E2), a double bromodomain (DBrD), zinc knuckle and winged-helix (WH) DNA binding domains, a RAP74 interacting domain (RAPiD), and has intrinsic kinase activity

facilitated by N- and C-terminal kinase domains (NTK and CTK) (Dikstein et al., 1996; Jacobson et al., 2000; Mizzen et al., 1996; Pham and Sauer, 2000; Curran et al., 2018). Some literature describes a large central domain in the middle of the protein as a domain of unknown function, DUF3591, which spans the HAT and winged-helix domains and shows strong evolutionary conservation in eukaryotes (Curran et al., 2018; Wang et al., 2014). In the N-terminus of the protein, 2 regions that have been identified to interact with TBP are sometimes referred to as TAND1 and TAND2 or together, simply, the TBP interacting domain (Kokubo et al., 1993; Liu et al). The TAF1 gene, located on the X chromosome, has 12 isoforms (in humans) due to splice variation, as well as 15 natural variants each consisting of just a single base pair change (The UniProt Consortium, 2019). Much of the TAF1 sequence is highly conserved in eukaryotes, however the C-terminus of the protein shows homology to a separate protein in yeast, BDF1, which through some evolutionary means has become attached to TAF1 in higher eukaryotes (Sawa et al., 2004).

TAF1 is a G1 progression-required protein which was first established in a temperature sensitive mutant hamster cell line (ts13) containing a single base pair mutation in the TAF1 homolog CCG1 HAT domain (Sekiguchi et al., 1995). It was found that when the temperature is raised from the permissive temperature (33°C) where the cells grow normally, to the restrictive temperature (39.5°C) the cells would arrest in G1. This phenotype could be rescued by expression of human TAF1, marking it as an important regulator of cell cycle (Wang and Tjian,

1994; Sekiguchi et al., 1995). Rescue efficiency is greatly reduced when TAF1 mutants deficient in kinase activity or HAT activity are transfected into the ts13 cells, suggesting that the kinase activity of TAF1 is important for cell cycle progression (O'Brien and Tjian, 1998). Similarly, mutations in either the WH or Zinc Knuckle DNA binding domains results in a drastic reduction in rescue efficiency (Curran et al., 2018). Taken together, these results suggest that both TAF1s ability to modify other proteins and its ability to bind DNA itself are critical in its role in transcription and subsequently, cell cycle progression.

TAF1 domains

Histone acetyltransferase (HAT) domain

TAF1 has histone acetyltransferase activity specific for histones H3 and H4, as well as a non-histone target, the β-subunit of TFIIE (Imhof et al., 1997; Mizzen et al., 1996). Thus far, there is no specific function associated with the acetylation of TFIIE by TAF1 and it has only been seen in vitro with purified proteins. Although TAF1 has not been found to be a major HAT responsible for histone acetylation, contributing very little to global acetylation levels, it has been found to be important for the regulation of specific genes (Durant and Pugh, 2006; Hilton et al., 2005). In the temperature sensitive ts13 mutant cell line discussed in the previous section, transcriptional activation of cell cycle regulatory proteins, cyclin A, cyclin D1, and cyclin E, is diminished when the cells

are grown at the restrictive temperature, and introduction of HAT-deficient TAF1 does not rescue activation of these genes (Hilton et al., 2005). Importantly, the HAT-deletion mutant TAF1 used in those experiments was still found to bind TFIID as efficiently as wildtype TAF1, further suggesting enzymatic HAT activity is vital for activation of those cell cycle genes and the domain is not merely a scaffold used in recruitment of the transcriptional machinery (Hilton et al., 2005).

Ubiquitin activating/conjugating (UBAC) domain

TAF1 has a ubiquitin-activating/conjugating domain that has been shown to monoubiquitinate the H1 linker histone (Pham and Sauer, 2000). Nuclear extract was subjected to SDS-page using gels containing histones, bound proteins were then renatured and tested for ubiquitination activity, leading to the identification of TAF1 as a E1/E2 UBAC enzyme (Pham and Sauer, 2000). Subsequent in vitro ubiquitination assays showed TAF1 specifically monoubiquitinates H1 histones and the domain responsible for this activity lies somewhere between amino acids 768-1218 (Pham and Sauer, 2000). It was also shown that cells expressing UBAC-deficient TAF1 had globally reduced levels of monoubiquitinated H1 (Pham and Sauer, 2000). While polyubiquitination of proteins is generally associated with proteasomal degradation, monoubiquitination can have other regulatory effects. Mutations in the UBAC domain of TAF1 in drosophila embryos is correlated to a reduction in

transcription, implying H1 monoubiquitination by TAF1 may be important for transcriptional activation (Pham and Sauer, 2000). It has also been shown that TAF1 contributes to monoubiquitination of Pax3, facilitating its proteasomal degradation; this interaction is linked to regulation of myogenic differentiation (Boutet et al., 2010).

Double bromodomain (DBrD)

The C-terminus of TAF1 contains two bromodomain motifs, together called a double bromodomain (DBrD) which binds to acetylated lysine residues on both histone and non-histone proteins (Jacobson et al., 2000; Li et al., 2007). Through isothermal titration calorimetry experiments, it was found that the DBrD binds to a fully acetylated H4 peptide (having 4 acetylation sites: K5, K8, K12, and K16) with an affinity of ~5μM and a stoichiometric ratio of 1:1 (Jacobson et al., 2000). Further experiments demonstrated that an H4 peptide with two of the sites acetylated binds with higher affinity than a peptide with only one acetyllysine, suggesting that binding of one of the bromodomain motifs may have a cooperative effect on binding of the other (Jacobson et al., 2000). Aside from binding to histones, TAF1 also binds to acetyllysines on non-histone proteins via the DBrD. Diacetylation of the tumor suppressor p53 by p300 serves as a transcriptional activation switch by recruiting TFIID to target genes through TAF1s DBrD (Li et al., 2007). The presence of both a HAT and DBrD domain

gives TAF1 the ability to recruit TFIID to activated regions of chromatin through acetyllysine binding and to then further modify the chromatin, promoting transcriptional activation. This theme is seen in other HATs, for example the CBP/P300 co-activator contains both HAT and bromo- domains (Chen et al., 2010).

Winged helix and zinc knuckle DNA binding domains

Recruitment of TFIID to gene promoters is facilitated in part by the ability of TAF1 to directly bind DNA. Two separate DNA binding domains have been identified and structurally characterized as a winged helix-turn-helix motif and a conserved zinc knuckle motif (Curran et al., 2018; Wang et al., 2014). Mutations in either DNA binding domain results in a decrease in TFIID occupancy on promoter DNA and a reduction in cell proliferation and viability (Curran et al., 2018; Wang et al., 2014). The ability of exogenously expressed TAF1 to rescue the ts13 mutant cell line grown at the restrictive temperature is equally reduced when mutations to either one of DNA binding domains are introduced and this is further exacerbated with mutations in both domains; demonstrating the critical function of both domains in TAF1-mediated TFIID recruitment (Curran et al., 2018; Wang et al., 2014).

Kinase Domains

In an attempt to determine if a kinase was associated with TFIID, the complex was immunopurified followed by an in vitro kinase assay which showed TAF1 becomes phosphorylated (Dikstein et al., 1996). Further analysis revealed TAF1 itself has intrinsic kinase activity, capable of autophosphorylation as well as specific transphosphorylation of the TFIIF subunit, RAP74 (Dikstein et al., 1996). Deletion analysis lead to the discovery of two separate kinase domains within TAF1, both separately capable of auto- and transphosphorylation of serine and threonine residues (Dikstein et al., 1996). Since then other substrates or TAF1 kinase activity has been discovered; TAF7, another subunit in TFIID, as well as the tumor suppressor p53 (Gegonne et al., 2005; Li et al., 2004). No mechanistic function of RAP74 phosphorylation by TAF1 has been published though it has been speculated that it's likely associated with transcriptional regulation (Dikstein et al., 1996). TAF7 and TAF1 interact very extensively in the TFIID complex, the structure of which has been analyzed crystallographically, and it was found that TAF7 represses TAF1 HAT activity until phosphorylation of TAF7 by TAF1 which leads to a conformational change that relieves this repression (Gegonne et al., 2001; Gegonne et al., 2005; Kloet et al., 2012) . This interaction links TAF1 kinase activity to the regulation of cell cycle specific genes, as the HAT activity was shown to be important for expression of a number of these (Kloet et al., 2012). TAF1 also phosphorylates tumor suppressor p53, a transcription factor heavily involved in cell cycle regulation (Li et al., 2004). It was shown that p53

phosphorylation by TAF1 causes dissociation of both proteins from the p21 promoter, acting as a shutoff switch for this cell cycle repressor (Li et al., 2004; Wu et al., 2014). To further demonstrate the importance of the kinase domains of TAF1 in cell cycle-specific transcriptional regulation, it was found that rescue of the ts13 temperature sensitive mutant cell line when grown at the restrictive temperature is greatly reduced upon addition of TAF1 deficient in kinase activity, a similar result as seen with HAT-deficient and DNA-binding-deficient TAF1 mutants (O'brien and Tjian, 1998). One known inhibitor of TAF1 kinase is the naturally occurring flavone, apigenin, though this compound has been shown to inhibit a handful of other ATP binding proteins, it has been used in research studies of TAF1 kinase (Critchfield et al., 1997; Liu et al., 2004).

TAF1 in disease

X-Linked dystonia parkinsonism and X-linked intellectual disability syndrome

Reduced expression of a neuron-specific variant, N-TAF1, as well as the presence of an SVA-type retrotransposon in exon 32, has been linked to a neurodegenerative disease, X-linked dystonia parkinsonism (XDP), a disease that has only been found in people of Filipino descent (Ito et al., 2016). Other TAF1 variants have been linked to a specific phenotype characterized by facial dysmorphology, intellectual disability, and other neurological effects; seen only in

males (O'Rawe et al., 2015). Bioinformatic studies revealed 8 single-nucleotide variants (SNVs) and TAF1 gene duplication specific to individuals displaying this phenotype (O'Rawe et al., 2015). Interestingly, 4 of the SNVs involve changes to amino acids within the region of TAF1 important for TAF7 binding and HAT regulation, and another SNV found to change an amino acid in one the bromodomains (O'Rawe et al., 2015).

TAF1 in cancer

Advancements in bioinformatics and computational biology have leaped forward in recent years as more tumor samples are sequenced and made available on platforms like The Cancer Genome Atlas (TCGA). A 2013 study analyzed mutations in 3281 tumors across 11 cancer types and found TAF1 to be one of the most frequently mutated general transcription factors (Table 1.1, Kandoth et al., 2013). The highest rates of TAF1 mutation were found in uterine corpus endometrial carcinoma (UCEC) and second highest in lung squamous cell carcinoma (LUSC) (Kandoth et al., 2013). A 2018 study analyzed over 9000 tumor samples across 33 different cancer types to identify cancer drivers; TAF1 was called in the analysis as a likely oncogenic driver of uterine corpus endometrial carcinoma (UCEC) (Bailey et al., 2018). In a more mechanistic study, increased TAF1 expression was shown to enhance androgen receptor (AR) transcriptional activity which is associated with the progression of prostate

cancer (Tavassoli et al., 2010). The Cancer Genome Atlas currently reports 475 TAF1 mutations found in 395 cancer cases, the most frequent mutation, R890C, being in the HAT domain. As TAF1 has been repeatedly shown to exert some control over cell cycle through its numerous domains and functions, it makes sense that mutations would be found in cancer.

1.3 TAF1 and p53

Tumor suppressor p53 is a transcription factor involved in activating a wide variety of genes involved in cell cycle regulation, DNA damage response, metabolism, and apoptosis (Kruse and Gu, 2009; Vousden and Prives, 2009). One mechanism by which p53 activates transcription of a target gene, p21, is by recruiting TAF1 and subsequently TFIID through interaction between acetylated lysines on p53 and the double bromodomain of TAF1 (Li et al., 2007). Diacetylation of p53 mediated by p300/CBP and PCAF in response to DNA damage facilitates recruitment of TAF1 in complex with a Holo-TFIID to the p21 promoter, activating transcription (Bode and Dong, 2004; Li et al., 2007). This activation can be switched off by another functional characteristic of TAF1, its kinase activity (Li et al., 2007). TAF1 phosphorylates p53 at thr55 which causes dissociation of both p53 and TAF1 from the p21 promoter, thus halting transcription (Li et al., 2007; Wu et al., 2014). Furthermore, it was shown that TAF1 mediated phosphorylation of p53 facilitates increased interaction with the

E3 ubiquitin ligase, MDM2, ultimately to leading to degradation of p53 (Li et al., 2004). Treatment with apigenin (a TAF1 kinase inhibitor) leads to a significant induction of p53 due to an increase in the unphosphorylated species which is less readily degraded (Liu et al., 2004).

TAF1 kinase is cellular ATP concentration dependent

Phosphorylation of p53 by TAF1 fluctuates with cellular ATP concentrations, which was shown under DNA damage conditions (Wu et al., 2014). Following DNA damage, PARP-1 is activated which results in a depletion of cellular ATP, and during this early stage of the DNA damage response, TAF1 kinase activity concurrently drops and p53 stabilizes on the p21 promoter, activating transcription and thus halting cell cycle (Carson et al., 1986; Wu et al., 2014). The ATP sensor and master metabolic regulator, AMPK, senses the increased AMP:ATP ratio and by the later stages of the DNA damage response, cellular ATP levels have recovered (Éthier et al., 2012). As the recovery of cellular ATP takes place, p53 thr55 phosphorylation returns to basal levels and cell cycle is allowed to resume as transcription of p21 is terminated (Figure 1.3; Wu et al., 2014). Using small molecules, 4AN and compound c, to inhibit PARP-1 and AMPK, respectively, cellular ATP levels can be stabilized resulting in little or no change to the status of TAF1-mediated p53 phosphorylation following DNA damage (Wu et al., 2014). TAF1 kinase was found to have a Km_{ATP} of 1.9mM,

which is inside the physiological range, making TAF1 kinase activity sensitive to fluctuations in cellular ATP (Wu et al., 2014). These results suggest TAF1 regulates cell cycle via p53 phosphorylation in response to changes to cellular ATP status (Wu et al., 2014).

1.4 E2F1

E2F1 is a transcription factor that plays a role in both activation and repression of transcription, and like p53 acts as a key regulator in cell cycle and the DNA damage response (Roworth et al., 2015). The E2F family consists of nine proteins, E2F1-8 (with two versions of 3; a,b) where 1-3a are considered to be activators of transcription while 3b-8 have been shown to act mainly as transcriptional repressors, though there are many exceptions to this categorization (Chong et al., 2009). The role of E2F1 in cell cycle has been well characterized however this protein's role in DNA damage response and in inducing apoptosis remains to be controversial and is a source of many conflicting reports. In G1, E2F1 protein levels are kept relatively low and it exists in a repressed state, bound to the pocket protein pRB, which binds to and sterically blocks the transactivation domain at the C-terminus of E2F1 (Flemington et al., 1993). Upon mitogenic signaling, Cyclin D becomes activated through the MYC and RAS pathways and subsequently associates with CDK 4/6 to phosphorylate pRB which lowers the binding affinity between pRB and E2F1

resulting in transcriptional activation of Cyclin E by E2F1 (Matsushime et al., 1992; Roworth et al., 2015). E2F1 itself also undergoes many post-translational modifications that result in enhanced activation including acetylation by P/CAF (Martínez-Balbás et al., 2000). Cyclin E associates with CDK2 which in turn facilitates hyperphosphorylation of pRB resulting in complete dissociation from E2F1 (Roworth et al., 2015). Since E2F1 upregulates its own transcription, as cells enter S-phase there is a rapid amplification of E2F1 protein levels as a result of the activated Cyclin E/CDK2 which in turn allows for further activation of Cyclin E and A transcription as well as many other genes required for DNA replication (Roworth et al., 2015). During S-phase E2F1 protein slowly returns to basal levels as Cyclin A binds and recruits SKP2 which facilitates E2F1 ubiquitination and subsequent degradation (Dubrez, 2017).

Under DNA damage conditions, E2F1 becomes stabilized and activated through phosphorylation by ATM/ATR and CHK2 (Lin et al., 2001; Stevens et al., 2003). Following these modifications, E2F1 transcriptionally activates a multitude of genes involved in repair pathways and apoptosis. E2F1 can induce apoptosis in a p53-dependent and p53-independent manner (Roworth et al., 2015). E2F1 activates p53 by upregulating expression of ARF which inhibits MDM2 thus stabilizing p53 (Berkovich et al., 2003). E2F1 also directly activates transcription of APAF1 and p73 as well as BH3-only proteins BIM and NOXA, all of which ultimately lead to apoptosis (Moroni et al., 2001; Roworth et al., 2015). TOPBP1 is a negative regulator or E2F1-induced apoptosis that binds to the N-terminal

region of E2F1 and recruits chromatin remodelers BRG1/BRM that work to shut down expression of E2F1 target genes (Liu et al., 2004). API5, a gene under E2F1 regulation acts as both an activator and repressor of E2F1. API5 is an apoptosis inhibitor that has been shown to repress E2F1-mediated apoptosis, interestingly, it has also been shown that API5 is required for E2F1-mediated activation of cell cycle genes (Morris et al., 2006; Navarro et al., 2013). In contrast to activating pro-apoptotic genes, E2F1 has also been shown to upregulate many genes involved in DNA repair such as RPA2, a ssDNA binding protein that localizes at sites of DNA damage (Chen et al., 2016). A highly tuned balance between E2F1, p53 and the factors that positively and negatively regulate them ultimately decide whether a cell will survive or not after DNA damage.

1.5 FOXM1

FOXM1 is one of roughly 100 FOX transcription factors that all share a common DNA binding domain, the forkhead box (sometimes called winged helix) domain (Liao et al., 2018). FOX proteins are characterized from FOXA to FOXR based on sequence homology, all playing unique roles in the cell (Laissue, 2019). ChIP-seq experiments revealed that FOXM1 uniquely binds to CHR elements (cell cycle genes homology region), whereas the majority of other FOX proteins are thought to bind to the RYAAAYA canonical forkead binding motif

(Chen et al., 2013). The binding of FOXM1 to CHR elements was then found to be reliant on interaction with the MMB activator complex (Chen et al., 2013). FOXM1 has been shown to be a key regulator in cell cycle, proliferation, and plays an important role in differentiation. Mutation, aberrant expression, and dysregulation of FOXM1 is linked to developmental defects and cancer; most significantly, overexpression of FOXM1a and b has been shown to drive progression of pancreatic cancer (Cui et al., 2014; Liao et al., 2018; Marchand et al., 2019).

FOXM1 is an 84 kDa protein that consists of an N-terminal repressor domain, the conserved forkhead DNA binding domain, and a C-terminal transactivation domain (Clark et al., 1993). There are 4 known isoforms, FOXM1a-d, as a result of alternative splicing that all have variations in the transactivation domain (The UniProt Consortium, 2019). FOXM1a shows no transactivation activity while all 3 other isoforms do (Kong at el., 2014). Direct interaction between the N-terminal repressor domain and the transactivation domain cause an autoinhibitory confirmation of the protein, which is relieved by phosphorylation by Cyclin-A/CDK complexes in G2 (Laoukili et al., 2008). FOXM1 expression is induced as cells enter S-phase and remains high through the rest of mitosis, activating a host of different target genes necessary for mitosis; two of the most well characterized targets being PLK1 and CCNB1 (Chai et al., 2018; Fu et al., 2010; Leung et al., 2001) . Expression of FOXM1 is tightly regulated by multiple transcription factors; transcriptionally activated by E2F1 and

suppressed by p53, among others (Millour et al., 2011). FOXM1 is subject to many posttranslational modifications including phosphorylation, acetylation, SUMOylation, and ubiquitination (Liao et al., 2018). Of the most prevalent PTMs made to FOXM1 is phosphorylation; multiple CDKS, PLK1, CHK2, and ERK1/2 have been shown to phosphorylate FOXM1at sites spanning the protein, some having activating effects and some repressing (Liao et al., 2018).

In the DNA damage response (DDR), FOXM1 plays an important role by activating transcription of many genes involved in the different mechanisms of DNA repair (Liao et al., 2018). In early stages of DNA damage there is an induction of FOXM1 via E2F1 driven transcriptional activation (Zona et al., 2014). An activated CHK2 phosphorylates and stabilizes FOXM1, followed by increased expression of FOXM1 target genes involved in DDR (Tan et al., 2007). It is necessary to downregulate FOXM1 once target DDR genes have been activated, which is achieved through p53-mediated repression of E2F1 (Barsotti and Prives, 2009; Millour et al., 2011). Downregulation of FOXM1 by p53 overpowers other activating signals like PTMs and direct E2F1 driven transcriptional activation, allowing for proper arrest of cell cycle, DNA damage repair, and apoptosis, as FOXM1 driven expression of cell cycle progression genes comes to a halt (Barsotti and Prives, 2009; Millour et al., 2011; Zona et al., 2014).

1.6 Summary of Chapters

In Chapter 2, I investigate whether TAF1 phosphorylates any transcription factors other than those already known. Given that TAF1 phosphorylates several GTFs as well as the cell cycle regulator, p53, we hypothesized there may be other sequence specific transcription factors that are targeted by TAF1 kinase as a mode of transcriptional regulation. Exploiting the fluctuations in cellular ATP following UV-induced DNA damage and the effect this has on TAF1 kinase activity, we employ bioinformatics to identify a list of potential phosphorylation targets and subsequently test them in vitro. We identify 2 novel targets for TAF1 phosphorylation and go on to show that these novel phosphorylation reactions effect transcriptional activity. This chapter is written as a manuscript to be submitted for publication.

Chapter 3 describes all the preparatory work that went into the Chapter 2 manuscript. Many proteins needed to be cloned and purified for these experiments using multiple expression systems (E. coli and insect cell); including many truncated and mutant proteins. Here I will describe the techniques used to generate all of these materials.

1.7 References

1. Andel III F. Three-Dimensional Structure of the Human TFIID-IIA-IIB Complex. Science. 1999 Dec 10;286(5447):2153–2156.

2. Barsotti AM, Prives C. Pro-proliferative FoxM1 is a target of p53-mediated repression. Oncogene. 2009 Dec;28(48):4295–4305.

3. Berkovich E, Lamed Y, Ginsberg D. E2F and Ras synergize in transcriptionally activating p14ARF expression. Cell Cycle. 2003 Apr;2(2):127–133.

4. Berkovich E, Lamed Y, Ginsberg D. E2F and Ras Synergize in Transcriptionally Activating p14 ^{ARF} Expression. Cell Cycle. 2003 Mar 7;2(2):127–134.

5. Bode AM, Dong Z. Post-translational modification of p53 in tumorigenesis. Nat Rev Cancer. 2004 Oct;4(10):793–805.

6. Boutet SC, Biressi S, Iori K, Natu V, Rando TA. Taf1 Regulates Pax3 Protein by Monoubiquitination in Skeletal Muscle Progenitors. Molecular Cell. 2010 Dec;40(5):749–761.

7. Buratowski S, Hahn S, Guarente L, Sharp PA. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell. 1989 Feb;56(4):549–561.

8. Burke TW, Kadonaga JT. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes & Development. 1996 Mar 15;10(6):711–724.

9. Carson DA, Seto S, Wasson DB, Carrera CJ. DNA strand breaks, NAD metabolism, and programmed cell death. Experimental Cell Research. 1986 Jun;164(2):273–281.

10. Chai N, Xie H, Yin J, Sa K, Guo Y, Wang M, et al. FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1. Biochemical and Biophysical Research Communications. 2018 Jun;500(4):924–929.

11. Chalkley GE. DNA binding site selection by RNA polymerase II TAFs: a TAFII250-TAFII150 complex recognizes the Initiator. The EMBO Journal.

1999 Sep 1;18(17):4835–4845.

12. Chen C-C, Juan C-W, Chen K-Y, Chang Y-C, Lee J-C, Chang M-C. Upregulation of RPA2 promotes NF-κB activation in breast cancer by relieving the antagonistic function of menin on NF-κB-regulated transcription. Carcinogenesis. 2017 Feb 1;38(2):196–206.

13. Chen J, Ghazawi FM, Li Q. Interplay of bromodomain and histone acetylation in the regulation of p300-dependent genes. Epigenetics. 2010 Aug 16;5(6):509–515.

14. Chen X, Muller GA, Quaas M, Fischer M, Han N, Stutchbury B, et al. The Forkhead Transcription Factor FOXM1 Controls Cell Cycle-Dependent Gene Expression through an Atypical Chromatin Binding Mechanism. Molecular and Cellular Biology. 2013 Jan 15;33(2):227–236.

15. Chong J-L, Wenzel PL, Sáenz-Robles MT, Nair V, Ferrey A, Hagan JP, et al. E2f1–3 switch from activators in progenitor cells to repressors in differentiating cells. Nature. 2009 Dec;462(7275):930–934.

16. Cianfrocco MA, Kassavetis GA, Grob P, Fang J, Juven-Gershon T, Kadonaga JT, et al. Human TFIID Binds to Core Promoter DNA in a Reorganized Structural State. Cell. 2013 Jan;152(1–2):120–131.

17. Clark KL, Halay ED, Lait E, Burley SK. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone HS. 1993;364:9.

18. Cohen S. The stimulation of epidermal proliferation by a specific protein (EGF). Developmental Biology. 1965 Dec 1;12(3):394–407.

19. Cohen S, Carpenter G, King L. Epidermal growth factor-receptor-protein kinase interactions. Co-purification of receptor and epidermal growth factor-enhanced phosphorylation activity. J. Biol. Chem. 1980 May 25;255(10):4834–4842.

20. Critchfield JW, Coligan JE, Folks TM, Butera ST. Casein kinase II is a selective target of HIV-1 transcriptional inhibitors. Proceedings of the National Academy of Sciences. 1997 Jun 10;94(12):6110–6115.

21. Cui J, Shi M, Xie D, Wei D, Jia Z, Zheng S, et al. FOXM1 Promotes the Warburg Effect and Pancreatic Cancer Progression via Transactivation of LDHA Expression. Clinical Cancer Research. 2014 May 15;20(10):2595–2606.

22. Curran EC, Wang H, Hinds TR, Zheng N, Wang EH. Zinc knuckle of TAF1

is a DNA binding module critical for TFIID promoter occupancy. Sci Rep. 2018 Dec;8(1):4630.

23. Dikstein R, Ruppert S, Tjian R. TAFII250 Is a Bipartite Protein Kinase That Phosphorylates the Basal Transcription Factor RAP74. Cell. 1996 Mar;84(5):781–790.

24. Dubrez L. Regulation of E2F1 Transcription Factor by Ubiquitin Conjugation. IJMS. 2017 Oct 19;18(10):2188.

25. Durant M, Pugh BF. Genome-Wide Relationships between TAF1 and Histone Acetyltransferases in Saccharomyces cerevisiae. MCB. 2006 Apr 1;26(7):2791–2802.

26. Éthier C, Tardif M, Arul L, Poirier GG. PARP-1 Modulation of mTOR Signaling in Response to a DNA Alkylating Agent. PLoS ONE. 2012 Oct 24;7(10):e47978.

27. Flemington EK, Speck SH, Kaelin WG. E2F-1-mediated transactivation is inhibited by complex formation with the retinoblastoma susceptibility gene product. Proceedings of the National Academy of Sciences. 1993 Aug 1;90(15):6914–6918.

28. Flores O, Lu H, Reinberg D. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. J. Biol. Chem. 1992 Feb 5;267(4):2786–2793.

29. Flores O, Maldonado E, Reinberg D. Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. J. Biol. Chem. 1989 May 25;264(15):8913–8921.

30. Fu Z, Malureanu L, Huang J, Wang W, Li H, van Deursen JM, et al. Plk1-dependent phosphorylation of FoxM1 regulates a transcriptional programme required for mitotic progression. Nat Cell Biol. 2008 Sep;10(9):1076–1082.

31. Garcia-Jove Navarro M, Basset C, Arcondéguy T, Touriol C, Perez G, Prats H, et al. Api5 Contributes to E2F1 Control of the G1/S Cell Cycle Phase Transition. PLoS ONE. 2013 Aug 7;8(8):e71443.

32. Ge H, Martinez E, Chiang C-M, Roeder RG. Activator-dependent transcription by mammalian RNA polymerase II: In vitro reconstitution with general transcription factors and cofactors Internet.. In: Methods in Enzymology. Academic Press; 1996 cited 2020 Sep 27.. p. 57–

71.Available from: http://www.sciencedirect.com/science/article/pii/S0076687996740089

33. Gegonne A, Weissman JD, Singer DS. TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. Proceedings of the National Academy of Sciences. 2001 Oct 23;98(22):12432–12437.

34. Gegonne A, Weissman JD, Zhou M, Brady JN, Singer DS. TAF7: A possible transcription initiation check-point regulator. Proceedings of the National Academy of Sciences. 2006 Jan 17;103(3):602–607.

35. Herr AJ. RNA Polymerase IV Directs Silencing of Endogenous DNA. Science. 2005 Apr 1;308(5718):118–120.

36. Hilton TL, Li Y, Dunphy EL, Wang EH. TAF1 Histone Acetyltransferase Activity in Sp1 Activation of the Cyclin D1 Promoter. MCB. 2005 May 15;25(10):4321–4332.

37. Imhof A, Yang X-J, Ogryzko VV, Nakatani Y, Wolffe AP, Ge H. Acetylation of general transcription factors by histone acetyltransferases. Current Biology. 1997 Sep;7(9):689–692.

38. Ito N, Hendriks WT, Dhakal J, Vaine CA, Liu C, Shin D, et al. Decreased N-TAF1 expression in X-linked dystonia-parkinsonism patient-specific neural stem cells. Dis. Model. Mech. 2016 Apr 1;9(4):451–462.

39. Jacobson RH. Structure and Function of a Human TAFII250 Double Bromodomain Module. Science. 2000 May 26;288(5470):1422–1425.

40. Jacobson RH. Structure and Function of a Human TAFII250 Double Bromodomain Module. Science. 2000 May 26;288(5470):1422–1425.

41. Kaestner KH. Unified nomenclature for the winged helix/forkhead transcription factors. :6.

42. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013 Oct 17;502(7471):333–339.

43. Kim Y-J, Björklund S, Li Y, Sayre MH, Kornberg RD. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell. 1994 May;77(4):599–608.

44. Kloet SL, Whiting JL, Gafken P, Ranish J, Wang EH. Phosphorylation-Dependent Regulation of Cyclin D1 and Cyclin A Gene Transcription by

TFIID Subunits TAF1 and TAF7. Molecular and Cellular Biology. 2012 Aug 15;32(16):3358–3369.

45. Kokubo T, Gong D-W, ROEDERt RG, HORIKOSHItt M, Nakatani Yoshihir. The Drosophila 11O-kDa transcription factor TFIID subunit directly interacts with the N-terminal region of the 230-kDa subunit. :5.

46. Koleske AJ, Young RA. An RNA polymerase II holoenzyme responsive to activators. Nature. 1994 Mar;368(6470):466–469.

47. Kolesnikova O, Ben-Shem A, Luo J, Ranish J, Schultz P, Papai G. Molecular structure of promoter-bound yeast TFIID. Nat Commun. 2018 Dec;9(1):4666.

48. Kong X, Li L, Li Z, Le X, Huang C, Jia Z, et al. Dysregulated Expression of FOXM1 Isoforms Drives Progression of Pancreatic Cancer. Cancer Res. 2013 Jul 1;73(13):3987–3996.

49. Kruse J-P, Gu W. Modes of p53 Regulation. Cell. 2009 May;137(4):609–622.

50. Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes & Development. 1998 Jan 1;12(1):34–44.

51. Laissue P. The forkhead-box family of transcription factors: key molecular players in colorectal cancer pathogenesis. Mol Cancer. 2019 Dec;18(1):5.

52. Laoukili J, Alvarez M, Meijer LAT, Stahl M, Mohammed S, Kleij L, et al. Activation of FoxM1 during G2 Requires Cyclin A/Cdk-Dependent Relief of Autorepression by the FoxM1 N-Terminal Domain. MCB. 2008 May 1;28(9):3076–3087.

53. Leung TWC, Lin SSW, Tsang ACC, Tong CSW, Ching JCY, Leung WY, et al. Over-expression of FoxM1 stimulates cyclin B1 expression. FEBS Letters. 2001 Oct 19;507(1):59–66.

54. Li AG, Piluso LG, Cai X, Gadd BJ, Ladurner AG, Liu X. An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment. Molecular Cell. 2007 Nov;28(3):408–421.

55. Li H-H, Li AG, Sheppard HM, Liu X. Phosphorylation on Thr-55 by TAF1 Mediates Degradation of p53: A Role for TAF1 in Cell G1 Progression. Molecular Cell. :12.

56. Liao G-B, Li X-Z, Zeng S, Liu C, Yang S-M, Yang L, et al. Regulation of the master regulator FOXM1 in cancer. Cell Commun Signal. 2018 Dec;16(1):57.

57. Lin W-C, Lin F-T, Nevins JR. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. :12.

58. Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, et al. Solution Structure of a TBP–TAFII230 Complex: Protein Mimicry of the Minor Groove Surface of the TATA Box Unwound by TBP. :11.

59. Liu K. TopBP1 recruits Brg1/Brm to repress E2F1-induced apoptosis, a novel pRb-independent and E2F1-specific control for cell survival. Genes & Development. 2004 Mar 15;18(6):673–686.

60. Liu K, Luo Y, Lin F-T, Lin W-C. TopBP1 recruits Brg1/Brm to repress E2F1-induced apoptosis, a novel pRb-independent and E2F1-specific control for cell survival. Genes Dev. 2004 Mar 15;18(6):673–686.

61. Marchand B, Pitarresi JR, Reichert M, Suzuki K, Laczkó D, Rustgi AK. PRRX1 isoforms cooperate with FOXM1 to regulate the DNA damage response in pancreatic cancer cells. Oncogene. 2019 May;38(22):4325–4339.

62. Martínez-Balbás MA, Bauer U-M, Nielsen SJ, Brehm A, Kouzarides T. Regulation of E2F1 activity by acetylation. EMBO J. 2000 Feb 15;19(4):662–671.

63. MatsuiS T, Segallg J, Weill PA, Roederll R. Multiple Factors Required for Accurate Initiation of Transcriptibony Purified RNA Polymerase 11. :5.

64. Matsushime H, Ewen ME, Strom DK, Kato JY, Hanks SK, Roussel MF, et al. Identification and properties of an atypical catalytic subunit (p34PSK-J3/cdk4) for mammalian D type G1 cyclins. Cell. 1992 Oct 16;71(2):323–334.

65. Millour J, de Olano N, Horimoto Y, Monteiro LJ, Langer JK, Aligue R, et al. ATM and p53 Regulate FOXM1 Expression via E2F in Breast Cancer Epirubicin Treatment and Resistance. Molecular Cancer Therapeutics. 2011 Jun 1;10(6):1046–1058.

66. Mizzen CA, Yang X-J, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, et al. The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity. Cell. 1996 Dec;87(7):1261–1270.

67. Moroni MC, Hickman ES, Denchi EL, Caprara G, Colli E, Cecconi F, et al. Apaf-1 is a transcriptional target for E2F and p53. Nature Cell Biology. 2001 Jun;3(6):552–558.

68. Moroni MC, Hickman ES, Denchi EL, Caprara G, Colli E, Cecconi F, et al. Apaf-1 is a transcriptional target for E2F and p53. Nat Cell Biol. 2001 Jun;3(6):552–558.

69. Morris EJ, Michaud WA, Ji J-Y, Moon N-S, Rocco JW, Dyson NJ. Functional identification of Api5 as a suppressor of E2F-dependent apoptosis in vivo. PLoS Genet. 2006 Nov 17;2(11):e196.

70. Morris EJ, Michaud WA, Ji J-Y, Moon N-S, Rocco JW, Dyson NJ. Functional Identification of Api5 as a Suppressor of E2F-Dependent Apoptosis In Vivo. PLoS Genet. 2006;2(11):e196.

71. Navarro MG-J, Basset C, Arcondéguy T, Touriol C, Perez G, Prats H, et al. Api5 Contributes to E2F1 Control of the G1/S Cell Cycle Phase Transition. PLOS ONE. 2013 Aug 7;8(8):e71443.

72. O'Brien T, Tjian R. Functional Analysis of the Human TAFII250 N-Terminal Kinase Domain. Molecular Cell. 1998 May;1(6):905–911.

73. Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS. Plant Nuclear RNA Polymerase IV Mediates siRNA and DNA Methylation-Dependent Heterochromatin Formation. Cell. 2005 Mar;120(5):613–622.

74. O'Rawe JA, Wu Y, Dörfel MJ, Rope AF, Au PYB, Parboosingh JS, et al. TAF1 Variants Are Associated with Dysmorphic Features, Intellectual Disability, and Neurological Manifestations. The American Journal of Human Genetics. 2015 Dec;97(6):922–932.

75. Patel AB, Louder RK, Greber BJ, Grünberg S, Luo J, Fang J, et al. Structure of human TFIID and mechanism of TBP loading onto promoter DNA. Science. 2018 Dec 21;362(6421):eaau8872.

76. Pham A-D. Ubiquitin-Activating/Conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila. Science. 2000 Sep 29;289(5488):2357–2360.

77. Poppy Roworth A, Ghari F, La Thangue NB. To live or let die – complexity within the E2F1 pathway. Molecular & Cellular Oncology. 2015 Jan 2;2(1):e970480.

78. Pugh BF, Tjian R. Transcription from a TATA-less promoter requires a

multisubunit TFIID complex. :12.

79. Roeder RG, Rutter WJ. Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms. Nature. 1969 Oct;224(5216):234–237.

80. Roeder RG, Rutter WJ. Specific Nucleolar and Nucleoplasmic RNA Polymerases. Proceedings of the National Academy of Sciences. 1970 Mar 1;65(3):675–682.

81. Samuels M, Sharp PA. Separation and Characterizationof Factors Mediating Accurate Transcription by RNA Polymerase 11. Transcription Factors. :9.

82. Sawa C, Nedea E, Krogan N, Wada T, Handa H, Greenblatt J, et al. Bromodomain Factor 1 (Bdf1) Is Phosphorylated by Protein Kinase CK2. MCB. 2004 Jun 1;24(11):4734–4742.

83. Sawadogo M, Roeder RG. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. Proceedings of the National Academy of Sciences. 1985 Jul 1;82(13):4394–4398.

84. Seger R, Krebs EG. The MAPK signaling cascade. FASEB j. 1995 Jun;9(9):726–735.

85. Smale ST, Kadonaga JT. The RNA Polymerase II Core Promoter. Annu. Rev. Biochem. 2003 Jun;72(1):449–479.

86. Stevens C, Smith L, La Thangue NB. Chk2 activates E2F-1 in response to DNA damage. Nat Cell Biol. 2003 May;5(5):401–409.

87. Tavassoli P, Wafa LA, Cheng H, Zoubeidi A, Fazli L, Gleave M, et al. TAF1 Differentially Enhances Androgen Receptor Transcriptional Activity via Its N-Terminal Kinase and Ubiquitin-Activating and -Conjugating Domains. Molecular Endocrinology. 2010 Apr 1;24(4):696–708.

88. Thomas MC, Chiang C-M. The general transcription machinery and general cofactors. Crit. Rev. Biochem. Mol. Biol. 2006 Jun;41(3):105–178.

89. Thomas MC, Chiang C-M. The General Transcription Machinery and General Cofactors. Critical Reviews in Biochemistry and Molecular Biology. 2006 Jan;41(3):105–178.

90. Van Dyke MW, Sawadogo M, Roeder RG. Stability of transcription

complexes on class II genes. Mol. Cell. Biol. 1989 Jan;9(1):342–344.

91. Vousden KH, Prives C. Blinded by the Light: The Growing Complexity of p53. Cell. 2009 May;137(3):413–431.

92. Wang H, Curran EC, Hinds TR, Wang EH, Zheng N. Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. Cell Res. 2014 Dec;24(12):1433–1444.

93. Weil PA, Segall J, Harris B, Ng SY, Roeder RG. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. J. Biol. Chem. 1979 Jul 10;254(13):6163–6173.

94. Weiss SB, Gladstone L. A MAMMALIAN SYSTEM FOR THE INCORPORATION OF CYTIDINE TRIPHOSPHATE INTO RIBONUCLEIC ACID [1]. J. Am. Chem. Soc. 1959 Aug;81(15):4118–4119.

95. Wu S-Y, Chiang C-M. Properties of PC4 and an RNA Polymerase II Complex in Directing Activated and Basal Transcription *in Vitro*. J. Biol. Chem. 1998 May 15;273(20):12492–12498.

96. Wu S-Y, Thomas MC, Hou SY, Likhite V, Chiang C-M. Isolation of Mouse TFIID and Functional Characterization of TBP and TFIID in Mediating Estrogen Receptor and Chromatin Transcription. J. Biol. Chem. 1999 Aug 13;274(33):23480–23490.

97. Wu Y, Lin JC, Piluso LG, Dhahbi JM, Bobadilla S, Spindler SR, et al. Phosphorylation of p53 by TAF1 Inactivates p53-Dependent Transcription in the DNA Damage Response. Molecular Cell. 2014 Jan;53(1):63–74.

98. Zona S, Bella L, Burton MJ, Nestal de Moraes G, Lam EW-F. FOXM1: An emerging master regulator of DNA damage response and genotoxic agent resistance. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2014 Nov;1839(11):1316–1322.

1.8 Tables and Figures

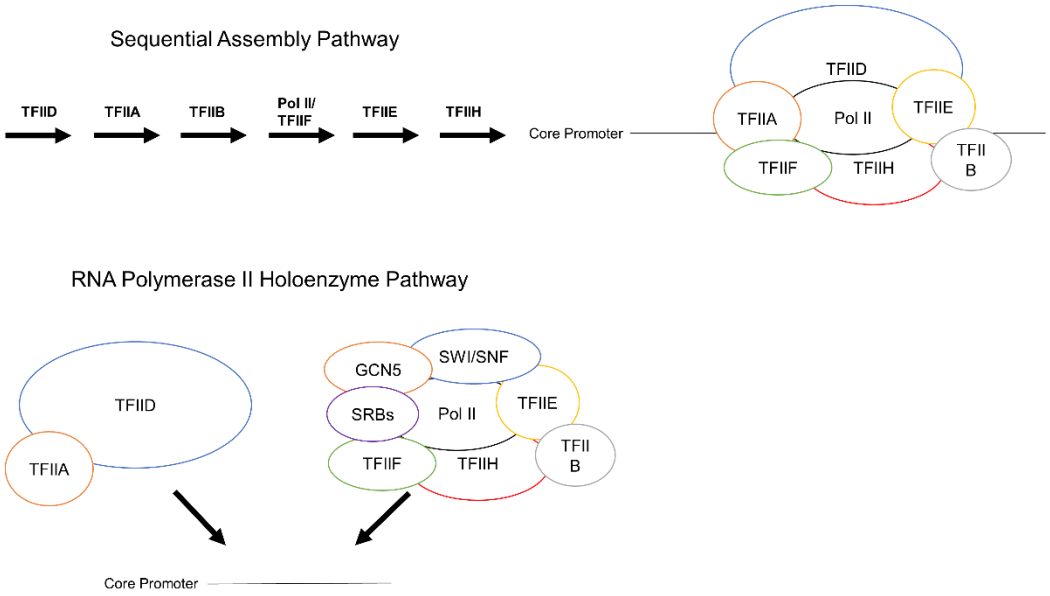


Figure 1.1 Formation of the preinitiation complex Schematic representation of sequential assembly pathway (top) and holoenzyme pathway (bottom) for forming the transcription preinitiation complex. (Thomas and Chiang, 2006).

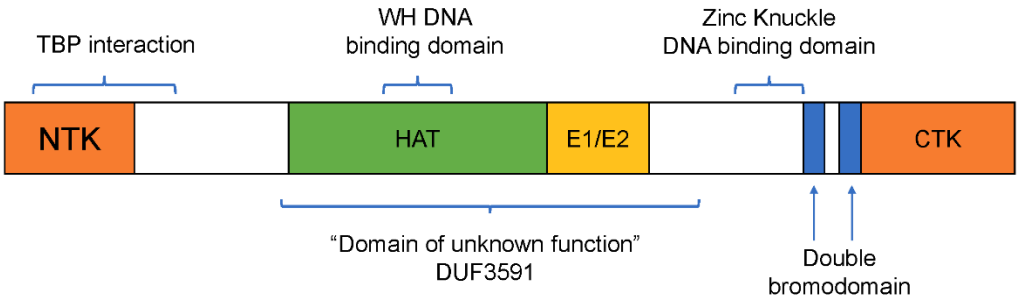


Figure 1.2 Domains of TAF1 Abbreviations are as follows: N-terminal kinase domain (NTK), C-terminal kinase domain (CTK), histone acetyltransferase domain, E1/E2 ubiquitin activating/conjugating domain (E1/E2).

	BLCA	BRCA	COAD	GMB	HNSC	KRC	AML	LUAD	LUSC	OV	UCEC	Pan-Cancer
VHL	0	0	0	0	0	52	0	0	0.6	0	0.9	6.9
GATA3	1	10.6	1	0	2	0	0	2.6	2.9	0.3	0.4	3.2
TSHZ3	2	0.7	3.1	0.7	1.3	1.2	0.5	14.9	6.3	1	3.9	2.6
EP300	17.4	0.8	2.1	0.3	8	1.4	0	0.9	4.6	0.3	5.2	2.5
CTCF	2	2.4	1.6	0	3.3	0.5	0.5	1.3	0	0.3	16.5	2.4
TAF1	2	1.1	1.6	1.4	2.3	1.2	0	4	6.9	1.9	8.7	2.3
TSHZ2	4.1	0.9	3.1	2.4	1.3	0.7	0	6.6	3.5	1	1.7	1.8
RUNX1	1	3.3	1	0	0.7	0	9	0.4	0	0	1.3	1.6
MECOM	5.1	0.5	1	1.4	1.7	1	0	3.5	4.6	0.6	3	1.5
TBX3	3.1	2.4	1	0	0.7	0	0	4.4	2.9	1	1.3	1.4

Table 1.1 10 most commonly mutated transcription factors across 11 cancer types From Kandoth et al., 2013: displays the percent of samples found to have mutations in each cancer type. Pan-Cancer refers to a statistical analysis that aims to compare similarities and differences across multiple tumor types. Abbreviated cancer types are: bladder urothelial carcinoma (BLCA), breast adenocarcinoma (BRCA), colon and rectal carcinoma (COAD, READ), glioblastoma multiforme (GMB), head and neck squamous cell carcinoma (HNSC), kidney renal clear cell carcinoma (KRC), acute myeloid leukaemia (AML), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), ovarian serous carcinoma (OV), and uterine corpus endometrial carcinoma (UCEC).

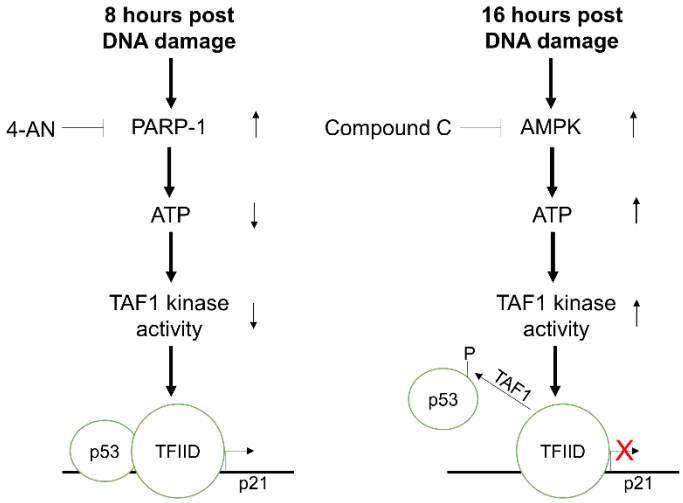


Figure 1.3 p53 regulation by TAF1 kinase Model for TAF1-mediated regulation of p53 in response to fluctuations in cellular ATP following DNA damage as described by Wu et al., 2014.

Chapter 2: Expression and Purification of TAF1, CDK8, RAP74, E2F1, FOXM1, ATF2 and SP1

2.1 Introduction

The work that will be described in Chapter 3 required protein purification using a variety of different cloning methods, expression systems, and purification techniques. While using E. coli to express proteins is sufficient in many cases to generate properly folded and functional molecules, complex human proteins often require chaperones and posttranslational modifications in order to achieve functionality, the machinery for which is not found in prokaryotic systems (Gräslund et al., 2008). A popular solution to this problem is using an insect cell expression system, where these eukaryotic cells can be cultured in suspension at large volumes to generate similar amounts of protein as can be achieved with E. coli. Large proteins, such as TAF1 (250kDa), are difficult to express in E. coli. and furthermore, (as will be described in this chapter), truncations of the N- and C-terminal kinase domains (NTK and CTK) purified from E. coli do not retain any activity. This is presumably because they require some eukaryotic translational or posttranslational machinery. To generate full length TAF1, as well as truncated NTK and CTK proteins, SF9 insect cells were used, as many other groups have done to study the enzymatic activity of TAF1 in vitro (Dikstein et al., 1996; Li et al., 2004; O'Brien and Tjian, 1998). Unlike with E. coli, where the gene of interest is inserted into the cells via plasmid transformation, baculovirus is generally used as the method of delivering recombinant DNA into insect cells. In this chapter I will describe the expression and purification of full length TAF1, truncated NTK

and CTK domains, and a kinase dead TAF1 using a baculovirus expression system.

To study TAF1 kinase activity in the experiments presented in Chapter 3, we also needed to generate the substrates to be tested; E2F1, FOXM1, ATF2, Sp1, and RAP74, which was done using E. coli. When applicable, bacterial expression of proteins is the most ideal method because it's not only cost efficient, but these cells are easy to culture and manipulate for large scale preparation with optimal expression. Generally, there is some optimization that must be done for bacterial protein expression including how long expression is induced for, when expression is induced, and the temperature at which the cells are grown during induction. For instance, inducing expression of the protein of interest at the normal growing temperature of 37°C will result in very rapid expression. However, this often results in problems with degradation, solubility, and folding. In this chapter I will describe the expression of these transcription factors in E. coli and purification using a variety of affinity chromatography techniques and our efforts to overcome the issues presented above. The data described in the next chapter show that TAF1 phosphorylates E2F1 and FOXM1. We also mapped the phosphorylation site on E2F1. This involved cloning, expressing, and purifying many truncations and mutant E2F1 proteins which will also be described in this chapter.

2.2 Results

Purification of full-length wildtype TAF1

To purify TAF1 from insect cells, we first needed to make a recombinant baculovirus housing the TAF1 gene that includes a tag for affinity purification. To ensure the protocols were working, a baculovirus capable of expressing green fluorescent protein (GFP) as a positive control was also generated. Through PCR cloning, a FLAG tag was added to the N-terminus of TAF1 and the recombinant gene was ligated into a baculovirus donor vector, a plasmid that has the viral polyhedrin promoter (Figure 2.1A). This donor plasmid also includes the right and left arms of the Tn7 transposon which flank the gene of interest (and promoter), this facilitates transposition of the gene into the baculoviral genome. This is an engineered technology that is now commonly used for inserting foreign DNA into the baculovirus genome (Luckow et al., 1993). The GFP gene was cloned separately into a donor vector. Following construction, the donor plasmid was transformed into commercially obtained DH10Bac E. coli cells. These cells house a large circular piece of DNA called a bacmid (~136kb) that includes the entire baculoviral genome with the *attTn7* target transposition site where the FLAG-TAF1 (or GFP) gene will be inserted (*Bac-to-Bac Baculovirus Expression System Manual,* 2015). These transformed cells were plated on agar containing X-gal and underwent blue/white color screening. A special feature of the bacmid in the DH10bac cells is that it is engineered so that successful transposition of the gene of interest into the bacmid leads to disruption of the lacZα gene causing the

colonies to remain white; whereas unaltered bacmid-containing cells grow as blue colonies (*Bac-to-Bac Baculovirus Expression System Manual,* 2015). White colonies were isolated and grown in liquid culture, followed by purification of the recombinant bacmid DNA using a midi-prep kit specific for large DNA molecules. The bacmid DNA constructs, now containing the FLAG-TAF1 gene or GFP gene, were then transfected into SF9 insect cells grown in monolayer culture using the cationic polymer, PEI, which allows the DNA to be endocytosed into the cells (Boussif et al., 1995; Sonawane et al., 2003). After transfection, the insect cells, now containing the full viral genome, produced a first-generation virus which buds out from the cells and into the growth media. Cells transfected with the GFP-containing bacmid were monitored using a fluorescent microscope to track progression of the viral production process by viewing expression of GFP (Figure 2.1B). We found that the transfection process was not very efficient, with maybe ~5% of the cells expressing GFP at 48 hours post transfection, which is not surprising given the size of the bacmid. However, as the virus buds out of the successfully transfected cells, it infects neighboring cells which propagates throughout the plate. By 96 hours post-transfection, nearly 100% of cells were expressing GFP. Additionally, the TAF1 baculovirus infected cells showed similar morphology to the GFP infected cells, indicating the TAF1 virus was also being produced. The virus containing media was collected and used to infect additional cells to amplify the virus to achieve a higher viral titer. To verify that the FLAG-TAF1 virus was functioning properly, some infected cells were lysed and protein

was analyzed via Western Blot, revealing robust expression of TAF1. At this point, the GFP baculovirus was no longer needed.

Expression conditions were optimized through time and virus titrations to find the best conditions with minimal protein degradation but high expression. We found that high concentrations of virus or increased duration of infection resulted in increased TAF1 degradation or cell death. The best expression conditions were found to be a ~1:50 dilution of the virus into the growth media and harvesting after 48 hours (Figure 2.1C). To use these conditions on a large scale, SF9 insect cells were adapted to suspension culture and grown up to 0.5L and infected at a cell density of ~3 million cells/mL followed by harvesting 48 hours later. Previous work in our lab found that TAF1 (a nuclear localized protein) has better activity if extracted from the nucleus of the insect cells as opposed to whole cell lysate. Presumably, this is because any incorrectly folded or degraded TAF1 protein does not get transported into the nucleus. Nuclei from infected cells were isolated using a Dounce homogenizer followed by high salt extraction to release nuclear proteins.

The purification of FLAG-TAF1 extracted from the nuclei of infected SF9 cells was performed using commercial anti-FLAG antibody conjugated agarose beads using a batch method (as opposed to a column). FLAG-TAF1-containing nuclear extract was incubated with the beads for 3 hours and then collected by centrifugation and washed 3 times each with high and low salt wash buffers. Following the wash step, the beads were incubated with a small volume of a

glycerol-containing storage buffer that had a high concentration of FLAG peptide, to elute the protein. The purified FLAG-TAF1 was analyzed for purity and concentration by SDS-PAGE and silver staining, using a standard curve of BSA on the gel (Figure 2.1D). It was found that the elution contained roughly 50ng/ul of FLAG-TAF1 with some minor degradation products or contaminant proteins giving rise to some fainter bands lower on the gel. As will be described in Chapter 3, we found this purified TAF1 did have kinase activity, and was able to autophosphorylate and trans-phosphorylate target proteins.

Purification of kinase dead TAF1

Since wildtype human TAF1 likely interacts with many proteins inside insect cells due to homology of the transcription machinery throughout eukaryotes, it is important to have a kinase dead mutant TAF1 as a negative control when using this expression system. This allows verification that any phosphorylation seen in a kinase assay is due to TAF1 and not some co-purified contaminating kinase. A previous publication describes a TAF1 mutant that contains 11 point mutations in the NTK, rendering the domain inactive though still allowing incorporation into TFIID (O'Brien and Tjian, 1998). These mutations occur clustered in two stretches in the NTK described as NT1 and N7, where multiple D, E, K, C, and S amino acids are mutated to A (Figure 2.2B) (O'Brien and Tjian, 1998). Our kinase dead TAF1 was made using this NTK mutant in

addition to deletion of the CTK, since the active site in that domain is unknown (Figure 2.2B). This construct, which we already had in our lab from previous experiments, was used to create a baculovirus to express kinase dead TAF1. Since the majority of this mutant TAF1 is identical to the wildtype, and can be still be incorporated in TFIID, we reasoned that it would likely co-purify with the same contaminants as the wildtype TAF1 from insect cells, an ideal negative control.

In the same way the baculovirus was made to express wildtype TAF1, we prepared a baculovirus for production of the kinase dead TAF1. The mutant gene was PCR cloned to include an N-terminal FLAG tag and ligated into the pFastBac1 vector (Figure 2.2A). Once the baculovirus was made, expression optimization was performed with the same time and dilution titrations as were done with the wildtype. We determined that the same conditions used for the wildtype would be suited for the mutant protein. After large scale infection, FLAG-kinase-dead TAF1 was extracted from the nuclei and purified in the same fashion as the wildtype, resulting in roughly the same purity and concentration (Figure 2.2C). As will be described in the next chapter, this kinase dead TAF1 was unable to phosphorylate any target proteins.

Purification of the N- and C-terminal kinase domains of TAF1

Because we are primarily interested in TAF1 kinase activity, truncated kinase domains (NTK and CTK, both ~50kDa) were purified. We first purified

these two truncations using bacteria. The NTK and CTK truncated genes (TAF1 amino acids 1-434 and 1425-1893, respectively) were cloned into the pET22b expression vector, which includes a C-terminal 6X His tag (Figure 2.3A,B). Expression tests using BL21 E. coli cells revealed the CTK is more efficiently expressed than the NTK which showed a high degree of degradation. This was improved by expressing the NTK at a low temperature,18°C in contrast to the 30°C initially used. Large scale cultures were grown and lysed via sonication, and batch purifications were carried out using nickel beads, eluting the proteins with imidazole (Figure 2.3C). As was expected from the expression tests, the CTK was purified to a higher concentration than the NTK. The two purified proteins were tested for kinase activity where we found neither were capable of auto- or transphosphorylation. This is most likely due to a lack of some required posttranslational modifications or chaperone proteins needed for the NTK and CTK to fold properly.

Since the attempt at applying a bacterial expression method for purifying the kinase domains of TAF1 lead to inactive protein, the insect cell expression system was deemed necessary. This method has been used to produce active NTK and CTK truncations in the literature as well in our lab previously (Dikstein et al., 1996). In the same way that the wildtype TAF1 and kinase dead TAF1 were generated (described in the previous sections of this chapter), baculoviruses were made to express HA tagged NTK and FLAG tagged CTK (Figure 2.3D and 2.3E). After baculoviruses were made and tested for

expression, large scale volumes of SF9 insect cells grown in suspension were infected and HA-NTK and FLAG-CTK protein was extracted from nuclei. To purify the NTK in batch, Anti-HA antibody was conjugated to Protein A agarose beads, by chemical crosslinking using dimethyl pimelimidate. The HA-NTK was eluted from the beads using HA peptide. CTK was purified in batch using anti-FLAG agarose beads in the same way as the wildtype TAF1 (Figure 2.3F). In agreement with the literature, both truncations showed autophosphorylation activity, and were able to phosphorylate target proteins (Figure 2.3F). These truncated TAF1 proteins were to be used for a project that became outside of the scope of this dissertation but can be used (along with the baculoviruses to make more protein) by future students.

Purification of RAP74

RAP74, a 74kDa protein, is a subunit of TFIIF and a known phosphorylation target of TAF1 (Dikstein et al., 1996; Flores et al., 1989). The next chapter of this dissertation describes an investigation of TAF1 kinase, aiming to identify novel targets. For those experiments, a positive control was needed to verify the purified TAF1 was functional, thus RAP74 was used for this purpose. Because this protein has previously been used in our lab to study TAF1 kinase activity, a bacterial expression plasmid containing a His-tagged RAP74 was available to use. His-RAP74 was expressed in BL21 E. coli cells at 30°C for

5 hours and lysed via sonication. Batch purification was carried out by incubating clarified lysate with Ni-NTA resin for 1 hour which was collected via centrifugation followed by 3 high salt and 3 low salt washed. His-RAP74 was eluted into a glycerol-containing storage buffer with 250mM imidazole (Figure 2.4).

Purification of E2F1

E2F1 is a 47kDa protein, consisting of a Cyclin A binding domain, a DNA binding domain, heptad repeat and marked box domains (which together make up its dimerization domain), and a transactivation domain (Cress and Nevins, 1994; Jost et al., 1996; Trouche and Kouzarides, 1996; Xu, et al., 1994; Zheng et al., 1999). A bacterial expression plasmid, pGEX, containing the E2F1 gene with an N-terminal GST tag was obtained (Figure 2.5A). GST-E2F1 was expressed in BL21 E. coli cells, lysed by sonication. Batch purification was carried out by incubating clarified lysate with glutathione agarose beads followed by 3 high salt and 3 low salt washes. GST-E2F1 was eluted from the beads using 20mM reduced glutathione in a 20% glycerol-containing storage buffer (Figure 2.5D). This purified protein was used in the experiments presented in Chapter 3.

Purification of 4 E2F1 truncations and 7 mutants for phosphorylation site mapping

To map the phosphorylation site(s) on E2F1, multiple E2F1 constructs including truncations and single amino acid mutants were purified. Truncations consisting of E2F1's four major functional domains were cloned via PCR into a bacterial expression vector, pGEX-6P-1, to be purified through a N-terminal GST-tag as was done with the full-length protein (Figure 2.5B). The truncations, amino acids 1-108, 109-198, 199-361, and 362-437 (the Cyclin A binding domain, the DNA binding domain, the dimerization domain, and the transactivation domain, respectively), were expressed in BL21 E. coli cells under the same conditions as the full length E2F1, and purified in the same way using glutathione agarose beads (Figure 2.5E). The only truncated protein to have some degradation problems was the transactivation domain, likely because of intrinsic disorder, something common in transcription factors' transactivation domains (Sigler, 1988). These purified E2F1 domain truncations were used in the phosphorylation site mapping experiments that will be discussed in the next chapter.

For mapping the phosphorylation site inside the transactivation domain of E2F1 (described in Chapter 3), we used phosphorylation deficient mutants to investigate each of the 7 possible sites (six serine and one threonine, Figure 2.5C). To do this, we utilized site-directed mutagenesis, a process that involves PCR amplifying an entire plasmid that contains the gene of interest using primers that contain the desired mutation. These PCR reactions were carried out using

the pGEX-6P-1 plasmid containing the E2F1 transactivation domain truncation as the template; seven different reactions were run with each primer set specific for mutating one of the seven possible phosphorylation sites to alanine. After the reactions were carried out, the resulting samples were digested with the methylation-sensitive restriction endonuclease, DpnI, to remove any remaining wildtype template plasmid, which is methylated (compared to the unmethylated mutant plasmid produced in the PCR). Finally, the DpnI digested PCR reactions were transformed into E. coli cells and the resulting colonies were screened for plasmids containing the correct mutation. Once all seven plasmids were confirmed to have the desired mutations (S364A, S375A, S382A, S392A, S401A, S403A, and T433A), the mutant proteins were all purified. All seven mutant proteins were expressed in BL21 E. coli under the same conditions as the wildtype truncation and purified in the same way using glutathione beads (Figure 2.5F). The resulting purified mutant E2F1 transactivation domain truncations were used for determining the site of phosphorylation by TAF1 which will be presented in Chapter 3.

Purification of the CDK8 kinase module

The phosphorylation site we mapped for E2F1 has previously been described as a CDK8 target site. In order to make sure we were not seeing a signal from co-purified CDK8 in our TAF1, we carried out a CDK inhibition assay

which required purified CDK8. This kinase requires two additional proteins to achieve activity, Cyclin C, and MED12. Baculoviruses for expression of FLAG-CDK8 and Cyclin C, as well as a bacterial expression vector for GST-MED12 were provided by Tom Boyer. The baculovirus stocks were amplified with insect cells and tested for expression. Insect cells grown in suspension were co-infected with both the FLAG-CDK8 and Cyclin C viruses followed by co-purification using anti-FLAG beads. (data shown in Figure 3.6 in the following Chapter). GST-MED12 was expressed in BL21 E. coli cells and purified using glutathione agarose beads. The co-purified CDK8 and Cyclin C and purified MED12 were combined and used for a kinase assay described in Chapter 3 where we found this CDK8 complex was able to autophosphorylate and transphosphorylate target protein.

Purification of FOXM1

A FOXM1 bacterial expression vector containing an N-terminal His-tag was purchased and expression tests were carried out (Figure 2.6A). FOXM1 did not express well in E. coli, there was little induction of the protein upon addition of IPTG, and what amount of protein that expressed was heavily degraded and largely insoluble. Though a range of induction temperatures and times were tested (from 18°C to 37°C), this had little effect on improving the expression of FOXM1. To resolve the solubility issue, different extractions buffers were tried,

varying in pH and addition of different non-ionic detergents and glycerol. We were able to modestly improve solubility and homogeneity of the protein, however, there still remained a very prominent degradation product at a higher concentration than the full-length protein (Figure 2.6B). This was used in preliminary kinase assays and we found FOXM1 was a target of TAF1-mediated phosphorylation (described in Chapter 3). However, because the degradation product was phosphorylated in our kinase assays, this reduced the signal for the full-length (data not shown). To continue, we needed to improve the purification so that the phosphorylation results would be clearer.

Next, we tried to express FOXM1 in insect cells, which involved generating a baculovirus to express FOXM1 using the same processes described previously in this chapter (Figure 2.6C). We found that FOXM1 did express much better in insect cells but instead of having degradation problems like we had in E. coli, we found that FOXM1 co-purified with some unknown kinase (Figure 2.6D). In testing the insect cell expressed FOXM1 in kinase assays, it showed very strong phosphorylation signal when there was no purified TAF1 added, owing to whatever co-purified with FOXM1 (data not shown). Additional rounds of purification were attempted but did little to reduce this strong background signal. While insect cell expressed purified FOXM1 had little degradation, it was not useful for our kinase assays due to high background signal.

Our last attempt to purify FOXM1, that would have minimal degradation and no contaminant kinases, was to truncate the FOXM1 protein, deleting the transactivation domain, and expressing it in E. coli (Figure 2.6E). As previously mentioned, transactivation domains tend to have disordered regions and can lead to expression and degradation difficulties in E. coli. This truncated FOXM1 construct was cloned and we found that it did express much better in E. coli compared to the full-length protein (Figure 2.6F). Since there still remained a small degree of degradation (though less than with the full-length construct), we added a second affinity tag, GST, to the C-terminus of the protein. Through this method, running tandem purifications, one for the N-terminal His-tag, followed by one for the C-terminal GST-tag, we were able to capture full length protein with much less degraded FOXM1. The purification of the His-ΔTAD-FOXM1-GST protein construct was carried out by both nickel affinity and glutathione affinity chromatography, and the resulting protein retained significantly less degradation products (Figure 2.6F). This was the protein used in the kinase assay that will be presented in Chapter 3.

Purification of ATF2

Another protein we identified as a potential TAF1 kinase target was ATF2. A pGEX bacterial expression vector containing the ATF2 gene with an N-terminal GST tag was obtained as a gift from a colleague and used to express the protein

in BL21 E. coli cells. ATF2 expressed very robustly and was highly soluble. Due to the large amount of protein obtained after growing a 0.5L culture, the purification was carried out using a gravity column (as opposed to the batch method used for the previously described purifications), to make it easier to capture all of the protein. A 2mL glutathione agarose column was prepared, and the purification was carried out at 4°C. Multiple elution fractions were collected, the third having the highest concentration of ATF2, this fraction was used for experiments in Chapter 3 (Figure 2.7).

Purification of Sp1

Sp1 was the fourth transcription factor that we tested for phosphorylation by TAF1. To purify this protein a pGEX bacterial expression vector containing the Sp1 gene with an N-terminal GST tag was obtained from a colleague. This protein was expressed in BL21 E. coli first at 30°C for ~5 hours but we found this resulted in prominent degradation products. It was then expressed at 18°C overnight which helped considerably with the degradation, though some remained. Sp1 was purified in batch using glutathione agarose beads (Figure 2.8). This purified protein was used in kinase assays as described in Chapter 3.

2.3 Discussion

There are many methods to express human proteins without using human cells, making it more accessible to study them as culturing human cells is more time consuming, costly, and challenging, especially when trying to generate a large amount of protein. E. coli expression systems are very useful and easy to work with, having a workflow from cloning to protein expression that can be done in a matter of days. However, for certain proteins, these cells are not sufficient for generating functional molecules. Insect cell expression systems are a powerful tool for generating hard-to-express proteins or those that require posttranslational modifications. The tools that are now commercially available for generating recombinant baculovirus make it more user friendly to use insect cell systems. One of the most common early methods of generating recombinant baculovirus consisted of co-transfecting insect cells with wildtype baculovirus and a transfer plasmid containing the gene of interest (Miller, 1988). The gene would then be incorporated into the viral genome via homologous recombination, however, this happened at very low efficiencies (<1%) and then required multiple rounds of plaque purification to isolate recombinant virus (Miller, 1988). This was a very time and labor-intensive process but revealed this expression system to be powerful in generating large amounts of complex proteins. Now, the site-specific transposition method as described in this chapter, is commercially available and relatively cheap, making this expression system much easier and faster to work with. In this chapter I've described the process of purifying wildtype human TAF1,

as well as a kinase dead mutant, and kinase domain truncations that retained proper activity using insect cells. This process was part of the preparatory work that went into the experiments described in Chapter 3.

Also described in this chapter was the purification of all the substrates needed for kinase assays that will be described in Chapter 3; RAP74, E2F1, FOXM1, ATF2, and Sp1. As phosphorylation data was collected and we found E2F1 and FOXM1 to be targets of TAF1 kinase, more proteins required purification for further investigation. Our original FOXM1 purifications contained degradation products, and since we saw phosphorylation of this protein (and the degradation product) we wanted to improve our FOXM1 expression method. After trying insect cells for FOXM1 expression and finding it co-purified with very active kinases that interfered with our TAF1 mediated phosphorylation results, we tried deleting a disordered domain, which improved our degradation issue in E. coli. The final improvement our purified FOXM1 was to include two separate affinity tags on either end of the protein and running tandem purifications to isolate the full-length protein. Additionally, we needed to generate multiple E2F1 truncations and then mutants for phosphorylation site mapping. This was an extremely involved process involving many rounds of cloning, including site-directed mutagenesis, and subsequently many rounds of purification. This was all successful and we were able to map the phosphorylation site (described in Chapter 3).

The purification methods described in this chapter included multiple affinity-based methods; GST fusion proteins, His-tags for nickel affinity chromatography, and FLAG and HA epitope tagging for antibody-based affinity purification. Many of these experiments were carried out using the batch method of protein purification which is easy and ideal for small amounts of protein, though ATF2 was purified using a column due to high amounts of protein.

2.4 Materials and Methods

Baculovirus protein expression and purification

Recombinant baculovirus for the expression of wildtype TAF1, kinase-dead TAF1, NTK, CTK, and FOXM1 were generated according to Gibco's *Bac-to-Bac Baculovirus Expression System USER GUIDE* using the pFastBac1 transfer plasmid (Gibco) and DH10bac E. coli cells (Gibco). Transfections of bacmid DNA were performed with PEI transfection reagent. SF9 insect cells were grown at 27°C either in monolayer or suspension culture using EX-CELL 420 serum free growth media. Protein expression conditions were optimized for each baculovirus. For expression of CDK8 and Cyclin C, baculovirus stocks were gifted to us from Thomas Boyer's lab at the University of Texas Health Science Center, San Antonio.

Nuclear extracts for baculovirus expressed FLAG-TAF1, FLAG-kinase-dead TAF1, FLAG-CTK, HA-NTK, and His-FOXM1 were prepared as described

previously. FLAG purifications were carried out using ANTI-FLAG M2 Agarose Affinity Gel (Sigma); nuclear extracts were incubated with beads for 3 hours and collected via centrifugation followed by 3 low salt washes, 3 high salt washes, and elutions using 1 mg/ml FLAG peptide (Sigma). Proteins were stored at -80°C in 20% glycerol. HA purifications were carried out in the same way but using anti-HA (12CA5) conjugated beads and elutions used 1 mg/ml HA peptide (Sigma). His-tag purifications were carried out in the same way but with Ni-NTA Resin (Thermo) and elutions carried out with 250mM Imidazole (Sigma).

Primers used for cloning into pFastBac1:

FLAG-TAF1 (wildtype): 5’ – GCGGGATCCACCATGGACTACAAAGACGATGACGACAAGATGGGACCCGG CTGCGATTT – 3’ AND 5’ – ATAAGAATGCGGCCGCTCATTCATCAGAGTCCAAGT – 3’

FLAG-kinase-dead-TAF1: 5’ – GCGGGATCCACCATGGACTACAAAGACGATGACGACAAGATGGGACCCGG CTGCGATTT – 3 and 5’ – ATAAGAATGCGGCCGCTCAAGCTGCTTCTTTAGCAGTAC – 3’

HA-NTK: 5’ -
GCGGGATCCACCATGTACCCATACGATGTTCCAGATTACGCTGGACCCGGC
TGCGATTTGCT – 3’ and 5’ –
TTTATAGCGGCCGCATCATCCTCCCAATGCAGCT – 3’

FLAG-CTK: 5’ –
GCGGGATCCACCATGGACTACAAAGACGATGACGACAAGTTCCACACTCCA
GTCAATGC – 3’ and 5’-
ATAAGAATGCGGCCGCTCATTCATCAGAGTCCAAGT – 3’

His-FOXM1: 5’ –
CGCGGATCCATGCGCGGATCCATGAAAACTAGCCCCCGTCG – 3’ and 5’ –
ATAAGAATGCGGCCGCCTAATGATGATGATGATGATGCTGTAGCTCAGGAA
TAAACT – 3’

Bacterial protein expression and purification

His-RAP74, GST-E2F1 (full length, 4 truncations, and 7 mutants), wildtype His-FOXM1, His-ΔTAD-FOXM1-GST, GST-ATF2, GST-Sp1, NTK-His and CTK-His, and MED12 were all expressed using BL21 DE3 E. coli cells. IPTG (GoldBio) was used to induce expression, OD at time of induction was optimized

for each protein as well as length of induction. All proteins were extracted using a 10% glycerol, 1% Triton X-100 (Sigma) Tris buffer and sonicated at 20% amplitude for ~ 4 minutes. Cell debris was cleared via high speed centrifugation. Purifications were either carried out in batch or via gravity column using either Glutathione Agarose (Pierce) or Ni-NTA Resin (Thermo), with low salt and high salt washes followed by elution with either 20mM L-Glutathione reduced (Sigma) or 250mM Imidazole (Sigma).

Plasmids used for bacterial expression:

His-RAP74 plasmid was a gift from A.J. Berk

HA-TAF1 pCMV plasmid was a gift from R. Tjian

E2F-1 wt-pGex2TK was a gift from William Kaelin (Addgene plasmid # 21668 ; http://n2t.net/addgene:21668 ; RRID:Addgene_21668)

pET-6xHis/hFOXM1 was purchased from VectorBuilder (ID: VB160702-1008wrc)

Sp1-pGEX plasmid was a gift from Thomas Gilmore

MED12-pGEX plasmid was a gift from Tom Boyer.

Primers used for cloning into pGEX-6P-1:

E2F1 aa 1-108: 5’ - CGCGGATCC ATGGCCTTGGCCGGGGCCCCT – 3’ and 5’ – ATAAGAATGCGGCCGCTCAAGCTGGCCCACTGCTCT – 3’

E2F1 aa 109-198: 5’ – CGCGGATCCCGGGGCAGAGGCCGCCATCC – 3’ and 5’ – ATAAGAATGCGGCCGCTCAGCCCACTGTGGTGTGGCTGC – 3’

E2F1 aa 199-361: 5’ – CGCGGATCCGTCGGCGGACGGCTTGAGGG – 3’ and 5’ – ATAAGAATGCGGCCGCTCACCGGGACAACAGCGGTTCTT – 3’

E2F1 aa 362-447: 5’ – CGCGGATCCGGCAGCCTGCGGGCTCCCGT – 3’ and 5’ – ATAAGAATGCGGCCGCTCAGAAATCCAGGGGGGTGA – 3’

FOXM1-His: 5’ – CGCGGATCCATGAAAACTAGCCCCCGTCG – 3’ and 5’ – ATAAGAATGCGGCCGCCTAATGATGATGATGATGATGCTGTAGCTCAGGAA TAAACT – 3’

Site-directed mutagenesis

Site-directed mutagenesis for creating the 7 Ser/Thr to Ala E2F1 mutant transactivation domain truncations was performed according to *QuickChange II Site-Directed Mutagenesis Instruction Manual* (Agilent). PCR reactions were carried out using Pfu polymerase (agilent) and digested with DpnI (NEB) followed by transformation into DH5alpha cells. Plasmid was extracted from colonies and mutations were confirmed via sanger sequencing.

Primers used for site-directed mutagenesis:

S364A: 5’ – TTGTCCCGGATGGGCGCCCTGCGGGCTCCCGTG – 3’ and 5’ – CACGGGAGCCCGCAGGGCGCCCATCCGGGACAA – 3’

S375A: 5’ – GACGAGGACCGCCTGGCCCCGCTGGTGGCGGCC – 3’ and 5’ – GGCCGCCACCAGCGGGGCCAGGCGGTCCTCGTC – 3’

S382A: CTGGTGGCGGCCGACGCGCTCCTGGAGCATGTG – 3’ and 5’ – CACATGCTCCAGGAGCGCGTCGGCCGCCACCAG – 3’

S392A: 5’ – GTGCGGGAGGACTTCGCCGGCCTCCTCCCTGAGGAGTT – 3’ and 5’ – CTCAGGGAGGAGGCCGGCGAAGTCCTCCCGCACATGCT – 3’

S401A: 5’ – CCTGAGGAGTTCATCGCCCTTTCCCCACCCCAC – 3’ and 5’ – GTGGGGTGGGGAAAGGGCGATGAACTCCTCAGG – 3’

S403A: 5’ – GAGTTCATCAGCCTTGCCCCACCCCACGAGGCC – 3’ and 5’ – GGCCTCGTGGGGTGGGGCAAGGCTGATGAACTC – 3’

T433A: 5’ – GACTTTGGGGACCTCGCCCCCCTGGATTTCTGA – 3’ and 5’ – TCAGAAATCCAGGGGGGCGAGGTCCCCAAAGTC – 3’

2.5 References

1. Boussif O, Lezoualc'h F, Zanta MA, Mergny MD, Scherman D, Demeneix B, et al. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proc Natl Acad Sci U S A. 1995 Aug 1;92(16):7297–7301.

2. Cress WD, Nevins JR. Interacting domains of E2F1, DP1, and the adenovirus E4 protein. J Virol. 1994 Jul;68(7):4213–4219.

3. Dikstein R, Ruppert S, Tjian R. TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. Cell. 1996 Mar 8;84(5):781–790.

4. Esposito D, Chatterjee DK. Enhancement of soluble protein expression through the use of fusion tags. Current Opinion in Biotechnology. 2006 Aug 1;17(4):353–358.

5. Flores O, Maldonado E, Reinberg D. Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. The Journal of Biological Chemistry. 1989 May 25;264(15):8913–8921.

6. Jost CA, Ginsberg D, Kaelin WG. A conserved region of unknown function participates in the recognition of E2F family members by the adenovirus E4 ORF 6/7 protein. Virology. 1996 Jun 1;220(1):78–90.

7. Kaelin WG, Krek W, Sellers WR, DeCaprio JA, Ajchenbaum F, Fuchs CS, et al. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. Cell. 1992 Jul 24;70(2):351–364.

8. Li H-H, Li AG, Sheppard HM, Liu X. Phosphorylation on Thr-55 by TAF1 Mediates Degradation of p53: A Role for TAF1 in Cell G1 Progression. Molecular Cell. 2004 Mar 26;13(6):867–878.

9. Luckow VA, Lee SC, Barry GF, Olins PO. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. J Virol. 1993 Aug;67(8):4566–4579.

10. Miller LK. Baculoviruses as Gene Expression Vectors. Annual Review of Microbiology. 1988;42(1):177–199.

11. O'Brien T, Tjian R. Functional analysis of the human TAFII250 N-terminal

kinase domain. Mol Cell. 1998 May;1(6):905–911.

12. Sigler PB. Acid blobs and negative noodles. Nature. 1988 May;333(6170):210–212.

13. Sonawane ND, Szoka FC, Verkman AS. Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. J. Biol. Chem. 2003 Nov 7;278(45):44826–44831.

14. Structural Genomics Consortium, China Structural Genomics Consortium, Northeast Structural Genomics Consortium, Gräslund S, Nordlund P, Weigelt J, et al. Protein production and purification. Nat Methods. 2008 Feb;5(2):135–146.

15. Trouche D, Kouzarides T. E2F1 and E1A(12S) have a homologous activation domain regulated by RB and CBP. Proc Natl Acad Sci U S A. 1996 Feb 20;93(4):1439–1442.

16. Xu M, Sheppard KA, Peng CY, Yee AS, Piwnica-Worms H. Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding activity of E2F-1/DP-1 by phosphorylation. Molecular and Cellular Biology. 1994 Dec;14(12):8420–8431.

17. Zheng N, Fraenkel E, Pabo CO, Pavletich NP. Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F–DP. Genes Dev. 1999 Mar 15;13(6):666–674.

18. Bac-to-Bac Baculovirus Expression System Manual. Invitrogen. 2015. http://tools.thermofisher.com/content/sfs/manuals/bactobac_man.pdf

2.6 Figures

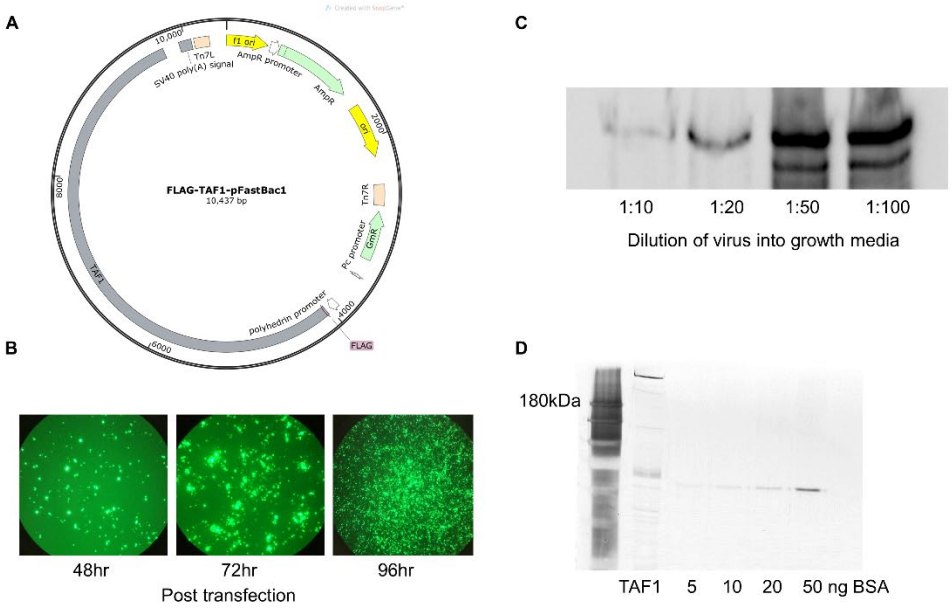


Figure 2.1 Purification of TAF1 using baculovirus. (A) Plasmid map of FLAG-TAF1 in pFastBac1 baculovirus donor vector. (B) Propagation of control GFP baculovirus at time points following transfection of the donor plasmid, done alongside transfection of TAF1 baculovirus. (C) Western blot of dilution titrations using amplified TAF1 baculovirus; 1:10 and 1:20 dilutions killed virtually all the cells within 24 hours. (D) Silver stained gel showing purified FLAG-TAF1 next to BSA standards for concentration determination.

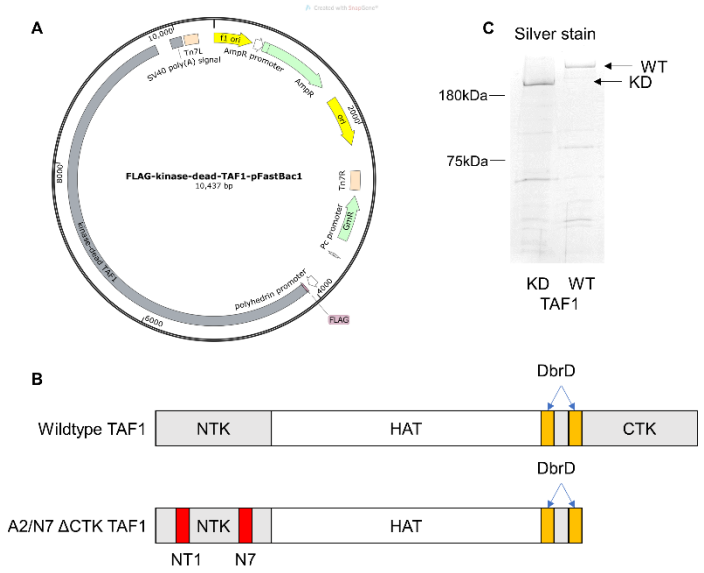


Figure 2.2 Purification of kinase dead A2/N7 ΔCTK TAF1. (A) Plasmid map of kinase dead TAF1 cloned into pFastBac1 baculovirus donor vector. (B) Schematic of wildtype versus kinase dead (A2/N7 ΔCTK) TAF1. Region labeled NT1 contains stretch of mutations: 4 D to S , 1 E to A, and 1 K to A. N7 contains second stretch of mutations: 1 C to A, 1 S to A, 2 D to A, and 1 E to K. (C) Silver stained gel showing purified kinase dead TAF1 next to wildtype.

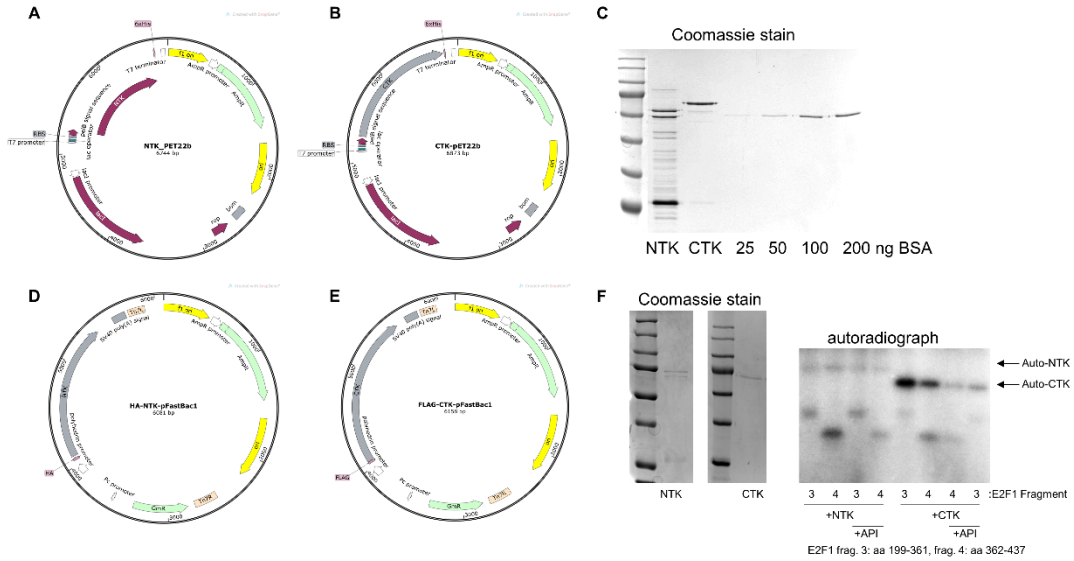


Figure 2.3 Purification of TAF1 NTK and CTK. (A and B) pET22b plasmid maps of bacterial expression vectors for expression of His-tagged NTK (A) and CTK (B). (C) Coomassie stained gel of purified bacterial expressed NTK and CTK with BSA standards. (D and E) pFastBac1 maps of donor plasmids for generating baculovirus for insect cell expression of HA-NTK and FLAG-CTK. (F) Coomassie stained gel of purified NTK and CTK from insect cells. Activity was assayed via in vitro kinase assay using different truncations of E2F1.

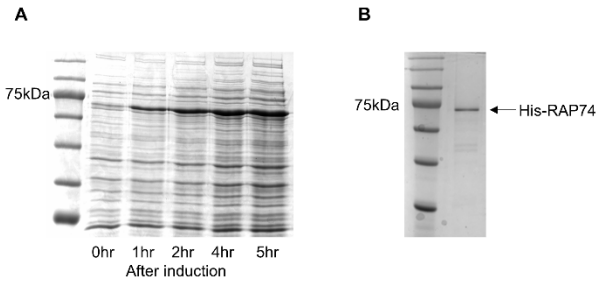


Figure 2.4 Purification of RAP74. (A) Expression tests for RAP74 in BL21 E. coli cells carried out at 30°C for 5 hours, time points indicated are hours after adding IPTG. (B) RAP74 elution from Ni purification.

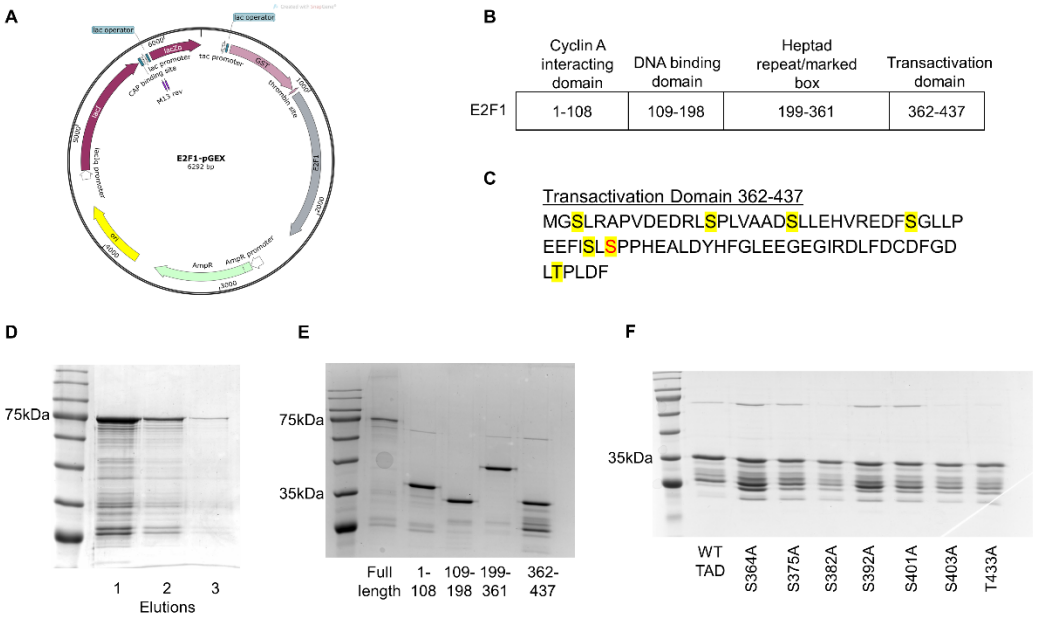


Figure 2.5 Purification of E2F1: full length, truncations, and mutants. (A) Map of full length E2F1-pGEX plasmid (Addgene plasmid #21668) used for expressing E2F1 and for subcloning all truncation and mutations. (B) Schematic of domains of E2F1 that were each separately purified. (C) Amino acid sequence of transactivation domain of E2F1 with possible phosphorylation sites highlighted. S375 is in red, as this was the phosphorylation site found to be specific for TAF1 kinase. (D) Purification of full length E2F1. (E) Purification of truncated domains of E2F1, including full length E2F1 on the gel. (F) Purification of E2F1 transactivation domain mutants after site-directed mutagenesis was performed, shown next to wildtype transactivation domain truncation.

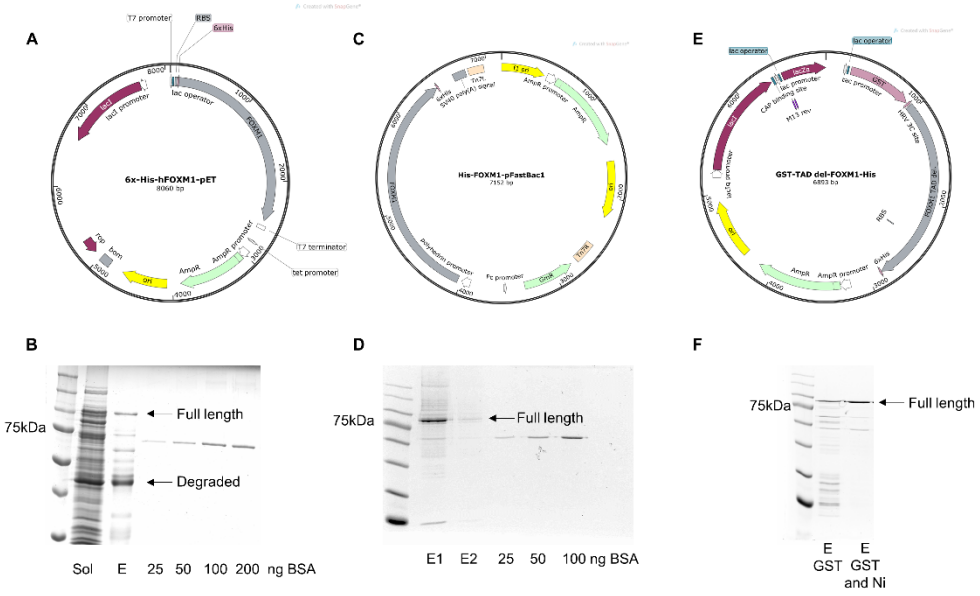


Figure 2.6 Purification of bacterial and insect cell expressed FOXM1. (A) Map of His-FOXM1-pET (full length) plasmid for bacterial expression. (B) Purification of His-FOXM1 from bacteria, shows soluble protein from cell extract, elution and BSA standards for concentration determination. (C) Map of His-FOXM1-pFastBac1 for baculovirus generation for insect cell expression of full length His-FOXM1. (D) Purification of His-FOXM1 from insect cell whole cell extract, two elutions shown. (E) pGEX-6P plasmid map with truncated FOXM1 (transactivation domain deleted), includes N-terminal GST tag and C-terminal His tag, for bacterial expression. (F) Purification of truncated FOXM1, first elution shown after GST purification only, second is from GST followed by Ni purification.

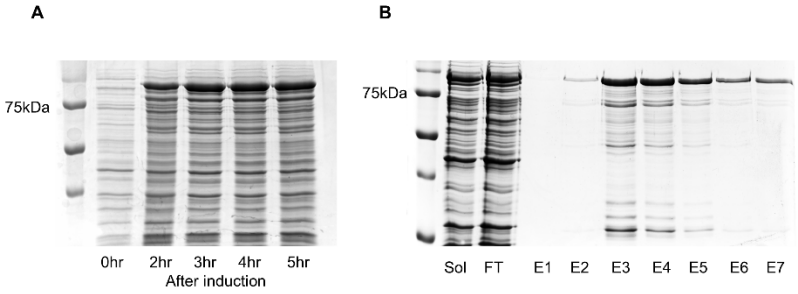


Figure 2.7 Purification of ATF2. (A) Expression test of ATF2 in E. coli, showing time points following addition of IPTG. (B) Purification of GST-ATF2 using a 2ml glutathione gravity column, showing soluble protein fraction from cell extract, flow through from column, and 7 elutions.

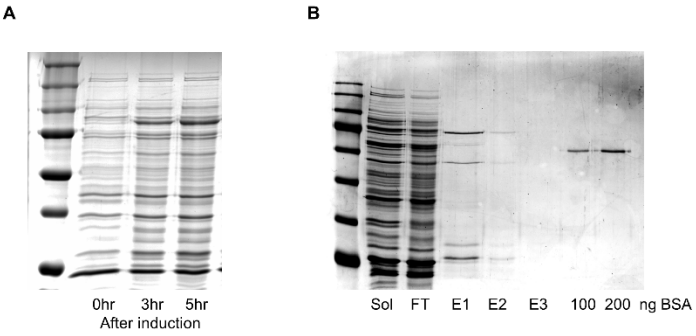


Figure 2.8 Purification of Sp1. (A) Expression test for Sp1 in E. coli cells checked at 0, 3, and 5 hours post induction. (B) Purification of GST-Sp1, showing soluble fraction from cell extract, flow through, 3 elutions, and BSA standards for concentration comparison.

Chapter 3: TAF1 regulates transcriptional activity of E2F1 and FOXM1 through phosphorylation

Chapter 3 is intended to become a publication. Work presented in this chapter is a collaboration of all co-authors for the publication. Dr. Joy Lin is responsible for the TAF1 ChIP-Seq which provided the data shown in Figure 3.1, Figure 3.2, and Table 3.1. Dr. Lily Maxham performed the bioinformatic analysis shown in Figure 3.1A, B, and E, Figure 3.2, and Table 3.1. Dr. Lily Maxham performed the TAF1, p53, FOXM1, Sp1, and ATF2 ChIP assays shown in Figure 3.3A and the FOXM1 ChIP assays shown in figure 3.4C

3.1 Abstract

TAF1 is the largest subunit of general transcription factor TFIID and possesses intrinsic protein kinase activity. We previously reported that TAF1 phosphorylates p53 on the p21 promoter, leading to dissociation of p53 and TAF1 from the promoter and inactivation of transcription. Because TAF1 binds to many promoters, we investigated whether it regulates other enhancer-bound transcription factors in a similar manner. ChIP-seq analysis revealed a list of promoters that TAF1 differentially binds under different kinase conditions. Through bioinformatics analysis, we identified a set of candidate transcription factors, including E2F1, FOXM1, Sp1 and ATF2. Upon performing in vitro phosphorylation, we found that TAF1 directly phosphorylates E2F1 and FOXM1, two transcription factors that are essential in cell cycle regulation. Furthermore, we mapped the site of E2F phosphorylation to Ser375 and show that mutation of the phosphorylation site affects E2F1 DNA binding and transcription. These data shed light on the dual role of TAF1 as both a general transcription factor involved in genome-wide regulation of transcription activation, and a protein kinase involved in negative regulation of specific transcription.

3.2 Introduction

Transcription activation is carefully orchestrated in order to maintain proper control of all cellular processes. It requires the assembly of the basic transcription machinery at the promoter which consists of enhancer-binding transcription factors, coactivators, and general transcription factors (GTFs) as well as the RNA polymerase II (Pol II) enzyme. As the first GTF assembled on the promoter, TFIID is a 1.3MDa complex consisting of the TATA-box binding protein (TBP) and 13 TBP associated factors (TAFs; Burley and Roeder, 1996; Tora, 2002). The largest TAF, TAF1, is a multifunctional protein enzyme that possesses a histone acetyl transferase (HAT) domain, an E1/E2 ubiquitin activating and conjugating (UBAC), a double bromodomain, multiple DNA binding domains, and N-terminal and C-terminal kinase domains (Dikstein et al., 1996; Jacobson et al., 2000; Mizzen et al., 1996; Pham and Sauer, 2000; Curran et al., 2018). TAF1 plays a critical role in transcription initiation with the ability to recognize and bind core promoter facilitating recruitment of other components of the basic transcription machinery (reviewed in Thomas and Chiang, 2006). Once all necessary factors are recruited to a promoter, transcription of the DNA template from a specific start site occurs rapidly and with high accuracy.

While TAF1's role in transcription initiation has long been studied, the mechanisms by which it modulates transcription via its kinase activity are less clear. Our previous findings revealed that, upon DNA damage, TAF1 is recruited to the p21 promoter through interaction with acetylated p53, activating

transcription and subsequently halting cell cycle (Li et al., 2007). However, TAF1 also turned off p21 transcription in the later stages of DNA damage by phosphorylating p53 on the promoter, causing dissociation of p53 and TAF1 from the promoter (Wu et al., 2014). Furthermore, we showed that, upon DNA damage, the ability of TAF1 to both positively and negatively regulate p53-mediated p21 transcription is dependent of its kinase activity (Wu et al., 2014). At early times of DNA damage, poly (ADP-ribose) polymerase-1 (PARP-1) activation causes a depletion of cellular ATP (Schreiber et al 2006), which inhibits TAF1 kinase activity and p53 phosphorylation. This enables TAF1 to support p53 mediated p21 transcription (Wu et al., 2014). Reduction of cellular ATP levels leads to activation of AMP-activated protein kinase (AMPK). Once activated, AMPK regulates several metabolic responses (Hardie, 2007) that recover ATP levels at later times of DNA damage. The ATP recovery increases the kinase activity of TAF1, which leads to p53 phosphorylation, and dissociation of p53 and TAF1 from the promoter (Wu et al., 2014).

As a subunit of TFIID, TAF1 is bound to many promoters (Kim et al, 2005) and interacts with a multitude of transcription factors, leading us to hypothesize that there may be others, like p53, that are regulated by TAF1-mediated phosphorylation in a similar manner. To search for new TAF1 phosphorylation targets, we employed ChIP-seq to identify promoters that TAF1 differentially binds under different kinase conditions. Through bioinformatics analysis, we identified a set of candidate transcription factors that bind to those promoters.

Upon testing phosphorylation of several transcription factor candidates from our analysis, we reveal that TAF1 phosphorylates both E2F1 and FOXM1. We also found that TAF1-mediated phosphorylation negatively regulates their transcriptional activity. Given both E2F1 and FOXM1 are involved in cell cycle regulation and, like p53, are dysregulated in many cancers, this study provides new insights into the role of TAF1 protein kinase in regulation of specific transcription.

3.3 Results

Identification of 421 promoters that show kinase dependent TAF1 occupancy

In an effort to find new phosphorylation targets of TAF1, we first determined whether the previously reported phosphorylation dependent TAF1 binding on the p21 promoter (Wu et al., 2014) occurs in a broader scale. TAF1 ChIP-seq was performed to assess TAF1's overall binding to the promoter in response to UV irradiation under two TAF1 kinase conditions: 8 hours after UV while TAF1 kinase activity is reduced due to ATP depletion, as well as 16 hours after UV while TAF1 kinase activity is recovered (Wu et al, 2014). To ensure the kinase activity plays a role in the promoter binding, we also treated cells with TAF1 kinase inhibitor apigenin at 16 hours after UV. Analysis of the sequencing data revealed TAF1 genome-wide binding is increased at 8 hours after UV, which

is followed by a decrease in binding at 16 hours after UV (Figure 3.1A). Furthermore, the decrease at 16 hours was rescued when recovery of TAF1 kinase activity was blocked by apigenin (Figure 3.1A). Utilizing the TAF1 ChIP-seq datasets available in the ENCODE database, we showed that the TAF1 peaks identified in our analyses overlap at least sixty percent in 6 other cell lines (Figure 3.1E).

An analysis of TAF1 peaks that showed significant increase from 0 to 8 hour and significant decrease from 8 to 16 hour after UV leads to 2352 resulting peaks within 3 kb of a transcription start site (Figure 3.1B). To assess the effect of TAF1 kinase activity, these peaks were then intersected with peaks that rescued by the TAF1 kinase inhibitor apigenin. This resulted in 634 rescuing peaks (Figure 3.1B). Since apigenin, which is a naturally occurring flavonoid compound, is known to inhibit other protein kinases, we also performed ChIP-seq, with or without apigenin, at 8 hours after UV while TAF1 kinase activity is already low to remove potential off-target effects (Figure 3.2). This approach filtered out 159 peaks form the previously identified 634 peaks (Figure 3.1B). Using ChIPpeakAnno, we found the remaining 475 peaks represented 421 unique promoters (Figure 3.1B). Genome browser view of TAF1 binding to RPA2 and API5, two representative promoters, illustrated TAF1 binding increases from 0 to 8 hour, reduces from 8 to 16 hours and is rescued by apigenin at 16 after UV (Figure 3.1D). A pathway enrichment analysis was performed on the 421 promoters, which revealed cell cycle and metabolism as the top two pathways

regulated by TAF1 kinase activity (Figure 3.1C). These results suggest that, upon UV, TAF1's binding to the promoter is selectively modulated by its kinase activity.

Identification of target transcription factors that are negatively regulated by TAF1 via phosphorylation

To determine potential transcription factors that bind to the 421 promoters, we utilized the ENCODE database which encompasses ChIP-seq data of 161 transcription factors from datasets from 91 different cell lines, and overlapped those bound genomic regions with 421 promoters, with a cutoff of +/- 0.2kb from the center of the TAF1 peaks. We then ranked the transcription factors by how many of promoters they bind to (Table 3.1). Not surprisingly among the top ranked were general transcription factors (GTFs) such as RNA Pol II. However, we also observed sequence specific transcription factors in this analysis. Based on their known interaction with TFIID or other members of the general transcription machinery and function in cell cycle regulation, four candidate transcription factors, FOXM1, Sp1, ATF2 and E2F1, were selected from the list for further analysis on their potential regulation by TAF1.

To begin, we first verified their binding to their corresponding target promoters by ChIP assays under the same conditions as the TAF1 ChIP-seq (Figure 3.3A and 3.3B). For FOXM1, we examined the binding to the BRIP1, SIRT1, and POLE2 promoters, all of which are involved in DNA damage response (DDR) (Alves-Fernandes and Jasiulionis, 2019; Frugoni et al., 2016;

Monteiro et al., 2012). For Sp1, we assayed the binding to the SOD2, RPS9, and PIGT promoters that are involved in reactive oxygen species, metabolism, cell cycle, and development, respectively (Kvarnung et al., 2013; Lindström and Nistér, 2010; Zelko et al., 2002). ATF2 binding was analyzed on the MSH6 and RAD23B promoters that are involved in DNA repair (Kunkel and Erie, 2005; Masutani et al., 1994). For E2F1, we examined its binding to the API5 and RPA2 promoters that are involved in apoptosis and DDR, respectively (Kunkel and Erie, 2005; Morris et al., 2006). The ChIP assay revealed that, like p53 binding on the p21 promoter, FOXM1, Sp1, ATF2, and E2F1 all bound to the target promoters with a binding pattern displaying an increase at 8 hours and a decrease at 16 hours, as well as a rescue by at 16 hours post UV apigenin. As expected, TAF1 shows the same binding pattern on these promoters. To assess functional significance of the differential binding of those transcription factors on their transcription activity, hnRNA was detected by qPCR at 0, 8, and 16 hours following UV as well as at 16 hours plus apigenin (Figure 3.3C). The assay indicated that hnRNA levels of target genes corresponded to the binding of transcription factors to the promoters. These results suggest that binding of FOXM1, Sp1, ATF2, and E2F1 to their target promoters and subsequent transcription are modulated by TAF1 kinase activity.

TAF1 directly phosphorylates E2F1 and FOXM1

To test whether TAF1 could directly phosphorylate candidate transcription factors like p53, E2F1, FOXM1, ATF2, and Sp1 were expressed in bacteria, purified, and then subjected to *in vitro* kinase assays with baculovirus expressed and purified TAF1 (Figure 3.4A, see figure 3.6 for purifications). As a measure of the specific kinase activity of TAF1, we include phosphorylation of RAP74 as a positive control and a TAF1 kinase-dead mutant (KD) as a negative control. Our results revealed that wild type TAF1 phosphorylates two of four transcription factors, E2F1 and FOXM1 (Figure 3.4A). To verify the specific phosphorylation, we also showed that the TAF1 kinase inhibitor apigenin blocked phosphorylation of E2F1 and FOXM1 by TAF1 (Figure 3.4B). We should note that the purified FOXM1 used in the kinase assays was a truncated protein (aa 1-599) with the transactivation domain deleted due to solubility and degradation of full-length protein.

We have previously shown that phosphorylation of p53 by TAF1, which leads to the dissociation of p53 and TAF1 from the p21 promoter, is cellular ATP level dependent (Wu et al., 2014). Following DNA damage, activation of PARP-1 leads to a decrease in cellular ATP level, which results in lower TAF1 kinase activity at 8 hours post UV. This decrease in ATP level is later sensed by AMPK, which allows for a recovery of cellular ATP level at 16 hours post UV (Wu et al., 2014). With two confirmed novel phosphorylation targets, we next test if their DNA binding is also regulated by cellular ATP level under identical conditions. We treat cells with the specific inhibitors of PARP-1 and AMPK, 4-AN and

compound C, to block PARP-1 mediated depletion of cellular ATP and AMPK mediated recovery of cellular ATP (Wu et al, 2014) and assayed for the effect on E2F1 and FOXM1 binding to their target promoters. As shown in Figure 3.4C, blocking PARP-1 and AMPK activity eliminated the binding pattern of E2F1 and FOXM1 after UV. These data provide further evidence that E2F1 and FOXM1 are regulated by TAF1 in a cellular ATP dependent manner.

TAF1 phosphorylates E2F1 at Ser375

To better understand TAF1-mediated phosphorylation, we aimed to identify the residue of E2F1 phosphorylated by TAF1 (Figure 3.5). We first made 4 truncated E2F1 constructs consisting of its 4 main functional domains, or group of domains: amino acids 1-108, 109-198, 199-361, and 362-347 (Figure 3.5A). These truncated E2F1 proteins were expressed in bacteria, purified, and then subjected to in vitro kinase assays with purified TAF1 (see Figure 3.6 for purifications). The assay shows that TAF1 specifically phosphorylates E2F1 transactivation domain (aa 362-347) (Figure 3.5B). Next, we individually mutated all 7 possible phosphorylation sites (6 serine and 1 threonine) within this domain to alanine and subjected them to an in vitro phosphorylation assay with purified TAF1 (see Figure 3.6 for purification). Results show that the only mutant unable to be phosphorylated by TAF1 was the S375A mutant, indicating that Ser375 may be the primary target for direct phosphorylation of E2F1 by TAF1 (Figure

3.5C). Interestingly, Ser375 has previously been described as a CDK8 phosphorylation site (Zhao et al., 2013). To rule out identified Ser375 phosphorylation is not due to a contaminant CDK8 activity under our assay condition, we showed the specific CDK8 inhibitor, SEL 120-34A, while inhibited CDK8-mediated E2F1 phosphorylation, did not have any effect on TAF1-mediated E2F1 phosphorylation (Figure 3.5E). These data indicate that TAF1 directly phosphorylates E2F1 at Ser375.

To examine if TAF1 could phosphorylate E2F1 at Ser375 in vivo, we overexpressed wildtype TAF1 and kinase-dead mutant TAF1 (KD) in U2OS and detected Ser375 phosphorylation of E2F1 using the phos-Ser375 specific antibody. As illustrated in Figure 3.5D, phosphorylation of endogenous E2F1 at Ser375 was clearly detected upon overexpression of TAF1, but not upon TAF1 KD mutant. Those data suggested that the kinase activity of TAF1 was responsible for Ser375 phosphorylation of endogenous E2F1.

TAF1 negatively regulates E2F1 and FOXM1 through phosphorylation

To examine the effect of TAF1 phosphorylation on E2F1 and FOXM1 binding to target promoter and transcription, we overexpressed wildtype TAF1 and kinase-dead mutant TAF1 in U2OS and examined the binding by ChIP assays and transcription by RT-qRCR (Figure 3.7). We found that, similar to p53, both E2F1 and FOXM1 show a decrease in binding to their target promoters

when wild type TAF1, but not the kinase-dead mutant, is overexpressed (Figure 3.7A). Consistent with this result, we also detected a reduction in the RNA level of E2F1 and FOXM1 target genes when wildtype TAF1, but not kinase dead TAF1, was overexpressed (Figure 3.7B). These data indicate that phosphorylation of E2F1 and FOXM1 by TAF1 negatively regulates their promoter occupancy and represses the transcription of target genes.

Since we identified the phosphorylation site on E2F1, we investigated the direct effect of this phosphorylation on E2F1 binding to a target promoter using phosphorylation mutants. We co-transfected cells with empty vector, wildtype TAF1 or kinase dead TAF1 overexpression plasmid with wildtype E2F1, S375A E2F1, or phosphor-mimetic S375D E2F1 overexpression plasmid, and assayed promoter binding with ChIP-PCR (Figure 3.7C). We observed an increase in binding with the S375A mutant compared to wildtype E2F1 when TAF1 is not overexpressed, and furthermore, we saw no change in binding upon overexpression of wildtype or kinase dead TAF1 (Figure 3.7C). Consistent with our model, the S375D phospho-mimetic E2F1 showed a decrease in binding compared to wildtype when co-transfected with empty vector and no change in binding upon overexpression of wildtype or kinase dead TAF1 (Figure 3.7C). These results provide further evidence that TAF1-mediated phosphorylation of E2F1 on Ser375 negatively regulates promoter binding.

2.4 Discussion

While TAF1's role in transcription initiation has long been studied, the molecular mechanisms by which it modulates transcription via its kinase activity are less clear. Our previous findings revealed a TAF1 kinase activity dependent mechanism for cell cycle regulation which functions through phosphorylation and inactivation of p53, suggesting a dual role for TAF1 in both positive and negative regulation of transcription (Li et al, 2004, Li et al, 2007). In addition, we reported TAF1 kinase activity was modulated by cellular ATP level (Wu et al., 2014). In this study, through ChIP-seq analysis, we were able to identify two more novel targets of TAF1 kinase. Furthermore, we've shown that TAF1 phosphorylates E2F1 and FOXM1 leading to reduction in transcription of target genes. These data provide additional evidence for TAF1 modulating transcription via its kinase activity.

E2F1 is a well-known regulator of cell cycle, responsible for transcriptional activation of a host of genes required for G1 progression and DNA replication (Roworth et al., 2015). Though the classically understood function of E2F1 is to promote cell cycle, it has also been shown to activate DNA damage repair genes and can induce cellular apoptosis; the complex role of this transcription factor in altering cell fate is tightly regulated by a host of posttranslational modifications that, through different mechanisms, alter its ability to transactivate target genes (Lin et al., 2001; Moroni et al., 2001; Roworth et al., 2015; Stevens et al., 2003).

TAF1 regulating E2F1 transcriptional activity through phosphorylation offers a new mechanism by which TAF1 kinase activity is involved in regulating cell cycle.

As previously described (Zhao et al., 2013), the kinase module of Mediator, CDK8, phosphorylates E2F1 at Ser375. Here we find that TAF1 also contributes to E2F1 phosphorylation on that residue. As Mediator and TFIID are both components of the general transcription machinery that incorporate kinase subunits, it's not surprising that there is some overlapping or redundant phosphorylation-mediated regulation. CDK8 can phosphorylate other proteins associated with the general transcription machinery as well as many sequence specific transcription factors which has been shown to have both positive and negative regulatory effects on transcription (Reviewed by Poss et al., 2013). This is similar to the functional diversity of TAF1 as a kinase, with phosphorylation of TAF7 and TFIIAβ associated with transcriptional activation, but phosphorylation of p53, E2F1 and FOXM1 is observed to have a negative effect on transcription (Kloet et al., 2012; Solow et al., 2001; Wu et al., 2014). Additionally, both CDK8 and TAF1 are important in the regulation of p53-dependent transcription of p21, demonstrating that these two kinase subunits of the general transcription machinery may work in coordination to orchestrate and fine-tune some gene-specific transcriptional regulation (Donner et al., 2010; Wu et al., 2014).

Among four candidate transcription factors we analyzed, ATF2 and Sp1 were found to display the same binding pattern following DNA damage as p53 on several target promoters, however, we did not detect their phosphorylation by

TAF1 in vitro. It's possible that these two transcription factors are regulated in a TAF1 kinase-dependent manner, but not through direct phosphorylation. Many of the promoters identified in our analysis are regulated by multiple sequence specific transcription factors, raising possibility for many modes of indirect regulation. Interestingly, it's been shown that E2F1 and Sp1 interact directly, which could potentially enable Sp1 dissociation through E2F1 phosphorylation. In fact, E2F1 seems to bind all of the Sp1 and ATF2 target promoters we identified for differential binding following DNA damage, which could potentially facilitate regulation of other transcription factors by TAF1. In addition, it is also possible that Sp1 and ATF2 may be phosphorylated by TAF1 in vivo, but not in vitro, due to lack of proper in vivo configuration and interaction partners.

Finally, both E2F1 and FOXM1 are known regulators of cell cycle, responsible for the transcription activation of a host of required genes. Though their classically understood function is to promote cell proliferation, they have also been implicated in the DNA damage response. It is well documented that the complex role of E2F1 in altering cell fate is tightly regulated by a range of posttranslational modifications that, through different mechanisms. TAF1 has previously been shown to play a critical role in cell G1 progression (Wang and Tjian 1994; Li et al 2004). Thus, the work presented here, revealing TAF1 as a negative regulator of E2F1 and FOXM1 under DNA damage conditions, expand our understanding of TAF1 kinase activity in regulating cell cycle. Furthermore, regulation by TAF1 is likely an important factor influencing the role of E2F1 and

FOXM1 in the DNA damage response. Interestingly, TAF1 has been suggested as a significantly mutated gene in multiple cancer types (Kandoth et al., 2013). A recent study predicted TAF1 as a likely oncogenic cancer driver in uterine corpus endometrial carcinomas (UCEC) (Bailey at al., 2018). Given its role in cancer, negative regulation of E2F1 and FOXM1 by TAF1 through phosphorylation could potentially be perturbed with mutation, contributing to cancer progression. Though this requires further investigation, the research presented here may help to unravel some of the complex regulatory mechanisms altered in cancer cells.

2.5 Materials and Methods

Reagents, transfection and antibodies

For inhibition of PARP-1, cells were treated with 1μM 4-AN (Trevigen) 30 minutes prior to UV irradiation. For inhibition of AMPK, cells were treated with 10μM Compound C (Millipore) 30 minutes prior to UV irradiation. For inhibition of TAF1, cells were treated with 40μM Apigenin (Sigma) 6 hours before harvesting. For inhibition of CDK8, Sel-120 34A (MedChemExpress) was added to in vitro phosphorylation reaction at the indicated concentrations.

Transfections were performed using BioT (Bioland Scientific) according to manufacturer's protocol. For overexpression of wildtype E2F1, S375A, or S375D, the genes were cloned into pcDNA3.1 vector with FLAG tag via PCR and prepared using ZymoPure II maxiprep kit (Zymo). For overexpression of wildtype

or kinase dead TAF1, pCMV-HAhTAF11 or p-LXSN-MT-TAF1N1398 A2/N7Ala2 plasmids were prepared using ZymoPure II maxiprep kit (Zymo).

Antibodies used for Western Blot: anti-TAF1 (6B3, Santa Cruz Biotechnologies), anti-Vinculin (Vin-11-5, Sigma), anti-E2F1 (KH95X, Santa Cruz Biotechnologies), anti-phospho (Ser375) E2F1 (Clone 71-10, Sigma). Antibodies used for ChIP assay were: anti-p53 (FL393, Santa Cruz Biotechnologies), anti-E2F1 (#3742, Cell Signaling), anti-FOXM1 (K-19X, Santa Cruz Biotechnologies), anti-ATF2 (C-19X, Santa Cruz Biotechnologies), anti-Sp1 (PEP2X, Santa Cruz Biotechnologies), anti-TAF1 (Ab1230), and anti-FLAG M2 (Sigma).

ChIP analysis

ChIP experiments were carried out as previously described (Li et a., 2007). U2OS cells were treated with 20 J/m² UVC, nuclear extracts were collected at indicated times post UV treatment, and sonicated to generate chromatin fragments of ~300 bp.

Primer set information: p21: 5'-GTGGCTCTGATTGGCTTTCTG and 5'-CTGAAAACAGGCAGCCCAAG, SOD1: 5'-ATTGGTTTGGGGCCAGAGTG and 5'-CTCGCAAACAAGCCTCCGTC, SIRT1: 5'- GGAGCGGTAGACGCAACA and 5'-CGTCCGCCATCTTCCAACT, BRIP1: 5'- CGTGGACTTCCCTCCGACTT and 5'-ATTCGTCTCGGGTTGTGTGG, POLE2: 5'- CTTCCCTCTCGCCCTTCAA and 5'- ACTTTCAGCCTACTCGGTCC, BRCA2: 5'-TGATAGAAGGTGGAAATGAGG and 5'- CATAAGGGGGCAGAATAAGAG, SOD2: 5'-

GGCTCAACATGCTGCTAGTG and 5’- CGCTTTCTTAAGGCCCGC, RPS9: 5’-CCTCTTTCTCAGTGACCGGG and 5’- GTTCAACCACCCTGCTCTGT, PIGT: 5’-CCTGCCTACTCCCTCTCGT and 5’- CCGGGATGCGGTTATCAGAG, MSH6: 5’-TTTAAATACTCTTTCCTTGCCTGG and 5’- TCTTCCGCTTTCGAGCAACT, RAD23B: 5’-CCTTGGGTTGGGCAGTAAATC and 5’-GCACTGGTGTGAAGTGTGAGA, API5: 5’ -GGTCAGGACAAGGATAGCGG and 5’ -CGGGGACTCAACTCACCTG, RPA2: 5’ -TGGTTTTCCGCTATTCCCCC and 5’ -GCTCGCCCTCTTGCTAAAAC.

All binding sites were amplified with 30-35 cycles PCR and analyzed by agarose gel electrophoresis and ethidium bromide staining.

Library generation and Illumina sequencing

Cells were fixed with 1% formaldehyde for 15 minutes and chromatin was sheared using Diagenode Bioruptor to a size of 200-400 bp. 45ug of sheared DNA was immunoprecipitated with anti-TAF1 antibody and Protein A beads (Pierce), washed, and eluted with 1% SDS. 5ug was set aside as input DNA. Following RNase A treatment (Qiagen) and reverse crosslinking, DNA was purified using QIAquick PCR purification kit (Qiagen). Libraries for Illumina sequencing were prepared using the Illumina-compatible NEXTflex ChIP-seq Kit (Bioo Scientific) as described (Li et al., 2007). The resulting DNA libraries were validated and quantified by checking the purity, concentration, and size of the amplicons on the Agilent Bioanalyzer High Sensitivity DNA Chip. Sequencing was performed on Illumina HiSeq 2000 instrument.

Data Analysis:

ChIP-Seq data sets were aligned using Bowtie2 (version 0.12.8) (Langmead et al., 2009) to the human reference genome (GRCh37/hg19). Alignment files were analyzed with MACS2 v. 2.0.10 using a 0.001 q-value cutoff (Zhang et al., 2008) to identify the TAF1 peaks. The R Bioconductor package ChIPpeakAnno (Zhu et al, 2010) was used to analyze peaks identified by ChIP-seq to acquire the closest Ensembl gene (10kb around transcription start site). MACS was also used for TAF1 differential binding analysis using the C3 stringency parameter ($-\log_{10}$(pvalue)=3, pvalue of < 0.001) to compare difference in binding between 2 conditions. This was used to find genes with significant increase in binding from 0 to 8 hours, decrease from 8 to 16 hours, and increase from 16 hours to 16 hours plus apigenin. We used the k-means clustering function of the Cistrome 'Heatmap' tool (Liu et al, 2011) to display TAF1 ChIP-Seq levels on heatmaps. In this analysis, the signal profiles from 0, 8, 16, and 16hr +API were entered into Cistrome along with a BED file containing the genomic regions centered at the summits of TAF1 peaks at 8hr after DNA damage to generate heatmaps. In the heatmap representation, each row represents the ±2.5 kb centered on the summit of TAF1 enriched peaks and ranked according to the enrichment of TAF1occupancy at 8 h after DNA damage.

TAF1 kinase assay

Flag tagged TAF1 and kinase-dead TAF1 were expressed in SF9 insect cells using baculovirus expression system, immunoprecipitated with ant-FLAG M2

Affinity Gel (Sigma) and eluted with FLAG peptide. His-RAP74, GST-E2F1, His-FOXM1, GST-Sp1, GST-ATF2 were expressed in E. coli and purified with either Ni-NTA resin (Qiagen) or Glutathione affinity gel (GE Healthcare). In vitro TAF1 phosphorylation reactions were carried out with ~50ng of TAF1 and 100-200ng of substrate as previously described (Li et al., 2004). Phosphorylation was detected by autoradiography.

Mapping E2F phosphorylation site

E2F1 serine-to-alanine mutants were made by site-directed mutagenesis using QuickChange II kit (Agilent) according the manufacturer's protocol. Primer information: S364A: 5' -TTGTCCCGGATGGGCGCCCTGCGGGCTCCCGTG and 5' -CACGGGAGCCCGCAGGGCGCCCATCCGGGACAA

S375A: 5' -GACGAGGACCGCCTGGCCCCGCTGGTGGCGGCC and 5' -GGCCGCCACCAGCGGGGCCAGGCGGTCCTCGTC

S382A: 5' -CTGGTGGCGGCCGACGCGCTCCTGGAGCATGTG and 5' -CACATGCTCCAGGAGCGCGTCGGCCGCCACCAG

S392A: 5' -GTGCGGGAGGACTTCGCCGGCCTCCTCCCTGAGGAGTT and 5' -CTCAGGGAGGAGGCCGGCGAAGTCCTCCCGCACATGCT

S401A: 5' -CCTGAGGAGTTCATCGCCCTTTCCCCACCCCAC and 5' -GTGGGGTGGGGAAAGGGCGATGAACTCCTCAGG

S403A: 5' -GAGTTCATCAGCCTTGCCCCACCCCACGAGGCC and 5' -GGCCTCGTGGGGTGGGGCAAGGCTGATGAACTC

T433A: 5’ -GACTTTGGGGACCTCGCCCCCCTGGATTTCTGA and 5’ -TCAGAAATCCAGGGGGGCGAGGTCCCCAAAGTC

RT-qPCR

RNA was extracted cells using TRIzol reagent (Sigma) according to manufacturer’s protocol. Reverse transcription was carried out using iScript Reverse Transcription Supermix (Bio-rad). RT-qPCR was performed using iQ SYBR Green Supermix (Bio-rad) on a CFX96 Real Time System (Bio-rad). Primer set information: SIRT1:5’-GGGAAGATTGCTCAGGGGTAA and 5’-TGAGGCACTTCATGGGGTATG

GAPDH: 5’-AGGTGAAGGTCGGAGTCAAC and 5’-GACAAGCTTCCCGTTCTCAG

p21: 5’-GACACAGCAAAGCCCGGCCA and 5’-CAACTCATCCCGGCCTCGCC

SOD2: 5’-GAAACCAAGCCAACCCCAAC and 5’-TCCAGGTGTCGCATTCTGAT

AP15: 5’-TCTCCAGGGTAAAACGGGTG and 5’-TGAAAAACTCCCAACACAAGTC MSH6: 5’ -TACAAGGACTGGCAGTCTGC and 5’ -CAGCTGGCAAACAGCACTAC

RPA2: 5’ -AACACTGTGGTTCCTCCAGAA and 5’ -ACTTCCCATTAAACAGGGAGAC

CTNNBIP1: 5’ – AGCTGCCTCCGCACTCCATC and 5’ -TGAGGAAGGAGATGGGATCA

2.6 References

1. Alves-Fernandes DK, Jasiulionis MG. The Role of SIRT1 on DNA Damage Response and Epigenetic Alterations in Cancer. Int J Mol Sci Internet.. 2019 Jun 28 cited 2020 Dec 4.;20(13). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6651129/

2. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018 Apr;173(2):371-385.e18.

3. Bieda M, Xu X, Singer MA, Green R, Farnham PJ. Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. Genome Res. 2006 May;16(5):595–605.

4. Carson DA, Seto S, Wasson DB, Carrera CJ. DNA strand breaks, NAD metabolism, and programmed cell death. Experimental Cell Research. 1986 Jun;164(2):273–281.

5. Chen X, Müller GA, Quaas M, Fischer M, Han N, Stutchbury B, et al. The Forkhead Transcription Factor FOXM1 Controls Cell Cycle-Dependent Gene Expression through an Atypical Chromatin Binding Mechanism. Molecular and Cellular Biology. 2013 Jan 15;33(2):227–236.

6. Curran EC, Wang H, Hinds TR, Zheng N, Wang EH. Zinc knuckle of TAF1 is a DNA binding module critical for TFIID promoter occupancy. Sci Rep. 2018 Dec;8(1):4630.

7. Dikstein R, Ruppert S, Tjian R. TAFII250 Is a Bipartite Protein Kinase That Phosphorylates the Basal Transcription Factor RAP74. Cell. 1996 Mar;84(5):781–790.

8. Donner AJ, Szostek S, Hoover JM, Espinosa JM. CDK8 Is a Stimulus-Specific Positive Coregulator of p53 Target Genes. Mol Cell. 2007 Jul 6;27(1):121–133.

9. Éthier C, Tardif M, Arul L, Poirier GG. PARP-1 Modulation of mTOR Signaling in Response to a DNA Alkylating Agent. PLoS ONE. 2012 Oct 24;7(10):e47978.

10. Fang Z, Lin M, Li C, Liu H, Gong C. A comprehensive review of the roles of E2F1 in colon cancer. Am J Cancer Res. 2020 Mar 1;10(3):757–768.

11. Frugoni F, Dobbs K, Felgentreff K, Aldhekri H, Al Saud BK, Arnaout R, et al. A Novel Mutation in the POLE2 Gene Causing Combined Immunodeficiency. J Allergy Clin Immunol. 2016 Feb;137(2):635-638.e1.

12. Hardie DG. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. Nat Rev Mol Cell Biol. 2007 Oct;8(10):774–785.

13. Hilton TL, Li Y, Dunphy EL, Wang EH. TAF1 Histone Acetyltransferase Activity in Sp1 Activation of the Cyclin D1 Promoter. MCB. 2005 May 15;25(10):4321–4332.

14. Hilton TL, Li Y, Dunphy EL, Wang EH. TAF1 Histone Acetyltransferase Activity in Sp1 Activation of the Cyclin D1 Promoter. Mol Cell Biol. 2005 May;25(10):4321–4332.

15. Hollern DP, Swiatnicki MR, Rennhack JP, Misek SA, Matson BC, McAuliff A, et al. E2F1 Drives Breast Cancer Metastasis by Regulating the Target Gene FGF13 and Altering Cell Migration. Scientific Reports. 2019 Jul 24;9(1):10718.

16. Imre G, Rajalingam K. Apoptosis inhibitor 5: Role in the inhibition of caspase-2. Cell Cycle. 2018 Feb 8;17(4):397–398.

17. Jacobson RH. Structure and Function of a Human TAFII250 Double Bromodomain Module. Science. 2000 May 26;288(5470):1422–1425.

18. Kaelin WG, Krek W, Sellers WR, DeCaprio JA, Ajchenbaum F, Fuchs CS, et al. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. Cell. 1992 Jul 24;70(2):351–364.

19. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013 Oct 17;502(7471):333–339.

20. Kloet SL, Whiting JL, Gafken P, Ranish J, Wang EH. Phosphorylation-dependent regulation of cyclin D1 and cyclin A gene transcription by TFIID subunits TAF1 and TAF7. Mol Cell Biol. 2012 Aug;32(16):3358–3369.

21. Kunkel TA, Erie DA. Dna mismatch repair. Annu. Rev. Biochem. 2005 Jun 1;74(1):681–710.

22. Kvarnung M, Nilsson D, Lindstrand A, Korenke GC, Chiang SCC, Blennow E, et al. A novel intellectual disability syndrome caused by GPI anchor deficiency due to homozygous mutations in PIGT. J Med Genet. 2013 Aug;50(8):521–528.

23. Li AG, Piluso LG, Cai X, Gadd BJ, Ladurner AG, Liu X. An acetylation switch in p53 mediates holo-TFIID recruitment. Mol Cell. 2007 Nov 9;28(3):408–421.

24. Li AG, Piluso LG, Cai X, Gadd BJ, Ladurner AG, Liu X. An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment. Molecular Cell. 2007 Nov;28(3):408–421.

25. Li H-H, Li AG, Sheppard HM, Liu X. Phosphorylation on Thr-55 by TAF1 mediates degradation of p53: a role for TAF1 in cell G1 progression. Mol Cell. 2004 Mar 26;13(6):867–878.

26. Liao G-B, Li X-Z, Zeng S, Liu C, Yang S-M, Yang L, et al. Regulation of the master regulator FOXM1 in cancer. Cell Communication and Signaling. 2018 Sep 12;16(1):57.

27. Lin WC, Lin FT, Nevins JR. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. Genes Dev. 2001 Jul 15;15(14):1833–1844.

28. Lindström MS, Nistér M. Silencing of Ribosomal Protein S9 Elicits a Multitude of Cellular Responses Inhibiting the Growth of Cancer Cells Subsequent to p53 Activation. PLOS ONE. 2010 Mar 8;5(3):e9578.

29. Lv C, Zhao G, Sun X, Wang P, Xie N, Luo J, et al. Acetylation of FOXM1 is essential for its transactivation and tumor growth stimulation. Oncotarget. 2016 Sep 13;7(37):60366–60382.

30. Martínez-Balbás MA, Bauer U-M, Nielsen SJ, Brehm A, Kouzarides T. Regulation of E2F1 activity by acetylation. The EMBO Journal. 2000 Feb 15;19(4):662–671.

31. Masutani C, Sugasawa K, Yanagisawa J, Sonoyama T, Ui M, Enomoto T, et al. Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. EMBO J. 1994 Apr 15;13(8):1831–1843.

32. Mizzen CA, Yang X-J, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, et al. The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity. Cell. 1996 Dec;87(7):1261–1270.

33. Monteiro LJ, Khongkow P, Kongsema M, Morris JR, Man C, Weekes D, et al. The Forkhead Box M1 protein regulates BRIP1 expression and DNA damage repair in epirubicin treatment. Oncogene. 2013 Sep;32(39):4634–4645.

34. Moroni MC, Hickman ES, Lazzerini Denchi E, Caprara G, Colli E, Cecconi F, et al. Apaf-1 is a transcriptional target for E2F and p53. Nat Cell Biol. 2001 Jun;3(6):552–558.

35. Morris EJ, Michaud WA, Ji J-Y, Moon N-S, Rocco JW, Dyson NJ. Functional Identification of Api5 as a Suppressor of E2F-Dependent Apoptosis In Vivo. PLOS Genetics. 2006 Nov 17;2(11):e196.

36. O'Brien T, Tjian R. Functional Analysis of the Human TAFII250 N-Terminal Kinase Domain. Molecular Cell. 1998 May;1(6):905–911.

37. Pearson A, Greenblatt J. Modular organization of the E2F1 activation domain and its interaction with general transcription factors TBP and TFIIH. Oncogene. 1997 Nov 27;15(22):2643–2658.

38. Pham A-D. Ubiquitin-Activating/Conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila. Science. 2000 Sep 29;289(5488):2357–2360.

39. Poppy Roworth A, Ghari F, La Thangue NB. To live or let die - complexity within the E2F1 pathway. Mol Cell Oncol. 2015 Mar;2(1):e970480.

40. Poss ZC, Ebmeier CC, Taatjes DJ. The Mediator complex and transcription regulation. Critical Reviews in Biochemistry and Molecular Biology. 2013 Nov;48(6):575–608.

41. Ruppert S, Wang EH, Tjian R. Cloning and expression of human TAF II 250: a TBP-associated factor implicated in cell-cycle regulation. Nature. 1993 Mar;362(6416):175–179.

42. Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, Balasubramanian S. Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. Genome Biology. 2013 Jan 24;14(1):R6.

43. Schreiber V, Dantzer F, Ame J-C, de Murcia G. Poly(ADP-ribose): novel functions for an old molecule. Nature Reviews Molecular Cell Biology. 2006 Jul;7(7):517–528.

44. Shi W, Feng Z, Zhang J, Gonzalez-Suarez I, Vanderwaal RP, Wu X, et al. The role of RPA2 phosphorylation in homologous recombination in response to replication arrest. Carcinogenesis. 2010 Jun;31(6):994–1002.

45. Solow S, Salunek M, Ryan R, Lieberman PM. Taf(II) 250 phosphorylates human transcription factor IIA on serine residues important for TBP binding

and transcription activity. J Biol Chem. 2001 May 11;276(19):15886–15892.

46. Stevens C, Smith L, La Thangue NB. Chk2 activates E2F-1 in response to DNA damage. Nat Cell Biol. 2003 May;5(5):401–409.

47. Wang EH, Tjian R. Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by hTAFII250. Science. 1994 Feb 11;263(5148):811–814.

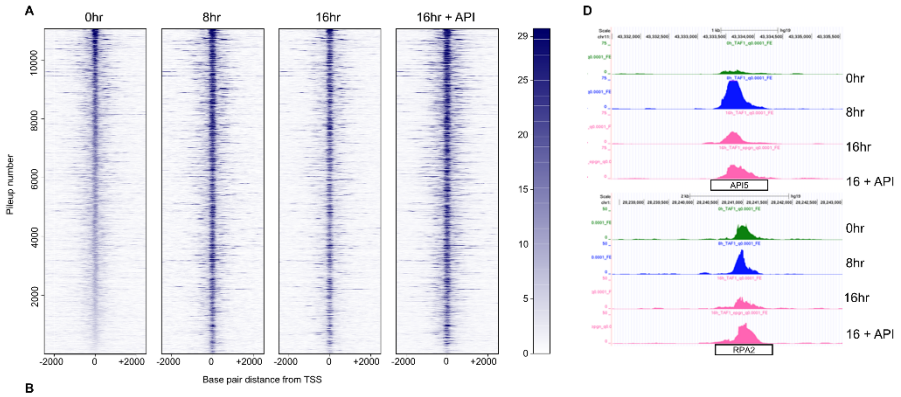
48. Wu Y, Lin JC, Piluso LG, Dhahbi JM, Bobadilla S, Spindler SR, et al. Phosphorylation of p53 by TAF1 Inactivates p53-Dependent Transcription in the DNA Damage Response. Molecular Cell. 2014 Jan;53(1):63–74.

49. Zelko IN, Mariani TJ, Folz RJ. Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. Free Radical Biology and Medicine. 2002 Aug 1;33(3):337–349.

50. Zhao J, Ramos R, Demma M. CDK8 regulates E2F1 transcriptional activity through S375 phosphorylation. Oncogene. 2013 Jul;32(30):3520–3530.

51. Zona S, Bella L, Burton MJ, Nestal de Moraes G, Lam EW-F. FOXM1: An emerging master regulator of DNA damage response and genotoxic agent resistance. Biochim Biophys Acta. 2014 Nov;1839(11):1316–1322.

2.7 Figures



B

	Peaks
0 ↑ 8 ↓ 16	2352
0 ↑ 8 ↓ 16 ↑ 16 + API	634
Remove 8 ↑ 8 + API	475

C

Overrepresented Reactome Pathways of the 421 promoters	P-value
Cell Cycle	1.96E-11
Metabolism of RNA	3.15E-10
Cell Cycle, Mitotic	2.69E-10
Metabolism of proteins	1.88E-10
Infectious disease	8.18E-10
Gene expression (Transcription)	8.21E-09
Cell Cycle Checkpoints	6.26E-08
Cellular responses to stress	0.000000113
G1/S Transition	0.000000143
Cellular responses to external stimuli	0.000000194
M Phase	0.000000366

Figure 3.1 TAF1 ChIP-seq and analysis (A) Heatmap of TAF1 ChIP-seq carried out at 0, 8, 16 hours after UV irradiation and 16 hours plus TAF1 kinase inhibitor, apigenin. Each row represents the ±2.5kb region centered on TAF1 peaks bound at 8 hours post UV and is ranked by TAF1 occupancy. X-axis represents distance from the transcription start site (TSS), and Y-axis shows pileup number. (B) TAF1 bound peaks that meet the C3 statistical stringency parameter for each indicated differential binding analysis overlap. (C) PANTHER Reactome pathway analysis of 421 promoters that are uniquely mapped from the 475 TAF1 bound peaks resulting from the differential binding analysis shown in panel B. Shows 10 most statistically overrepresented pathways. (D) UCSC genome browser shots of TAF1 peaks on API5 and RPA2 promoter at indicated time points following UV treatment from ChIP-seq. binding analysis shown in panel B. Shows 10 most statistically overrepresented pathways. (E) Overlap of TAF1 ChIP-seq dataset with ENCODE datasets from 6 other indicated cell lines.

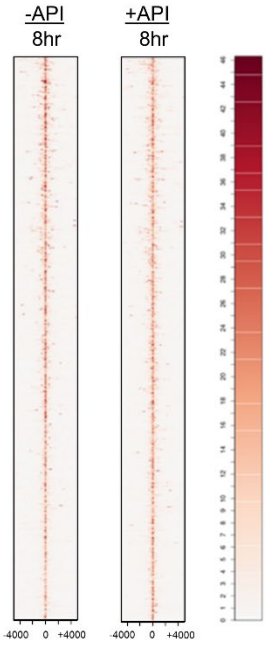


Figure 3.2 TAF1 ChIP-seq at 8 hours following UV and 8 hours plus apigenin Heatmap of 8kb regions of TAF1 bound peaks at 8 hours following UV treatment and 8 hours plus apigenin. Heatmap ranked by TAF1 occupancy at 8 hours following UV.

Rank	ChIP-seq	Promoter (421)
1	Pol2	398
2	Pol2	366
3	Pol2	356
4	Pol2	346
5	Pol2	324
6	Pol2	324
7	Taf1	316
8	Pol2	306
9	Pol2	304
10	Pol2	301
11	Pol2	288
12	Pol2	288
13	Pol2	287
14	Pol2	284
15	Pml	283
16	Pol2	282
17	Elf1	280
18	Mxi1	276
19	Pol2	276
20	Pol2	274
21	Yy1	271
22	Yy1	269
23	E2f1	266
24	Taf1	265
25	Sin3ak20	259
26	Mxi1	257
27	Tbp	249
28	Pol2	248
29	Taf1	244
30	Cmyc	239
31	Maz	236
32	Runx3	230
33	Creb1Dex100nm	229
34	Pol2	228
35	Cmyc	227
36	Sin3	227
37	Yy1c	222
38	Nrsf	219
39	Tbp	207
40	Gabp	204
41	Chd2	203
42	Max	202
43	Gabp	198
44	Elf1	197
45	Pou2f2	189
46	Sp1	173
47	Egr1	172
48	Gabp	166
49	Max	158
50	Mybl2sc81192	153

Rank	ChIP-seq	Promoter (421)
51	Maz	150
52	Pax5c20	146
53	Nrsf	145
54	Elf1	144
55	Foxm1	136
56	Ets1	135
57	Tblr1	133
58	Max	132
59	Znf1	126
60	Sp1	125
61	Elk1	124
62	Jund	123
63	Atf2	122
64	Ctcf	120
65	Six5	120
66	Tcf12	118
67	Tcf12	117
68	Nrf1	116
69	Cebpdsc636	115
70	Nfic	115
71	Tcf3	115
72	Nfic	114
73	Whip	112
74	Cmyc	109
75	Pax5n19	109
76	Bhlhe40c	106
77	GrDex100nm	106
78	Bhlhe40c	105
79	Chd1	105
80	Zbtb33	104
81	Cmyc	103
82	Chd2	102
83	Six5	102
84	Srf	102
85	Nfyb	101
86	Bclaf1	100
87	Rfx5	95
88	NfkbTnfa	94
89	Ctcf	92
90	Ctcf	89
91	E2f4	88
92	Mta3	87
93	Ctcf	85
94	Ebf1	85
95	Kap1	82
96	Pu1	81
97	Ctcf	80
98	Ets1	79
99	Sin3ak20	79
100	Hdac2	78

Table 3.1 Top 100 ranked transcription factors bound to 421 promoters 100 transcription factors from ChIP-seq datasets on ENCODE database that bind to the 421 promoters found in TAF1 differential binding analysis. Proteins are ranked by the number of the 421 promoters they bind to from TAF1 analysis.

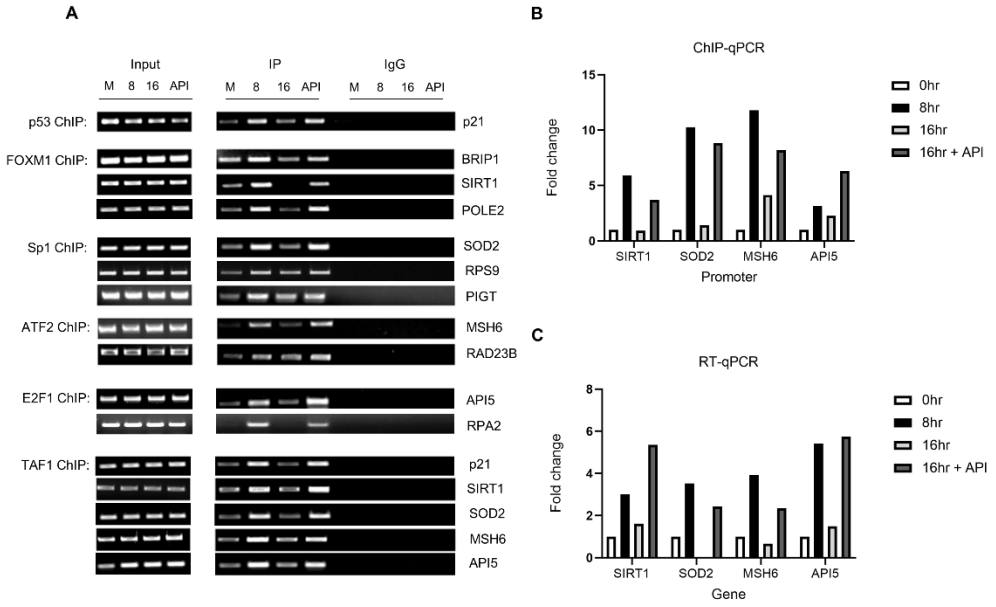


Figure 3.3 Binding patterns of FOXM1, Sp1, ATF2, and E2F1 following UV treatment (A) ChIP-PCR for p53, FOXM1, Sp1, ATF2, E2F1, and TAF1, on multiple promoters from differential binding analysis. Carried out at 0, 8, 16 hours post UV, and 16 hours + apigenin. (B) ChIP-qPCR validation for FOXM1, Sp1, ATF2, and E2F1 binding after UV treatment. (C) hnRNA fold change for target genes that ChIP-qPCR were carried out on.

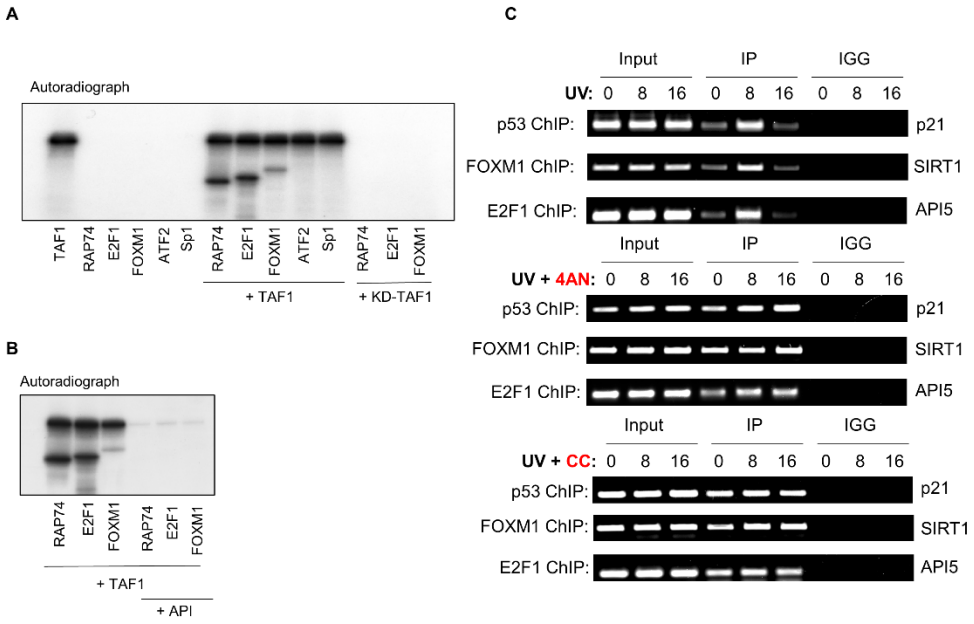


Figure 3.4 TAF1 phosphorylates E2F1 and FOXM1 (A) *In vitro* kinase assay with purified RAP74, E2F1, FOXM1, ATF2, and Sp1 with TAF1, visualized by autoradiography. Top bands are TAF1 autophosphorylation. (B) *In vitro* kinase assay with purified RAP74, E2F1, and FOXM1 with TAF1, showing inhibition by apigenin, visualized by autoradiography. (C) Promoter binding for p53, FOXM1, and E2F1 with stabilized ATP levels at 0, 8, and 16 hours post UV by inhibiting PARP-1 and AMPK with 4AN and Compound C, respectively, analyzed by ChIP-PCR.

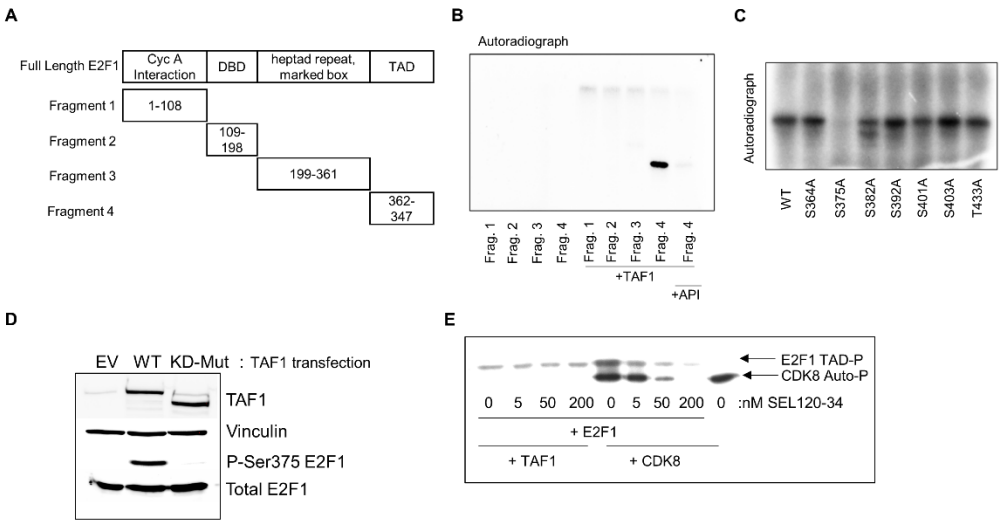


Figure 3.5 TAF1 phosphorylates E2F1 at Ser375 (A) Schematic of 4 domain truncations of E2F1. (B) TAF1 *in vitro* kinase assay with 4 E2F1 truncations, including inhibition by apigenin of TAD domain, phosphorylation visualized by autoradiography. (C) TAF1 *in vitro kinase assay* with E2F1 TAD domain phosphorylation-deficient mutants, visualized by autoradiography (D) Phosphorylation of E2F1 on Ser375 with transfection of empty vector, TAF1, or kinase-dead TAF1 overexpression plasmids, visualized by Western Blot using phospho-S375 E2F1 antibody. (E) *In vitro* kinase assay showing phosphorylation of E2F1 by purified TAF1 or CDK8 with increasing concentration of CK8 inhibitor, SEL 120-34, visualized by autoradiography.

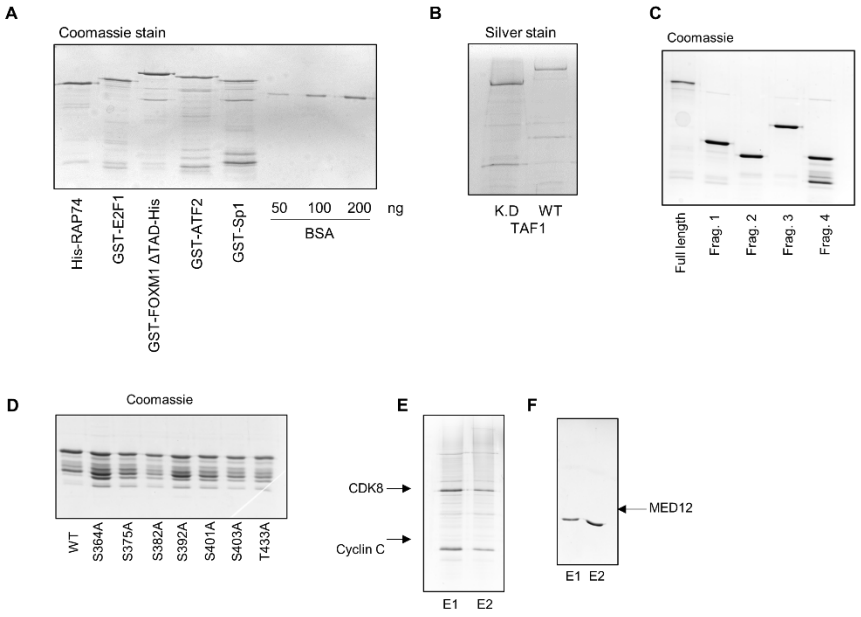


Figure 3.6 Purification of RAP74, E2F1, FOXM1, ATF2, Sp1, TAF1, CKD8, Cyclin C, and MED12. (A) Coomassie stained SDS-PAGE gel of RAP74, E2F1, FOXM1, ATF2, and Sp1, expressed in E. coli and purified using indicated affinity tags. (B) Silver stained SDS-PAGE gel of wildtype TAF1 and kinase dead TAF1 (K.D.) expressed in insect cells and purified via FLAG-tag. (C) Coomassie stained SDS-PAGE gel showing truncated domains of E2F1 purified from E. coli via GST-tag. (D) Purified bacteria expressed mutants of E2F1 TAD domain visualized via Coomassie stained SDS-PAGE. (E) Coomassie stained SDS-PAGE gel showing co-purified CDK8 and Cyclin C co-expressed in insect cells and co-purified via FLAG-tag on CDK8. (F) Coomassie stained SDS-PAGE gel on purified MED12 expressed in bacteria and purified via GST-tag.

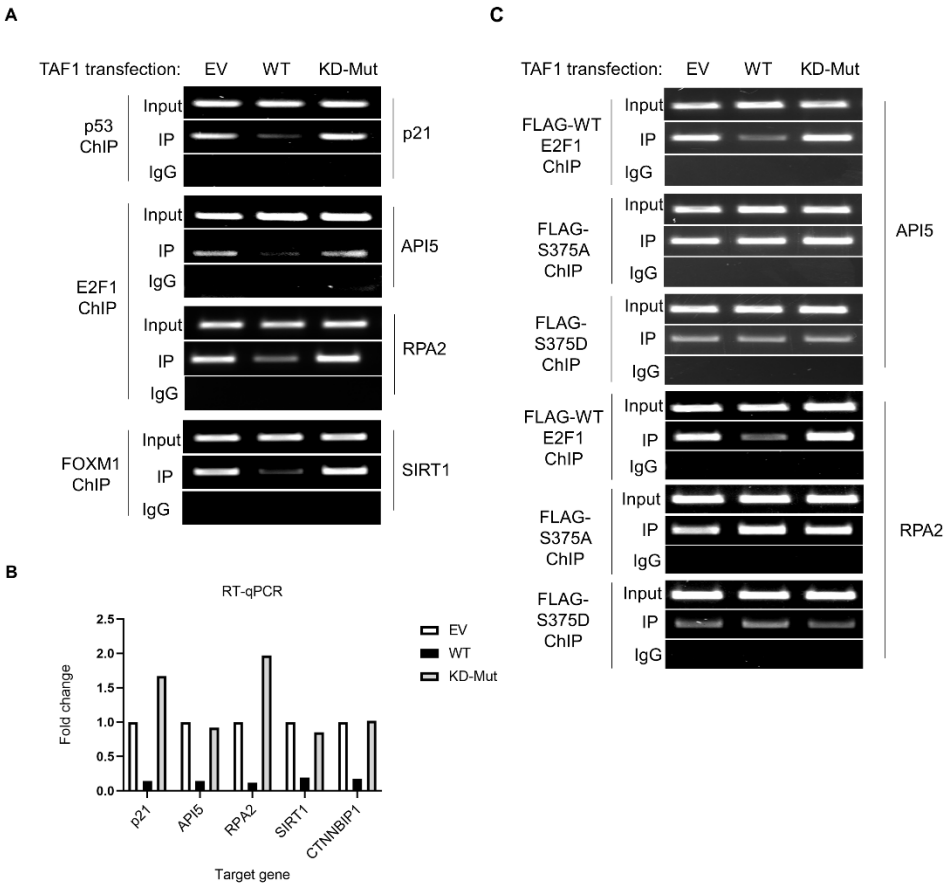


Figure 3.7 Phosphorylation of E2F1 and FOXM1 leads to reduced promoter occupancy and target gene expression. (A) Promoter binding of p53, E2F1, and FOXM1 with overexpression of wildtype or KD-TAF1, analyzed by ChIP-PCR. (B) hnRNA fold change of p53, E2F1, and FOXM1 target genes with overexpression of wildtype or KD-TAF1 analyzed by RT-qPCR. (C) Cells were co-transfected with TAF1 (empty vector, wildtype TAF1, or KD-TAF1) and FLAG-E2F1 (wildtype E2F1, S375A, or S375D) and promoter binding was analyzed by ChIP-PCR using anti-FLAG antibody to pull down exogenously expressed E2F1.

Chapter 4: Conclusion

4.1 Conclusions:

As a member of the general transcription machinery, TAF1 plays a critical role in the expression of protein coding genes. TFIID, the complex TAF1 associates with, is first to recognize and bind to a gene promoter, seeding the transcription preinitiation complex (Thomas and Chiang, 2006). However, TAF1 is a unique protein among the general transcription machinery as it has a diverse range of enzymatic functions which gives rise to gene specific regulation. TAF1's HAT and kinase activity have both been shown to have a transcriptional consequence specifically for cell cycle related genes and furthermore, loss of either domain leads to cell cycle arrest (Noguchi et al., 1994; O'Brien and Tjian, 1998). Our lab has previously reported a mechanism by which TAF1 regulates cell cycle through phosphorylation of tumor suppressor p53 (Wu et al., 2014). The results described in this dissertation offer new insights into how TAF1 kinase activity affects cell cycle regulation.

E2F1 and FOXM1 are both master regulators of cell cycle and the DNA damage response (Liao et al., 2018; Roworth et al., 2015). We found that phosphorylation of these two transcription factors by TAF1 leads to a reduction in promoter binding and expression levels of target genes. Given the nature of TAF1 kinase being cellular ATP dependent, this implies that FOXM1 and E2F1 transcriptional activity can be regulated by metabolic ques. As our lab has

extensively studied, p53 is regulated during the DNA damage response by ATP fluctuations sensed by TAF1 kinase (Wu at al., 2014). The findings presented here suggest that this is also the case for E2F1 and FOXM1. In the early stages of DNA damage there is a reduction in TAF1 kinase activity caused by PARP-1 mediated ATP depletion, during this time we see a corresponding increase in E2F1 and FOXM1 binding to promoters. When ATP levels recover and TAF1 kinase becomes active again we see a drop in E2F1 and FOXM1 binding. This is the same trend seen with p53 on the p21 promoter. This suggests that TAF1 plays an important role in orchestrating the DNA damage response with the ability to mediate stabilization of p53, E2F1, and FOXM1 on target genes and later terminate transcription through the effect that the ATP status of the cell has on its kinase domains.

The E2F1 target genes focused on in this study were API5 and RPA2. ChIP-seq analysis showed a high degree of differential TAF1 binding on these promoters during times of altered cellular ATP, and in the presence of the TAF1 kinase inhibitor, apigenin. The function of API5 (apoptosis inhibitor 5) remains largely ambiguous, particularly in DNA damage conditions; it has been shown to inhibit E2F1-induced apoptosis in response to certain stressors like α-toxin but, despite its name, does not seem to have an anti-apoptosis effect in cells that have undergone DNA damage (Imre and Rajalingam, 2018). RPA2 is a single stranded DNA binding protein involved in DNA replication and DNA damage repair which becomes phosphorylated following DNA damage and facilitates homologous

recombination (Shi et al., 2010). Given that the role of API5 in DNA damage conditions is not well understood, it is hard to speculate whether the phosphorylation of E2F1 which leads to a reduction in API5 expression would have a pro-growth effect or the opposite, or what the consequence for the DNA damage response might be. However, for RPA2, which is involved in DNA repair, it makes sense that the TAF1-E2F1 interaction leads to an induction of expression in the early stages of DNA damage and later facilitates termination of transcription.

The FOXM1 targets we looked at were BRIP1, SIRT1, and POLE2. These are all heavily involved in DNA damage response. With low TAF1-mediated phosphorylation in the early stages of DNA damage, FOXM1 stabilizes on these promoters to induce expression, allowing these DNA repair genes to function at the appropriate time. And like with p53, and E2F1, at the later stages of DNA damage TAF1 mediates termination of transcription of these genes. This makes physiological sense and suggests that TAF1 is one of the many proteins involved in coordinating the appropriate transcriptional response to DNA damage.

From an evolutionary standpoint, it makes sense that that cells would have evolved a mechanism to harness this fluctuation in cellular ATP caused by one component of the DNA damage response as a signal to elicit another. This also brings up the question of how individuals with metabolic disorders, such as diabetes, would be affected by this regulatory mechanism. A major phenotype of diabetes is hyperglycemia, which results in increased ATP production in certain tissues; this could result in aberrant activation of TAF1 kinase leading to reduced

levels of p53, FOXM1, and E2F1 target genes. As these transcription factors are all regulators of cell cycle as well as the DNA damage response, this regulation may play a part in diabetes leading to an increased risk of cancer, a phenomenon that has been extensively studied (Giovannucci et al., 2010). Some experiments currently ongoing in our lab involving high glucose treatment of cells may be able to shed light on these questions.

Another potential avenue of study that the work described in this dissertation leads to, is the investigation of the role that acetylation may play in the regulation of E2F1 and FOXM1 by TAF1. Our lab has previously shown that TAF1 is recruited to the p21 promoter by acetylated p53 through interaction with TAF1's double bromodomain (Li et al.,2007). E2F1 and FOXM1 are both acetylated by P300, the same enzyme responsible for p53 acetylation (Lv et al., 2016; Marzio et al., 2000). It's possible that TAF1 is recruited to FOXM1 and E2F1 target promoters through its double bromodomain. This could give rise to a similar mechanism of regulation as is seen with p53 where acetylation-mediated recruitment of TAF1 leads to transcriptional activation which is later terminated through TAF1's kinase domains (Li et al., 2007; Wu et al., 2014).

On a genome wide scale, E2F1, FOXM1, and p53 combined bind to an enormous set of gene promoters. The ability of TAF1 to regulate the DNA binding of these transcription factors provides new explanations for the results of older reports that the kinase domains of TAF1 are necessary for cell cycle progression. It would be interesting to examine this more quantitatively by doing performing an

RNA-seq experiment with a cell line that expresses a kinase dead TAF1. This would allow us to see exactly what genes are affected by TAF1-mediated phosphorylation. It would also be interesting to do something similar with the other domains of TAF1 such as the double bromodomain, HAT, and UBAC domain, assuming altering these wouldn't completely disturb basal transcription.

TAF1 is one of the most commonly mutated general transcription factors in cancer (Kandoth et al., 2013). While the most common mutations to TAF1 in cancer occur within the HAT domain, there are many that have appeared in cases, according to The Cancer Genome Atlas, that occur within both the NTK and CTK. The kinase activity of TAF1 has the ability to regulate p53, E2F1, and FOXM1 which are major effectors in cell cycle. It's possible that mutations and subsequent alterations to the kinase domains of TAF1 could have an effect on regulation of these cell cycle proteins and confer an oncogenic phenotype. Further analysis would need to be performed to investigate this possibility. Interestingly, apigenin, a known TAF1 kinase inhibitor which we've employed in many experiments described in this dissertation, has been shown by many publications to have an anti-cancer effect (Sung et al., 2016; Yan et al., 2017). While apigenin, a naturally occurring ATP analog, does have multiple targets other than TAF1, the data presented in this dissertation may provide further insight into how this molecule confers anti-cancer effects. Furthermore, this implicates the kinase domains of TAF1 as a possible drug target for cancer therapy.

The implications of phosphorylation-mediated transcriptional regulation facilitated by TAF1 are uniquely interesting when viewed in the context of brain cells. Mutations in the TAF1 gene, and a neuron-specific isoform, N-TAF1, have been found to be associated with intellectual disability and X-linked dystonia-parkinsonism (Gudmundsson et al., 2019; Sako et al., 2011). The TAF1 mutations observed in patients with intellectual disability are most commonly found in the HAT domain, double bromodomain, and RAP74-interacting domain (O'Rawe et al., 2015). N-TAF1 has an insertion of a 6 nucleotide microexon as a result of a neuron-specific splicing factor, SRRM4/nSR100 (Capponi et al., 2019). It is not known how these variations to the TAF1 gene affect kinase activity but it's conceivable that alterations to the protein could have an effect on the kinase domains or interaction with phosphorylation targets. This also begs the question of how, or if, neuron-specific TAF1 variations play any part in brain cancer progression. As we've demonstrated in this dissertation, TAF1 not only regulates p53, but also FOXM1 and E2F1, all of which are dysregulated across many cancer types; variations in TAF1 could have an effect on this regulation in neural tissues. Furthermore, metabolic pathways have been suggested as therapeutic targets for glioblastoma; inhibiting metabolic pathways could alter cellular ATP levels and have an effect on TAF1 kinase-mediated regulation of p53, FOXM1 and E2F1, perhaps playing into the benefit of these drugs (Marie and Shinjo, 2011).

4.2 References:

1. Capponi S, Stöffler N, Irimia M, Schaik FMAV, Ondik MM, Biniossek ML, et al. Neuronal-specific microexon splicing of TAF1 mRNA is directly regulated by SRRM4/nSR100. RNA Biology. 2020 Jan 2;17(1):62–74.

2. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and Cancer. Diabetes Care. 2010 Jul;33(7):1674–1685.

3. Gudmundsson S, Wilbe M, Filipek-Górniok B, Molin A-M, Ekvall S, Johansson J, et al. TAF1 , associated with intellectual disability in humans, is essential for embryogenesis and regulates neurodevelopmental processes in zebrafish. Scientific Reports. 2019 Jul 24;9(1):10730.

4. Imre G, Rajalingam K. Apoptosis inhibitor 5: Role in the inhibition of caspase-2. Cell Cycle. 2018 Feb 8;17(4):397–398.

5. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013 Oct;502(7471):333–339.

6. Li AG, Piluso LG, Cai X, Gadd BJ, Ladurner AG, Liu X. An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment. Molecular Cell. 2007 Nov 9;28(3):408–421.

7. Liao G-B, Li X-Z, Zeng S, Liu C, Yang S-M, Yang L, et al. Regulation of the master regulator FOXM1 in cancer. Cell Commun Signal Internet.. 2018 Sep 12 cited 2020 Nov 7.;16. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6134757/

8. Lv C, Zhao G, Sun X, Wang P, Xie N, Luo J, et al. Acetylation of FOXM1 is essential for its transactivation and tumor growth stimulation. Oncotarget. 2016 Aug 17;7(37):60366–60382.

9. Marie SKN, Shinjo SMO. Metabolism and Brain Cancer. Clinics (Sao Paulo). 2011 Jun;66(Suppl 1):33–43.

10. Marzio G, Wagener C, Gutierrez MI, Cartwright P, Helin K, Giacca M. E2F family members are differentially regulated by reversible acetylation. J Biol Chem. 2000 Apr 14;275(15):10887–10892.

11. Noguchi E, Sekiguchi T, Nohiro Y, Hayashida T, Hirose E, Hayashi N, et al.

Minimum essential region of CCG1/TAFII250 required for complementing the temperature-sensitive cell cycle mutants, tsBN462 and ts13 cells, of hamster BHK21 cells. Somat Cell Mol Genet. 1994 Nov;20(6):505–513.

12. O'Rawe JA, Wu Y, Dörfel MJ, Rope AF, Au PYB, Parboosingh JS, et al. TAF1 Variants Are Associated with Dysmorphic Features, Intellectual Disability, and Neurological Manifestations. Am J Hum Genet. 2015 Dec 3;97(6):922–932.

13. Poppy Roworth A, Ghari F, La Thangue NB. To live or let die - complexity within the E2F1 pathway. Mol Cell Oncol. 2015 Mar;2(1):e970480.

14. Sako W, Morigaki R, Kaji R, Tooyama I, Okita S, Kitazato K, et al. Identification and localization of neuron-specific isoform of TAF1 in rat brain: implications for neuropathology of DYT3 dystonia. Neuroscience. 2011 Aug 25;189:100–107.

15. Shi W, Feng Z, Zhang J, Gonzalez-Suarez I, Vanderwaal RP, Wu X, et al. The role of RPA2 phosphorylation in homologous recombination in response to replication arrest. Carcinogenesis. 2010 Jun;31(6):994–1002.

16. Sung B, Chung HY, Kim ND. Role of Apigenin in Cancer Prevention via the Induction of Apoptosis and Autophagy. J Cancer Prev. 2016 Dec;21(4):216–226.

17. Thomas MC, Chiang C-M. The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol. 2006 Jun;41(3):105–178.

18. Wu Y, Lin JC, Piluso LG, Dhahbi JM, Bobadilla S, Spindler SR, et al. Phosphorylation of p53 by TAF1 inactivates p53-dependent transcription in the DNA damage response. Mol Cell. 2014 Jan 9;53(1):63–74.

19. Yan X, Qi M, Li P, Zhan Y, Shao H. Apigenin in cancer therapy: anti-cancer effects and mechanisms of action. Cell Biosci Internet.. 2017 Oct 5 cited 2020 Nov 7.;7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5629766