THE YEAST TRANSCRIPTION FACTOR GAL4:

A MODEL FOR UNDERSTANDING EUKARYOTIC TRANSCRIPTION

APPROVED BY SUPERVISORY COMMITTEE

Thomas Kodadek, Ph.D.

George DeMartino, Ph.D.

Eric Olson, Ph. D.

Raymond MacDonald, Ph.D.

DEDICATION

At the beginning of my graduate studies I joined the lab of Dr. Thomas Kodadek. He has mentored my growth as a scientist while also providing a laboratory environment that promotes creativity. Thank you. I also want to acknowledge my current and former committee members; Dr. George DeMartino, Dr. Raymond MacDonald, Dr. Eric Olson and Dr. Stephen Johnston. I appreciate the suggestions and support they have provided throughout my graduate career.

The Kodadek laboratory has seen many scientists come and go over the years; all of which have contributed to numerous scientific discussions. I especially want to thank my bay mates; Dr. Chase Archer, Dr. Anne Gocke and Dr. Lyle Burdine. Through-out my time in Tom's laboratory these three have contributed to many invaluable discussions and suggestions. I will miss you and our many happy hours.

Most importantly, I want to acknowledge my husband Justin, his friendship and support are what helped me survive the last six years. Without you, I would not be where I am today. I also want to thank the rest of my family and friends for their support. Even though they did not always understand my research, our parents would always ask how things were going and were nice enough to listen to me. Our friends are scattered across the Dallas-Fort Worth area and they always found time to meet and hang-out; for that I am very grateful.

THE YEAST TRANSCRIPTION FACTOR GAL4

A MODEL FOR UNDERSTANDING EUKARYOTIC TRANSCRIPTION

By

Melissa Ann O'Neal

DISSERTATION

Presented to the Faculty of the Graduate School of Biological Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

May, 2009

THE YEAST TRANSCRIPTION FACTOR GAL4:
A MODEL FOR UNDERSTANDING EUKARYOTIC TRANSCRIPTION

Melissa Ann O'Neal, Ph.D.

University of Texas Southwestern Medical Center at Dallas, 2009

Thomas Kodadek, Ph.D.

The 26S proteasome regulates numerous cellular pathways, including transcription, through proteolytic and non-proteolytic methods. The Kodadek and Johnston laboratories recently established a novel function for the proteasomal ATPases: the destabilization of activator-DNA complexes. This is independent of proteolysis but requires direct activator-ATPase interactions as well as ATP hydrolysis. The Gal4 mutant Gap71, which is hyper-sensitive to destabilization from a *GAL* promoter, was instrumental to this discovery. Gal4, but not Gap71,

was mono-ubiquitylated in a HeLa nuclear extract and *in vivo*, suggesting that mono-ubiquitylation of an activator is critical to resisting destabilization by the proteasomal ATPases. To gain a better understanding of these events, the three amino acid substitutions in the Gap71 DNA-binding domain were individually cloned and analyzed for their contributions to the function of Gal4. The data showed that Serine 22 and Lysine 23 but not Lysine 25 were important for the efficiency of the activator. The charge at Lysine 23 was found to be important for Gal4-based transcription and subsequent *in vitro* work revealed that Gal4 was not only phosphorylated at Serine 22 but that this phosphorylation event was essential for the function of the activator. Many times a phosphorylation event precedes a mono-ubiquitylation event on an activator. Knowing the kinase and ligase machinery that modifies Gal4 would permit us to further test our model. As a result, I designed selection screens in an attempt to isolate the kinase and/or ligase machinery components that modify Gal4. While these particular enzymes were not identified, other novel genes were found to negatively affect the galactose utilization pathway, *MSU1* and *SPS1*. Altogether, the data demonstrated that two post-translation events, phosphorylation and mono-ubiquitylation, prevent an activator-DNA complex from being disrupted, leading to an elegant model in which the proteasomal ATPases act as an important check point in transcription.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Ferdous A, O'Neal M, Nalley K, Sikder D, Kodadek T, Johnston SA. Phosphorylation of the Gal4 DNA-binding domain was essential for activator mono-ubiquitylation and efficient promoter occupancy. Mol Biosystems. 2008 Nov;4(11):1116-25. Epub 2008 Aug 26.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

6AU – 6-Azauracil

19S RP – 19S regulatory particle

20S CP – 20S core particle

AAA-ATPases – (ATPases Associated with various cellular Activity)-ATPases

Ab – Antibody

AD – activation domain

bp – base pair

ChIP – Chromatin Immunoprecipitation

CIP – Calf Intestinal Phosphatase

CTD – C-terminal domain of RNA polymerase II

DBD – DNA-binding domain

DIM – dimerization domain

E1 – ubiquitin activating enzyme

E2 – ubiquitin conjugating enzyme

E3 – ubiquitin ligase

E4 – multiple ubiquitin chain assembly factor

ERα – estrogen receptor alpha

GTF – general transcription factor

HA tag – hemagglutinin tag

HECT domain – homologous to E6-AP carboxyl terminal domain

IP – immunoprecipitation

IPC – Internal Primer Control

Kb – kilobase

ORF – open reading frame

PI – protease inhibitors

PIC – preinitiation complex

PMSF – phenylmethanesulfonyl fluoride

RNAP II – RNA polymerase II

RT – room temperature

UAS – consensus upstream activating sequence

UAS_G – consensus upstream activating sequence 5'-CGG-$N_{(11)}$-CCG-3' for the activator Gal4

Ub – ubiquitin

WCE – whole cell extract

CHAPTER I

INTRODUCTION

Transcription

Transcription is an elaborate, multi-step process that creates RNA in a DNA-template fashion. A large amount of the regulation of gene expression occurs at the level of transcription. The core component of the "protein machine" that creates messenger RNA (mRNA) is RNA polymerase II. This enzyme transcribes the DNA template with the assistance of several complexes termed general transcription factors (GTF) [1-3].

The transcription cycle proceeds through four distinct stages (for a review see [4]). The first is the formation of a pre-initiation complex (PIC) on the promoter of a gene. This complex is comprised of the general transcription factors TFIIB, D, E, F and H along with the RNA polymerase (RNAP) II enzyme [5]. The TFIIE and TFIIH complexes are thought to be recruited last to the PIC and are involved in the 'melting' of the DNA at the promoter [6, 7]. This intermediary phase breaks the hydrogen bonds between the base pairs in the duplex DNA, which moves the transcriptional machinery into the second stage called initiation. At this point RNAP II begins generating short transcripts as it is

still tethered to the PIC. RNAP II must clear the promoter so it can move along the remainder of the DNA template. The mechanism by which promoter clearance occurs is still unclear but recent literature suggests Ctk1 is a major contributor to the release of RNAP II from the PIC [8]. Once the RNAP II disengages the promoter, it proceeds along the length of the gene for the third stage termed elongation. While the promoter clearance mechanism is ambiguous, one distinction between initiation and elongation is found on the carboxy-terminal domain (CTD) of RNAP II. This domain contains a heptapeptide repeat (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) that occurs as many as 26 times in yeast and is only hyper-phosphorylated upon entering the elongation phase of transcription (for review see [9]). As with initiation, elongation is augmented with additional complexes such as FACT to improve the efficiency and specificity of RNAP II [10]. These are thought to be largely involved in the monumental task of helping the polymerase move through the many nucleosomes in its path during elongation. Termination is the last stage that occurs when the full-length transcript is completed. The nascent transcript is then cleaved and processed in preparation for translation by the ribosome.

Most eukaryotic transcription events are activated by one or more gene-specific transcription factors that have two main components. One is a DNA-binding domain (DBD) that recognizes an upstream activation sequence (UAS) or enhancer. The DNA-binding domain dictates the genes that are to be regulated by

the activator and gives the cell finite control of the transcriptional process. The second component is an activation domain (AD) within the transcription factor. This domain is responsible for recruiting the transcriptional machinery once the protein has bound the UAS site within the targeted promoter [11].

Gal4 protein is a well studied transcription factor from *Saccharomyces cerevisiae* that has become a paradigm for understanding the mechanism of eukaryotic transcriptional activators. This protein is comprised of 881 amino acids and multiple domains (Figure 1). It has a zinc-finger DNA-binding domain located in the N-terminal region of the protein that recognizes the consensus upstream activating sequence 5'-CGG-$N_{(11)}$-CCG-3' (UAS_G: N = any nucleotide) [12-15]. Known *GAL* promoters contain 1-4 copies of this UAS. When multiple UAS_G sites are present, the activator can bind to the promoter cooperatively [16]. Gal4 dimerizes before it binds a UAS_G site through a small dimerization domain downstream the DNA-binding domain. The first activation domain is located downstream the dimerization domain and is relatively uncharacterized. The second activation domain is about 34 amino acids in length and is located near the C-terminal region of the Gal4 protein [17]. The latter is the activation domain responsible for recruiting the transcriptional machinery to the promoter.

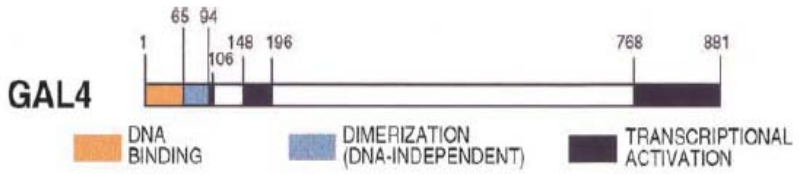

Figure 1: Diagram of the domains in the transcription factor Gal4 [18]. This protein is 881 amino acids long and has four characteristic domains: a DNA-binding domain, a dimerization domain, and two activation domains. Copy-Right Permission Granted.

Gal4 is essential for yeast survival when the sole carbon source for the organism is the sugar galactose. This is because the activator regulates the majority of the genes in the galactose utilization pathway in *S. cerevisiae*. A simplistic model of Gal4-based regulation is shown in Figure 2. The expression of genes in the galactose utilization pathway is suppressed in glucose-containing media. To ensure the system remains off, the protein Mig1 is expressed and it binds near the UAS_G sites to ultimately prevent binding by any residual Gal4 protein in the cell (Top) [19, 20]. When raffinose is the only sugar source present, the activator becomes poised at the UAS_G sites, however transcription is repressed by another interaction that occurs between the protein Gal80 and the Gal4 C-terminal activation domain (Middle) [17, 21]. When galactose is the sole carbon source, the protein Gal3 quickly releases the Gal80 based-repression and transcription occurs within minutes (Bottom) [22, 23]. While research has shown the importance of Gal3 in the regulation of Gal4 the mechanism by which this occurs is still being debated. This well defined and easily manipulated regulation pathway is what makes Gal4 amenable to transcription studies.

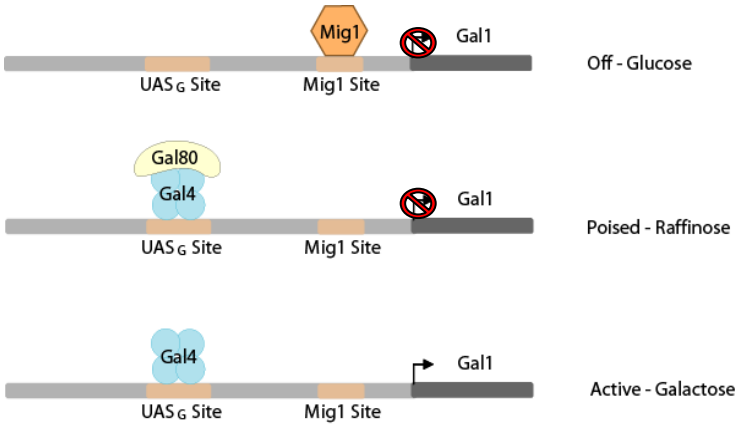

Figure 2: The Gal4-based transcription in yeast. When glucose is present in the media Mig1 prevents Gal4 from occupying the UAS_G sites. Raffinose permits Gal4 to bind the UAS_G sites but the activator is muzzled by the protein Gal80. The addition of galactose relieves Gal80 suppression and permits active transcription.

When *GAL* gene expression is induced with galactose, the Leloir pathway is activated to convert the galactose into glucose-6-phosphate so that the molecule can more efficiently generate energy for the cell (Figure 3, for a review see [24]). Gal2 is a permease that transports galactose across the membrane into the cell [25, 26]. Once in the cell, the enzyme galactose mutarotase (Gal10) converts the sugar population into the α-D-galactose anomer so that the molecule can be recognized and phosphorylated by the enzyme galactokinase (Gal1). This product is then modified by galactose-1-phosphate uridyl transferase (Gal7) to become glucose-1-phosphate. Gal10 provides two distinct activities in the Leloir pathway. The galactose mutarotase domain mentioned previously is located at the N-terminal of the protein while a UDP-glucose-4-epimerase domain is at the C-terminal of the Gal10 protein. The UDP-glucose-4-epimerase recycles by-products from the

Gal7-mediated reaction. Phosphoglucomutases (Gal5 and PGM1) then convert the Gal7 product glucose-1-phosphate into glucose-6-phosphate. The genes *Gal1, Gal2, Gal7,* and *Gal10* are controlled by UAS_G sites within their promoter. Therefore, deficiencies associated with the Leloir pathway or the activator Gal4 lead to decreased metabolism of galactose. This in turn causes a decrease in ATP production and ultimately cell death when galactose is the sole carbon source for the cells.

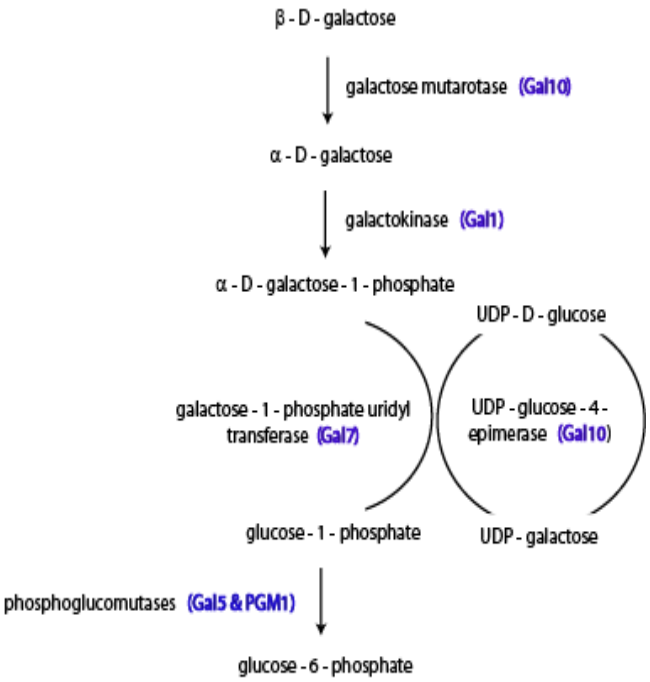

Figure 3: The Leloir Pathway. This diagram shows the mechanism for conversion of galactose into glucose-6-phosphate. This figure is modified from www.yeastgenome.org.

As with all transcription factors the activator Gal4 requires co-activators to function efficiently. There are two main categories of transcriptional co-activators. One class is defined as those that interact with RNAP II and the GTFs, an example being the Mediator complex. This co-activator comprises multiple proteins that arrange into three distinct sub-complexes: the head, the middle and the tail module. The head sub-complex binds the carboxyl-terminal domain (CTD) of RNAP II [27]. The middle sub-complex bridges the head and tail sub-complexes and binds both the CTD of RNAP II and the general transcription factor TFIIE [28]. The tail sub-complex consists of four proteins, including Gal11, that together recognize and bind activators, thus bridging the transcriptional machinery to the transactivator [29-31]. The Mediator has many layers of regulation that control how the complex contributes to transcription as both a co-activator and co-repressor (for reviews see [32, 33]). The other class of co-activators, which includes SAGA, is comprised of enzymes that modify chromatin. The ~1.8 MDa SAGA complex is recruited to the promoter of a gene by an activation domain within a transcription factor. Once established at the promoter, the enzymes within SAGA can induce several histone modifications such as acetylation and phosphorylation. In addition, it can remove histone modifications such ubiquitin. It is thought that the pattern of modifications established by this complex assists in making the chromatin accessible to the transcriptional machinery (for a review see [34]).

Over a decade ago experiments by Johnston and co-workers led to the identification of a novel co-activator in yeast, the proteasomal ATPases. The original study established an activator-proteasome link by using a mutant of *Gal4* called $gal4^{D}$ that lacked two-thirds of the C-terminal activation domain. Figure 4 highlights the amino acid composition of the Gal4 AD and illustrates the truncation associated with the $gal4^{D}$ genotype. $Gal4^{D}$ exhibits only 3% of the activity of the wild-type activator *in vivo* when expressed at physiological levels [35, 36]. *In vitro*, the protein binds to UAS_G sites as well as the wild-type protein, suggesting that the basis of its low level of activity is a poorly functional activation domain. A selection was carried out in an attempt to identify extragenic suppressors of $Gal4^{D}$. The idea was that an extragenic mutation might increase the ability of a co-activator to bind the fragment of the AD present in $Gal4^{D}$ and thus restore *GAL* gene transcription. Only two extragenic suppressor mutations were isolated, *sug 1-1* and *sug 2-1*. The genes were later determined to encode ATPases within the 26S proteasome, suggesting that the proteasome might be intimately involved in RNA polymerase II transcription [37-40].

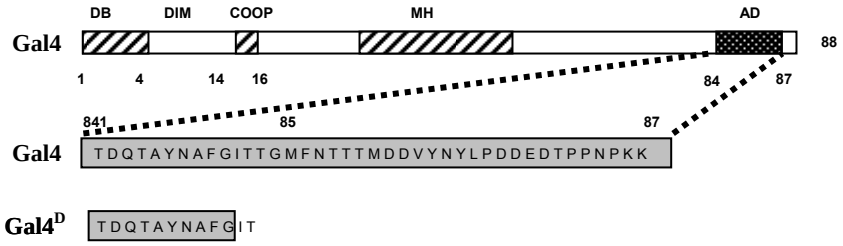

Figure 4: Diagram of Gal4^{D} and how the protein compares to wild-type Gal4 protein. The Gal4 protein contains 881 amino acids while the Gal4^{D} protein contains 853 amino acids. Two-thirds of the classical acidic-activation domain on the Gal4 C-terminal is missing.

The Ubiquitin-Proteasome System

The 26S proteasome is a large (~2 MDa), multi-protein complex that is responsible for most non-lysosomal degradation in a cell (Figure 5) [41, 42]. The proteolytic function of this massive complex is linked to the regulation of many cellular processes including cell cycle progression and transcription [35, 43].

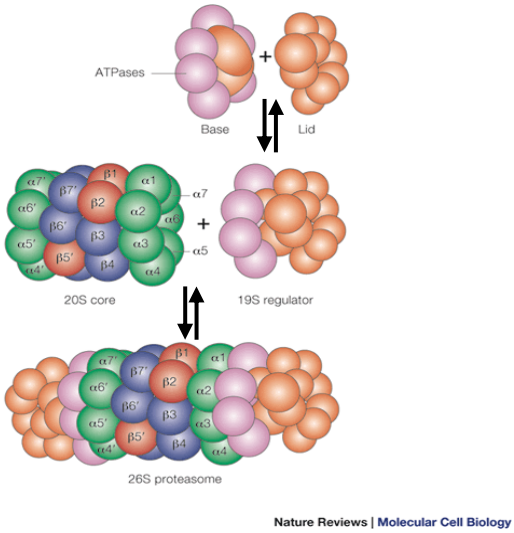

Figure 5: The yeast 26S proteasome. The classical representation of the proteasome is a 20S catalytic core flanked by two 19S regulatory particles. The 19S regulatory particle can be further broken down biochemically into two sub-complexes, the base and lid. This figure is modified from [44]. Copy-Right Permission Granted (#2158860756338).

The 26S proteasome is composed of two main subunits, the 19S regulatory particle (RP) and the 20S core particle (CP). The 20S CP is comprised of two copies each of seven alpha (α) proteins and seven beta (β) proteins assembled into four stacked heteroheptameric rings to create a barrel-shaped assembly (Figure 5).

The enzymatic activities of the 20S CP that are responsible for protein degradation are housed within the two interior β-rings [45]. The outer α-rings have a centrally located pore that permits the entrance of unstructured proteins and the subsequent release of degradation products [46]. This 20S CP is flanked by capping particles that regulate the entrance of proteins into the 20S CP pore. The most common capping particle is the 19S RP, which can be found on one or both ends of the 20 CP [47]. While the 20S CP is viewed as one complex, the 19S RP can be further broken down into a lid and base sub-complex (Figure 5). The lid sub-complex comprises nine Rpn (Regulatory particle non-ATPase) proteins. The base sub-complex of this multi-protein complex has a hexameric ring of AAA (ATPases Associated with various cellular Activities)-class ATPases [48, 49]. As a whole, the 19S RP dictates the proteins that will be degraded through the recognition and binding of poly-ubiquitylated chains linked to the proteins destined for destruction [50]. Rpn10 is one of the proteins involved in substrate recognition [51, 52]. For a protein to be recognized by Rpn10 it has to be modified with a Lys^{48}-linked poly-ubiquitin chain containing at least four ubiquitin monomers [53]. This poly-ubiquitin chain is subsequently removed by the 19S RP so that the ubiquitin can be recycled back into the cellular milieu [54, 55]. The doomed protein is then denatured by the proteasomal ATPases and fed into the pore of 20S CP for degradation [56].

Ubiquitin is a small protein (76 amino acids) that meidates several functions throughout the cell. The yeast genome contains four genes that encode ubiquitin: *UBI1*, *UBI2*, *UBI3* and *UBI4*. The first three are single ubiquitin sequences with intermittent introns and are fused to unrelated genes. *UBI4* has five ubiquitin moieties tethered in a row without any unrelated genes and is induced with various forms of stress [57]. In all cases, the mono-ubiquitin moiety is generated by cleavage of the original fusion protein. A multi-step reaction then attaches the ubiquitin molecule to proteins via an isopeptide bond between a lysine on a substrate protein and the carboxyl terminal of ubiquitin (Figure 6) [58, 59]. The reaction can cease at this point or the mono-ubiquitin can be extended into various forms of poly-ubiquitin chains [60, 61].

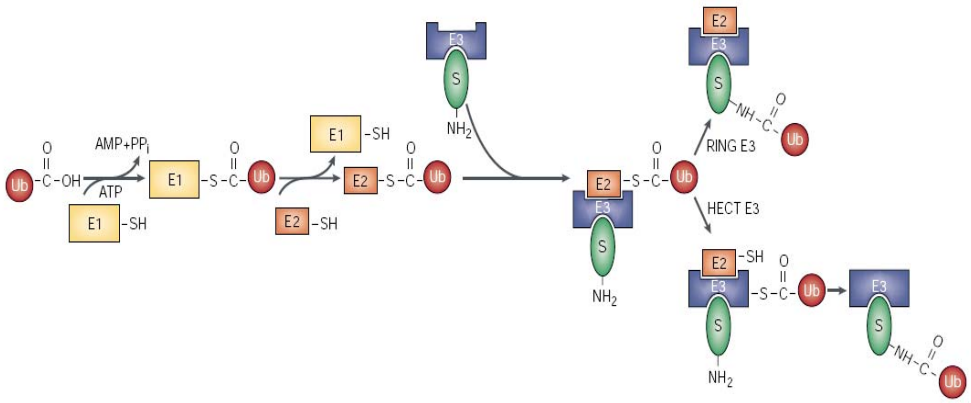

Figure 6: The ubiquitylation reaction. The diagram shows the various steps that contribute to a single ubiquitin moiety becoming covalently attached to the targeted protein [62]. Copy-Right Permission Granted (#2158860448231).

The ubiquitylation reaction begins with an ubiquitin activating enzyme (E1). The cysteine in the active site of this enzyme becomes covalently attached to the ubiquitin through a thioester linkage. Once this bond is established, the ubiquitin is transferred through transthioesterification to the active site cysteine of a ubiquitin conjugating enzyme (E2). The last step involves an ubiquitin ligase (E3 ligase), of which there are two classes. One is represented by the class of ring finger E3 ligases. These complexes recognize the substrate and interact with the E2 enzyme. The protein arrangement then facilitates the transfer of ubiquitin from the E2 enzyme to the targeted protein [63]. The other class includes the HECT (homologous to E6-AP carboxyl terminal) domain E3 ligases. This type of ligase facilitates the transfer of the ubiquitin from the E2 into a third thioester bond with itself. The HECT domain E3 ligase then directly ubiquitylates the targeted protein [64]. There are more than two classes of E3 ligases, however it is predominantly thought that all of the classes behave in one of the two ways stated above. In all causes the mono-ubiquitin becomes attached covalently to the ε-amino group of a lysine side chain [58, 59].

UBA1 is the only known yeast ubiquitin activating enzyme (E1) and is thought to interact with all of the ubiquitin conjugating enzymes (E2) [65]. A database analysis revealed eleven ubiquitin conjugating enzymes that facilitate interactions with the hundreds of ubiquitin E3 ligases found in the cell. This multitude of E3 ligases is what ultimately gives the ubiquitylation reaction

specificity, which is important because the this reaction regulates many cellular functions such as EGFR internalization and DNA repair (for a review see [66]).

A protein can maintain a mono-ubiquitylated status or a chain can be extended through the action of a multi-ubiquitin chain assembly factor (E4). Multiple forms of poly-ubiquitin chains are found within the cell and are formed by the conjugation of the C-terminus of one ubiquitin to one of the seven lysine residues (6, 11, 27, 29, 33, 48 and 63) within another ubiquitin (for reviews see [67, 68]). In addition, the lysines used to generate these linkages direct the activity of the modified protein. Two of the more prevalent poly-ubiquitin linkages in the cell are at the ubiquitin residues Lys^{48} and Lys^{63}. The addition of at least four or more ubiquitin molecules in a Lys^{48}-linked poly-ubiquitin chain is the most common method to tag proteins for destruction by the 26S proteasome [53]. A Lys^{63}-linked poly-ubiquitin chain is the other prominent linkage found in the cell and has been implicated in numerous functions including receptor internalization and signaling [69, 70].

The multi-step ubiquitylation reaction conjugates ubiquitin to a protein, which can then be reversed in the presence of a de-ubiquitylase (DUB), a hydrolase that cleaves the isopeptide bond. The 19S regulatory particle of the proteasome maintains DUB activity to de-ubiquitylate a protein substrate before it is denatured and fed into the 20S core particle, thus recycling ubiquitin back into the cellular milieu [51, 52]. Another aspect of de-ubiquitylases is represented by

CYLD. This type of de-ubiquitylase antagonizes signaling pathways that modulate processes such as mitosis and T-cell development (for a review see [71]). The ubiquitylation reaction can lead to the mono-, multi- or poly-ubiquitylation of a protein. In addition, the ubiquitylation pattern of the protein can be altered by de-ubiquitylase enzymes. This in turn creates a dynamic protein modification system that gives the cell yet another level of regulation.

The Ubiquitin-Proteasome System and Transcription

The primary function of the proteasome is the degradation of proteins. Research over the last decade has shown that the proteasome plays important roles in transcription in part through this proteolytic activity, but also via novel non-proteolytic mechanisms.

Some activators are degraded continually by the proteasome, which keeps genes targeted by that activator silent. Stimulated signaling pathways can in turn trigger the poly-ubiquitylation and ultimate degradation of inhibitory proteins within a pathway to prevent the degradation of the transcription factor, leading to the activation of target genes. This is a more classical example of how proteasome function is used to activate transcription and it can be found in both the WNT signaling cascade and the NF-κB pathway (for a review see [72]). There is also evidence that proteasome-mediated turnover of activators is essential for their function. This was first suggested by the discovery that 'degron' sequences overlap with the activation domains of transcription factors such as Myc [73, 74]. These data suggested that an activator was destroyed as a result of transcriptional activation and more specifically that poly-ubiquitylation of the activator was a necessary consequence of driving the transcription cycle. This theory was further supported by studies of the yeast transcription factor GCN4 and the mammalian transcription factor estrogen receptor alpha (ERα) [75, 76]. In the case of ERα, ChIP assays determined that the transcriptional machinery

and the proteasomal ATPases were cycling on and off the promoter in response to estradiol. When the cells were treated with estradiol and MG132, the half-life of ERα was prolonged, and surprisingly, transcription of ERα–responsive genes was decreased strongly. Moreover, MG132 abolished the cyclic pattern of protein association with the promoter, suggesting this dynamic behavior is essential for transcription. Why this would be the case is not clear, but one idea is that the activity of the activator is reduced by a modification that occurs during transcription, which in turn requires proteolysis to clear the promoter and thus allow additional transcription cycles [75].

While studies have shown a connection between the ubiquitin-proteasome system and active transcription, components of the two processes can also interact independent of proteolysis. Some of the first evidence to suggest this occurred was a study of the artificial activator LexA-VP16 that suggested the active form of the fusion protein was mono-ubiquitylated. In an effort to identify the E3 ubiquitin ligase that was responsible for transcription-dependent turnover of LexA-VP16 in yeast, Tansey and co-workers examined the activity of LexA-VP16 in yeast strains deleted for known E3 ubiquitin ligases and found that this activator worked poorly in the absence of the Met30 protein [77]. The level of LexA-VP16 was much higher in this strain than in wild-type yeast, arguing that Met30-mediated LexA-VP16 ubiquitylation was essential for both activator function and turnover. A major surprise came when these researchers probed the

nature of the ubiquitylation process necessary for LexA-VP16 function. They discovered that the genetic fusion of a single ubiquitin molecule to the N-terminus of the protein substantially reactivated LexA-VP16-driven transcription, but that the protein level remained high. They suggested that the genetic Ub-LexA-VP16 fusion was a mimic of natively mono-ubiquitylated LexA-VP16 and that this was the functional form of the activator. It is important to point out that no direct evidence was presented for natively mono-ubiquitylated LexA-VP16 in this study. However, mono-ubiquitylated mammalian activators such as CIITA and TAT were later characterized directly in other systems, lending credence to this hypothesis [78, 79]. It is not clear how the mono-ubiquitylation of LexA-VP16 stimulated its function, though it was reported later that this form of the activator more efficiently recruits the elongation factor P-TEFb to promoters [80]. While the function of an activator may be coupled tightly to it turnover, this study demonstrated that it is not a mechanistic requirement for LexA-VP16. This is because in the *Δmet30* strain Ub-LexA-VP16 levels are high, indicating poor turnover of the protein, yet it exhibits relatively robust activity.

The breakthrough discovery that mono-ubiquitin can contribute to the function of an activator led to an explosion in the field of ubiquitin. One recent study analyzed the co-activator SRC3 which is over-expressed in about 60% of tumors [81]. This oncogenic protein is regulated by a phosphorylation event that triggers mono-ubiquitylation of SRC3 at several different residues. It was found

that the E3 ligase Fbw7α mono-ubiquitylated a lysine in the degron region and that this event subsequently led to the degradation of the protein. Two lysines in the receptor interaction domain were also mono-ubiquitylated by Fbw7α and were shown to improve the efficiency of SRC3 as a co-activator [82]. These studies showed how mono-ubiquitin positively affects the efficiency of LexA-VP16 and SRC-3.

Another interesting aspect of mono-ubiquitylation is that the modification can induce the translocation of a protein from the nucleus to the cytoplasm, which has the effect of down-regulating an activator. Smad4 facilitates the induction of target genes in the TGF-β signaling pathway. It was recently discovered that Smad4 was mono-ubiquitylated and that this modified form of the protein was not found in the co-activator complex that induces target genes of the TGF-β signaling pathway. It was determined that the mono-ubiquitylation event caused the Smad4 protein to be exported to the cytoplasm, preventing the completion of the TGF-β signaling pathway [83]. In parallel, the oncogenic protein p53 also undergoes nuclear export when mono-ubiquitylated, though different mono-ubiquitylation events have a stimulatory effects [84, 85]. Taken together, these studies highlight the complexity of activator regulation through mono-ubiquitylation, which can either positively or negatively effect activator function. While activator mono-ubiquitylation has only been studied carefully for a handful of a transcription factors, it is likely to be widespread within the cell. The

prevalence of this modification makes determining how mono-ubiquitylation affects co-activators and activators such as SRC-3 and p53 even more intriguing.

An explanation for the positive regulation of activators by mono-ubiquitylation lays in their interaction with the proteasomal ATPases. As mentioned previously, a functional link between Gal4 and the proteasomal ATPases was suggested by the discovery of specific alleles of *SUG1* and *SUG2* that could suppress the growth on galactose deficiency exhibited by $gal4^{D}$ strains [35-39]. Based on these results, experiments were designed to test the hypothesis that the 19S regulatory particle played a role in transcription at the *GAL1-10* promoter. These studies found that strains harboring certain alleles of *Sug1* were sensitive to 6-azauracil (6AU), an inhibitor of transcriptional elongation, suggesting a defect in this process. More importantly, it was found that shifting a *sug1-20* strain, which carries a temperature-sensitive allele of the gene, to the non-permissive temperature resulted in strong inhibition of *GAL* gene transcription. The same procedure using a strain carrying a temperature-sensitive allele of the 20S sub-unit Pre1 did not affect *GAL*-based transcription. This novel discovery suggested that the ATPases may function in transcription independently of the proteolytic core complex. To probe the possible mechanistic basis of the reliance of *GAL* transcription on Sug protein function, pulse-chase experiments were conducted with whole cell extracts (WCE) prepared from a strain carrying the temperature-sensitive *sug1-20* allele. The data showed that inactivation of the

Sug1-20 protein resulted in the production of primarily short transcripts, again suggesting a role for the ATPase in elongation [86, 87]. The production of long transcripts was rescued by the addition of purified 19S RP to the extract containing inactivated Sug1-20 protein, demonstrating that this is likely to be a direct effect of this complex on the elongation process.

The Gal4^{D} phenotype was the result of a truncation of the C-terminal activation domain (AD) of Gal4 (Figure 4). Mutant alleles of two different proteasomal ATPases rescued this phenotype, suggesting an interaction occurred between the proteasomal ATPases and the full-length C-terminal AD of Gal4. To test this hypothesis, pull down assays were performed using a fusion protein with GST and the Gal4 C-terminal activation domain (GST-AD). The data from this *in vitro* study showed that the Gal4 AD interacts with components of the 19S base sub-complex (Rpts 1-6) but not the 19S lid sub-complex (Rpns 9 and 12) or the 20S CP (Pre1). To verify that these interactions were biologically relevant, *in vivo* data were collected through chromatin immuno-precipitation assays (ChIP) analysis of the *GAL1-10* promoter during inducing conditions. These results found that the base sub-complex of the 19S RP (Rpts 5 and 6) was recruited to the *GAL1-10* promoter but not components from the lid sub-complex or the 20S CP [88]. In addition, subsequent cross-linking studies demonstrated an interaction between the Gal4 activation domain and the two proteasomal ATPases Rpt4 and Rpt6 [89]. The conclusions deduced from these data were that the proteasomal

ATPases play a positive role in transcriptional elongation and that the Gal4 AD interacts with Rpt4 and Rpt6 to recruit the proteasomal ATPases to the promoter.

The base sub-complex of the 19S RP was characterized as playing a positive role in transcription. While these results were intriguing, the data did not explain the initial analysis that showed the recessive mutations *sug1-1* or *sug2-1* rescued the $Gal4^{D}$ phenotype, which was no growth on galactose. Based on that study, the 19S RP should exhibit an inhibitory role in transcription. As a result, additional experiments were performed and the data revealed that the proteasomal ATPases can destabilize transcriptional activators from a promoter [90]. This work also found that the artificial activator GST-Gal4(DBD)-VP16 was removed from UAS_G sites in the presence of competitor DNA *in vitro*. To augment these findings, the *in vitro* assay was repeated with a proteasomal ATPase preparation that had been pre-incubated with an antibody to the ATPase Sug1. If the proteasomal ATPases were disrupting the protein-DNA interaction, then tittering out one of the ATPases should prevent the destabilization of the activator-DNA complex. This modified preparation of the 19S RP was unable to destabilize the GST-Gal4-VP16 fusion protein from the DNA template, confirming the idea that the proteasomal ATPases negatively control transcription by affecting the stability of the activator-DNA interaction. Additional experiments found that the dissociation of an activator from a DNA template by the proteasomal ATPases was dependent on an activation domain and ATP. Surprisingly, it was determined

that the GST-Gal4-VP16 fusion protein that was not destabilized from the DNA template was mono-ubiquitylated. The conclusions from these data sets were that the proteasomal ATPases participated in transcriptional elongation and the formation or stabilization of the pre-initiation complex [90]. The idea that the proteasomal ATPases can function independent of the 20S CP was intriguing and was further supported by a ChIP on chip analysis done by a fellow member of the Kodadek laboratory. As expected, this global study revealed that many genes had a signal for both a 19S (Sug1) and a 20S (Pre1) component, suggesting the presence of 26S. A second and third set of genes demonstrated a signal for either the 19S or 20S sub-complex but not both. Additional ChIP on chip experiments utilized the temperature sensitive mutants *sug1-20* and *pre1-1/4-1*. These strains further demonstrated that a subset of the genes with both 19S and 20S signal were functioning independent of the intact 26S proteasome [91].

One of the key reagents in discovering the destabilization of activators by the proteasomal ATPases was a mutant of *GAL4* termed *gap71*. This mutant was originally created to study how zinc finger transcription factors behave. The *gap71* allele was developed by changing three amino acids within the *GAL4* DBD to create a sequence similar to the DBD of *PPR1*; serine 22 to aspartic acid (S22D), lysine 23 to glutamine (K23Q) and lysine 25 to phenylalanine (K25F) [92]. When GST-Gap71-VP16 was used in the destabilization assay it was hyper-sensitive to dissociation from the UAS_G sites by the proteasomal ATPases and

was not mono-ubiquitylated, making the differences between Gal4 and Gap71 ideal for studying the stability of the activator-DNA interaction [90]. The original phenotype of Gap71 proved interesting, and at first confusing. *In vitro* analysis of the Gap71 protein showed that it bound DNA as well as the wild-type protein, but was none the less inactive *in vivo* [93]. An explanation for the Gap71 phenotype is that the activator-DNA complex is disrupted by the proteasomal ATPases, however no *in vivo* studies had been done. To test this hypothesis, chromatin immuno-precipitation (ChIP) assays were performed to analyze the occupancy of the full-length proteins Gal4 and Gap71 on UAS_G sites under induced and non-induced conditions. The wild-type activator is poised on the promoter when raffinose is the sole carbon source for the cells and transcription is induced upon induction with galactose-containing media, which results in exposure of the activation domain as Gal80-mediated repression is suppressed. Gal4-dependent ChIP signals were monitored as a function of times after induction with the sugar galactose (Figure 7). The Gal4 protein showed a robust signal on the promoter when the yeast were grown on raffinose or galactose (middle). The No Ab control showed no signal for occupancy on the promoter (bottom). The mutant Gap71 showed a unique profile. It occupied well under raffinose conditions but was quickly destabilized from the promoter upon induction by galactose (top). The ChIP assay was repeated in a strain with the *sug1-20* allele. When that strain was shifted to a non-permissive temperature, the Gap71 protein was destabilized

less quickly, suggesting that the dissociation of activator-DNA complexes by the proteasomal ATPases occurred *in vivo* [90].

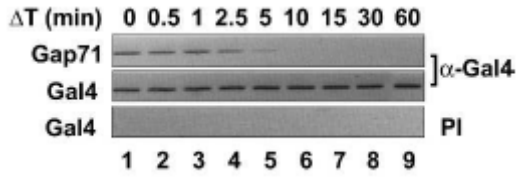

Figure 7: ChIP assay for activators on the *GAL1-10* promoter. The wild-type activator Gal4 maintained occupancy of the promoter. The Gap71 protein was removed from the promoter upon induction with galactose. Experiment was done by Dr. Devanjan Sikder [90]. Copy-Right Permission Granted.

The *in vivo* and *in vitro* data suggested that Gap71, but not mono-ubiquitylated Gal4, was destabilized from a DNA template by the proteasomal ATPases. To understand the process by which the activator-DNA interaction is disrupted, better characterization of Gap71 was necessary. A simplistic model is that mono-ubiquitin and other post-translational modifications contribute to the stability of the activator-DNA interaction (Figure 8). The activation domain recruits the proteasomal ATPases that destabilize the activator from the promoter. Gal4 protein, which is mono-ubiquitylated, would remain bound to the DNA so that the activator appears to maintain constant occupancy of the promoter (left). An activator that is not mono-ubiquitylated would be dissociated from the DNA by the proteasomal ATPases (right).

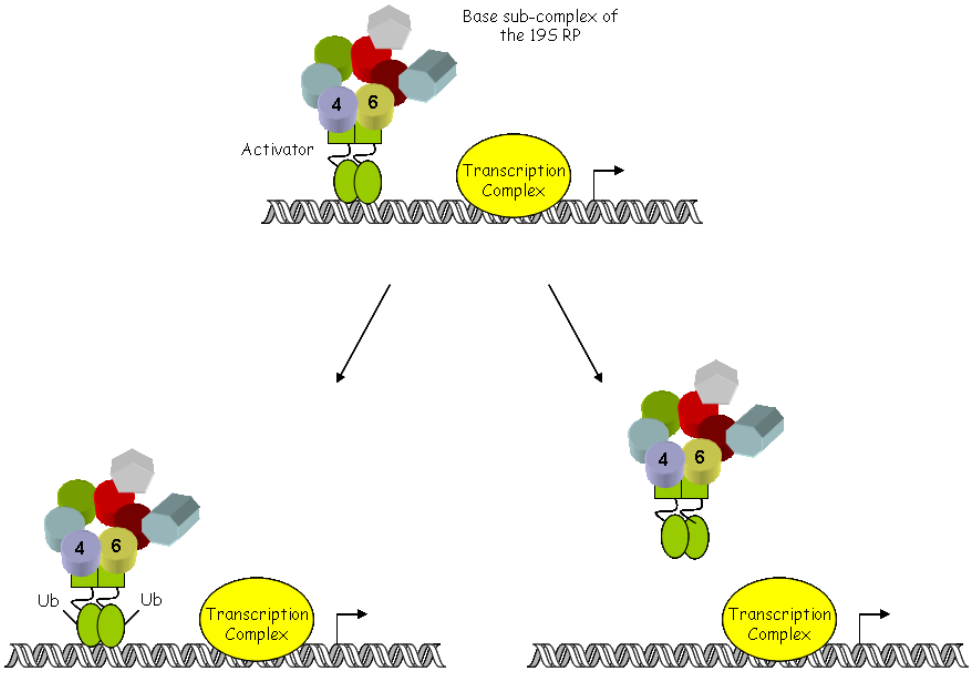

Figure 8: A model for destabilization of the activator Gal4 by the proteasomal ATPases. The Gal4-based proteins occupy the UAS_G sites and the activation domain then recruits the proteasomal ATPases under inducing conditions. Once recruited, the proteasomal ATPases destabilize activators that are not mono-ubiquitylated (right). Those activators that are mono-ubiquitylated some how prevent dissociation of the activator from the promoter (left).

CHAPTER II

ANALYSIS OF THE GAL4 MUTANT GAP71

INTRODUCTION

RNA is generated from a DNA template through the highly regulated process of transcription, which typically begins when an activator recruits the transcriptional machinery to a promoter to nucleate the pre-initiation complex [11]. Transcriptional activators bind to a promoter through an interaction between the DNA-binding domain (DBD) of an activator and a defined consensus site within the DNA-template. Inhibitors can prevent the activator from binding the recognition sequence in the DNA-template. In the case of the *GAL1-10* promoter, the protein Mig1 behaves binds the promoter and physically prevents Gal4 from occupying that region of the DNA [19, 20]. In addition, mutations in an activator itself can directly alter the binding properties of the activator-DNA interaction or permit the activator to be more easily displaced by indirect mechanisms such as that seen by the proteasomal ATPases [90]. Once an activator has bound the DNA it must then be able to activate transcription [11]. This activity of the activator is controlled by numerous co-activators and inhibitors. For instance, the

Mediator complex is recruited to the promoter through an interaction with the Gal4 AD to play a positive role in transcription (for reviews see [32, 33]). However, that co-activator and others will not be recruited when Gal80 interacts with and muzzles the activation domain of the bound Gal4 [17, 21]. While many times the transcription of a gene is defined by the activator present, it is a fine-tuned symphony of events that ultimately controls when DNA is to be transcribed.

One of the more interesting aspects of activator regulation is that mono-ubiquitylation can influence the efficiency of some activators. While this mechanism is not fully understood, it is established that mono-ubiquitylation events and their subsequent effects can occur independent of proteolysis [77-79]. An explanation for at least one of the roles ubiquitin plays in transcription is found in the base sub-complex of the 19S RP that was recently discovered to function independent of the lid sub-complex and the 20 CP to dissociate an activator-DNA interaction [88, 90]. Initial experiments characterized this process as being dependent on a functional activation domain and ATP. Studies have now shown that the dissociation of the activator-DNA interaction is prevented by a 'protective' mono-ubiquitylation event on an activator [86, 94, 95].

The *Gal4* mutant *gap71* was fundamental to the discovery that the activator-DNA interaction could be destabilized by the proteasomal ATPases. This mutant has three amino acid substitutions (S22D, K23Q and K25F) in the DNA-binding domain of Gal4, is hyper-sensitive to dissociation from a *GAL-*

based promoter and is not mono-ubiquitylated [96]. While it is still unclear exactly how the proteasomal ATPases regulate an activator, these studies did establish that the mono-ubiquitylation event on the Gal4 DNA-binding domain protects the activator-DNA complex from being disrupted and that the residues 22, 23 and 25 are influential in the stability of the activator-DNA interaction, warranting additional studies of Gap71.

RESULTS

In Vivo Analysis of Mutants of Gal4

The experiments that discovered the proteasomal ATPases destabilized activator-DNA interactions were done *in vitro* with artificial activators and *in vivo* with full-length protein. To bridge the data sets, I began my studies of Gap71 by analyzing the artificial activators *in vivo*. The fusion proteins contained the first 147 amino acids of the Gal4 or Gap71 DNA-binding domain and the acidic activation domain VP16. The sequences encoding these two fusion proteins were excised from the bacterial plasmid pGEX-CS and inserted behind the native *GAL4* promoter in the single copy plasmid pSB32, as well as the multi-copy plasmid YEp351.

The ability of the fusion proteins to activate transcription *in vivo* was assessed by a *MEL1*-based reporter gene assay (Figure 9, top panel). The negative control was a *Δgal4* strain. The first two columns represent the data collected for the single-copy and multi-copy version of Gap71-VP16; neither construct was able to activate the reporter gene. The last two columns represent the data collected for the Gal4-VP16 fusion protein. While the single-copy plasmid did not express enough protein to activate transcription in our assay, expression by the multi-copy plasmid produced a robust signal.

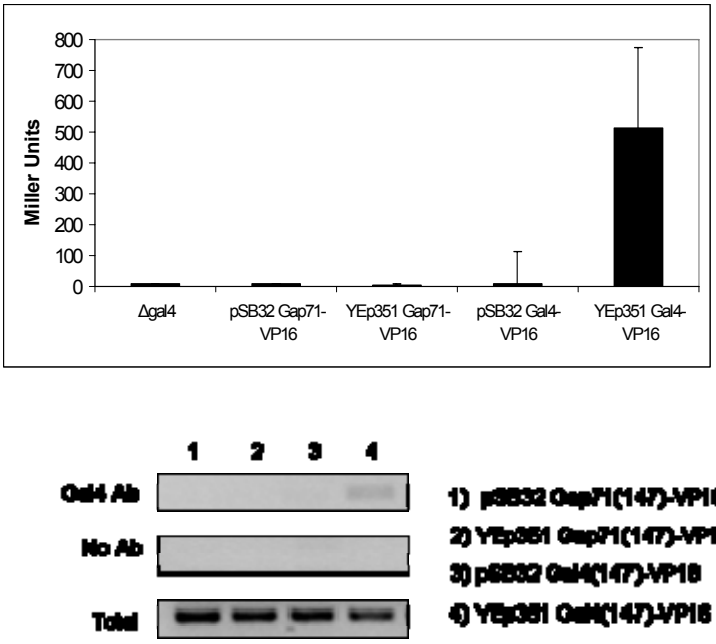

Figure 9: Analysis of the Gal4(147)-VP16 and Gap71(147)-VP16 phenotypes *in vivo*. The data were graphed to show the activity profile of the various fusion proteins. pSB32 was a single copy plasmid and YEp351 was a multi-copy plasmid that should lead to the increased expression of the constructs ~25 fold. The results from the activity assay were graphed according to their value in Miller units (top). The results from ChIP assays analyzing the occupancy level of the fusion proteins were shown for the *Gal1-10* promoter (bottom).

The large difference in the activities of Gal4-VP16 and Gap71-VP16, when expressed from a multi-copy vector, could be explained in two different ways. The multi-copy versions of Gal4-VP16 and Gap71-VP16 could have resulted in robust occupancy, leaving the data for Gap71-VP16 to represent an activation deficiency. The other possibility was that I was unable to force occupancy of the UAS_G sites by Gap71-VP16. To determine the ability of the artificial activators to occupy a *GAL* promoter, I performed a chromatin-immunoprecipitation assay (ChIP) (Figure 9, bottom panel). The cells for the

ChIP assay were harvested simultaneously with the samples for the activity assay. An immunoprecipitation from each chromatin solution was done with a Gal4 N-terminal antibody or with no antibody added. Totals represented $1/20^{th}$ of the starting material used for the pull-downs. Surprisingly, not even the wild-type constructs were able to robustly occupy the *GAL1-10* promoter. There is a faint band indicating that the Gal4(147)-VP16 fusion protein occupied the promoter. The original question was whether or not the fusion proteins behaved similarly *in vitro* and *in vivo*. The *in vitro* experiments showed the Gap71(147)-VP16 fusion protein was released from a UAS_G template while the wild-type fusion protein remained bound to the template [90]. Even though the ChIP signals are weak, my data suggest the fusion proteins behave the same *in vivo* and *in vitro*.

To determine the residue that elicits the Gap71 phenotype, site-directed PCR mutagenesis was used to deconstruct *gap71* into three individual amino acid substitutions: Ser22 to aspartic acid (S22D), Lys23 to glutamine (K23Q) and Lys25 to phenylalanine (K25F). The full-length Gal4-based proteins performed better in our assays than the artificial activators; therefore the parent vector for these new derivatives was the plasmid pSB32-$S10_3$-Gal4, which expresses epitope-tagged full-length protein from the native *GAL4* promoter. Additional sub-cloning put the new mutants of *Gal4* into the multi-copy plasmid YEp351 again under the control of the native promoter. These six new constructs were made by Dr. Kip Nalley in the Kodadek laboratory. To analyze the phenotype of

the new Gal4 derivatives, I used a *MEL1*-based reporter gene assay (also known as an α-galactosidase assay) to compare the efficiency of the new activators against Gap71 (negative control) and Gal4 (positive control). In an attempt to mimic an endogenous protein expression profile, transcription was initially driven by the native *GAL4* promoter from the single-copy plasmid pSB32 (Figure 10, top panel). The negative controls were a *Δgal4* strain and a strain transformed with Gap71. Both showed negligible activity. All of the data were normalized to the activity of wild-type Gal4. The S22D mutant was approximately 20% as potent as Gal4, while the mutant K23Q resembled the negative controls. K25F showed an intermediate activity level, about 50% of Gal4.

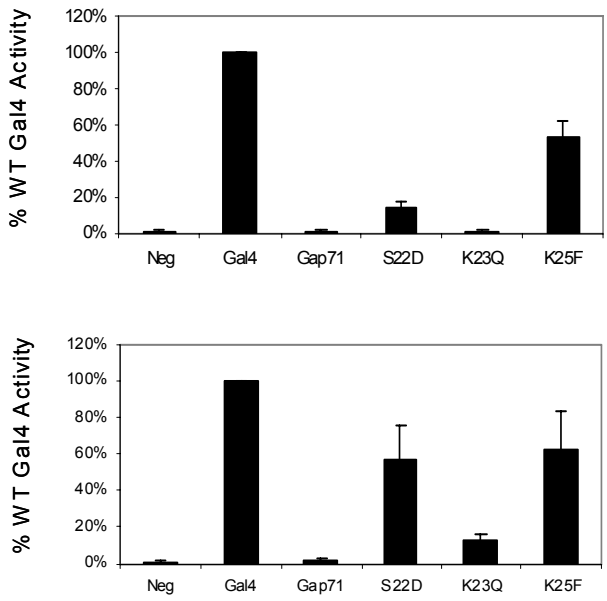

Figure 10: Alpha-galactosidase activity assay comparing the activity of various mutants of the activator Gal4. Values were normalized to Gal4 being 100%. Proteins were expressed by the native *GAL4* promoter by the single-copy plasmid pSB32 (top panel) or by the multi-copy plasmid YEp351 (bottom panel) [96]. Copy-Right Permission Granted.

Decreased activity could mean the activator was not fully occupying the promoter or that it was bound, but the downstream functions of the activator had been compromised. If a deficient activator was fully occupying the promoter, then over-expression of that activator should not increase its efficiency since the recognition sites are presumed to be saturated in this situation. However, if promoter occupancy was limiting the function of the activator then over-expression of the activator should rescue the efficiency of the activator. In theory, this scenario forces the binding kinetics to favor occupancy by the activator. In an attempt to determine the basis of the deficiencies in the activity assay, I over-expressed the same constructs with the multi-copy plasmid YEp351 and repeated the *MEL1*-based reporter gene assay (Figure 10, bottom panel). Again the negative controls were a *Δgal4* strain and a strain over-expressing Gap71. The positive control strain was transformed with YEp351 $S10_3$-Gal4, whose activity was normalized to 100%. Over-expression of the mutant S22D recovered the activity of the protein to about 50% that of wild-type Gal4. The mutant K23Q maintained minimal activity while the K25F mutant again exhibited an activity ~ 60% that of the wild-type protein.

Studies have shown that Gap71 can efficiently bind a DNA template *in vitro* but not during inducing conditions *in vivo* [90, 93]. The question posed was whether or not the individual mutations inherent to Gap71 also had an occupancy deficit. A chromatin immunoprecipitation (ChIP) assay was done to ascertain the

occupancy levels of the various activators on the *GAL1-10* promoter (Figure 11). These cells were grown on raffinose-containing media to poise the activators on the promoter. The cultures were then induced with galactose-containing media to remove the Gal80-based repression, thereby exposing the activation domain of the activator to recruit the proteasomal ATPases and the transcriptional machinery. Again, the constructs were expressed at low levels by the native *GAL4* promoter by the single-copy plasmid pSB32. The data shown represent the occupancy level of the activator on the *GAL1-10* promoter (Figure 11, top panel). The experiment used a *Δgal4* strain and a strain expressing Gap71 as the negative controls. The positive control was wild-type Gal4, which was normalized to 100%. The mutant activators Gap71, S22D and K23Q did not occupy the *GAL1-10* promoter. The K25F mutant had variability but appeared to occupy the *GAL 1-10* promoter well. To see if over-expression of the activators rescued their occupancy deficiency, I repeated the experiment using constructs expressing the protein ~25 fold higher with the multi-copy plasmid YEp351 (Figure 11, bottom panel). Again, a *Δgal4* strain and a strain expressing Gap71 were the negative controls and the data were normalized to Gal4 representing 100% occupancy of the *GAL1-10* promoter. The mutants Gap71 and K23Q showed limited occupancy of the promoter despite the increase in their protein levels. The occupancy level of the S22D mutant was partially rescued to about 50% of wild-type Gal4 levels while the K25F mutant showed occupancy levels surpassing those of wild-type Gal4.

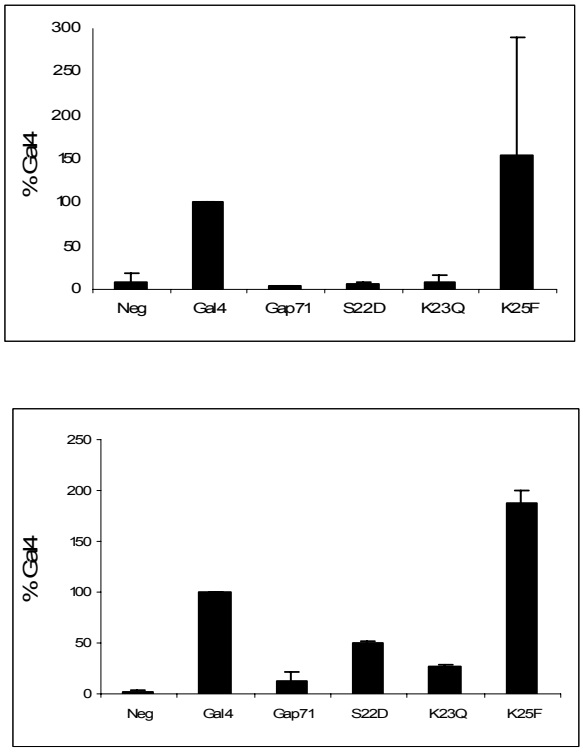

Figure 11: ChIP analysis of the panel of Gal4 derivatives on the *GAL1-10* promoter. Cultures were induced for 45 minutes with 2% galactose. Pull downs are done with an N-terminal, poly-clonal Gal4 antibody. Values were normalized to Gal4 being 100%. Proteins were expressed under the native *GAL4* promoter in the single copy plasmid pSB32 (top panel) or the multi-copy plasmid YEp351 (bottom panel).

The goal of these experiments was to determine the individual amino acids responsible for the Gap71 phenotype. The Lys25 mutant showed a more robust activity profile than Gap71 and did not appear to have an occupancy deficiency of the *GAL1-10* promoter. However, the amino acids substitutions at Ser22 and Lys23 were found to be the main contributors to the Gap71 phenotype.

The Ser22 and Lys23 residues were intriguing because these amino acids are prime targets for post-translational modifications. The most comon modification at a serine residue is phosphorylation. Since the S22D mutant retains some activity and is known to act as a phosphoserine mimic in many cases, it is tempting to speculate that S22 is a site for an essential phosphorylation event (Figures 10 and 11). Therefore Ser22 was mutated into an alanine (S22A) to remove the side chain, thereby eliminating the putative phosphoserine mimic. If Ser22 is a target of phosphorylation, then the S22A mutant phenotype should be more severe than the S22D mutant phenotype. The structures of the amino acid side chains of alanine, serine and aspartic acid along with a phosphorylated serine residue are diagrammed in Figure 12.

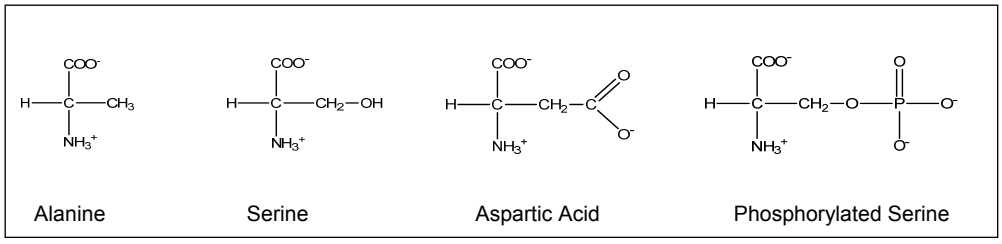

Figure 12: Structures of several amino acid residues and a phosphorylated serine residue. This diagram shows the various amino acids analyzed at residue twenty-two and how they compare to a phosphorylated serine residue.

In addition to Ser22, the analysis of the individual mutations in Gap71 warranted further investigation of Lys23. Several modifications can occur at this residue such as acetylation, ubiquitylation and methylation. Unlike Gap71, Gal4 is mono-ubiquitylated and is able to maintain the activator-DNA interaction long enough to initiate transcription. As a result, it seemed likely that Lys23 was the site of mono-ubiquitylation in the Gal4 DNA-binding domain.

As well as being a site for modifications, the side chain of a lysine residue has a positive charge that is absent in glutamine, the residue substituted for K23 in Gap71. To restore this charge, residue twenty three was mutated to an arginine (K23R). A comparison of the structures of the side chains of the amino acids mentioned can be seen in Figure 13. If the charge of the side chain is important at position twenty-three, then the phenotype of the K23R mutant more closely resemble that of Gal4 than K23Q. If residue twenty-three is the site of a post-translational modification, then the mutants K23Q and K23R should have a similar phenotype.

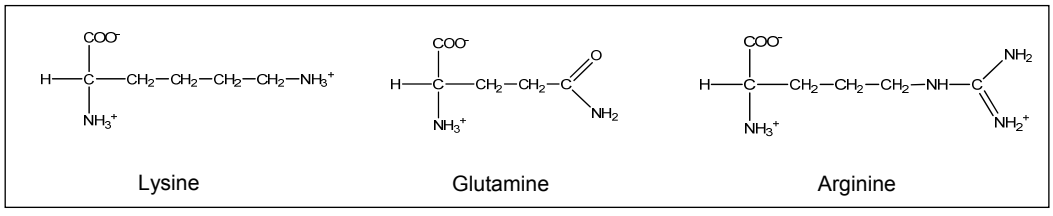

Figure 13: Structures of amino acid side chains. This diagram shows various amino acid side chains and their corresponding charges that were analyzed at Lys23.

The new Gal4 derivatives S22A and K23R were generated using site-directed PCR mutagenesis. To maintain consistency across the panel of mutants, the template was full-length $S10_3$-Gal4 driven by the native *GAL4* promoter by the single copy vector pSB32. The new constructs were confirmed by sequencing and were subsequently sub-cloned into the multi-copy vector YEp351. I began my analysis with the *MEL1*-based activity assay used previously. The proteins were expressed at low levels and a sample was taken after induction with galactose for 45 minutes (Figure 14, top panel). The negative controls were a *Δgal4* strain and strain expressing Gap71. The data were normalized to the activity of Gal4. The Gap71, S22A and K23Q mutants were inactive in this assay. The S22D mutant had marginal activity while the mutant K23R had partial activity that significantly surpassed the activity of the K23Q mutant. The K25F mutant had 50% of the activity seen in the positive control. As before, the panel of Gal4 activators was also analyzed when over-expressed by the multi-copy plasmid YEp351 (Figure 14, bottom panel). The negative controls were a *Δgal4* strain and a strain expressing Gap71, while the activity level of wild-type Gal4 was normalized to 100%. The phenotype of the over-expressed mutant S22A was not rescued and resembled both the K23Q and Gap71 data. The phenotype of the over-expressed mutant S22D was partially rescued while K23R exhibited an activity level similar to that of wild-type Gal4. The K25F mutant phenotype remained constant.

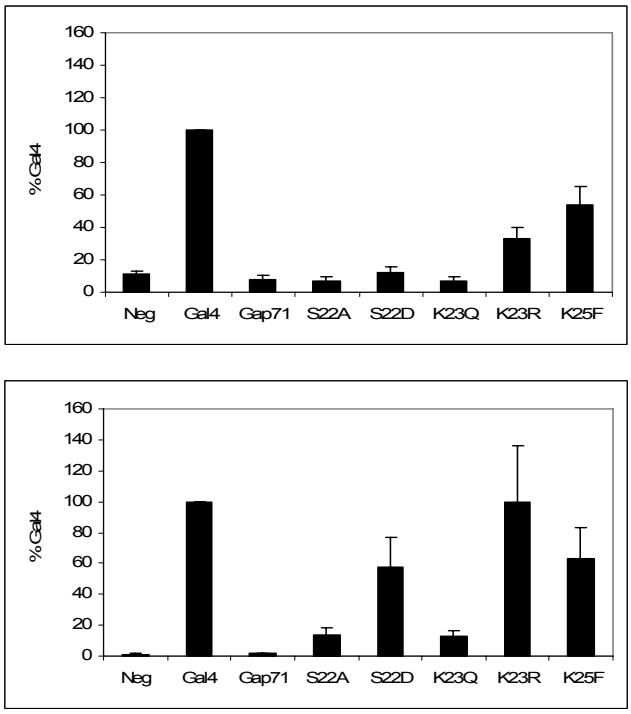

Figure 14: Activity assays testing the new S22A and K23R substitutions. The full-length proteins were expressed under the native *GAL4* promoter by the single-copy plasmid pSB32 (top panel) or the multi-copy plasmid YEp351 (bottom panel).

To determine if the phenotype associated with the new constructs resulted from an occupancy deficiency, chromatin-immunoprecipitation (ChIP) assays were done to evaluate the *GAL1-10* promoter for the presence of the Gal4 derivatives. I used the same full-length panel of pSB32 based constructs used in Figure 14. The negative controls were a *Δgal4* strain and a strain expressing Gap71. The value for Gal4 occupancy at the *GAL1-10* promoter was normalized to 100%. The Gap71, S22A, S22D and K23Q mutants did not occupy the *GAL1-10* promoter when exposed to galactose-containing media. The K23R and K25F mutants occupied the UAS_G sites significantly better than the wild-type Gal4

control (Figure 15, top panel). The same panel of Gal4 derivatives was over-expressed by the multi-copy plasmid YEp351 to see if over-expression would improve the S22A phenotype (Figure 15). The negative control was a *Δgal4* strain and wild-type Gal4 data was normalized to 100%. The Gap71, S22A and K23Q mutations showed limited occupancy of the *GAL1-10* promoter. The S22D mutation shifted to an occupancy level 50% that of wild-type Gal4, which was significantly higher than the value shown for the S22A mutant. The profile of the K23R and K25F mutants again surpassed the level of wild-type Gal4 occupancy.

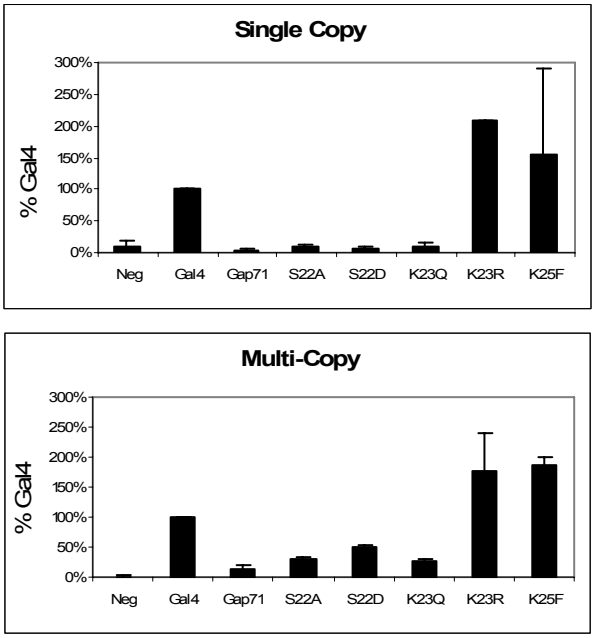

Figure 15: ChIP analysis of the Gal4 derivatives on the *GAL1-10* promoter. Pull downs were done with an antibody raised against the N-terminal of Gal4. Wild-type Gal4 was normalized to 100%. The full-length Gal4 derivatives were expressed under the native *GAL4* promoter by the single-copy plasmid pSB32 (top panel) or the multi-copy plasmid YEp351 (bottom panel).

The original hypothesis was that the S22D substitution was behaving as a phosphoserine mimic and that the S22A substitution would abolish the function of the activator. The data collected from the activity and occupancy assays showed the substitutions at residue twenty two, S22A and S22D, behaved the same when expressed at low levels, yet the two mutants have distinct phenotypes when over-expressed. While the S22A mutant remained unable to occupy the prmoter, both the activity level and the occupancy level of the S22D mutant was 50% that of Gal4, supporting the idea that S22D is a phosphoserine mimic.

The other assumption going into this study was that Lys23 is mono-ubiquitylated. The data showed that the over-expressed K23R mutant rescued the activity and occupancy deficiencies of the activator to a level similar to wild-type Gal4 while the over-expressed K23Q mutation remained relatively inactive. The data suggest the charge at residue twenty three is important for Gal4 function, not a mono-ubiquitylation event. Overall, these preliminary results suggested that Ser22 is a site of phosphorylation. In addition, the data implied that Lys23 is most likely part of a recognition pattern that is vital for Gal4 function. While the data found the charge at Lys23 is important, the results do not rule out the possibility that Lys23 is modified.

In Vivo Analysis of Gal4-Based Fusion Proteins

While it was determined that Ser22 and Lys23 mutations have occupancy deficiencies, possible effects of these mutations on events in the transcription cycle downstream of promoter occupancy were unclear. In an attempt to probe this issue, I decided to analyze the individual amino acid substitutions in the artificial activator Gal4-Gal11. This construct was shown to robustly activate *GAL* promoters despite lacking a classical activation domain [97]. The fusion protein has the first 147 amino acids of the Gal4 DNA-binding domain fused to the C-terminal portion of Gal11 (amino acids 799-1081). Since the Gal4-Gal11 fusion lacks a classical activation domain, but does couple to the Mediator complex via the fused Gal11 domain, we were hopeful that Gal4-Gal11 would not interact directly with the proteasomal ATPases complex. If so, this would allow proteins, like Gap71, with mutations that abrogate mono-ubiquitylation to bind promoters *in vivo*, thereby allowing analysis of potential downstream effects of the lack of activator mono-ubiquitylation. I used the multi-copy plasmid YEp351 to over-express the mutant Gal4-Gal11-based proteins in an attempt to force complete occupancy of the UAS_G sites. The parental plasmid had the Gal4-Gal11 fusion protein under control of the endogenous *GAL4* promoter in the multi-copy plasmid YEp351. Site-directed PCR mutagenesis was used to create the individual amino acid substitutions while the YEp351 Gap71-Gal11 construct was sub-cloned from another vector.

The initial experiment evaluated the efficiency of the activators in the *MEL1*-based reporter gene assay. A *Δgal4* strain and two strains expressing Gal4-Gal11 or Gap71-Gal11 were used as controls (Figure 16, Top). The data were normalized to the activity of Gal4-Gal11. The K25F mutant behaved like the wild-type Gal4 control. The S22D and K23R mutants showed partial activity while the S22A and K23Q mutants mimicked the Gap71 phenotype. The ability of the activator to occupy a *GAL* promoter was assessed by a chromatin-immunoprecipitation assay (Figure 16, Bottom). The negative control was *Δgal4* strain. The data were normalized to the occupancy level of the wild-type construct. The K25F mutant behaved like wild-type Gal4. The individual mutants S22A and K23Q mimicked the Gap71 phenotype, while both the S22D and K23R mutants displayed an intermediate occupancy level.

In essence, the ChIP assay data mimics the profile seen when the full-length constructs were used in Figure 11, even though these constructs lack a classical acidic activation domain. This demonstrates that they still were destabilized from the promoter, contrary to our expectations. Therefore, this panel of Gal4-Gal11 derivatives does not permit the analysis of potential effects of mutations at the residues Ser22, Lys23 or Lys25 that are downstream promoter occupancy.

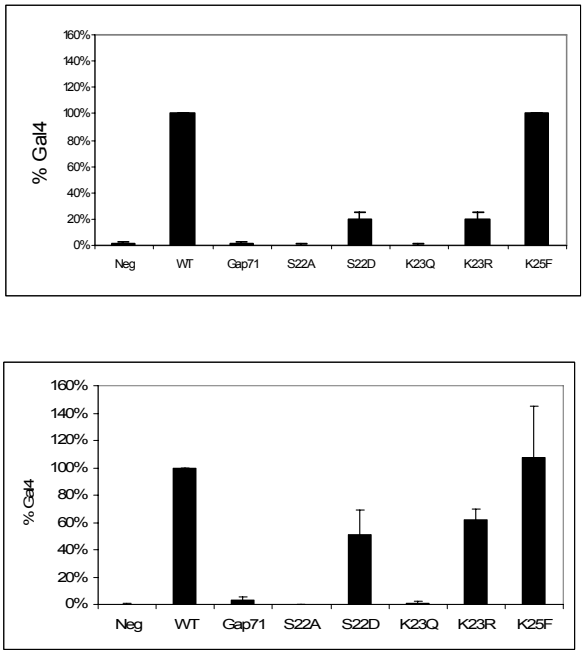

Figure 16: Profile of YEp351 Gal4-Gal11 panel of constructs. The graph shows the activity profile of the Gal4-Gal11 derivatives (top panel). The samples were standardized to the wild-type construct being 100%. The graph shows the occupancy profile of the same Gal4-Gal11 cultures on a *GAL* promoter (bottom panel). The samples were also standardized to the wild-type construct being 100%.

In an attempt to circumvent this problem through sheer mass action, I decided to redo the experiment using a different artificial activator whose expression was driven by the powerful *ADH1* promoter on a multi-copy plasmid. The fusion protein used was a gift from the laboratory of Dr. C. Picard and was already established in the Kodadek laboratory. It has the first 93 amino acids of the Gal4 DNA-binding domain fused to the acidic activation domain VP16 [98]. Previous data established that the K23Q mutation maintains the same phenotype

as Gap71, therefore the original construct from Dr. Picard was modified using site-directed PCR mutagenesis to have the K23Q substitution.

I repeated the line of experiments using the new constructs in an attempt to achieve higher expression levels and ultimately full occupancy of a *GAL* promoter by a Gal4 mutant. The negative control was a *Δgal4* strain. The activity assay revealed that the wild-type Gal4(93)-VP16, but not the mutant Gal4(93)K23Q-VP16, has a robust signal (Figure 17, left panel). As before, I had collected ChIP samples alongside the samples for the activity analysis. The wild-type Gal4(93)-VP16 construct had a robust signal indicating a high level of occupancy of the UAS_G sites (Figure 17, right panel). The higher expression of the mutant Gal4(K23Q)-VP16 was unable to improve the occupancy deficiency associated with this mutation.

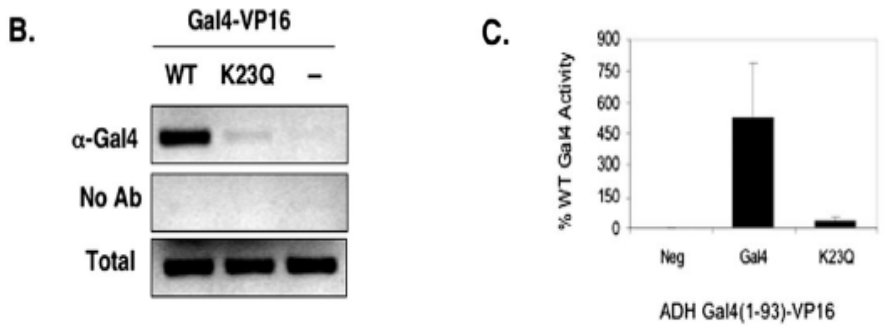

Figure 17: Analysis of artificial activators when highly expressed by the ADH1 promoter in the multi-copy plasmid YEp351[96]. Copy-Right Permission Granted.

DISCUSSION

A collaboration between the Kodadek and Johnston laboratories previously identified a novel function of the proteasomal ATPases that occurred independent of proteolysis to destabilize activator-DNA complexes. A mutant of Gal4 termed Gap71 was key to this discovery as it was found to be hyper-sensitive to dissociation from a DNA template. The initial study used the artificial activators Gal4(147)-VP16 and Gap71(147)-VP16 for the *in vitro* experiments but assessed the full-length Gal4 and Gap71 proteins *in vivo*. To bridge the data sets, I sub-cloned the artificial activators into yeast expression vectors for a subsequent *in vivo* analysis. While the function of the artificial activators was not as robust at the full-length counterparts *in vivo*, the results for the data sets had the same conclusions, the artificial activator Gal4-VP16 was able to occupy a *GAL* promoter but the activator Gap71-VP16 was not.

To gain a better understanding of how the proteasomal ATPases destabilize activators from a DNA template, the Gap71 protein needed to be further characterized. Therefore, the three amino acid substitutions within the DNA-binding domain of Gap71 were individually cloned into Gal4, generating the mutants S22D, K23Q and K25F. These clones were constructed to determine the role each residue plays in the function of the activator Gal4 and ultimately their contribution to the stability of the Gal4-DNA interaction.

The initial study of the individual mutations implied that Ser22 and Lys23 were the main contributors to the Gap71 phenotype. The side chain of an aspartic acid can often behave as a phosphoserine mimic. Therefore, a serine to alanine (S22A) substitution was made to remove the aspartic acid side chain. The Gal4 derivative with the S22A mutation did not bind well to the *GAL1-10* promoter when exposed to galactose-containing media, thereby rendering the protein inactive. The data for the activator with the S22D mutation showed the protein bound the *GAL1-10* promoter at a level half that of wild-type Gal4, which correlated with the data from the *MEL1*-based reporter gene assay. Comparison of the S22A and S22D phenotypes augmented the hypothesis that Ser22 was phosphorylated because the S22D mutant was significantly more active than the S22A mutant. An *in vitro* phosphorylation assay subsequently revealed that Ser22 in the Gal4 DNA-binding domain was indeed phosphorylated [96]. Previous work by other groups had shown that the C-terminus of Gal4 was phosphorylated at three other positions and that these modifications were nonessential for transcription [99-102]. What was unique about the modification at Ser22 was that it represented the first phosphorylation event to be essential for the function of the activator Gal4.

In addition to Ser22, Lys23 and Lys25 were also mutated in Gap71. The analysis of these two mutations suggested Lys23 but not Lys25 was a substantial contributor to the Gap71 phenotype. The K25F mutant was able to bind the

GAL1-10 promoter well and had an activity level half that of Gal4, which is dissimilar to the inability of Gap71 to occupy a *GAL* promoter. The mutation of Lys23 to glutamine (K23Q) was found to have an occupancy profile similar to that of Gap71. Do to the severity of the K23Q mutant phenotype, the residue at position twenty three was mutated to an arginine (K23R), thus restoring a positive charge to the side chain. The K23R phenotype was a robust improvement over the K23Q phenotype, suggesting the charge of lysine twenty three is important for proper Gal4 function. However, these data do not rule out the possibility that the charge at Lys23 is important for Gal4 function as well being a site of mono-ubiquitylation.

Previous studies from the collaboration had shown that Gap71 was hypersensitive to destabilization from a DNA template by the proteasomal ATPases and that one difference between the two proteins was that the Gal4 DNA-binding domain (DBD) was mono-ubiquitylated whereas that of Gap71 was not. In addition, the study discovered the mono-ubiquitylation event on the Gal4 DBD was dependent on the presence of an activation domain, suggesting the C-terminal of Gal4 houses the recognition sequence for the corresponding E3 ligase machinery [90]. Considering the correlation between mono-ubiquitin and the increased stability of the activator-DNA complex, an *in vitro* ubiquitylation assay was performed to probe the individual mutants of Gap71 (S22D, K23Q and K25F) [96]. Gap71 and the K23Q mutation were not mono-ubiquitylated while

the Gal4 K25F was mono-ubiquitylated at a level similar to that of wild-type Gal4. Surprisingly, the other mutation inherent to Gap71, S22D, had only a faint mono-ubiquitylation signal [96]. Many studies have shown that a phosphorylation event often triggers an ubiquitylation event elsewhere on the same protein, thus explaining how a mutation of a serine residue could affect the mono-ubiquitylation status of a protein [103, 104]. Based on the *in vitro* data our laboratory has produced and the *in vivo* results shown here, I speculate that the phosphorylation event at Ser22 precedes a mono-ubiquitylation event elsewhere on the Gal4 DBD. Instead of being the site of mono-ubiquitylation, it is more probable that the charge at Lys23 is part of a recognition pattern for the kinase that phosphorylates Ser22.

To further analyze the importance of Ser22 and Lys23 I established the panel of Gal4 derivatives in the artificial activator Gal4-Gal11. This construct was chosen because it was a potent AD-less activator, thereby preventing the interaction between the activator and the proteasomal ATPases [97]. The hypothesis was that these activators would be less sensitive to dissociation from a *GAL* promoter, thus permitting the study of additional affects caused by the three individual amino acid substitutions. Surprisingly, the Gap71-Gal11 mutant and various amino acid substitutions at Ser22 and Lys23 were still destabilized from a *GAL* promoter. The results were unexpected but most likely can be explained through established associations with the transcriptional machinery. The

Mediator complex was identified previously as a transcriptional co-activator with three sub-complexes; the head, middle and tail modules (for reviews see [32, 33]). The portion of Gal11 used to create the pseudo activation domain is the same part found to assemble within the tail sub-complex of the co-activator Mediator, thus permitting an AD-less activator the ability to initiate transcription [105]. The artificial activators Gap71(147)-VP16 and Gal4K23Q(93)-VP16 could not bind a *GAL* promoter either. The acidic activation domain VP16 has been shown to associate with the TFIIB and TFIID subunits of the RNAP II pre-initiation complex [106-108]. Both the Gal4-Gal11 and Gal4-VP16 constructs were able to activate *GAL*-based transcription, therefore the hypothesis is that the transcriptional machinery, which includes Mediator, facilitates destabilization of the activator-DNA interaction by bringing the proteasomal ATPases into proximity with an un-ubiquitylated, mutant activator.

The idea that mono-ubiquitin 'protects' the activator-DNA interaction was not only intriguing but is currently the only known way to stabilize the activator-DNA interaction [90]. Experiments developed to determine how the proteasomal ATPases destabilize an activator-DNA complex discovered free mono-ubiquitin can prevent the dissociation of the activator-DNA complex *in vitro* [94]. Furthermore, cross-linking studies revealed that mono-ubiquitin interacts with the 19S RP base sub-complex proteins Rpt1 and Rpn1, which are not the same base sub-complex proteins (Rpt4 and Rpt6) that interact with the Gal4 activation

domain [89, 94]. The resulting model can be found in Figure 18. The proteasomal ATPases are recruited to the *GAL1-10* promoter in a Gal4 activation domain-dependent manner [88]. The the interaction between the Gal4 activation domain and the 19S RP base sub-complex is disrupted by the affinity of ubiquitin for the 19S RP proteins Rpt1 and Rpn1, leading to a rearrangement of the proteins at the *GAL1-10* promoter that prevents the dissociation of the activator from the DNA template (Figure 18). In addition, this rearrangement of transcriptional machinery could be priming the proteasomal ATPases for the positive role the sub-complex plays in transcriptional elongation.

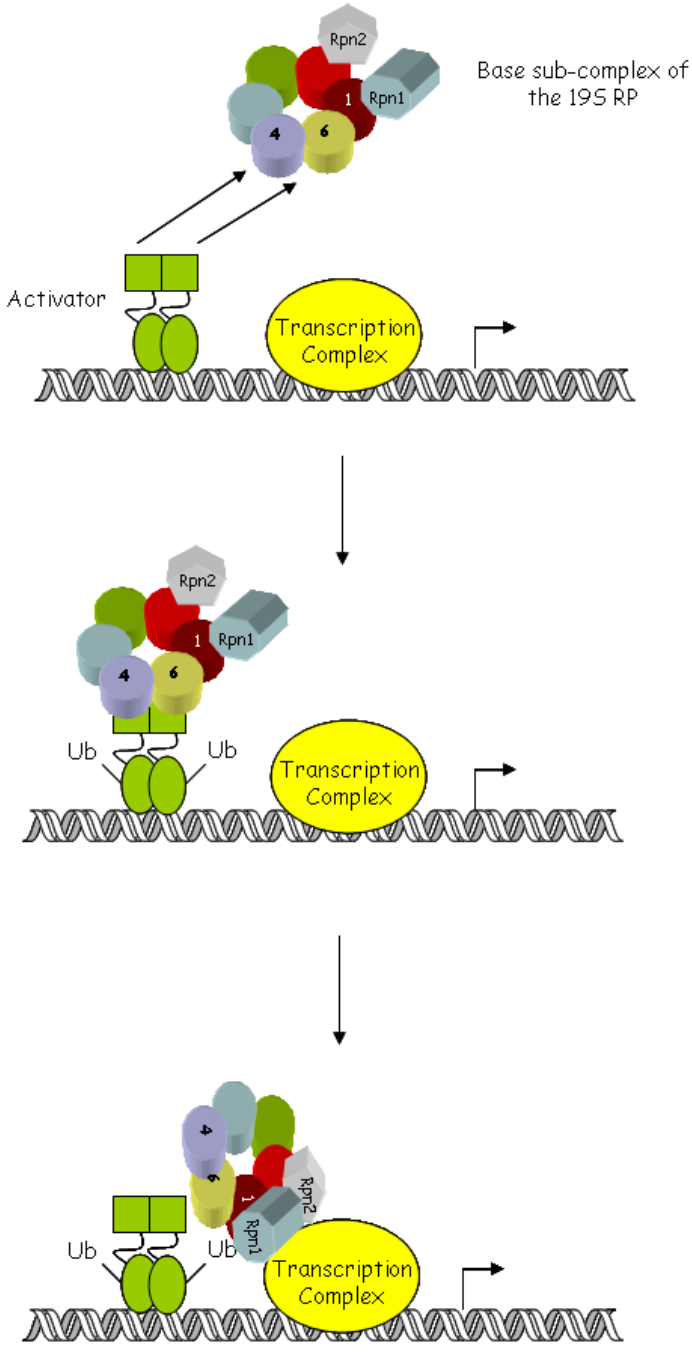

Figure 18: A model depicting how the proteasomal ATPases affect the stability of activator-DNA complexes. The Gal4 AD recruits the proteasomal ATPases to the promoter. The ATPases preferentially bind mono-ubiquitin over the activation domain and permit Gal4-based transcription. The other possibility is that the mono-ubiquitin is not present and the proteasomal ATPases remove the activator from the promoter.

CHAPTER III

ANALYSIS OF THE PROTEASOME-BASED PROCESS THAT DISSOCIATES AN ACTIVATOR-DNA INTERACTION

INTRODUCTION

As reviewed in the preceding chapters, mono-ubiquitylation of transactivators is likely to be of general importance in eukaryotic cells [78, 79, 90]. Since Gal4 has proven to be a paradigm for the mechanism of action of many mammalian activators, it seems likely that further detailed studies of Gal4 mono-ubiquitylation would provide information relevant to other transactivators of greater clinical interest.

Therefore, I designed a set of experiments to determine the lysine within the Gal4 DNA-binding domain that is modified by the protective mono-ubiquitin species, thus permitting a focused analysis of this post-translational modification. In addition to understanding where post-translational events occur, the identity of the enzymes that facilitate those modifications might also provide useful insights. As a result, I also pursued experiments designed to identify the protein kinase(s) and ubiquitin E3 ligase(s) that target Gal4.

RESULTS

Lysine Scan of the Gal4 DNA-Binding Domain

Our laboratory has shown that when the DNA-binding domain (DBD) of the transcription factor Gal4 was not mono-ubiquitylated, the activator-DNA interaction became hyper-sensitive to destabilization by the proteasomal ATPases [90]. Although the activation domain is essential for this ubiquitylation event, the lysine that is modified must occur within the first 147 amino acids of Gal4 since the acidic activation domain VP16 does not contain lysines [90]. I have also done experiments with constructs containing only the first 93 amino acids of the Gal4 DBD fused to the same acidic activation domain VP16. When Lys23 of this fusion protein was mutated to a glutamine I saw a loss of function in the *MEL1* based activity assay. This phenotype was augmented by an occupancy deficiency at the *GAL1-10* promoter. Both results are similar to the phenotype of full-length Gap71 *in vivo*, arguing that the site of Gal4 mono-ubiquitylation is located within these residues [96]. Based on the content of the artificial activators and their corresponding data, I focused my analysis on the thirteen lysines found in the first 93 amino acids of the Gal4 DBD (Figure 19).

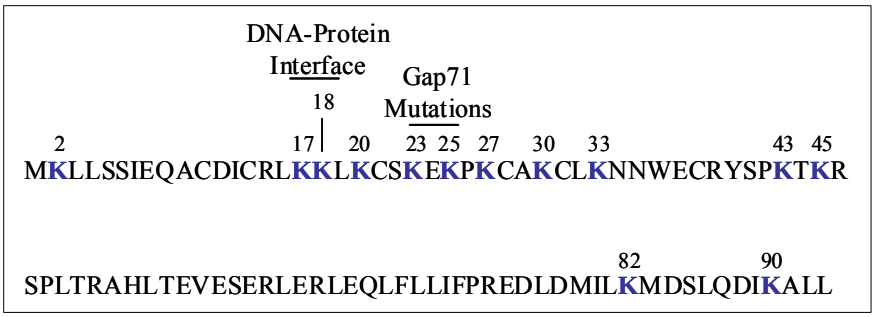

Figure 19: This is a representation of the first 93 amino acids of the Gal4 protein. This portion of the protein contains the DNA-binding domain and demonstrates the location of the lysine residues within that domain. The lysines are highlighted in blue and the number is that of the residue within the Gal4 DBD.

I was able to eliminate a priori some of the thirteen lysine residues as potential targets of the ligase machinery. Two of the lysines are known to be part of the Gal4-DNA interface, Lys17 and Lys18 [109]. In addition, Lys23 and Lys25 have been evaluated previously [96]. The model deduced from those experiments is that Gap71 is dissociated from the promoter under inducing conditions because the mutation at residue twenty-two prevents the phosphorylation event that precedes the 'protective' mono-ubiquitylation event elsewhere in the DNA-binding domain. This leaves eleven lysines in the Gal4 DBD as possible sites for a mono-ubiquitylation event.

Lysine scans can be a powerful tool in the identification of sites of modification. This method substitutes a lysine with an arginine, thereby eliminating the site of conjugation for a modification while maintaining the

positive charge of the original lysine residue. The modified proteins are then analyzed to see if the substitution yields a phenotype of interest.

To do a lysine scan on the Gal4 DNA-binding domain, I chose to use a construct that had already been established in the laboratory, *ADH1* driven Myc-Gal4(93)-ER-VP16-Flag [110]. This fusion protein contains the first 93 amino acids of the Gal4 DNA-binding domain and the acidic activation domain VP16 as before, however the activator is now flanked by an N-terminal Myc-tag and a C-terminal Flag-tag. One potential difficulty is that the high levels of the Gal4-VP16 fusion protein expressed by the *ADH1* promoter will induce squelching within the yeast cells, making regulation by the ERα ligand binding domain (LBD) invaluable. This domain will sequester the fusion protein in the cytosol by binding to the chaperone protein HSP90 [111, 112]. Only when the interaction is disrupted by the addition of β-oestradiol will the fusion protein be permitted access to the nucleus, thus creating a way to circumvent squelching of Gal4 [113].

To generate the individual lysine to arginine mutations (KXR), I used site-directed PCR mutagenesis. The new constructs were mini-prepped and verified through sequencing. The yeast strain SC41, which lacks endogenous Gal4 and has an endogenous *GAL*-driven reporter gene *MEL1*, was transformed by the new clones and two controls. Once transformed, western blots were done to make sure the fusion proteins were expressed (data not shown). One of the problems I encountered early on was that the cloned fusion proteins were not always

expressed. In particular, the substitutions at Lys27, Lys30 and Lys33 were cloned and sequence verified numerous times but failed to produce a band on a western blot. From the clones that did show expression, the cells were grown to an OD_{600} ~ 0.6 and then the nucleus was flooded with the given fusion protein using 1 μM estradiol for 45 minutes. Samples were harvested after induction, frozen with liquid nitrogen and stored in the freezer. Once all the samples were collected, they were thawed on ice and evaluated using a standard α-galactosidase assay (Figure 20).

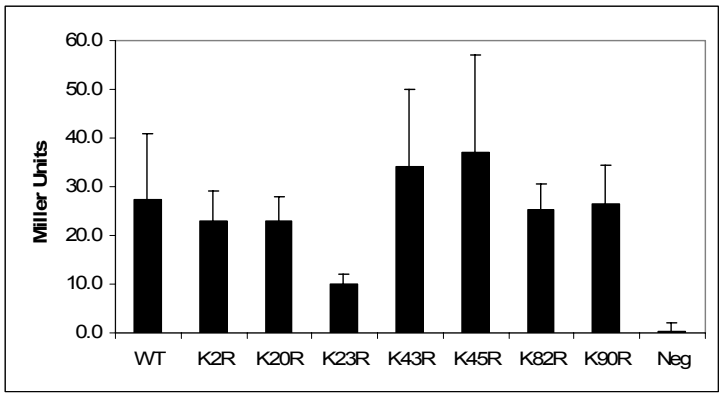

Figure 20: This graph depicts the values from the α-galactosidase assay in Miller Units. Three individual samples were evaluated and averaged. The x-axis labels show which mutation was being analyzed. 'WT' was the wild-type positive control and 'Neg' was a strain transformed with the empty vector.

The column labeled 'WT' is the wild-type form of the fusion protein, pRS313 Myc-Gal4-ER-VP16-Flag. The column labeled 'Neg' is the yeast strain transformed with the empty vector pRS313. Lys23 has previously been analyzed in the context of full-length protein and the mutant K23R is included as an

internal control. No conclusions could be determined from this experiment because the values were too low and very erratic. I mentioned previously that squelching occurs when Gal4 is expressed in cells at high levels and can lead to both lower expression levels of Gal4 and erratic data; therefore it is possible that sequestering the fusion protein in the cytosol through the ER LBD did not prevent the squelching mechanism from being engaged. Based on the nature of the data presented here I decided not to reattempt cloning mutants K27R, K30R and K33R and try another method to identify the site of the 'protective' mono-ubiquitylation event on the Gal4 DBD. Attempts to determine the ubiquitylation site of Gal4-VP16 by mass spectrometry also failed.

Search for Novel Proteins that Interact with Gal4

The identification of the kinase and ligase machinery that modifies the Gal4 DNA-binding domain was a problem of major importance. Therefore, I chose to start with a candidate gene approach. I began my study by analyzing *CKS1* and *RAD18*. The protein Cks1 was reported to regulate transcription and also to interact directly with the proteasome, thereby presenting the possibility that this kinase adaptor protein contributes to the phosphorylation of Gal4 at Ser22 [114]. The other protein of interest was Rad18, a ring-finger E3 ligase that heterodimerizes with Rad6 to mono-ubiquitylate histones, which can be important for the initiation of transcription [115]. If these two proteins were contributing to the function of the activator Gal4 then the 'protective' post-translational events on the DBD would be abrogated. As a result, a chromatin immuno-precipitation (ChIP) assay was done to analyze the occupancy profile of Gal4 in the *Δcks1* and *Δrad18* strain backgrounds (Figure 21). Samples grown in raffinose-containing medium were collected (-1) as well as samples 30 and 60 minutes following induction with galactose. The ChIP assays were done with a C-terminal Gal4 antibody or no antibody. The samples were then analyzed with primers that recognized the *GAL1-10* promoter. These data show that deletion of either gene does not impede Gal4 from occupying the *GAL1-10* promoter, suggesting these two proteins are not the enzymes of interest. However these proteins cannot be discounted completely because many genes in the yeast genome are redundant.

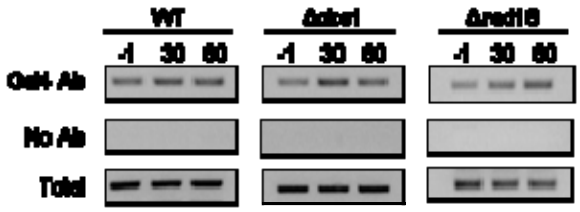

Figure 21: The data represent the results for the ChIP analysis of *Δcks1* and *Δrad18*. In the presence of raffinose- or galactose-containing media, the Gal4 protein was able to occupy the *GAL1-10* promoter in all three strain backgrounds.

Next, I took a broad candidate gene approach to identify the kinase responsible for the phosphorylation event at Ser22 and also the ligase machinery associated with mono-ubiquitylation of the Gal4 DBD. This protocol will screen yeast knock-out strains for their ability to grow on galactose-containing media. Cell cultures were washed to remove residual media and re-suspended in dH_2O to an $OD_{600} \sim 0.2$. A 10X dilution series was then made and the cultures were plated onto media containing glucose or galactose. All of the dilution series should grow on the plates with glucose in the media as they are a positive control and reference point for the strains inherent ability to grow (Figure 20, left panel). The deletion strains involved in galactose metabolism will grow well on glucose- but not galactose-containing media because the cell will not be able to produce enough energy to survive. The phenotype of interest is demonstrated with a *Δgal4* strain in Figure 22 (right panel). The strains with the same phenotype as *Δgal4* will give us our candidate genes for further testing.

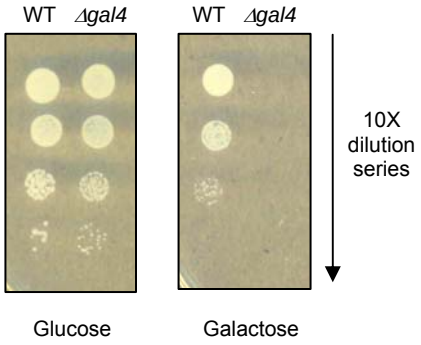

Figure 22: Growth-on-galactose assay. The left panel is the positive control that all of the cultures should grow on (glucose). The right panel is looking for a lack of growth-on-galactose phenotype (galactose).

In an attempt to identify the kinase that phosphorylates Ser22 of the Gal4 DNA-binding domain, I searched a yeast database for all known, non-essential kinases. This list was extensive so I narrowed my search to those that were characterized to localize in the nucleus. The 31 kinases of interest are listed in Table 1.

ALK1/ YGL021W	GAL83/ YER027C	PKH2/ YOL100W	SWE1/ YJL187C
BUB1/ YGR188C	HOG1/ YLR113W	PRR2/ YDL214C	TEL1/ YBL088C
BUD32/ YGR262C	IME2/ YJL106W	PTK2/ YJR059W	YAK1/ YJL141C
CHK1/ YBR274W	IRE1/ YHR079C	RIM15/ YFL033C	YBL009W
CTK1/ YKL139W	KSP1/ YHR082C	SLT2/ YHR030C	YCK3/ YER123W
CTK3/ YML112W	KSS1/ YGR040W	SNF1/ YDR477W	YMR291W
DUN1/ YDL101C	MEK1/ YOR351C	SPS1/ YDR523C	YPK2/ YMR104C
FUS3/ YBL016W	PHO85/ YPL031C	SSN3/ YPL042C	

Table 1: A list of the thirty-one, non-essential kinases that are characterized to localize in the yeast nucleus. They are listed by a common name and yeast accession number.

Several enzymes are associated with the conjugation of ubiquitin to a target protein. As a result, I searched a yeast database for those characterized as an E1, E2, E3 or E4 enzyme. E4 enzymes are typically characterized as contributing to the extension of poly-ubiquitin chains; however this group of enzymes is not well characterized so they were included in the search. The list was then expanded to include those proteins that contained domains commonly associated with the ubiquitylation machinery; an F-Box domain, a ring finger domain, a HECT domain or a WW domain. Several deletion strains for ligase components had already been analyzed in house by a growth-on-galactose assay (data not shown). These deletion strains did not show the phenotype of interest and were excluded from my study (Table 2). Candidate genes were also eliminated if they had been evaluated previously in a study of F-box proteins by the Tansey group (Table 3) [100]. Altogether, eighty-two genes were selected for analysis of their contributions to the *GAL* pathway (Table 4).

APC9/ YLR102C	HUL4/ YJR036C	TOM1/ YDR457W	SIZ1/ YDR409W
ASI1/ YMR119W	HUL5/ YGL141W	UBR1/ YGR184C	UBC8/ YEL012W
CDC26/ YFR036W	UBC7/ YMR022W	UBR2/ YGR024C	ELC1/ YPLO46C
GRR1/ YJR090C	SSM4/ YIL030C	UFD4/ YKL010C	RAD6/ YGL058W
HRD1/ YOL013C			

Table 2: This list shows which proteins were previously analyzed in our laboratory for their contributions to the *GAL* pathway.

COS111/YBR203W	MDM30/ YLR368W	SKP2/ YNL311C	YDR306C
DIA2/ YOR080W	MFB1/ YDR219C	UFO1/ YML088W	YJL149W
ELA1/YNL230C	RAX1/ YOR301W	YBR280C	YLR224W
GRR1/ YJR090C	RCY1/ YJL204C	YDR131C	YLR352W
HRT3/ YLR097C	REV7/ YIL139C		

Table 3: This table list F-Box proteins already analyzed by the growth-on-galactose method. The strains were shown to have no phenotype [100].

ASI3/ YNL008C	MAG2/YLR427W	RTT101/ YJL047C	UBS1/ YBR165W
ASR1/YPR093C	MDM30/ YLR368W	SAF1/YBR280C	UFD2/YDL190C
BUL1/ YMR275C	MFB1/ YDR219C	SAN1/ YDR143C	UFO1/ YML088W
BUL2/ YML111W	MMS2/ YGL087C	SIP5/YMR140W	VPS8/YAL002W
CDH1/ YGL003C	MND2/ YIR025W	SLX1/YBR228W	WSS1/ YHR134W
COS111/YBR203W	MOT2/YER068W	SLX8/YER116C	YBL104C
CST9/YLR249W	NFI1/ YOR156C	STE5/YDR103W	YBR062C
CUL3/ YGR003W	PEP5/YMR231W	STP22/YCL008C	YDR128W
DMA1/YHR115C	PEX2/YJL210W	SWM1/ YDR260C	YDR131C
DMA2/YNL116W	PEX4/ YGR133W	TUL1/ YKL034W	YDR266C
DIA2/ YOR080W	PEX10/YDR265W	UBC4/ YBR082C	YDR306C
ELA1/YNL230C	PEX12/YMR026C	UBC5/ YDR059C	YHL010C
FAR1/YJL157C	PIB1/ YDR313C	UBC11/ YOR339C	YJL149W
FMP27/YLR454W	PSH1/YOL052W	UBC12/YLR306W	YKR017C
GGA1/ YDR358W	QRI8/YMR022W	UBC13/ YDR092W	YLR224W
HRD1/ YOL013C	RAD5/YLR032W	REV7/ YIL139C	YLR352W
HRD3/ YLR207W	RAD16/YBR114W	RIS1/YOR191W	YMR187W
HRT3/ YLR097C	RAX1/ YOR301W	RKR1/YMR247C	YOL138C
IRC20/YLR247C	RCY1/ YJL204C	RMD5/YDR255C	YPR117W
ITT1/YML068W	RPL40A/ YIL148W	SKP2/ YNL311C	
LIP2/ YLR239C	RPL40B/ YKR094C	UBI4/ YLL039C	

Table 4: This list represents the 82 proteins I analyzed for contributions to the *GAL* pathway. This list was created from genes known to be a part of the ubiquitylation machinery and unknown genes with specified domains.

The list of candidate genes resulted in 123 yeast knock-out strains being purchased from the Open Biosystems yeast knock-out library. The growth-on-galactose assay referenced in Figure 22 was used to analyze each strain. The deletion strains *Δbud32*, *Δpho85, Δpkh2* and *Δsps1* were knock-outs of kinases that grew well on glucose but not galactose-containing media. None of the deletion strains representing the ligase machinery components showed the phenotype of interest (data not shown). Growth assays were repeated to analyze the four deletion strains of interest on raffinose-containing media (data not shown). Both *Δbud32* and *Δpho85* had deficiencies for growth on raffinose, suggesting these phenotypes were not caused by problems specific to the galactose utilization pathway. This left Pkh2 and Sps1 as my remaining candidates from the screen (Figure 23). The left panel shows the growth patterns of the strains on glucose-containing media and demonstrated that the strains grew similarly well. The right panel shows that *Δpkh2* and *Δsps1* mimicked the *Δgal4* phenotype when galactose was the sole carbon source in the media. The *Δseh1* strain was being re-checked for growth on galactose-containing medium. The phenotype was not as severe *Δpkh2* or *Δsps1* and was therefore not analyzed further.

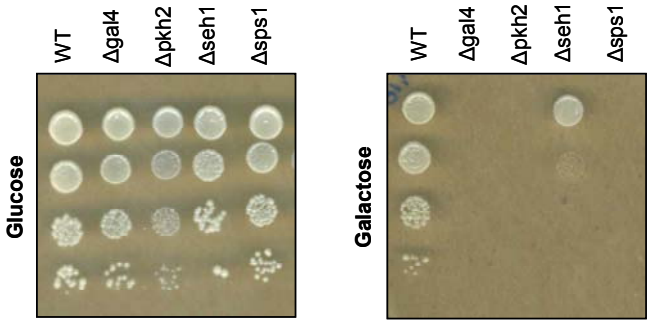

Figure 23: Growth-on-galactose assays. Wild-type (WT) was the positive control and *Δgal4* was the negative control within each plate. The left panel was the positive control for strain viability. The right panel was the test for the phenotype of interest.

A candidate gene approach is a simple and quick method that is commonly used in *S. cerevisiae* to narrow down genes that warrant additional studies. In my study, the deletion strains that showed a lack of growth on the sugar source galactose merely indicate a deficiency somewhere in the *GAL* pathway. Based on our model, the kinase I am looking for will inhibit Gal4-based transcription by decreasing the activators ability to occupy a *GAL* promoter. The *Δpkh2* and *Δsps1* strains were tested for their ability to activate *GAL*-based transcription. This was done by growing the strains on raffinose-containing media and inducing transcription with galactose for 45 minutes. An aliquot was taken to represent the strain growing on each sugar source. Both the *Δpkh2* and *Δsps1* strain were able to activate *GAL1* transcription (Figure 24), therefore neither of these kinases is likely responsible for phosphorylation of Gal4 at Ser22 but rather presumably play some role in galactose metabolism different from supporting Gal gene transcription.

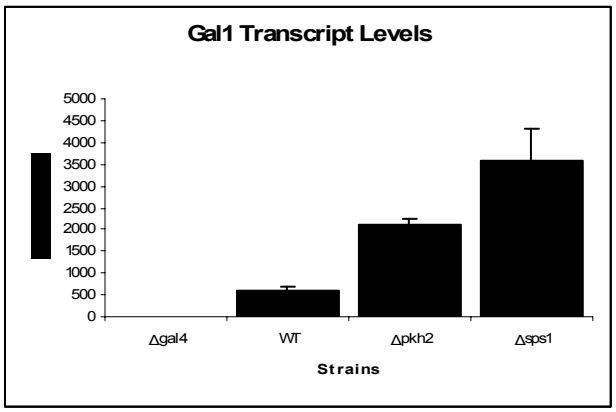

Figure 24: *Gal1* transcript levels of deletion strains. Data were graphed as fold increase of *GAL1* transcripts in galactose-containing media versus raffinose-containing media.

In parallel to the candidate gene approach, I designed a genetic selection to identify the E3 ubiquitin ligase that targets the Gal4 DNA-binding domain (DBD). The rationale behind the selection was as follows. Our current model is that this phosphorylation event at Ser22 triggers the recruitment of the ligase machinery that mono-ubiquitylates the Gal4 DBD (Figure 25, Top). When full-length Gal4(S22D) was over-expressed *in vivo*, the efficiency of the activator was partially restored (Figure 10), in contrast to the case of Gap71 or K23Q. This suggested that the aspartate poorly mimics the natively phosphorylated Ser22 residue when expressed at normal levels, weakening the ability of the E3 ligase to target Gal4. However, when the protein is expressed at high levels, binding to the E3 ligase may be restored, leading to activity. If true, this line of thinking suggests that over-expression of the E3 ligase in the presence of low levels of S22D might also rescue activity (Figure 25).

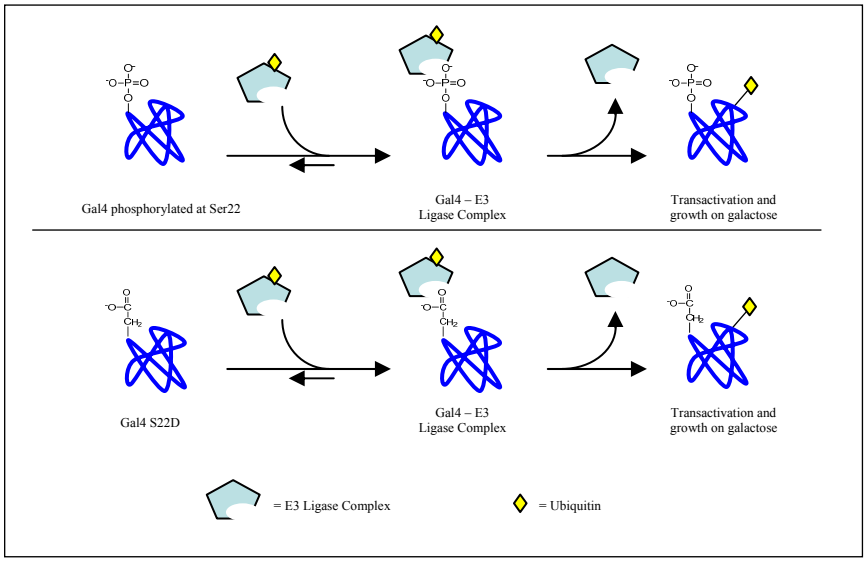

Figure 25: This diagram shows the similarities between a phosphorylation event at Ser22 and the aspartic acid (S22D) substitution. This diagram is meant to highlight the structures behind the hypothesis that aspartic acid behaves as a phosphoserine mimic.

To carry out this selection, a strain expressing full-length Gal4(S22D) was transformed with yeast genomic library (ATCC #37323). The library was established from a wild-type yeast genome, which will cause false positives resulting from plasmids harboring the *Gal4* gene. The panel of full-length Gal4 constructs analyzed previously was in the single copy plasmid pSB32 (Figures 10 and 11). This plasmid utilizes the same selection marker as the plasmid YEp13, which is the parent plasmid used for construction of the yeast genomic library. To obtain a different selection marker, I sub-cloned the full-length Gal4(S22D) mutant as well as the rest of the panel of Gal4 derivatives into the single copy plasmid YCp50. I transformed a *Δgal4* strain with the new YCp50 Gal4(S22D) construct. Plate derived growth assays have a threshold around 15% to 20% of

Gal4 activity. Gal4-based activity over the threshold will show growth while those with lesser expression will not survive. The α-galactosidase assay in Figure 10 shows the activity of the Gal4(S22D) mutant protein to be about 15% that of wild-type Gal4, which is just below the assays detection threshold. I did the growth-on-galactose assay comparing several colonies with the Gal4(S22D) construct and all appeared to have a similar no growth on galactose phenotype (data not shown). Once I had established my strain background, a yeast slurry was made from the parental strain that had been transformed with YCp50 Gal4(S22D). I then proceeded to set up three transformations: (1) with 400 μL of H_2O, (2) with 40 μg of empty YEp13 vector and (3) with 40 μg of the YEp13-based yeast over-expression library. A small amount of each transformation was plated on Ura^- media (the YCp50 selection marker) to make sure the strain had remained viable throughout the protocol. If the yeast were transformed with YEp13 then the transformed culture would grow on plates with Leu^-/Ura^- media. Therefore, all three transformations had a dilution series plated on Leu^-/Ura^- media with glucose to calculate the efficiency of the transformation protocol. The remainder of each transformation was screened on Leu^-/Ura^- plates with galactose. All three cultures grew equally well on the positive control test plates with media that lacked URA^-. Set (1) was transformed with water and showed no growth on Leu^-/Ura^- media with either glucose or galactose as the sugar source. Set (2) was transformed with empty vector to determine background levels.

Based on the dilution series, I estimated that set (2) had ~134,000 cells transformed successfully. Of these, only 110 colonies grew on the Leu^-/Ura^- media with galactose, giving a background level of < 0.1%. Set (1) and set (2) were controls while set (3) addressed the main question by challenging the Gal4(S22D) phenotype with the yeast over-expression library. Based on the dilution series of set (3), I calculated ~13,000 yeast cells had been transformed. This number was ten fold less than the empty vector control. This discrepancy most likely can be explained by the size of the empty vector versus the size of vectors found in the library. The vector backbone was already large at 10 kilobases (kb) while the library was the 10 kb backbone with fragments that are 5 kb to 20 kb. From the estimated 13,000 transformations in set (3), I had 343 colonies that grew on Leu^-/Ura^- media with galactose.

The yeast library used in the transformation did not exclude Gal4. Therefore my initial goal was to remove all colonies transformed with vectors containing *GAL4* fragments from the data set by identifying the genome fragment represented in each vector. After trying numerous enzymes, I found that the iProof High Fidelity DNA polymerase by BioRad was most efficient for yeast colony PCR. I generated a yeast colony PCR fragment for analysis by using primers based in the YEp13 vector that flanked the inserted fragment. I then prepped these samples and sequenced them in house using our ABI Prism® 3100 Genetic Analyzer. Sequencing reactions for both ends of the 343 fragments were

set up to determine what portion of the yeast genome they represented. The sequencing reactions typically resulted in 400-600 bases that were subsequently subjected to a BLAST search analysis. Those vectors that were shown to contain a portion of *GAL4* were removed from the data set.

The phenotype of the remaining strains required verification. The colonies were grown in liquid culture and then plated in a standard growth-on-galactose assay (Figure 26). The picture shows the *Δgal4* control in the upper left side. As seen in Figure 26, many false positives did exist in the original data set. False positives and strains transformed with a vector containing a section of *GAL4* were removed from the data set. This left 42 strains for further analysis.

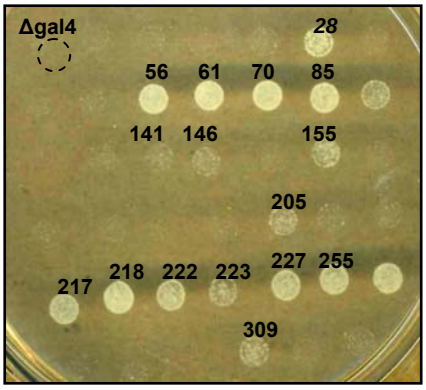

Figure 26: This is a sample of the growth-on-galactose assay that was used for re-verification of the strains identified in the over-expression screen. This picture shows strains plated on Leu^-/Ura^- media with galactose. All of the strains grew well on glucose (data not shown). The numbers indicate the order in the strains were collected.

Once the phenotypes of the strains had been confirmed, I went back to the data collected for the genomic content of each vector to determine what genomic fragments were in each vector. I analyzed the various sections of DNA to see what genes were present and if the gene had a promoter present. From the original 42 vectors I found 37 of those were unique. One fragment of the yeast genome was identified in 3 of the 42 vectors. Two additional fragments of the yeast genome were found in 2 vectors each. The candidate vectors were then analyzed for their ability to rescue several *GAL*-based phenotypes. I transformed a *Δgal4* strain with YCp50 Gal4, YCp50 Gal4(S22A), YCp50 Gal4(S22D) or empty YCp50 vector. The candidate vectors were then transformed into these strains and the growth-on-galactose assay was repeated. A candidate vector was removed from consideration if the phenotype of YCp50 (*Δgal4*) or YCp50 Gal4(S22A) was rescued by the over-expression of that particular vector since these were highly unlikely to be the E3 ubiquitin ligase I sought. I was left with 18 unique genomic fragments. The candidate vectors that contained the same genomic fragment did not vary by many bases and their content is listed in the Table 5 with the remaining hits. The 18 unique fragments were also tested for their ability to rescue *GAL*-based transcription. The preliminary data suggest all eighteen do promote transcription.

Hit Number	Accession	Name	Viability	Function
#21, #43, #146	YMR288W	HSH155	Essential	Splicing Factor
	YMR287C	MSU1	Non-Essential	RNase
#61, #123	YBR257W	POP4	Essential	Subunit of RNase P and RNase MRP
	YBR256C	RIB5	Essential	Riboflavin Synthase
#155, #171C	YDR217C	RAD9	Non-Essential	DNA Damage Dependent Check Point Protein
	YDR218C	SPR28	Non-Essential	Nucleotide Hydrolase
#28	YIL084C	SDS3	Non-Essential	Deacetylase Complex
	YIL083C		Essential	Homolog to hPPCS
#56	YDL169C	UGX2	Non-Essential	Hypothetical Protein
#70	YGL236C	MTO1	Non-Essential	Mitochondrial tRNA Modifications
	YGL235W		Non-Essential	Hypothetical Protein
	YGL234W	ADE5,7	Non-Essential	Adenine Synthesis
#85	YDR421W	ARO80	Non-Essential	Transcription Factor in Aromatic Amino Acid Catabolism
#141	YAR015W	ADE1	Non-Essential	Adenine Synthesis
	YAR018C	KIN3	Non-Essential	Kinase
#149	YOR311C	HSD1	Non-Essential	ER Membrane Protein
#205	YGL180W	ATG1	Non-Essential	Kinase Involved in Autophagy
#212	YOR021C		Non-Essential	Hypothetical Protein
#217	YNL020C	ARK1	Non-Essential	Kinase that contributes to cytoskeleton regulation
	YNL019C		Non-Essential	Hypothetical Protein
#218	YBR058C	UBP14	Non-Essential	Ubiquitin Protease
	YBR059C	AKL1	Non-Essential	S/T Kinase
	YBR058C-A	TSC3	Non-Essential	Sphingolipid Biosynthesis
#222	YCR088W	ABP1	Non-Essential	Actin Binding Protein
	YCR087C-A		Non-Essential	Hypothetical Protein with two zinc fingers
#223	YDL197C	ASF2	Non-Essential	Anti-Silencing Protein
	YDL198C	YHM1	Non-Essential	Mitochondrial GDP/GTP Transporter
	YDL199C		Non-Essential	Hypothetical Sugar Transporter
#227	YKR054C	DYN1	Non-Essential	Dynein
#255	YNL082W	PMS1	Non-Essential	ATP Binding Protein
#309	YMR241W	YHM2	Non-Essential	Mitochondrial DNA-Binding Protein

Table 5: This data set shows the genes that were found in the vectors that rescued the mutant Gal4(S22D) phenotype but not the Gal4(S22A) or empty vector (*Δgal4*) phenotypes. The 'hit' number indicates the order the colonies were originally collected.

Once the list of candidate vectors had been trimmed down, I wanted to characterize the individual genes within each candidate vector. In theory, if the over-expression of a particular protein rescues the Gal4(S22D) phenotype then the knock-out of the same gene could distress the galactose utilization pathway. Therefore, I assayed yeast strains lacking the non-essential genes found in Table 5 for their ability to grow on galactose (Figure 27). The left panel and right panels display the abilities of the strains to grow on glucose and galactose. Not all of the strains are shown, but those not pictured here exhibited a phenotype similar to the wild-type strain when grown on galactose.

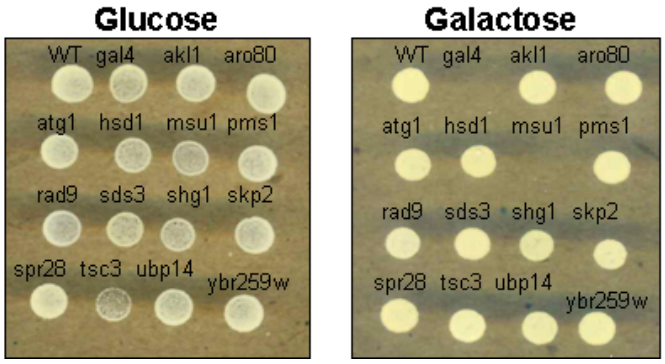

Figure 27: Growth-on-galactose assay. The left panel was the positive control plate with glucose-containing media. The right panel was the same strains being screened on galactose-containing media. The name above the spot indicates what gene in the strain was deleted. WT was the wild-type positive control sample and *Δgal4* is the negative control.

From the non-essential genes listed in Table 5, only the *Δmsu1* strain showed the preferred phenotype when compared to the *Δgal4* negative control strain (Figure 27, right). Even more interesting was that *MSU1* is from the

candidate vector that was isolated three times. *HSH155* was the other gene found on that plasmid but it could not be tested in this manner because it is essential for yeast survival.

I also need to parse out which specific gene in the candidate vector rescues the Gal4(S22D) phenotype since all of the 'hits' in the selection contained >1 gene. To do this, I attempted to clone the genes found in the candidate vectors into a new plasmid under the control of the *ADH1* promoter. The new constructs were designed to include a Myc-tag at the N-terminus of the protein for identification purposes. Primers were designed for each gene based on the annotation provided by the Saccharomyces Genome Database at www.yeastgenome.org. DNA agarose gels confirmed PCR products were produced for every primer set and that they were approximately the correct size. A majority the genes were successfully sub-cloned into the pRS313 vector backbone and were sequence confirmed to be cloned in frame with the already present Myc-tag. However, only five genes were shown to be expressed *in vivo*: *ATG1*, *DUT1*, *RIB5*, *SPR28* and *TSC3* (data not shown). Growth assays were then done to determine if the individually cloned genes could rescue the Gal4(S22D) phenotype (Figure 28). The wild-type strain is the positive control. Gal4(S22D) is a *Δgal4* strain transformed with the YCp50 Gal4(S22D) plasmid used in the selection screen. The remaining strains tested were the Gal4(S22D) strain transformed with the indicated plasmid. The plasmid pRS313 already had a

Myc tag inserted behind the *ADH1* promoter, pRS313 Myc (Empty). This was used as an additional control.

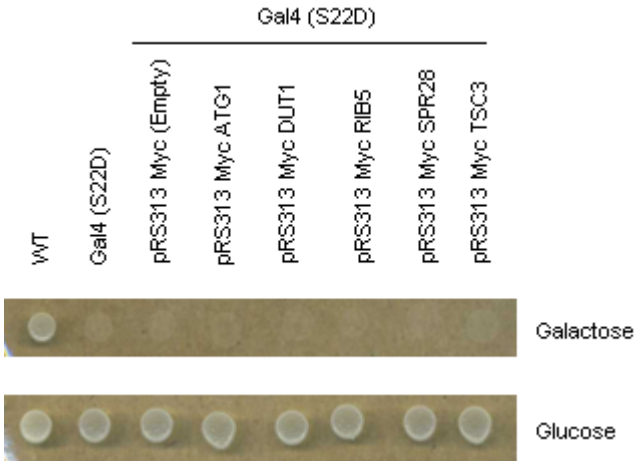

Figure 28: Growth assays testing the ability of the new Myc tagged clones to rescue the Gal4(S22D) no growth on galactose phenotype. WT is the positive control. Gal4(S22D) demonstrates the phenotype to be rescued.

Over-expression of these five genes was unable to rescue the Gal4(S22D) phenotype. The hit #205 only contained the gene *ATG1*. Therefore the original plasmid with the genomic DNA fragment has some unknown property rescuing growth on galactose as isolation of the plasmid and transformation into a new Gal4(S22D) strain also permitted growth on galactose. The remaining four genes were on 'hits' with more than one gene, suggesting the other gene found in the genomic fragment of the plasmid was causing the rescue of Gal4(S22D); however it is possible that it is the same situation as for hit #205.

The main 'hit' of interest was identified three times and contained the genes *MSU1* and *HSH155*. *MSU1* was cloned countless times but the plasmid never expressed the protein, while the HSH155 clone remained elusive. The screen also identified two fragments of the genome twice each. One contained the genes RIB5 and POP4, while the other contained the genes RAD9 and SPR28. All four genes were cloned from these spans of DNA but only RIB5 and SPR28 were expressed, suggesting POP4 and RAD9 were rescuing the growth on galactose phenotype. For the clones that were not expressed, PCR fragments were generated with at least three different primers sets using various digestion enzymes for insertion into the pRS313 plasmid. In addition, each fragment was cloned into the TOPO® TA Cloning system by Invitrogen. This provided a few additional clones but did not generate all the plasmids desired.

Analysis of *MSU1* and *SPS1*

The lack of growth on galactose phenotype encompasses the multi-step galactose utilization pathway that converts galactose into a usable substrate for entry into glycolysis. The candidate gene approach I discussed earlier revealed the putative serine/threonine kinase *SPS1* as a contributor to galactose metabolism [116, 117]. The over-expression screen data showed that a strain lacking *MSU1* had a phenotype similar to *Δsps1* and *Δgal4* strains. *MSU1* is a 3'-5' exonuclease that localizes in the mitochondria, where it interacts with the RNA helicase Suv3 to form the mitochondrial degradasome [118]. Many approaches to identify *GAL* genes have been reported over the years but neither *MSU1* nor *SPS1* was found in these studies. As a result, I set out to determine where these two genes contribute to the process of galactose utilization.

The initial study was a control experiment to make sure the deletion in the strain (*Δmsu1* or *Δsps1*) was generating the phenotype and not an unknown mutation elsewhere in the genome. The growth-on-galactose assay was redone with modifications. This time the deletion strain in question was transformed with a vector that adds the gene back in an attempt to rescue the no growth phenotype. The *Δmsu1* strain was not rescued by the only commercially available vector containing MSU1 because the protein was not expressed (data not shown). I also cloned Myc-MSU1 into pRS313 based on the information at the yeast database www.yeastgenome.org. This vector was sequence confirmed to be in

frame with the already present Myc tag but also did not express protein. Considering neither plasmid expresses the MSU1 protein, it is possible the start site is mis-annotated. The phenotype of the *Δsps1* strain was rescued when it was transformed with a plasmid expressing Myc-tagged *SPS1* (Figure 29). Expression of the Myc-tagged *SPS1* construct was confirmed by western blotting with a Myc antibody (data not shown).

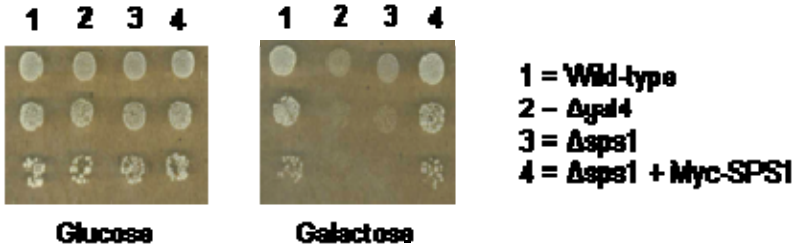

Figure 29: This growth assay shows the expression of Myc-SPS1 can rescue the *Δsps1* strain.

The original hypothesis was that the candidate gene approach and the over-expression screen would identify genes that regulate the activator Gal4. If the either the *Δmsu1* or *Δsps1* strain prevent Gal4 from functioning properly then Gal4-based transcription should be reduced. I began my study by growing the *Δmsu1* or *Δsps1* strains along side a wild-type and *Δgal4* control strain in raffinose-containing media to an OD_{600} ~ 0.6. A sample was collected and the cells were then split and induced with either glucose or galactose. Samples were harvested at 15, 30 and 60 minutes after induction with galactose and 60 minutes after induction with glucose (Figure 30). The *Δgal4* strain was included as a

negative control and showed no expression of the *GAL1* transcript. Surprisingly the wild-type, *Δmsu1* and *Δsps1* strain all showed a robust induction of *GAL1* transcripts upon induction with galactose-containing media. This result indicates that Msu1 and Sps1 play important roles in galactose utilization at some post-transcriptional step.

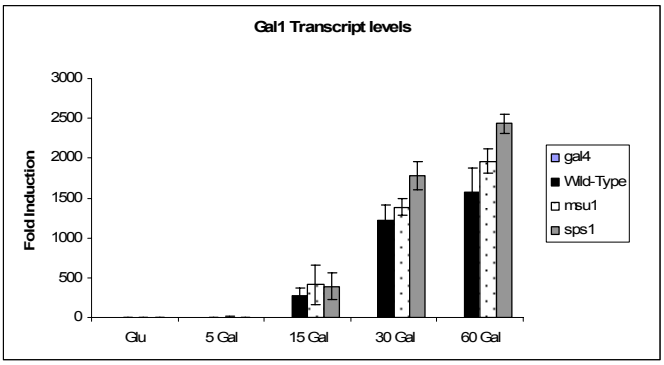

Figure 30: Transcript levels of the *GAL1* gene in the strains indicated. RNA samples were prepped and cDNA was generated from full-length transcripts. The resulting product was quantitated using qPCR.

Since the phenotype of the deletion strains was not rooted in *GAL*-based transcription, the next step was determining whether or not enzymes in the galactose utilization (Leloir) pathway were being expressed (Figure 3). Gal2 is a galactose permease that transports galactose in the cell. Once in the cell, the sugar is fed into the Leloir pathway, which is comprised of proteins regulated by Gal4. Once this process is complete, the modified sugar is fed into glycolysis as glucose-6-phosphate. Considering the cells grow well on glucose and have active

Gal4-based transcription, my hypothesis is that the problem lies in the Leloir pathway. Either a protein is not being expressed properly or the enzymes are being mis-regulated in these strain backgrounds.

One benefit of *S. cerevisiae* is that a researcher can easily create a tagged protein to be expressed in a yeast strain, however this has led to a paucity of commercially available antibodies targeted to native yeast proteins. As a result, none of the enzymes in the Leloir pathway have commercially available antibodies. To circumvent the antibody issue, I chose to knock-out the *GAL1*, *GAL5*, *GAL7* and *GAL10* genes in the wild-type, *Δgal4*, *Δmsu1* and *Δsps1* strain backgrounds. Once these genes are removed from the genome, I intended to add back the genes with purchased vectors from Open Biosystems that include an HA-His_6 tag. Because Gal4-based transcription is well characterized, most commercially available vectors are under the *GAL1-10* promoter. I have shown that Gal4-based transcription was not decreased in the *Δmsu1* and *Δsps1* deletion strains, thereby permitting use of the *GAL*-based promoter. The wild-type and *Δgal4* strain will serve as controls. I developed primers that would knock-out my gene of interest and insert the $HIS3^+$ gene with its own promoter through homologous recombination (Figure 31).

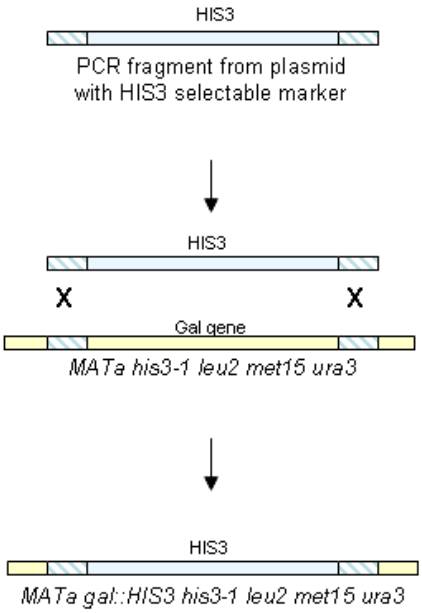

Figure 31: A diagram depicting how the gene of interest was replaced with a wild-type *HIS3*$^{+}$ allele using homologous recombination. A PCR generated fragment of *HIS3*$^{+}$ was flanked with sequences within the targeted gene. Typically the N-terminal of the fragment was designed to cross-over around the start site of the targeted gene and the C-terminal of the fragment should cross-over about 500 base pairs before the termination site in the targeted gene.

Primers were developed to generate a PCR fragment of the promoter and coding region for the *HIS3*$^{+}$ allele. The primers were also designed to include overhangs on each end of the PCR product that match interior portions of the gene being knock-out. These overhangs will facilitate homologous recombination to remove the targeted gene from the genome and replace it with the promoter and coding region of the *HIS3*$^{+}$ allele. Once the fragment was amplified and concentrated, it was used to transform the strain of choice. To screen for the transformed yeast, the samples were plated on HIS^{-} media. The Open Biosystems parental strain for the knock-out collection (BY4741) was not a clean deletion for *HIS3*$^{+}$ but a 187 base pair (bp) excision, *his3-1*. As a result, I designed an array of primers that allowed me to determine if the gene was knocked-out and if the linear fragment had also inserted into the *his3-1* site. The *his3-1* allele was a

deletion for base pairs 305 to 492. Therefore I designed primers within the deleted segment of *HIS3* (forward at bp 347 and reverse at bp 365) and primers that flank the deleted portion (forward at bp 9 and reverse at bp 608). When the primers are paired for PCR analysis of the putative knock-out strains, the primers flanking the deleted segment should result in a doublet that represents the innate *his3-1* allele and the newly added *HIS3*$^{+}$ allele (Figure 32, lane 3). The larger band represents the insertion of the wild-type *HIS3*$^{+}$ gene. The smaller band represents *his3-1* and shows that the recombination was specific to our targeted site. If the recombination occurred at both sites the lane would have one band representing two wild-type copies of *HIS3*$^{+}$. Lanes 2 and 4 are another confirmation that the *HIS3*$^{+}$ allele was integrated into the genome as they pair a primer within the excised *his3-1* segment with a primer that flanks the same excised segment.

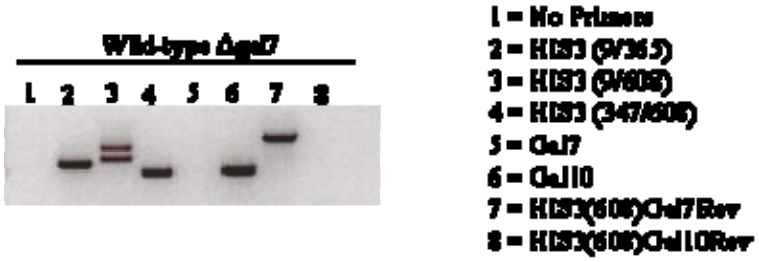

Figure 32: Genotypic profile of a wild-type strain background that was modified to be *Δgal7*. The PCR products show how the primers were paired to determine the correct genotype of the strain.

Along with identifying the *HIS3* profile of the strain I needed to verify the deletion of the targeted *GAL* gene. To analyze the genotype of the strain, internal primers were designed for each gene that was to be knocked out. For the wild-type background strain with *Δgal7* I did PCR to assay for the genes *GAL7* and *GAL10* (Figure 32, lanes 5 and 6). The expected genotype was shown with a band present for *GAL10* but not *GAL7*. To further confirm the genotype of the strain, I used a reverse primer for *HIS3*$^{+}$ at base pair 365 and a forward primer from the *GAL7* or *GAL10* promoter region. If the targeted gene was properly excised and replaced with *HIS3*$^{+}$, then the primer set with *GAL7* should yield a band while the primer set with *GAL10* should not. Again the results demonstrate the required genotype (Figure 32, lanes 7 and 8). The genotype of each strain was confirmed in a similar manner to the primer analysis demonstrated here.

Once the knock-out strains were established, I introduced tagged versions of the deleted genes back into the knock-out strains to monitor the expression profile of the proteins in the Leloir pathway in the *Δmsu1* and *Δsps1* deletion strains. Open Biosystems maintains a group of yeast tagged strains termed the Yeast ORF Collection. These genes are incorporated in the BG1805 vector backbone under the *GAL1-10* promoter. This vector fuses a 19 kD tag to the C-terminal of the gene and is comprised of a protein A domain, HIS_6 domain and a hemaglutination (HA) domain. Each knock-out strain was subsequently transformed by the BG1805 vector that contained the strains missing gene. These

strains were then grown with raffinose-containing media to an OD_{600} ~ 0.6. At that point a sample was taken as a negative control for the plasmid-based gene expression. The cells were then induced for 1 hour with galactose and another sample was harvested. The samples were all frozen with liquid nitrogen. Lysates were then made and equal amounts of total protein were loaded onto a SDS-Page gel for western blot analysis. The lysates made from the sample harvested under raffinose growing conditions showed no expression of the tagged proteins (data not shown). However, induction with galactose induced the expression of the tagged proteins in the wild-type, *Δmsu1* and *Δsps1* strain backgrounds (Figure 33). The strain background is indicated directly above the lane on the western blot. The first four lanes are the *Δgal5* knock-outs in the indicated strain background that are transformed with a plasmid expressing Gal5-HA-His_6 (~82 kDa). The next four lanes are the *Δgal7* knock-out strains in the given strain background that have been transformed with a vector encoding Gal7-HA-His_6 (~61 kDa). The final four lanes are the *Δgal10* knock-out strains that are transformed with a plasmid expressing Gal10-HA-His_6 (97 kDa).

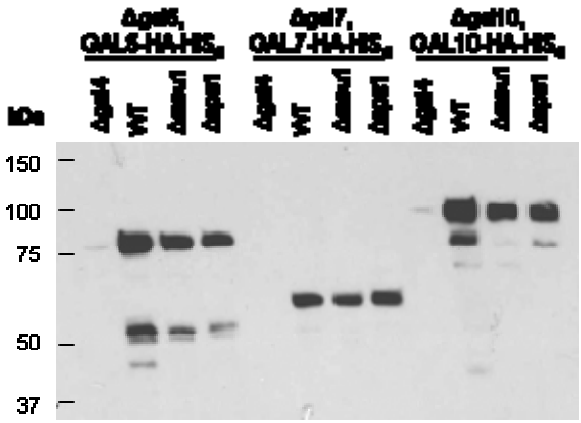

Figure 33: Protein expression profile of the knock-out strains. Within each strain background, the knocked-out gene was transformed back into the strain with a plasmid that encoded the missing *GAL* gene with a HA-His_6 tag. The samples represent the expression pattern of the strains after being induced 1 hour with galactose.

The hypothesis was that the phenotype associated with *Δmsu1* and *Δsps1* resulted from a deficiency in the translation of a protein(s) in the Leloir pathway. These data disprove that hypothesis, suggesting the proteins in the Leloir pathway are being positively regulated by MSU1 and SPS1. However, it is also possible that the problem lies in another unknown pathway that intersects with the Leloir pathway.

DISCUSSION

The model for *GAL4*-based transcription states that a phosphorylation event on Ser22 triggers a mono-ubiquitylation event elsewhere on the Gal4 DNA-binding domain (DBD) and that this sequential process positively affects the function of the activator. While the phosphorylation event has been characterized, less is known about the 'protective' mono-ubiquitylation event on the Gal4 DBD. The hypothesis was that the mutation of the mono-ubiquitylated lysine in the Gal4 DBD would result in a deficiency in the α-galactosidase assay. As a result, several lysines were individually mutated to arginine because this residue maintained the positive charge and cannot be ubiquitylated. However the data were too erratic to make any firm conclusions. In the end, this experiment did not delineate the lysine within the Gal4 DBD that was mono-ubiquitylated. Even if the data had been of higher quality, the lysines in the Gal4 DBD are clustered, opening the possibility for the mono-ubiquitylation event to shift to another lysine residue. In addition to the lysine scan, I did a mass spectrometry experiment in an attempt to identify the site of mono-ubiquitylation. The initial problem with this study was the lack of identifiable fragments from the fusion protein GST-Gal4(147)-VP16. The lysines within the DBD are clustered and 100% digestion would lead to very small, hard to detect, fragments. It is also possible that some secondary structure was not reduced in the SDS-Page gel,

which prevents trypsin from accessing the fusion protein. Several enzymes are now available that digest residues other than lysine and arginine in the protein. Based on the amino acid profile of the Gal4 DBD switching to the Glu-C (Asp and Glu) or Chymotrypsin-C (Phe, Trp, Tyr, Leu and Met) enzymes might be better. In addition, the discovery that ubiquitin does more than signal a protein for degradation has led to an explosion of studies in the field of ubiquitylation, which has in turn initiated advances in mass spectroscopy. The first identification of an ubiquitylation site *in vivo* was accomplished in 2002 [119]. Once a protocol is established that can identify fragments from the Gal4 DBD, I feel identification of the site of mono-ubiquitylation will be possible and that the answers would prove invaluable.

The relationship between the proteasomal ATPases and the stability of the activator-DNA interaction is still largely undefined. Determining the kinase and ligase machinery associated with these post-translation events on the Gal4 DBD will not only expand our knowledge of the process of destabilization but will give us additional tools to determine how the process is regulated. I began with a candidate gene approach that analyzed non-essential, nuclear kinases and numerous ligase machinery components. From these 123 strains, a few showed a no growth on galactose phenotype but permitted Gal4 to occupy the *Gal1-10* promoter and activate transcription, suggesting they did not contribute to the two post-translational modifications we were interested in. I also designed an over-

expression screen that was biased towards identifying the ligase machinery that modifies the Gal4 DBD. From this screen, 18 unique plasmids that rescued the Gal4 mutant S22D were isolated. The genes in these vectors were analyzed but no ligase machinery components were identified as modifiers of Gal4. While the screens did not determine the modification enzymes interact with Gal4, they did identify genes that affect the *GAL* pathway. Specifically, I analyzed the genes *MSU1* and *SPS1*. The deletion of these two genes resulted in a no growth on galactose phenotype. Most interestingly, the deletion strains permitted *GAL*-based transcription and did not repress the expression of the proteins in the Leloir pathway. These two proteins are most likely affecting the pathway through events independent of *GAL* gene transcription and translation.

CHAPTER IV

CONCLUSIONS AND DISCUSSION

The focus of this dissertation was to gain a better understanding of how the base sub-complex of the 19S RP regulates the stability of an activator-DNA complex. The initial model stated that post-translational modification events on Gal4 but not Gap71 were pertinent to the activators ability to resist dissociation from the promoter region. The dissection and analysis of the individual amino acid substitutions Chapter II suggested that Ser22 and Lys23 were the main contributors to the cumulative Gap71 phenotype. Subsequent work determined Ser22 is phosphorylated and that the charge at Lys23 is crucial for this event to occur. While Gal4 has been shown to be phosphorylated at three other positions, this was the first phosphorylation event shown to be required for the function of the activator. In addition, the phosphorylation of Ser22 triggers mono-ubiquitylation of the Gal4 DBD, which has been shown to 'protect' the integrity of the activator-DNA interaction [90, 94, 95]. The site of mono-ubiquitylation is still unknown, however several lysines in the Gal4 DNA-binding domain can be ruled out as ubiquitylation sites, leaving eleven possibilities. The proteasomatl ATPase-driven dissociation of an activator-DNA complex is not specific to Gal4 as the process has been shown to occur on Pho4 [90]. Therefore I did a sequence

alignment of four yeast zinc transcription factors and found that Lys18, Lys20 and Lys33 are conserved residues (Figure 34). The crystal structure of the Gal4 DNA-binding domain bound to DNA showed that Lys18 was part of the activator-DNA interface but that the Lys20 and Lys33 residues are not occluded by the tertiary structure of the protein [109]. As experiments proceed to analyze the lysine residues in Gal4, Lys20 and Lys33 would be good candidates for a focused study.

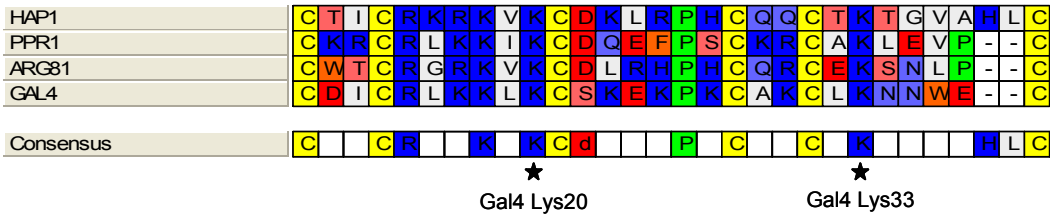

Figure 34: Amino acid alignment of yeast zinc finger transcription factors in yeast by Accelrys Gene. The Gal4 residues Lys18, 20 and 33 are conserved between the four activators.

Yeast screens were done in an unsuccessful attempt to identify the kinase and ligase machinery that modify the Gal4 DNA-binding domain. It is possible that the kinase of interest was among those tested, but has a redundant counterpart, making it impossible to identify through analysis of deletion mutants. Alternatively, the kinase and ligase of interest was not tested because it has not been characterized or because it is essential for yeast survival. While the first two possibilities are difficult to assess, the third could be tested by a method developed by Shokat and co-workers [120]. This study used chemical genetics to

create conditional alleles of individual tyrosine and serine/threonine kinases. This was done by modifying a residue in the active site of an expressed kinase so that it could be selectively and reversibly inhibited by a chemical analog of PP1, a Src-family-selective-inhibitor. All of the 20 essential, nuclear kinases in *S. cerevisiae* have been modified and are stably expressed *in vivo*, making this an excellent system to follow up the kinase-based candidate gene approach. A yeast two-hybrid assay could also identify the kinase and ligase machinery that modifies Gal4, however I believe a more biochemical approach is necessary. The lab has established *in vitro*-based assays that use yeast nuclear extracts to generate the phosphorylation event at Ser22 or the 'protective' mono-ubiquitylation event on the Gal4 DNA-binding domain. By testing aliquots from biochemical fractionation experiments in the phosphorylation and ubiquitylation assays, the list of potential proteins would be greatly diminished. In concert with a mass spectrometry analysis, a short list of candidate kinase and ligase enzymes could be assimilated.

The candidate gene approach and the over-expression screen in Chapter III resulted in two genes whose deletion prevented a yeast strain from growing on galactose-containing media, *MSU1* and *SPS1*. Msu1 is an exoribonuclease that complexes with Suv3 to create the mitochondrial degradasome. While no obvious connection existed between Msu1 and transcription, we found the protein interesting because it was not previously characterized to affect Gal4 function.

The other protein of interest was Sps1, a putative serine/threonine kinase whose deletion was linked to a transcription deficiency for a group of genes important to spore wall synthesis [116]. The individual deletions of these two genes were originally characterized to inhibit growth on galactose-containing media while growth on glucose-containing media was robust. Disappointingly, both deletion strains were fully functional with respect to *GAL*-based transcription and robustly express the enzymes in the galactose utilization pathway.

The cause of the phenotype associated with deletion of the putative kinase Sps1 is still unclear. However, one could speculate that this kinase regulates one of the enzymes involved in the galactose utilization pathway. If this were true, Sps1 would positively regulate an enzyme in the pathway or negatively regulate an inhibitor of the pathway; with either scenario the possibilities are numerous.

As for Msu1, a study by Chandel and co-workers suggested the mitochondria is hyper-active in yeast cells grown on galactose-containing media to compensate for the inefficient production of energy [121]. Therefore, it is possible that the *Δmsu1* phenotype is the result of the mitochondria being unable to up-regulate energy production when the cells are grown on galactose-containing media instead of inhibiting the galactose utilization pathway.

CHAPTER V

MATERIAL AND METHODS

Site-Directed PCR Mutagenesis:

Full Length Gal4: $S10_3$-Gal4 and $S10_3$-Gap71 were under the control of the native *GAL4* promoter in the single copy vector pSB32. Specific mutations in the Gal4 DNA sequence were generated by site-directed PCR mutagenesis of pSB32 $S10_3$-Gal4 using oligo pairs S22A and K23Q in Table 6.

Gal4(93)-VP16: pHCA ADH driven Gal4(93)-VP16 was the parental vector and was a generous gift from Dr. C. Picard [98]. pHCA Gal4(93)K23Q -VP16 was generated by site-directed PCR mutagenesis with the K23Q oligo pair found in Table 6.

Myc-Gal4(93)-ER-VP16-Flag: The parental vector was the pRS313 *ADH1* driven Myc-Gal4(1-93)-ER-VP16-Flag construct. ER symbolizes the ER ligand binding domain. The vector was modified using site-directed PCR mutagenesis with the primer pairs in Table 6, all sets.

Gal4-Gal11 Panel: The parental vector YEp351 Gal4-Gal11 was under the control of the *GAL4* promoter. The individual mutations were created using site-directed PCR mutagenesis with oligo pairs in Table 6 (S22A, S22D, K23Q, K23R and K25F).

Confirmation of Mutations: The site-directed PCR mutagenesis products were amplified with iProof High Fidelity DNA Polymerase (BioRad #172-5301). The samples were then exposed to the enzyme Dpn1 for 1 hour at 37° C to digest the original, unmodified parental vector. The remaining vector product was transformed into DH5α cells. The candidate colonies were grown, collected and their plasmids were extracted with the Wizard Plus SV Miniprep Kit (Promega #A1460). The resulting plasmid preps were screened by sequencing with the primer Gal4R (Table 6). The new constructs were stocked down in the Kodadek freezer system (Table 7).

Oligo Pair	Forward	Reverse
S22A	GCTCAAGTGCGCCAAAGAAAAACCGAAGTGCGCC	GGCGCACTTCGGTTTTTCTTTGGCGCACTTGAGC
S22D	GCTCAAGTGCGACAAAGAAAAACCGAAGTGCGCC	GGCGCACTTCGGTTTTTCTTTGTCGCACTT GAGC
K23Q	GCTCAAGTGCTCCCAAGAAAAACCGAAGTGCGCC	GGCGCACTTCGGTTTTTCTTGGGAGCACTTGAGC
K23R	GGCGCACTTCGGTTTTTCTCGGGAGCACTTGAGC	GCTCAAGTGCTCCCGAGAAAAACCGAAGTGCGCC
K25F	GCTCAAGTGCTCCAAAGAATTTCCGAAGTGCGCC	GGCGCACTTCGGAAATTCTTTGGAGCACTTGAGC
K2R	CGATAGAAGACAGTAGCCGCGAATTCCCGAGGTCC	GGACCTCGGGAATTCGCGGCTACTGTCTTCTATCG
K17R	GCGATATTTGCCGACTTAGAAAGCTCAAGTGCTCC	GGAGCACTTGAGCTTTCTAAGTCGGCAAATATCGC
K18R	GCGATATTTGCCGACTTAAAAGGCTCAAGTGCTCC	GGAGCACTTGAGCCTTTTAAGTCGGCAAATATCGC
K20R	GCCGACTTAAAAAGCTCCGGTGCTCC	GGAGCACCGGAGCTTTTTAAGTCGGC
K23R	GGCGCACTTCGGTTTTTCTCGGGAGCACTTGAGC	GCTCAAGTGCTCCCGAGAAAAACCGAAGTGCGCC
K25R	GCTCAAGTGCTCCAAAGAACGACCGAAGTGCGC	GCGCACTTCGGTCGTTCTTTGGAGCACTTGAGC
K27R	GCTCAAGTGCTCCAAAGAAAAACCGAGGTGCGC	GCGCACCTCGGTTTTTCTTTGGAGCACTTGAGC
K30R	CCGAAGTGCGCCCGGTGTCTGAAGACAACTGGG	CCCAGTTGTCTTCAGACACCGGGCGCACTTCGG
K33R	GTGCGCCAAGTGTCTGCGGACAACTGGG	CCCAGTTGTCCGCAGACACTTGGCGCAC
K43R	GCTACTCTCCCCGAACCAAAAGGTCTCCG	CGGAGACCTTTTGGTTCGGGGAGAGTAGC
K45R	GCTACTCTCCCAAAACCCGAAGGTCTCCG	CGGAGACCTTCGGGTTTTGGGAGAGTAGC
K82R	CCTTGACATGATTTTGCGAATGGATTCTTTACAGG	CCTGTAAAGAATCCATTCGCAAAATCATGTCAAGG
K90R	GCAGGGGTACCCAACAATGCTCGTATATCC	GGATATACGAGCATTGTTGGGTACCCCTGC
GAL4R		GGCTAGAAAGACTGGAACAG

Table 6: Primers used for site-directed PCR mutagenesis and subsequent sequencing for genotype verification.

Cloning:

Full Length Gal4-Based Constructs: The full-length Gal4-based constructs generated by site-directed PCR mutagenesis were sub-cloned into the multi-copy vector YEp351. The parental construct $S10_3$-Gal4 was under the native *GAL4* promoter. The promoter and gene were flanked by a unique BamHI site at the N-terminal and a unique HindIII site at the C-terminal. The single-copy vectors and empty YEp351 were digested with both enzymes for 2 hours in a 37° C incubator. The YEp351 vector was also digested for 1 hour with calf-intestinal phosphatase (CIP) to prevent the vector from re-annealing to itself. The bands were extracted from a 1% agarose gel and the fragments were purified using the Qiagen gel extraction kit (#28706). Ligations were done with NEB T4 Ligase over-night at 16° C per the given protocol (#M0202S).

Single-Copy Gal4-VP16 fusion protein: The original vectors were pGEX-CS-GST-Gal4-VP16, pGEX-CS-GST-Gap71-VP16 and pSB32-$S10_3$-Gal4. The DBD-VP16 and $S10_3$-Gal4 fragments were excised from the vectors with NcoI and PvuII. The pSB32-$S10_3$-Gal4 vector was also digested for 1 hour with CIP as before. The digestion of pSB32-$S10_3$-Gal4 left the native *GAL4* promoter and termination sites intact. The bands were extracted from a 1% agarose gel and the fragments were purified using the Qiagen gel extraction kit. The DBD-VP16 fragments were ligated into the digested pSB32 vector backbone with NEB T4

Ligase over-night at 16°C per the given protocol (#M0202S). This generated the DBD-VP16 constructs under the native *GAL4* promoter in a single-copy vector.

Multi-Copy Gal4-VP16 fusion proteins: Within the pSB32-S10$_3$-Gal4 vector the promoter and termination sites were flanked by BamHI and HindIII. These restriction sites remained in the new pSB32-DBD-VP16 constructs. The DBD-VP16 clones and YEp351 were digested with the BamHI and HindIII digestion pair. The YEp351 vector was also digested for 1 hour with CIP as before. The DBD-VP16 fragments were ligated into the digested YEp351 vector backbone with NEB T4 ligase as before.

YEp351-Gap71-Gal11: The Gal4 DBD has a unique SphI site. The *GAL4* promoter was maintained in the constructs by using a BamHI site inherent to the promoter region along with SphI. The YEp351 Gal4-Gal11 and YEp351 Gap71 (full-length) constructs were both digested with BamHI and SphI. The YEp351 Gal4-Gal11 vector backbone was also digested for 1 hour with CIP as before. The bands were extracted from a 1% agarose gel and the fragments were purified using the Qiagen gel extraction kit. The fragment from YEp351 Gap71 was ligated into the YEp351 Gal4-Gal11 backbone to yield YEp351 Gap71-Gal11.

YCp50 Panel of Constructs: The pSB32 vector backbone uses *LEU* as a selection marker. I needed the full-length Gal4 proteins in the context of a different selection marker. The vector YCp50 was deemed appropriate and the selection marker on the vector was *URA*. The full-length Gal4 clones driven by

the endogenous *GAL4* promoter can be excised from the pSB32 vector with co-digestion by BamHI and HindIII. YCp50 and the panel of pSB32 full-length Gal4-based constructs were digested with both enzymes. The YCp50 sample was also digested with CIP as before. The bands were extracted from a 1% agarose gel and the fragments were purified using the Qiagen gel extraction kit. The Gal4-based fragments were ligated in the YCp50 vector backbone with NEB T4 ligase as before.

Confirmation of Clones: Ligation products were transformed into DH5α cells. Colonies were grown, mini-prepped and analytical digestion was used to check for the appropriate insert. The DNA was then sequenced with the Gal4R primer to confirm the genotype.

pTK427	pSB32	pTK426	YEp351-$S10_3$-Gal4 (K25F)
pTK428	pSB32-$S10_3$-Gal4	pTK368	YEp351 Gal4 - Gal11
pTK429	pSB32-$S10_3$-Gap71	pTK498	YEp351 Gal4(K23R) – Gal11
pTK430	pSB32-$S10_3$-Gal4 (S22A)	pTK499	YEp351 Gal4(K25F) – Gal11
pTK431	pSB32-$S10_3$-Gal4 (S22D)	pTK500	YEp351 Gap71 – Gal11
pTK432	pSB32-$S10_3$-Gal4 (K23Q)	pTK 521	YCp50-$S10_3$-Gal4
pTK527	pSB32-$S10_3$-Gal4 (K23R)	pTK 522	YCp50-$S10_3$-Gap71
pTK433	pSB32-$S10_3$-Gal4 (K25F)	pTK 523	YCp50-$S10_3$-Gal4 (S22A)
pTK420	YEp351	pTK 524	YCp50-$S10_3$-Gal4 (S22D)
pTK421	YEp351-$S10_3$-Gal4	pTK 525	YCp50-$S10_3$-Gal4 (K23Q)
pTK422	YEp351-$S10_3$-Gap71	pTK 526	YCp50-$S10_3$-Gal4 (K25F)
pTK423	YEp351-$S10_3$-Gal4 (S22A)	pTK 513	Myc-Gal4-ER-VP16-Flag
pTK424	YEp351-$S10_3$-Gal4 (S22D)	pTK 565	Myc-Gal4-ER-VP16-Flag (S22A)
pTK425	YEp351-$S10_3$-Gal4 (K23Q)	pTK 566	Myc-Gal4-ER-VP16-Flag (S22D)
pTK528	YEp351-$S10_3$-Gal4 (K23R)	pTK 567	Myc-Gal4-ER-VP16-Flag (K23Q)
pTK495	YEp351 Gal4(S22A) – Gal11	pTK 568	Myc-Gal4-ER-VP16-Flag (K23R)
pTK496	YEp351 Gal4(S22D) – Gal11	pTK 569	Myc-Gal4-ER-VP16-Flag (K25F)
pTK497	YEp351 Gal4(K23Q) – Gal11		

Table 7: Location of new constructs in the Kodadek laboratory freezer system.

Yeast Transformations (Short Method):

Yeast cultures were grown over-night to saturation with glucose-containing media. Each sample started with 1 mL of culture being pelleted. The pellet was washed with dH_2O and was re-suspended in 10 μL of 10 mg/ml sonicated salmon sperm (Stratagene #201190). A 5-10 μL aliquot of plasmid was then mixed into the slurry. Then 600 μL of filtered Plate solution (0.1 M lithium acetate (pH 7.5), 0.1 M Tris (pH 7.5), 0.01 M EDTA and 20% PEG 3350) was added and the sample was vortexed. The samples were then incubated over-night at room temperature. In this method the cells settle so the following day 40 μL from the Bottom of each tube was plated on the appropriate drop-out media plate.

Yeast Transformation (Long Method):

This method is the yeast transformation protocol in the Hybrizap Manual (Stratagene Catalog #211341). A modified version of this method was used for the generation of the yeast knock-out strains. The Preparation of Yeast Competent Cells was scaled-down. In step two, the protocol adds 50 mL of saturated culture to 300 mL of fresh culture. I added 17 mL of saturated culture to 100 mL of fresh culture. I followed the remainder of this protocol and then modified step seven by re-suspending in 500 μL of TE-LiAc solution.

Alpha-Gal Assay:

TKY226 (*MATa Δgal4 ura3-52 leu2 MEL1*) was used for this assay. Cells were grown in raffinose-containing media at 30° C until an OD_{600} ~ 0.6. Cells were then induced with galactose for 30 minutes; at which point 50 ml aliquots were pelleted and frozen with liquid nitrogen. The α-galactosidase assay was based on one previously described [14]. Cells were disrupted by vortexing with zirconia/silica beads (BioSpec #11079105z) at 4° C in citric acid buffer (31mM citric acid, 39mM KH_2PO_4, pH 4.0). The cells were vortexed for 3 minutes in 1 minute increments with a 1 minute rest in between. The beads were removed from the sample and the lysates were spun at maximum speed for 30 minutes at 4° C. The cleared lysates were collected and the protein concentration was measured by Coomassie Plus protein assay (Pierce #1856210). Each reaction began with 50 μg of total protein. Warmed citrate buffer (30° C) was added to the lysate for a final volume of 200 μL. The reaction was started with 200 μL of PNP-Gal stock (6 mM citric acid in 7.75 mM Na_2HPO_4). The samples were rotated at 30° C. A 100 μL sample was removed at the given time points and the reaction was stopped with 900 μL of 0.1 M Na_2CO_3. The samples were evaluated using a spectrophotometer set at 400 nm. The final value was calculated by dividing the spectrophotometer reading by mg of protein used, minutes of incubation and the extinction coefficient 0.0182 (value = spec reading/mg protein/mins/0.0182).

ChIP Assay: The protocol is based on work by C.D. Allis [122, 123].

Fixing the Cells: The samples were grown over-night with raffinose-containing media. At OD_{600} ~ 0.6 the cells were induced with galactose for 1 hour. Aliquots of 100 mL each were collected and immediately fixed with a final concentration of 1% formaldehyde (Sigma #F-1635) for 10 minutes with constant agitation. The cross-linking reaction was quenched with a final concentration of 0.2 M glycine. The samples were mixed and then were spun down at 4000 RPM at 4° C for 5 minutes. The pellets were combined into one 50 mL conical tube with a dH_2O wash and spun again. The pellets were then frozen with liquid nitrogen and stored in a -80° C freezer.

Chromatin Preparation: All buffers are chilled and the samples are kept on ice throughout the protocol. Samples were thawed on ice, re-suspended in 5 mL fresh 0.1 M Tris, pH 8.0 with 10mM DTT and incubated on ice for 20 minutes. The samples were then spun at 4000 RPM at 4° C for 5 minutes. The pellets were each re-suspended in 5 mL Hepes/Sorb buffer (20 mM Hepes, pH 7.4 with 1.2 M sorbitol) and spun as before. The pellets were then re-suspended in 5 mL Hepes/Sorb buffer that contained 2 mg zymolase and 0.5 mM phenylmethanesulfonyl fluoride (PMSF). The mixtures were incubated in a 30° C shaker for 45 minutes. The reaction was stopped with 10 mL Pipes/Sorb buffer (20 mM Pipes, pH 6.8, 1 mM $MgCl_2$, 1.2 M Sorbitol and 0.5 mM PMSF). The pellets were spun as before and re-suspended with 5 mL of PBS with 0.5 mM

PMSF. The pellets were again spun as before and re-suspended with 5 mL Triton-X/Hepes buffer (0.25% Triton-X 100, 10 mM EDTA, 10 mM Hepes, pH 6.5 and 0.5 mM PMSF) with protease inhibitors (Roche #11873580001). The pellets were spun as before and re-suspended in 5 mL NaCl buffer (200 mM NaCl, 1 mM EDTA, 10 mM Hepes (pH6.5), 0.5 M PMSF and protease inhibitors). The samples were then transferred to a 1.5 mL eppendorf tubes with NaCl buffer. The samples were spun down for 30 seconds in a table top centrifuge at maximum speed. The pellets were then re-suspended in 0.5 mL of SDS lysis buffer (1% SDS, 10 mM EDTA, 0.1 % DOC, 50 mM Tris (pH 8) and 1.0 mM PMSF and protease inhibitors). At this point, the pellets were ready for sonication to shear the chromatin. Machines vary and will need to be optimized every time you switch sonicators to generate an average fragment size ~ 500 base pairs long. I used the Ultrasonic Homogenizer Model 150 V/T by Biologics, Inc. The machine was run at 40% power. Samples were kept in an ice-water bath, sonicated 10 seconds, allowed to rest 1 minute and sonicated for 5 more seconds. The resulting slurry was then spun at 4° C at maximum speed for 10 minutes. I kept the supernatant and added IP Dilution Buffer (0.01% SDS, 1.1% Triton-X 100, 1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris, pH 8 and 0.1 mM PMSF with protease inhibitors) to a final volume of 1 mL. Samples could be stored in the freezer at this point.

Immuno-precipitations: Reactions were typically set up the day the chromatin solutions were made. Three reactions were set up per chromatin solution sample (Gal4 Ab, No Ab and Total). Gal4 Ab samples were 100 μL of chromatin solution, 2 μL sonicated salmon sperm (Stratagene #201190), 12 μL Gal4 antibody and 985 μL IP Dilution Buffer. The No Ab controls were the same composition minus the antibody. The totals have 50 μL of chromatin solution, 2 μL sonicated salmon sperm and 50 μL IP Dilution Buffer. All samples were incubated over-night on a nutator at 4°C. The following day 40 μL of the ImmunoPure Protein A Plus bead slurry (Pierce #22811) was added to the Gal4 Ab and No Ab samples. All three set of samples were then incubated on a nutator at room temperature (RT) for 2 hours. The totals were then set aside while the other sample sets were washed and eluted. Each sample with Protein A beads had a 10 minute wash with TSE-150 (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 150 mM NaCl and 20 mM Tris, pH 8), TSE-500 (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 500 mM NaCl and 20 mM Tris, pH 8), LiCl Det (0.25 M lithium chloride, 1 mM EDTA, 1% IGEPAL, 1% DOC and 10 mM Tris, pH 8) and twice with TE (1 mM EDTA, 10 mM Tris, pH 7.5). The product was then recovered using 500 μL of elution buffer (1% SDS and 0.1 M $NaHCO_3$) for 30 minutes on a nutator at RT. The product was collected into new tubes and incubated over-night in a 65° C heat block to reverse cross-linking. The totals were also incubated over-night at

65° with 400 μL of elution buffer. The next day the samples were set up to precipitate in the freezer over-night with 50 μL sodium acetate and 1 mL cold ethanol. The following day the samples were spun down at 4° C at maximum speed for 30 minutes. The precipitated DNA was then washed with cold 70% ethanol and spun at maximum speed for 15 minutes at RT. Pellets were allowed to dry and then re-suspended in 1X Proteinase K buffer (5X buffer = 25 mM EDTA, 1.25 % SDS and 50 mM Tris, pH 7.5). To that mixture, 1 μL Proteinase K (Roche #13532400) was added and the samples were incubated in a 55° C water bath for 2 hours. The DNA was extracted with 500 μL phenol:chloroform:isoamyl (Sigma #P3803) and cleaned-up with 450 μL of chloroform. The DNA was precipitated and pelleted as before. The Ab and No Ab samples were re-suspended in 100 μL TE. The totals were re-suspended in 500 μL TE.

Sample Analysis: qPCR was used to quantitate the samples. Occupancy of the *Gal7* or *GAL1-10* promoter UAS_G sites was detected using the primers in Table 8. Another primer set was designed to detect background levels and were specific to a portion of Chromosome IX, the internal primer control (IPC). This section of the chromosome is non-coding and the primers are also listed in Table 8. The reaction was set up with iQ Sybr Green Supermix (Bio-Rad) and 2.5 μL of sample. The reactions were analyzed with the Bio-Rad ABI Prism 7700.

	Forward	Reverse
Gal7 Promoter	TTGCTTTGCCTCTCCTTTTG	TGCTCTGCATAATAATGCCC
GAL1-10 Promoter	GTGGAAATGTAAAGAGCCCC	CTTTATTGTTCGGAGCAGT
IPC	GGAAGGTGACGATAATACG	CGGACATCCTAAATCTTTGG

Table 8: List of primers used to analyze chromatin-immunoprecipitation samples.

Growth Assays:

Cultures were grown in raffinose-containing media until the cells were in mid-log phase. The cultures were then washed and re-suspended in dH_2O. Samples were then diluted to OD_{600} ~ 0.2. A 10 fold dilution series was generated with dH_2O. From each dilution set, I plated 5 μL of the sample. The dilutions were plated on yeast 'complete' media plates with raffinose, galactose or glucose as the sole sugar source. The plates were incubated at 30° until the glucose plate showed saturated growth for the OD_{600} ~ 0.2 dilution from the positive control strain. Then the plates were removed and photographed to determine the individual phenotypes of the strains plated. Normally the plates were allowed to incubate a day or two longer with new pictures taken each day.

Over-expression Screen:

The *Δgal4* strain YRG2 (MATa *ura3-52 his3-200 ade2-101lys2-801 trp1-901 leu2-3 112 gal4-542 gal80-538 L*YS2::UAS_{GAL1}-$TATA_{GAL1}$-HIS3 URA3::$UAS_{GAL4\ 17mers(x3)}$-$TATA_{CYC1}$-*lacZ*) was from the Hybrizap kit (Stratagene #978000) and was transformed with YCp50 Gal4(S22D). The new strain was

then transformed again with a yeast over-expression library from ATCC (#37323). The transformed yeast were plated on Leu^-/Ura^- drop-out plates with galactose as the sugar source to screen for vectors that rescued the Gal4(S22D) phenotype.

Sequencing:

Sequencing was done using the ABI Prism® 3100 Genetic Analyzer. PCR generated fragments were not purified. Instead 8 μL of PCR product was treated with 1 μL of SAP1 and 1 μL of Exo1 to remove residual primers and dNTPs. The mix was incubated for 30 minutes at 37°C and then for 15 minutes at 80°C to heat inactivate the enzymes. The Big Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems #4336917) was used to generate the fragment for sequencing. The samples were set up with 2 μL of Big Dye, 3 μL of 5X Buffer, 1 μL of 10 μM primer, 10 μL of PCR and 4 μL of dH_2O. The DNA was precipitated with 80 μL of precipitation cocktail (29 mL H_2O, 6 mL 3M sodium acetate, 125 mL ethanol). This suspension was then moved into the plate used by the sequencer. The plate was spun for 30 minutes at 4°C at 3000g. The plate was then briefly inverted on a paper towel to drain. The pellets were washed with 150 uL of chilled 70% ethanol and re-spun for 10 minutes at 4°C at 3000g. Again the plate was briefly inverted on a paper towel to drain. The pellets were

then covered with 24 uL Dimethylform and denatured for five minutes at 95°C in a PCR block. The samples were then loaded onto the machine and for analysis.

gDNA Extractions:

A yeast culture was grown to saturation in glucose-containing media and a 10 mL sample was harvested. The pellets were washed with dH_2O and re-spun. The pellet was re-suspended in 200 µL Lysis Buffer (10 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 2% Triton X-100 and 1% SDS). This was followed by the addition of 150 µL of silica beads and 200 µL PCI (Phenol:Chloroform:Isoamyl, 25:24:1). The samples were vortexed at room temperature (RT) for 1 minute and then spun for 5 minutes at maximum speed at RT. The aqueous phase was collected and mixed with 200 µL of chloroform. The samples were then spun again as before. The DNA was precipitated with 1mL cold EtOH and 300 mM NaCl. The samples were spun for 10 minutes at maximum speed at 4° C. Pelleted DNA was then washed with chilled 70% ethanol and spun again for 10 minutes at maximum speed at RT. The wash was removed and the DNA pellets were allowed to dry in a 37° C incubator. Samples were re-suspended in 50 µL of TE.

Plasmid extractions from yeast:

I began with the gDNA extraction protocol. The gDNA was then transformed into DH5α cells. Colonies were grown, mini-prepped, and sequenced to verify their content.

Gene Knock-Out Method:

I followed the method described at the website http://www.fhcrc.org/science/labs/breeden/Methods/Gene%20Deletion.html. The primers used to generate the knock-out fragments are in Table 9. The fragments were transformed into the appropriate yeast strain using the 'long method' because it was more efficient. The transformed yeast were plated on plates with selection media. Candidate colonies were grown and gDNA was extracted from each for verification of the genotype. PCR analysis of the gDNA determined which strains were correct. The strains used were obtained from Open Biosystems and turned out to not be true deletions of *HIS3*. Therefore the fragments generated could have inserted into the *HIS3* gene or the targeted gene of interest. As a result four primers were designed that were located in the *HIS3* allele. These primers along with *Gal* gene primers allowed analysis for the correct genotype (Table 9). The numbers next to the primer name indicate the location of the primer in the gene. A '+' sign indicates primers that are upstream the start site.

Several strains were generated in this process (Table 10). These new knock-out strains were then transformed by a vector carrying a tagged version of the gene that was previously deleted. These vectors were purchased from the Yeast ORF Collection maintained by Open Biosystems. Those new strains were also stocked down in the Kodadek Laboratory yeast strain stocks (Table 10).

Gal1ForKO2(34)	CCTGAGTTCAATTCTAGCGCAAAGGAATTACCAAG ACCATTGGCCGAAAAGTGCCCGTTTCGGTGATGAC
Gal1RevKO2(1481)	GCAAGGGCTTCTTTTACCTTTTCTATGTTGCCATTT GGGCCCCCTGGTTCCTGATGCGGTATTTTCTCCT
Gal5ForKO2(20)	CGGTTCCCACCAAACCATATGAAGACCAAAAGCCT GGTACCTCTGGTTTGCGCGTTTCGGTGATGAC
Gal5RevKO2(1633)	GCTTCAAGTATTCTTCAGCTGTCTTTTGGTATTGTG ATTTATCATCGTTCCTGATGCGGTATTTTCTCCT
Gal7ForKO2(7)	GCTGAAGAATTTGATTTTTCTAGCCATTCCCATAG ACGTTACAATCCACTAACCGCGTTTCGGTGATGAC
Gal7RevKO2(1039)	CCGAAGTTAAATCTCTTTGAGGCTCACCTAACAAT TCAAAACCAACCTTCCTGATGCGGTATTTTCTCCT
Gal10ForKO2(7)	GCTCAGTTACAAAGTGAAAGTACTTCTAAAATTGT TTTGGTTACAGGTGGTGCCGTTTCGGTGATGAC
Gal10RevKO2(2060)	CCGTTTTTCAAGGTTACACAATCTTTCCAGTTCTCT TGATTGATAGCTTCCTGATGCGGTATTTTCTCCT
HIS3For(9)	GCAGAAAGCCCTAGTAAAG
HIS3Rev(608)	GCTTCTCTTATGGCAACC
HIS3For(347)	GGTTTGGATCAGGATTTGC
HIS3Rev(365)	GCAAATCCTGATCCAAACC
Gal1For(+548)	GTGGAAATGTAAAGAGCCCC
Gal1Rev(1665)	CCATGTCTTTCTTCAGCG
Gal5For(+94)	CCATGTCTTTCTTCAGCG
Gal5Rev(1746)	GGTGAAAATCATTAAGCC
Gal7For(+311)	TTGCTTTGCCTCTCCTTTTG
Gal7Rev(1190)	CCGTTCAAGTCGACAACC
Gal10For(+342)	CTTTATTGTTCGGAGCAGT
Gal10Rev(2175)	GCAAATCTATGTGTTGCA

Table 9: List of primers used to knock-out specified yeast genes (GalXKO: x = gene number). These primers inserted *HIS3*$^{+}$ into the genome at the targeted locus. The list also includes those primers used to verify the genotype.

	Knock-Out Strains		Transformed Knock-Out Strains
TKY223	*WT*	TKY380	*WT, Δgal5, Gal5-HIS$_6$-HA*
TKY359	*WT, Δgal1*	TKY381	*WT, Δgal7, Gal7-HIS$_6$-HA*
TKY360	*WT, Δgal5*	TKY382	*WT, Δgal10, Gal10-HIS$_6$-HA*
TKY361	*WT, Δgal7*	TKY384	*Δgal4, Δgal5, Gal5-HIS$_6$-HA*
TKY362	*WT, Δgal10*	TKY385	*Δgal4, Δgal7, Gal7-HIS$_6$-HA*
TKY199	*Δgal4*	TKY386	*Δgal4, Δgal10, Gal10-HIS$_6$-HA*
TKY363	*Δgal4, Δgal1*	TKY388	*Δmsu1, Δgal5, Gal5-HIS$_6$-HA*
TKY364	*Δgal4, Δgal5*	TKY389	*Δmsu1, Δgal7, Gal7-HIS$_6$-HA*
TKY365	*Δgal4, Δgal7*	TKY390	*Δmsu1, Δgal10, Gal10-HIS$_6$-HA*
TKY366	*Δgal4, Δgal10*	TKY396	*Δsps1, Δgal5, Gal5-HIS$_6$-HA*
TKY232	*Δmsu1*	TKY397	*Δsps1, Δgal7, Gal7-HIS$_6$-HA*
TKY367	*Δmsu1, Δgal1*	TKY398	*Δsps1, Δgal10, Gal10-HIS$_6$-HA*
TKY368	*Δmsu1, Δgal5*		
TKY369	*Δmsu1, Δgal7*		
TKY370	*Δmsu1, Δgal10*		
TKY186	*Δsps1*		
TKY375	*Δsps1, Δgal1*		
TKY376	*Δsps1, Δgal5*		
TKY377	*Δsps1, Δgal7*		
TKY378	*Δsps1, Δgal10*		

Table 10: Strains generated from the yeast knock-out method were stocked down in the Kodadek laboratory yeast freezer system.

RNA Production and Analysis:

Cells were harvested and pelleted at the indicated time points. A typical collected sample was 1 mL of culture. The samples were frozen with liquid nitrogen and put in the -20° C freezer for one night. The cells were then thawed on ice and re-suspended in 400 μL of TES (10 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH 8) and 0.5 % SDS). The samples were mixed with 400 μL of acid-phenol and incubated for 45 minutes in the 65° C heat block with occasional agitation. The samples were then iced 5 minutes and spun at maximum speed at

4° C for 10 minutes. The supernatant was collected and remixed with 400 μL of acid-phenol. The samples were spun as before. The supernatant was again collected and this time mixed with 400 μL of chloroform. The samples were again spun as before. The supernatant was collected into a fresh tube. The RNA was precipitated by adding 40 uL of 3M sodium acetate (pH 5.3) and 1 mL cold 100% EtOH. The samples were spun at maximum speed at 4° C for 5 minutes. The samples were aspirated and the RNA pellet was washed with chilled 70% ethanol. The samples were spun again at maximum speed at 4° C for 5 minutes. The liquid was removed and the pellets were allowed to dry at room temperature for 10 minutes. Samples were re-suspended in 50μL of RNAse free water (Invitrogen #10977-015). Sample concentration was determined with 1.5 μL on the NanoDrop®ND-1000 Spectrophotometer. cDNA was then prepared using the ImProm-II™ Reverse Transcriptase kit (Promega #A3800). The 20 μL product was diluted with 80 μL water. From this 100 μL stock of cDNA 2.5 μL was used per PCR reaction or qPCR reaction.

BIBLIOGRAPHY

1. Matsui, T., et al., *Multiple factors required for accurate initiation of transcription by purified RNA polymerase II.* J Biol Chem, 1980. **255**(24): p. 11992-6.

2. Segall, J., T. Matsui, and R.G. Roeder, *Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III.* J Biol Chem, 1980. **255**(24): p. 11986-91.

3. Weil, P.A., et al., *Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA.* Cell, 1979. **18**(2): p. 469-84.

4. Orphanides, G., T. Lagrange, and D. Reinberg, *The general transcription factors of RNA polymerase II.* Genes Dev, 1996. **10**(21): p. 2657-83.

5. Buratowski, S., et al., *Five intermediate complexes in transcription initiation by RNA polymerase II.* Cell, 1989. **56**(4): p. 549-61.

6. Flores, O., I. Ha, and D. Reinberg, *Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF.* J Biol Chem, 1990. **265**(10): p. 5629-34.

7. Parvin, J.D. and P.A. Sharp, *DNA topology and a minimal set of basal factors for transcription by RNA polymerase II.* Cell, 1993. **73**(3): p. 533-40.

8. Ahn, S.H., M.C. Keogh, and S. Buratowski, *Ctk1 promotes dissociation of basal transcription factors from elongating RNA polymerase II.* EMBO J, 2009.

9. Dahmus, M.E., *Phosphorylation of the C-terminal domain of RNA polymerase II.* Biochim Biophys Acta, 1995. **1261**(2): p. 171-82.

10. Formosa, T., *FACT and the reorganized nucleosome.* Mol Biosyst, 2008. **4**(11): p. 1085-93.

11. Ptashne, M., *How eukaryotic transcriptional activators work.* Nature, 1988. **335**(6192): p. 683-9.

12. Giniger, E., S.M. Varnum, and M. Ptashne, *Specific DNA binding of GAL4, a positive regulatory protein of yeast.* Cell, 1985. **40**(4): p. 767-74.

13. Vashee, S., et al., *How do "Zn2 cys6" proteins distinguish between similar upstream activation sites? Comparison of the DNA-binding specificity of the GAL4 protein in vitro and in vivo.* J Biol Chem, 1993. **268**(33): p. 24699-706.

14. Johnston, S.A. and J.E. Hopper, *Isolation of the yeast regulatory gene GAL4 and analysis of its dosage effects on the galactose/melibiose regulon.* Proc Natl Acad Sci U S A, 1982. **79**(22): p. 6971-5.

15. Lohr, D., P. Venkov, and J. Zlatanova, *Transcriptional regulation in the yeast GAL gene family: a complex genetic network.* FASEB J, 1995. **9**(9): p. 777-87.

16. Vashee, S. and T. Kodadek, *The activation domain of GAL4 protein mediates cooperative promoter binding with general transcription factors in vivo.* Proc Natl Acad Sci U S A, 1995. **92**(23): p. 10683-7.

17. Ma, J. and M. Ptashne, *The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80.* Cell, 1987. **50**(1): p. 137-42.

18. Hidalgo, P., et al., *Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain.* Genes Dev, 2001. **15**(8): p. 1007-20.

19. Carlson, M., *Glucose repression in yeast.* Curr Opin Microbiol, 1999. **2**(2): p. 202-7.

20. Schuller, H.J., *Transcriptional control of nonfermentative metabolism in the yeast Saccharomyces cerevisiae.* Curr Genet, 2003. **43**(3): p. 139-60.

21. Johnston, S.A., J.M. Salmeron, Jr., and S.S. Dincher, *Interaction of positive and negative regulatory proteins in the galactose regulon of yeast.* Cell, 1987. **50**(1): p. 143-6.

22. Leuther, K.K. and S.A. Johnston, *Nondissociation of GAL4 and GAL80 in vivo after galactose induction.* Science, 1992. **256**(5061): p. 1333-5.

23. Sil, A.K., et al., *The Gal3p-Gal80p-Gal4p transcription switch of yeast: Gal3p destabilizes the Gal80p-Gal4p complex in response to galactose and ATP.* Mol Cell Biol, 1999. **19**(11): p. 7828-40.

24. Holden, H.M., I. Rayment, and J.B. Thoden, *Structure and function of enzymes of the Leloir pathway for galactose metabolism.* J Biol Chem, 2003. **278**(45): p. 43885-8.

25. Tschopp, J.F., et al., *GAL2 codes for a membrane-bound subunit of the galactose permease in Saccharomyces cerevisiae.* J Bacteriol, 1986. **166**(1): p. 313-8.

26. Maier, A., et al., *Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters.* FEMS Yeast Res, 2002. **2**(4): p. 539-50.

27. Lee, Y.C. and Y.J. Kim, *Requirement for a functional interaction between mediator components Med6 and Srb4 in RNA polymerase II transcription.* Mol Cell Biol, 1998. **18**(9): p. 5364-70.

28. Kang, J.S., et al., *The structural and functional organization of the yeast mediator complex.* J Biol Chem, 2001. **276**(45): p. 42003-10.

29. Jiang, Y.W., P.R. Dohrmann, and D.J. Stillman, *Genetic and physical interactions between yeast RGR1 and SIN4 in chromatin organization and transcriptional regulation.* Genetics, 1995. **140**(1): p. 47-54.

30. Bhoite, L.T., Y. Yu, and D.J. Stillman, *The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II.* Genes Dev, 2001. **15**(18): p. 2457-69.

31. Jiang, Y.W. and D.J. Stillman, *Regulation of HIS4 expression by the Saccharomyces cerevisiae SIN4 transcriptional regulator.* Genetics, 1995. **140**(1): p. 103-14.

32. Myers, L.C. and R.D. Kornberg, *Mediator of transcriptional regulation.* Annu Rev Biochem, 2000. **69**: p. 729-49.

33. Biddick, R. and E.T. Young, *Yeast mediator and its role in transcriptional regulation.* C R Biol, 2005. **328**(9): p. 773-82.

34. Baker, S.P. and P.A. Grant, *The SAGA continues: expanding the cellular role of a transcriptional co-activator complex.* Oncogene, 2007. **26**(37): p. 5329-40.

35. Swaffield, J.C., J.F. Bromberg, and S.A. Johnston, *Alterations in a yeast protein resembling HIV Tat-binding protein relieve requirement for an acidic activation domain in GAL4.* Nature, 1992. **360**(6406): p. 768.

36. Matsumoto, K., et al., *Function of positive regulatory gene gal4 in the synthesis of galactose pathway enzymes in Saccharomyces cerevisiae: evidence that the GAL81 region codes for part of the gal4 protein.* J Bacteriol, 1980. **141**(2): p. 508-27.

37. Swaffield, J.C., K. Melcher, and S.A. Johnston, *A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein.* Nature, 1995. **374**(6517): p. 88-91.

38. Russell, S.J., U.G. Sathyanarayana, and S.A. Johnston, *Isolation and characterization of SUG2. A novel ATPase family component of the yeast 26 S proteasome.* J Biol Chem, 1996. **271**(51): p. 32810-7.

39. Rubin, D.M., et al., *Identification of the gal4 suppressor Sug1 as a subunit of the yeast 26S proteasome.* Nature, 1996. **379**(6566): p. 655-7.

40. Swaffield, J.C., K. Melcher, and S.A. Johnston, *A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein.* Nature, 1996. **379**(6566): p. 658.

41. Baumeister, W., et al., *The proteasome: paradigm of a self-compartmentalizing protease.* Cell, 1998. **92**(3): p. 367-80.

42. Coux, O., K. Tanaka, and A.L. Goldberg, *Structure and functions of the 20S and 26S proteasomes.* Annu Rev Biochem, 1996. **65**: p. 801-47.

43. Pagano, M., *Cell cycle regulation by the ubiquitin pathway.* FASEB J, 1997. **11**(13): p. 1067-75.

44. Kloetzel, P.M., *Antigen processing by the proteasome.* Nat Rev Mol Cell Biol, 2001. **2**(3): p. 179-87.

45. Groll, M., et al., *Structure of 20S proteasome from yeast at 2.4 A resolution.* Nature, 1997. **386**(6624): p. 463-71.

46. Groll, M., et al., *A gated channel into the proteasome core particle.* Nat Struct Biol, 2000. **7**(11): p. 1062-7.

47. Glickman, M.H., et al., *The regulatory particle of the Saccharomyces cerevisiae proteasome.* Mol Cell Biol, 1998. **18**(6): p. 3149-62.

48. Confalonieri, F. and M. Duguet, *A 200-amino acid ATPase module in search of a basic function.* Bioessays, 1995. **17**(7): p. 639-50.

49. Glickman, M.H., et al., *A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3.* Cell, 1998. **94**(5): p. 615-23.

50. Deveraux, Q., et al., *A 26 S protease subunit that binds ubiquitin conjugates.* J Biol Chem, 1994. **269**(10): p. 7059-61.

51. Elsasser, S., et al., *Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome.* J Biol Chem, 2004. **279**(26): p. 26817-22.

52. Verma, R., et al., *Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system.* Cell, 2004. **118**(1): p. 99-110.

53. Thrower, J.S., et al., *Recognition of the polyubiquitin proteolytic signal.* Embo J, 2000. **19**(1): p. 94-102.

54. Verma, R., et al., *Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome.* Science, 2002. **298**(5593): p. 611-5.

55. Guterman, A. and M.H. Glickman, *Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome.* J Biol Chem, 2004. **279**(3): p. 1729-38.

56. Benaroudj, N., et al., *ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation.* Mol Cell, 2003. **11**(1): p. 69-78.

57. Ozkaynak, E., et al., *The yeast ubiquitin genes: a family of natural gene fusions.* EMBO J, 1987. **6**(5): p. 1429-39.

58. Ciechanover, A., et al., *ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation.* Proc Natl Acad Sci U S A, 1980. **77**(3): p. 1365-8.

59. Hershko, A., et al., *Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis.* Proc Natl Acad Sci U S A, 1980. **77**(4): p. 1783-6.

60. Hershko, A., et al., *Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins.* J Biol Chem, 1982. **257**(23): p. 13964-70.

61. Hershko, A., et al., *Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown.* J Biol Chem, 1983. **258**(13): p. 8206-14.

62. Weissman, A.M., *Themes and variations on ubiquitylation.* Nat Rev Mol Cell Biol, 2001. **2**(3): p. 169-78.

63. Lorick, K.L., et al., *RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination.* Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11364-9.

64. Scheffner, M., U. Nuber, and J.M. Huibregtse, *Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade.* Nature, 1995. **373**(6509): p. 81-3.

65. McGrath, J.P., S. Jentsch, and A. Varshavsky, *UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme.* EMBO J, 1991. **10**(1): p. 227-36.

66. Haglund, K. and I. Dikic, *Ubiquitylation and cell signaling.* EMBO J, 2005. **24**(19): p. 3353-9.

67. Peng, J., et al., *A proteomics approach to understanding protein ubiquitination.* Nat Biotechnol, 2003. **21**(8): p. 921-6.

68. Li, W. and Y. Ye, *Polyubiquitin chains: functions, structures, and mechanisms.* Cell Mol Life Sci, 2008. **65**(15): p. 2397-406.

69. Duncan, L.M., et al., *Lysine-63-linked ubiquitination is required for endolysosomal degradation of class I molecules.* EMBO J, 2006. **25**(8): p. 1635-45.

70. Krappmann, D. and C. Scheidereit, *A pervasive role of ubiquitin conjugation in activation and termination of IkappaB kinase pathways.* EMBO Rep, 2005. **6**(4): p. 321-6.

71. Glittenberg, M. and P. Ligoxygakis, *CYLD: a multifunctional deubiquitinase.* Fly (Austin), 2007. **1**(6): p. 330-2.

72. Crosetto, N., M. Bienko, and I. Dikic, *Ubiquitin hubs in oncogenic networks.* Mol Cancer Res, 2006. **4**(12): p. 899-904.

73. Salghetti, S.E., S.Y. Kim, and W.P. Tansey, *Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc.* Embo J, 1999. **18**(3): p. 717-26.

74. Salghetti, S.E., et al., *Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3118-23.

75. Reid, G., et al., *Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling.* Mol Cell, 2003. **11**(3): p. 695-707.

76. Lipford, J.R., et al., *A putative stimulatory role for activator turnover in gene expression.* Nature, 2005. **438**(7064): p. 113-6.

77. Salghetti, S.E., et al., *Regulation of transcriptional activation domain function by ubiquitin.* Science, 2001. **293**(5535): p. 1651-3.

78. Greer, S.F., et al., *Enhancement of CIITA transcriptional function by ubiquitin.* Nat Immunol, 2003. **4**(11): p. 1074-82.

79. Bres, V., et al., *A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter.* Nat Cell Biol, 2003. **5**(8): p. 754-61.

80. Kurosu, T. and B.M. Peterlin, *VP16 and ubiquitin; binding of P-TEFb via its activation domain and ubiquitin facilitates elongation of transcription of target genes.* Curr Biol, 2004. **14**(12): p. 1112-6.

81. Anzick, S.L., et al., *AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer.* Science, 1997. **277**(5328): p. 965-8.

82. Wu, R.C., et al., *SRC-3 coactivator functional lifetime is regulated by a phospho-dependent ubiquitin time clock.* Cell, 2007. **129**(6): p. 1125-40.

83. Wang, B., H. Suzuki, and M. Kato, *Roles of mono-ubiquitinated Smad4 in the formation of Smad transcriptional complexes.* Biochem Biophys Res Commun, 2008.

84. Shan, J., et al., *Dissecting roles of ubiquitination in the p53 pathway.* Ernst Schering Found Symp Proc, 2008(1): p. 127-36.

85. Li, M., et al., *Mono- versus polyubiquitination: differential control of p53 fate by Mdm2.* Science, 2003. **302**(5652): p. 1972-5.

86. Ferdous, A., et al., *The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II.* Mol Cell, 2001. **7**(5): p. 981-91.

87. Ferdous, A., T. Kodadek, and S.A. Johnston, *A nonproteolytic function of the 19S regulatory subunit of the 26S proteasome is required for efficient activated transcription by human RNA polymerase II.* Biochemistry, 2002. **41**(42): p. 12798-805.

88. Gonzalez, F., et al., *Recruitment of a 19S proteasome subcomplex to an activated promoter.* Science, 2002. **296**(5567): p. 548-50.

89. Archer, C.T., L. Burdine, and T. Kodadek, *Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking.* Mol Biosyst, 2005. **1**(5-6): p. 366-72.

90. Ferdous, A., et al., *The role of the proteasomal ATPases and activator monoubiquitylation in regulating Gal4 binding to promoters.* Genes Dev, 2007. **21**(1): p. 112-23.

91. Sikder, D., S.A. Johnston, and T. Kodadek, *Widespread, but non-identical, association of proteasomal 19 and 20 S proteins with yeast chromatin.* J Biol Chem, 2006. **281**(37): p. 27346-55.

92. Liang, S.D., et al., *DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA.* Mol Cell Biol, 1996. **16**(7): p. 3773-80.

93. Corton, J.C., E. Moreno, and S.A. Johnston, *Alterations in the GAL4 DNA-binding domain can affect transcriptional activation independent of DNA binding.* J Biol Chem, 1998. **273**(22): p. 13776-80.

94. Archer, C.T., et al., *Physical and functional interactions of mono-ubiquitylated transactivators with the proteasome.* J Biol Chem, 2008. **283**: p. 21789-98.

95. Archer, C.T., et al., *Activation Domain-dependent Monoubiquitylation of Gal4 Protein Is Essential for Promoter Binding in Vivo.* J Biol Chem, 2008. **283**(18): p. 12614-23.

96. Ferdous, A., et al., *Phosphorylation of the Gal4 DNA-binding domain is essential for activator mono-ubiquitylation and efficient promoter occupancy.* Mol Biosyst, 2008. **4**(11): p. 1116-25.

97. Gaudreau, L., et al., *Transcriptional activation by artificial recruitment in yeast is influenced by promoter architecture and downstream sequences.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2668-73.

98. Louvion, J.F., B. Havaux-Copf, and D. Picard, *Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast.* Gene, 1993. **131**(1): p. 129-34.

99. Sadowski, I., et al., *GAL4 is phosphorylated as a consequence of transcriptional activation.* Proc Natl Acad Sci U S A, 1991. **88**(23): p. 10510-4.

100. Muratani, M., et al., *The F box protein Dsg1/Mdm30 is a transcriptional coactivator that stimulates Gal4 turnover and cotranscriptional mRNA processing.* Cell, 2005. **120**(6): p. 887-99.

101. Mylin, L.M., J.P. Bhat, and J.E. Hopper, *Regulated phosphorylation and dephosphorylation of GAL4, a transcriptional activator.* Genes Dev, 1989. **3**(8): p. 1157-65.

102. Hirst, M., et al., *GAL4 is regulated by the RNA polymerase II holoenzyme-associated cyclin-dependent protein kinase SRB10/CDK8.* Mol Cell, 1999. **3**(5): p. 673-8.

103. Chi, Y., et al., *Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase.* Genes Dev, 2001. **15**(9): p. 1078-92.

104. Wu, G., et al., *Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase.* Mol Cell, 2003. **11**(6): p. 1445-56.

105. Li, Y., et al., *Yeast global transcriptional regulators Sin4 and Rgr1 are components of mediator complex/RNA polymerase II holoenzyme.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10864-8.

106. Lin, Y.S. and M.R. Green, *Mechanism of action of an acidic transcriptional activator in vitro.* Cell, 1991. **64**(5): p. 971-81.

107. Stringer, K.F., C.J. Ingles, and J. Greenblatt, *Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID.* Nature, 1990. **345**(6278): p. 783-6.

108. Lin, Y.S., et al., *Binding of general transcription factor TFIIB to an acidic activating region.* Nature, 1991. **353**(6344): p. 569-71.

109. Marmorstein, R., et al., *DNA recognition by GAL4: structure of a protein-DNA complex.* Nature, 1992. **356**(6368): p. 408-14.

110. Nalley, K., S.A. Johnston, and T. Kodadek, *Proteolytic turnover of the Gal4 transcription factor is not required for function in vivo.* Nature, 2006. **442**(7106): p. 1054-7.

111. Fankhauser, C.P., P.A. Briand, and D. Picard, *The hormone binding domain of the mineralocorticoid receptor can regulate heterologous activities in cis.* Biochem Biophys Res Commun, 1994. **200**(1): p. 195-201.

112. Picard, D., *Posttranslational regulation of proteins by fusions to steroid-binding domains.* Methods Enzymol, 2000. **327**: p. 385-401.

113. Wehrman, T.S., et al., *Enzymatic detection of protein translocation.* Nat Methods, 2005. **2**(7): p. 521-7.

114. Morris, M.C., et al., *Cks1-dependent proteasome recruitment and activation of CDC20 transcription in budding yeast.* Nature, 2003. **423**(6943): p. 1009-13.

115. Bailly, V., et al., *Yeast DNA repair proteins Rad6 and Rad18 form a heterodimer that has ubiquitin conjugating, DNA binding, and ATP hydrolytic activities.* J Biol Chem, 1997. **272**(37): p. 23360-5.

116. Friesen, H., et al., *Mutation of the SPS1-encoded protein kinase of Saccharomyces cerevisiae leads to defects in transcription and morphology during spore formation.* Genes Dev, 1994. **8**(18): p. 2162-75.

117. Iwamoto, M.A., et al., *Saccharomyces cerevisiae Sps1p regulates trafficking of enzymes required for spore wall synthesis.* Eukaryot Cell, 2005. **4**(3): p. 536-44.

118. Dziembowski, A., et al., *The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism.* J Biol Chem, 2003. **278**(3): p. 1603-11.

119. Marotti, L.A., Jr., et al., *Direct identification of a G protein ubiquitination site by mass spectrometry.* Biochemistry, 2002. **41**(16): p. 5067-74.

120. Bishop, A.C., et al., *A chemical switch for inhibitor-sensitive alleles of any protein kinase.* Nature, 2000. **407**(6802): p. 395-401.

121. Bell, E.L., et al., *The Qo site of the mitochondrial complex III is required for the transduction of hypoxic signaling via reactive oxygen species production.* J Cell Biol, 2007. **177**(6): p. 1029-36.

122. Dedon, P.C., et al., *A simplified formaldehyde fixation and immunoprecipitation technique for studying protein-DNA interactions.* Anal Biochem, 1991. **197**(1): p. 83-90.

123. Braunstein, M., et al., *Transcriptional silencing in yeast is associated with reduced nucleosome acetylation.* Genes Dev, 1993. **7**(4): p. 592-604.