UC Riverside
UC Riverside Electronic Theses and Dissertations

Title
Genome-Wide Analysis of Transcriptional Regulation via TAF1 Kinase

Permalink
https://escholarship.org/uc/item/55g6w5h5

Author
Maxham, Lily Ann

Publication Date
2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Genome-Wide Analysis of Transcriptional Regulation via TAF1 Kinase

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

Doctor of Philosophy

in

Cell, Molecular, and Developmental Biology

by

Lily Ann Maxham

December 2017

Dissertation Committee:
Dr. Xuan Liu, Chairperson
Dr. Frances Sladek
Dr. Weifeng Gu

Copyright by
Lily Ann Maxham
2017

The Dissertation of Lily Ann Maxham is approved:

____

____

____

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to acknowledge the many great scientists that have aided and supported me in my PhD research. I would like to begin by acknowledging the laboratory of Dr. Xuan Liu. First and foremost, Dr. Xuan Liu has been my mentor and PI throughout my PhD. She has guided me and provided substantial insight that has led me to the following results. Further, I would like to acknowledge Dr. Yong Wu for his mentorship in laboratory techniques and the RNA-seq data used in Chapter 3. Dr. Joy Lin performed the TAF1 ChIP-seq that was used in Chapter 2 and Chapter 3 analyses. Thomas Benedict purified many bacterial proteins for this study as well as the TAF1 kinase protein via insect cells. Thomas has been a very valuable collaborator for the TAF1 kinase experiments utilized in Chapter 2. Selene Bobadilla has also acted as a mentor in laboratory techniques and provided insight into the projects presented here. Further, other members of Dr. Xuan Liu's lab - including Yumi Nobumori, Carina Edmondson, and Jacob Schrier - have contributed conceptually to the data presented.

As well, Dr. Sladek and Dr. Gu have acted as members on my guidance committee and have offered guidance, advice, and support throughout my PhD. Dr. Sladek has provide valuable insight and offered suggestions that led to the success discussed here. Dr. Gu provided bioinformatic assistance and mentorship that aided in the identification of TAF1 kinase targets.

My husband, Dr. Patrick Schreiner, has been supportive and encouraging throughout my PhD. Further, Dr. Schreiner provided bioinformatic guidance utilized in Chapter 3 to help determine core promoter element abundance in TAF1 bound peaks. Lastly, I would like to acknowledge my family for the continued support throughout my scientific career: Susan Maxham, Robert Maxham, Erik Hausler, Heidi Hausler, and Danielle Hausler.

Dedication:

To my darling daughter, Eva:

Curiosity is one of the forms of feminine bravery

- Victor Hugo

ABSTRACT OF THE DISSERTATION

Genome-Wide Analysis of Transcriptional Regulation via TAF1 Kinase

by

Lily Ann Maxham

Doctor of Philosophy, Graduate Program in Cell, Molecular,
and Developmental Biology
University of California, Riverside, December 2017
Dr. Xuan Liu, Chairperson

Gene regulation is a complex mechanism and our knowledge of molecular pathways remains limited. In particular, current research examining transcriptional regulation emphasizes events that precede the initiation of transcription. However, correct regulation also requires turning off transcription when it’s no longer needed. Because DNA-bound activators interact with and recruit components of the basal transcription apparatus to the promoter, these basal factors could, in theory, mark the activators. Presumably, through post-translational modification after transcription, this process could lead to transcription termination. To determine this mechanism, previous studies have elucidated the relationship between the TATA-Binding Protein Associated Factor 1 (TAF1) and p53 in response to Ultra Violet (UV) DNA damage. The p53 tumor suppressor is a transcription factor that plays a critical role in guarding cell genomes against DNA damage. Among the genes induced by p53 is p21, which induces cell cycle arrest. TAF1 phosphorylates p53 at Thr55 on the p21 promoter, leading to p53 dissociation from the promoter and transcription

termination. TAF1 is the largest subunit of general transcription factor TFIID and possesses intrinsic protein kinase activity. Because TAF1 is a component of general transcription machinery bound to many promoters, the following research was designed to elucidate how TAF1 contributes to transcriptional regulation via it's DNA binding activities and kinase activities. This dissertation aims to reveal TAF1's genome-wide response to UV induced DNA damage. My studies included an in-depth look into TAF1's recruitment to the promoter via core promoter elements, novel TAF1 kinase targets, and the regulation of gene expression upon DNA damage via TAF1. These studies will provide important insights into how TAF1 impacts genome-wide transcription regulation and explores how to regulate transcription termination. Since the regulation of gene expression is a vital process for the integrity of organisms, this discovery provides an in-depth understanding of the cellular development of a wide range of species.

Table of Contents: Page:

Chapter 1: Introduction

1.1 General transcription machinery and gene regulation 2
1.1a Assembly pathway 2
1.1b TFIID 4
1.2 Core promoter elements 5
1.2a Core promoter elements and the regulation of gene expression 6
1.3 TAF1 8
1.3a TAF1 and core promoter elements 9
1.3b TAF1 structure 10
1.3c TAF1 interaction and modification of general transcription machinery 13
1.3d TAF1 interaction and modification of histones 15
1.3e TAF1 regulation through activators 16
1.4 Interplay between TAF1 and p53 in response to DNA damage 17
1.4a TAF1 recognition of acetylation p53 17
1.4b TAF1 regulation of p53 in response to UV via kinase activity 18
1.4c TAF1 kinase activity is cellular [ATP] dependent 19
1.5 Genomic integrity and disease 21
1.5a TAF1 mutations in cancer 21
1.5b TAF1 mutations in disease 22
1.6 Summary of Chapters 22
1.7 References 24
1.8 Figures and tables 33

Chapter 2: Genome-wide analysis of transcriptional regulation by TAF1 in response to UV induced DNA damage

2.1 Abstract 38
2.2 Introduction 38
2.3 Results 42
2.3a Genome-wide TAF1 binding to the promoters upon DNA damage 42
2.3b Optimization of TAF1 ChIP-seq analysis to identify potential TAF1 kinase regulated promoters 43
2.3c Identification and characterization of TAF1 kinase regulated promoters 45
2.3d Utilization of the ENCODE database in the identification of potential transcription factor targets 47
2.3e Transcription factors FOXM1, ATF2, E2F1, and SP1 show 48

similar binding pattern and target expression to p53 in response to UV treatment	
2.3f Transcription factor dissociation via TAF1 kinase is ATP-level dependent	50
2.3g TAF1 directly phosphorylates FOXM1and E2F1	51
2.4 Discussion	52
2.5 Materials and methods	60
2.6 References	64
2.7 Figures and tables	72

Chapter 3: Core promoter element profile of TAF1 bound peaks in response to UV DNA-induced DNA damage and the correlation with expression of target genes.

3.1 Introduction	86
3.2 Results	89
3.2a Core promoter element profile of TAF1 bound promoters upon DNA damage.	89
3.2b Differential peak analysis of TAF1 ChIP-seq revealed a higher abundance of TATA elements present in TAF1-recruited promoters following DNA damaging conditions	90
3.2c RNA-Seq analysis indicates gene expression following UV treatment is partially reliant on the core promoter elements present	91
3.3 Discussion	94
3.4 Methods	96
3.5 References	99
3.6 Figures and tables	103

Chapter 4: Conclusion

4.1 Conclusion	110
4.2 References	117

List of Figures: Page:

Chapter 1:

Figure 1.1: Recognition of TFIID and TFIIB to core promoter elements 33
Figure 1.2: TAF1 domains and substrates 34
Figure 1.3: TAF1 regulation of p53 in response to UV-induced DNA damage 35

Chapter 2:

Figure 2.1: TAF1 ChIP-seq analysis and comparison to other cell lines 72
Figure 2.2: Comparison of C2 and C3 TAF1 ChIP-seq analysis 73
Figure 2.3: Identification of potential promoters under TAF1 kinase regulation 74
Figure 2.4: PANTHER go-term analysis of top 421 promoters 75
Figure 2.5: Schematic of target TAF1 transcription factor identification 76
Figure 2.6: ENCODE database shows known transcription factor binding in the genome browser 79
Figure 2.7: ChIP reveals positive correlation to p53 binding and target expression 80
Figure 2.8: Not all transcription factors correlated with p53 trend following UV treatment 81
Figure 2.9: String analysis of protein interactions 82
Figure 2.10: Dissociation of FOXM1 on BRCA2 is [ATP] dependent 83
Figure 2.11: TAF1 phosphorylation of FOXM1 and E2F1 84

Chapter 3:

Figure 3.1: Higher core promoter abundance in the TAF1 bound promoters upon DNA damage 103
Figure 3.2: Genome-wide TAF1 recruitment to core promoter elements following UV induced DNA damage 104
Figure 3.3: RNA-seq analysis following UV treatment 105
Figure 3.4: Functional annotation of UV activated or repressed genes 106
Figure 3.5: Core promoter elements present in activated or repressed genes following DNA damage that were TAF1 bound in mock conditions or TAF1 recruited following DNA damaging conditions 108

List of Tables: Page:

Chapter 1:

Table 1.1: Top 10 transcription factor SMGs in cancer identified in 12 cancer types. 36

Chapter 2:

Table 2.1: ENCODE database cell line selection and total available ChIP-seq's 76
Table 2.2: Top 100 ChIP-seq proteins bound to the 421 promoters 77
Table 2.3: Select transcription factor targets and their known functions 78

Chapter 3:

Table 3.1: Genes found to significantly increase or decrease in expression bound by TAF1 following UV treatment 107

Chapter 1: Introduction to TAF1 and its role in transcriptional regulation

1.1 General transcription machinery and gene regulation

The fundamentals of life depend on the carefully orchestrated timing of gene expression within cells. The expression of genes is a complex and highly-regulated process vital and universal in all cells. Regulation of gene expression is critical for proper development and the lack of proper expression of genes may cause lethality or disease. This regulation is heavily dependent on the general transcription machinery (GTM) comprising the preinitiation complex (PIC). The protein RNA polymerase II transcribes protein coding genes in eukaryotic cells and, in conjunction with other general transcription machinery, allows for gene-specific expression. In order to identify other members of the pre-initiation complex, RNA polymerase II was purified and subcellular fractionation was performed. The fractions were named TF (for transcription factor) roman numeral II (for RNA Pol II driven) and then fraction they were found in (A-F), resulting in TFIIA, TFIID, TFIIB, TFIIE, and TFIIF. Further, they were found to be crucial in the regulation of gene expression (Sawadogo and Roeder, 1985; Reinberg and Roeder, 1987, Flores et al., 1989; Flores et al., 1992, Ge et al., 1996).

1.1a Assembly pathway:

Transcription factor assembly, a critical step in the process of transcription, has been thoroughly studied. There are two main theories in PIC assembly, either a sequential assembly or a holoenzyme assembly. The

sequential assembly pathway states that TFIID is the first GTF to recognize the core promoter, the site of transcription initiation. This is followed by TFIIA, TFIIB, Pol II/TFIIF, TFIIE, and lastly TFIIH (Thomas and Chiang, 2006). This sequential pathway was determined via in vitro transcription experiments with chromatographic fractions (Fire et al., 1984; Hawley and Roeder, 1985; Samuels and Sharp 1986), native gel electrophoresis and DNase I footprinting (Buratowski et al., 1989).

The holoenzyme pathway states that RNA Pol II forms a complex with some members of the GTM and proteins involved in chromatin remodeling, DNA repair, and mRNA processing (Thomas and Chiang, 2006). Upon purifying human RNA pol II, it was determined that TFIIB, TFIIE, TFIIF, TFIIH, GCN5 (HAT), SWI/SNF (chromatin remodelers), and SRBs (suppressors of RNA polymerase B mutations) were all co-purified (part of the holoenzyme). Importantly, this holoenzyme did not contain TFIID or TFIIA (Wu and Chiang, 1998; Wu et al., 1999). This alludes to the idea that TFIID is involved in promoter identification and plays a role in the RNA pol II holoenzyme recruitment to the promoter for transcription (Thomas and Chiang, 2006).

While both pathways may occur in an *in vivo* setting, it is important to note that TFIID seems to play an important role in promoter identification, recruitment of the PIC, and thus, transcription.

1.1b TFIID

The general transcription factor, TFIID, is a multiprotein complex comprised of the TATA-binding protein (TBP) and 13 TATA-binding protein Associated Factors (termed TAF 1-13). TBP and the TAF cofactors were found to be necessary for transcription (Dynlacht et al., 1991; Tanese et al., 1991). Within TFIID, how the TAFs interact was further elucidated upon cryo-electron microscopy (Louder et al., 2016). They determined that TAF8 interacts with TAF2, aiding in incorporation into TFIID. As well, TAF2 and TAF8 interact with TAF6 (Louder et al., 2016). Further TAF1 is known to interact with TAF2 to promote DNA binding (Chalkey and Verrijzer, 1999) and is also known to interact with TAF7 (Wang et al., 2014) (discussion in section 1.3b/c). The remaining TAFs and their association to one another still remains to be determined. Louder et al., further show evidence that TAFs are contact points for other members of the GTM such as TFIIF and TFIIE, indicating TAFs may be responsible for PIC recruitment.

Further, TFIID is known to recognize and bind to many components of a promoter, a fixed sequence that identifies the transcription start site of a gene (Thomas and Chiang, 2006). Recently, Louder et al., via single particle cryo-electron microscopy, determined the structure of TFIID in complex with core promoter DNA. They previously determined that TFIID has two conformations, 'canonical', in which TFIID is unbound and the 'rearranged' state in which TFIID is bound to the core promoter (Cianfrocco.et al., 2013*)*. They suggest that in the

canonical formation the TAND domain of TAF1 interacts with TBP (described in section 1.3b/c), inhibiting TBP's ability to bind to DNA. However, in the rearranged state, TAND and TBP are at opposite ends of TFIID, allowing TBP- and, presumably, TFIID- bind to promoter DNA (Louder et al., 2016).

As the name preludes, TBP recognizes the TATA-containing promoters, however, TAFs recognize other core promoter elements (covered in section 1.2). It is important to note that while TFIID was found to bind to TATA elements *in vitro*, *in vivo* studies suggest that TFIID primarily binds TATA-less promoters (Basehoar et al., 2004; Kim et al., 2005). Our data, discussed in Chapter 3, shows that TAF1, the largest subunit of TFIID (Kokubo et a., 1993), is found to bind to primarily TATA-less promoters, supporting these previous findings.

1.2 Core promoter elements

Core promoter elements are fixed sequences within the DNA that mark the site of transcription. These elements are recognized by TFIIB and members of TFIID, which allows for the recruitment of the PIC to begin transcription. The seven known core promoter elements, their sequences, location, and the protein(s) found to bind are outlined in Figure 1.1.

Core promoter elements vary in composition at a given promoter, ranging from a single core promoter element to all seven core promoter elements (Thomas and Chiang, 2006). The INR sequence is the most frequent core

promoter element (Ohler et al., 2002; Fitzgerald et al., 2006, Gershonzon et al., 2006) and is found in approximately ~70% of promoters (Carninci et al., 2006; Kim et al., 2005; Cooper et al., 2006). Further, it is the site of transcription initiation and is primarily recognized by TAF1/TAF2 proteins (Chalkey and Verrijzer, 1999). TFIIB is found to recognize BRE sequences upstream and downstream of the INR (Lagrange et al., 1998) and are believed to be present in approximately 30% of human promoters (Jin et al., 2006). Recently, TAF1 has been determined to bind to the motif ten element (MTE) which is found in approximately 64% of promoters (Jin et al., 2006). The DPE is recognized by TAF6/TAF9 and TAF1 (Burke and Kadonaga, 1997; Louder et al., 2016) and is estimated to be present in ~20% of promoters (Jin et al., 2006). Lastly, the TBP is found to recognize TATA elements, which are found to be present in 10-30% of all human promoters (Carninci et al., 2006; Suzuki et al., 2001).

1.2a Core promoter elements and the regulation of gene expression

It is believed that the composition of core promoter elements can dictate the rate and intensity of transcription (Bae et al., 2015; Juven-Gershon and Kadonaga, 2010). Through studies of the Caudal protein in *Drosophila* (Juven-Gershon et al., 2008), it was shown that when a DPE is present, there was strong activation of targets via Caudal. However, there only was weak activation in the presence of a TATA element and this was further reduced in the presence of the BREu. In this specific example, it goes to show that certain genes are

strongly regulated via the presence of certain core promoter elements (Juven-Gershon and Kadonaga, 2010).

The TATA element, specifically, is believed to play an important role in stress response (Bae et al., 2015) and can influence the intensity of gene expression (Morachis et al., 2009). The TATA element has been found to be enriched in stress related genes, whereas growth-promoting genes are often TATA-less (Bae et al., 2015), indicating a potential role of TATA in response to stress. This indicates that the composition of core promoter elements and the factors responsible for recognizing core promoter elements, may play a vital role in the DNA damage response.

TFIID binding to a TATA only promoter is TBP dependent. However, TFIID with a mutated TBP is still able to bind to TATA-less promoters when an INR sequence is present (Marinez et al., 1994). This suggests that there is an interplay between TATA elements and INR elements and thus further emphasizes the importance of TBP and TAF1/TAF2 proteins in the regulation of gene expression (Wasserman and Sauer, 2000).

While TFIID binding to the core promoter has been thoroughly shown, whether the composition of core promoter elements bound by GTM changes in response to stress requires further investigation.

1.3 TAF1

TAF1 is the largest subunit of TFIID (Kokubo et a., 1993) and weighs approximately 250 kDa. There are 12 known splice variants of TAF1 produced by alternative splicing and 15 natural variants (The UniProt Consortium, 2017). While the gene name, TAF1, is the most common name used, this gene is also known as BA2R, CCG1, CCGS, TAF2A and is located on chromosome X. Further, TAF1 protein is alternatively named p250 or TAFII250 (The UniProt Consortium, 2017). Due to TAF1's role in transcription as a member of TFIID, it makes sense that TAF1 is found only in the nucleus (Sekiguchi et al., 1991). TAF1 has many identified post translational modifications via mass spectrometry, including 35 known phosphorylation sites, 2 ubiquitin sites, 4 acetylation sites, and 1 sumoylation site (Hornbeck, 2015; PhosphoSitePlus).

It has been determined that TAF1 is an essential gene for most eukaryotic organisms. The importance of TAF1 is evident since TAF1 inactivation in yeast cells resulted in G1 arrest and TAF1 null mutations in *Drosophila* resulted in lethality in early development (Talavera and Basilico, 1977; Nishimoto et al., 1982; Walker et al., 1996; Wassarman et al., 2000). TAF1 function was further investigated, via microarray analysis, and it was revealed that upon TAF1 inactivation, that genes involved in cell cycle and growth regulation were TAF1 dependent (Wang and Tjian, 1994; Walker et al., 1997; O'Brien and Tjian, 2000).

A more in-depth look into TAF1 structure, enzymatic activities, and interactions with other proteins provide a clearer understanding of TAF1 function (covered in section 1.3a-e).

1.3a TAF1 and core promoter elements

Since most promoters are TATA-less, the PIC relies on TAF proteins, rather than TBP, to direct assembly properly on the promoter. As stated previously, it has been determined that TAF1 binds to core promoter elements INR (Chalkey and Verrijzer, 1999) and DPE (Louder et al., 2016). TAF1 is the largest subunit of TFIID (Kokubo et a., 1993) and is found to interact with other members of TFIID. Specifically, it has been shown that TAF1 is in complex with TAF2 when interacting with promoter DNA (Chalkey and Verrijzer, 1999). More recent studies have shown that TAF1 interacts with downstream promoter elements (DPE) (Louder et al., 2016) and the importance of this TAF1 binding with DPE was investigated. TAF1 is found to bind to positions -3 to +31 of the TSS through TAF1 winged helix (WH) domain (discussed further in section 1.3b). Since TAF1 was found to bind to downstream promoter elements as part of TFIID and the PIC, it makes sense that for RNA-pol II to transcribe genes, TAF1 must release from the DPE. This release of TAF1 is thought to lead to the isomerization of TFIID allowing for the activation of transcription (Louder et al., 2016). Therefore, TAF1 plays a clear role in TFIID recruitment, PIC orientation to the promoter, and the regulation of transcription initiation.

1.3b TAF1 structure

TAF1 contains bipartite kinase domains (N-terminal kinase (NTK) and C-terminal kinase (CTK)), HAT domain, E1/E2 domain, DNA binding abilities, and a double bromodomain (DBD) (Wassarman and Sauer, 2000) (Figure 1.2). TAF1 further has a domain DUF3591 which comprises TAF1 HAT domain and winged helix domain (WH). These domains are described next and how these domains are involved in the modifications of GTM, histones, and activators are described in sections 1.3c-1.3e (respectively).

Bipartite kinase domains:

There are two independent serine/threonine protein kinase domains of TAF1, the NTK and CTK (Dikstein et al., 1996; O'Brien and Tjian, 1998). The NTK is located approximately between positions 1-414 (a.a) and the CTK is located approximately located between positions 1425-1872 (a.a) (Figure 1.2). Phosphorylation via TAF1 has been demonstrated *in vitro* in *Drosophila*, human, and yeast proteins (Wassarman and Sauer, 2000). Further kinase activity has been shown *in vivo* in the human cell line, U2OS (Li et al, 2004; Wu et al., 2014; discussed further section 1.4). These kinase domains are responsible for both TAF1 autophosphorylation and transphosphorylation of other proteins both *in vitro* and *in vivo* (*in vitro*; Dikstein et al., 1996; Solow et al., 2001; *in vivo*; Wu et al., 2014; Maile et al., 2004).

While few targets of the NTK and CTK have been identified (Figure 1.2) (discussed in sections 1.3c/e, and 1.4), their importance has been evident. CTK deletion results in lethality in Drosophila (V. Jo, J. G. Shanklin, E. M. Schlag and D.A.W., unpublished) and when NTK is deleted, TAF1 cannot rescue phenotypes previously shown, including cell cycle and cell growth regulation in t13, a hamster cell line (O'brien and Tjian, 1998; O'Brien and Tjian, 2000).

Double bromodomains:

The double bromodomains (DBD) are located at approximately 1397-1467 and 1520-1590 (a.a) (Figure 1.2). These double bromodomains are responsible for recognizing acetylated lysine residues on other proteins (Dhalluin et al., 1999a; Jacobson et al., 2000a). It has been shown that the DBD recognizes acetylated histones, a marker for transcription activation (Jacobson et a., 2000) (discussed in section 1.3d). Moreover, TAF1 BDB recognizes bi-acetylated p53, an activator, recruiting TAF1 to the promoter (Li et al., 2007) (discussed in section 1.4). This emphasizes the importance of the DBD in TFIID recruitment for transcriptional activation.

HAT domain:

TAF1 was found to contain HAT activity (Mizzen et al., 1996) and the domain is located at approximately 517 -976 (a.a) (Figure 1.2). TAF1 has the ability to acetylate both histones (Mizzen et a., 1996; Wassarman et al., 2000)

(discussed in section 1.3d) and general transcription machinery (Imhof et al., 1997) (discussed in section 1.3c) as shown in *in vitro* studies. A study done in vivo, in the hamster cell line ts13, showed that a mutation in the HAT domain (G716D) led to late G1 arrest and reduced cyclin D1 transcription (Dunphy et al., 2000). Thus, suggesting that the HAT activity of TAF1 was also necessary *in vivo.*

Ubiquitin conjugating domain:

TAF1 contains ubiquitin conjugating activity (Pham and Sauer, 2000). The E1/E2 domain of TAF1 is believed to be between amino acids 738-1259 (Tavassoli et al., 2010). TAF1 was found to monoubiquitinate histone 1 using a drosophila extract in *in vitro* studies. When this domain was mutated and was inactivated, there was less ubiquitination of histone H1 *in vitro*. Using this mutant construct in *Drosophila* embryos via *in situ* hybridization, they witnessed a decrease in genes normally upregulated by TAF1 ubiquitin conjugating activity (Pham and Sauer, 2000). This data suggests that TAF1 ubiquitin conjugating activity may be involved in gene activation.

DUF3591 domain:

The DUF3591 domain is located between amino acids 600 and 1109 (Louder et al., 2016). This domain of TAF1 has been found to be highly conserved between yeast and human. Further, this domain is found to

encompass a winged helix (WH) domain responsible for DNA binding activity and interactions with TAF7 (discussed in section 1.3b/c). Since this domain overlaps with the HAT domain, this domain also processes HAT activity (Figure 1.2).

The combination of these domain structures of TAF1 allows TAF1 to act as a master regulator of transcription. This includes TAF1's regulation through modifications of histones, modifications of the general transcription machinery and modifications of activators.

1.3c TAF1 interaction and modification of GTM

TAF1 acts genome-wide as a transcriptional regulator and its functions are carefully orchestrated via many variables. TAF1 is found to interact with, and modify other members of the general transcription machinery as a member of the TFIID complex.

While TAF1 is a well-known member of TFIID, how TAF1 interacts with other members is still poorly understood. Recently, a crystal structure for how TAF1 interacts with TAF7 has been determined (Wang et al., 2014). In this study, Wang et al., showed that the WH domain of TAF1 is necessary for DNA binding and the WH domain along with the heterodimeric triple barrel interact with TAF7. While a TAF1 mutant (G716D) in the hamster cell line (ts13) leads to cell cycle arrest and downregulation of genes (Dunphy et al., 2000), the mechanism behind this was still poorly understood. This mutation (G716D), is located between

TAF1's winged helix (WH) domain and heterodimeric triple barrel in the DUF3591 domain (Wang et al., 2014). This experiment could explain the effects of the G716D TAF1 mutation which led to cell cycle arrest and downregulation of its target genes. This implicates the DUF3591 domain as an important domain for TAF1 function. Further, TAF1 interaction with TAF7 is thought to inhibit TAF1 HAT activity (Gegonne, et al., 2001; Chiang and Roeder,1995). Inhibiting TAF1 HAT activity may lead to the inhibition histone acetylation, lowering transcription activity (discussed in section 1.3d).

TAF1 N-terminal contains two domains, which can interact with TBP; TAND1 and TAND2. TBP binds to TATA-box elements, facilitated via TFIIA, to nucleate the PIC and initiate transcription. TAND1 inhibits TBP DNA binding domain (Liu et al., 1998) and TAND2 inhibits binding of TFIIA (Kokubo et al., 1998). Therefore, TAF1 can act as a negative regulator of transcription through N-terminal interactions.

In addition to TAF1 interacting with other members of the general transcription machinery, TAF1 is also found to modify these members. Through TAF1's NTK, TAF1 has been shown to phosphorylate both TFIIFalpha (RAP74) and TFIIA *in vitro* (Wu et al., 2014; Ruppert and Tjian, 1995; Dikstein et al., 1996; Solow et al., 2001). RAP74, when hyperphosphorylated, leads to an increase in transcription elongation in vitro. Therefore, it is reasonable to suggest that TAF1 phosphorylation of RAP74 contributes to transcription elongation. TAF1 phosphorylation of TFIIA encourages the formation of a TFIIA-TBP-TATA

element complex. This would lead to an increase in nucleation of the PIC-allowing for increase in transcription. TAF1 has also been shown to acetylate RAP74 and TFIIE *in vitro* (Imhof et al., 1997), however, further research into the effects *in vivo* still remains to be elucidated.

These findings further support the theory of TAF1 involvement in DNA binding and nucleation of the PIC leading to the activation of transcription.

1.3d TAF1 interaction and modification of histones

TAF1 has also been shown act as a coactivator mediating transcriptional activation through histone-modifying activities. TAF1 double bromodomains can recognize acetylated K14 of H3 and K5, K8, K12, or K16 of H4 (Jacobson et al., 2000), thereby recruiting TAF1 to acetylated histones, a marker of activation. The HAT domain of TAF1 can then multiply the acetylation of the histones (Jacobson et al., 2000). TAF1 can acetylate histone H3 and H4 (Mizzen et. al., 1996) and ubiquitinate linker histone H1 (Pham and Sauer, 2000). This acetylation leads to a less-positive histone, creating a more relaxed connection between the histone-DNA. This allows TFIID to nucleate/recruit the remaining initiation complex for transcription activation. These activities suggest that TAF1 interaction with histones may play a significant role in genome-wide transcriptional activation.

1.3e TAF1 regulation through activators

TAF1 also shows an ability to interact with activators which can increase promoter occupancy and thus control TAF1 regulatory activities (Pugh and Tjian, 1990; Dynlacht et al., 1991). TAF1's physical interaction with an activator tethers TFIID to a particular promoter leading to the regulation of gene expression. For example, TAF1's direct interacting activators include HIV Tat (Weissman et. al., 1998), Adenovirus E1A (Geisberg et al., 1995), and Herpes simplex virus type 1 ICP4 (Carozza and DeLuca, 1996). Recently, studies have shown that TAF1 is involved in recognition and regulation of p53, an important tumor suppressor protein (Li et al., 2004; Wu et al., 2014) (discussed in section 1.4). These interactions lead to the recruitment of the PIC allowing for activation of transcription. However, TAF1 can interact with retinal blastoma (RB) protein, this leads to an inhibition of TAF1 kinase activity (Siegert and Robbins, 1999; Solow et al., 2001). This inhibition can be overcome via E1A and cyclin D1 (Siegert et al., 2000). These data suggest that activators can regulate TAF1 activity and thus affect transcriptional activities of TFIID. As well, TAF1 has also shown to be involved in regulating the activator, as discussed further in section 1.4 and Chapter 2.

1.4 Interplay between TAF1 and p53 in response to DNA damage

1.4a TAF1 recognition of acetylated p53

The double bromodomain of TAF1 has been found to recognize bi-acetylated lysine residues on p53 (Li et al., 2007). p53 is a tumor suppressor protein involved in cellular response to DNA damage (Vousden and Prives, 2009; Kruse and Gu, 2009). It has been shown that TAF1 interacts with the C-terminal domain of p53 (Li et al., 2004). Upon DNA damage, p53 becomes phosphorylated and acetylated which leads to p53 stabilization and accumulation in the nucleus, allowing for p53 to transcriptionally regulate genes to prevent cell proliferation (Bode and Dong, 2014). p53 becomes acetylated upon DNA damage at K305, K320, K370, K372, K373. K382 and K386 (Bode and Dong, 2004). It was previously reported that upon DNA damage, TAF1 recognizes and binds to p53 through acetylated K373 and K382 via its double bromodomain (Li et. al., 2007). It is this recognition of p53 acetylated lysine's that recruits TAF1 to the p21 promoter. p21 is responsible for binding to and inhibiting all currently known cyclin-dependent protein kinases that are required for G1-S phase transition, thereby enforcing cell cycle arrest. When either the double bromodomain of TAF1 or K373 and k382 of p53 were mutated, this recruitment was no longer observed (Li et a., 2007).

1.4b TAF1 regulation of p53 in response to UV via kinase activity

Ultra violet (UV) radiation induces DNA damage (Sinha and Häder, 2002; Review). Specifically, the most common types of UV-induced DNA damage are cyclobutane-pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs), pyrimidine adducts (Clingen et al., 1995; Mitchell and Karentz, 1993). To protect against UV-induced DNA damage, our cells have adapted a robust system of defense including TAF1's involvement in transcriptional regulation following DNA damage. TAF1 was found to regulate transcription factor p53, a quintessential tumor suppressor protein that aids in maintaining genomic integrity in response to cell stress.

It was determined that at early stages of UV-induced DNA damage, acetylated p53 recruits TAF1 to the p21 promoter to activate transcription of p21, inducing cell cycle arrest. However, during the later stages of DNA damage, p53 is phosphorylated at Thr55 by TAF1 kinase activity, releasing p53 from the p21 promoter, terminating transcription, and allowing persistence of the cell cycle (Wu et. al., 2014) (Figure 1.3). This regulation via TAF1 kinase activity was further verified through the usage of apigenin: a TAF1 kinase inhibitor. In the presence of the inhibitor, TAF1 was unable to phosphorylate p53 at Thr55, thereby not releasing p53 from the p21 promoter, causing transcription of p21 to continue. Further, this phosphorylation of p53 at Thr55 leads to p53 degradation (Li et al., 2004). Therefore, these results suggest that TAF1 phosphorylation of p53 in a timely manner is necessary for the proper response to cell stress.

This example elucidates that while TAF1 is responsible for activating transcription through its participation in the initiation complex, TAF1 is also responsible for ceasing transcription. Since TAF1 is a general transcription factor bound to many promoters, TAF1 may also target transcription factors other than p53. While TAF1's regulation of p53 in response to UV induced damage has been determined, other possible transcription factors under this same mode of regulation have yet to be elucidated.

1.4c Kinase activity is cellular [ATP] dependent

As previously mentioned, TAF1 has been shown to be recruited by p53 on the p21 promoter via the recognition of acetylated lysine's in early stages of DNA damage. This leads to the transcriptional activation of p21, ceasing the cell cycle. In later stages of DNA damage, TAF1 phosphorylates p53, leading to p53 dissociation from the p21 promoter and subsequent transcription termination and degradation (Wu et al., 2014). This timely regulation of p53 by TAF1 is crucial for the proper response to DNA damage, to allow for genomic integrity and cell recovery. This prompts a question, however, of 'how does TAF1 know when to phosphorylate p53?'.

TAF1's ability to act as this molecular switch for activation/termination of transcription relies on its unique feature of its kinase activity being cellular ATP dependent. U2OS cells under basal conditions exhibit a cellular ATP concentration of 1.13 +/- 0.01 mM, however, TAF1 kinase has a Km of 1.9 mM

(Wu et al., 2014). Upon DNA damage induced by UV radiation, ATP levels drop initially, rendering the kinase activity of TAF1 inactive. Responsible for this drop of ATP is PARP-1, a nuclear enzyme responsible for adding ADP-ribose units on Glu residues of acceptor proteins. PARP-1 uses NAD+ as a substrate to modify numerous proteins, which leads to the depletion of cellular ATP (Schreiber et al., 2006). However, in later stages of DNA damage this depletion in ATP leads to the activation of AMPK, which allows for ATP recovery (Wu et al., 2014). AMPK detects and reacts to the AMP:ATP ratio by inhibiting energy consuming pathways and stimulating energy producing pathways (Hardie et al, 2007). In this later stage of DNA damage, cellular ATP levels show ~2.5 fold increase (Wu et al., 2014). This ATP recovery leads to TAF1 kinase activation, thereby phosphorylating target protein, p53, releasing p53 from the promoter and ceasing transcription (Wu et. al., 2014) (Figure 1.3). While this mechanism of regulation of p53 in response to DNA damage has been well described, other possible TAF1 targets under this same mechanism have not been determined. Since TAF1 is a general transcription factor bound to many promoters, TAF1 may negatively regulate other transcription factor targets in a similar manner to p53 via its kinase activity. A genome-wide analysis of TAF1 regulation in response to DNA damage still remains to be elucidated.

1.5 Genomic integrity and disease

1.5a TAF1 mutations in cancer

Due to TAF1's important role in transcription regulation, cell cycle progression, as well as response to DNA damage, it is not surprising to find TAF1 as one of the most mutated transcription factors in cancers (Kandoth et al., 2013) (Table 1.1). A large systematic analysis was performed on 3,281 tumors from 12 different cell types and found 127 significantly mutated genes (SMG). Of these significantly mutated genes was TAF1. TAF1 was one of the highest ranked transcription factor SMG's, with the highest mutations found in uterus corpus endometrial carcinoma (UCEC), lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC) (Kandoth et al., 2013) (Table 1.1).

TAF1 has 155 known mutations including 119 missense mutations, 24 synonymous mutations, and 12 truncation mutations (Rubio-Perez et al., 2015; Gonzalez-Perez, 2013). Whole-exome sequencing of 57 uterine serous carcinoma (USC), identified TAF1 as one of the most frequently mutated genes. Of the mutations in TAF1, most were identified in the HAT domain (Zhao et al., 2013). Further, TAF1 was reported to be overexpressed in human lung and breast carcinoma and was associated with high mitotic activity and poor tumor differentiation (Wada et al., 1992).

1.5b TAF1 mutations in disease

In addition to cancers, TAF1 mutations are also found in other diseases. For example, TAF1 is thought to contribute to X-linked dystonia-parkinsonism (Herzfeld et al., 2013; Makino et al., 2007; Domingo et al., 2015). This disease is fatal (Nolte et al., 2003) and is characterized by adult-onset dystonia and parkinsonism. More recently, TAF1 variants have been identified as leading to dysmorphic features, neurological manifestations, and intellectual disability, known as X-linked intellectual disability syndrome (O'Rawe et al., 2015). Dysmorphic features include sagging cheeks, low-set and protruding ears, a long face, a pointed chin, among others. TAF1 sequence variants were identified in all 14 affected individuals examined. Most of these variants were missense variants while few were duplication variants. Half of the missense variants were found in domain DUF3591, which is important for HAT activity, TAF7 interaction, and DNA binding activity. Indicating that these functions of TAF1 are necessary to avoid X-linked dystonia-parkinsonism (O'Rawe et al., 2015).

1.6 Summary of Chapters

In Chapter 2, I investigate TAF1's genome-wide role in transcriptional regulation via it's kinase activity. Following the discovery of TAF1's regulation of p53 in response to DNA damage, I aimed to identify other transcription factors and promoters under a similar mode of regulation via TAF1. To accomplish this, I

describe a pipeline utilizing bioinformatics alongside experimental verification. We identify novel TAF1 kinase targets following UV-induced DNA damage.

To further investigate the importance of TAF1 in response to DNA damage, Chapter 3 investigates the role of core promoter elements in TAF1 recruitment to promoters following UV- induced DNA damage. To identify core promoter elements, I utilized sequence specific searches in TAF1 bound peaks. Following UV-induced DNA damage, we witnessed an increase in TATA element abundance in promoters where TAF1 was recruited to and whose gene expression was significantly increased.

1.7 References

1. Bae, S.-H., Han, H.W., and Moon, J. (2015). Functional analysis of the molecular interactions of TATA box-containing genes and essential genes. *PloS One* 10, e0120848.

2. Basehoar, A.D., Zanton, S.J., and Pugh, B.P. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* 116: 699-709. [PubMed: 15006351]

3. Bode, A.M., and Dong, Z. (2004). Post-translational modification of p53 in tumorigenesis. 4, 793.

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

5. Burke TW, Kadonaga JT. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev* 1996;10:711–724. [PubMed: 8598298]

6. Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A.M., Taylor, M.S., Engström, P.G., Frith, M.C., et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nature Genetics 38, 626–635.

7. Carrozza, M., and N. DeLuca. "Interaction of the Viral Activator Protein ICP4 with TFIID through TAF250." Molecular and Cellular Biology16 (1996): 3085-093.

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

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

10. Cianfrocco, M.A. *et al.* Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell* **152,** 120–131 (2013).

11. Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Res 2006;16:1–10. [PubMed: 16344566]

12. D. L. Mitchell and D. Karentz, The induction and repair of DNA photodamage in the environment, in Environmental UV Photobiology, eds. A. R. Young, L. Björn, J. Moan and W. Nultsch, Plenum, New York, 1993, pp. 345–377.

13. Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., Zhou, M.-M., and Zhou, M.-M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature 399, 491–496.

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

15. Domingo A., Westenberger A., Lee L.V., Brænne I., Liu T., Vater I., Rosales R., Jamora R.D., Pasco P.M., Cutiongco-Dela Paz E.M. New insights into the genetics of X-linked dystonia-parkinsonism (XDP, DYT3) Eur. J. Hum. Genet. 2015;23:1334–1340.

16. Dunphy E.L., Johnson T., Auerbach S.S., Wang E.H., Johnson T., Auerbach S.S., Wang E.H., Auerbach S.S., Wang E.H., Wang E.H. Requirement for TAF(II)250 acetyltransferase activity in cell cycle progression. Mol. Cell. Biol. 2000;20:1134–1139

17. Dynlacht, Brian David, Timothy Hoey, and Robert Tjian. "Isolation of Coactivators Associated with the TATA-binding Protein That Mediate Transcriptional Activation." Cell 66.3 (1991): 563-76.

18. Fire, A., Samuels, M., and Sharp, P.A 1984. Interactions between RNA polymerase II, factors, and template leading to accurate transcription. J Biol Chem 259: 2509-2516.

19. FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. Comparative genomics of Drosophila and human core promoters. Genome Biol 2006;7:R53. [PubMed: 16827941]

20. Flores, O., Lu, H., and Reinberg, D. 1992. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. J Biol Chem 267:2786-2793.

21. Flores, O., Maldonado, E., and Reinberg, D. 1989. Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. J Biol Chem 264: 8913-8921.

22. Ge, H., Martinez, E., Chiang, C.-M., and Roeder, R.G. 1996. Activator-dependent transcription by mammalian RNA polymerase II: in vitro reconstitution with general transcription factors and cofactors. Methods Enzymol 274:57-71.

23. Gegonne, A., Weissman, J.D., and Singer, D.S. (2001). TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. Proc. Natl. Acad. Sci. USA 98, 12432–12437.

24. Geisberg, J., J. Chen, and R. Ricciardi. "Subregions of the Adenovirus E1A Transactivation Domain Target Multiple Components of the TFIID Complex." Molecular and Cellular Biology 15.11 (1995): 6283-290.

25. Gershenzon NI, Trifonov EN, Ioshikhes IP. The features of Drosophila core promoters revealed by statistical analysis. BMC Genomics 2006;7:161. [PubMed: 16790048]

26. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, Santos A & Lopez-Bigas N IntOGen-mutations identifies cancer Nature Methods 2013; doi:10.1038/nmeth.2642

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

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

29. Herzfeld T., Nolte D., Grznarova M., Hofmann A., Schultze J.L., Müller U. X-linked dystonia parkinsonism syndrome (XDP, lubag): disease-specific sequence change DSC3 in TAF1/DYT3 affects genes in vesicular transport and dopamine metabolism. Hum. Mol. Genet. 2013;22:941–951.

30. Hornbeck PV, et al (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 43:D512-20.

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

32. Jacobson, R. H. "Structure and Function of a Human TAFII250 Double Bromodomain Module." Science 288.5470 (2000): 1422-425.

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

34. Jin, V.X., Singer, G.A., Agosto-Pérez, F.J., Liyanarachchi, S., and Davuluri, R.V. (2006). Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs. BMC Bioinformatics 7, 114.

35. Juven-Gershon T, Hsu J-Y, Kadonaga JT. Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. Genes Dev 2008;22:2823–2830.

36. Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Developmental Biology 339, 225–229.

37. Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature 502, 333–339.

38. Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. A high- resolution map of active promoters in the human genome. Nature 2005;436:876–880.

39. Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A. Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. 2005. A high-resolution map of active promoters in the human genome. Nature 436: 876-880.

40. Kokubo, T., Swanson, M.J., Nishikawa, J.-I., Hinnebusch, A.G., and Nakatani, Y. 1998. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. Mol Cell Biol 18:1003– 1012.

41. Kruse, Jan-Philipp, and Wei Gu. "Modes of P53 Regulation." Cell 137.4 (2009): 609-22.

42. 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 Dev 1998;12:34–44. [PubMed: 9420329]

43. Li, A.G., Piluso, L.G., Cai, X., Gadd, B.J., Ladurner, A.G., and Liu, X. (2007). An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment. Molecular Cell 28, 408–421.

44. Li, Heng-Hong, Andrew G., Hilary M. Sheppard, and Xuan Liu. "Phosphorylation on Thr-55 by TAF1 Mediates Degradation of P53."Molecular Cell 13.6 (2004): 867-78.

45. Li, Heng-Hong, Andrew G., Landon G. Piluso, Xin Cai, Brian J. Gadd, Andreas G. Ladurner, and Xuan Liu. "An Acetylation Switch in P53 Mediates Holo-TFIID Recruitment." Molecular Cell 28.3 (2007): 408-21.

46. Liu, D.,Ishima,R.,Tong,K.I.,Bagby,S.,Kokubo,T.,Muhandiram,D.R., Kay, L.E., Nakatani, Y., and Ikura, M. 1998. Solution structure of a TBP-TAFII230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. Cell 94:573–583

47. Louder, R.K., He, Y., López-Blanco, J.R., Fang, J., Chacón, P., and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature 531, 604–609.

48. Maile, T., Kwoczynski, S., Katzenberger, R.J., Wassarman, D.A., and Sauer, F. 2004. TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. Science 304:1010–1014.

49. Makino S., Kaji R., Ando S., Tomizawa M., Yasuno K., Goto S., Matsumoto S., Tabuena M.D., Maranon E., Dantes M. Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. Am. J. Hum. Genet. 2007;80:393–406.

50. Mizzen, Craig A., Xiang-Jiao Yang, Tetsuro Kokubo, James E. Brownell, Andrew J. Bannister, Tom Owen-Hughes, Jerry Workman, Lian Wang, Shelley L. Berger, Tony Kouzarides, Yoshihiro Nakatani, and C.david Allis. "The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity." Cell 87.7 (1996): 1261-270.

51. Morachis, J.M., Murawsky, C.M., and Emerson, B.M. (2010). Regulation of the p53 transcriptional response by structurally diverse core promoters. Genes & Development 24, 135–147

52. Nishimoto, T., Sekiguchi, T., Kai, R., Yamashita, K., Takahashi, T. and Sekiguchi, M. (1982). Large-scale selection an danalysis of temperature-sensitive mutants for cell reproduction from BHK cells. Somatic cell genet. 8, 811-824.

53. Nolte D., Niemann S., Müller U. Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. Proc. Natl. Acad. Sci. USA. 2003;100:10347–10352.

54. O'Brien, T. and Tjian, R. (1998). Functional analysis of the human TAFII250 N-terminal kinase domain. Mol. Cell 1, 905-911.

55. O'Brien, T. and Tjian, R. (2000). Different functional domains of TAFII250 modulate expression of distinct subsets of mammalian genes. Proc. Natl. Acad. Sci. USA 97, 2456-2461.

56. O'Rawe, J.A., Wu, Y., Dörfel, M.J., Rope, A.F., Au, P.Y.B., Parboosingh, J.S., Moon, S., Kousi, M., Kosma, K., Smith, C.S., et al. (2015). TAF1 Variants Are Associated with Dysmorphic Features, Intellectual Disability, and Neurological Manifestations. The American Journal of Human Genetics 97, 922–932.

57. Ohler U, Liao GC, Niemann H, Rubin GM. Computational analysis of core promoters in the Drosophila genome. Genome Biol 2002;3:RESEARCH0087. [PubMed: 12537576]

58. P. H. Clingen, C. F. Arlett, L. Roza, T. Mori, O. Nikaido and M. H. L. Green, Induction of cyclobutane pyrimidine dimers, pyrimidine (6-4)pyrimidone photoproducts, and Dewar valence isomers by natural sunlight in normal human mononuclear cells, Cancer Res., 1995, 55, 2245–2248.

59. Pham, A.-D., and F. Sauer. "Ubiquitin-Activating/Conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila." Science 289.5488 (2000): 2357-360.

60. Reinberg, D. and Roeder, R.G 1987. Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. J Biol Chem 262:3310-3321.

61. Rubio-Perez, C., Tamborero, D., Schroeder, MP., Antolín, AA., Deu-Pons,J., Perez-Llamas, C., Mestres, J., Gonzalez-Perez, A., Lopez-Bigas, N. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals novel targeting opportunities. Cancer Cell 27 (2015), pp. 382-396

62. Ruppert, S., and R. Tjian. "Human TAFII250 Interacts with RAP74: Implications for RNA Polymerase II Initiation." Genes & Development9.22 (1995): 2747-755.

63. Samuels, M., and Sharp, P.A. 1986. Purification and characterization of specific RNA polymerase II transcription factor. J Bio Chem 261: 2003-20013.

64. Sawadogo and Roeder, R.G. 1985. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. Proc Natl Acad Sci USA 82:4394-4398.

65. Schreiber, Valérie, Françoise Dantzer, Jean-Christophe Ame, and Gilbert De Murcia. "Poly(ADP-ribose): Novel Functions for an Old Molecule."Nature Reviews Molecular Cell Biology 7.7 (2006): 517-28.

66. Sekiiguchi T., Nohiro Y., Nakamura Y., Hisamoto N., Nichimoto T. (1991) The human CCG1 gene, essential for pression of the G1 phase, encodes a 210-kilodalton nuclear DNA-binding protein. Mol. Cell. Bio. 11:3317-3325.

67. Sinha, R.P., and Häder, D.-P. (2002). UV-induced DNA damage and repair: a review. Photochemical & Photobiological Sciences 1, 225–236.

68. Solow, S., Salunek, M., Ryan, R. and Lieberman, P. M. (2001). TAFII250 phosphorylates human TFIIA on serine residues important for TBP binding and transcription activity. J. Biol. Chem. 276, 15886-15892.

69. Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., and Nakamura, Y. (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Research 11, 677–684.

70. Talavera, A. and Basilico, C. (1997). Temperature sensitive mutants of BHK cells affected in cell cycle progression. J. Cell. Physiol. 92, 425-436.

71. Tavassoli, P., Wafa, L. A., Cheng, H., Zoubeidi, A., Fazli, L., Gleave, M., Rennie, P. S. (2010). TAF1 Differentially Enhances Androgen Receptor Transcriptional Activity via Its N-Terminal Kinase and Ubiquitin-Activating and -Conjugating Domains. Molecular Endocrinology, 24(4), 696–708.

72. The UniProt Consortium (2017). UniProt: the universal protein knowledgebase. Nucleic Acids Research 45, D158–D169.

73. Thomas, M.C., and Chiang, C.-M. (2006). The General Transcription Machinery and General Cofactors. Critical Reviews in Biochemistry and Molecular Biology 41, 105–178.

74. Vousden, Karen H., and Carol Prives. "Blinded by the Light: The Growing Complexity of P53." Cell 137.3 (2009): 413-31.

75. Wada C, Kasai K, Kameya T, Ohtani H (1992) A general transcription initiation factor, human transcription factor IID, overexpressed in human lung and breast carcinoma and rapidly induced with serum stimulation. Cancer Res 52(2):307–313.

76. Walker, S. S., Reese, J.C., Apone, L.M. and Green, M.R. (1996). Transcription activation in cells lacking TAFIIS. Nature 383, 185-188.

77. Walker, S. S., Shen, W.-C., Reese, J.C., Apone, L. M. and Green, M.R. (1997). Yeast TAFII145 required for transcription of G1/S cyclin genes and regulated by the cellular growth state. Cell 90, 607-614.

78. Wang, E.H. and Tjian, R. (1994). Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by TAFII250. Science 263, 811-814.

79. Wang, H., Curran, E.C., Hinds, T.R., Wang, E.H., and Zheng, N. (2014). Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. Cell Res 24, 1433–1444.

80. Wassarman, D. A., Aoyagi, N., Pile, L. A. and Chlag, E. M. (2000). TAF250 is required for multiple developmental events in drosophila. Proc. Natl. Acad. Sci. USA 97, 1154-1159.

81. Wassarman, D.A., and Sauer, F. (2001). TAFII250. Journal of Cell Science 114, 2895–2902.

82. Weissman, Jocelyn D., Jae Ryoung Hwang, and Dinah S. Singer. "Extensive Interactions between HIV TAT and TAFII250." Biochimica Et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology1546.1 (2001): 156-63.

83. Wu, S.-Y. and Chiange, C.-M. 1998. Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. J Biol Chem 273: 12492-12498.

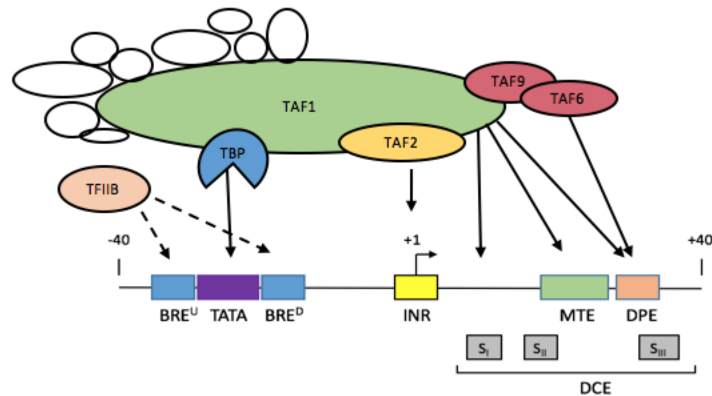
84. Wu, S.-Y., Thomase, M.C., Hou, S.Y., Likhite, V. and Chiang, C.-M. 1999. Isolation of mouse TFIID and functional characterization of TBP and TFIID in mediating estrogen receptor and chromatin transcription. J Biol Chem 274: 23480-23490.

85. Wu, Yong, Joy C. Lin, Landon G. Piluso, Joseph M. Dhahbi, Selene Bobadilla, Stephen R. Spindler, and Xuan Liu. (2014) "Phosphorylation of P53 by TAF1 Inactivates P53-Dependent Transcription in the DNA Damage Response." Molecular Cell 53.1 (2014): 63-74.

86. Zhao, S., Choi, M., Overton, J.D., Bellone, S., Roque, D.M., Cocco, E., Guzzo, F., English, D.P., Varughese, J., Gasparrini, S., et al. (2013). Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma. Proceedings of the National Academy of Sciences 110, 2916–2921.

1.8 Figures and Tables

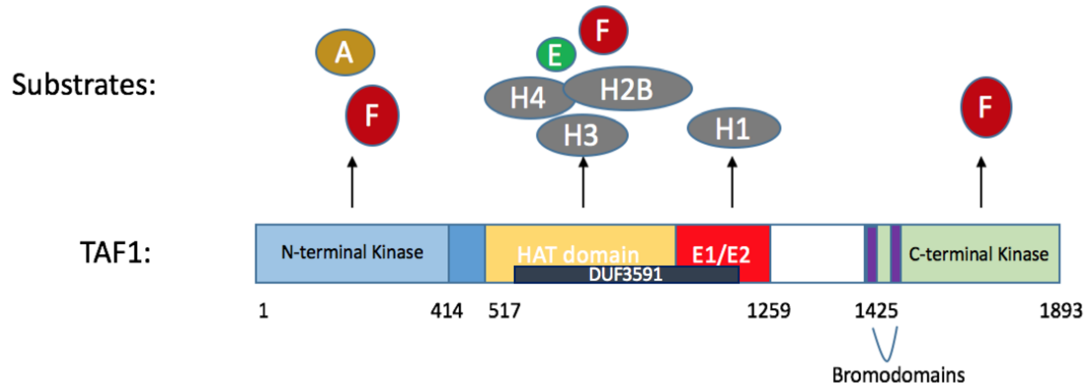
Figure 1.1: Recognition of TFIID and TFIIB to core promoter elements



Core Promoter Element	Position	Consensus sequence (5' to 3')	Bound Protein	Estimated abundancy
BRE^{U}	-38 TO -32	(G/C)(G/C)(G/A)CGCC	TFIIB	~ 30%
TATA	-31 TO -24	TATA(A/T)A(A/T)(A/G)	TBP	~ 10-30%
BRE^{D}	-23 TO -17	(G/A)T(T/G/A)(T/G)(G/T)(T/G)(T/G)	TFIIB	~ 30%
INR	-2 TO +5	PyPyN(T/A)PyPy	TAF1/TAF2	~ 70%
MTE	+18 TO +29	C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C)	TAF1	~ 60%
DPE	+28 TO +34	(A/G)G(A/T)CGTG	TAF6/TAF9 and TAF1	~ 20%
DCE	3 Subelements: +6 to +11 +16 to + 21 +30 to +34	Core sequence: S_{I} CTTC S_{II} CTGT S_{III} AGC	TAF1	n.a

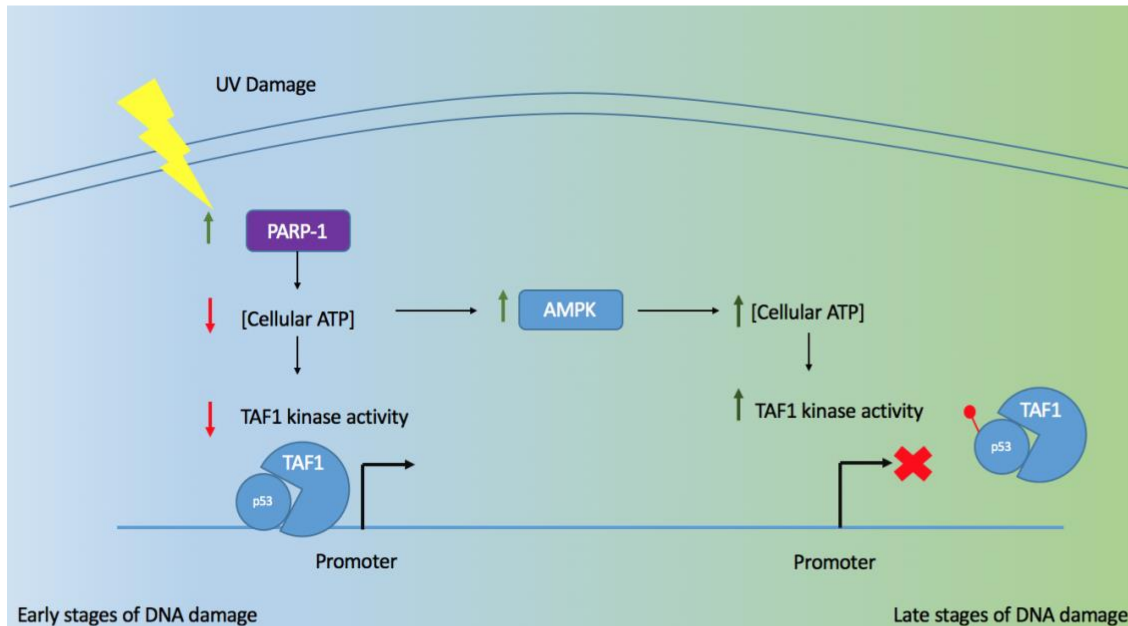
Legend: (Top) Figure representing known general transcription factor binding to core promoter elements. (Bottom) Table of core promoter elements, their known position relative to TSS, consensus sequence, and proteins found to bind. Py = pyrimidines (A or G), N = any base, / = or.

Figure 1.2: TAF1 domains and substrates



Legend: Domains of TAF1 protein and their known substrates. Abbreviations are as follows: TFIIA (A), TFIIF (F), TFIIE (E), Histone 4 (H4) Histone 3 (H3), Histone 2B (H2B), and Histone 1 (H1).

Figure 1.3: TAF1 regulation of p53 in response to UV-induced DNA damage



Legend: TAF1 kinase regulation of p53 following UV damage as described by Wu et al., 2014. In early stages of DNA damage, PARP-1 becomes activated leading to a decrease in cellular ATP concentration. Since TAF1 kinase activity is cellular ATP concentration dependent, when ATP is low, TAF1 kinase activity is null. This allows p53 to remain bound to the promoter allowing for active transcription. In later stages of DNA damage, AMPK senses this drop of ATP and leads to cellular ATP recovery. This increase in cellular ATP concentration activates TAF1 kinase activity allowing TAF1 to phosphorylate p53. This leads to p53/TAF1 dissociation from the promoter, ceasing transcription.

Table 1.1: Top 10 transcription factor SMGs in cancer identified in 12 cancer types.

	BLCA	BRCA	COAD	GMB	HNSC	KRC	AML	LUAD	LUSC	OV	UCEC	Pan-Cancer
VHL	0.0	0.0	0.0	0.0	0.0	52.3	0.0	0.0	0.6	0.0	0.9	6.9
GATA3	1.0	10.6	1.0	0.0	2.0	0.0	0.0	2.6	2.9	0.3	0.4	3.2
TSHZ3	2.0	0.7	3.1	0.7	1.3	1.2	0.5	14.9	6.3	1.0	3.9	2.6
EP300	17.4	0.8	2.1	0.3	8.0	1.4	0.0	0.9	4.6	0.3	5.2	2.5
CTCF	2.0	2.4	1.6	0.0	3.3	0.5	0.5	1.3	0.0	0.3	16.5	2.4
TAF1	2.0	1.1	1.6	1.4	2.3	1.2	0.0	4.0	6.9	1.9	8.7	2.3
TSHZ2	4.1	0.9	3.1	2.4	1.3	0.7	0.0	6.6	3.5	1.0	1.7	1.8
RUNX1	1.0	3.3	1.0	0.0	0.7	0.0	9.0	0.4	0.0	0.0	1.3	1.6
MECOM	5.1	0.5	1.0	1.4	1.7	1.0	0.0	3.5	4.6	0.6	3.0	1.5
TBX3	3.1	2.4	1.0	0.0	0.7	0.0	0.0	4.4	2.9	1.0	1.3	1.4

Legend: Data from Kandoth et al., 2013 depicting top 10 significantly mutated transcription factors across 11 cancer types. Percentages of samples with mutation in each tumor type is shown. Pan-cancer refers to a statistic that includes commonalities and differences across multiple sets of tumor types. Cancer types abbreviated as follows: breast adenocarcinoma (BRCA), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), uterine corpus endometrial carcinoma (UCEC), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), colon and rectal carcinoma (COAD, READ), bladder urothelial carcinoma (BLCA), kidney renal clear cell carcinoma (KIRC), ovarian serous carcinoma (OV) and acute myeloid leukaemia (LAML; conventionally called AML).

Chapter 2: TAF1 as a genome-wide transcriptional regulator in response to DNA damage

The TAF1-ChIP sequencing utilized was performed by Dr. Joy Lin and I was responsible for the analysis. The TAF1 peak overlaps with TAF1 ChIP sequencing from other cell lines via ENCODE was also done by Dr. Joy Lin (Figure 2.1C). The proteins utilized in the kinase assays were purified by Thomas Benedict. I was responsible for all else described in this chapter.

2.1 Abstract

Gene regulation is a complex mechanism and our knowledge of molecular pathways remains limited. In particular, the mechanism behind turning off transcription when it's no longer needed is poorly understood. TAF1, a member of the basal transcription machinery, has been previously implicated in the turning off of p53 transcription of p21 following DNA damage. However, TAF1's role in genome-wide transcription regulation had yet to be elucidated. Since TAF1 is a component of general transcription machinery bound to many promoters, we investigated how TAF1 contributes to the turning-off of overall transcription. Utilizing TAF1 ChIP-Seq and the ENCODE database, we identified TAF1 target transcription factors, E2F1, SP1, FOXM1, and ATF2, and verified their regulation via TAF1 in response to DNA damage using ChIP assays and RT-PCR. Further we show direct phosphorylation of E2F1 and FOXM1 via TAF1 utilizing kinase assays. These studies provided important insights into how TAF1 impacts genome-wide transcription regulation.

2.2 Introduction

The TATA-Binding Protein Associated Factor 1 (TAF1) is a member of the general transcription machinery that plays a pivotal role in the regulation of gene expression. Current research examining transcriptional regulation emphasizes

the events that precede the initiation of transcription and while these events play a crucial role in regulating gene expression, correct regulation also requires turning off transcription when it's no longer needed. While TAF1 is a member of the general transcription factor machinery responsible for regulating transcription, it is the interaction with DNA bound activators that signals the initiation of transcription. Presumably, through post-translational modification after transcription, this process could lead to turning off transcription. Therefore, we aimed to determine which transcription factors TAF1 could potentially be interacting with and regulating their activity.

TAF1 is the largest subunit of TFIID, (Burley and Roeder, 1996; Tora 2002; Thomas and Chiang, 2006) and contains kinase activity (Dikstein et al., 1996), histone acetylation transferase (HAT) activity (Mizzen et. al., 1996), ubiquitination activity (Crane-Robinson, 1999; Pham and Sauer, 2000), and a double bromodomain (Jacobson, 2000). TAF1 acts genome-wide as a transcriptional regulator and its functions are carefully orchestrated via many variables, including its ability to regulate transcription through the recognition of the activator. TAF1 has been found to recognize acetylated lysine groups on activators, such as p53. p53 is a tumor suppressor protein involved in cellular response to DNA damage (Vousden and Prives, 2009; Kruse and Gu, 2009) and is known to activate p21 that is involved in cell cycle arrest (Wu et al., 2014). p53 becomes acetylated upon DNA damage at K305, K320, K370, K372, K373. K382 and K386 and phosphorylated at 17 different residues (Bode and Dong, 2004)

leading to its stabilization and activation. Upon DNA damage, TAF1 recognizes and binds to p53 through acetylated K373 and 382 via its double bromodomain (Li et. al., 2007). p21 is responsible for binding to and inhibiting all currently known cyclin-dependent protein kinases that are required for G1-S phase transition, thereby enforcing cell cycle arrest. However, it has also been found that TAF1 phosphorylates p53 at Thr55, which leads to p53 inactivation (Li et al., 2004). It was determined that at early stages of DNA damage, acetylated p53 recruits TAF1 to the p21 promoter to activate transcription of p21. However, during the later stages of DNA damage, p53 is phosphorylated at Thr55 by TAF1 kinase activity, releasing p53 from the p21 promoter, terminating transcription (Wu et. al., 2014). This was further verified through the usage of apigenin, a TAF1 kinase inhibitor. In the presence of the inhibitor, TAF1 was unable to phosphorylate p53 at Thr55, thereby not releasing p53 from the p21 promoter, causing transcription of p21 to continue. This study elucidates that while TAF1 is responsible for activating transcription through its participation in the initiation complex, TAF1 is also responsible for ceasing transcription.

TAF1's ability to act as this molecular switch for activation/termination of transcription relies on its unique feature of its kinase activity being cellular ATP dependent. Upon DNA damage induced by ultra violet (UV) radiation, ATP levels drop initially, rendering the kinase activity of TAF1 inactive (Wu et al., 2014). Responsible for this drop of ATP is PARP-1, a nuclear enzyme responsible for adding ADP-ribose units on Glu residues of acceptor proteins. PARP-1 uses

NAD+ as a substrate to modify numerous proteins, which leads to the depletion of cellular ATP (Schreiber et al., 2006). However, in later stages of DNA damage this depletion in ATP leads to the activation of AMPK, which allows for ATP recovery (Wu et al., 2014). AMPK is a master regulator of cell metabolism (Hardie et al, 2007) which detects and reacts to the AMP:ATP ratio by inhibiting energy consuming pathways and stimulating energy producing pathways. ATP recovery leads to TAF1 kinase activation thereby phosphorylating target protein, p53, releasing p53 from the promoter and ceasing transcription (Wu et. al., 2014). While this mechanism of regulation of p53 in response to DNA damage has been well described, other possible TAF1 targets under this same mechanism had not yet been determined.

While previous studies have investigated the relationship between TAF1 and p53 in response to ultra violet (UV) DNA damage (Wu et. al, 2014), a genome wide role in turning off transcription by TAF1 had yet to be understood. Here, we show that TAF1 plays a role in genome-wide transcription termination following DNA damage. Through bioinformatics analysis utilizing TAF1 ChIP-seq and the ENCODE database, potential transcription factor targets of TAF1 regulation were determined including E2F1, FOXM1, SP1, ATF2, EGR-1 and Creb1. Since the regulation of gene expression is a vital process for the integrity of organisms, this discovery provides a transformative and in-depth understanding of for a wide range of species.

2.3 Results

2.3a Genome-wide TAF1 binding to the promoters upon DNA damage

It has previously been reported that TAF1 is recruited to the p21 promoter in correlation with p53 at 8 hours of UV treatment and dissociates from the promoter at 16 hours of UV treatment in a kinase dependent manner (Wu et al., 2014). TAF1 recruitment and dissociation following DNA damage has been shown to regulate p53, a tumor suppressor protein responsible for regulating the expression of p21.

Because TAF1 is a general transcription factor, we reasoned TAF1 may be responsible for regulating other transcription factors in a similar manner to p53. In order to investigate whether this was an isolated event or was observable on a genome-wide scale, TAF1-ChIP sequencing in U2OS cells was performed by Dr. Joy Lin after 0, 8, 16hrs of UV treatment and 16hrs of UV treatment plus apigenin, a TAF1 kinase inhibitor. TAF1 peaks were identified using MACS, which "empirically models the length of the sequenced ChIP fragments and uses it to improve the spatial resolution of predicted binding sites" (Liu, 2014). The total number of TAF1 peaks found in each ChIP sample is shown in Figure 2.1B.

To elucidate a TAF1 binding pattern following UV induced DNA damage, a Galaxy analysis of the TAF1 ChIP-seq via Cistrome (Liu et. al., 2011) was performed using the 8hr bound peaks as the centered genomic regions. This analysis elutes that TAF1 genome-wide binding is increased after 8hrs of UV

treatment followed by a decrease in binding after 16 hours of DNA damage (Figure 2.1A). Furthermore, this decrease at 16hrs of UV treatment was rescued when TAF1 kinase activity was inhibited via apigenin (Figure 2.1A). To ensure the results of the TAF1 ChIP-seq in our U2OS system were comparable amongst other cell lines, a percentage of overlap with several TAF1 ChIP-seq data sets was determined via ENCODE. Utilizing the TAF1 ChIP-seq datasets available via the ENCODE database, we determined that in our mock TAF1 ChIP-seq sample, there was at least a sixty percent overlap in TAF1 bound peaks in six other cell lines (Figure 2.1C). This result indicates the TAF1 kinase regulated transcription factors identified in this study may be applicable to other cell lines.

2.3b Optimization of TAF1 ChIP-seq analysis to identify potential TAF1 kinase regulated promoters

Following UV induced DNA damage, TAF1 showed an overall increase in DNA binding at 8hrs and an overall reduction at 16hrs (Figure 2.1A). To identify peaks that are significantly increased at 8hrs and decreased at 16hrs, a peak-calling program, MACS (Liu, 2014) was utilized. MACS has the ability to perform differential binding analysis by comparing two different conditions, allowing one to identify peaks that TAF1 increased or decreased in binding to upon UV treatment. Using the MACS differential binding program, we compared differential binding that showed an increase of TAF1 binding between 0 and 8 hours of UV treatment as well as differential binding that showed a decreased

between 8 and 16 hours of UV treatment. To ensure genome-wide identification of TAF1 transcriptionally regulated promoters, the stringency parameters (C) used in the MACS differential peak calling analysis required optimization.

In order to call differentially bound peaks between two conditions, we needed to determine a stringency cutoff. Do to this, MACS allows us to set a log 10 likelihood ratio, which can be used to determine how significant the differences are between the peak pileups. The higher the C value, the more stringent the differential peak calling is. For example, analysis of TAF1 data using C3 ($-\log_{10}$(pvalue)=3, pvalue of < 0.001) results in 5065 peaks that significantly increased from 0 to 8hrs of UV treatment while analysis using C2 ($-\log_{10}$(pvalue)=2, pvalue of <0.01) resulted in 6393 peaks (Figure 2.2B). We needed to determine the optimal C value in order to identify novel TAF1 kinase target promoters. Therefore, we inspected individual peaks under two stringency parameters, C3 and C2, in order to call TAF1 bound peaks that show a strong increase in binding at 8hrs and a drop at 16hrs. A random selection of 15 peaks were inspected from C3 analysis output using the UCSC genome browser (Kent et. al., 2002). Our results reveal all inspected elements appeared to be strong peaks that show a clear trend of increasing at 8hrs and decreasing at 16hrs of UV treatment. Further, when inspecting peaks that were not called by C3 but called by C2 analysis, it was apparent that C2 data did not show peaks with as strong of a correlation of increasing at 8hrs and decreasing at 16hrs of UV treatment (Figure 2.2A). To show that MACS differential binding analysis called

appropriate peaks, peaks that were not called were also inspected (Figure 2.2A).

Together, these results indicate that C2 will provide a more comprehensive analysis of genome-wide transcriptional targets of TAF1 regulation, however, C3 data will provide a more stringent list of targets to be tested. In order to determine the top transcription factors and promoters under TAF1 kinase regulation, a more stringent analysis is necessary.

2.3c Identification and characterization of TAF1 kinase regulated promoters

To determine which subset of promoters correlated strongest with TAF1 binding following UV damage, of increasing at 8hrs and dropping at 16hrs, we utilized the schematic outlined in Figure 2.3A. An intersection of TAF1 bound genomic regions (Kent et al., 2002) that significantly increased from 0 to 8hrs and decreased from 8 to 16hrs of UV treatment within 3kb of a transcription start site lead to 2352 resulting peaks (Figure 2.3B). In order to determine promoters under TAF1 kinase regulation, these peaks were then intersected with peaks that significantly rescued from 16hrs UV treatment to 16hrs of UV treatment plus apigenin, a TAF1 kinase inhibitor. This resulted in 634 remaining peaks (Figure 2.3B).

Apigenin is a naturally occurring flavonoid believed to be a promising cancer therapeutic drug (Shukla and Gupta 2010). Apigenin leads to the inhibition of protein kinase C (Lin et al., 1997) and MAPK (Shukla and Gupta, 2007) and can inhibit casein kinase expression (Hessenauer et al., 2003).

Further, it was shown that apigenin decreased the levels of phosphorylated EGFR tyrosine kinases and c-myc (Yin et al., 2009) leading to programmed cell death (Yin et al., 2009). Since apigenin is known to have other cellular effects, we took into account possible background effects of apigenin. Therefore, to make this list of potential TAF1-kinase regulated promoters more specific, we aimed to remove possible background effects of apigenin treatment. To accomplish this, TAF1 ChIP-seq was performed at 8hrs of UV treatment plus apigenin by Dr. Joy Lin (Figure 2.3C). Since ATP levels are low following 8hrs of UV (Wu et al., 2014), TAF1 kinase activity should be low given its known Km of 1.9 mM (Wu et. al., 2014). Therefore, any TAF1 peaks that increased from 8hrs of UV to 8hrs of UV plus apigenin may be caused by a TAF1-kinase independent mechanism. There were 2854 TAF1 peaks that significantly increased from 8hrs of UV to 8hrs of UV plus apigenin and of these, 159 peaks overlapped with the 634 peaks identified previously (Figure 2.3B). These 159 potential TAF1-kinase independent peaks were removed from our 634 peaks resulting in 475 TAF1 peaks. The 475 peaks were analyzed using ChIPpeakAnno (Zhu et al., 2010) and determined there was 421 unique promoters (Figure 2.3B). TAF1 binding to PIGT, SOD2, and SMPD1, three promoters from the 421 identified promoters, are shown in the genome browser (Figure 2.3D) (Kent et al., 2002).

To further understand the possible effects of TAF1 kinase activity, a functional analysis was performed on the 421 identified genes. Since TAF1 was found to increase in its binding occupancy to these promoters upon UV

treatment, it was not surprising to find that the regulation of cellular processes was highly enriched in this analysis. Within the cellular processes function, cell communication (41.9%) and cell cycle (32.3%) were dominant categories (Figure 2.4) (Thomas et al., 2003).

2.3d Utilization of the ENCODE database in the identification of potential transcription factor targets

Next, we aimed to determine which transcription factors TAF1 could potentially be interacting with and regulating their activity. To accomplish this, we utilized the database ENCODE which encompasses ChIP-seq data with 161 transcription factors in 91 different cell lines (Dunham et al., 2012) and overlapped it with our 421 identified promoters (Figure 2.5). We decided to first narrow our analysis to 6 cell lines which all have a wild-type p53, in order to be consistent with our analysis in U2OS (Table 2.1). The resulting 227 ChIP-Seq data sets were analyzed by overlapping transcription factor binding regions with the 421 TAF1 bound promoters that TAF1 increased in binding at 8hrs of UV, decreased at 16hrs, and showed recovery at 16hrs in the presence of apigenin. A visual representation is shown in Figure 2.6, in which the ChIP-seq from the ENCODE database is shown as a grey bar under the TAF1 bound peak in the genome browser. To identify transcription factors binding to the TAF1 bound peaks, we analyzed -3kb and +1kb from the TSS. The transcription factors were then ranked based on how many of the 421 promoters it was found to bind. As

expected, the top ranked transcription factors were factors of general transcription machinery (Table 2.2). However, some sequence specific transcription factors were also highly ranked. These sequence specific transcription factors were then further investigated by determining the total number of peaks in each individual ChIP-seq sample (MACS; Liu et al., 2014) and how many promoters this corresponded to (ChIPpeakAnno; Zhu et al., 2010) (Table 2.2, columns 4 &5). A ratio was then determined for promoters the transcription factor was bound to in our analysis (out of 421) to total promoters in the ChIP-seq sample (Table 2.2, column 6). This ratio was used to determine transcription factors that were more specific to the potential TAF1-kinase regulated promoters rather than ubiquitously bound. We also investigated interactions of identified transcription factor with general transcription factor machinery to aid in selection of potential TAF1-kinase regulated targets (Table 2.3; column 7). Many of the identified transcription factors were known to be involved in stress response and cell cycle regulation including E2F1, FOXM1, ATF2, and SP1. This result was expected since DNA damaging conditions were used in this analysis.

2.3e Transcription factors FOXM1, ATF2, E2F1, and SP1 show similar binding pattern and target expression to p53 in response to UV treatment

In order to verify transcription factor binding, ChIP assay was performed for select transcription factors on promoters identified in this analysis as well as

known target promoters. We chose to investigate E2F1, FOXM1, SP1, ATF2, EGR1 and Creb1 transcription factors due to their ranking in the aforementioned analysis as well as their known role in DNA damage response. Transcription factors E2F1, FOXM1, SP1, and ATF2 showed a binding pattern to target promoters similar to p53 binding to the p21 promoter (Figure 2.7A). These transcription factors showed an increase in binding following 8hrs of UV exposure, and the dissociation at 16hrs was rescued following apigenin treatment. We further verified the TAF1 ChIP-seq results via TAF1 ChIP assay at identified promoters (Figure 2.7A). ChIP results were also verified using qPCR for select promoters (Figure 2.7B).

To address functional significance of the transcription factor binding following UV exposure, hnRNA levels of target promoters were investigated. These patterns of transcription factor binding to target promoters showed a positive correlation to hnRNA expression indicating a strong role in transcription factor regulation of these promoters upon UV damage (Figure 2.7C). These results suggest that transcription factors E2F1, FOXM1, SP1, and ATF2 may be under the same mode of regulation as p53 via TAF1 kinase activity.

It is important to note that not all transcription factors tested showed this same trend in binding following UV exposure. We showed that Creb1 bound to promoters APEX1 and PIGT and that EGR-1 bound to PIGT and API5 promoters. However, CREB1 and EGR-1 did not exhibit the same trend following UV induced damage as p53 (Figure 2.7A). While CREB1 and EGR-1 were strong

candidates in this analysis, they did not show the expected DNA binding trend following UV. This may be due to the fact that there is intricate interplay between these transcription factors at various promoters as shown by their known interactions (Figure 2.9). Therefore, while they are found to bind to the identified promoters in this analysis, TAF1 may be interacting with, and regulating, another transcription factor bound to the same promoter.

2.3f Transcription factor dissociation via TAF1 kinase is ATP-level dependent

Previous studies have shown that TAF1's phosphorylation of p53 at Thr55, which leads to the dissociation from the p21 promoter, is cellular ATP concentration dependent (Wu et al., 2014). We have previously shown that FOXM1 exhibited a similar binding pattern to p53 following DNA damage and that the dissociation of FOXM1 from the promoter in late stages of damage is TAF1 kinase dependent via the use of apigenin. Therefore, we aimed to determine if the dissociation of FOXM1 from the BRCA2 promoter was dependent on cellular ATP concentration. To accomplish this, PARP-1 activity was inhibited using 4-AN treatment (Horton et al, 2005) and AMPK activity was inhibited using compound C treatment (Zhou et al., 2010). Following DNA damage, PARP-1 leads to a drop in cellular ATP concentration. This decrease in ATP is later sensed by AMPK, which allows for a recovery in cellular ATP concentration (Wu et al., 2014). This recovery in ATP activates TAF1 kinase

activity. Therefore, by inhibiting either PARP-1 or AMPK activity, ATP levels do not recover in late stages of DNA damage. Since TAF1 kinase activity is dependent on high cellular ATP concentrations (Km = 1.9mM) (Wu et al., 2014), TAF1 will exhibit low kinase activity when PARP-1 or AMPK are inhibited. While upon UV treatment alone we observe that FOXM1 dissociates from the promoter at 16hrs of UV treatment (Figure 2.7A), with the addition of 4-AN or compound C, the dissociation is no longer found (Figure 2.10). This indicates that the dissociation of FOXM1 from BRCA2 is ATP concentration dependent, further implicating TAF1 kinase dependency in FOXM1 transcriptional regulation.

2.3g TAF1 directly phosphorylates FOXM1 and E2F1

TAF1 is known to autophosphorylate and transphosphorylate other proteins such as RAP74, TFIIA (Dikstein et al., 1996; Solow et al., 2001), p53 (Li et al., 2004; Wu et al., 2014), TAF7 (Kloet et al., 2012) and histone H2B (Maile, 2004). Further, this phosphorylation has been shown to play a role in transcriptional regulation (Wu et al., 2014). The importance of TAF1 kinase activity was emphasized when TAF1 was found to phosphorylate p53 at Thr55. This phosphorylation was deemed necessary for TAF1/p53 dissociation from the p21 promoter, ceasing p21 transcription (Wu et al., 2014). Therefore, we tested if TAF1 is responsible for phosphorylating the identified transcription factor targets FOXM1, ATF2, E2F1, and Sp1 (Figure 2.11B) in an in vitro phosphorylation assay. As a control, we show bacculovirus expressed TAF1 was able to

phosphorylate known targets p53 and RAP74. Further, this phosphorylation was ceased in the presence of apigenin, a TAF1 kinase inhibitor (Figure 2.11B). We show here, for the first time, that TAF1 directly phosphorylates FOXM1 and E2F1 (Figure 2.11C). We further demonstrated this phosphorylation was not due to contaminate kinase via the use of apigenin, a TAF1 kinase inhibitor (Figure 2.11D). The bacterially expressed and purified transcription factors used in the kinase assays are shown in Figure 2.11A. Since TAF1 phosphorylates p53 in late stages of DNA damage and leads to the dissociation of p53 from the p21 promoter, this data may suggest that the phosphorylation of FOXM1 and E2F1 is also leading to their dissociation.

2.4 Discussion

TAF1 is an essential protein in many, if not most, eukaryotic organisms and understanding its role in transcriptional regulation will provide insights into the basics of gene regulation and can be applied to a variety of cell types in a wide range of species. Correct gene expression requires timely activation as well as turning off transcription when gene expression is no longer needed. This regulation is particularly important in the case of DNA damage to maintain genomic integrity and avoid disease. While TAF1 is known to play an important role in transcriptional activation, we show here, that TAF1 kinase activity plays an

important role in genome-wide transcriptional termination of gene expression in response to UV DNA damage.

We found via ChIP-sequencing that TAF1 is recruited to promoter's genome-wide in response to UV induced DNA damage. Further, we show that at late stages of UV induced DNA damage, TAF1 dissociates from the promoter and that this dissociation appears to be TAF1 kinase dependent as emphasized through the use of apigenin. This not only implicates TAF1 as an important protein in DNA damage response, but further implicates TAF1 kinase activity in turning of transcription following DNA damage.

Following the discovery that TAF1 phosphorylates p53 on the p21 promoter, leading to p53 dissociation and turning-off of p21 transcription (Wu et al., 2014), we investigated if TAF1 can regulate other transcription factors in a similar manner. Using the pipeline described in Figure 2.3A and Figure 2.5, potential TAF1 kinase-regulated transcription factors and promoters were identified. The transcription factors identified and investigated include FOXM1, E2F1, ATF2, and SP1, all transcription factors well known for their involvement in cell cycle regulation or DNA damage response.

E2F1 and response to DNA damage:

E2F1 was one of the top transcription factors identified in this analysis as potentially TAF1 kinase regulated. Known for its role in DNA damage response,

E2F1 was found to be recruited to promoters upon DNA damage and effected transcription of its targets.

E2F1 is a member of the E2F family which plays a central role in the regulation of cell cycle progression from the G1 to S phase as well as plays a role in stress response by promoting cell survival. Upon UV DNA damage, E2F1 is known to promote cell survival by inhibiting apoptosis and cell cycle promoting genes (Biswas et al., 2011), allowing for DNA repair (Biswas et al., 2014). UV induced DNA damage leads to the increase in E2F1 protein expression, analogous to p53, reaching a peak after 6-12 hours of UV treatment (O'Connor and Lu, 2000). However, while protein levels are increased, E2F1 is transcriptionally inactive, independent of p53 and Rb (O'Connor and Lu, 2000). UV damage induces phosphorylation of both E2F1 and TopBP1 via ATM/ATR, which leads to stabilization of E2F1 and TopBP1 complexes (Lin et al., 2001; Liu et al., 2003). When TopBP1 binds to E2F1, E2F1 transcriptional activity is repressed (Liu et al., 2003, Liu et al., 2004). TopBP1 is known for its important role in suppressing E2F1 mediated cell cycle progression and apoptosis following DNA damage (Biswas et al., 2011). Based on these previous studies, it is possible E2F1 protein levels are increased to induce repression of cell cycle progression genes in a TopBP1 complex.

Our finding have shown that E2F1 is recruited to promoters such as API5 and RPA2 upon UV induced DNA damage (Figure 2.8A). However, upon measuring hnRNA levels of API5, we observed a decrease in expression

following 8 hrs of UV treatment (Figure 2.8C). In later stages of DNA damage (16hrs), we witnessed a drop in E2F1 binding and a recovery in API5 expression (Figure 2.8A, Figure 2.8C). API5, Apoptosis Inhibitor 5, inhibits apoptosis after growth factor deprivation. Further, API5 is known to be required for E2F1 transcriptional activation of cell cycle progression genes (Garcia-Jove et al., 2013). Therefore, it's conceivable that upon DNA damage, E2F1 is negatively regulating API5, to inhibit cell cycle progression. RPA2, Replication Protein A2, is a member of the replication protein A complex that is involved in DNA replication. It has been previously shown that RPA2 expression is E2F1 dependent (Chao-Chung et. Al., 2017). Taken together, our findings prove to be consistant with the theory of E2F1 repression upon DNA damage and may act through the interaction with TOPBP1.

FOXM1 and DNA damage response:

FOXM1 is well known for its role in cell cycle and DNA damage response (Zona et al.,2014). Thus, FOXM1 was chosen to be investigated as potentially TAF1 kinase regulated. FOXM1 is a member of the Forkhead family (Lam et al., 2013) which is also known to play a central role in cell cycle progression of the G2 to mitosis transition as well as plays a role in DNA damage response by promoting genes involved in DNA repair. It has been shown that upon DNA damaging conditions, such as UV, FOXM1 is phosphorylated via CHK2 in U2OS cells leading to its stabilization and subsequent activation of its downstream DNA

repair targets (Tan et al., 2007). FOXM1 is a well-known transcription factor responsible for the regulation of many DNA repair genes. This central role in DNA repair was emphasized following the observation that FOXM1 depleted cells showed increased DNA damage (Tan et al., 2007). FOXM1 is known to transcriptionally activate DNA repair genes including BRCA1-interacting protein-terminal helicase 1 (BRIP1)(Monteiro et al., 2013), SIRT1(Yuan et al., 2007), POLE2 (Zhou et al., 2014), and BRCA2 (Tan et al., 2007) in response to DNA damage.

Upon UV treatment in U2OS, we observed an increase in FOXM1 binding to the promoters of BRIP1, SIRT1, POLE2, and BRCA2(Figure 2.8A), consistent with previous findings (Monteiro et al., 2013; Yuan et al., 2007; Zhou et al., 2014; Tan et al., 2007). In later stages of DNA damage (16hrs), we observed a drop in FOXM1 binding to these promoters (Figure 2.8A). We further observed a corresponding trend in SIRT1 expression (Figure 2.8C) to FOXM1 binding, indicating SIRT1 expression is regulated via FOXM1.

ATF2 and DNA damage response:

Since the Activating Transcription Factor 2 (ATF2) is also known for its role in cell response to stress, it made ATF2 a likely candidate for being TAF1 kinase regulated. ATF2 is a member of the CREB protein family (Nomura et al., 1993) and is transcriptionally activated in response to ROS stress, cytokine exposure, and UV irradiation (Van Dam et al., 1995). ATF2 transcriptional targets

are typically categorized as being involved in the regulation response to stress and DNA damage, regulation of growth and tumorigenesis, and the maintenance of homeostasis (Bhoumik et al., 2007). The phosphorylation of ATF2 via ATM is required for its response to DNA damage. In a microarray study, ATF2 regulated genes were significantly enriched for genes involved in DNA repair (Hayakawa et al., 2004). Among ATF2 regulated genes were RAD23B and MSH6. Our findings suggest that ATF2 increases in promoter occupancy at RAD23B and MSH6 following UV treatment. We further show ATF2 known targets MSH6, ERCC3, and XPA show a similar trend in hnRNA expression to ATF2 binding. This indicates that these genes are under ATF2 transcriptional regulation in response to UV induced DNA damage.

SP1 and DNA damage response:

The Specificity Protein 1 (SP1) plays a diverse role in cell functions, including cell response to stress. SP1 activities in response to UV induced DNA damage was identified in this analysis as potentially TAF1 kinase regulated.

SP1 is one of the most well characterized transcription factors and is known to bind to GC rich regions necessary for the regulation of a plethora of genes (Kadonaga et al., 1986; Letovsky & Dynan, 1989; Azizkhan et al., 1993; Suske, 1999; Vizcaíno et al., 2015). SP1 may undergo many post-translational modifications including acetylation, sumoylation, phosphorylation, and glycosylation. It was previously shown that phosphorylation of SP1 will affect

DNA binding activities, including A-CDK phosphorylation of SP1 (Fojas-deBorja et al., 2001). SP1 has been implicated in the regulation of tumorigenesis and cell growth primarily due to its target genes involvement in apoptosis, angiogenesis, metastasis, and metabolism reprogramming (Wierstra, 2008; Beishline & Azizkhan-Clifford, 2015, Hanahan & Weinberg, 2011). Of such targets, SP1 has been shown to be critical in the regulation of SOD2, a member of the superoxide dismutase family, responsible for the regulation of ROS in response to oxidative stress (Yong et al., 2002). SP1 has been implicated in many cancers (Safe & Abdelrahim, 2005; Wierstra, 2008; Guan et al. et al., 2012; Sankpal et al., 2012) and has been shown to correlate with SOD2 levels. We show here, that SP1 binding increases at the SOD2 promoter and see an increase in SOD2 expression upon UV treatment. As well, we observed increase in binding to SOD1, RPS9, PIGT, and USP3 promoters. This increase in binding following DNA damage was found to decrease in later stages (16hrs). Further, we observed a correlating trend in target gene expression (Figure 2.8C). This data reveals SP1 transcriptional regulation of target promoters following UV induced DNA damage.

In later stages of DNA damage (16hrs), we observed a decrease in the aforementioned transcription factor binding, and this decrease was rescued upon apigenin treatment (Figure 2.8A). Since these transcription factors showed a decrease in binding following 16hrs of UV treatment and this dissociation could be rescued upon TAF1 kinase inhibition, FOXM1 was further investigated to

determine whether this dissociation was ATP concentration dependent. Using PARP1 and AMPK inhibitors with UV treatment, FOXM1 dissociation from the BRCA2 promoter at late stages of DNA damage was no longer seen. Since TAF1 kinase activity is dependent on cellular ATP concentrations (Km of 1.9 mM (Wu et al., 2014)), this further solidifies that TAF1 kinase is playing a role in transcription factor dissociation and transcriptional termination of its targets. Kinase assays revealed that TAF1 is directly phosphorylating E2F1 and FOXM1, presumably leading to their dissociation from DNA. SP1 and ATF2 showed no phosphorylation, however, a possible explanation may be TAF1 is regulating a co-factor of these proteins leading to subsequent dissociation (Vizcaíno et al., 2015, Duyndam, 1999).

This data reveals that TAF1 acts as a genome-wide transcriptional regulator in response to UV induced DNA damage. Through this analysis, we identified many transcription factors whose transcriptional termination may be TAF1 kinase regulated. Therefore, it’s conceivable that TAF1 may be acting as a master regulator in response to DNA damage. This is further exemplified via that TAF1 is one of the most mutated transcription factors in cancers (Kandoth et al., 2013). While we show TAF1 transcriptional regulation of transcription factors FOXM1, E2F1, SP1, and ATF2- our analysis indicates that there are many other transcription factor candidates. This alludes to the fact that TAF1 may play a more influential role in DNA damage response than previously thought.

2.5 Materials and methods

Chromatin Immunoprecipitation:

ChIP analysis was carried out as described previously (Li et al, 2007). Nuclear extracts of U2OS were collected at indicated time points after mock or 20 J/m2 UVC treatment and sonicated to generate chromatin fragments of 300 bp.

Antibody information: Anti-p53 antibody (FL393, Santa Cruz), anti-TAF1 antibody (Ab1230 against TAF1 DBrD), anti-FOXM1 (K-19x, Santa Cruz), anti-E2F1 (KH95x, Santa Cruz), anti-ATF2 (C-19x, Santa Cruz), anti-SP1 (PEP2x, Santa Cruz), anti-Creb1(C-21x, Santa Cruz), and anti-Egr1 (588x, Santa Cruz). 2 ug of each antibody used in chromatin immunoprecipitation.

Primers used are as follows:

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

USP3: 5'- GCGGCGTGACTAGGAAAGTC and 5'-

CACCCCTAGACTGGAGGCTT

PIGT: 5'- CCTGCCTACTCCCTCTCGT and 5'- CCGGGATGCGGTTATCAGAG

MSH6: 5'- TTTAAATACTCTTTCCTTGCCTGG and 5'-

TCTTCCGCTTTCGAGCAACT

RAD23B: 5'-CCTTGGGTTGGGCAGTAAATC and 5'-

GCACTGGTGTGAAGTGTGAGA

Compound C and 4-AN treatment were performed as previously described (Wu et. al., 2014).

All binding sites were amplified with 30–35 cycles of PCR. The PCR products were electrophoresed by agarose gels and visualized by ethidium bromide.

Data Analysis:

All ChIP-Seq data sets were aligned using Bowtie2 (version 0.12.8) (Langmead et al., 2009) to the human reference genome (GRCh37/hg19). The 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 binding peaks. The ratio of the peak pileup between 0 and 8hr samples was used to estimate activated TAF1 bindings at 8hr after DNA damage. The ratio of the peak pileup between 16 and 8hr was used to determine decreased TAF1 binding at 16hr after DNA damage. The

peaks identified by ChIP-Seq were analyzed with the R Bioconductor package ChIPpeakAnno (Zhu et al, 2010) to retrieve the nearest Ensembl gene (10 kb around transcription start site, TSS).

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 peak and ranked according to the enrichment of TAF1occupancy at 8 h after DNA damage.

RT-PCR and RT-qPCR:

Total RNA was extracted using TRIzol reagent (Sigma) according to manufacturer's protocol. RT-PCR was performed using SuperScript One-Step RT-PCR kit (Invitrogen). RT-qPCR was performed using iQ SYBR Green Supermix and iScript cDNA synthesis kit on CFX96 Real Time System (Bio-Rad Laboratories).

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

XPA: 5’- CGGGGAGAATCTGCACACATA and 5’-GAGCTAGGTCCTCGGAGTGG

ERCC3: 5’-ACATCAGGGTGGCAACTTTCA and 5’-CAGATTTGGCGAAGGGTTGTG

TAF1 kinase Assay:

In vitro TAF1 phosphorylation assay was carried out using 120 ng of purified substrate with 350 uM ATP (^{32}p) in 20 µl of phosphorylation buffer under conditions as described (Li et al, 2004). TAF1 was expressed using bacculovirus system in sf9 cells and substrate proteins via *Escherichia coli* BL21 strain (Thomas Benedict). TAF1 was immunoprecipitated with anti-HA antibodies (12CA5), and phosphorylation was carried out for 1 hour at 30 °C. Reactions were stopped by the addition of SDS sample buffer. Then samples were heated for 5 min to 95 °C before analysis by SDS–PAGE and autoradiography. Phosphorylation of substrates was detected by autoradiography.

2.6 References

1. Azizkhan, J. C., Jensen, D. E., Pierce, A. J., & Wade, M. (1993). Transcription from TATA-less promoters: dihydrofolate reductase as a model. Crit Rev Eukaryot Gene Expr 3, 229–254.

2. Beishline, K., & Azizkhan-Clifford, J. (2015). Sp1 and the 'hallmarks of cancer'. FEBS J 282, 224–258.

3. Bhoumik, A., Lopez-Bergami, P., and Ronai, Z. (2007). ATF2 on the double – activating transcription factor and DNA damage response protein. Pigment Cell Research 20, 498–506.

4. Biswas, A.K., and Johnson, D.G. (2012). Transcriptional and Nontranscriptional Functions of E2F1 in Response to DNA Damage. Cancer Research 72, 13–17.

5. Biswas, Anup K., David L. Mitchell, and David G. Johnson. "E2F1 Responds to Ultraviolet Radiation by Directly Stimulating DNA Repair and Suppressing Carcinogenesis." Cancer Research 74 (2014)

6. Bode, A.M., and Dong, Z. (2004). Post-translational modification of p53 in tumorigenesis. 4, 793.

7. Burley, S. K., and R. G. Roeder. "Biochemistry and Structural Biology of Transcription Factor IID (TFIID)." Annual Review of Biochemistry 65.1 (1996): 769-99.

8. Buschmann T, Potapova O, Bar-Shira A, Ivanov VN, Fuchs SY, Henderson S, Fried VA, Minamoto T, Alarcon-Vargas D, Pincus MR, Gaarde WA, Holbrook NJ, Shiloh Y and Ronai Z (2001) Jun NH2-terminal kinase phosphorylation of p53 on Thr-81 is important for p53 stabilization and transcriptional activities in response to stress. Mol. Cell. Biol. 21: 2743–2754.

9. Chao-Chung Chen, Chi-Wen Juan, Kuan-Yu Chen, Yi-Chien Chang, Janq-Chang Lee, Ming-Chung Chang; Upregulation of RPA2 promotes NF-κB activation in breast cancer by relieving the antagonistic function of menin on NF-κB-regulated transcription, Carcinogenesis, Volume 38, Issue 2, 1 February 2017, Pages 196–206

10. Crane-Robinson, Colyn. "How Do Linker Histones Mediate Differential Gene Expression?" BioEssays 21.5 (1999): 367-71.

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

12. Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.

13. Duyndam, M.C., van Dam, H., Smits, P.H., Verlaan, M., van der Eb, A.J., and Zantema, A. (1999). The N-terminal transactivation domain of ATF2 is a target for the co-operative activation of the c-jun promoter by p300 and 12S E1A. Oncogene 18, 2311–2321.

14. Dyson, N. “The regulation of E2F1 by pRb-family proteins.” Genes Dev (1998)12,2245-2262

15. E.W. Lam, J.J. Brosens, A.R. Gomes, C.Y. Koo, Forkhead box proteins: tuning forks for transcriptional harmony, nature reviews, Cancer 13 (2013) 482–495.

16. Fojas-de-Borja, P., Collins, N. K., Du, P., Azizkhan-Clifford, J., & Mudryj, M. (2001). Cyclin A-CDK phosphorylates Sp1 and enhances Sp1-mediated transcription. EMBO J 20, 5737–5747.

17. Garcia-Jove Navarro, M., Basset, C., Arcondéguy, T., Touriol, C., Perez, G., Prats, H., and Lacazette, E. (2013). Api5 Contributes to E2F1 Control of the G1/S Cell Cycle Phase Transition. PLoS ONE 8, e71443.

18. Gentleman, R., V. Carey, D. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, and Others. "Bioconductor: Open Software Development for Computational Biology and Bioinformatics." Genome Biology 5 (2004)

19. Gu W., Roeder R.G. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell. 1997;90:595–606. doi: 10.1016/S0092-8674(00)80521-8.

20. Guan, H., Cai, J., Zhang, N., Wu, J., Yuan, J., Li, J., et al. (2012). Sp1 is upregulated in human glioma, promotes MMP-2-mediated cell invasion and predicts poor clinical outcome. Int J Cancer 130, 593–601.

21. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell 144, 646–674.

22. Harbour, J.william, Robin X. Luo, Angeline Dei Santi, Antonio A. Postigo, and Douglas C. Dean. "Cdk Phosphorylation Triggers Sequential Intramolecular Interactions That Progressively Block Rb Functions as Cells Move through G1." Cell 98.6 (1999): 859-69

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

24. Hayakawa J, Mittal S, Wang Y, Korkmaz KS, Adamson E, English C, Ohmichi M, McClelland M, Mercola D. Identification of promoters bound by c-Jun/ATF2 during rapid large-scale gene activation following genotoxic stress. Mol Cell 2004;16:521–535.

25. Hessenauer A, Montenarh M, Gotz C. Inhibition of CK2 activity provokes different responses in hormone-sensitive and hormone-refractory prostate cancer cells. Int J Oncol 2003;22:1263–1270. [PubMed: 12738992]

26. Horton JK, Stefanick DF, Naron JM, Kedar PS, Wilson SH. Poly (ADP-ribose) polymerase activity prevents signaling pathways for cell cycle arrest after DNA methylating agent exposure. J. Biol. Chem. 2005;280:15773–15785

27. J. Zhou, Y. Wang, Y. Wang, X. Yin, Y. He, L. Chen, W. Wang, T. Liu, W. Di, FOXM1 modulates cisplatin sensitivity by regulating EXO1 in ovarian cancer, PLoS One 9 (2014) e96989.

28. Jacobson, R. H. "Structure and Function of a Human TAFII250 Double Bromodomain Module." Science 288.5470 (2000): 1422-425.

29. Kadonaga, J. T., Jones, K. A., & Tjian, R. (1986). Promoter-specific activation of RNA poly- merase II transcription by Sp1. Trends Biochem Sci 11, 20–23.

30. Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature 502, 333–339.

31. Kent W., C.W Sugnet, T.S Furey, K.M Rosekin, T.H Pringle, A.M Zahler, D. Haussler. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006

32. Kloet, S.L., Whiting, J.L., Gafken, P., Ranish, J., and Wang, E.H. (2012). Phosphorylation-Dependent Regulation of Cyclin D1 and Cyclin A Gene Transcription by TFIID Subunits TAF1 and TAF7. Molecular and Cellular Biology 32, 3358–3369.

33. Kruse, Jan-Philipp, and Wei Gu. "Modes of P53 Regulation." Cell 137.4 (2009): 609-22.

34. L.J. Monteiro, P. Khongkow, M. Kongsema, J.R. Morris, C. Man, D. Weekes, C.Y. Koo, A.R. Gomes, P.H. Pinto, V. Varghese, L.M. Kenny, R. Charles Coombes, R. Freire, R.H. Medema, E.W. Lam, The Forkhead Box M1 protein regulates BRIP1 expression and DNA damage repair in epirubicin treatment, Oncogene 32 (2013) 4634–4645.

35. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

36. Letovsky, J., & Dynan, W. S. (1989). Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. Nucleic Acids Res 17, 2639–2653.

37. Li, Heng-Hong, Andrew G., Hilary M. Sheppard, and Xuan Liu. "Phosphorylation on Thr-55 by TAF1 Mediates Degradation of P53."Molecular Cell 13.6 (2004): 867-78.

38. Li, Heng-Hong, Andrew G., Landon G. Piluso, Xin Cai, Brian J. Gadd, Andreas G. Ladurner, and Xuan Liu. "An Acetylation Switch in P53 Mediates Holo-TFIID Recruitment." Molecular Cell 28.3 (2007): 408-21.

39. Lin JK, Chen YC, Huang YT, Lin-Shiau SY. Suppression of protein kinase C and nuclear oncogene expression as possible molecular mechanisms of cancer chemoprevention by apigenin and curcumin. J Cell Biochem Suppl 1997:28–29. 39–48.

40. Lin, Weei-Chin, Fang-Tsyr Lin, and Joseph Nevins. "Selective Induction of E2F1 in Response to DNA Damage, Mediated by ATM-dependent Phosphorylation." Genes Dev. 15.14 (2001): 1833-844.

41. Liu T, et al. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 2011;12:R83.

42. Liu, K., F.-T. Lin, J. M. Ruppert, and W.-C. Lin. "Regulation of E2F1 by BRCT Domain-Containing Protein TopBP1." Molecular and Cellular Biology 23.9 (2003): 3287-304.

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

44. Liu, Tao, Jorge A. Ortiz, Len Taing, Clifford A. Meyer, Bernett Lee, Yong Zhang, Hyunjin Shin, Swee S. Wong, Jian Ma, Ying Lei, Utz J. Pape, Michael Poidinger, Yiwen Chen, Kevin Yeung, Myles Brown, Yaron Turpaz, and X. Shirley Liu. "Cistrome: An Integrative Platform for Transcriptional Regulation Studies." Genome Biology 12.8 (2011): R83.

45. Liu, Tao. "Call Differential Binding Events." GitHub. N.p., 20 June 2014. Web. 03 Sept. 2014. <https://github.com/taoliu/MACS/wiki/Call-differential-binding-events>.

46. Maile, T., Simona Kwoczynski, Rebeccah Katzenberger, David Wassarman, and Frank Sauer (2004). TAF1 Activates Transcription by Phosphorylatino of Serine 33 in Histone H2B. Science 304, 1008–1010.

47. Meng, P., and R. Ghosh. "Transcription Addiction: Can We Garner the Yin and Yang Functions of E2F1 for Cancer Therapy?" Cell Death and Disease (2014).

48. Mizzen, Craig A., Xiang-Jiao Yang, Tetsuro Kokubo, James E. Brownell, Andrew J. Bannister, Tom Owen-Hughes, Jerry Workman, Lian Wang, Shelley L. Berger, Tony Kouzarides, Yoshihiro Nakatani, and C.david Allis. "The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity." Cell 87.7 (1996): 1261-270.

49. Nomura N, Zu YL, Maekawa T, Tabata S, Akiyama T, Ishii S. Isolation and characterization of a novel member of the gene family encoding the cAMP response element-binding protein CRE- BP1. J Biol Chem 1993;268:4259–4266.

50. O'connor, Daniel J., and Xin Lu. "Stress Signals Induce Transcriptionally Inactive E2F-1 Independently of P53 and Rb." Oncogene 19.20 (2000): 2369-376.

51. Paul D. Thomas, Michael J. Campbell, Anish Kejariwal, Huaiyu Mi, Brian Karlak, Robin Daverman, Karen Diemer, Anushya Muruganujan, Apurva Narechania. 2003. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res., 13: 2129-2141

52. Pham, A.-D., and F. Sauer. "Ubiquitin-Activating/Conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila." Science 289.5488 (2000): 2357-360.

53. Safe, S., & Abdelrahim, M. (2005). Sp transcription factor family and its role in cancer. Eur J Cancer 41, 2438–2448.

54. Sankpal, U. T., Maliakal, P., Bose, D., Kayaleh, O., Buchholz, D., & Basha, R. (2012). Expression of specificity protein transcription factors in pancreatic cancer and their association in prognosis and therapy. Curr Med Chem 19, 3779–3786.

55. Schreiber, Valérie, Françoise Dantzer, Jean-Christophe Ame, and Gilbert De Murcia. "Poly(ADP-ribose): Novel Functions for an Old Molecule."Nature Reviews Molecular Cell Biology 7.7 (2006): 517-28.

56. Shukla S, Gupta S. (2007) Apigenin-induced cell cycle arrest is mediated by modulation of MAPK, PI3K- Akt, and loss of cyclin D1 associated retinoblastoma dephosphorylation in human prostate cancer cells. Cell Cycle; 6:1102–1114.

57. Shukla, S., and Gupta, S. (2010). Apigenin: A Promising Molecule for Cancer Prevention. Pharmaceutical Research 27, 962–978.

58. Solow, S., Salunek, M., Ryan, R. and Lieberman, P. M. (2001). TAFII250 phosphorylates human TFIIA on serine residues important for TBP binding and transcription activity. J. Biol. Chem. 276, 15886-15892.

59. Suske, G. (1999). The Sp-family of transcription factors. Gene 238, 291–300.

60. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017 Jan; 45:D362-68

61. Thomas, Mary C., and Cheng-Ming Chiang. "The General Transcription Machinery and General Cofactors." Critical Reviews in Biochemistry and Molecular Biology 41.3 (2006): 105-78.

62. Tora, L. "A Unified Nomenclature for TATA Box Binding Protein (TBP)-associated Factors (TAFs) Involved in RNA Polymerase II Transcription." Genes & Development 16.6 (2002): 673-75.

63. Van Dam H, Wilhelm D, Herr I, Steffen A, Herrlich P, Angel P. ATF-2 is preferentially activated by stress-activated protein kinases to mediate c-jun induction in response to genotoxic agents. Embo J 1995;14:1798–1811.

64. Vizcaíno, C., Mansilla, S., and Portugal, J. (2015). Sp1 transcription factor: A long-standing target in cancer chemotherapy. Pharmacology & Therapeutics 152, 111–124.

65. Vousden, Karen H., and Carol Prives. "Blinded by the Light: The Growing Complexity of P53." Cell 137.3 (2009): 413-31.

66. Wierstra, I. (2008). Sp1: emerging roles—beyond constitutive activation of TATA-less housekeeping genes. Biochem Biophys Res Commun 372, 1–13.

67. Wu, Yong, Joy C. Lin, Landon G. Piluso, Joseph M. Dhahbi, Selene Bobadilla, Stephen R. Spindler, and Xuan Liu. "Phosphorylation of P53 by TAF1 Inactivates P53-Dependent Transcription in the DNA Damage Response." Molecular Cell 53.1 (2014): 63-74.

68. Y. Tan, P. Raychaudhuri, R.H. Costa, Chk2 mediates stabilization of the FoxM1 tran- scription factor to stimulate expression of DNA repair genes, Mol. Cell. Biol. 27 (2007) 1007–1016.

69. Yin F, Giuliano AE, Van Herle AJ. Signal pathways involved in apigenin inhibition of growth and induction of apoptosis of human anaplastic thyroid cancer cells (ARO). Anticancer Res 1999;19:4297–4303. [PubMed: 10628390]

70. Yong, X.U., Porntadavity, S., and St Clair, D.K. (2002). Transcriptional regulation of the human manganese superoxide dismutase gene: the role of specificity protein 1 (Sp1) and activating protein-2 (AP-2). Biochemical Journal 362, 401–412.

71. Z. Yuan, X. Zhang, N. Sengupta, W.S. Lane, E. Seto, SIRT1 regulates the function of the Nijmegen breakage syndrome protein, Mol. Cell 27 (2007) 149–162.

72. Zhang Y, et al. Modelbased analysis of ChIP-Seq (MACS) Genome Biol. 2008;9:R137.

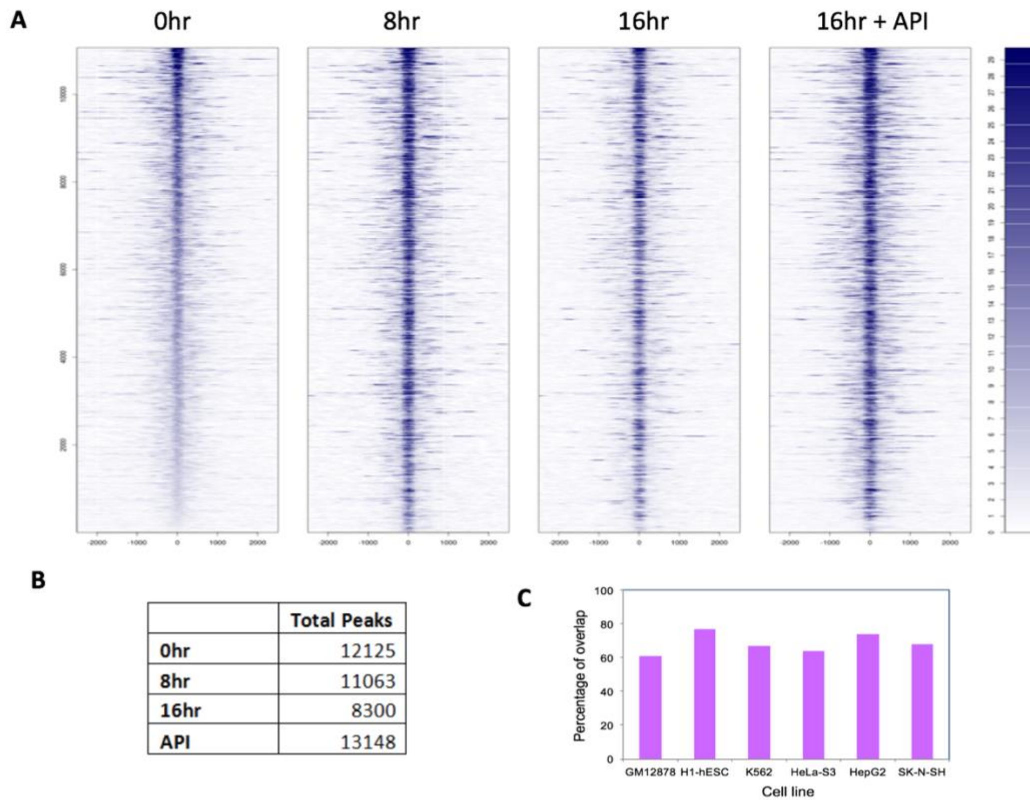
73. Zhu L, Gazin C, Lawson N, Pages H, Lin S, Lapointe D and Green M (2010). "ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data." BMC Bioinformatics, 11(1), pp. 237. ISSN 1471-2105, doi: 10.1186/1471-2105-11-237, http://www.biomedcentral.com/1471-2105/11/237.

74. Zhu LJ, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010;11:237.

75. Zona, S., Bella, L., Burton, M.J., Nestal de Moraes, G., and Lam, E.W.-F. (2014). FOXM1: An emerging master regulator of DNA damage response and genotoxic agent resistance. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1839, 1316–1322.

2.7 Figures and Tables

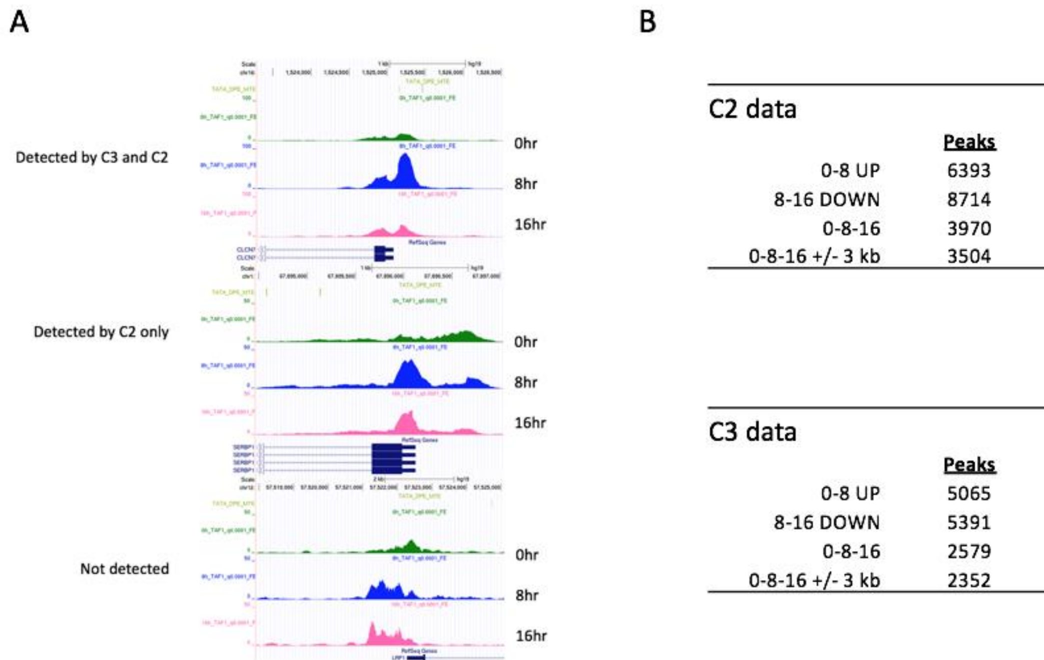
Figure 2.1: TAF1 ChIP-seq analysis and comparison to other cell lines



	Total Peaks
0hr	12125
8hr	11063
16hr	8300
API	13148

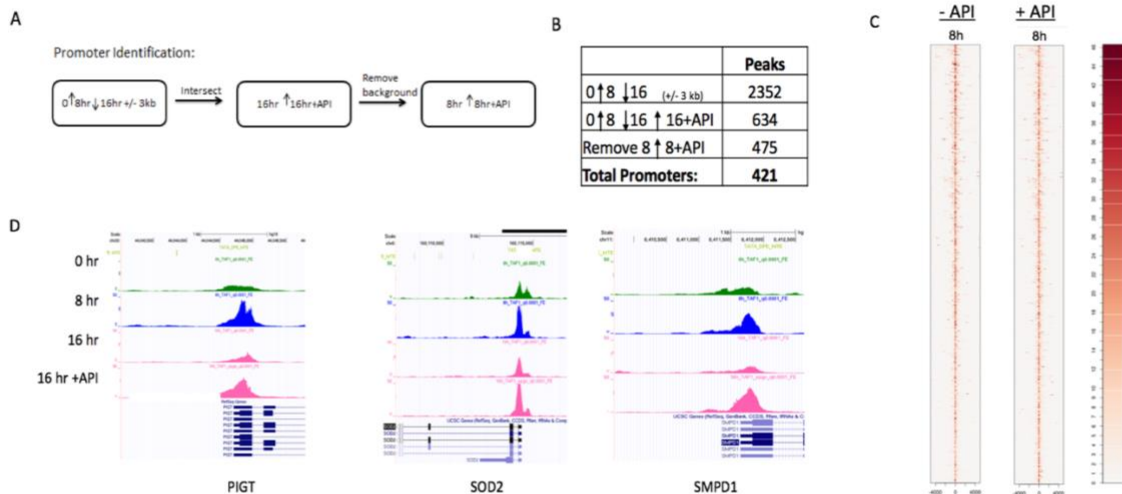
Legend: (A) Heatmap of 5 kb regions of TAF1 bound peaks genome-wide in U2OS cells following 0hrs, 8hrs, 16hrs, or 16hrs +apigenin of UV treatment (ChIP-seq performed by Dr. Joy Lin). Heatmap is ranked according to TAF1 occupancy following 8hr of UV treatment. Y-axis shows peaks ordered by pileup number, X-axis shows distance from the transcription start site. (B) Table of total TAF1 bound peaks under indicated conditions using MACS2 v. 2.0.10. (C) Percentage of overlapping TAF1 bound peaks under mock conditions in various cell lines using ENCODE.

Figure 2.2: Comparison of C2 and C3 TAF1 ChIP-seq analysis



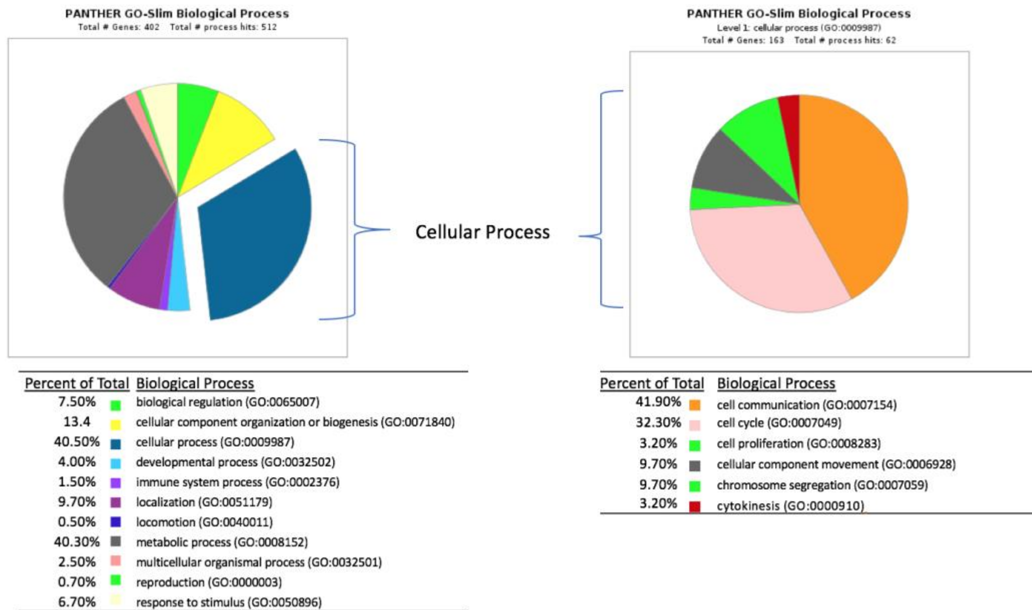
Legend: (A) Genome browser shot examples of peaks called by C3 and C2 (CLCN7), C2 only (SERBP1), or not detected (LRP1). TAF1 peaks at 0 hrs (green), 8hrs of UV (blue) and 16hrs of UV (pink) are shown. (B) Table of C2 (top) and C3 (bottom) differentially bound TAF1 peaks under mock, 8hrs, or 16hrs of UV treatment.

Figure 2.3: Identification of potential promoters under TAF1 kinase regulation



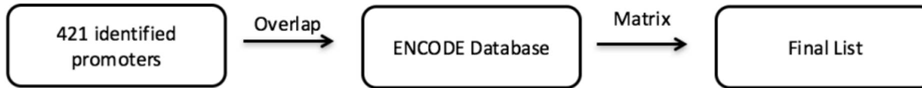
Legend: (A) Schematic of promoter identification. (B) Table of TAF1 bound peaks peaks found to increase after 8hrs, decrease at 16hrs, and recover in the presence of apigenin. Removal of peaks found to increase from 8hrs of UV to 8hrs +apigenin and resulting promoters are shown. Analyzed using C3 stringency. (C) Heatmap of 8 kb regions of TAF1 bound peaks genome-wide in U2OS cells following 8hrs and 8hrs +apigenin of UV treatment (ChIP-seq performed by Dr. Joy Lin). Heatmap is ranked according to TAF1 occupancy following 8hr of UV treatment. Y-axis shows peaks ordered by pileup number, X-axis shows distance from the transcription start site. (D) Genome browser shots of TAF1 ChIP-seq example peaks found in the 421 promoters.

Figure 2.4: PANTHER go-term analysis of top 421 promoters



Legend: PANTHER go-term analysis (Thomas et al., 2003) on biological processes (shown left) of top 421 promoters. Cellular process was further categorized using go-term analysis (shown right).

Figure 2.5: Schematic of target TAF1 transcription factor identification



Legend: The 421 promoters will be analyzed using selected ChIP-seq data through the ENCODE database to identify potential transcription factors under TAF1 kinase regulation

Table 2.1: ENCODE database cell line selection and total available ChIP-seqs

Cell Line	Total ChIP-Seq
GM12878	90
A549	35
MCF7	18
Imr90	5
HepG2	77
U2OS	2
Total ChIP's:	227
Unique TF's:	108

Legend: Table of selected cell lines and their total number of ChIP-seqs available through the ENCODE database.

Table 2.2: Top 100 ChIP-seq proteins bound to the 421 promoters

Rank	ChIP-seq	Promoter (421)	Rank	ChIP-seq	Promoter (421)
1	Pol2	398	51	Maz	150
2	Pol2	366	52	Pax5c20	146
3	Pol2	356	53	Nrsf	145
4	Pol2	346	54	Elf1	144
5	Pol2	324	55	Foxm1	136
6	Pol2	324	56	Ets1	135
7	Taf1	316	57	Tblr1	133
8	Pol2	306	58	Max	132
9	Pol2	304	59	Znf1	126
10	Pol2	301	60	Sp1	125
11	Pol2	288	61	Elk1	124
12	Pol2	288	62	Jund	123
13	Pol2	287	63	Atf2	122
14	Pol2	284	64	Ctcf	120
15	Pml	283	65	Six5	120
16	Pol2	282	66	Tcf12	118
17	Elf1	280	67	Tcf12	117
18	Mxi1	276	68	Nrf1	116
19	Pol2	276	69	Cebpdsc636	115
20	Pol2	274	70	Nfic	115
21	Yy1	271	71	Tcf3	115
22	Yy1	269	72	Nfic	114
23	E2f1	266	73	Whip	112
24	Taf1	265	74	Cmyc	109
25	Sin3ak20	259	75	Pax5n19	109
26	Mxi1	257	76	Bhlhe40c	106
27	Tbp	249	77	GrDex100nm	106
28	Pol2	248	78	Bhlhe40c	105
29	Taf1	244	79	Chd1	105
30	Cmyc	239	80	Zbtb33	104
31	Maz	236	81	Cmyc	103
32	Runx3	230	82	Chd2	102
33	Creb1Dex100nm	229	83	Six5	102
34	Pol2	228	84	Srf	102
35	Cmyc	227	85	Nfyb	101
36	Sin3	227	86	Bclaf1	100
37	Yy1c	222	87	Rfx5	95
38	Nrsf	219	88	NfkbTnfa	94
39	Tbp	207	89	Ctcf	92
40	Gabp	204	90	Ctcf	89
41	Chd2	203	91	E2f4	88
42	Max	202	92	Mta3	87
43	Gabp	198	93	Ctcf	85
44	Elf1	197	94	Ebf1	85
45	Pou2f2	189	95	Kap1	82
46	Sp1	173	96	Pu1	81
47	Egr1	172	97	Ctcf	80
48	Gabp	166	98	Ets1	79
49	Max	158	99	Sin3ak20	79
50	Mybl2sc81192	153	100	Hdac2	78

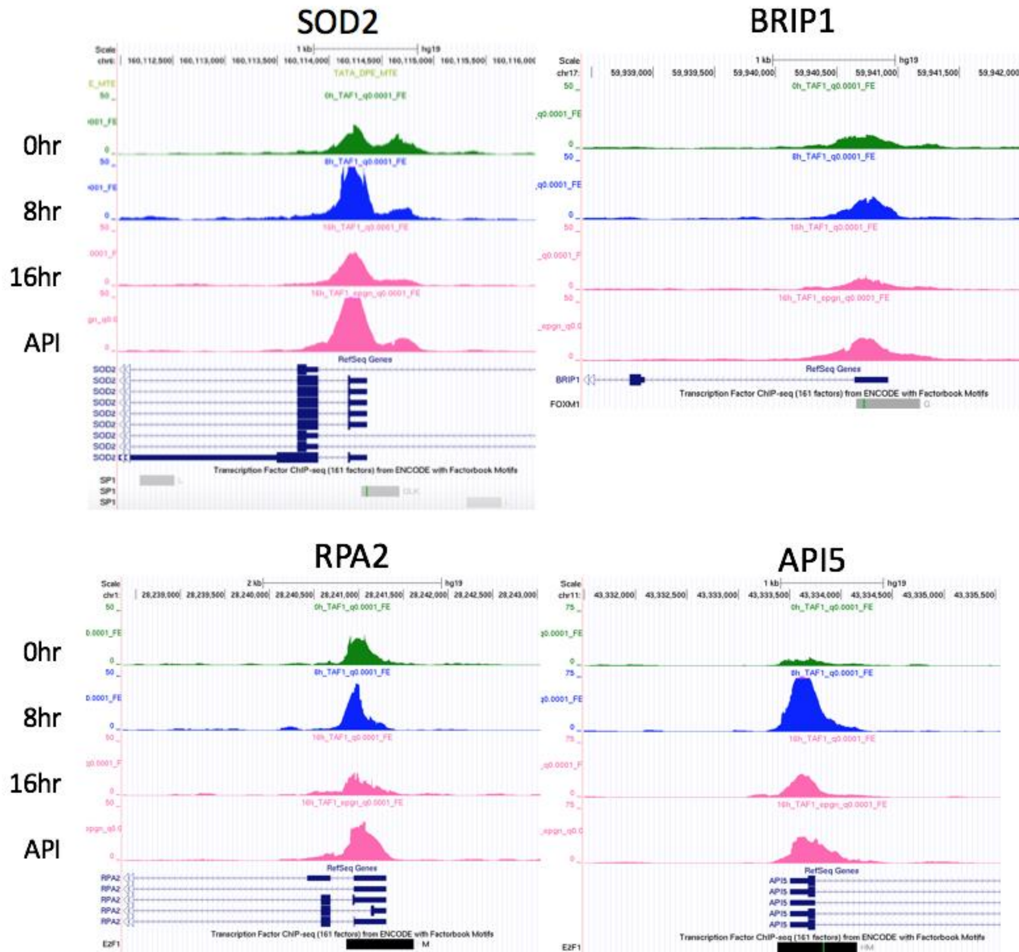
Legend: Table of the top 100 proteins bound to the 421 promoters. ChIP-seq's are ranked by the number of promoters they were found to bind to out of 421.

Table 2.3: Select transcription factor targets and their known functions

RANK	CHIP	PROMOTER (421)	Total Peaks:	Total Promoters:	Promoters:Total promoters	Interaction with GTM:	Function:
17	Elf1	280	23008	10816	0.025887574	no	Immune Response
21	Yy1	271	17876	9368	0.028928266	TAF1 (Y), TAF2(Y), TAF3(Y), TAF4 (Y+DM), TAF7(H + Y), TAF9/12 (H), TBP (Y,M,+D)	Regulating diverse number of promoters
23	e2f1	266	21946	9396	0.028309919	TBP (H)	Control of cell cycle
32	Runx3	230	67965	13564	0.01695665	no	Activate or suppress transcription of various promoters
33	Creb1	229	15913	9121	0.025106896	TAF4 (H)	Response to hormonal stimulation of cAMP pathway
38	Nrsf/REST	219	12828	7367	0.029727162	TBP(H), TAF1 (Y), TAF2(Y+D), TAF4 (D+Y), TAF6 (Y)	represses neuronal genes in non-neuronal tissues
40	Gabp	204	10109	7001	0.029138694	TBP (Y+D), TAF 8 (Y)	Activation of cytochrome oxidase expression and nuclear control of mitochondrial function
46	Sp1	173	18248	8564	0.020200841	TBP (H), TAF1 (Y), TAF4 (H), TAF2 (Y+D)	cell differentiation, cell growth, apoptosis, immune responses, response to DNA damage, and chromatin remodeling
47	Egr1	172	16331	8555	0.020105202	TBP (M, Y, D), TAF1 (Y), TAF2 (Y+D), TAF4 (Y+D)	Differentitation and mitogenesis
55	Foxm1	136	22926	7346	0.018513477	TBP(H)	Transcriptional activator involved in cell proliferation
61	Elk1	124	5584	4083	0.030369826	TBP(H)	Binds serum response factor in the promoter of the c-fos proto-oncogene. Is a nuclear target for the ras-raf-MAPK signaling cascade
62	Jund	123	21614	7832	0.015704801	TBP (H,Y), TAF1 (M,Y), TAF12(Y), TAF3 (Y,C), TAF5(H,Y)	protect cells from p53-dependent senescence and apoptosis
63	ATF2	122	23467	6993	0.017446017	TBP(Y)	binds to the cAMP-responsive element (CRE), HAT activity, DNA damage response
65	Six5	120	4839	3842	0.031233732	no	function in the regulation of organogenesis
67	TCF12	117	20896	7721	0.015153478	TBP(H), TAF4(H)	participate in regulating lineage-specific gene expression through the formation of heterodimers with other bHLH E-proteins
71	Tcf3	115	16021	6824	0.016852286	TAF4(H)	lymphopoiesis, and the encoded protein is required for B and T lymphocyte development
75	Pax5n19	109	19740	7167	0.015208595	TBP(H)	Paired box transcription factors are important regulators in early development
76	Bhlhe40c	106	14628	6967	0.015214583	no	This gene is believed to be involved in the control of circadian rhythm and cell differentiation
84	Srf	102	8544	5136	0.019859813	TBP(D), TAF1(Y,M), TAF6(Y), TAF7(Y)	Nuclear protein that is involved in cell cycle regulation, apoptosis, cell growth, and cell differentiation
85	Nfyb	101	13295	6542	0.015438704	TBP(H), TAF11(H), TAF12(H,Y), TAF6(H,Y), TAF1(Y)	can function as both an activator and a repressor, depending on its interacting cofactors

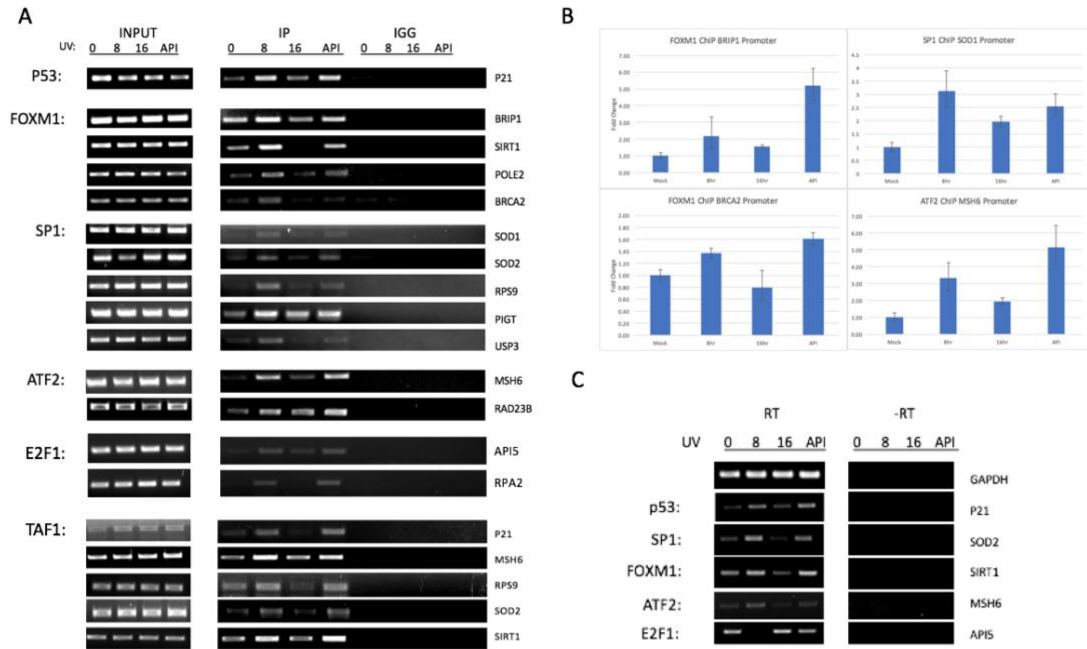
Legend: Sequence specific transcription factors ranked based the number of promoters they were found to bind to out of 421. Each ChIP-seq was analyzed for the total number of peaks in that sample and total promoters. A ratio of promoters/421 to total promoters in that sample is shown. Known interaction with general transcription machinery (Y=yeast, H= human, M= mouse, D= drosophila), and known function are shown.

Figure 2.6: ENCODE database shows DNA binding of representative transcription factors



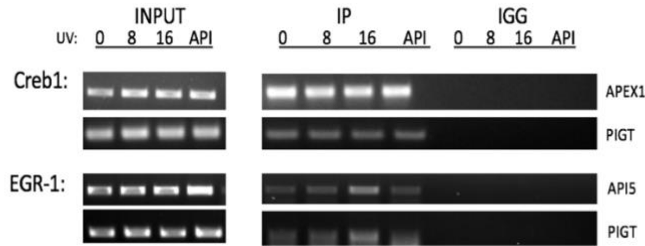
Legend: Genome browser shots of TAF1 ChIP seq at SOD2, BRIP1, RPA2, and API5 promoters. ENCODE database input shows known transcription factor binding as grey bars under promoter regions.

Figure 2.7: ChIP reveals positive correlation to p53 binding and target expression



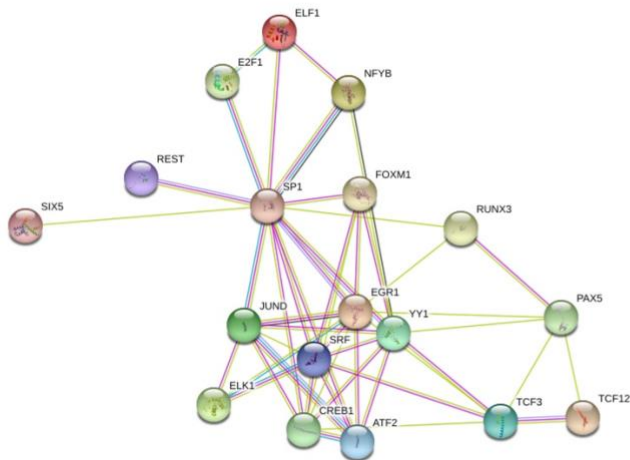
Legend: (A) ChIP was performed on proteins p53, FOXM1, SP1, ATF2, and TAF1 in U2OS cells at the indicated promoters under 0hr, 8hr, 16hr, and 16hrs of UV treatment + apigenin treatment. (B) qPCR verification of FOXM1, ATF2, and SP1 ChIP results following 0hr, 8hr, 16hr, and 16hrs of UV treatment + apigenin. (C) hnRNA levels of p53, SP1, FOXM1, ATF2, and E2F1 target genes following 0hr, 8hr, 16hr, and 16hrs UV treatment + apigenin. P21 was used as positive control and GAPDH used as loading control.

Figure 2.8: Not all transcription factors correlated with p53 trend following UV treatment



Legend: ChIP was performed on E2F1, Creb1, and EGR-1 proteins in U2OS cells at the indicated promoters following 0hr, 8hr, 16hr, and 16hrs of UV treatment + apigenin treatment.

Figure 2.9: String analysis of protein interactions

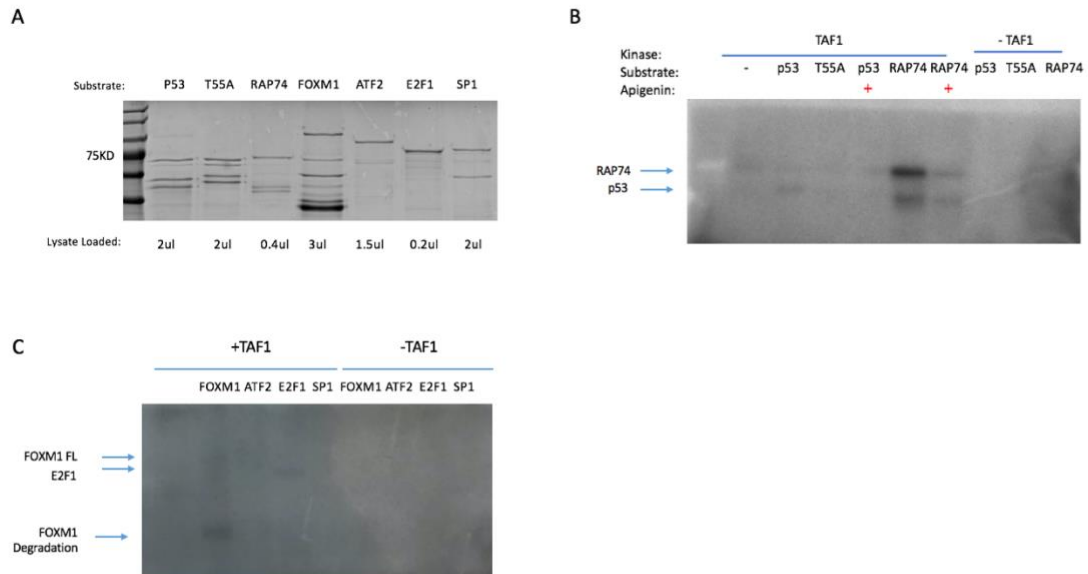


Legend: String analysis (version 10.5) of 18 transcription factors (Szklarczyk et. al., 2017) shows interaction between top candidate transcription factors.

Figure 2.10: Dissociation of FOXM1 on BRCA2 is [ATP] dependent

Legend: ChIP was performed on p53 and FOXM1 proteins in U2OS cells under indicated UV time titration with the addition of 4-AN or compound C (CC) on the indicated promoters.

Figure 2.11: TAF1 phosphorylation of FOXM1 and E2F1



Legend: HA-TAF1 kinase was purified via bacculovirus infection of sf9 cells. Protein substrates were expressed via BL21 and purified via GST or HIS-tag. (A) Coomassie blue staining of relative substrate concentrations used in kinase assay reactions. (B) Control panel of TAF1 in vitro phosphorylation assay of p53, T55A, and RAP74 with apigenin treatment as indicated. Phosphorylation was detected by autoradiograph. (C) TAF1 in vitro phosphorylation assay with substrates FOXM1, E2F1, ATF2, and Sp1. Phosphorylation was detected by autoradiography.

Chapter 3: Potential role for TATA element in response to DNA damage

The TAF1 ChIP-sequencing data utilized in this analysis was performed by Dr. Joy Lin. The RNA-sequencing was performed by Dr. Yong Wu. Scripts for the core promoter element analysis were written with the assistance of Dr. Patrick Schreiner. I was responsible for the ChIP-seq analysis, core promoter element abundance determination, and analysis of RNA-seq data overlapped with TAF1 ChIP-seq data.

3.1 Introduction

The regulation of transcription is a vital process for all species and the activation of transcription requires the coordinated recruitment of basal transcription machinery to the promoter. Basal transcription machinery consists of general transcription factor TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, and RNA polymerase II enzyme. The general transcription factor TFIID is composed of over 13 different TATA-binding protein Associated Factors (TAF) proteins and the TATA binding protein (TBP) (Burley and Roeder, 1996; Tora, 2002). TAF1 is the largest subunit of TFIID and has both direct and indirect interactions with specific sequences found at promoter regions, known as core promoter elements (Thomas and Chiang, 2006).

Core promoter elements are specific sequences located at fixed locations within the DNA and indicate the start site of transcription. The core promoter elements play an important role in recruitment of general transcription machinery. Therefore, understanding the composition of core promoter elements may provide insight to the regulation of a given promoter. While there are many steps for gene regulation in response to DNA damage, the role of core promoter elements tends to be overlooked. Therefore, investigating the core promoter element composition following DNA damage may provide further insight into this area.

Promoter regions may consist of the initiator (INR), TATA box, downstream promoter element (DPE), the TFIIB response element (BRE), and the motif ten element (MTE). The INR sequence is the most prevalent core promoter element (Ohler et al., 2002; Fitzgerald et al., 2006, Gershonzon et al., 2006), is the site of transcription initiation and is primarily recognized by TAF1/TAF2 proteins (Chalkey and Verrijzer, 1999). The TATA box is arguably the most known core promoter element, is found upstream of INR (Carninci et al., 2006, Ponjavic et al., 2006) and is recognized by TBP. The BRE element is recognized by the TFIIB upstream of the TATA box (Lagrange et al., 1998) and the DPE is found downstream of the INR and interacts with TFIID (Burke and Kadonaga, 1996). Lastly, the MTE is located upstream of the DPE and may also be a recognition site for TFIID as indicated by DNAse I footprinting (Thomas and Chiang, 2006).

Promoters vary in the composition of their core promoter elements, ranging from only containing one element sequence to containing all elements. Evidence has shown that the composition of core promoter elements has an association with specific gene function (Bae et al., 2015), rate of activation, the intensity, and the longevity the gene is expressed (Morachis et al., 2010). Specifically, the TATA element is found to be enriched in stress related genes, whereas growth-promoting genes are often TATA-less (Bae et al., 2015), indicating a potential role of TATA in response to stress. These findings indicate the importance of the composition of core promoter elements and that the

general transcription factors responsible for recognizing core promoter elements, may play a vital role in the DNA damage response.

Since TFIID recognizes core promoter elements and is essential in the nucleation of the preinitiation complex, necessary for transcription, we chose to investigate how the core promoter element profile effects the recruitment of TAF1 following DNA damage. While TAF1's role in transcription initiation has been determined, investigation into its recruitment to promoters through core promoter elements in response to DNA damage still remains to be elucidated. We thus investigated TAF1's global binding to core promoter elements as well as its effect on gene expression in response to UV-induced DNA damage. Our analysis indicates that genes that significantly increased in expression upon DNA damage and showed TAF1 recruitment were more likely to contain a TATA box. Further, genes that significantly decreased in expression upon DNA damage had a lower abundance of TATA-boxes. These data suggest that the presence of a TATA-box plays a key role in the TAF1 recruitment and provides insight into the gene regulation upon stress. Thus, by analyzing core promoter elements we may further be able to unravel the role of core promoter elements in the DNA damage response.

3.2 Results

3.2a Core promoter element profile of TAF1 bound promoters upon DNA damage.

The core promoter plays a role in general transcription factor machinery recruitment and transcription initiation. To study the composition of core promoter elements bound by TAF1 upon DNA damage, we performed TAF1 ChIP-seq under mock and DNA damage conditions (previously described in Dissertation Chapter 2). Analysis of TAF1 ChIP-seq results revealed 4187 TAF1 bound peaks within 0.2 kb of the transcription start site (TSS) under mock conditions and 3086 peaks following UV exposure (Figure 3.1B). This reduction in total TAF1 bound peaks following DNA damage may elute to the fact that transcription is ceasing on many promoters. We next performed the core promoter analysis via sequence specific searches using Seqinr (Charif and Lobry, 2007) through the Bioconductor (Gentleman, 2004). Our analysis revealed a general trend of increasing abundance of TATA, INR, BRE upstream, BRE downstream, DPE, and MTE present in the TAF1 bound promoter following UV treatment (Figure 3.1 A, B). The presence of MTE element was very low, making the analysis insignificant. However, in general, our results show there is a moderate increase in core promoter elements present at TAF1 bound promoters following DNA conditions indicating a possible role of core promoter elements in DNA damage response. While we observed an increase in core promoter elements following

DNA damage, we also observed a decrease in the number of total TAF1 bound peaks following DNA damage. This may indicate that while TAF1 dissociates from promoters in which transcription is ceased following DNA damage, core promoter elements may play a role in TAF1 recruitment to promoters following DNA damage.

3.2b Differential peak analysis of TAF1 ChIP-seq revealed a higher abundance of TATA elements present in TAF1-recruited promoters following DNA damaging conditions

Following the observation of higher core promoter element abundance in TAF1-bound promoters after DNA damaging conditions, we next analyzed the percentage of core promoter elements at promoters in which TAF1 significantly increased in binding from mock conditions to DNA damage conditions via differential peak calling analysis. This TAF1 recruitment was determined using differential peak binding analysis, MACS2 (Liu, 2014), for significant increase in binding between mock and DNA damage conditions. In total, there were 1923 TAF1 bound peaks that significantly increased from mock conditions to DNA damage conditions within 0.2 kb of the TSS. A visualization of TAF1 ChIP-seq result in the form of a heatmap (Liu et. al., 2011) is shown in Figure 3.2A and an example from the genome browser (Kent et. al., 2002) is shown in Figure 3.2C.

Intriguingly, the TATA element was the only core promoter element that exhibited a change in abundance in TAF1 recruited promoters as compared to

total promoters bound following DNA damage (Figure 3.2B). Those promoters in which TAF1 significantly increased in binding to following DNA damage showed a 1.7-fold higher abundance of TATA elements as compared to total TAF1 bound promoters following DNA damage (Figure 3.2 B). Since our results show an increase in TATA elements present when TAF1 is recruited to promoters following UV treatment, it indicates a possible role in TATA elements in the recruitment of TAF1 in response to stress.

The genes in which TAF1 was recruited to upon DNA damage were annotated using DAVID (Huang et al., 2009 (A); Huang et al., 2009 (B)). The top clusters of gene annotation (Figure 3.3D) showed notable categories such as metabolism, DNA repair, cell cycle regulation, and chromatin assembly. These clusters are expected following DNA damaging conditions. This data emphasizes the role of the TATA element and TAF1 recruitment upon DNA damage and possibly the regulation of gene expression.

3.2c RNA-Seq analysis indicates gene expression following UV treatment is partially reliant on the core promoter elements present

To further investigate the role of core promoter elements bound by TAF1, we aimed to determine the relationship between the core promoter elements and the increase or decrease in gene expression via RNA-seq. RNA-seq was performed by Dr. Yong Wu in U2OS cells under identical mock and DNA damage conditions (UV) and analyzed for genes that are significantly increased or

decreased upon DNA damage. Our analysis revealed 1636 genes that significantly increased and 2841 that were significantly decreased upon DNA damage (Figure 3.3A). While UV treatment has been shown to lead to RNA stability (Gowrishankar et al., 2005), possibly accounting for the increase in RNA levels, transcription also plays a role. For example, gene CDKN1A, also known as p21, is well known for its role in cell cycle arrest and shows significant activation following DNA damage via hnRNA levels (Wu et al., 2014) and shows an increase in total RNA levels in this RNA-seq data. Meanwhile, TGFB3, a gene known to be involved in cell growth, shows a significant reduction in RNA levels (Figure 3.3B). A deeper look into those genes affected by UV induced DNA damage revealed that activated or stabilized genes showed major categories including catabolism, oxidative phosphorylation, and response to stress. Alternatively, reduced or repressed genes included categories such as metabolism, regulation of splicing, and cell cycle progression (Figure 3.4) (DAVID; Hoang et al., 2009 (A), Hoang et al., 2009 (B))

Genes that increased or decreased in expression following UV treatment were then further analyzed by overlapping with TAF1 bound promoters following DNA damage. This resulted in 322 promoters that showed an increase in expression and were TAF1 bound following UV and 613 promoters that showed a decrease in expression and bound by TAF1 within 0.2 kb of the transcription start site (TSS) (Table 3.1). Intriguingly, TAF1 bound promoters following DNA damage, in which the RNA expression showed a significant increase, were

nearly 2-fold times more likely to contain a TATA element than total TAF1 bound promoters following DNA damage (Figure 3.5A). Further, genes in which their expression was reduced upon DNA damage and TAF1 was bound following DNA damage, were nearly 32 percent less likely to contain a TATA element.

This finding prompted investigation into the core promoter element abundance in promoters that TAF1 was recruited to following DNA damage and the corresponding gene expression. TAF1 peaks that significantly increased in binding from mock to DNA damage were identified, resulting in 1923 peaks within 0.2 kb of the TSS (Figure 3.2B). These TAF1 recruited promoters were then analyzed for their core promoter element abundance when gene expression was found to be increased (257 genes) or decreased (380 genes) in response to UV induced DNA damage (Table 3.1). We observed a greater than 4.5-fold increase in the presence of TATA elements when TAF1 was recruited to promoters and gene expression increased compared to TAF1 bound promoters in control conditions (Figure 3.5B). Therefore, we speculate that the presence of a TATA element may in involved in TAF1 recruitment to the promoter following DNA damage leading to the activation of a subset of genes. Further, DPE showed nearly 2.5-fold increase in TAF1 recruited promoters that showed an increase in expression. This striking increase in TATA element abundance in TAF1 recruited promoters upon cellular stress indicates the TATA elements involvement in cell stress response.

3.3 Discussion

The recruitment of basal transcription machinery to the start site of transcription plays an important role in gene expression. The carefully orchestrated recruitment is essential for rapid response in times of stress to ensure proper genome stability and cell survival. TFIID plays a major role in the recognition of core promoter elements and the diversity in core promoter structures plays a role in transcriptional regulation. Thus, we aimed to elucidate the role of core promoter elements in TAF1 recruitment to promoters following UV induced DNA damage. Further, whether core promoter elements could provide insight into the activation or repression of genes following TAF1 recruitment.

Genome-wide promoters bound by TAF1 in mock conditions comprised 1% TATA, nearly 50% BRE elements, 52% INR sequences and less than 10% DPE elements. However, we observed an increase in all core promoter elements bound by TAF1 following UV induced DNA damage. It has been shown that a single promoter containing a TATA box, Inr, MTE, and DPE, known as a super promoter, yields the highest levels of transcription (Juven-Gershon et al., 2006). Therefore, it is thought that the more core promoter elements present, the more highly regulated said gene is. This supports our finding that there are more core promoter elements present in TAF1 bound peaks following UV damage since

upon DNA damaging conditions, proper regulation of genes is vital for the integrity of the genome.

Previous studies have investigated the core promoter element abundance and have varying results. For example, one study found that about 32% of human genes contain a TATA box (Suzuki et al., 2001) verses, another found about 24% (Yang et. Al., 2007) and the most recent found that many previous analyses overestimated the percentage of promoters' dependent on a TATA element, and estimated only about 10% of human promoters are TATA dependent (Carninci et al., 2006). Further, it is suggested the mammalian genome contains ~30% BRE (Jin et al., 2006), ~70% INR elements (Carninci et al., 2006; Kim et al., 2005; Cooper et al., 2006) and an estimated ~20% DPE (Jin et al., 2006). While the percentage of core promoter elements seen in this analysis is low in comparison to the previously described genome wide core promoter element profile, this analysis shows only the percent bound by TAF1 which may account for the difference seen.

To investigate TAF1 recruitment to promoters upon DNA damaging conditions, we analyzed the promoters for core promoter elements present in which TAF1 significantly increased in binding to from mock to UV treated conditions. Of these 1923 promoters, the only core promoter element that showed a change in comparison to total TAF1 bound promoters following DNA damage, was the TATA element. The TATA element increased by 2.5-fold, indicating a possible role of the TATA element and TAF1 recruitment following

UV treatment. The involvement of the TATA element was further emphasized when gene expression upon UV treatment was investigated. Genes that significantly increased in expression and TAF1 was recruited to following UV treatment showed as much as 4.5 times higher likelihood of containing a TATA element. A previous study of the p21 promoter showed that containing a TATA box allows for a faster activation of transcription, a quintessential reaction to DNA damage. This rapid reaction is due to the preassembly of the pre-initiation complex already bound to the p21 promoter, more specifically, the TATA element of the core promoter (Morachis, 2010). Furthermore, promoters that do not contain a TATA element are slower to fire transcription (Morachis, 2010). Given this finding, it makes sense that upon DNA damage, we witness an increase in TATA elements bound by TAF1.

These research findings further emphasize the role of core promoter elements in stress response. By further unraveling the mechanisms behind cells response to stress, we can gain insight into many diseases.

3.4 Methods

ChIP-seq:

ChIP-seq was performed by Dr. Joy Lin. All ChIP-Seq data sets were aligned using Bowtie2 (version 0.12.8) (Langmead et al., 2009) to the human reference genome (GRCh37/hg19). The 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 binding peaks. The ratio of the peak pileup between 0 and 8hr samples was used to estimate activated TAF1 bindings at 8hr after DNA damage. The ratio of the peak pileup between 16 and 8hr was used to determine decreased TAF1 binding at 16hr after DNA damage. The peaks identified by ChIP-Seq were analyzed with the R Bioconductor package ChIPpeakAnno (Zhu et al, 2010) to retrieve the nearest Ensembl gene (10 kb around transcription start site, TSS).

RNA-Seq:

RNA-seq was performed by Dr. Yong Wu. Read alignments were performed via cufflinks (Trapnell et al., 2010) and differential expression analysis was determined using CuffDiff (Trapnell et al., 2013).

Core promoter identification:

The TAF1 peaks bound +/- .2kb from TSS were analyzed for core promoter element abundance. Core promoter elements were detected using Bioconductor (Gentleman, 2004) and the seqinr (Charif and Lobry, 2007) package. The core promoter elements were defined as follows:

BRE upstream: [GC][GC][GA]CGCC

TATA: TATA[AT]A[AT][AG]

BRE downstream: [GA]T[TGA][TG][GT][TG][TG]

MTE: C[GC]A[AG]C[GC][GC]AACG[GC]

DPE: [AG]G[AT]CGTG

INR: [CT][CT]A[ATCG][TA][CT][CT]

Generation of Heatmap:

We used the max-min 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 mock and UV treatment were entered into Cistrome along with a BED file containing the genomic regions centered at the summits of TAF1 peaks following UV induced DNA damage to generate heatmaps. In the heatmap representation, each row represents the ±2.5 kb centered on the summit of TAF1 enriched peak and ranked according to the enrichment of TAF1occupancy after DNA damage.

3.5 References

1. Bae, S.-H., Han, H.W., and Moon, J. (2015). Functional analysis of the molecular interactions of TATA box-containing genes and essential genes. PloS One 10, e0120848.

2. Burke TW, Kadonaga JT. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev 1996;10:711–724. [PubMed: 8598298]

3. Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A.M., Taylor, M.S., Engström, P.G., Frith, M.C., et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nature Genetics 38, 626–635.

4. Chalkley, G.E. and Verrijzer, C.P. 1999. DNA binding site selection by RNA polymerase II TAFs: a TAFII250-TAFII150 complex recognizes the initiator. EMBO J 18:4835–4845.

5. Charif, D., and Lobry, J.R. (2007). SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. In Structural Approaches to Sequence Evolution: Molecules, Networks, Populations, U. Bastolla, M. Porto, H.E. Roman, and M. Vendruscolo, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 207–232.

6. Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A., and Roeder, R.G. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. EMBO J 12:2749–2762.

7. Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Res 2006; 16:1–10. [PubMed: 16344566]

8. Dynlacht, Brian David, Timothy Hoey, and Robert Tjian. “Isolation of Coactivators Associated with the TATA-binding Protein That Mediate Transcriptional Activation.” Cell 66.3 (1991): 563-76

9. FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. Comparative genomics of Drosophila and human core promoters. Genome Biol 2006;7:R53. [PubMed: 16827941]

10. Gentleman, R., V. Carey, D. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, and Others. "Bioconductor: Open Software Development for Computational Biology and Bioinformatics." Genome Biology 5 (2004)

11. Gershenzon NI, Trifonov EN, Ioshikhes IP. The features of Drosophila core promoters revealed by statistical analysis. BMC Genomics 2006;7:161. [PubMed: 16790048]

12. Gowrishankar, G., Winzen, R., Bollig, F., Ghebremedhin, B., Redich, N., Ritter, B., Resch, K., Kracht, M., and Holtmann, H. (2005). Inhibition of mRNA deadenylation and degradation by ultraviolet light. Biological Chemistry 386.

13. A) Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57.

14. B) Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

15. Jin, V.X., Singer, G.A., Agosto-Pérez, F.J., Liyanarachchi, S., and Davuluri, R.V. (2006). Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs. BMC Bioinformatics 7, 114.

16. Juven-Gershon T, Cheng S, Kadonaga JT. Rational design of a super core promoter that enhances gene expression. Nat. Methods. 2006;3:917–922.

17. Kent W., C.W Sugnet, T.S Furey, K.M Rosekin, T.H Pringle, A.M Zahler, D. Haussler. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006

18. Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. A high- resolution map of active promoters in the human genome. Nature 2005;436:876–880.

19. 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 Dev 1998;12:34–44. [PubMed: 9420329]

20. Liu T, et al. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol. 2011;12:R83.

21. Liu, Tao, Jorge A. Ortiz, Len Taing, Clifford A. Meyer, Bernett Lee, Yong Zhang, Hyunjin Shin, Swee S. Wong, Jian Ma, Ying Lei, Utz J. Pape, Michael Poidinger, Yiwen Chen, Kevin Yeung, Myles Brown, Yaron Turpaz, and X. Shirley Liu. "Cistrome: An Integrative Platform for Transcriptional Regulation Studies." Genome Biology 12.8 (2011): R83.

22. Morachis, J.M., Murawsky, C.M., and Emerson, B.M. (2010). Regulation of the p53 transcriptional response by structurally diverse core promoters. Genes & Development 24, 135–147.

23. Nakajima, N., Horikoshi, M., and Roeder, R.G. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. Mol Cell Biol 8:4028–4040.

24. Ohler U, Liao GC, Niemann H, Rubin GM. Computational analysis of core promoters in the Drosophila genome. Genome Biol 2002;3:RESEARCH0087. [PubMed: 12537576]

25. Ponjavic J, Lenhard B, Kai C, Kawai J, Carninci P, Hayashizaki Y, Sandelin A. Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. Genome Biol 2006;7:R78. [PubMed: 16916456]

26. Purnell, B.A., Emanuel, P.A., and Gilmour, D.S. 1994. TFIID sequence recognition of the initiator and sequences farther downstream in Drosophila class II genes. Genes Dev 8:830–842.

27. Sawadogo, M. and Roeder, R.G. 1985a. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. Proc Natl Acad Sci USA 82:4394–4398.

28. Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., and Nakamura, Y. (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Research 11, 677–684.

29. Tanese, N., Pugh, B.F., and Tjian, R. 1991. Coactivators for a proline- rich activator purified from the multisubunit human TFIID complex. Genes Dev 5:2212–2224.

30. Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotech 31, 46–53.

31. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology 28, 511–515.

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

33. Zhou, Q., Lieberman, P.M., Boyer, T.G., and Berk, A.J. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. Genes Dev 6:1964–1974.

3.6 Figures and Tables

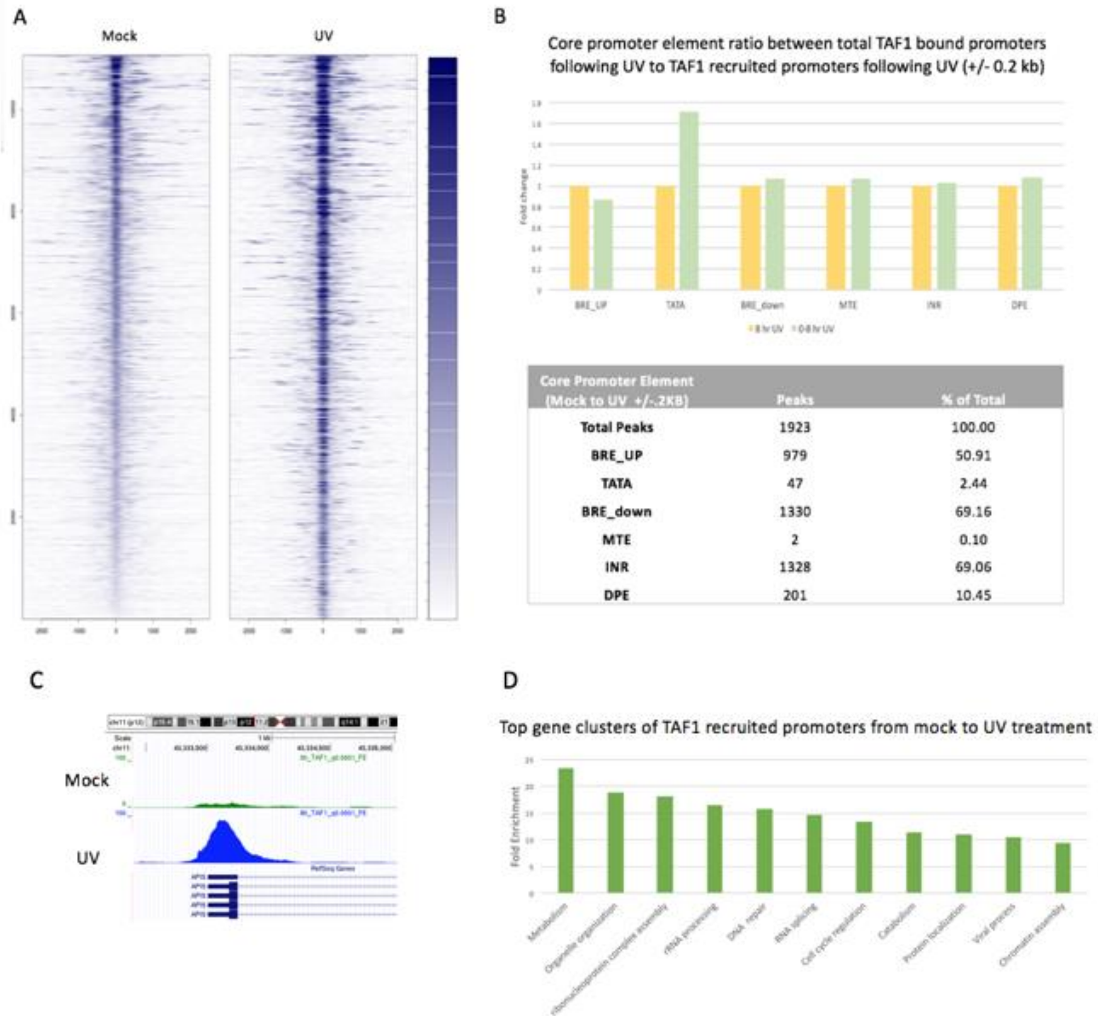
Figure 3.1: Higher core promoter abundance in the TAF1 bound promoters upon DNA damage

Core Promoter Element (0hr +/-.2KB)	Peaks	% of Total
Total peaks	4187	!00
BRE_UP	2006	47.91
TATA	42	1.00
BRE_down	2013	48.08
MTE	2	0.05
INR	2191	52.33
DPE	247	5.90

Core Promoter Element (8hr +/-.2KB)	Peaks	% of Total
Total peaks	3086	
BRE_UP	1817	58.88
TATA	44	1.43
BRE_down	2005	64.97
MTE	3	0.10
INR	2066	66.95
DPE	299	9.69

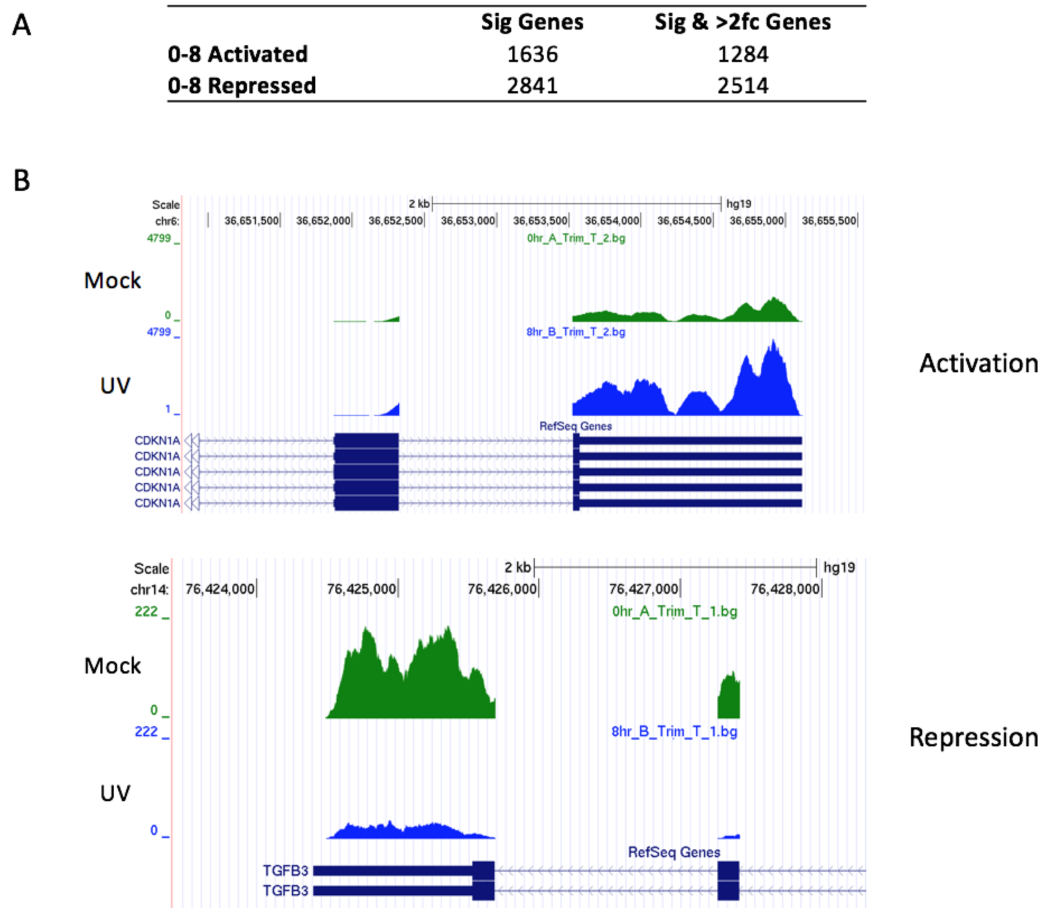
Legend: (A) Analysis of TAF1 ChIP seq reveals composition of each core promoter elements (BREu, BREd, TATA, MTE, INR, and DPE) on the TAF1-bound promoters. (B) The presence of core promoter elements (BREu, BREd, TATA, MTE, INR, and DPE) on the TAF1-bound promoters before (mock) and after UV treatment (UV) (B).

Figure 3.2: Genome-wide TAF1 recruitment to core promoter elements following UV induced DNA damage



Legend: (A) Heatmap of TAF1 ChIP-seq before (Mock) and after UV treatment. X-axis shows distance from the transcription start site and Y-axis is TAF1 peaks ordered by pileup number. (B) Top: Ratio of the presence of core promoter elements in total TAF1 bound promoters bound after UV compared to TAF1 recruited (Mock to UV) promoters. (C) Genome browser shot of TAF1 bound peak at API5 promoter before and after UV treatment. (D) DAVID analysis of 1923 TAF1 recruited promoters upon UV treatment ranked based on enrichment score.

Figure 3.3: RNA-seq analysis following UV treatment

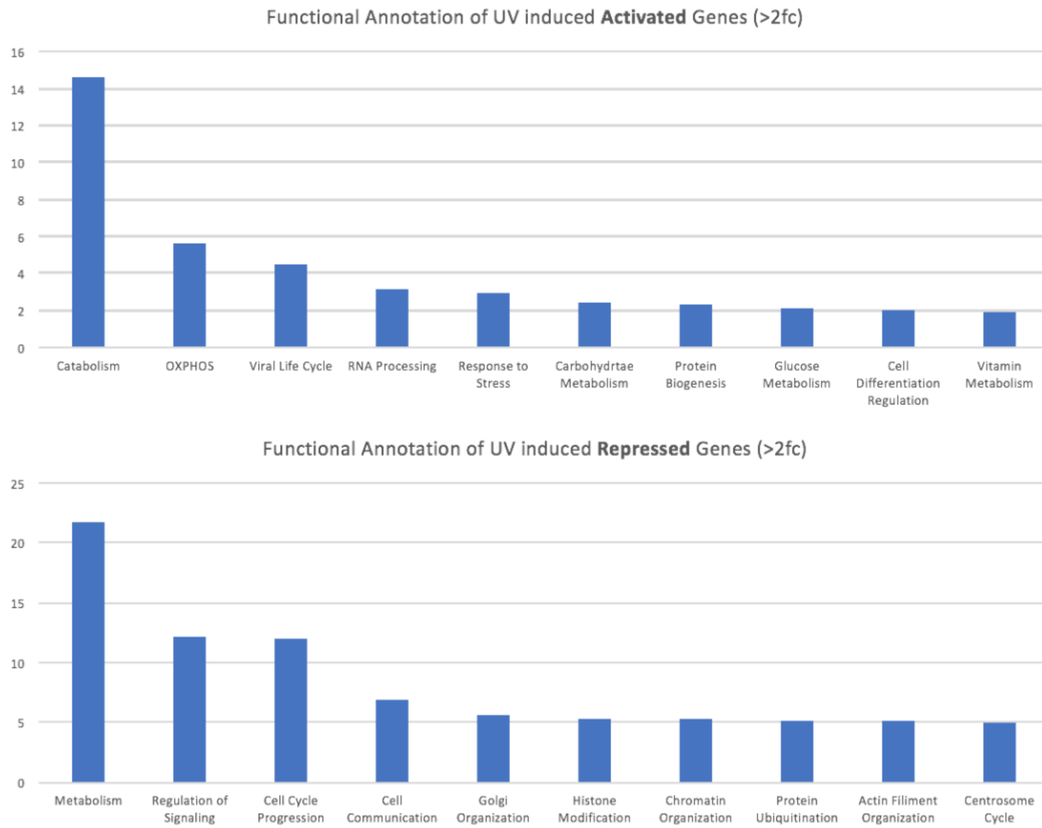


A

	Sig Genes	Sig & >2fc Genes
0-8 Activated	1636	1284
0-8 Repressed	2841	2514

Legend: (A) Summary of genes found to either significantly increase or decrease in response to UV in U2OS cells. (B) Genome browser shot shows activated (CDN1A) or repressed (TGFB3) genes.

Figure 3.4: Functional annotation of UV activated or repressed genes



Legend: DAVID functional analysis of activated or repressed genes (>2fc) following UV induced DNA damage. Function annotation categories are ranked from left to right based on enrichment score.

Table 3.1: Genes found to significantly increase or decrease in expression bound by TAF1 following UV treatment

	Sig. Genes
0-8 Increase in expression	1636
Increased/8hr TAF1 ChIP-seq	322
Increased/0-8hr TAF1 ChIP-seq	257
0-8 Reduced in expression	2841
Reduced/8hr TAF1 ChIP-seq	613
Reduced/0-8hr TAF1 ChIP-seq	380

Legend: RNA-seq was performed by Dr. Yong Wu. Total genes found to be significantly increased or decreased following UV treatment are identified. TAF1-ChIP seq was performed by Dr. Joy Lin. Analysis of data sets shows the number activated or repressed genes bound by TAF1 following UV treatment.

Figure 3.5: Core promoter elements present in activated or repressed genes following DNA damage that were TAF1 bound in mock conditions or TAF1 recruited following DNA damaging conditions

Increase in RNA (8hr +/-.2KB)

Core Promoter Element	Peaks	% of Total
Total	322	
BRE_UP	182	56.52
TATA	9	2.79
BRE_down	211	65.53
MTE	0	0
INR	229	71.12
DPE	32	9.94

Decrease in RNA (8hr +/-.2KB)

Core Promoter Element	Peaks	% of Total
Total	613	100.00
BRE_UP	376	61.34
TATA	6	0.98
BRE_down	390	63.62
MTE	1	0.16
INR	393	64.11
DPE	56	9.14

Increase in RNA (0-8 +/-.2KB)

Core Promoter Element	Peaks	% of Total
Total	257	100.00
BRE_UP	127	49.42
TATA	12	4.67
BRE_down	179	69.65
MTE	0	0
INR	183	71.21
DPE	36	14.01

Decrease in RNA (0-8 +/-.2KB)

Core Promoter Element	Peaks	% of Total
Total	380	100.00
BRE_UP	197	51.84
TATA	5	1.32
BRE_down	260	68.42
MTE	2	0.53
INR	258	67.89
DPE	32	8.42

Legend: (A) Frequency of core promoter elements present in all TAF1 peaks following UV treatment compared to TAF1 peaks shown to be activated or repressed. (B) Frequency of core promoter elements present in TAF1 recruited promoters (before or after UV) that were activated or repressed as compared to mock TAF1 bound peaks.

Chapter 4: Conclusion

4.1 Conclusions:

The TATA box binding protein (TBP) associated factor 1 (TAF1) is the largest member of TFIID (Kokubo et a., 1993) and contains intrinsic kinase activity (Dikstein et al., 1996). As a member of the general transcription machinery, TAF1 has been implicated as playing an important role in transcriptional regulation through its promoter binding activity and its interaction and modifications of activators. Here we show supporting evidence that TAF1 acts as a genome-wide transcriptional regulator through the phosphorylation of activators in response to UV-induced DNA damage. Further, we show that TAF1 was more likely to be recruited to promoters that contain a TATA element upon DNA damage.

The implications of this study support TAF1's known role in transcriptional regulation. Previous research has shown that TAF1 is recruited to promoters through the recognition of an activator (Wassarman and Sauer, 2001). In-depth studies have been performed on TAF1's interaction with p53, a tumor suppressor protein that aids in genomic stability in response to cell stress. Upon UV-induced DNA damage, p53 becomes acetylated. TAF1, through its double bromodomain, recognizes acetylated K373 and k382 of p53 (Li et al., 2007), recruiting TAF1 to the p21 promoter in early stages of DNA damage. p21 is responsible for binding to and inhibiting all cyclin dependent kinases responsible for cell cycle progression, thus inducing cell cycle arrest. In later stages of DNA damage,

TAF1 phosphorylates p53 at Thr55, causing p53/TAF1 to dissociate from the promoter, ceasing transcription of p21 allowing for cell cycle progression (Wu et al., 2014). TAF1's ability to act as this molecular switch in response to DNA damage relies on TAF1s intrinsic kinase activity. Basal levels of cellular ATP in U2OS cells are 1.13 +/- 0.01 mM, however, TAF1 has a Km of 1.9 mM (Wu et al., 2014), thus, requiring high cellular ATP concentrations. In late stages of DNA damage, cellular ATP concentrations increase allowing TAF1 kinase to be activated. This regulation via TAF1 kinase activity was further verified when TAF1 kinase activity was inhibited via the use of apigenin. With the addition of apigenin, phosphorylation of p53 at Thr55 nor dissociation from the p21 promoter was observed (Wu et al., 2014). While TAF1's regulation of p53 in response to DNA damage has been determined, a genome-wide analysis of TAF1's kinase targets would provide a more comprehensive understanding of TAF1s importance in DNA damage response.

Our genome-wide analysis of TAF1 kinase transcriptional regulation in response to DNA damage supports the aforementioned research. We conclude that the transcription factors E2F1, FOXM1, ATF2, and SP1 follow a similar binding pattern following DNA damage as p53. Further, following the addition of apigenin dissociation from target promoters in late stages of DNA damage was no longer seen. Each of the aforementioned transcription factors are known to play an important role in cell cycle regulation or DNA damage response. We further show, via *in vitro* kinase assays, that TAF1 is directly phosphorylating

E2F1 and FOXM1. While we did not observe phosphorylation of SP1 or ATF2, TAF1 may be regulating transcription via other interacting factors. This finding promotes the idea that TAF1 is acting as a genome-wide regulator of transcriptional termination in response to DNA damage through its kinase activity.

TAF1 is also known to be a regulator of transcription through its DNA binding activity. The composition of core promoter elements, found in promoter regions, is thought to regulate gene expression (Juven-Gershon and Kadonaga, 2010). Promoters may consist of the initiator (INR), motif ten element (MTE), TFIIB response element (BREupstream, BREdownstream), downstream promoter elements (DPE), and the TATA element (Thomas and Chiang, 2006). It is believed that the more core promoter elements present, the higher the level of regulation that is associated with the gene (Juven-Gershon et al., 2006). Our data supports that theory, in that all core promoter elements bound by TAF1 were found in higher abundance following DNA damage, as opposed to mock conditions. Further, the TATA element, specifically, has been implicated in playing a role in gene expression in response to cellular stress, and is found enriched in genes known to be involved in cellular stress (Bae et al., 2015). It was also found that containing a TATA element led to increased intensity of gene expression (Morachis et al., 2010). Our data suggests that upon DNA damage, promoters in which TAF1 was recruited to and showed increased gene expression were 4.5-fold times more likely to contain a TATA element than TAF1

promoters under mock conditions. Therefore, we conclude that TAF1 is bound to more core promoter elements following DNA damage, and most strikingly, the TATA element. These findings support this idea of core promoter elements playing a role in stress response and leads to the further understanding of TAF1's role in DNA binding following DNA damage. This further implicates core promoter elements in stress response and TAF1 recruitment.

Insights provided by this investigation could lead to further interesting studies. While we showed that TAF1 directly phosphorylates E2F1 and FOXM1 via *in vitro* studies and showed their dissociation from promoters was TAF1 kinase dependent in *in vivo* experiments, *in vitro* binding assays (such as EMSA) would show conclusive evidence of TAF1's direct regulation of transcription factor binding. This would eliminate the possibility of TAF1 acting through other factors to regulate the transcription factor binding and show a more direct regulation via TAF1 kinase. Further, since TAF1 contains two kinase domains, the NTK and CTK (Dikstein et al., 1996; O'Brien and Tjian, 1998), it would be interesting to determine which kinase domain was responsible for the phosphorylation of FOXM1 and E2F1. It has been previously shown that the NTK domain is responsible for the phosphorylation of RAP74 and TFIIA (Wu et al., 2014; Ruppert and Tjian, 1995; Dikstein et al., 1996; Solow et al., 2001). While the CTK shows a lesser ability to transphosphorylate but exhibits strong autophosphorylation (unpublished data; Landon Piluso). This data indicates that

the NTK domain is likely responsible for the phosphorylation of E2F1 and FOXM1.

In order to determine a physiological relevance of TAF1's phosphorylation of E2F1 and FOXM1, identification of the phosphorylation sites would be necessary. A mutation located at this phosphorylation site of FOXM1 and E2F1 could then lead to the discovery of genes negatively regulated via TAF1 kinase activity. E2F1 is known to inhibit apoptosis and cell cycle promoting genes (Biswas et al., 2011) to allow for DNA repair (Biswas et al., 2014) in response to DNA damage. Our results witnessed a recruitment of E2F1 to promoters, however, we observed a decrease in their expression, supporting previous findings. TAF1 phosphorylation may play an important role in the activation of these genes in a genome-wide manner to allow for cell death or cell cycle continuation. Further, since FOXM1 is known to play a role in the regulation of many DNA repair genes (Zona et al., 2014), it would be interesting to determine the genome-wide importance of TAF1 phosphorylation of FOXM1 in DNA repair. Since TAF1 is a significantly mutated gene across several cancer types (Kandoth et al., 2013), these findings could lead to the further understanding of TAF1's role in tumorigenesis. As well, this would provide a more global understanding of the effect of TAF1 kinase activity.

It was shown via studies of TAF1's regulation of p53 in response to DNA damage that the timing of TAF1's phosphorylation of p53 at Thr55 is important to maintaining genomic integrity. This was exhibited through p53's regulation of the

gene p21, responsible for enforcing cell cycle arrest in response to UV-induced DNA damage. This timing of TAF1's phosphorylation of p53 is quintessential to the DNA damage response and relies on TAF1's kinase activity acting as a molecular switch in response to fluctuating cellular ATP concentrations (Wu et al., 2014). This leads to the possibility that conditions that may affect cellular ATP levels will affect TAF1 kinase activity and thus response to DNA damage response. Diseases such as hyperglycemia, a hallmark sign of diabetes, is characterized as having high blood glucose levels. It was shown that high glucose levels lead to an increase in cellular ATP levels (Wu et al., 2014). Such high ATP levels activate TAF1 kinase activity, leading to the phosphorylation of p53 and subsequent transcriptional termination of p21 (Wu et al., 2014). There is a known association between diabetic patients and increased risk in certain cancers such as liver cancer and colon cancer (Harding et al., 2015). Therefore, it is plausible that this increased risk of cancers in diabetic patients may be caused by TAF1. If ATP levels are already high, caused by the hyperglycemia, then this may cause premature phosphorylation via TAF1. This would terminate the transcription of genes important in maintaining genomic integrity. According to our data, there are likely many TAF1 kinase targets including the already identified p53, RAP74, TFIIA, H3/H4, E2F1, and SP1.

This work could be further extrapolated to mouse models. Upon identification of TAF1 phosphorylation sites on identified transcription factors E2F1 and FOXM1, xenografts of stable phospho-site mutant cell lines on diabetic

mice could be performed. By comparing xenograft tumors in normal vs. diabetic mice we could potentially determine the gravity that TAF1 kinase has in tumor formation in diabetic states.

If TAF1 is to be regarded as a master transcriptional regulator in response to DNA damage via its kinase activity we would expect that TAF1 is responsible for the modification of multiple transcription factor targets, beyond what was discovered in this study. Our lab has previously described the importance of TAF1 modification of p53 in response to DNA damage (Wu et al., 2014), and here we describe other TAF1 kinase targets E2F1 and FOXM1. It would provide insight to mutate the TAF1 phosphorylation sites on E2F1 and FOXM1 and determine the effect on cell cycle regulation and DNA repair. Taken together, these studies could provide transformative insight into DNA damage response and the correlation between diabetes and cancer.

4.2 References

1. Bae, S.-H., Han, H.W., and Moon, J. (2015). Functional analysis of the molecular interactions of TATA box-containing genes and essential genes. PloS One 10, e0120848.

2. Biswas, A.K., and Johnson, D.G. (2012). Transcriptional and Nontranscriptional Functions of E2F1 in Response to DNA Damage. Cancer Research 72, 13–17.

3. Biswas, Anup K., David L. Mitchell, and David G. Johnson. "E2F1 Responds to Ultraviolet Radiation by Directly Stimulating DNA Repair and Suppressing Carcinogenesis." Cancer Research 74 (2014)

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

5. Harding, J.L., Shaw, J.E., Peeters, A., Cartensen, B., and Magliano, D.J. (2015). Cancer Risk Among People With Type 1 and Type 2 Diabetes: Disentangling True Associations, Detection Bias, and Reverse Causation. Diabetes Care 38, 264–270.

6. Juven-Gershon T, Cheng S, Kadonaga JT. Rational design of a super core promoter that enhances gene expression. Nat. Methods. 2006;3:917–922

7. Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Developmental Biology 339, 225–229.

8. Kokubo, T., Swanson, M.J., Nishikawa, J.-I., Hinnebusch, A.G., and Nakatani, Y. 1998. The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. Mol Cell Biol 18:1003– 1012.

9. Li, Heng-Hong, Andrew G., Landon G. Piluso, Xin Cai, Brian J. Gadd, Andreas G. Ladurner, and Xuan Liu. "An Acetylation Switch in P53 Mediates Holo-TFIID Recruitment." Molecular Cell 28.3 (2007): 408-21.

10. Morachis, J.M., Murawsky, C.M., and Emerson, B.M. (2010). Regulation of the p53 transcriptional response by structurally diverse core promoters. Genes & Development 24, 135–147.

11. O'Brien, T. and Tjian, R. (1998). Functional analysis of the human TAFII250 N-terminal kinase domain. Mol. Cell 1, 905-911.

12. Ruppert, S., and R. Tjian. "Human TAFII250 Interacts with RAP74: Implications for RNA Polymerase II Initiation." Genes & Development9.22 (1995): 2747-755.

13. Solow, S., Salunek, M., Ryan, R. and Lieberman, P. M. (2001). TAFII250 phosphorylates human TFIIA on serine residues important for TBP binding and transcription activity. J. Biol. Chem. 276, 15886-15892.

14. Thomas, M.C., and Chiang, C.-M. (2006). The General Transcription Machinery and General Cofactors. Critical Reviews in Biochemistry and Molecular Biology 41, 105–178.

15. Wassarman, D.A., and Sauer, F. (2001). TAFII250. Journal of Cell Science 114, 2895–2902.

16. Wu, Y., Lin, J.C., Piluso, L.G., Dhahbi, J.M., Bobadilla, S., Spindler, S.R., and Liu, X. (2014). Phosphorylation of p53 by TAF1 Inactivates p53-Dependent Transcription in the DNA Damage Response. Molecular Cell 53, 63–74.

17. Zona, S., Bella, L., Burton, M.J., Nestal de Moraes, G., and Lam, E.W.-F. (2014). FOXM1: An emerging master regulator of DNA damage response and genotoxic agent resistance. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1839, 1316–1322.