Implementation and Optimization of an in vivo Photo-crosslinking Methodology to Define Direct Targets of Transcriptional Activators

by

Adaora Nwokoye

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Medicinal Chemistry)
in The University of Michigan
2012

Doctoral committee:

Professor Anna K. Mapp, Chair
Professor Ronald W. Woodard
Associate Professor George A. Garcia
Assistant Professor Garry D. Dotson

2012

Dedicated to my Parents, Husband, and Daughter

Acknowledgements

I will like to use this medium to thank everyone that has assisted me toward the accomplishment of my graduate degree. Starting with my research advisor Dr. Anna Mapp, I thank you for helping me to develop as an independent scientist. Anna, you challenged me to think extensively about my project while providing a welcoming environment that has facilitated the achievement of my research goals. Also, thank you for the emotional support and generosity towards my family. I truly appreciate it. In addition, I would like to thank my committee members for their time and for helping me think about my project with different perspectives. Thank you.

I want to thank all the members of the Mapp lab that have assisted me in anyway. Specifically, I would like to thank Dr. Chinmay Majmudar for helping me in my first few years of graduate school and Amanda Dugan for helping to proofread my thesis document. I have been privileged to work with some of the best scientists, who have serves as collaborators on various projects. To my collaborators, Dr. Lori Lee, Dr. Jody Lancia, Dr. Chinmay Majmudar, Dr. Malathy Krishnamurthy, Amanda Dugan and Hugo Fung, thank you for the expertise you brought to our projects. To all other members of the Mapp lab thank you for always making the lab fun and interesting.

I would like to thank all my friends for their support and encouragements throughout this endeavor. In particular, I will like to thank Dr Cherie Dotson for all the love she has showered on my family and the support to me personal. Cherie, I thank you, for the constant encouragement to push myself. Much thanks to Dr Helmy, my undergraduate MARC director, for providing the leverage to be successful in graduate school. Thanks Dr Helmy for your support throughout.

I can truly say that without the support from my family I will not have accomplished my goals in graduate school. To my husband, Anosike, words cannot articulate what your love and support has meant to me all these years. I would not have been able to achieve my goals in graduate school without your constant support. I love you. Somtochukwu, thank you helping me to remember to laugh and have fun every day I come home. You are the best daughter anyone can ask for. I love you. To my parents, thank you for being my cheerleaders, I am truly grateful, from the bottom of my heart, for your support. In particular, I want to thank my mom who moved to the USA to take care of my daughter so that I can finish graduate school. Thank you mom for always being there for me, and supporting me in every way with patience. You are the best mom ever, thank you.

And last but not the least; I thank God for the strength and enablement throughout my graduate experience.

Table of Contents

List of Figures

Chapter 1

Chapter 2

List of Tables

Chapter 2

Chapter 3

List of Abbreviations

A	alanine
a.a	amino acid
aaRS	tRNA synthetase
ATP	adenosine-5'-triphosphate
b-ME	mercaptoethanol
BSA	bovine serum albumin
C	carbon
CBP	CREB binding protein
CD	circular dichroism
Cdk8	cyclin dependent kinase
ChIP	chromatin immunoprecipitation
CREB	cAMP response element-binding
cryo-EM	cryo-electron microscopy
Cys	cysteine
D	aspartic acid
DBD	DNA-binding domain
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E	glutamic acid
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
F	phenylalanine
Fmoc	9-fluorenylmethoxycarbonyl
FP	fluorescence polarization
FRET	fluorescence resonance energy transfer
G	gram
Gal4	galactose protein 4
Gal80	galactose protein 80
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GCN4	general control nonderepressible 4
GTF	general transcription factor
H	hydrogen
HCl	hydrochloric acid
HER2	human Epidermal growth factor Receptor 2
His	histidine
hMDM2	human double-minute 2

HPLC	high pressure liquid chromatography
HRP	horseradish peroxidase
hSOD	human Superoxide dismutases
IPTG	isopropyl ß-D-1-thiogalactopyranoside
K	lysine
K_D	dissociation constant at equilibrium
kDa	kilodalton
KIX	kinase-inducible domain interacting domain of CBP
KOAc	potassium acetate
L	leucine
LiCl	lithium Chloride
M	methionine
MDM2	murine double minute 2
Med	mediator
MeOH	methanol
MES	2-(N-morpholino)-ethane sulfonic acid
MgOAc	magnesium acetate
min	minutes
MP	masking protein
mRNA	messenger RNA
MudPIT	multidimensional protein identification technology
N	asparagine
NaCl	sodium Chloride
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NRSF	neuron-restrictive silencer factor
NUA4	nucleosome acetyltransferase of histone H4
ONPG	ortho-Nitrophenyl-ß-galactoside
P	proline
p53	protein 53
pAzpa	p-Azido-L-phenylalanine
pBpa	p-benzoyl-L-phenylalanine
PBS	phosphate Buffered Saline
PDB	protein data bank
PEAS	N-((2-pyridyldithio) ethyl)-4-azidosalicylamide
PEG	polyethylene glycol
Phe	phenylalanine
PIC	pre-initiation complex
PNA	peptide nucleic acids
PTC124	premature termination codon 124
PVDF	polyvinyl difluoride
REST	repressor element 1 silencing transcription factor
RNA	ribonucleic acid

RNA Pol II	RNA Polymerase II
RP-HPLC	reverse-phase HPLC
SAGA	Spt-Ada-Gcn5-Acetyltransferase
SCX	strong cation exchange
SDOM	standard deviation of the mean
SnRNA	small nuclear ribonucleic acid
Srb4	subunit of the RNA polymerase II mediator complex 4
Ste12	transcription factor that is activated by a MAP kinase signaling cascade
Sug1 (RPT)	regulatory Particle Triple-A protein 1
Sug2 (RPT)	regulatory Particle Triple-A protein 2
SUR2	sulfonylurea receptors
SWI/SNF	switch/Sucrose nonfermentable
T	threonine
TAD	transcriptional activation domain
TAF	TATA binding protein associated factor
TBP	TATA-binding protein
TBS	tert-butyldimethylsilane
TFA	trifluoroacetic acid
TFIIA	transcription factor II A
TFIIB	transcription factor II B
TFIID	transcription factor II D
TFIIE	transcription factor II E
TFIIF	transcription factor II F
TFIIH	transcription factor II H
TFO	triplex-forming oligonucleotide
UAS	upstream activation sequence
VEGF-A	vascular endothelial growth factor A
VP16	Viral protein 16
W	tryptophan
WT	Wild type
Y	tyrosine
ZFP	zinc finger protein
Zn	zinc

Abstract

Implementation and Optimization of an
in vivo Photo-crosslinking Methodology to Define Direct Targets of
Transcriptional Activators

by

Adaora Nwokoye

Chair: Anna K. Mapp

Protein-protein interactions are used to accomplish many biological processes. Understanding the identity of and mechanisms governing directly interacting proteins is instrumental to offering therapeutic interventions for diseases they facilitate. For example, in transcription many human diseases have strong correlations with alterations in gene expression. Thus, there is intense interest in the development of chemical agents to restore aberrant gene expression to normal levels. However, in order to ultimately find therapies for diseases associated with altered transcription profiles, there needs to be an in-depth understanding of how transcriptional activators interact with their transcriptional machinery partners (coactivators). Currently, very few direct binding partners of transcriptional activators are known, let alone structurally characterized, making the generation of tailored screens for inhibitors of

activator−coactivator interactions challenging. To better understand activator−coactivator interactions, we probed for direct binding partners of activators in vivo, using an enhanced tRNA/tRNA synthetase pair developed to site specifically incorporate the nonnatural amino acid *p*-benzoyl-L-phenylalanine (pBpa) into the amphipathic activators. Initially we started with the model prototypical yeast transcriptional activator, Gal4, and later expanded our studies to two other prototypical activators, Gcn4 and VP16.

First, we used a powerful method, nonsense suppression, to incorporate pBpa, which has a crosslinking moiety, into Gal4. Using pBpa-containing constructs of Gal4 we carried out in vivo photo-crosslinking experiments in the yeast strain LS41. Crosslinked activator−coactivator complexes were immunoprecipitated and analyzed by Western blotting. Before identifying the binding partners of Gal4, we determined whether pBpa was readily incorporated into the Gal4 TAD and if these photo-crosslinkable constructs were transcriptionally active. Results showed that all Gal4 pBpa constructs were permissive for the incorporation of pBpa, produced the full length protein and were transcriptionally functional. Our initial in vivo crosslinking experiments revealed a well-characterized binding partner of Gal4, the masking protein Gal80. Further, using in vivo photo-crosslinking again, we were able to capture other targets that engage in modest-affinity and/or transient interactions with transcriptional activators (Gal4, Gcn4 and VP16) including Med15, Taf12, Tra1 and Snf2. In the future, in vivo photo-crosslinking methodology can be used to define both tight and modest-affinity protein-protein interactions.

CHAPTER 1
INTRODUCTION

A. Summary

In response to signaling cues, a needed set of genes is transcribed to RNA, which is then translated into, proteins. Precise regulation of this process allows specific functions to be efficiently carried out for the maintenance of numerous cellular processes required within an organism. Since transcription of specific genes is a vital process required to maintain normal human physiology, misregulation of transcription can produce severe consequences. In fact, gene misregulation has been linked to almost all human diseases, either as a cause or an effect.[1-3] The need to discover small molecules that can modulate transcription and in doing so address disease-associated transcriptional misregulation continually builds, but first, scientists have to identify the specific protein interactions needed for transcription to occur.

As alluded to above, a major problem blocking the development of small molecule transcriptional modulators is that direct protein-protein interactions (PPIs) for the process of transcription to occur are not fully understood. Knowing direct proteins that interact with transcriptional activators, researchers can further study their binding interfaces. Information from these types of studies can be used to generate tailored-screens for small molecules that can modulate

transcription. My dissertation will focus on developing and implementing an in vivo photo-crosslinking methodology that is used to tease out and characterize the direct binding partners of the prototypical transcriptional activator, Gal4. Furthermore, the direct partners of Gal4 will be compared to two other well-studied prototypical activators to determine if these activators use similar mechanisms for transcription. The major implication when activators have similar mechanisms is that the same small molecule may be used to modulate the transcription of those activators.

B. Overview of transcription

Transcription is an essential process needed for the viability of an organism. During transcription, the cell's genetic material (DNA) is transcribed to RNA. Generally, transcription is initiated by internal and external cellular stimuli, which causes a class of proteins know as transcriptional activators to undergo post-translational modifications and sometimes translocate to the nucleus, localize there, and bind specific DNA sites upstream of a gene (Figure 1-1).[4]

Once localized to their genes, transcriptional activators recruit numerous multi-protein coactivator complexes in order to assemble the transcriptional machinery at the promoter and initiate transcription.[5-7] In eukaryotes, of the 3 types of RNA polymerases, pol II is responsible for transcribing all protein encoding mRNA and SnRNA; subsequently mRNA is translated into protein by the ribosome. [8-10] The focus of this chapter is to review the current model of the protein-protein contacts needed for pol II transcription.

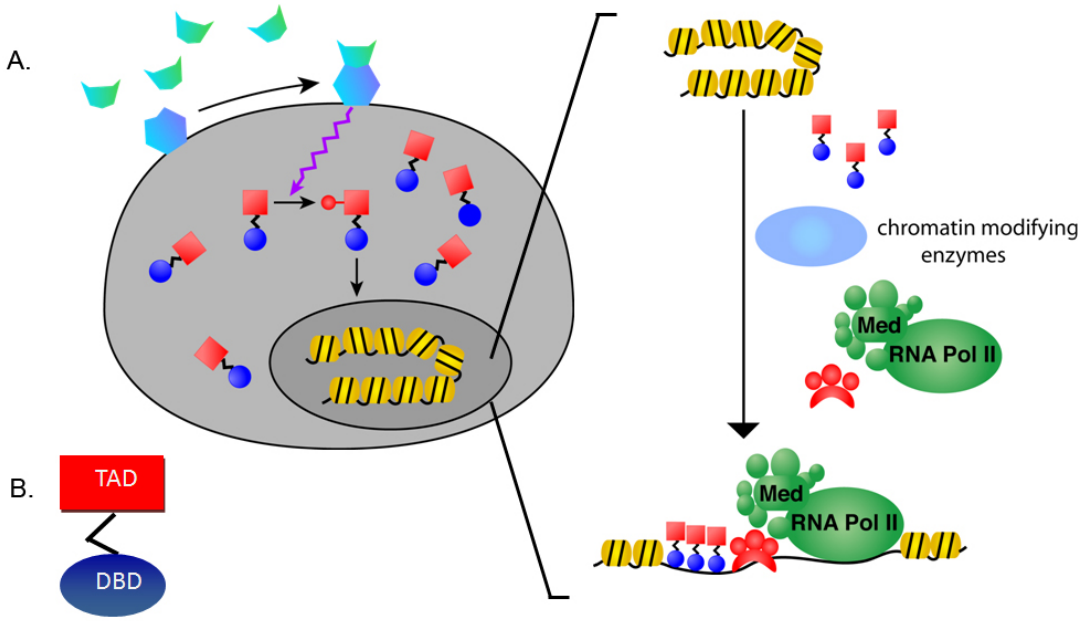


Figure 0-1-1: A general schematic of endogenous transcription. (A) Sequence of events that transcriptional activators participate in and protein complexes they interact with for transcription to be achieved. (B) Architecture of transcriptional activators, which function as modular proteins, and contain a transcriptional activator domain (TAD) and/or a DNA binding domain (DBD).

The core promoter within the gene serves as a platform for transcriptional activator-stimulated assembly of the transcriptional pre-initiation complex (PIC) - TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH and RNA polymerase II, and other proteins complexes like the mediator, chromatin remodeling complexes, SAGA and Swi/Snf complexes as shown in Figure 1-2.[11]

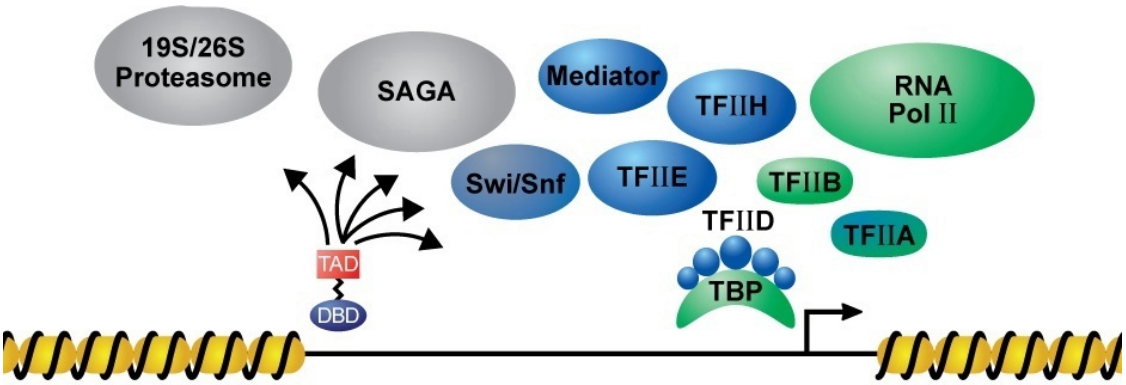


Figure 1-2: Protein complexes (over 3 decades) that have been associated with TAD transcription and/or have been proposed as targets of TADs. Transcriptional activators interact with coactivators to assemble the PIC and specify the transcription initiation site. This figure has been slightly modified from the original publication.[4]

Collectively coactivators specify the transcriptional start site.[11-13] It is generally accepted that transcriptional activators interact with specific proteins within the complexes that make up the PIC. Data from chromatin immunoprecipitation (ChIP) and in vitro fluorescence resonance energy transfer (FRET) experiments have supported a sequential recruitment model in which the prototypical yeast activator Gal4 first recruits SAGA, a chromatin modifying complex, to the promoter, followed by the Mediator scaffolding complex, RNA Polymerase II, TFIIH, TFIIE, THIIB and TBP roughly at the same time.[14-19] However, while activators such as Gal4 stimulate PIC formation through direct binding interactions with the transcriptional machinery, the identities of the binding partners within the PIC and other transcription related complexes remains hotly debated.[4]

C. Transcriptional activators

A transcriptional activator may be described as a protein that has the ability to bridge two chemically different complex molecules, DNA and proteins, for the transcription of a gene to occur. In order to accomplish this, transcriptional activators are modular proteins that minimally contain two domains that function independently (Figure 1-1B).[20, 21] The two domains, the DNA binding domain (DBD) and the transcriptional activation domain (TAD) together respond to extracellular and intracellular signals to bind to specific sequences on DNA, and then recruit proteins and protein complexes that are important for high-level transcription.[20, 21]

Transcriptional activators interestedly have the ability to swap these domains and still maintain their function as demonstrated in a domain swapping experiments carried out by Brent and Ptashne (Figure 1-3).[22, 23] In these experiments, a fusion protein between the DBD of a bacterial transcription repressor LexA and the TAD of a yeast transcription activator was able to function as a transcriptional activator at the LexA binding sites.[22] Results from the above experiment have opened the door to possible development of artificial transcription activators where the functional domains of endogenous transcriptional activators are swapped out for their artificial equivalents.[24]

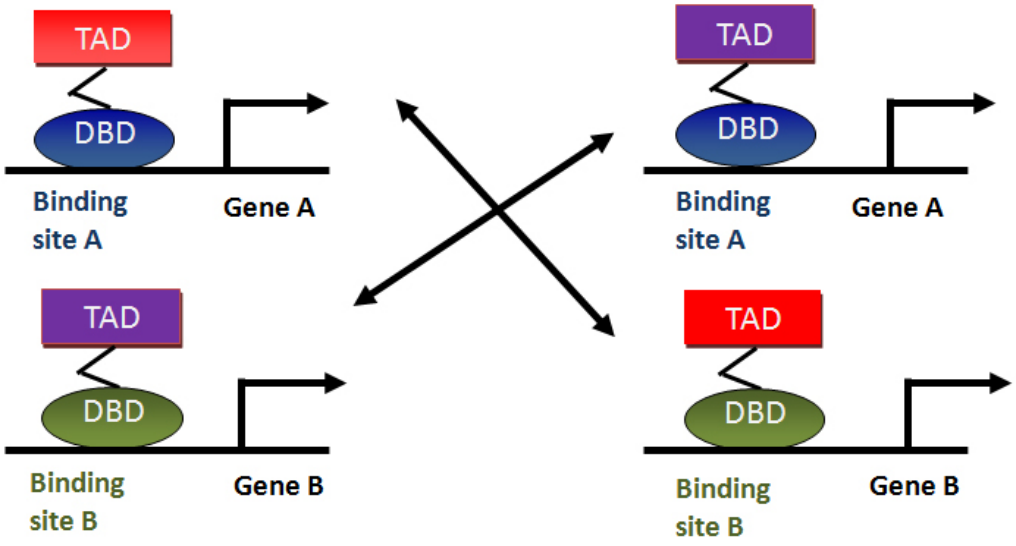


Figure 1-3: The modules of transcriptional activators in domain swapping experiments, work independently to accomplish transcription. Different TADs when linked to the same DBD were able to activate the same gene.

C.1. DNA binding domain

The DNA binding domain (DBD) is responsible for gene-targeting specificity. It localizes the transcriptional activator by binding to specific DNA sequences. Through a series of favorable electrostatic and/or Van der Waals interactions the DBD makes contact with base pairs and sugar phosphate backbones protruding from recognition motifs within the DNA matrix.[25, 26] Historically DBDs have been amenable to structural characterization by both solution and solid-state techniques, and thus DBDs are more extensively structurally characterized than the TAD. DBDs in transcriptional activators have been shown to use several structural folds including zinc fingers, zinc clusters, and leucine zippers for DNA binding.[27] For example, the alpha helixes of the yeast transcription activator Gcn4 are positioned to interact with adjacent

pseudo-palindromic 4 bp half sites so that it forms a coiled coil homodimer as seen in figure 1-4A.[25, 28] As another example, the $Zn(II)_2Cys_6$ binuclear clusters of the yeast activator Gal4 specifically interact with the highly conserved CGG triplets located at the ends of a 17 bp recognition sequence (Figure 1-4B).[29, 30] The DBDs interact with DNA through highly specific interactions with affinities ranging from low to high nM.[26, 27, 29, 31] Furthermore, techniques like chromatin immunoprecipitation and in vivo footprinting have been used to identify the DNA binding sites for DBDs within the genome.[32-34] Therefore, there is a better understanding of how the DBD of transcription activators function in the context of activator localization to cognate DNA sequences.

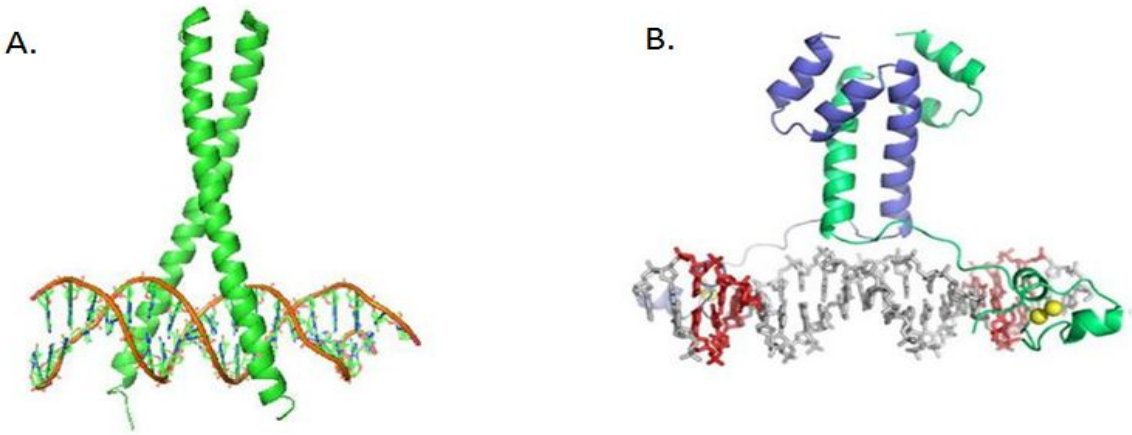


Figure 1-4: Crystal structure of transcriptional activator DBDs in complex with DNA. (A) The basic region leucine zipper (bZIP) DNA-binding motif of Gcn4 bound to DNA as a dimer from PDB 1YSA and (B) the zinc cluster DNA-binding motif of Gal4 bound to DNA as a dimer from PBD 3COQ.

C. 2 Transcriptional activation domain

The transcriptional activation domain (TAD) of an activator modulates the level of transcription of a transcribed gene by facilitating the assembly of the PIC;

however, the mechanism by which the TAD accomplishes this remains ill-defined. It is known that the TAD participates in many interactions throughout the course of transcription (Figure 1-5). Some of these interactions with transcriptional activators yield post-translational modifications such as ubiquitylation, acetylation, phosphorylation, sumoylation, and glycosylation.[35-43] When transcriptional activators are not actively participating in transcription, they may be making regulatory contacts with their masking proteins.

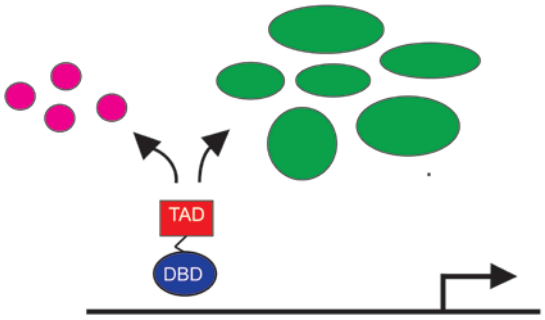


Figure 1-5: A schematic showing the different class of interactions of transcriptional activators. TADs can participate in high affinity interactions with their masking protein (pink circles), these interactions are better characterized than those made with coactivator complexes (green circles) that are moderate in affinity, transient in nature and not well characterized.

These masking interactions are important because the TAD is abrogated from interacting with its partners until it is time for transcription.[44, 45] For example, the TAD of the amphipathic yeast activator Gal4 makes a key masking interaction with the repressor protein Gal80 under conditions where glucose is readily available, Figure 1-6.[46]

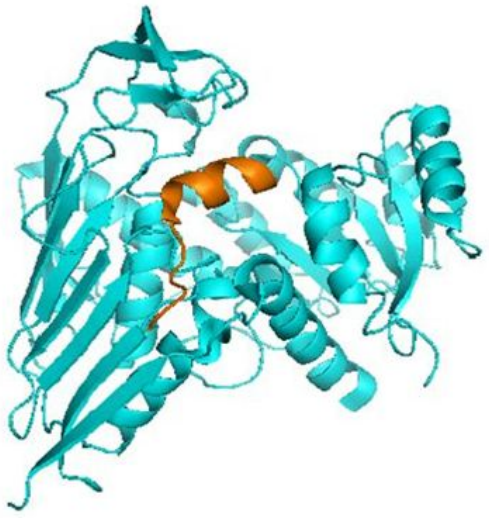


Figure 1-6: Example of a masking interaction between the Gal4 TAD, which is involved with galactose catabolism and its masking protein Gal80. In this crystal structure, the Gal4 TAD (orange) is bound as an α-helix to its masking protein Gal80 (cyan), from PDB 3E1K.

However, in the presence of an inducing sugar such as galactose, a conformational change between Gal4 and Gal80 occurs such that Gal4 is able to up regulate transcription of galactose inducible genes by making contacts with needed partners.[47] Because, the TAD is not as well-characterized as the DNA-binding domain it is primarily characterized by its amino acid content. Such class-characterizations include proline-rich, glutamine-rich and amphipathic, which is the largest and well-studied of all the TAD classes see Figure 1-7.[4, 21, 24]

Figure 1-7: Examples of TAD classifications. Amphipathic TADs like the yeast activator protein (Gal4) contain acidic residues that are interspersed with hydrophobic residues (highlighted in red) Proline rich TADs like the Human CAAT box transcription Factor (hCTF) contains several prolines (highlighted in red) residues. Glutamine rich TADs like the Drosophila Antennapedia protein (Antp) TAD contains several glutamines (highlighted in red) residues.

The amphipathic TADs have hydrophobic amino acids interspersed with polar amino acids. It has been demonstrated experimentally that the acidic residues within amphipathic TADs contributes to the overall acidity needed for initial electrostatic binding to its targets.[48-58] On the other hand, results from mutagenesis studies show that the hydrophobic residues play a crucial role in function. For example, in vitro studies have revealed several hydrophobic residues within the amphipathic VP16 TAD, were shown to be critical for interaction with transcription proteins as well as for function.[48, 53]

Circular dichroism (CD) and NMR studies with isolated TADs show that they are largely unstructured in aqueous solutions at neutral pH.[51, 57, 59-69] However, when TADs are in complex with their binding partners (like masking proteins), in hydrophobic solvent or lower pH conditions they have a propensity to form alpha-helixes or beta-sheets.[64, 65, 69] For example Figure 1-6 already above shows the Gal4 TAD as an alpha-helix in complex with its masking protein Gal80.[47, 70, 71] It is still unclear what common structural motifs amphipathic TADs must utilize for contacting their binding partners. A general understanding is that the coactivator-binding motifs located within activators are transiently-structured.[52, 72-74] Put together, the lack of structural detailed information available for TADs further demonstrates the complexity encountered in the mechanistic understanding of TAD function and further, the discovery of artificial transcriptional modulators.

D. The binding partners of TADs: the great unknown

Amphipathic transcriptional-activation domains are known to exhibit promiscuous binding profiles, a trait that comes as a result of their need to interact with multiple proteins in order to assemble transcriptional complexes at the promoter (Figure 1-2). For example, many proteins have been shown to interact with the Gal4 TAD in vitro.[75] Biochemical and genetic experimental evidence have suggested that Gal4 has several targets including TBP, TFIIB, Med15, Cdk8, Tra1, Taf12, Srb4 and Sug1, among others.[14, 17, 18, 76-84] However, while much is known about interactions between TADs and isolated, purified protein partners in vitro, little is known about the direct, in vivo interactions that occur as an activator recruits large, multi-protein complexes throughout transcription initiation.[4] The next few paragraphs highlight the difficulty encountered so far by scientists while defining targets of transcriptional activators and the many possible targets that have been proposed.

The discovery of the direct binding partners of transcriptional activation domains has been the focus of intense study since the early 1980s. The earliest experiments focused upon components of the transcription machinery - general transcription factors- and RNA pol II as candidates for TAD interactions.[85, 86] TBP, a component of the general transcription factor complex TFIID was shown by in vitro experiments to interact with amphipathic TADs.[81, 87, 88] Soon after, TFIIB recruitment was also shown to be important for activator-dependent transcription.[89, 90] Additional studies produced a growing list of suggested target complexes including TFIIH, TFIIA and RNA pol II. Further evaluation of these

suggested targets in the form of in vitro transcription assays, where these components were supplemented in the experiments, did not stimulate activated transcription and TAD function.[91-94] [91-94] A little later, TBP-associated factors (TAFs), components of the TFIID, were speculated to contact TADs during TFIID recruitment and thus joined the list of putative targets of TADs.[95-97] This belief was questioned in 1996 after reports from both the Struhl and Green groups indicated that TAFs were not essential for transcription in yeast.[98, 99]

Soon after, it was discovered that other protein complexes were needed for activated transcription and possibly could serve as TAD binding partners. The Mediator is a modular complex, consisting of a head, a middle module and a tail. The head interacts with RNA pol II while the tail region has been predicted to interact with TADs.[100] Apart from TFIID, the Mediator complex has been shown to be recruited to gene promoters.[101] For example, Gal1, a tightly regulated promoter and developmentally regulated promoters utilize Mediator recruitment and possibly TFIID to synergistically recruit RNA pol II and the transcriptional machinery.[101] Although the Mediator complex is not utilized at all promoters, it has been shown to be important for activated transcription.[101-106] In fact, a series of biochemical and genetic experiments suggest that the Med17 (Srb4), CDK8 (Srb10), Med15 (Gal11), and Med2 subunits of the Mediator complex interact with TADs (Figure 1-8).[46, 78, 107]

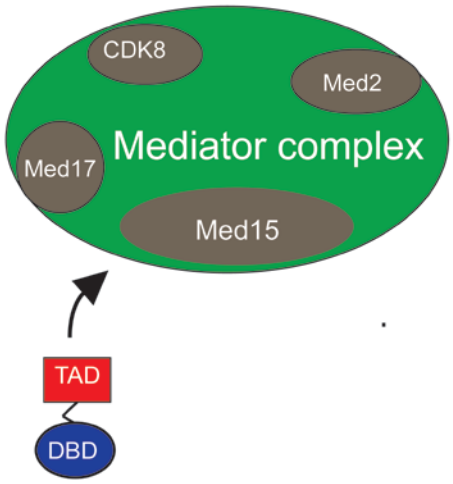


Figure 1-8: A schematic showing some subunits within the Mediator complex that have been named potential binding partners of TADs.

Additionally, chromatin remodeling and modifying complexes have equally been considered as TAD targets because of their role in transcription initiation.[108-112] The ATP-dependent nucleosome remodeling complex, SWI/SNF, is required for function and binding of Gal4 to nucleosomal binding sites in vivo at low affinities.[113] Other TADs like VP16, Gcn4, Swi5, and the mammalian glucocorticoid receptor, through in vitro experiments could interact with ATP-dependent nucleosome remodeling complexes. Another complex, SAGA, is a chromatin modifying complex that contains histone modifying enzymes as well as a subset of TAFs. This complex has been shown through cell-based experiments to be recruited by transcriptional activators.[117, 118] In vivo FRET experiments have shown that within the SAGA complex, the protein Tra1, which is required for Mediator recruitment to the upstream activating sequence (UAS) of the endogenous Gal1 Gene directly, interacts with the Gal4 TAD.[17, 18] This result indicated that Tra1 may be a direct, in vivo target of Gal4.

Both the 26S and the 19S proteasome have been thought to play an important role in regulating transcription. Although the proteasome is not considered a member of the transcription machinery, it is thought to control activator stability and abundance by proteolytically degrading the activator after it initiates transcription.[119] Specifically, within the 19S component, the proteins, Sug1 and Sug2 through in vitro and in vivo experimental evidences are thought to be transcriptional targets, since the proteasome is associated with proteolysis mediated transcriptional regulation.[77, 120-124]

In spite of the success made so far, much is yet to be determined about the direct targets of activators. Decades of studies have very successfully identified the complexes that are recruited by activators, but much is still unknown about their direct binding partners within those complexes. To fully appreciate transcriptional mechanisms, there has to be a better understanding of the direct PPIs needed for transcriptional regulation. Importantly, this understanding is instrumental to therapeutic intervention for transcription related diseases.

E. Transcription-related diseases

Numerous diseases are associated with malfunctioning transcriptional regulators, resulting in much interest in identifying molecules that can correct misregulated pathways.[45, 125, 126] A number of activator●co-activator interactions are implicated in human diseases and are thus potential targets for therapeutic intervention.[127, 128] As an example, the oncogenic tyrosine kinase HER2 is regulated via the activator-coactivator interaction between ESX and SUR2, and

disruption of this interaction with a small molecule blocks proliferation of cells exhibiting this phenotype.[127] However, the broader discovery of inhibitors of transcriptional protein interactions has been problematic, largely due to the lack of knowledge about which contacts TADs make with coactivators as they are turning a gene on.[129-134]

The modular architecture of transcriptional activators simplifies the general strategy for designing artificial replacements. One can imagine, for example, creating an artificial transcriptional activator by replacing one or both of the native DNA-binding and transcriptional activation domains with non-natural counterparts, a modular replacement strategy that has been successfully used in a number of applications. For example the nonnatural zinc finger protein (ZFP), which serves as a modular replacement for DBD, has been successfully engineered to recognize specific DNA sequences required for gene activation (Figure 1-9). In one particular study, the target gene for which the non-natural zinc finger protein was designed encoded for vascular endothelial growth factor A (VEGF-A), which when aberrantly expressed has been linked to tumorgenesis, diabetic retinopathy, ischemic heart and limb disease.[135] The designed ZFP, which was engineered to bind a specific DNA sequence within the VEGF-A promoter, was fused to either VP16 or p53 transcriptional activator to produce expression of VEGF-A.[135] In another study, ZFP dependent transcriptional activation was used to induce angiogenesis in infected mice.[136] Currently, ZFPs are in clinical trials and are being evaluated for efficacy in treating diabetic related diseases, neuropathy, cancer, and HIV infection.[137-139] Other artificial

DBD replacements to include the polyamides, triplex forming oligonucleotides (TFOs) and peptide nucleic acids (PNAs) have also been designed and utilized in in vivo and in vitro experiments.[140-144]

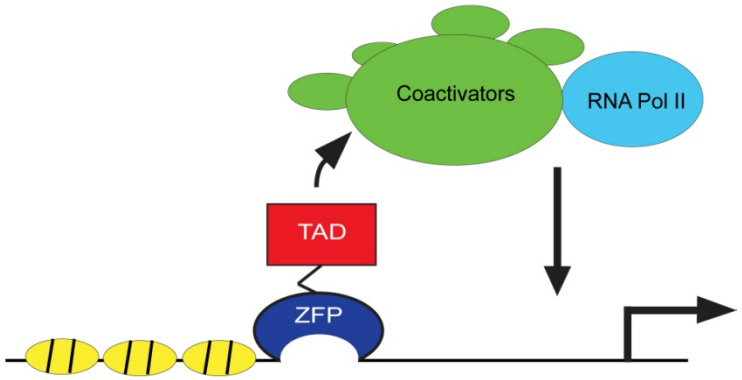


Figure 1-9: A schematic of zinc finger protein as a modular replacement for DBD. ZFP binds a specific DNA sequence where the tethered VP16 TAD recruits coactivators and RNA polymerase II, leading to increased levels of gene expression.

Similar to DNA binding domains, transcriptional activation domains can also serve as replacement therapies for inducing transcription. For example, the TAD VP16 when fused to REST DBD was able to activate transcription by offsetting the repressor element-1silencing transcription/neuron-restrictive silencer factor (REST/NRSF) repression. In abnormal cerebellum, REST/NRSF inhibits transcription by repressing a set of neuronal differentiation genes by binding the same promoter regions as transcriptional activators, thus blocking activator occupancy.[139, 145] In this experiment, a REST DBD-VP16 TAD fusion protein was used to de-repress REST/NRSF regulated genes in mice with pre-formed medulloblastoma tumors. Results showed that not only were the terminal neuronal differentiation genes being expressed, but cells expressing REST-VP16

displayed signs of apoptosis, and importantly, there was evident reduction in the tumor size after four weeks (Figure 1-10).[139]

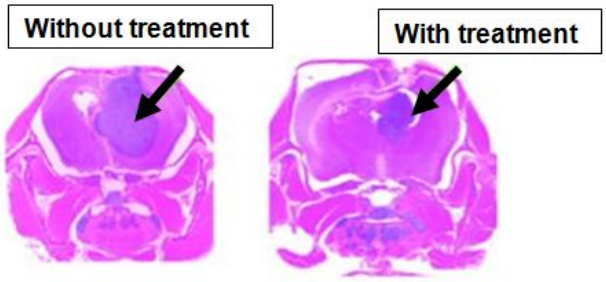


Figure 1-10: 4 weeks of treatment with adenovirus containing REST-VP16 fusion protein resulted in a reduction in tumor size compared to the control.

Although therapeutic TAD replacements are desired, there are a limited number of non-protein-based TADs. A greater number of the commonly used TADs are short sequences from their endogenous counterparts. A major limitation when using protein-based TADs is proteolysis, this lack of stability results to reduced cellular delivery. Therefore, a reasonable approach will be to generate synthetic TADs like small molecules that do not have the limitations mentioned above. The problem with using small molecules is that they cannot easily replicates PPIs, because these types of interactions entail large surface contacts, which are required for specificity.[146-151] The fact that irrespective of the extensive research employed thus far to generate synthetic TADs, there are only a hand full of examples see Figure 1-11, further underscores the difficulty encountered in generating synthetic TADs. In fact, the first synthetic small molecule TAD (Figure 1-11A) to have come so far to having a close binding

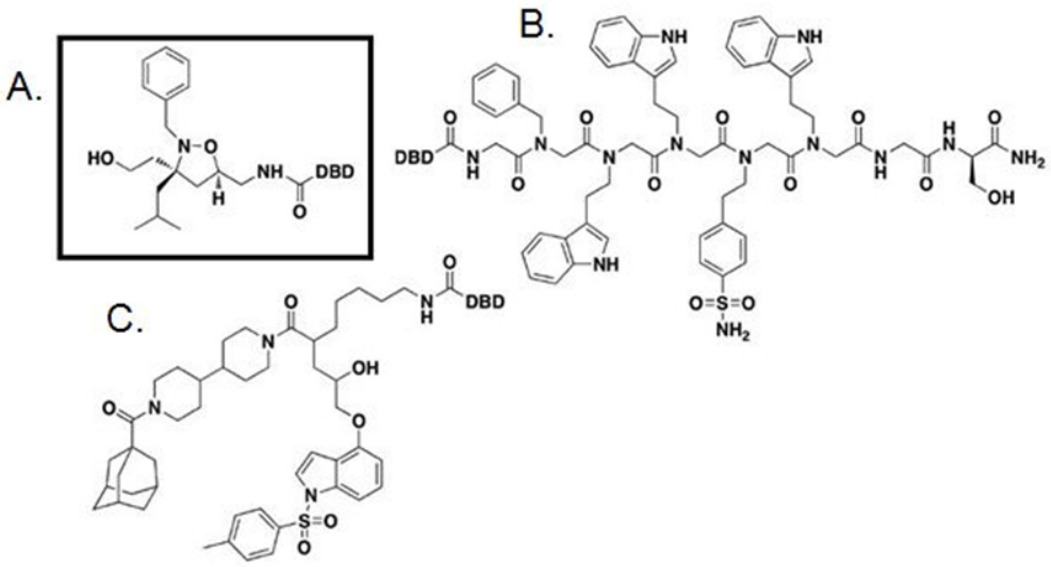


Figure 1-11: Artificial small molecule TADs. (A) Isoxazolidine, made by Mapp and coworkers, (B) a peptoid made by Kodadek and coworkers and (C) Wrenchnolol, made by Uesugi and coworkers.

fingerprint with natural TADs, and have been shown to exhibit close function in in vitro and in vivo experiments was reported by the Mapp laboratory.[152, 153] At this rate, better understanding of transcriptional PPIs with emphasis on the direct-partner interactions, will help to better design small molecules that can suffice for larger protein surface contacts and further serve as both mechanistic tools and prototype transcription-based therapeutic agents.

Furthermore, the array of transcriptional proteins available for possible binding, and the possibility that a protein may bind to more than one protein to facilitate assembly of a large (>50 proteins) complex needed for transcription has made the understanding of activator-target interactions difficult. Without knowing exactly which targets interact with a specific activator for transcriptional activation to occur, it is difficult to design screens for small molecule modulators that exhibit

the needed multi-partner binding profile. As discussed in detail in the next section, a major hindrance to the identification and characterization of direct PPIs involved in transcription is that the methods that have been employed to study these types of interactions have shortfalls. A better understanding of activator-target interactions will help to forge towards the development of efficient transcriptional modulators particularly those that can function as inhibitors.

F. Strategies for capturing protein-protein interactions

To efficiently understand transcriptional PPIs, one has to ensure that the methodologies employed for studying these types of interactions are appropriate. Over three decades, researchers have used traditional biochemical and genetic techniques to study protein-protein interactions required for transcription. Using these methodologies and techniques has revealed several possible targets that interact with transcription activators. Many questions are still unanswered especially, what are the relevant direct binding partners of transcriptional activators? Due to the complex nature of transcription, results from methodologies like traditional genetic deletions are not readily understood and interpreted. This is because of the pleiotropic effects observed when critical proteins and/or sequences of critical transcription proteins are altered.[154, 155] Transcriptional machinery proteins like Med15 and CBP can employ the same binding surface to interact with multiple activators and other proteins.[78, 82, 156-158] This functional redundancy creates the possibility for the disruption of an interaction(s) to be compensated for, to some extent, by other interactions that may or may not be physiologically relevant. In like manner TAD sequences

contain overlapping binding sites. For example, the VP16 TAD sequence can interact with a number of transcription protein targets, as well as the ubiquitnating E3 ligase complex.[159-161] Therefore, deletion or mutagenesis within the VP16 TAD can affect more than one single binding event making results from such experiments hard to interpret.

Biochemical techniques have revealed a significant amount of information regarding transcriptional network facilitators; however, little progress has been achieved in discovering the direct binding partners of transcriptional activators. For example, chromatin immunoprecipitation (ChIP) is an extensively used method that uses formaldehyde to crosslink proteins to DNA, thus revealing a snapshot of all the protein complexes interacting at a particular DNA location. ChIP has been a fundamentally important technique in showing what coactivator complexes co-localize with activators at a given promoter, and additionally, delineating the general time of co-localization. In two examples, in vivo ChIP assays were used to reveal that SAGA was localized to the Gal4 upstream activation sequence (UAS), and that the SAGA and ADA complexes were targeted to the promoter of a Gcn4 regulated gene.[18, 159, 162] Although ChIP has been useful in identifying the proteins that associate with activators at the promoter, one major limitation of this method is that is does not readily distinguish between proteins that are directly contacting one another from those that are interacting indirectly. In addition, ChIP cannot provide information on the spatial arrangement of coactivators or the relative distances between associated proteins.

While ChIP cannot be used to determine the relative distances between proteins within a complex or between complexes, fluorescence resonance energy transfer (FRET) has proven to be a useful tool for studying molecular dynamics such as protein-protein interactions and protein conformational changes. When used to study two proteins that are in close proximity, the first and second proteins are labeled with donor and acceptor fluorophores, respectively. If these proteins come in close proximity to one another, the donor's excitation energy is transferred to the acceptor's fluorophore so that the acceptor's emission is predominantly observed. Using in vivo FRET, Bhaumik et. al. found that Tra1, a subunit of the SAGA complex interacts with Gal4 in yeast.[17] Even though FRET gives spatial information about coactivators and TADs in comparison to ChIP, its limitations include a high probability of false negatives as well as false positives. Fluorophores that are poorly positioned in such a manner that they are incapable of interacting gives rise to false negatives. Another complexity is that TAD structures are not well defined, so, determining where to place fluorophores when studying transcription protein interactions is even more difficult. False positives occur when the two proteins containing fluorophores are brought together by a bridging protein even though they are not directly interacting. False positives can easily be a problem in a transcriptional context, since several protein complexes continually interacts with one another to achieve transcription.

Another technique, in vitro photo-crosslinking has two major advantages. Not only can it be used to identify the direct binding partners of a TAD but also it

can specify the binding site, and even, amino acids involved with binding when paired with other methods. This experiment requires that a label which has crosslinking capabilities is placed on the protein of choice like a TAD, and upon UV activation the protein forms covalent bonds with any proteins that are in close proximity. Using this technique in the presence of purified PIC in vitro, the Hahn group identified several potential direct binding partners of the yeast activator Gcn4 to include Tra1, a component of SAGA, Med15, a component of Mediator, and Taf12, part of the SAGA and TFIID complexes.[131] In this experiment, Gcn4 was labeled with the aryl azido photo-crosslinker ^{125}I-PEAS, which interacts with proteins through a cysteine mediated disulfide bond formation.[131] A similar result was reported by the same group for yeast Gal4 TAD, which was shown to target Tra1, Taf12 and Med15 also within the purified PIC in vitro conditions.[163]

Similarly, other crosslinking agents have been used to investigate the binding behavior of isolated TADs. Another, aryl azide crosslinker, Sulfo-SBED, has been utilized to study the direct in vitro binding partners of the transcriptional activators VP16 and Gcn4. Sulfo- SBED also contains a biotin label, which when transferred to the target is useful for Western blot analysis. After crosslinking, the biotin handle is transferred to the target through a DTT mediated disulfide bond disruption. Using this crosslinker, two TADs VP16 and Gcn4 were thought to interact with three subunits of the Swi/Snf complex namely, Swi1, Snf2 and Snf5.[164] Although in vitro photo crosslinking experiments are better poised to present direct binding partners of TADs, these experiments have flaws as well.

One flaw of in vitro crosslinking experiments is that the linker lengths of photo crosslinkers utilized for in vitro photo crosslinking increases the number of false positives associated with these experiments. For example, ^{125}I-PEAS, has a linker length that makes it possible for crosslinking to extend 14 Å beyond the alpha carbon of the cysteine.[165] Therefore, increasing its ability to capture indirect binding partners that may be serving as bridge proteins to the actual relevant direct partner, leads to false positives. Another limitation is that some photo cross-linkers can nonspecifically label the TAD. For example, Sulfo-SBED carries out nonspecifically labeling of lysines, this again, increases the probability of false positives as well as false-negatives, which further complicates interpretations from such experiments. Therefore, there is a pressing need for a technique (discussed in chapter 2) that satisfies the limitations presented in others.

G. Thesis overview

Chapter 1 presents a brief description of transcription, the numerous research endeavors of scientists seeking to understand the mechanisms involved with transcription, the modular nature of transcriptional activators, the types of interactions and interacting complexes associated with transcriptional activators, and the methodologies that have been used to identify these activator targets, and finally the health benefits associated with understanding transcription activator interactions with strides made so far towards this end. Chapter 2 starts out making the case that there is an immense need for better methodologies that has to be employed to better understand transcription PPIs. Furthermore,

chapter 2 describes the development, optimization, implementation, and validation of an in vivo photo-crosslinking strategy that can be utilized for studying transcriptional PPIs. Chapter 3 further pushes the envelope by determining whether in vivo photo-crosslinking strategy can be employed to capture weak affinity targets as well as endogenous targets. Chapter 4 concludes by proposing some future experiments that can be carried out to accurately characterize activator-target binding interfaces. The work presented in this dissertation is the much-needed step that has to be accomplished before the effective design and generation of small molecule- screens appropriate for the search of transcriptional PPI modulators.

H. References

1. Chen, X., S.T. Cheung, S. So, S.T. Fan, C. Barry, J. Higgins, K.M. Lai, J. Ji, S. Dudoit, I.O. Ng, M. Van De Rijn, D. Botstein, and P.O. Brown, *Gene expression patterns in human liver cancers.* Molecular Biology of the Cell, 2002. **13**(6): p. 1929-39.
2. Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L.

Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H.M. Yang, J. Yu, J. Wang, G.Y. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S.Z. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H.Q. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G.R. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W.H. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J.R. Schultz, G. Slater, A.F.A. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan and I.H.G.S. Conso, *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

3. Perou, C.M., T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, and D. Botstein, *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
4. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
5. Ptashne, M., *How eukaryotic transcriptional activators work.* Nature, 1988. **335**(6192): p. 683-9.
6. Ptashne, M., *1997 Albert Lasker Award for Basic Medical Research. Control of gene transcription: an outline.* Nat Med, 1997. **3**(10): p. 1069-72.
7. Ptashne, M. and A. Gann, *Transcriptional activation by recruitment.* Nature, 1997. **386**(6625): p. 569-77.
8. Paule, M.R., White, R. J., *Survey and summary: transcription by RNA polymerases I and III.* Nucleic Acids Res, 2000. **28**(6): p. 1283-1298.
9. Geiduschek, E.P. and G.A. Kassavetis, *The RNA polymerase III transcription apparatus.* J Mol Biol, 2001. **310**(1): p. 1-26.
10. Cramer, P., *Recent structural studies of RNA polymerases II and III.* Biochemical Society Transactions, 2006. **34**: p. 1058-1061.

11. Thomas, M.C. and C.M. Chiang, *The general transcription machinery and general cofactors.* Crit Rev Biochem Mol Biol, 2006. **41**(3): p. 105-78.
12. Martel, L.S., H.J. Brown, and A.J. Berk, *Evidence that TAF-TATA box-binding protein interactions are required for activated transcription in mammalian cells.* Mol Cell Bio, 2002. **22**: p. 2788-2798.
13. Roeder, R.G., *The role of general initiation factors in transcription by RNA polymerase II.* Trends Biochem Sci, 1996. **21**(9): p. 327-35.
14. Bryant, G.O. and M. Ptashne, *Independent recruitment in vivo by Gal4 of two complexes required for transcription.* Mol Cell, 2003. **11**(5): p. 1301-9.
15. Cosma, M.P., T. Tanaka, and K. Nasmyth, *Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter.* Cell, 1999. **97**(3): p. 299-311.
16. Bhaumik, S.R., *Analysis of in vivo targets of transcriptional activators by fluorescence resonance energy transfer.* Methods, 2006. **40**(4): p. 353-9.
17. Bhaumik, S.R., T. Raha, D.P. Aiello, and M.R. Green, *In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer.* Genes Dev, 2004. **18**(3): p. 333-43.
18. Bhaumik, S.R. and M.R. Green, *SAGA is an essential in vivo target of the yeast acidic activator Gal4p.* Genes Dev, 2001. **15**: p. 1935-1945.
19. Bhaumik, S.R. and M.R. Green, *Interaction of Gal4p with components of transcription machinery in vivo.* Methods Enzymol, 2003. **370**: p. 445-54.
20. Ma, J. and M. Ptashne, *Deletion analysis of GAL4 defines two transcriptional activating segments.* Cell, 1987. **48**(5): p. 847-53.
21. Ptashne, M., *Genes and Signal.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
22. Brent, R. and M. Ptashne, *A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor.* Cell, 1985. **43**(3 Pt 2): p. 729-36.
23. Keegan, L., G. Gill, and M. Ptashne, *Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein.* Science, 1986. **231**(4739): p. 699-704.
24. Ansari, A.Z. and A.K. Mapp, *Modular design of artificial transcription factors.* Curr Opin Chem Biol, 2002. **6**(6): p. 765-72.
25. Keller, W., P. Konig, and T.J. Richmond, *Crystal structure of a bZIP/DNA complex at 2.2 A: determinants of DNA specific recognition.* J Mol Biol, 1995. **254**(4): p. 657-67.
26. Luscombe, N.M., Austin, S. E., Berman, H. M., Thornton, J. M., *An overview of the structures of protein-DNA complexes.* Genome Biol, 2000. **1**(1): p. REVIEWS001.
27. Garvie, C.W. and C. Wolberger, *Recognition of specific DNA sequences.* Mol Cell, 2001. **8**(5): p. 937-46.
28. Ellenberger, T.E., C.J. Brandl, K. Struhl, and S.C. Harrison, *The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex.* Cell, 1992. **71**(7): p. 1223-37.

29. Marmorstein, R., M. Carey, M. Ptashne, and S.C. Harrison, *DNA recognition by GAL4: structure of a protein-DNA complex.* Nature, 1992. **356**(6368): p. 408-14.
30. Hong, M., M.X. Fitzgerald, S. Harper, C. Luo, D.W. Speicher, and R. Marmorstein, *Structural basis for dimerization in DNA recognition by Gal4.* Structure, 2008. **16**(7): p. 1019-26.
31. Littlefield, O. and H.C. Nelson, *A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocrystal.* Nat Struct Biol, 1999. **6**(5): p. 464-70.
32. Jothi, R., S. Cuddapah, A. Barski, K. Cui, and K. Zhao, *Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data.* Nucleic Acids Res, 2008. **36**(16): p. 5221-31.
33. Molina, M.P., C. Cain, and J. Bargonetti, *In vivo footprinting and DNA affinity chromatography for analysis of p53 DNA binding ability.* Methods Mol Biol, 2003. **234**: p. 151-70.
34. Johnson, D.S., A. Mortazavi, R.M. Myers, and B. Wold, *Genome-wide mapping of in vivo protein-DNA interactions.* Science, 2007. **316**(5830): p. 1497-502.
35. Ansari, A.Z., S.S. Koh, Z. Zaman, C. Bongards, N. Lehming, R.A. Young, and M. Ptashne, *Transcriptional activating regions target a cyclin-dependent kinase.* Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14706-9.
36. Bouwman, P. and S. Philipsen, *Regulation of the activity of Sp1-related transcription factors.* Mol Cell Endocrinol, 2002. **195**(1-2): p. 27-38.
37. Chrivia, J.C., R.P. Kwok, N. Lamb, M. Hagiwara, M.R. Montminy, and R.H. Goodman, *Phosphorylated CREB binds specifically to the nuclear protein CBP.* Nature, 1993. **365**(6449): p. 855-9.
38. Jackson, S.P. and R. Tjian, *O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation.* Cell, 1988. **55**(1): p. 125-33.
39. Khidekel, N. and L.C. Hsieh-Wilson, *A 'molecular switchboard'--covalent modifications to proteins and their impact on transcription.* Org Biomol Chem, 2004. **2**(1): p. 1-7.
40. Lamarre-Vincent, N. and L.C. Hsieh-Wilson, *Dynamic glycosylation of the transcription factor CREB: a potential role in gene regulation.* J Am Chem Soc, 2003. **125**(22): p. 6612-3.
41. Lamph, W.W., V.J. Dwarki, R. Ofir, M. Montminy, and I.M. Verma, *Negative and positive regulation by transcription factor cAMP response element-binding protein is modulated by phosphorylation.* Proc Natl Acad Sci U S A, 1990. **87**(11): p. 4320-4.
42. Sadowski, I., D. Niedbala, K. Wood, and M. Ptashne, *GAL4 is phosphorylated as a consequence of transcriptional activation.* Proc Natl Acad Sci U S A, 1991. **88**(23): p. 10510-4.
43. Ferdous, A., D. Sikder, T. Gillette, K. Nalley, T. Kodadek, and S.A. Johnston, *The role of the proteasomal ATPases and activator monoubiquitylation in regulating Gal4 binding to promoters.* Genes Dev, 2007. **21**(1): p. 112-23.

44. des Etages, S.A., D.A. Falvey, R.J. Reece, and M.C. Brandriss, *Functional analysis of the PUT3 transcriptional activator of the proline utilization pathway in Saccharomyces cerevisiae.* Genetics, 1996. **142**(4): p. 1069-82.
45. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing, the p53 network.* Nature, 2000. **408**: p. 307-310.
46. Ansari, A.Z., R.J. Reece, and M. Ptashne, *A transcriptional activating region with two contrasting modes of protein interaction.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13543-8.
47. Thoden, J.B., C.A. Sellick, R.J. Reece, and H.M. Holden, *Understanding a transcriptional paradigm at the molecular level. The structure of yeast Gal80p.* J Biol Chem, 2007. **282**(3): p. 1534-8.
48. Sullivan, S.M., P.J. Horn, V.A. Olson, A.H. Koop, W. Niu, R.H. Ebright, and S.J. Triezenberg, *Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus.* Nucleic Acids Res, 1998. **26**(19): p. 4487-96.
49. Callus, B.A. and B. Mathey-Prevot, *Hydrophobic residues Phe751 and Leu753 are essential for STAT5 transcriptional activity.* J Biol Chem, 2000. **275**(22): p. 16954-62.
50. Chen, R.H., S. Fields, and J.S. Lipsick, *Dissociation of transcriptional activation and oncogenic transformation by v-Myb.* Oncogene, 1995. **11**(9): p. 1771-9.
51. Leuther, K.K., J.M. Salmeron, and S.A. Johnston, *Genetic evidence that an activation domain of GAL4 does not require acidity and may form a beta sheet.* Cell, 1993. **72**(4): p. 575-85.
52. Lin, J., J. Chen, B. Elenbaas, and A.J. Levine, *Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein.* Genes Dev, 1994. **8**(10): p. 1235-46.
53. Regier, J.L., F. Shen, and S.J. Triezenberg, *Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator.* Proc Natl Acad Sci U S A, 1993. **90**(3): p. 883-7.
54. Sainz, M.B., S.A. Goff, and V.L. Chandler, *Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo.* Mol Cell Biol, 1997. **17**(1): p. 115-22.
55. Drysdale, C.M., E. Duenas, B.M. Jackson, U. Reusser, G.H. Braus, and A.G. Hinnebusch, *The transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids.* Mol Cell Biol, 1995. **15**(3): p. 1220-33.
56. Hermann, S., K.D. Berndt, and A.P. Wright, *How transcriptional activators bind target proteins.* J Biol Chem, 2001. **276**(43): p. 40127-32.
57. Ferreira, M.E., S. Hermann, P. Prochasson, J.L. Workman, K.D. Berndt, and A.P. Wright, *Mechanism of transcription factor recruitment by acidic activators.* J Biol Chem, 2005. **280**(23): p. 21779-84.

58. Jonker, H.R., R.W. Wechselberger, R. Boelens, G.E. Folkers, and R. Kaptein, *Structural properties of the promiscuous VP16 activation domain.* Biochemistry, 2005. **44**(3): p. 827-39.
59. Biddick, R. and E.T. Young, *The disorderly study of ordered recruitment.* Yeast, 2009. **26**(4): p. 205-20.
60. Dyson, H.J. and P.E. Wright, *Coupling of folding and binding for unstructured proteins.* Curr Opin Struct Biol, 2002. **12**(1): p. 54-60.
61. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
62. Wells, M., H. Tidow, T.J. Rutherford, P. Markwick, M.R. Jensen, E. Mylonas, D.I. Svergun, M. Blackledge, and A.R. Fersht, *Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5762-7.
63. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Curr Opin Struct Biol, 2009. **19**(1): p. 31-8.
64. Van Hoy, M., K.K. Leuther, T. Kodadek, and S.A. Johnston, *The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets.* Cell, 1993. **72**(4): p. 587-94.
65. Schmitz, M.L., M.A. dos Santos Silva, H. Altmann, M. Czisch, T.A. Holak, and P.A. Baeuerle, *Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation.* J Biol Chem, 1994. **269**(41): p. 25613-20.
66. Dahlman-Wright, K., H. Baumann, I.J. McEwan, T. Almlof, A.P. Wright, J.A. Gustafsson, and T. Hard, *Structural characterization of a minimal functional transactivation domain from the human glucocorticoid receptor.* Proc Natl Acad Sci U S A, 1995. **92**(5): p. 1699-703.
67. Shen, F., S.J. Triezenberg, P. Hensley, D. Porter, and J.R. Knutson, *Critical amino acids in the transcriptional activation domain of the herpesvirus protein VP16 are solvent-exposed in highly mobile protein segments. An intrinsic fluorescence study.* J Biol Chem, 1996. **271**(9): p. 4819-26.
68. Warnmark, A., A. Wikstrom, A.P. Wright, J.A. Gustafsson, and T. Hard, *The N-terminal regions of estrogen receptor alpha and beta are unstructured in vitro and show different TBP binding properties.* J Biol Chem, 2001. **276**(49): p. 45939-44.
69. Cress, W.D. and S.J. Triezenberg, *Critical structural elements of the VP16 transcriptional activation domain.* Science, 1991. **251**(4989): p. 87-90.
70. Thoden, J.B., L.A. Ryan, R.J. Reece, and H.M. Holden, *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. **283**(44): p. 30266-72.
71. Kussie, P.H., S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A.J. Levine, and N.P. Pavletich, *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. **274**(5289): p. 948-53.

72. Fields, S. and S.K. Jang, *Presence of a potent transcription activating sequence in the p53 protein.* Science, 1990. **249**(4972): p. 1046-9.
73. Dawson, R., L. Muller, A. Dehner, C. Klein, H. Kessler, and J. Buchner, *The N-terminal domain of p53 is natively unfolded.* J Mol Biol, 2003. **332**(5): p. 1131-41.
74. Teufel, D.P., S.M. Freund, M. Bycroft, and A.R. Fersht, *Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53.* Proc Natl Acad Sci U S A, 2007. **104**(17): p. 7009-14.
75. Traven, A., B. Jelicic, and M. Sopta, *Yeast Gal4: a transcriptional paradigm revisited.* EMBO Rep, 2006. **7**(5): p. 496-9.
76. 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.
77. Chang, C., F. Gonzalez, B. Rothermel, L. Sun, S.A. Johnston, and T. Kodadek, *The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro.* J Biol Chem, 2001. **276**(33): p. 30956-63.
78. Jeong, C.J., S.H. Yang, Y. Xie, L. Zhang, S.A. Johnston, and T. Kodadek, *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. **40**(31): p. 9421-7.
79. Klein, J., M. Nolden, S.L. Sanders, J. Kirchner, P.A. Weil, and K. Melcher, *Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA.* J Biol Chem, 2003. **278**(9): p. 6779-86.
80. Koh, S.S., A.Z. Ansari, M. Ptashne, and R.A. Young, *An activator target in the RNA polymerase II holoenzyme.* Mol Cell, 1998. **1**(6): p. 895-904.
81. Li, X.Y., A. Virbasius, X. Zhu, and M.R. Green, *Enhancement of TBP binding by activators and general transcription factors.* Nature, 1999. **399**(6736): p. 605-9.
82. Melcher, K. and S.A. Johnston, *GAL4 interacts with TATA-binding protein and coactivators.* Mol Cell Biol, 1995. **15**(5): p. 2839-48.
83. 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.
84. Xie, Y., C. Denison, S.H. Yang, D.A. Fancy, and T. Kodadek, *Biochemical characterization of the TATA-binding protein-Gal4 activation domain complex.* J Biol Chem, 2000. **275**(41): p. 31914-20.
85. Horikoshi, M., T. Hai, Y.S. Lin, M.R. Green, and R.G. Roeder, *Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex.* Cell, 1988. **54**(7): p. 1033-42.
86. 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.
87. Nedialkov, Y.A. and S.J. Triezenberg, *Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region.* Arch Biochem Biophys, 2004. **425**(1): p. 77-86.

88. Wu, Y., R.J. Reece, and M. Ptashne, *Quantitation of putative activator-target affinities predicts transcriptional activating potentials.* Embo J, 1996. **15**(15): p. 3951-63.
89. Lin, Y.S., I. Ha, E. Maldonado, D. Reinberg, and M.R. Green, *Binding of general transcription factor TFIIB to an acidic activating region.* Nature, 1991. **353**(6344): p. 569-71.
90. Choy, B. and M.R. Green, *Eukaryotic activators function during multiple steps of preinitiation complex assembly.* Nature, 1993. **366**(6455): p. 531-6.
91. Stargell, L.A. and K. Struhl, *The TBP-TFIIA interaction in the response to acidic activators in vivo.* Science, 1995. **269**(5220): p. 75-8.
92. Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J.L. Regier, S.J. Triezenberg, D. Reinberg, O. Flores, C.J. Ingles, and et al., *Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53.* Mol Cell Biol, 1994. **14**(10): p. 7013-24.
93. Ozer, J., A.H. Bolden, and P.M. Lieberman, *Transcription factor IIA mutations show activator-specific defects and reveal a IIA function distinct from stimulation of TBP-DNA binding.* J Biol Chem, 1996. **271**(19): p. 11182-90.
94. Tan, Q., K.L. Linask, R.H. Ebright, and N.A. Woychik, *Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria.* Genes Dev, 2000. **14**(3): p. 339-48.
95. Burley, S.K. and R.G. Roeder, *Biochemistry and structural biology of transcription factor IID (TFIID).* Annu Rev Biochem, 1996. **65**: p. 769-99.
96. Dynlacht, B.D., T. Hoey, and R. Tjian, *Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation.* Cell, 1991. **66**(3): p. 563-76.
97. Tjian, R., *Molecular machines that control genes.* Sci Am, 1995. **272**(2): p. 54-61.
98. Moqtaderi, Z., Y. Bai, D. Poon, P.A. Weil, and K. Struhl, *TBP-associated factors are not generally required for transcriptional activation in yeast.* Nature, 1996. **383**(6596): p. 188-91.
99. Walker, S.S., J.C. Reese, L.M. Apone, and M.R. Green, *Transcription activation in cells lacking TAFIIS.* Nature, 1996. **383**(6596): p. 185-8.
100. Asturias, F.J., Y.W. Jiang, L.C. Myers, C.M. Gustafsson, and R.D. Kornberg, *Conserved structures of mediator and RNA polymerase II holoenzyme.* Science, 1999. **283**(5404): p. 985-7.
101. Fan, X., D.M. Chou, and K. Struhl, *Activator-specific recruitment of Mediator in vivo.* Nat Struct Mol Biol, 2006. **13**(2): p. 117-20.
102. Koleske, A.J. and R.A. Young, *An RNA polymerase II holoenzyme responsive to activators.* Nature, 1994. **368**(6470): p. 466-9.
103. Flanagan, P.M., R.J. Kelleher, 3rd, M.H. Sayre, H. Tschochner, and R.D. Kornberg, *A mediator required for activation of RNA polymerase II transcription in vitro.* Nature, 1991. **350**(6317): p. 436-8.

104. Kim, Y.J., S. Bjorklund, Y. Li, M.H. Sayre, and R.D. Kornberg, *A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II.* Cell, 1994. **77**(4): p. 599-608.
105. Govind, C.K., S. Yoon, H. Qiu, S. Govind, and A.G. Hinnebusch, *Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo.* Mol Cell Biol, 2005. **25**(13): p. 5626-38.
106. Park, J.M., H.S. Kim, S.J. Han, M.S. Hwang, Y.C. Lee, and Y.J. Kim, *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. **20**(23): p. 8709-19.
107. Myers, L.C., C.M. Gustafsson, K.C. Hayashibara, P.O. Brown, and R.D. Kornberg, *Mediator protein mutations that selectively abolish activated transcription.* Proc Natl Acad Sci U S A, 1999. **96**(1): p. 67-72.
108. Fry, C.J. and C.L. Peterson, *Chromatin remodeling enzymes: who's on first?* Curr Biol, 2001. **11**(5): p. R185-97.
109. Kadonaga, J.T., *Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines.* Cell, 1998. **92**(3): p. 307-13.
110. Workman, J.L. and R.E. Kingston, *Alteration of nucleosome structure as a mechanism of transcriptional regulation.* Annu Rev Biochem, 1998. **67**: p. 545-79.
111. Tsukiyama, T. and C. Wu, *Chromatin remodeling and transcription.* Curr Opin Genet Dev, 1997. **7**(2): p. 182-91.
112. Carpenter, A.E., S. Memedula, M.J. Plutz, and A.S. Belmont, *Common effects of acidic activators on large-scale chromatin structure and transcription.* Mol Cell Biol, 2005. **25**(3): p. 958-68.
113. Burns, L.G. and C.L. Peterson, *The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo.* Mol Cell Biol, 1997. **17**(8): p. 4811-9.
114. Neely, K.E., A.H. Hassan, A.E. Wallberg, D.J. Steger, B.R. Cairns, A.P. Wright, and J.L. Workman, *Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays.* Mol Cell, 1999. **4**(4): p. 649-55.
115. Yudkovsky, N., C. Logie, S. Hahn, and C.L. Peterson, *Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators.* Genes Dev, 1999. **13**(18): p. 2369-74.
116. Wallberg, A.E., K.E. Neely, J.A. Gustafsson, J.L. Workman, A.P. Wright, and P.A. Grant, *Histone acetyltransferase complexes can mediate transcriptional activation by the major glucocorticoid receptor activation domain.* Mol Cell Biol, 1999. **19**(9): p. 5952-9.
117. Kwok, R.P., J.R. Lundblad, J.C. Chrivia, J.P. Richards, H.P. Bachinger, R.G. Brennan, S.G. Roberts, M.R. Green, and R.H. Goodman, *Nuclear protein CBP is a coactivator for the transcription factor CREB.* Nature, 1994. **370**(6486): p. 223-6.

118. Grant, P.A., D.E. Sterner, L.J. Duggan, J.L. Workman, and S.L. Berger, *The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes.* Trends Cell Biol, 1998. **8**(5): p. 193-7.
119. Muratani, M. and W.P. Tansey, *How the ubiquitin-proteasome system controls transcription.* Nat Rev Mol Cell Biol, 2003. **4**(3): p. 192-201.
120. Gonzalez, F., A. Delahodde, T. Kodadek, and S.A. Johnston, *Recruitment of a 19S proteasome subcomplex to an activated promoter.* Science, 2002. **296**(5567): p. 548-50.
121. Ferdous, A., F. Gonzalez, L. Sun, T. Kodadek, and S.A. Johnston, *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.
122. Sulahian, R., D. Sikder, S.A. Johnston, and T. Kodadek, *The proteasomal ATPase complex is required for stress-induced transcription in yeast.* Nucleic Acids Res, 2006. **34**(5): p. 1351-7.
123. 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.
124. Reid, G., M.R. Hubner, R. Metivier, H. Brand, S. Denger, D. Manu, J. Beaudouin, J. Ellenberg, and F. Gannon, *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.
125. Pandolfi, P.P., *Transcription therapy for cancer.* Oncogene, 2001. **20**(24): p. 3116-27.
126. Mapp, A.K., *Regulating transcription: a chemical perspective.* Org Biomol Chem, 2003. **1**(13): p. 2217-20.
127. Uesugi, M., *Modulation of gene expression by targeting a protein-protein interaction.* Albany Molecular reserach International, Technical report, 2005. **10**: p. 19-22.
128. Pupa, S.M., E. Tagliabue, S. Menard, and A. Anichini, *HER-2: a biomarker at the crossroads of breast cancer immunotherapy and molecular medicine.* J Cell Physiol, 2005. **205**(1): p. 10-8.
129. Majumdar, C.Y. and A.K. Mapp, *Chemical approaches to transcriptional regulation.* Curr Opin Chem Biol, 2005. **9**: p. 467-477.
130. Xiao, X., P. Yu, H. Lim, and T. Kodadek, *A cell permeable synthetic transcription factor mimic.* Angew Chem Inc Ed, 2007. **46**: p. 2865-2868.
131. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. **18**(3): p. 369-78.
132. Asada, S., Y. Choi, M. Yamada, S.C. Wang, M.C. Hung, J. Qin, and M. Uesugi, *External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12747-52.
133. Shimogawa, H., Y. Kwon, Q. Mao, Y. Kawazoe, Y. Choi, S. Asada, H. Kigoshi, and M. Uesugi, *A wrench-shaped synthetic molecule that*

modulates a transcription factor-coactivator interaction. J Am Chem Soc, 2004. **126**(11): p. 3461-71.
134. Asada, S., Y. Choi, and M. Uesugi, *A gene-expression inhibitor that targets an alpha-helix-mediated protein interaction.* J Am Chem Soc, 2003. **125**(17): p. 4992-3.
135. Liu, P.Q., E.J. Rebar, L. Zhang, Q. Liu, A.C. Jamieson, Y. Liang, H. Qi, P.X. Li, B. Chen, M.C. Mendel, X. Zhong, Y.L. Lee, S.P. Eisenberg, S.K. Spratt, C.C. Case, and A.P. Wolffe, *Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A.* J Biol Chem, 2001. **276**(14): p. 11323-34.
136. Rebar, E.J., Y. Huang, R. Hickey, A.K. Nath, D. Meoli, S. Nath, B. Chen, L. Xu, Y. Liang, A.C. Jamieson, L. Zhang, S.K. Spratt, C.C. Case, A. Wolffe, and F.J. Giordano, *Induction of angiogenesis in a mouse model using engineered transcription factors.* Nat Med, 2002. **8**(12): p. 1427-32.
137. Urnov, F.D. and E.J. Rebar, *Designed transcription factors as tools for therapeutics and functional genomics.* Biochem Pharmacol, 2002. **64**(5-6): p. 919-23.
138. Urnov, F.D., J.C. Miller, Y.L. Lee, C.M. Beausejour, J.M. Rock, S. Augustus, A.C. Jamieson, M.H. Porteus, P.D. Gregory, and M.C. Holmes, *Highly efficient endogenous human gene correction using designed zinc-finger nucleases.* Nature, 2005. **435**(7042): p. 646-51.
139. Fuller, G.N., X. Su, R.E. Price, Z.R. Cohen, F.F. Lang, R. Sawaya, and S. Majumder, *Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST)/neuron-restrictive silencer factor, which can be functionally countered by REST-VP16.* Mol Cancer Ther, 2005. **4**(3): p. 343-9.
140. Mapp, A.K., A.Z. Ansari, M. Ptashne, and P.B. Dervan, *Activation of gene expression by small molecule transcription factors.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 3930-5.
141. Dervan, P.B. and B.S. Edelson, *Recognition of the DNA minor groove by pyrrole-imidazole polyamides.* Curr Opin Struct Biol, 2003. **13**(3): p. 284-99.
142. Nielsen, P.E., *Peptide nucleic acid targeting of double-stranded DNA.* Methods Enzymol, 2001. **340**: p. 329-40.
143. Faria, M. and C. Giovannangeli, *Triplex-forming molecules: from concepts to applications.* J Gene Med, 2001. **3**(4): p. 299-310.
144. Uil, T.G., H.J. Haisma, and M.G. Rots, *Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities.* Nucleic Acids Res, 2003. **31**(21): p. 6064-78.
145. Lawinger, P., R. Venugopal, Z.S. Guo, A. Immaneni, D. Sengupta, W. Lu, L. Rastelli, A. Marin Dias Carneiro, V. Levin, G.N. Fuller, Y. Echelard, and S. Majumder, *The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells.* Nat Med, 2000. **6**(7): p. 826-31.

146. Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream.* Nat Rev Drug Discov, 2004. **3**(4): p. 301-17.
147. Pagliaro, L., J. Felding, K. Audouze, S.J. Nielsen, R.B. Terry, C. Krog-Jensen, and S. Butcher, *Emerging classes of protein-protein interaction inhibitors and new tools for their development.* Curr Opin Chem Biol, 2004. **8**(4): p. 442-9.
148. Stockwell, B.R., *Exploring biology with small organic molecules.* Nature, 2004. **432**(7019): p. 846-54.
149. Fletcher, S. and A.D. Hamilton, *Protein surface recognition and proteomimetics: mimics of protein surface structure and function.* Curr Opin Chem Biol, 2005. **9**(6): p. 632-8.
150. Yin, H. and A.D. Hamilton, *Strategies for targeting protein-protein interactions with synthetic agents.* Angew Chem Int Ed Engl, 2005. **44**(27): p. 4130-63.
151. Fletcher, S. and A.D. Hamilton, *Targeting protein-protein interactions by rational design: mimicry of protein surfaces.* J R Soc Interface, 2006. **3**(7): p. 215-33.
152. Minter, A.R., B.B. Brennan, and A.K. Mapp, *A small molecule transcriptional activation domain.* J Am Chem Soc, 2004. **126**(34): p. 10504-5.
153. Buhrlage, S.J., C.A. Bates, S.P. Rowe, A.R. Minter, B.B. Brennan, C.Y. Majmudar, D.E. Wemmer, H. Al-Hashimi, and A.K. Mapp, *Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors.* ACS Chem Biol, 2009. **4**(5): p. 335-44.
154. Hidalgo, P.A., A. Z.; Schmidt, P.; Hare, B.; Simkovich, N.; Farrell, S.; Shin, E. J.; Ptashne, M.; Wagner, G.,, *Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain.* Genes & Development, 2001. **15**(8): p. 1007-1020.
155. Himmelfarb, H.J., J. Pearlberg, D.H. Last, and M. Ptashne, *GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators.* Cell, 1990. **63**(6): p. 1299-309.
156. Bourbon, H.M., A. Aguilera, A.Z. Ansari, F.J. Asturias, A.J. Berk, S. Bjorklund, T.K. Blackwell, T. Borggrefe, M. Carey, M. Carlson, J.W. Conaway, R.C. Conaway, S.W. Emmons, J.D. Fondell, L.P. Freedman, T. Fukasawa, C.M. Gustafsson, M. Han, X. He, P.K. Herman, A.G. Hinnebusch, S. Holmberg, F.C. Holstege, J.A. Jaehning, Y.J. Kim, L. Kuras, A. Leutz, J.T. Lis, M. Meisterernest, A.M. Naar, K. Nasmyth, J.D. Parvin, M. Ptashne, D. Reinberg, H. Ronne, I. Sadowski, H. Sakurai, M. Sipiczki, P.W. Sternberg, D.J. Stillman, R. Strich, K. Struhl, J.Q. Svejstrup, S. Tuck, F. Winston, R.G. Roeder, and R.D. Kornberg, *A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II.* Mol Cell, 2004. **14**(5): p. 553-7.
157. Ruas, J.L., L. Poellinger, and T. Pereira, *Role of CBP in regulating HIF-1-mediated activation of transcription.* J Cell Sci, 2005. **118**(Pt 2): p. 301-11.

158. McManus, K.J. and M.J. Hendzel, *CBP, a transcriptional coactivator and acetyltransferase.* Biochem Cell Biol, 2001. **79**(3): p. 253-66.
159. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo.* J Biol Chem, 2002. **277**(48): p. 46043-50.
160. Salghetti, S.E., A.A. Caudy, J.G. Chenoweth, and W.P. Tansey, *Regulation of transcriptional activation domain function by ubiquitin.* Science, 2001. **293**(5535): p. 1651-3.
161. Salghetti, S.E., M. Muratani, H. Wijnen, B. Futcher, and W.P. Tansey, *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.
162. Kuo, M.H., E. vom Baur, K. Struhl, and C.D. Allis, *Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription.* Mol Cell, 2000. **6**(6): p. 1309-20.
163. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. **25**(20): p. 9092-102.
164. Neely, K.E., A.H. Hassan, C.E. Brown, L. Howe, and J.L. Workman, *Transcription activator interactions with multiple SWI/SNF subunits.* Mol Cell Biol, 2002. **22**(6): p. 1615-25.
165. Ebright, Y.W., Y. Chen, Y. Kim, and R.H. Ebright, *S-[2-(4-azidosalicylamido)ethylthio]-2-thiopyridine: radioiodinatable, cleavable, photoactivatible cross-linking agent.* Bioconjug Chem, 1996. **7**(3): p. 380-4.

CHAPTER 2
CAPTURING A KEY BINDING PARTNER OF GAL4 IN S. cerevisiae*

A. Chapter overview

Transcriptional activators are proteins that modulate transcription, this requires that they bind to several other proteins to assemble the pre-initiation complex (PIC) for transcription to occur [1, 2] Due to several decades of research, there is a greater understanding of the complexes that co-localize with transcriptional activators at gene promoters. And yet not much has been revealed regarding the direct binding partners within those complexes.[3-9] This lack of understanding of transcriptional protein-protein interactions, and correspondingly the mechanisms used to successfully turn on genes, can in part be attributed to the methodologies that have been employed to study transcription PPIs. Importantly, a major roadblock towards the development of transcriptional therapeutics has been the lack of detailed mechanistic understanding of interactions between activators. Towards the identification of direct partners of activators, in this Chapter I outline the development of a strategy, in vivo photo-crosslinking that will be tested and evaluated for capturing a key direct binding partner of the prototypical transcription activator Gal4.

* Portions of this Chapter have been published. Chinmay Y. Majmudar, Lori W. Lee, Jody K. Lancia, Adaora Nwokoye, Qian Wang, Amberlyn M. Wands, Lei Wang and Anna K. Mapp, *J. Am Chem Soc*. **2009**, 40, 14240-42.

B. In vivo photo crosslinking strategy

All of the techniques for studying protein-protein interactions surveyed in Section E of Chapter 1 have helped to deepen our understanding of transcriptional protein-protein interactions. However, to correctly define activator binding partners, we need methodologies and/or strategies that capture relevant direct protein partners of transcriptional activation domains (TADs) in an in vivo context. Satisfying the major limitations presented in other techniques would mean that such a methodology would be able to covalently capture the direct-binding protein partners of transcriptional activators in their native environment.

One problem encountered while studying transcriptional activator interactions is the low abundance of transcription proteins. A detailed report from Erin O'Shea and coworkers revealed that transcriptional proteins are generally less than 5000 copies per cell (Table 2-1)[10]. Another major challenge is that transcriptional activators only transiently interact with proteins in order to assemble the PIC.[11] With these considerations in mind, we wanted to develop a strategy for studying transcriptional protein-protein interactions that would capture direct binding partners. This methodology should possess the ability to capture all the interactions that activators carry out, even transient interactions that are modest in strength, with little to no perturbation of the proteins interacting interfaces.

Parent-complex/protein function	Transcriptional proteins	Copies per cell
ADA and SAGA	GCN5	1180
Ccr4-Not complex	Caf1	1520
Ccr4-Not complex	Ada3	2470
Mediator complex	Med15	606
Mediator complex	Med16	1720
Paf1 complex	Cdc73	538
Proteasome	Sug1	4700
Rsc chromatin remodeling complex	Rsc1	259
SAGA	Spt3	1710
SAGA	Spt7	2360
SAGA	Taf10	3370
SAGA	Spt20	4150
SAGA	Ada5	4150
Swi/Snf	Swi1	92.3
Swi/Snf	Snf2	217
Swi/Snf	Snf5	217
TFIID and SAGA	Taf12	930 + 45
TFIID and SAGA	Ccr4	2780
Transcriptional activator	Gal4	166
Transcriptional regulator	Gal80	784
Transcriptional regulator	Ngg1	1720

Table 2-1: Transcriptional proteins classified by either the complex they associate with or the function they perform in the cell with the number of protein copies per cell based on the yeast GFP fusion database (http://www.yeastgenome.org).

Figure 2-1 shows a schematic of the strategy conceived. In vivo photo-crosslinking strategy, as defined here, includes using site-specific incorporation to label the TAD with a photo-reactive nonnatural amino acid that has the ability to form covalent crosslinks to the nearest binding partners. Following crosslinking, the crosslinked complexes were immunopurified and the identities of the binding partners were then determined by Western blot analysis. Indeed, the in vivo photo crosslinking strategy outlined in Figure 2-1 holds great promise for tackling the problems encountered so far when studying direct TAD binding

partners. This strategy does not only expand the utility of mutagenesis but can covalently capture direct proteins in live cells, with a molecule that minimally perturbs the protein surfaces.

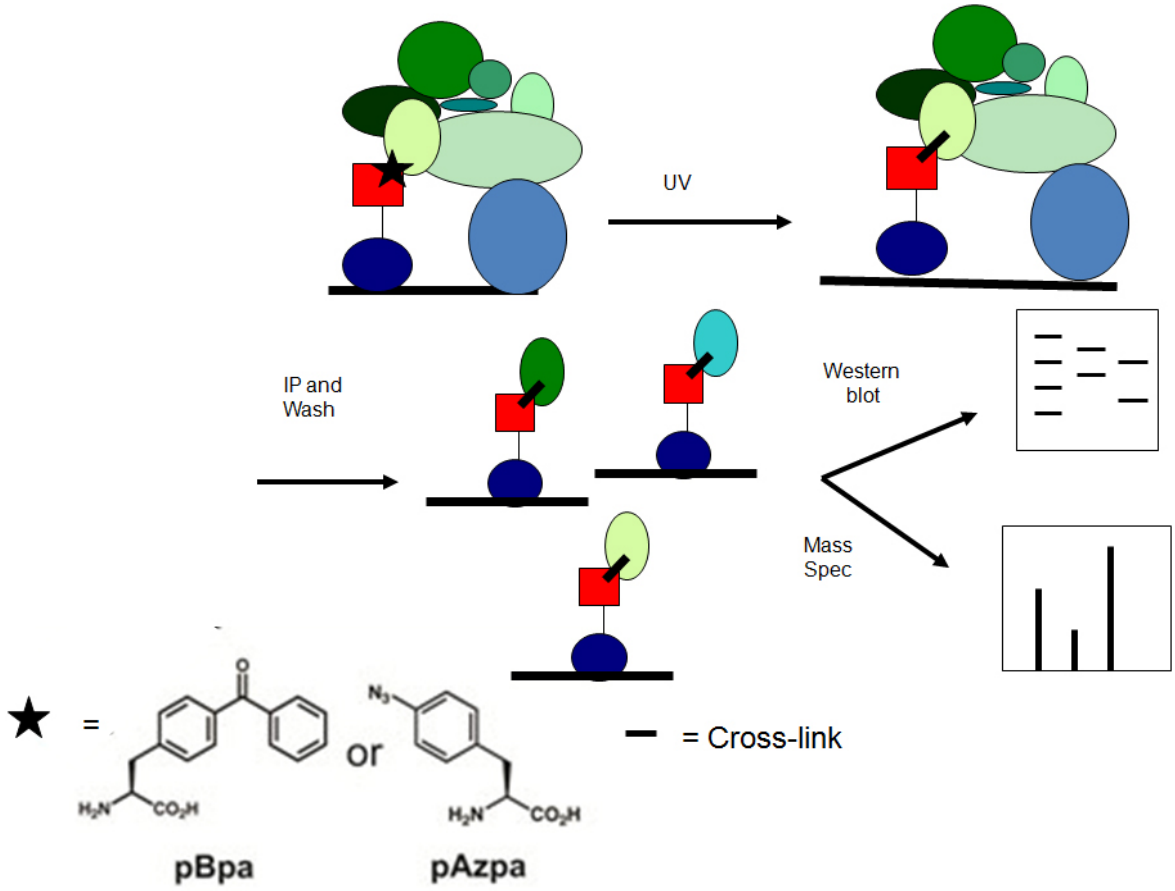


Figure 2-1: A schematic of the in vivo photo crosslinking strategy. In this strategy, A TAD protein (red box) with a site specifically incorporated photo-crosslinker is used to accomplish crosslinking and further crosslinked complexes are immunopurified before the identities of the binding partners of the TAD are determined by Western blotting analysis or other comparable techniques.

C. Nonsense suppression strategy and nonnatural amino acids

One way to achieve site specific incorporation of an amino acid with unique crosslinking abilities is to use nonsense suppression strategy. In the sub-sections below, I will provide a brief description of nonsense suppression

strategy and the two commonly used nonnatural amino acids with unique abilities that are suited for the studies described in this chapter.

C.1. Nonsense suppression strategy

The site-specific incorporation of nonnatural amino acids has facilitated the expansion of the genetic code beyond the twenty naturally occurring amino acids.[12-14] One of the most powerful methods, nonsense suppression, uses the cell's translational machinery to incorporate nonnatural amino acids into a protein of choice.[14] Standard site-directed mutagenesis techniques are used to incorporate a stop codon like an amber stop codon (TAG, the least used stop codon in most organisms) into the DNA sequence encoding the protein of choice.[13, 15] A tRNA that has been acylated with the desired nonnatural amino acid recognizes the amber stop codon during translation and incorporates the desired non-natural amino acid into the protein sequence. tRNA acylation is accomplished in the cell by using a modified aminoacyl-tRNA synthetase, which is capable of charging the tRNA with the nonnatural amino acid (Figure 2-2). Nonnatural amino acid incorporation has been accomplished in eukaryotes and bacteria in vitro and in vivo.[16, 17] In eukaryotes, especially S. *cerevisiae*, there are only a handful of examples with successful incorporation of non-natural amino acids.

A long list of tRNA modifications have been carried out to achieve an appreciable amount of protein that contains site specifically incorporated nonnatural amino acid in bacteria, yeasts and mammalian systems.

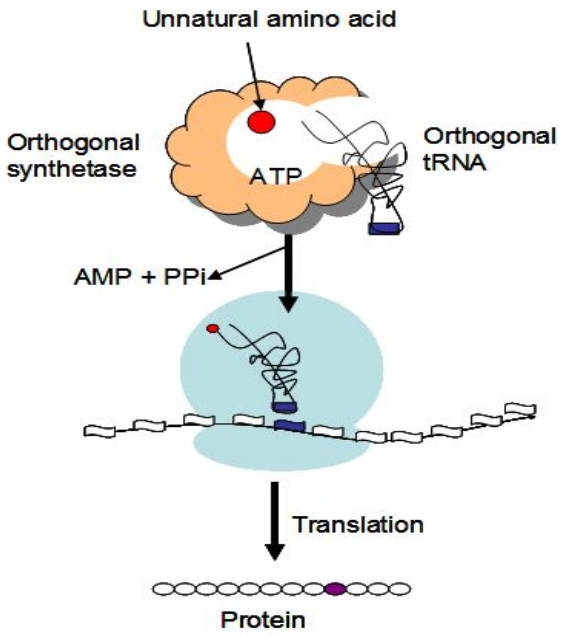


Figure 2-2: Site-specific incorporation of nonnatural amino acids in vivo, using nonsense suppression strategy. By means of the endogenous translational machinery, an orthogonal tRNA/synthetase pair recognizes the amber stop codon (designated here as a blue box) and charges the tRNA with a nonnatural amino acid.

For efficient incorporation of a nonnatural amino acid, the modified tRNA needs to be paired with a tRNA synthetase (orthogonal pair), which will only recognize that tRNA and incorporate the nonnatural amino acid instead of stopping translation or incorporating one of the naturally occurring amino acids.[13, 18] The first orthogonal tRNA/synthetase pair in E. coli was developed by the Schultz group. By modifying the $tRNA^{Tyr}$/tyrosyl-tRNA synthetase from Methanococcus jannaschii (M. jannaschii) they were able to realize low aminoacylation across kingdoms.[18] Through a general scheme of negative and positive selections, an efficient pair of orthogonal tRNA and tRNA synthetase that recognizes the amber stop codon was identified. The negative selection retains the orthogonal tRNAs from the library, and the subsequent positive selection retains the functional tRNAs with high affinity for the cognate synthetase. Briefly, first, a library of

tRNAs was generated by mutagenizing specific nucleotides on the tRNA and a library of tRNA synthetases were generated by randomizing some amino acid residues in the active site of the tRNA synthetase before the selection rounds, which yielded the efficient orthogonal tRNA and tRNA synthetase.[18]

Using the orthogonal tRNA/tRNA synthetase pairs designed for each unique amino acid, over 30 nonnatural amino acids have been successfully incorporated into proteins in bacteria. Also in bacteria, site-specific nonnatural amino acids have been successfully incorporated into proteins with yields of several milligrams to tens of milligrams per liter.[19-22] In contrast, the yields in eukaryotic systems have been significantly lower, ranging from 5ng/L to 8mg/L, which has been attributed to the differences in tRNA processing between prokaryotes and eukaryotes.[15, 23, 24] As an ongoing effort, several other groups have worked towards improving mutant (protein with the nonnatural amino acid incorporated) protein expression in eukaryotes by several modifications including: i) adding elements to the tRNA expression plasmid that are found naturally in eukaryotic tRNAs such as the A and B box and removing the 3'-CCA found in prokaryotic tRNAs, to generate a mature and functional tRNA (ii) testing expression of the tRNA under the control of different promoters and (iii) inactivating the rapid nonsense mediated mRNA degradation mechanism.[23, 25, 26]

Toward increasing protein expression yields in yeast, various groups have made key modifications to the first tRNA/tRNA synthetase reported by the Schultz group. Figure 2-3 is a summary of the different modifications that have been made to the orthogonal tRNA reported by Schultz. Starting with the Hahn

group, the effect of different promoters needed to drive tRNA production was tested. In these experiments, Hahn reported that the medium-strength N(GTT)PR promoter was able to produce “normal” levels of proteins with the nonnatural amino acid incorporated while the strong promoter fell short.[24, 26]

Further modifications were made by the Schultz laboratory where the orthogonal tRNA was expressed by a strong Pol II promoter. pPGK1 drove the expression of 3(SUP4-tRNA) chimeras, which are tRNAs lacking the 3’-CCA, but having SUP4 5' and 3' -flanking sequences.[15] SUP4 is a pol III promoter that contains internal A and B boxes, which is required for yeast tRNA transcription, unlike bacteria tRNAs.[23]

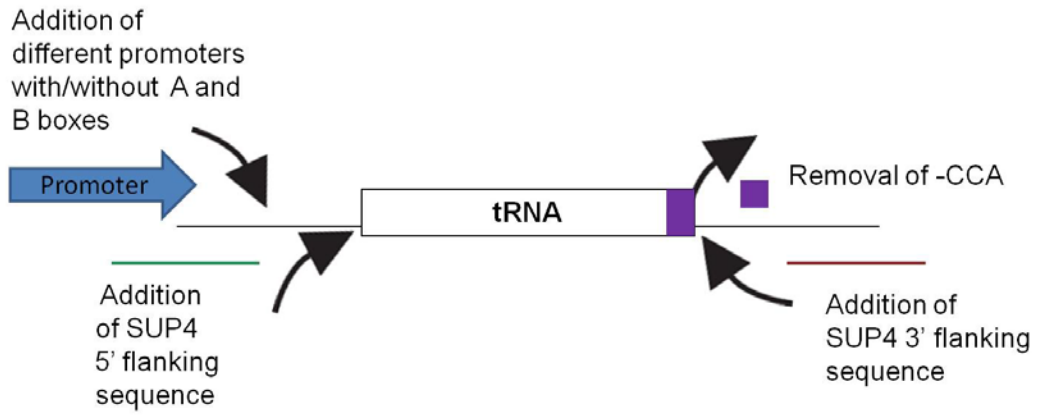


Figure 2-3: A summary of modifications made to the original tRNA developed by the Shultz group.

Additionally they removed the 3’ CCA nucleotides which is normally lacking in yeast. The series of modifications reported here, which was made to accommodate the difference in how bacteria and yeast process tRNA, were used to produce the fully processed tRNA.[27, 28] The new tRNA/tRNA synthetase

reported in these experiments was used to yield hSOD protein containing p-propargyloxyphenylalanine ranging 3-10mg/L.[23]

More recently, Wang and co-workers further improved the tRNA/synthetase expression plasmid, such that the E. coli tRNA may be post-transcriptionally cleaved to yield the final tRNA.[25] Using two yeast genes, SNR52 and RPR1 that are transcribed by Pol III and contain A and B-box promoter elements, they achieved enhanced production of the functional tRNA.[25] Similar to the second generation Schultz tRNAs, the Wang laboratory inserted the E. coli tRNA lacking the 3'-CCA but the SUP4 3'-flanking sequence into the plasmid.[23, 25] Their experiments showed that this new tRNA/synthetase expression plasmid yielded the mature tRNA and a subsequent increase in protein yield, 6-9 fold compared to the constructs with the SUP4 5'-flanking sequence alone.[25] They also showed that SNR52 promoter driving tRNA transcription produced more incorporated protein product even though having SUP4 5'-flanking sequence alone produced more tRNA, highlighting that tRNA processing possibly impacts tRNA function.[25]

Increasing nonnatural amino acid incorporation in eukaryotes opens the door to endless opportunities, since proteins with novel functionalities in appreciable amounts can be used to probe unanswered biological questions. For example nonnatural amino acids with crosslinking abilities can be used to covalently capture undefined direct binding partners of transcriptional activators. In the next sub-section, I will describe the two commonly used nonnatural amino

acids, which possess photo-reactive moieties that are employed in this chapter for studying transcriptional protein-protein interactions.

C.2. Nonnatural photo-reactive amino acids

Photo-active nonnatural amino acids have been used in vitro and in vivo for photo-crosslinking and photo-affinity labeling while studying protein-protein and protein/ ligand interactions.[14, 29, 30] The strength of the covalent bond formed after UV irradiation maintains the ligand/protein-target complexes, which are otherwise lost because of weak binding affinities seen with conventional methods used for studying ligand/protein interactions.[14, 31-36] Using nonsense suppression strategy, two primary nonnatural amino acids (pBpa and Azpa) that can participate in covalent capture have been successfully incorporated into proteins in vivo.[23 37, 38] The first, p-benzoyl-l-phenylalanine (pBpa) when activated, forms a diradical upon UV irradiation at 350-365 nm (Figure 2-4) making it able to insert into C-H bonds found in amino acids chains on the protein backbone of a close-by protein. pBpa has been shown to preferentially insert into C-H bonds that are connected to nitrogen or sulphur atoms, followed by triple and double bonded carbon bonds.[39-41]

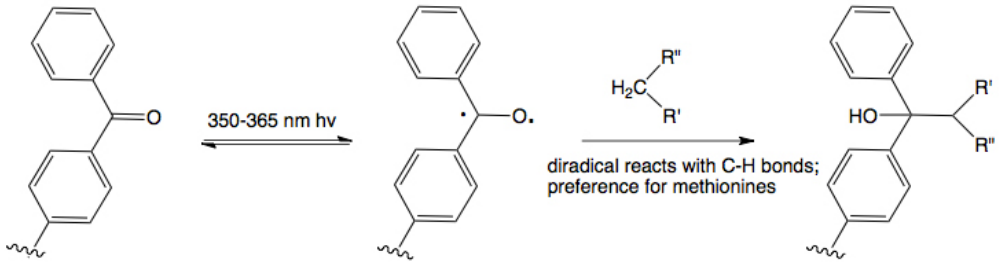


Figure 2-4: pBpa crosslinking reactive mechanism

In particular, pBpa shows a preference for reacting with methionine. When two helical peptides containing pBpa were used to label the hydrophobic cavity on calmondulin, pBpa selectively labeled two separate methionines in this cavity.[42] Other experiments have equally shown similar results similar to those reported by DeGrado and coworkers and further, that pBpa can react with methionine beyond the 3.1 Å reactive radius.[43] pBpa C-H insertion may not occur even in the presence of methionine because it was determined that the ideal angle of pBpa diradical attack is 108.9º, which occurs almost in-plane.[41] Advantages for using pBpa as a crosslinker include that its excitation wavelength (350-365 nm) is less damaging to biomolecules than Azpa. [15, 40] Another advantage is that the diradical induced by UV light can be maintained for up to 120 µs. If C-H insertion was not achieved, pBpa can relaxes back down and then re-initiate diradical formation, thereby increasing its chances of crosslinking.[44]

p-Azido-L-phenylalanine (pAzpa) is another nonnatural amino acid that has been used for studies in in vivo systems. It is believed that aryl azides like pAzpa convert to the more stable ketenimine upon UV irradiation at ~254 nm, and then crosslink to nucleophilic side chains in the protein-binding surface (Figure 2-5).[32, 45, 46] In one experiment the aryl azide located within a synthetic phospholipid bilayer was able to crosslink to tryptophan present on the polypeptide gramicidin, which was positioned at almost the opposite end.[46] Additionally, pAzpa can insert into either C-H or heteroatom-H bonds.[40] Unsuccessful insertion causes the nitrene to rearrange to a more stable ketenimine, which can react with nucleophiles like water.[40, 47]

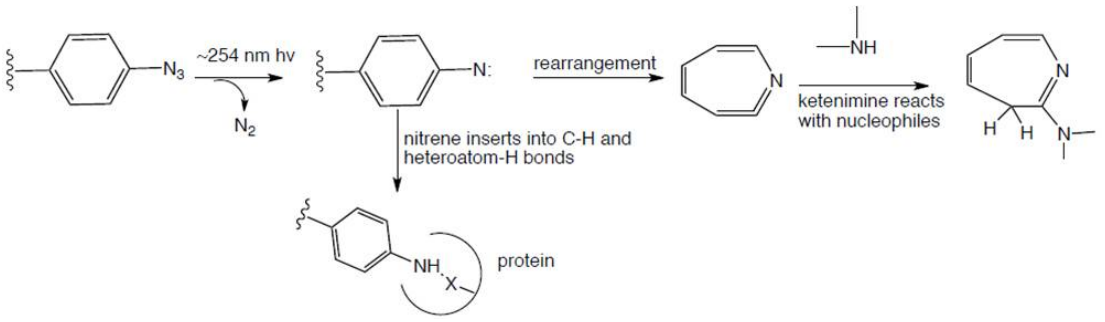


Figure 2-5: pAzpa crosslinking reactive mechanisms

Overall, incorporation of either pBpa or pAzpa in TADs will be useful for probing and targeting known and undiscovered direct binding partners of transcriptional activators. Making this route particularly promising is the fact that a large number of proposed TAD targets possess methionine and/or tryptophan and histidine residues, preferred by pBpa and pAzpa, respectively. For example, key residues needed for interaction between the well-characterized TAD, p53 a tumor suppressor protein, and its repressor protein, MDM2 includes methionine, tryptophan and histidine. Therefore, a p53 TAD containing either pBpa or pAzpa can be predicted to result in efficient crosslinking.[48] In summary, proteins containing pBpa and pAzpa may be efficient tools for studying transcriptional protein-protein interactions. In the next section, both pBpa and pAzpa will first be evaluated for efficient incorporation in the yeast transcriptional activator Gal4, followed by further development and optimization of the in vivo photo-crosslinking strategy, leading to the identification of a known direct binding partner of Gal4 in live yeast.

D. Results and discussion

One issue that remains unresolved in transcriptional protein-protein interactions is the identity of the direct binding partners of TADs during the process of transcription. To address this issue we used *S*. *cerevisiae* as the model organism because its transcriptional machinery is highly homologous to that of higher eukaryotes.[49] Importantly, genetic manipulations are fairly straightforward[50] in yeast relative to mammalian systems. Also we used the prototypical activator Gal4 for studies reported here to examine if an in vivo photo-crosslinking strategy can be utilized to identify direct TAD binding targets.

As the most characterized of the amphipathic activators, Gal4 is responsible for up-regulating expression of genes whose protein products are responsible for galactose catabolism in yeast.[51] Gal4 function is tightly regulated by the inhibitory protein Gal80.[52, 53] In the presence of glucose, Gal4 is engaged in an interaction with Gal80, making Gal4 unable to contact the necessary proteins for transcription initiation. Conversely, in the absence of glucose and in the presence of the inducing sugar galactose, Gal4 is no longer inhibited by Gal80 and is able to activate transcription (Figure 2-6).[52, 53]

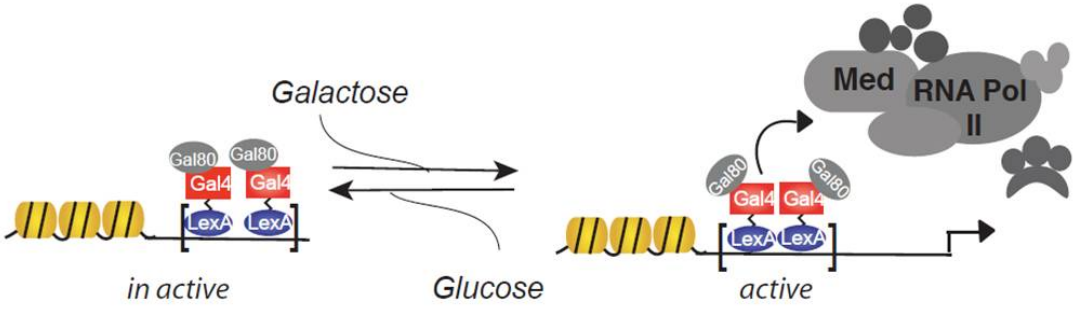


Figure 2-6: The general mechanism of Gal4 activation. In the present of glucose transcription is repressed by Gal80 and under galactose growth conditions Gal4 is activated for transcription.

Gal4 contains an N-terminal DNA binding and dimerization domain (DBD residues 1-147) and a C- terminal transcriptional activating (TAD residues 840-881) domain.[54] The Gal4 TAD has been shown to stimulate transcription in all eukaryotes tested when linked to a heterologous DBD, such as LexA.[55, 56] Here, all Gal4 TADs were linked to LexA DBD and constructed for expression in *S. cerevisiae* as previously described.[57] All protein expression was carried out in a yeast strain (LS41) containing an integrated LacZ reporter downstream of two LexA binding sites and the Gal1 promoter.

D.1. Optimization experiments for in vivo photo-crosslinking studies

Initially, our first goal was to develop a yeast system that could be used to successfully incorporate and express modified proteins (Gal4 having a nonnatural amino acid incorporated) in good yields. However, one primary problem of site-specific nonnatural amino acid incorporation is that there is often both read-through and high amounts of truncated protein. Here, read-through is defined as the incorporation of any other amino acid except for the nonnatural amino acid at the amber stop codon, which results in a full length transcriptional

activator that lacks the unnatural amino acid. Truncated proteins are defined as the resulting protein product when either a natural or nonnatural amino acid is not inserted during translation. A chimeric construct, LexA-Gal4 with a C-terminal FLAG tag, was constructed for initial yeast expression studies. In this construct the well-studied bacterial repressor protein, LexA was to serve as the DBD for localization to DNA, Gal4(840-881) as the TAD for interaction with the transcriptional machinery and FLAG was added for Western blot analysis.[58] In initial studies the construct in Figure 2-7, LexA-Gal4Phe849TAG, achieved by standard site directed mutagenesis, was used to develop conditions needed for successful incorporation.

Figure 2-7: The chimeric LexA-Gal4 construct used for initial optimization experiments.

D.1a. tRNA/synthetase optimization experiments

To achieve the best tRNA/synthetase plasmid for expressing full-length protein containing pBpa, Dr. Majmudar and I carried out a side-by-side comparison of four eukaryotic tRNA/tRNA synthetase (tRNA/aaRS) systems that have been previously published by other research groups. The first system obtained from the Schultz group was a single copy tRNA/aaRS plasmid from which two variants, expressing either 2 or 3 copies of the tRNA, were generated in our laboratory.[24] The second system, which was an improvement reported by

the Hahn laboratory, had a medium-strength N(GTT)PR promoter driving a single copy E. coli $tRNA_{CUA}$.[26] The third system, also developed by Schultz laboratory, involved a modification to their first system in which a strong two Pol II promoter, pPGK1 drove the expression of 3(SUP4-tRNA) chimeras, which are tRNAs lacking the 3'-CCA, but having SUP4 5' and 3' -flanking sequences.[23, 24] The final system was more recently improved by the Wang laboratory. In their system, a Pol III promoter (pSNR52) containing consensus A and B box sequences were used to enhance the production of tRNA lacking the 3'-CCA, but maintaining the SUP4 3'-flanking sequence.[25] These four tRNA/aaRS systems were used to test incorporation of the nonnatural amino acid pBpa in vivo into Gal4 TAD at Phe849. In order to test this, the tRNA/aaRS and the LexA-Gal4849TAG coding plasmids were co-transformed into *S. cerevisiae* (LS41 strain). Successful incorporation yielding full-length protein (~30 kDa) was evaluated by Western blotting to detect the C-terminal FLAG tag. Results from these experiments showed that the Wang system that used a pSNR52 promoter yielded the most mutant protein output with minimal read-through (Figure 2-8). This result also indicates that the addition of the A and B boxes, which influence the production of functional tRNA, is important for efficient incorporation.

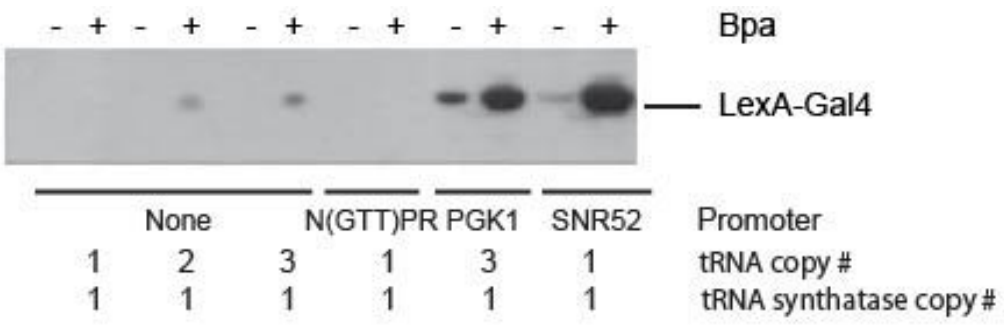


Figure 2-8: Using 4 different tRNA/synthetase systems (described above), the pSNR system was shown to yield the highest amount of full-length LexA-Gal4(849 TAG) protein with minimal amount of read-through.

D.1.b Optimization experiment with pBpa and pAzpa

To further optimize the yield of non-natural amino acid incorporation within the Gal4 TAD, we assessed which nonnatural amino acid between pBpa and pAzpa resulted in increased incorporation as well as for crosslinking studies. The first experiment in this sub-section was carried out to determine if using a pSNR tRNA/aaRS, specific for incorporating pAzpa could produce full length LexA-Gal4 similarly as seen for pBpa. LexA-Gal4849TAG was co-expressed with either the Wang group's pSNR tRNA/aaRS plasmid specific for either pBpa incorporation or pAzpa incorporation. Results from Figure 2-9 showed that both pBpa and pAzpa were readily incorporated to produce full-length protein. Because the results from this experiment were encouraging, the two nonnatural amino acids were further explored to assess which would produce efficient crosslinking with modified Gal4.

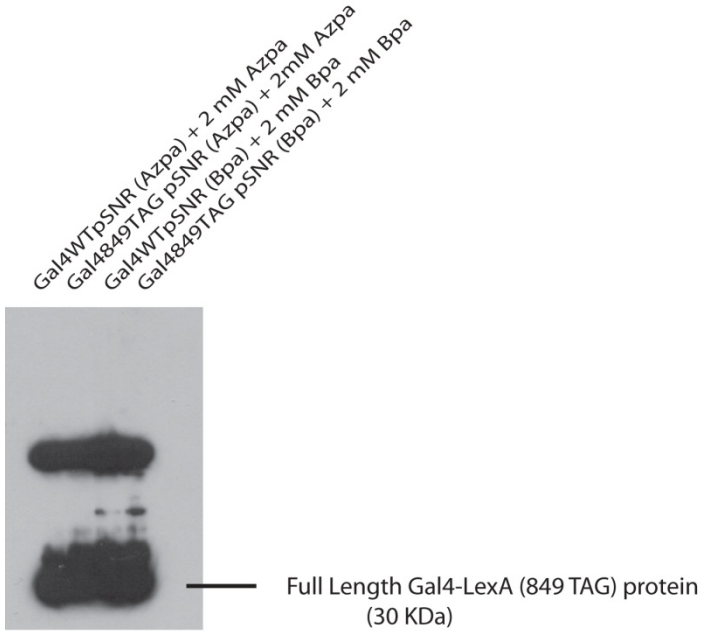


Figure 2-9: Comparison of pAzpa and pBpa incorporation in to LexA-Gal4 to yield the highest amount of full-length LexA-Gal4 (849 TAG) protein. Visualized with Flag HRP. Gal4WTpSNR(Azpa) and Gal4WTpSNR(Bpa) are controls.

A unique problem with transcriptional proteins is that they are not abundant proteins;[10] thus, we wanted to ensure that the nonnatural amino acid that would be selected for studying direct binding partners of Gal4 TAD would be efficient for capturing relevant target proteins. Towards that end optimization experiments summarized in Table 2-2 were carried out to ascertain the conditions needed for optimal crosslinking. After incorporation was achieved, LexA-Gal4 proteins containing pBpa or pAzpa were used to test whether UV light from either a Rayonet photo reactor at 4°C or an Enrosolar 15W UV lamp with cooling would produce crosslinked Gal4 complexes. Among other optimization experiments, we carried out a time course experiment to determine the length of time required for efficient crosslinking under UV light and we assessed the best stage in the population growth phase to harvests our yeast cells (Table 2-2).

From optimization experiments, starting with the UV source experiment, we found that using the Enrosolar 15W UV lamp as a UV light source produced a greater number of crosslinked Gal4 complexes for both LexA-Gal4 pBpa and pAzpa containing proteins after quantification by western blotting.

Variables	Tested conditions	Optimal conditions
Nonnatural amino acids for croslinking	pBpa and pAzpa were tested	pBpa
Concentration of pBpa and pAzpa	Not tested	2 mM (at this time)
OD (growth population)	0.5, 0.8, 1, 1.5, 2. 2.5 3, and 4	0.8-1.5
Crosslinking times	15 min, 30 min, 1 hr 1.5 hr and 2 hr	30 min to 1 hr
Immuno-precipitation of crosslinking complexes	IP with LexA or Flag	LexA
Antibodies fo Western blotting	different antibodies for LexA and Flag	Flag HRP

Table 2-2: A summary of the experimentally determined optimal conditions for carrying out in vivo photo-crosslinking experiments, based on results from all our extensive optimization experiments.

Overall, from crosslinking experiments, unlike pBpa, pAzpa exhibited a poor crosslinking profile. The low level of crosslinking observed could have been because pAzpa ketenimine intermediates readily interacts with water, thereby decreasing its capturing efficiency.[40, 59] From our time course experiment, results showed that increased crosslinking time yielded greater number of crosslinked Gal4 complexes after quantification by western blotting. To lessen the negative impact of UV exposure on the yeast cells, crosslinking time was limited to 1h. Extensive optimization experiments were carried out to determine other optimal conditions required for crosslinking experiments, see Table 2-2. All crosslinking experiments from this point were accomplished using these optimal conditions summarized in Table 2-2. Optimal crosslinking was achieved when: pBpa was

used for incorporation, yeast cells were harvested between 0.8-1.5 OD, crosslinking with UV light source (15W UV lamp) was between 30-60 min, LexA was used for immuno-precipitation and FLAG-HRP was employed for Western blot analysis.

*D.2. p-benzoyl-L-phenylalanine incorporation across the Gal4 transcriptional activation domain**

Transcriptional activator function requires the contact of several transcriptional proteins; however, molecular understanding of these interactions is limited.[60] Knowing that we could incorporate pBpa successfully at Gal4 F849TAG, we hypothesized that in vivo photo-crosslinking with Gal4 containing pBpa at various positions along the TAD would give a clearer understanding of proteins that interact with Gal4. To test our hypothesis, twelve additional point mutations, constructed in the same fashion as LexA-Gal4 F849TAG, were made throughout the Gal4 TAD (Figure 2-10).

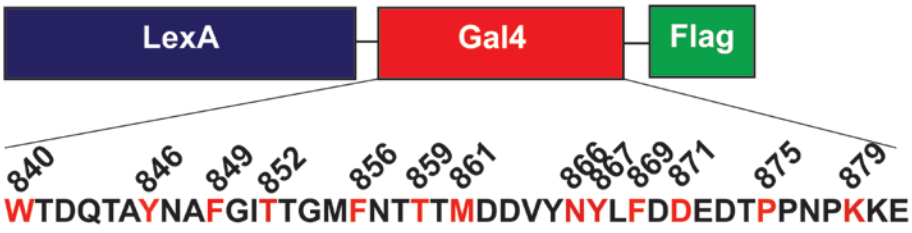


Figure 2-10: Thirteen amino acid residues (highlighted in red) selected for pBpa incorporation along Gal4 TAD.

Some of the residues mutated for the incorporation of pBpa in the Gal4 TAD have been shown through in vitro experiments to be important for binding to a well-studied binding partner like Gal80 or punitive binding partners like Tra1, Gal11, Ste12, and Taf12. Starting from the N terminus of the Gal4 TAD, amino

acids (a.a) 840 and 846 are thought to be outside Gal80 binding site while residues at positions, a.a 859 and a.a 856 have been suggested by protein gel shift experiments to be important for Gal4-Gal80 binding.[61] Furthermore, residues at positions a.a 860, a.a 864 and a.a 868, have been implied to engage in mediating the Gal4-Gal80 interaction. With the exception of positions a.a 861 and a.a 867, in vitro crosslinking indicates that a.a 869, a.a 870 and a.a 873 may contact Tra1, Gal11, Ste12, Taf12.[62] In addition, the last three positions, were selected to test if they were permissive for the incorporation of pBpa. The profile of all the residues selected was such that every amino acid class would be represented (see Table 2-3).

Amino Acid classes	TAG mutations made
Nonpolar, aliphatic	Yes (M, N)
Aromatic	Yes (F, Y & W)
Polar, uncharged	Yes (T)
Negatively charged	Yes (D)
Positively charged	Yes (K)
Cyclic (proline)	Yes (P)

Table 2-3: Selected residues for TAG mutation were representative of all the amino acid classes.

D.2.a. p-benzoyl-L -phenylalanine incorporation experiments along the Gal4 TAD[1]

To determine if pBpa can be successfully incorporated throughout the Gal4 TAD, eleven pLexA high copy expression plasmids expressing the wild type

(Wt) LexA(1-202)+Gal4(840-881)+FLAG and ten separate TAG mutations, were individually co-transformed with the already tested pSNR tRNA/synthetase pair plasmid in LS41 yeast strain. pSNRtRNA-pBpaRS incorporates pBpa at the amber stop codon mutations in the LexA-Gal4TAG constructs. Individual yeast colonies were grown in selection media containing 2% glucose. Following incubation, these cultures were used to inoculate 5 mL cultures induced with or without 2 mM pBpa for protein expression. The cells were lysed and analyzed by Western blotting with a FLAG-HRP conjugated antibody. All mutants of Gal4 were observed to express the full length protein only in the presence of pBpa (Figure 2-11).

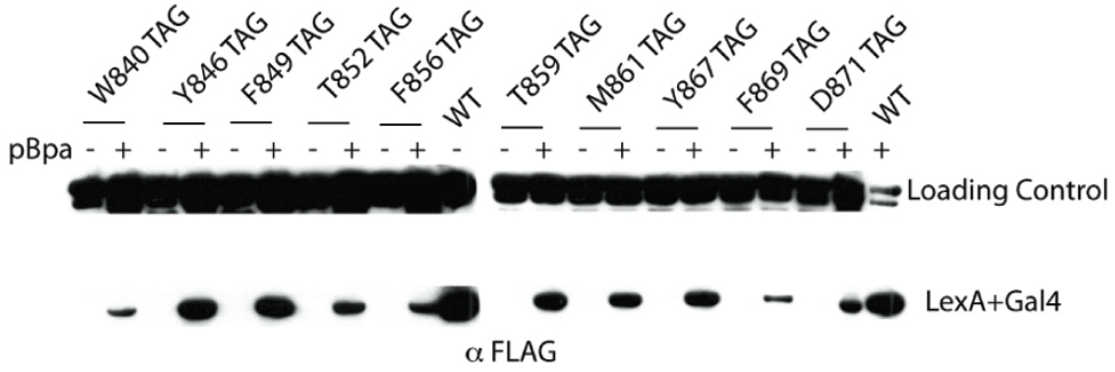


Figure 2-11: Evaluation of site-specific incorporation of pBpa at 10 positions along the Gal4 TAD.

Extending this study to the broader range of mutations throughout the Gal4 TAD resulted in improved incorporation at all positions, although the efficiency varied.

D.2.b. Transcriptional activity of Gal4 containing p-benzoyl-L -phenylalanine proteins[*]

In our next set of experiments the functional impact of incorporating pBpa across the Gal4 TAD at all thirteen positions using quantitative ß-galactosidase assays was determined. The repressor protein Gal80 inhibits Gal4 TAD activation in the presence of glucose. Gal4's inactivation by Gal80 can be overcome when galactose is present. Therefore, what we expected to see in our experiments was Gal4 activation in raffinose/galactose condition and the opposite in glucose conditions.

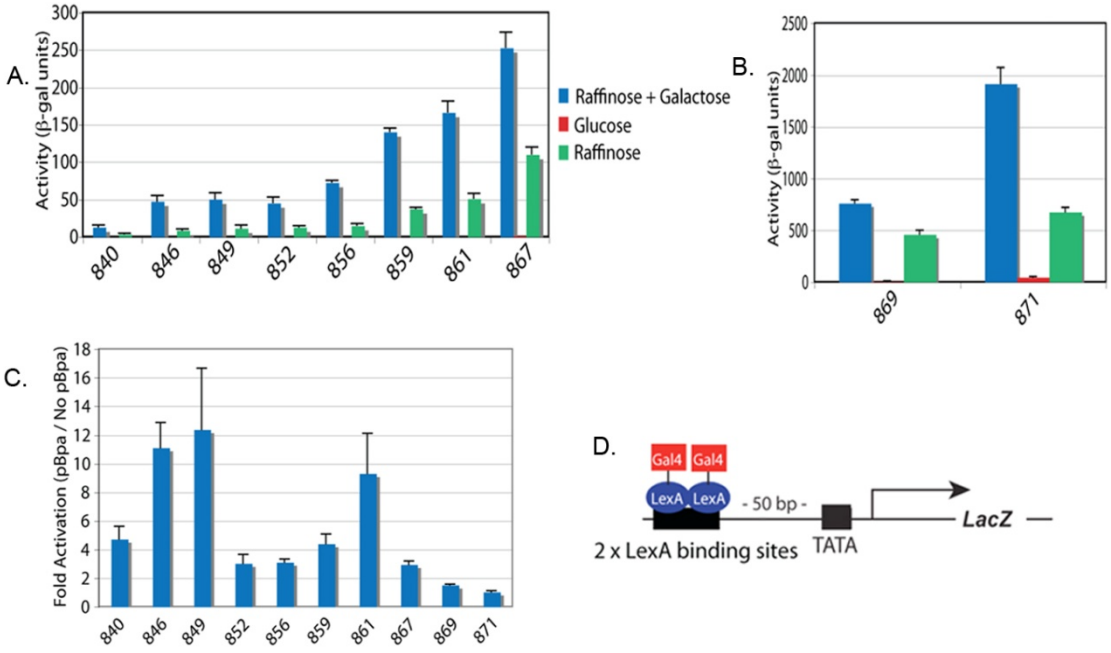


Figure 2-12: The functional impact of incorporating pBpa into the Gal4 TAD. (A.) β-galactosidase assays in LS41 (yeast strain), testing for pBpa incorporation with different sugars. (B.) Further analysis of β-galactosidase assays carried out in A. to assess fold activation (C.) LS41 contains an integrated LacZ reporter downstream of two LexA binding sites and the Gal1 promoter.

To determine if pBpa was transcriptionally functional and also, the efficiency of incorporating pBpa in all the Gal4 mutants, β-Galactosidase assays[63] were

performed using yeast cells expressing the LexA-Gal4 mutants. The yeast cells were grown in media containing either 2% raffinose + 2% galactose, 2% raffinose or 2% glucose and each sugar with or without 2 mM pBpa. As expected in Figure 2-12 under raffinose/galactose growth conditions, all the Gal4mutant proteins showed activity, and it seemed that activity was enhanced when the TAG mutation was closer to the C-terminus of Gal4TAD. Conversely, under glucose or raffinose growth conditions minimal transcriptional activity was observed, therefore the Gal4 protein containing pBpa was able to maintain its environmental sensitivity and be repressed by Gal80. Further analysis to assess the effect of truncation on activation (Figure 2-12C), indicates that Gal4 mutant protein incorporated from TAG positions closer to the N terminus, possibly have low protein truncations since they showed greater fold activation over their counterparts on the C terminus.

D.2.c. In vivo photo-crosslinking across the Gal4 transcriptional activation domain[*]

Once it was known that pBpa could be successfully incorporated at various positions along the Gal4 TAD and that all the mutants maintained functional activity, we carried out in vivo photo-crosslinking experiments to determine the binding profile of the Gal4 TAD. After pBpa was incorporated, yeast cells were crosslinked under optimized conditions. The cells were subsequently lysed and analyzed by Western blotting. Excitingly, all Gal4 mutants formed several crosslinked products, consistent with the multi-partner binding nature of Gal4.[62, 64, 65] Also, different binding profiles were exhibited

across the entire Gal4 TAD, so that there were some shared and unshared targets. For example, all the mutant proteins except for positions a.a 840 , a.a 856 and a.a 879, formed crosslinked products with a particular protein, adding up to the molecular weight of ~80 kDa (highlighted in the red box, Figure 2-13). We hypothesized that the boxed band was Gal80, which has a molecular weight of ~50 kDa.

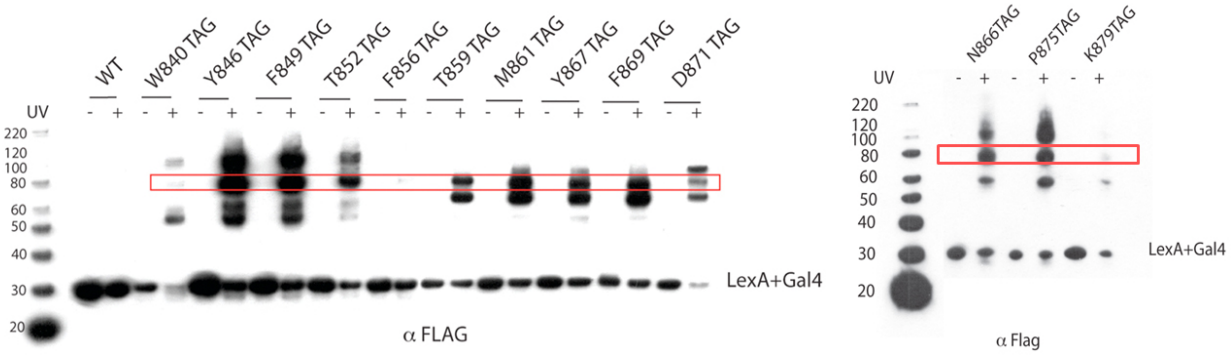


Figure 2-13: In vivo photo-crosslinking experiments with 13 Gal4 mutants containing pBpa. Crosslinked complexes are only observed with application of UV light. A band with a sum molecular weight of 80 kDa (red box) is shared by most mutants.

Being able to capture Gal80 as a direct binding partner of Gal4 in vivo, would demonstrate the usefulness of photo-crosslinking strategy for identifying direct transcription protein-protein interactions.

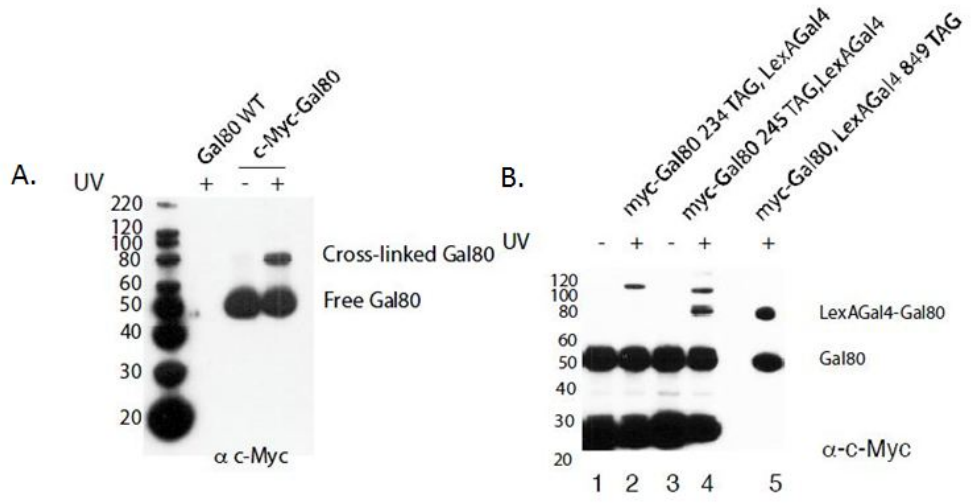


Figure 2-14: Crosslinking experiments to determine whether Gal4 crosslinks to c-Myc-Gal80 in vivo. (A.) Yeast cell expressing c-Myc-Gal80 and LexA-Gal4849Bpa. Only in the presence of c-Myc-Gal80 and UV light is Gal4-Gal80 crosslink observed. The experiment with Gal80 Wt is control. (B.) Yeast cells expressing c-Myc-Gal80 containing pBpa at positions 234 (outside Gal4 interacting site) and 245 (within Gal4 interacting site). Only position 245 shows crosslinking Gal80 (lane 4), lane 5 is used as control.

To determine if the 80 kDa protein complex was an interaction between Gal4-Gal80, Dr. Majmudar carried out crosslinking experiments in live yeast cells. In this experiment a c-Myc-tagged version of Gal80 was introduced into yeast and subsequently the cross-linked products between Gal4 and Gal80 was visualized with a c-Myc antibody (Figure 2-14A). Similarly, pBpa was incorporated into Gal80 at residue a.a 245 and crosslinking experiments with LexA-Gal4 confirmed that Gal4 directly interacts with Gal80 (Figure 2-14B).

We know that a.a 856 on the Gal4 TAD has been suggested to be important for Gal4-Gal80 interaction, but crosslinking experiments, showed what seemed to be attenuated binding due to pBpa incorporation at that position.[61, 66] To clarify if pBpa incorporation was responsible for attenuated interaction or loss of binding, Dr Majmudar carried out fluorescence polarization experiments to

access the binding of Gal4 TAD pBpa containing peptides, with Gal80. This experiment revealed that on the contrary, the Gal4856pBpa TAD exhibits a K_D of 0.7 ± 0.2 µM for Gal80, and this K_D is nearly identical to 1.2 ± 0.1 µM, (which is for the native Gal4 TAD with Gal80). Similarly, Gal4852pBpa and Gal4867pBpa TAD mutants, which seemed to show better crosslinking in Figure 2-13, had K_Ds that were nearly identical to that of Gal4856pBpa TAD. A possible explanation for the differences seen across the TAD for pBpa crosslinking is that the benzophenone moiety on pBpa may have been better positioned for crosslinking.[43]

Conventional mutagenesis and structural studies have suggested that the amino acids 851-871 on the Gal4 TAD comprise the binding site for Gal80. It is somewhat surprising that this short sequence incredibly impacts the function within Gal4-Gal80 complex interaction.[61, 66, 67] Using crosslinking experiments we were also able to see, that positions beyond 851-871, showed crosslinking to Gal80 (Figure 2-13). This result suggests that Gal4-Gal80 interaction extends to positions 846 on the N terminus, and 875 on the C terminus, but not as far as 840 and 879 (Figure 2-13)

We also demonstrated that pBpa could be incorporated in the Gal4 TAD to replace residues from all amino acid classes. Furthermore, an in vivo photo crosslinking strategy was used to capture Gal80, a well-defined binding partner of Gal4. Using this strategy also revealed an extended Gal4-Gal80 contact interface, a finding that was not uncovered with traditional biochemical methods. The important information from the experiments reported in this sub-section is

that an in vivo photo cross-linking strategy is appropriate for capturing direct protein-protein interactions in live cells.

E. Impacting crosslinking efficiency by increasing mutant protein output

Even though in vivo photo-crosslinking is a powerful method, it has not yet been successfully used to report discovery of novel PPIs, which would be the most powerful implementation of this strategy. Discovering unknown PPIs using this method require that ample amount of protein containing a crosslinker is produced and the crosslinker placement and reactivity are ideal for target identification. Some barriers to crosslinking experiments with site specifically incorporated nonnatural amino acids often include, low incorporation efficiencies and the inherent selectivities of the crosslinking reactions, all of which could perturb the crosslinked protein product distribution.[15, 40] Significant optimization of mutant protein expression is required to effectively maximize the chances for producing crosslinks in eukaryotes.

To further improve the incorporation efficiency of the photo-crosslinker, pBpa, we first determined the optimal pBpa concentration needed to achieve the highest amount of incorporation, a variable we had not addressed in our previous optimization experiments. Secondly, the copy number of orthogonal tRNA and aaRS was optimized. Finally, a chemical agent that aids in mis-sense codon read-through was evaluated for pBpa incorporation. Again, using the yeast transcriptional activator Gal4 as a model system we were able to show the optimal conditions for expressing the highest yield of the protein with the crosslinker incorporated, thus increasing the chances of capturing a direct

binding partner. The Gal4 model system was used not only because it is a well-studied transcriptional activator, but it interacts with low abundant proteins involved in transcription that have a history of being difficult to define.[2, 10, 68]

E.1. Effect of tRNA and Synthetase copy number on pBpa incorporation efficiency

Since our earlier optimization experiments (Section D.1) showed that the Wang pSNR52 system yielded the most protein activity and mutant protein output while minimizing read-through, we focused our optimization on the pSNR52 system. We evaluated the concentration of pBpa required for yielding the highest amount of mutant protein by varying the concentration of pBpa present in the growth media. Our results show that an optimal concentration of pBpa of 1 mM, was sufficient and afforded yields of modified protein that were ~64% of the WT construct (Figure 2-15A). pBpa concentrations greater than 1 mM did not show increase in the yield of full-length LexA-Gal4 Phe849pBpa protein. This may be because increasing pBpa concentration in the media may not necessarily increase the amount present in the cell as demonstrated previously with other nonnatural amino acids in mammalian systems.[70] Decreasing pBpa concentration to 0.5 mM resulted in a slight decrease in pBpa incorporation.

Knowing the optimal concentration for pBpa incorporation, we further fine-tuned the conditions in the pSNR52 system to improve mutant protein output, by increasing the copies of gene for both the tRNA and tRNA synthetase. Additionally, we evaluated the effect of varying the copy number of the aaRS and

its significance on the incorporation efficiency. In this set of experiments, we compared pBpa incorporation under the following tRNA/aaRS expression conditions: 1 copy pSNR52 tRNA/1 copy aaRS; 2 copies pSNR52 tRNA-1 copy aaRS; 1 copy pSNR52 tRNA-2 copies aaRS and 2 copies pSNR52 tRNA-2 copies aaRS. Results from the β-galactosidase assays, which also compared transcriptional activity of three additional tRNA/aaRS systems, revealed a slight gain in activity for the 2 copies pSNR52 tRNA-1 copy aaRS and 2 copies pSNR52 tRNA-2 copies aaRS constructs (Figure 2-15B). To determine if the slight gain in activity was due to an increase in incorporation efficiency, LexA-Gal4 Phe849pBpa protein expression was measured by Western blot analysis for the same tRNA/aaRS combinations. Figure 2-15C showed an increase in pBpa incorporation with 2 copies pSNR52 tRNA-2 copies aaRS, but a lowered expression with 2 copies pSNR52 tRNA-1 copy aaRS. Contrary to previous publications, these results with 2 copies pSNR52 tRNA-2 copies aaRS, shows that the additional copy of tRNA and synthetase resulted in an increase in full length mutant protein output, as well as statistically significant gain in activity (Figure 2-15B, blue box).[70, 71]

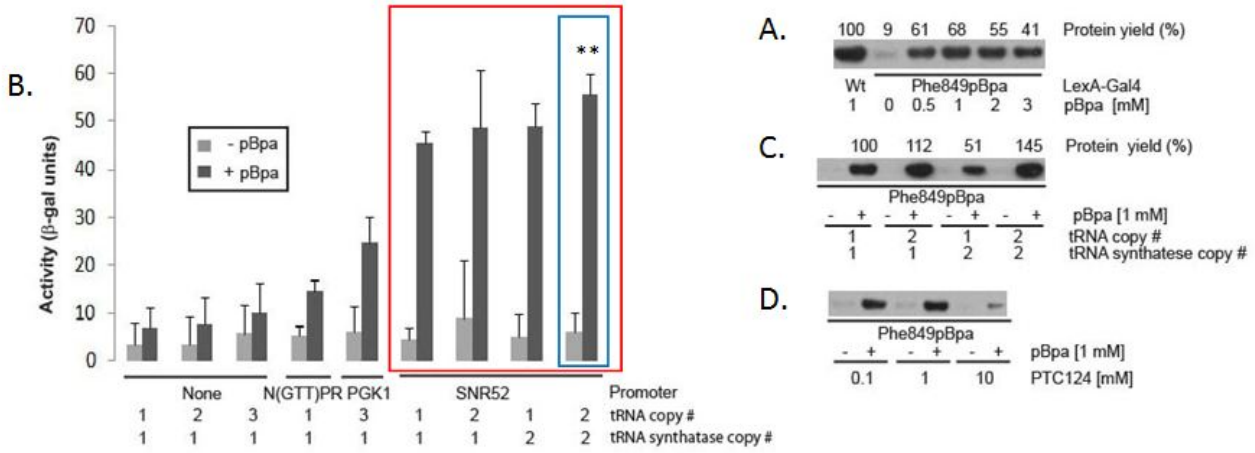


Figure 2-15: Determining the optimal orthogonal tRNA/aaRS pair, pBpa concentration, tRNA/aaRS copy number and the effect of PTC124 on LexA-Gal4 Phe849pBpa protein output. (A.) A quantitated Western Blot of lysed yeast generated using the pSNR52 tRNA/aaRS pair. Yeast cultures where grown in the presence of varying concentrations of pBpa. The % protein yield is calculated relative to WT LexA-Gal4. To account for loading variations, each band was normalized to α-tubulin. (B.) The optimal tRNA/aaRS pair and copy numbers was determined by β-galactosidase assays. The amount of activity, which is the average values of three independent experiments with the indicated error (SDOM), is related to functional protein output. Within the pSNR52 tRNA/aaRS system (red box), The construct with 2 tRNA and 2 aaRS copy numbers showed statistically significant activity with a p value of 0.0066 (** = p) (C) A quantitated Western Blot of lysed yeast generated with varying copy numbers of the tRNA/aaRS. The % protein yield is calculated relative to optimal LexA-Gal4 Phe849pBpa (pSNR 1 copy of tRNA and aaRS each). To account for loading variations, each band was normalized to α-tubulin and ~10% read-through was observed in the absence of pBpa. (d) A Western Blot of lysed yeast generated using the pSNR52 tRNA/aaRS in the presence of varying concentration of PTC124. All quantitated Western Blots were analyzed by Adobe Photoshop software to determine the density of the protein bands.

E.2. Effect of PTC124 on pBpa incorporation efficiency

The final experimental strategy employed to increase protein output was to use Premature termination codon 124 (PTC124) to create a loose translational mechanism that would permit an increase of amino acid incorporation at our engineered amber stop codon (TAG). PTC124 (Figure 2-16) in chemical footprinting studies have been shown to bind to the 60s ribosomal subunit thereby inducing read-through of nonsense mutations by allowing the ribosome

to insert an amino acid at a premature stop codon.[72-74] PTC124, functioning as a nonsense codon suppressor, results in read-through of the mutation and synthesis of the mature protein.[74]

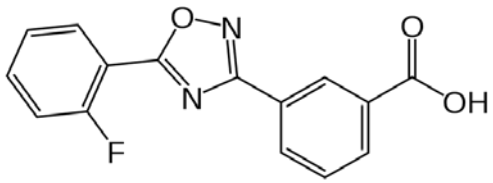


Figure 2-16: PTC124

To test if the presence of PTC124 will increase pBpa incorporation, LexA-Gal4 Phe849TAG and pSNRtRNA-pBpaRS plasmids were expressed in yeast at varying concentrations of PTC124. The addition of 0.1 µM to 10 µM PTC124 did not show increase in either full-length mutant protein output (Figure 2-15D) or transcriptional activity (data not shown). Results seen here may be because PTC124 propagated read-through levels for the UAG termination signals are generally less than UGA.[75] Also, since propensity of stop codon suppression depends on the nucleotide at the +1 position of the coding sequence, using PTC124 in these experiments may not have been optimal.[75-77] Based on these results it appears that the use of PTC124 as a vehicle to increase read-through did not result in increased protein output.

Collectively, the work in this section addresses some key variables that affect the experimental outcome when using photo-crosslinking in live cells. Several factors need to be considered when using in vivo photo-crosslinking for studying PPIs. We were able to demonstrate here that increasing the copy

number of tRNAs and aaRS present in the vector produced more protein, and a significant gain in activity was observed. Applying the considerations presented here would aid the use of in vivo photo-crosslinking as a tool for studying PPIs involved in a variety of biological processes.

F. Methods

LS41 [JPY9::pZZ41, *Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41] yeast was used for all experiments. pBpa was purchased from Chem-Impex International (Wood Dale, IL). All plasmids described below were constructed using standard molecular biology techniques and the sequences of all the isolated plasmids were verified by sequencing at the University of Michigan Core Facility (Ann Arbor, MI).

Table 2-4: Plasmids used for study in Chapter 2

Plasmid name	**Function**
pLexAGal4	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag
pLexAGal4 840TAG, pLexAGal4 846TAG, pLexAGal4 849TAG, pLexAGal4 852TAG, pLexAGal4 856TAG, pLexAGal4 859TAG, pLexAGal4 861TAG, pLexAGal4 867TAG, pLexAGal4 869TAG, pLexAGal4 871TAG	Express LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid
pSNRtRNA-pBpaRS	Expresses tRNA under the control of the SNR52 promoter and contains synthetase specific for pBpa
ptRNA-pBpaRS	Expresses tRNA with no eukaryotic Pol III promoter and contains synthetase specific for pBpa[24]
pMycGal80	Expresses Gal80 fused to c-Myc

	tag
pMycGal80 234 TAG, pMycGal80 245 TAG	Expresses Gal80 fused to c-Myc tag with a TAG replacing the codon of the existing amino acid
pmOCR-Gal80	Expresses Gal80 fused to the mOCR solubility tag in E. coli

pLexAGal4

A high copy plasmid expressing LexA(1-202)+Gal4(840-881)+FLAG tag under the control of the ADH1 promoter was created from pNLexA (Origene). First, site-directed mutagenesis was used to mutate the existing EcoRI and BamHI sites at the N-terminus of LexA and subsequently to insert EcoRI and BamHI sites at the C-terminus of LexA, producing the plasmid pCLexA. Primers (5'- TTA CGA ATT CTG GAC GGA CCA AAC TG -3') and (5'- AGT GGA TCC TTA TTT GTC GTC GTC GTC TTT ATA GTC CTC TTT TTT TGG G -3') were used to amplify Gal4(840-881) from EGY48 yeast genomic DNA. The amplified PCR product was digested with EcoRI and BamHI and inserted into pCLexA digested with EcoRI and BamHI and calf intestinal phosphate treated to create pLexAGal4. pLexAGal4 840TAG, pLexAGal4 846TAG, pLexAGal4 849TAG, pLexAGal4 852TAG, pLexAGal4 856TAG, pLexAGal4 859TAG pLexAGal4 861TAG, pLexAGal4 867TAG, pLexAGal4 869TAG and pLexAGal4 871TAG plasmids containing various amber mutants in the Gal4 TAD were derived from pLexAGal4. To create each plasmid, site-directed mutagenesis was used to replace an existing amino acid codon with TAG within the Gal4 TAD. In general, PCR primers were designed to have ~15 bases of homology on either side of the

TAG mutation. QuikChange (Stratagene, La Jolla, CA) was used to incorporate the TAG mutants using manufacturer recommended conditions.

pSNRtRNA-pBpaRS

The previously described pSNRtRNA-TyrRS plasmid incorporating tyrosine at the amber position was used to generate a plasmid pSNRtRNA-pBpaRS to incorporate pBpa at the amber position.[25] This plasmid was generated by insertion of the pBpa-specific *E. coli* tyrosyl synthetase (amplified from ptRNA-pBpaRS (p-benzoylPheRS-2)[24] obtained from Dr. P. G. Schultz, Scripps Research Institute, La Jolla, CA) using primers 5'- AGT TCA ACT AGT ATG GCA AGC AGT AAC TTG ATT -3' and 5'-TCG ATC TCG AGT TAT TTC CAG CAA ATC AGA CA-3') into SpeI and XhoI digested and calf intestinal phosphate treated pSNR TyrRS.

pMycGal80, pMycGal80234TAG and pMycGal80245TAG

A plasmid expressing Gal80 fused to the c-Myc tag under the control of the ADH1 promoter was generated by insertion of DNA encoding *S. cerevisiae* Gal80 into the high copy plasmid pADT7 (Clontech, Mountain View, CA). Primers 5'-TGT GAA GCT TAT GGA ACA AAA GTT GAT TTC TGA AGA AGA TTT GGA CTA CAA CAA GAG ATC TTC G -3' and 5' CGT CAA GCT TTT ATA AAC TAT AAT GCG AG -3' were used to amplify Gal80 from LS41 yeast genomic DNA. The amplified PCR product was digested with HindIII and inserted into HindIII digested and calf intestinal phosphate treated pADT7. pMycGal80234TAG and

pMycGal80245TAG were constructed by performing site-directed mutagenesis to incorporate a TAG codon to replace the codon of the existing amino acid

Incorporation of pBpa into LexA(1-202)+Gal4(840-881)

LS41 yeast was transformed with various pLexAGal4 TAG mutant plasmids and pSNRtRNA-pBpaRS or ptRNA-pBpaRS. Individual colonies were grown in 5 mL SC media containing 2% glucose but lacking histidine and tryptophan for selection. The cultures were incubated overnight at 30 °C and agitated at 250 rpm. Following incubation, these cultures were used to inoculate 5 mL cultures of SC media containing 2% glucose, with or without varying amounts of pBpa (dissolved in 50 µL of 1M NaOH), and 50 mL 1M HCl, except for the PTC124 experiments were varying concentrations of PTC124 (0, 0.1,1 and 10µM) were first dissolved in DMSO and then added, were subsequently incubated overnight at 30 °C with agitation to an OD_{660} of 1.5. 3 ODs of cells were isolated, washed with cold, sterile water and stored dry at -20 °C. The samples were lysed in 10 mL 4x NuPAGE LDS Sample buffer (Invitrogen), 15 mL Lysis Buffer (50 mM Tris-Acetate, pH 7.9, 150 mM KOAc, 20% glycerol, 0.2% Tween-20, 2 mM b-mercaptoethanol, 2 mM MgOAc) 5 mL 1M DTT and analyzed using Western blot with the anti-FLAG(M2) antibody (Sigma).

In vivo cross-linking

To perform in vivo cross-linking, individual colonies of each pLexAGal4 TAG mutant were grown in 5 mL SC media containing 2% glucose but lacking histidine and tryptophan for selection. The cultures were incubated overnight at

30 °C and agitated at 250 rpm. Following incubation, these cultures were used to inoculate 50 mL cultures of SC media containing 2% glucose, with 2 mM pBpa (dissolved in 0.5 mL of 1M NaOH), and 0.5 mL 1M HCl, which were subsequently incubated overnight at 30 °C with agitation to an OD_{660} of 1.5. For each mutant, 50 ODs of cells were isolated, washed with water and either resuspended in 2 mL water and irradiated for 1 h with 365 nm light (Eurosolar 15 W UV lamp) with cooling or kept in the dark at 4 °C. Following irradiation, all the cells were pelleted and stored at -80 °C until lysis. For lysis, cells were resuspended in 600 mL Lysis buffer (50 mM Hepes-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-Deoxycholate and 2X Complete Mini, EDTA Free Protease Inhibitor (Roche) and lysed using glass beads by vortexing at 4 °C. Subsequently, the lysate was pelleted and the supernatant incubated with 10 mL of LexA antibody (N-19, Santa Cruz Biotechnologies) for 2 h at 4 °C for immunoprecipitation. The protein bound to the antibody was isolated by incubation for 1 h with ~50 mL of prewashed protein G magnetic beads (Dynal Corporation, Invitrogen, Carlsbad, CA) at 4 °C. The beads were washed 3X with 1 mL Wash Buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5% NP-40, 1% Na-Deoxycholate and 1 mM EDTA) and stored dry at -20 °C. The protein was eluted from the beads by heating at 95 °C for 10 min in NuPAGE LDS Sample buffer (Invitrogen, Carlsbad, CA) containing DTT and probed using Western Blot analysis using anti-FLAG (M2) antibody (Sigma, St. Louis, MO).

β-Galactosidase assays

To evaluate the ability of each LexA+Gal4 TAG mutant in the absence or presence of 2 mM pBpa to activate transcription, saturated cultures (SC media + 2% raffinose) of each mutant were used to inoculate 5 mL SC media lacking histidine and tryptophan supplemented with 2% glucose or 2% raffinose + 2% galactose or 2% raffinose and grown to an OD of 1.5-2.0 before being harvested. The activity of each construct was monitored using β-galactosidase assays as previously described.[63]

Expression of Gal80

Expression of Gal80 fragments fused to the His6-mOCR solubility tags was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen). Briefly, cultures (50 mL) from single colonies were grown overnight at 37°C (300 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL) before addition to 1L of Select APS Super Broth supplemented with ampicillin (100 μg/mL). After an OD600 of 0.8 was reached, the cultures were cooled on ice for 30 m, and expression was induced with IPTG (final concentration 0.1 mM) for ~14 h at 20 °C. The cell pellet was resuspended in Buffer A (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME and Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the His-tagged protein was isolated using Ni NTA-Agarose (Qiagen). The Ni-NTA beads were washed 6 times with Wash buffer (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME, 30 mM imidazole). The protein was eluted from the beads at 4 ºC 3 times using Elution buffer (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-

ME, 300 mM imidazole). The resulting mixture was placed in dialysis tubing (Pierce) and the buffer exchanged to a low salt buffer (50 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 50 mM NaCl, 10 mM β-ME) overnight at 4 °C. The solution thus obtained was loaded onto an anion-exchange column (Q sepharose, GE Healthcare) and eluted with a NaCl gradient from 0-1 M. Fractions containing Med15(1-416) were pooled and buffer exchanged to Storage buffer (10 mM PBS pH 7.4, 10% glycerol, 0.01% NP-40, 1 mM DTT) using a PD-10 column (GE Healthcare) and concentrated using a 30K centrifugal filter device. The protein concentration was measured using absorbance at 280 nm. The identity and purity of the protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

Peptide Synthesis

Gal4(840-870) peptides with pBpa at various positions were synthesized using solid phase peptide synthesis in accordance with standard protocols. FMOC-pBpa was purchased from Chem-Impex International. The peptides contained a b-alanine linker at N-terminus and they were labeled using fluorescein isothiocynate on solid phase and subsequently cleaved using 95% TFA, 2.5% TIS, 2.5% H_20. The products were purified to homogeneity using reversed-phase HPLC on a C18 column with a gradient solvent system (buffer A: 20 mM Ammonium Acetate, buffer B: MeOH) and stored at -80 ºC. The identity was verified using electrospray mass spectrometry (LCT Micromass).

Fluorescence Polarization

Dissociation constants for fluorescein labeled Gal4 TADs and Gal80 were determined using fluorescence polarization as previously described. For each experiment, 50 nM TAD was incubated with varying concentrations of Gal80 in 100 mM PBS pH 7.2, 10% glycerol, 0.01% NP-40, 1 mM DTT and incubated for 10 min in 384 well low volume plate (Corning) before being detected using a TECAN Genios Pro plate reader and the dissociation constant calculated as previously described.

G. References

1. Thomas, M.C. and C.M. Chiang, *The general transcription machinery and general cofactors.* Crit Rev Biochem Mol Biol, 2006. **41**(3): p. 105-78.
2. Traven, A., B. Jelicic, and M. Sopta, *Yeast Gal4: a transcriptional paradigm revisited.* EMBO Rep, 2006. **7**(5): p. 496-9.
3. Horikoshi, M., T. Hai, Y.S. Lin, M.R. Green, and R.G. Roeder, *Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex.* Cell, 1988. **54**(7): p. 1033-42.
4. Lin, Y.S., I. Ha, E. Maldonado, D. Reinberg, and M.R. Green, *Binding of general transcription factor TFIIB to an acidic activating region.* Nature, 1991. **353**(6344): p. 569-71.
5. Stargell, L.A. and K. Struhl, *The TBP-TFIIA interaction in the response to acidic activators in vivo.* Science, 1995. **269**(5220): p. 75-8.
6. Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J.L. Regier, S.J. Triezenberg, D. Reinberg, O. Flores, C.J. Ingles, and et al., *Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53.* Mol Cell Biol, 1994. **14**(10): p. 7013-24.
7. Tan, Q., K.L. Linask, R.H. Ebright, and N.A. Woychik, *Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria.* Genes Dev, 2000. **14**(3): p. 339-48.
8. Dynlacht, B.D., T. Hoey, and R. Tjian, *Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation.* Cell, 1991. **66**(3): p. 563-76.
9. Fan, X., D.M. Chou, and K. Struhl, *Activator-specific recruitment of Mediator in vivo.* Nat Struct Mol Biol, 2006. **13**(2): p. 117-20.
10. Ghaemmaghami, S., W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'Shea, and J.S. Weissman, *Global analysis of protein expression in yeast.* Nature, 2003. **425**(6959): p. 737-41.

11. Bryant, G.O. and M. Ptashne, *Independent recruitment in vivo by Gal4 of two complexes required for transcription.* Mol Cell, 2003. **11**(5): p. 1301-9.
12. Wang, L. and P.G. Schultz, *Expanding the genetic code.* Chem Commun (Camb), 2002(1): p. 1-11.
13. Wang, L., J. Xie, and P.G. Schultz, *Expanding the genetic code.* Annu Rev Biophys Biomol Struct, 2006. **35**: p. 225-49.
14. Chin, J.W. and P.G. Schultz, *In vivo photocrosslinking with unnatural amino Acid mutagenesis.* Chembiochem, 2002. **3**(11): p. 1135-7.
15. Wang, Q., A.R. Parrish, and L. Wang, *Expanding the genetic code for biological studies.* Chem Biol, 2009. **16**(3): p. 323-36.
16. Xie, Y., C. Denison, S.H. Yang, D.A. Fancy, and T. Kodadek, *Biochemical characterization of the TATA-binding protein-Gal4 activation domain complex.* J Biol Chem, 2000. **275**(41): p. 31914-20.
17. Allard, S., R.T. Utley, J. Savard, A. Clarke, P. Grant, C.J. Brandl, L. Pillus, J.L. Workman, and J. Cote, *NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p.* Embo J, 1999. **18**(18): p. 5108-19.
18. Wang, L., A. Brock, B. Herberich, and P.G. Schultz, *Expanding the genetic code of Escherichia coli.* Science, 2001. **292**(5516): p. 498-500.
19. Shimogawa, H., Y. Kwon, Q. Mao, Y. Kawazoe, Y. Choi, S. Asada, H. Kigoshi, and M. Uesugi, *A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction.* J Am Chem Soc, 2004. **126**(11): p. 3461-71.
20. Park, J.M., H.S. Kim, S.J. Han, M.S. Hwang, Y.C. Lee, and Y.J. Kim, *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. **20**(23): p. 8709-19.
21. Wu, Z., G. Belanger, B.B. Brennan, J.K. Lum, A.R. Minter, S.P. Rowe, A. Plachetka, C.Y. Majmudar, and A.K. Mapp, *Targeting the transcriptional machinery with unique artificial transcriptional activators.* J Am Chem Soc, 2003. **125**(41): p. 12390-1.
22. Santoro, S.W., L. Wang, B. Herberich, D.S. King, and P.G. Schultz, *An efficient system for the evolution of aminoacyl-tRNA synthetase specificity.* Nat Biotechnol, 2002. **20**(10): p. 1044-8.
23. Chen, S., P.G. Schultz, and A. Brock, *An improved system for the generation and analysis of mutant proteins containing unnatural amino acids in Saccharomyces cerevisiae.* J Mol Biol, 2007. **371**(1): p. 112-22.
24. Chin, J.W., T.A. Cropp, J.C. Anderson, M. Mukherji, Z. Zhang, and P.G. Schultz, *An expanded eukaryotic genetic code.* Science, 2003. **301**(5635): p. 964-7.
25. Wang, Q. and L. Wang, *New methods enabling efficient incorporation of unnatural amino acids in yeast.* J Am Chem Soc, 2008. **130**(19): p. 6066-7.
26. Chen, H.T., L. Warfield, and S. Hahn, *The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex.* Nat Struct Mol Biol, 2007. **14**(8): p. 696-703.

27. Drabkin, H.J., H.J. Park, and U.L. RajBhandary, *Amber suppression in mammalian cells dependent upon expression of an Escherichia coli aminoacyl-tRNA synthetase gene.* Mol Cell Biol, 1996. **16**(3): p. 907-13.
28. Kowal, A.K., C. Kohrer, and U.L. RajBhandary, *Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria.* Proc Natl Acad Sci U S A, 2001. **98**(5): p. 2268-73.
29. Chin, J.W., A.B. Martin, D.S. King, L. Wang, and P.G. Schultz, *Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11020-4.
30. Chin, J.W., S.W. Santoro, A.B. Martin, D.S. King, L. Wang, and P.G. Schultz, *Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli.* J Am Chem Soc, 2002. **124**(31): p. 9026-7.
31. Brunner, J., *New photolabeling and crosslinking methods.* Annu Rev Biochem, 1993. **62**: p. 483-514.
32. Brunner, J., H. Senn, and F.M. Richards, *3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents.* J Biol Chem, 1980. **255**(8): p. 3313-8.
33. Kalkhof, S., C. Ihling, K. Mechtler, and A. Sinz, *Chemical cross-linking and high-performance Fourier transform ion cyclotron resonance mass spectrometry for protein interaction analysis: application to a calmodulin/target peptide complex.* Anal Chem, 2005. **77**(2): p. 495-503.
34. Schmidt, A., S. Kalkhof, C. Ihling, D.M. Cooper, and A. Sinz, *Mapping protein interfaces by chemical cross-linking and Fourier transform ion cyclotron resonance mass spectrometry: application to a calmodulin / adenylyl cyclase 8 peptide complex.* Eur J Mass Spectrom (Chichester, Eng), 2005. **11**(5): p. 525-34.
35. Sinz, A., *Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes.* J Mass Spectrom, 2003. **38**(12): p. 1225-37.
36. Sinz, A., *Investigation of protein-ligand interactions by mass spectrometry.* ChemMedChem, 2007. **2**(4): p. 425-31.
37. Galardy, R.E., L.C. Craig, and M.P. Printz, *Benzophenone triplet: a new photochemical probe of biological ligand-receptor interactions.* Nat New Biol, 1973. **242**(117): p. 127-8.
38. Galardy, R.E., L.C. Craig, J.D. Jamieson, and M.P. Printz, *Photoaffinity labeling of peptide hormone binding sites.* J Biol Chem, 1974. **249**(11): p. 3510-8.
39. Yin, H. and A.D. Hamilton, *Strategies for targeting protein-protein interactions with synthetic agents.* Angew Chem Int Ed Engl, 2005. **44**(27): p. 4130-63.
40. Tanaka, Y., M.R. Bond, and J.J. Kohler, *Photocrosslinkers illuminate interactions in living cells.* Mol Biosyst, 2008. **4**(6): p. 473-80.
41. Dorman, G. and G.D. Prestwich, *Benzophenone photophores in biochemistry.* Biochemistry, 1994. **33**(19): p. 5661-73.

42. O'Neil, K.T., S. Erickson-Viitanen, and W.F. DeGrado, *Photolabeling of calmodulin with basic, amphiphilic alpha-helical peptides containing p-benzoylphenylalanine.* J Biol Chem, 1989. **264**(24): p. 14571-8.
43. Wittelsberger, A., B.E. Thomas, D.F. Mierke, and M. Rosenblatt, *Methionine acts as a "magnet" in photoaffinity crosslinking experiments.* FEBS Lett, 2006. **580**(7): p. 1872-6.
44. Mohibullah, N. and S. Hahn, *Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3.* Genes Dev, 2008. **22**(21): p. 2994-3006.
45. Staros, J.V., *Aryl Azide Photolabels in Biochemistry.* Trends Biochem. Sci., 1980. **5**: p. 320-. 322.
46. Weinstein, S., B.A. Wallace, E.R. Blout, J.S. Morrow, and W. Veatch, *Conformation of gramicidin A channel in phospholipid vesicles: a 13C and 19F nuclear magnetic resonance study.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4230-4.
47. Knowles, J.R., *Photogenerated Reagents for Biological Receptor-Site Labeling.* Acc. Chem. Res., 1972. **5**: p. 155-160.
48. Kussie, P.H., S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A.J. Levine, and N.P. Pavletich, *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. **274**(5289): p. 948-53.
49. Hughes, T.R., *Yeast and drug discovery.* Funct Integr Genomics, 2002. **2**(4-5): p. 199-211.
50. Burke, D., D. Dawson, and T. Stearns, *Methods in Yeast Genetics. 1 ed. .* Cold Spring Harbor Laboratory Press: Cold Spring Habor, 2000.
51. 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.
52. Ma, J. and M. Ptashne, *The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80.* Cell, 1987. **50**(1): p. 137-42.
53. Ansari, A.Z., R.J. Reece, and M. Ptashne, *A transcriptional activating region with two contrasting modes of protein interaction.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13543-8.
54. Ma, J. and M. Ptashne, *Deletion analysis of GAL4 defines two transcriptional activating segments.* Cell, 1987. **48**(5): p. 847-53.
55. Brent, R. and M. Ptashne, *A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor.* Cell, 1985. **43**(3 Pt 2): p. 729-36.
56. Kakidani, H. and M. Ptashne, *GAL4 activates gene expression in mammalian cells.* Cell, 1988. **52**(2): p. 161-7.
57. Majmudar, C.Y., L.W. Lee, J.K. Lancia, A. Nwokoye, Q. Wang, A.M. Wands, L. Wang, and A.K. Mapp, *Impact of nonnatural amino acid mutagenesis on the in vivo function and binding modes of a transcriptional activator.* J Am Chem Soc, 2009. **131**(40): p. 14240-2.
58. Ebright, Y.W., Y. Chen, Y. Kim, and R.H. Ebright, *S-[2-(4-azidosalicylamido)ethylthio]-2-thiopyridine: radioiodinatable, cleavable,*

photoactivatible cross-linking agent. Bioconjug Chem, 1996. **7**(3): p. 380-4.
59. Geiger, M.W., Elliot, M. M., Karacostas, V. D., Moricone, T. J., Salmon, J. B., Sideli , V. L.,St. Onge, M. A., *ARYL AZIDES AS PROTEIN PHOTOLABELS: ABSORPTION SPECTRAL PROPERTIES AND QUANTUM YIELDS OF PHOTODISSOCIATION.* Photochemistry and Photobiology, 1984. **40**: p. 545-548.
60. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. **18**(3): p. 369-78.
61. Ansari, A.Z., S.S. Koh, Z. Zaman, C. Bongards, N. Lehming, R.A. Young, and M. Ptashne, *Transcriptional activating regions target a cyclin-dependent kinase.* Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14706-9.
62. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. **25**(20): p. 9092-102.
63. Majmudar, C.Y., J.K. Lum, L. Prasov, and A.K. Mapp, *Functional specificity of artificial transcriptional activators.* Chem Biol, 2005. **12**(3): p. 313-21.
64. Jeong, C.J., S.H. Yang, Y. Xie, L. Zhang, S.A. Johnston, and T. Kodadek, *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. **40**(31): p. 9421-7.
65. Melcher, K. and S.A. Johnston, *GAL4 interacts with TATA-binding protein and coactivators.* Mol Cell Biol, 1995. **15**(5): p. 2839-48.
66. Ptashne, M. and A. Gann, *Gene and Signal.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2002.
67. Thoden, J.B., L.A. Ryan, R.J. Reece, and H.M. Holden, *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. **283**(44): p. 30266-72.
68. Goodman, R.H. and S. Smolik, *CBP/p300 in cell growth, transformation, and development.* Genes Dev, 2000. **14**(13): p. 1553-77.
69. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
70. Takimoto, J.K., Z. Xiang, J.Y. Kang, and L. Wang, *Esterification of an unnatural amino acid structurally deviating from canonical amino acids promotes its uptake and incorporation into proteins in mammalian cells.* Chembiochem, 2010. **11**(16): p. 2268-72.
71. Brustad, E., M.L. Bushey, A. Brock, J. Chittuluru, and P.G. Schultz, *A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins.* Bioorg Med Chem Lett, 2008. **18**(22): p. 6004-6.
72. Young, T.S., I. Ahmad, J.A. Yin, and P.G. Schultz, *An enhanced system for unnatural amino acid mutagenesis in E. coli.* J Mol Biol, 2010. **395**(2): p. 361-74.

73. Kerem, E., S. Hirawat, S. Armoni, Y. Yaakov, D. Shoseyov, M. Cohen, M. Nissim-Rafinia, H. Blau, J. Rivlin, M. Aviram, G.L. Elfring, V.J. Northcutt, L.L. Miller, B. Kerem, and M. Wilschanski, *Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial.* Lancet, 2008. **372**(9640): p. 719-27.
74. Finkel, R.S., *Read-through strategies for suppression of nonsense mutations in Duchenne/ Becker muscular dystrophy: aminoglycosides and ataluren (PTC124).* J Child Neurol, 2010. **25**(9): p. 1158-64.
75. Rowe, S.M. and J.P. Clancy, *Pharmaceuticals targeting nonsense mutations in genetic diseases: progress in development.* BioDrugs, 2009. **23**(3): p. 165-74.
76. Welch, E.M., E.R. Barton, J. Zhuo, Y. Tomizawa, W.J. Friesen, P. Trifillis, S. Paushkin, M. Patel, C.R. Trotta, S. Hwang, R.G. Wilde, G. Karp, J. Takasugi, G. Chen, S. Jones, H. Ren, Y.C. Moon, D. Corson, A.A. Turpoff, J.A. Campbell, M.M. Conn, A. Khan, N.G. Almstead, J. Hedrick, A. Mollin, N. Risher, M. Weetall, S. Yeh, A.A. Branstrom, J.M. Colacino, J. Babiak, W.D. Ju, S. Hirawat, V.J. Northcutt, L.L. Miller, P. Spatrick, F. He, M. Kawana, H. Feng, A. Jacobson, S.W. Peltz, and H.L. Sweeney, *PTC124 targets genetic disorders caused by nonsense mutations.* Nature, 2007. **447**(7140): p. 87-91.
77. McCaughan, K.K., C.M. Brown, M.E. Dalphin, M.J. Berry, and W.P. Tate, *Translational termination efficiency in mammals is influenced by the base following the stop codon.* Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5431-5.
78. Manuvakhova, M., K. Keeling, and D.M. Bedwell, *Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system.* RNA, 2000. **6**(7): p. 1044-55.

CHAPTER 3
DISCOVERY OF ENDOGENOUS TARGETS OF TRANSCRIPTIONAL ACTIVATORS*

A. Chapter overview

Transcriptional activators are engaged in numerous interactions with a variety of protein partners, ranging from masking proteins or suppressors to multi-protein coactivator complexes (Figure 3-1). While the interactions between an activator and a masking protein are often high affinity, activators also engage in a series of moderate affinity, transient interactions with the array of protein targets within the transcriptional machinery.[1, 2] The majority of the current methodologies available to study protein-protein interactions are ideally suited for high-affinity interactions; therefore, the relationships between activators and their masking proteins are well-characterized. However, few methods exist that allow for the capture of transient protein-protein interactions in their native context. As a result, the interactions between activators and the transcriptional machinery are still poorly defined, thus demonstrating clearly the need for new methodologies for studying transient and/or moderate affinity PPIs in vivo.

* Portions of this Chapter have been published. Krishnamurthy, M., A. Dugan, A. Nwokoye, Y.H. Fung, J.K. Lancia, C.Y. Majmudar, and A.K. Mapp, *Caught in the act: covalent crosslinking captures activator-coactivator interactions in vivo.* ACS Chem Biol, 2011.

In Chapter 2, in vivo photo-crosslinking was used to capture the high-affinity and well-characterized binding interaction between the Gal4 TAD and its inhibitory protein Gal80. In this chapter, we will use in vivo photo-crosslinking again to capture the moderate-affinity binding interactions of Gal4 with Med15, Tra1, Taf12 and Snf2. Additionally, we extend these studies to include two other amphipathic activators, VP16 and Gcn4, and examine if these activators target the same subunits within shared coactivator complexes.

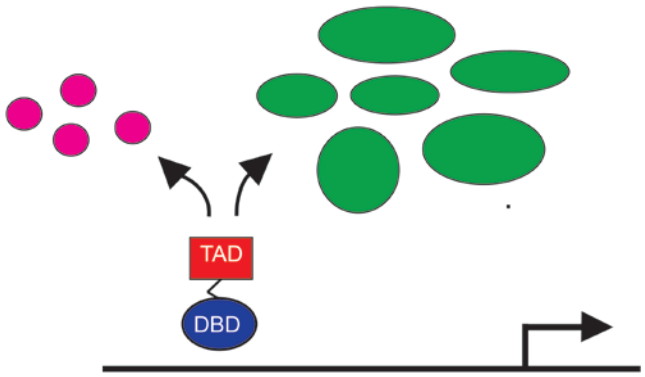


Figure 3-1: A schematic showing the different classes of interactions of transcriptional activators. TADs can participate in high affinity interactions with repressive masking protein (pink circles) as well as with coactivator complexes (green circles) to initiate transcription. TAD-masking protein interactions are generally well-characterized; However, the more moderate -affinity, transient interactions between TADs and the transcriptional machinery are not well characterized.

B. Significance

Many diseases result from misregulated gene transcription, often the consequence of a faulty protein network.[3-5] Therefore, researchers have a vested interest in designing small molecules that serve as transcriptional modulators to reconstitute lost interactions.[6-9] Nevertheless, the discovery of small molecule modulators has been hindered by a lack of structural and mechanistic information regarding the transient and multi-partner binding profile of transcriptional activators. Techniques such as co-crystallization and co-

purification have limited utility for studying transient PPIs in their native environments. This is because these techniques are best suited for probing stably associated proteins and are less ideal for studying proteins that engage in modest-affinity and/or transient multi-protein binding interactions, such as those between an activator and the transcriptional machinery.[10-12] Biochemical and genetic experiments have been used to identify several targets of the amphipathic activator Gal4, including TBP, TFIIB, Med15, Cdk8, Tra1, Taf12, Srb4 and Sug1 amongst others. [13-24] However, these studies have not successfully distinguished which of these speculated targets are indeed the binding partners in vivo. Thus, there is a clear need for in vivo methodologies that can capture transient activator-coactivator interactions in their native environment.

C. Targeting the transcriptional machinery during recruitment

As discussed previously in Chapter 1, transcriptional activators are modular proteins with a DNA binding domain (DBD) that localizes the activator to DNA and a transcriptional activation domain (TAD) that mediates the majority of contacts with various coactivator complexes. Activators assemble the transcriptional machinery at the promoter of a gene through a series of binding interactions with a variety of protein complexes (coactivators), including scaffolding proteins, helicase and chromatin-modifying complexes.[25, 26]

Chromatin immunoprecipitation (ChIP) experiments focused on the Gal1 promoter have demonstrated a sequential recruitment of coactivator complexes by the amphipathic activator, Gal4. First, Spt-Ada-Gcn5-acetyltransferase

(SAGA), a chromatin remodeling complex is recruited, followed by the SAGA-independent recruitment of Mediator and finally RNA polymerase II along with the general transcriptional machinery.[16] Additionally in other experiments, TADs were able to recruit the SWI/SNF complex, an ATP-dependent nucleosome remodeling enzyme, to promoters.[27, 28] Thus, from all these studies, it is apparent that activators like Gal4 recruit complexes such as Mediator, SAGA and SWI/SNF. However, the identity of the direct targets within these complexes and the extent of these interactions in cells remain unclear.

C.1 Mediator as a target

The Mediator scaffolding complex (Figure 3-2) is modular, consisting of a head (8 proteins), middle (8 proteins) and tail module (5 proteins) and serves as a bridge between activators and RNA polymerase II.[29 30, 31] The Mediator was originally discovered in yeast and is conserved in metazoans. [30, 31, 29 27, 28] Studies from Asturias and coworkers and Tijian and coworkers strongly support that the Mediator undergoes several conformational changes when contacted by activators.[32-34] In addition to the head domain component Med17(Srb4), several other proteins from the tail module including Med14(Rgr1), Med3(Pgd1), Med2 and Med15(Gal11) have been identified to influence transcription.[35-37]

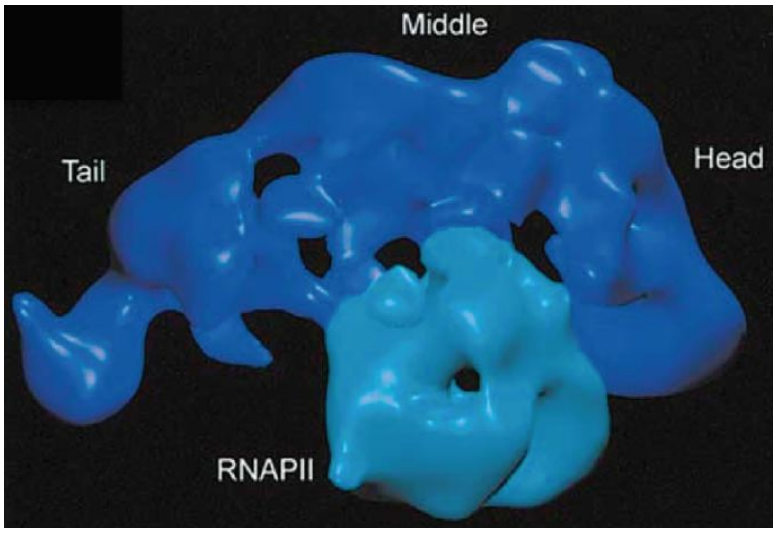


Figure 3-2: An EM image of Mediator in complex with RNA pol II.[59] The Mediator complex is composed of ~30 proteins divided into three regions, the head, middle, and tail.

Of this group of proteins, Med15 in particular stands out as a likely direct binding partner of transcriptional activators. Med15 is a 120 kDa protein that contains a N-terminus GACKIX domain between residues 2-93, and a glutamine rich region.[38] Importantly, the C-terminus of Med15 has been observed to associate with general transcription factors.[38, 39] In a study by Gaudreau et. al., various components of Mediator was localized to DNA to determine which Mediator proteins are sufficient for transcription to occur. From this elegant 'activator by-pass' experiment, it was revealed that Med15 was able to activate transcription two orders of magnitude higher than any other Mediator component and to similar levels as the potent yeast activator Gal4.[40]

Although the function of Med15 is not fully known, a number of genetic and biochemical studies have led scientists to believe that it is a likely target of DNA-bound transcriptional activators.[18] In addition to Gal4, Med15 has been

shown to be essential for activated transcription by the activators Gcn4, Swi5, Msn2, VP16 and Met4.[41-46] In fact, using direct binding experiments and integrated yeast reporter assays, fragments of Med15 were shown to interact with a number of natural TADs in vitro and deleting the fragments of Med15 resulted in a decreased activation in yeast.[18, 35, 36] Supporting this, Hahn and coworkers performed in vitro crosslinking studies with yeast whole cell extracts and identified Med15 to be a target of both Gal4 and Gcn4.[47, 48] Furthermore, Med15 has been implicated as a target of amphipathic activators through in vitro crosslinking and pull-down assays.[20, 47-49] Together, these studies suggest that Med15 is a probable target of amphipathic activators.

C.2 SAGA as a target

Besides the Mediator complex, the complex SAGA (Figure 3-3) has been implicated to interact with transcriptional activators. The SAGA complex is made up of TBP-associated factors (TAFs) as well as histone acetyltransferase (HATs) which modify histone proteins in the nucleosome.[50] The SAGA complex is also involved in recognition of methylated histone H3, contacting TATA-binding protein (TBP) and deubiquitinating several transcription proteins.[51]

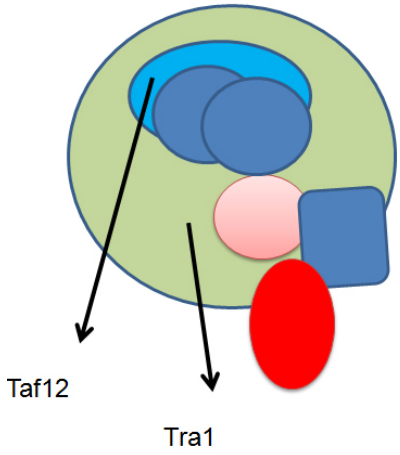


Figure 3-3: A schematic of the SAGA complex, which houses the proteins Tra1 and Taf12, implicated as TAD binding partners.

During transcription, the SAGA subunits Tra1 and Taf12 are believed to directly contact amphipathic activators. [47, 48] In fact, the TADs of Hap4, Gal4, Gcn4 and VP16 where shown to interact with purified Tra1, as determined by GST-pull-down assays.[52] Furthermore, the activator Hap4 was incubated with reconstituted SAGA complex in a cell free system and it specifically crosslinked to Tra1.[52] Additional experiments from Hahn and coworkers helped to support that Tra1 was a target of activators. In this in vitro crosslinking experiment, using yeast nuclear extracts, the Gal4 and Gcn4 TADs were shown to interact with Tra1.[47, 48] Furthermore, experiments carried out by the Green laboratory using in vivo FRET-based experiments identified Tra1, in the context of SAGA, as the cellular target of the activator Gal4 in live yeast.[15, 53] These experiments strongly support Tra1 to be a target of activators.

Another SAGA protein, Taf12, was shown to crosslink to TADs of both Gal4 and Gcn4.[47, 48] Taf12 is made up of 539 amino acids and is believed to interact with binding partners through a sequence spanning residues 250-260.[54]

Although much is not known about the role Taf12 plays to influence the output of transcription, experimental evidence suggests that this subunit is important for recruitment of SAGA and TFIID, of which it is also a subunit.

C.3 SWI/SNF as a target

The Swi/Snf chromatin-modifying complex (Figure 3-4) is an ATP-dependent nucleosome remodeling enzyme that has also been proposed to be important for activator function and has been shown to directly contact activators in vitro.[27, 37, 55-61 62] Studies carried out by Workman and coworkers found that Gcn4, Gal4 and VP16 were able to recruit the SWI/SNF complex to several promoters in vivo.[27, 28] Additional in vivo and in vitro colocalization studies show that VP16 recruits the Swi/Snf chromatin-remodeling complex early in transcription initiation.[27, 62, 63] In addition to the amphipathic activators Gal4 and VP16, Gcn4 has been shown through in vivo and in vitro binding studies to also recruit the Swi/Snf complex to a variety of promoters, implying that these activators are making contacts within the Swi/Snf complex.[27, 37, 56-58, 60, 61, 64] Experimental evidence through in vitro assays exists to further support that within the Swi/Snf complex, several subunits are possible targets of transcriptional activators.[56, 57, 64]

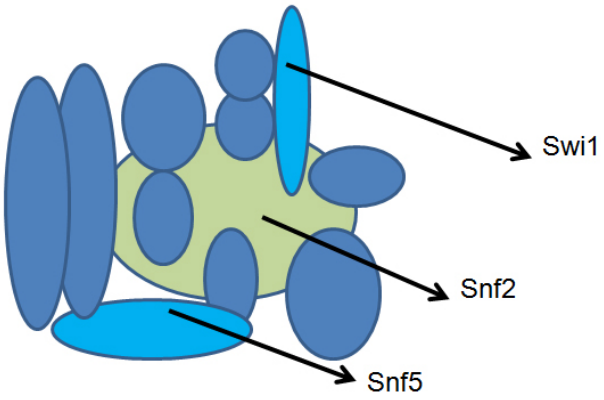


Figure 3-4: A schematic diagram of the Swi/Snf complex, three subunits from this complex - Snf2, Swi1 and Snf5 have been implicated as TAD binding partners.

One such possible target is the Snf2 ATPase, which is vital for Swi/Snf function and is highly conserved among eukaryotes.[65] This catalytic subunit, through structural studies of Swi/Snf in complex with the nucleosome, has been suggested to be positioned close to the activator.[28, 66, 67] In vitro studies have further suggested that the subunits Swi1 and Snf5 are additional possible activator targets.[56, 57, 64] Even though several subunits have been proposed based on in vitro studies, there is little support as to which subunits serves as the activator-binding subunit in vivo.

D. Results and Discussions

D.1. Med15 is a direct partner of Gal4

Activators participate in dynamic binding interactions with coactivator complexes that are characterized as modest-affinity and/or transient in nature.[55, 64, 68, 69] One of the Mediator proteins, Med15, has been backed by a strong body of in vitro and in vivo experimental evidence from our laboratory that suggests it has direct contacts with activators.[68, 70] Furthermore, the interaction between

Med15 and activators such as Gal4, Gcn4 and VP16 has been characterized as being moderate in affinity (Figure 3-5).[68, 70, 71]

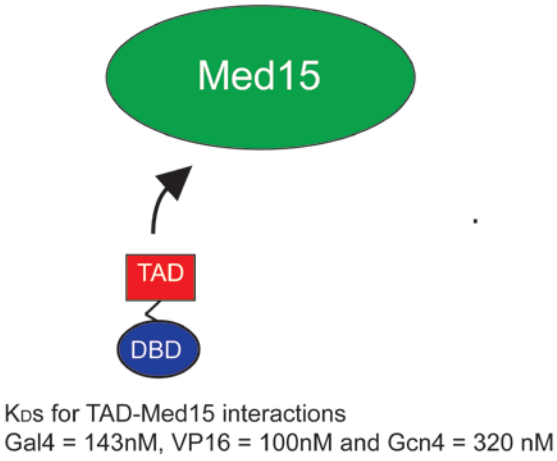


Figure 3-5: Med15 interactions with isolated TADs. Kinetic rate constants have determined that TADs interact transiently with the coactivator Med15 and equilibrium binding measurements place the affinity of the TAD for Med15 in the moderate category.[72, 88]

Thus, Med15 appeared to be an excellent test case to determine if the in vivo photo-crosslinking strategy will be effective for capturing moderate affinity binding interactions. Because in vitro crosslinking experiments performed by Dr Majmudar from our lab revealed contacts between the N-terminus fragment of Med15, (residues 1-416) and TADs of Gal4 and Gcn4, we hypothesized that Med15(1-416) is a probable target of Gal4 in vivo.

Before determining the ability of the Gal4 TAD to bind Med15(1-416), we carried out crosslinking experiments under raffinose and galactose conditions to find out if pBpa is readily incorporated under raffinose and galactose conditions and also to assess change in the pattern of formed crosslink-complexes as compared to growth under glucose conditions. In this experiment, plasmids expressing the LexA-Gal4867TAG mutation were co-transformed with pSNR

tRNA/synthetase pair plasmid in LS41 yeast strain. The yeast cells were grown in media containing 2% raffinose + 2% galactose with or without 2 mM pBpa. Our anticipation was that under raffinose/galactose growth conditions pBpa would be successfully incorporated into the Gal4 TAD. Additionally, the crosslinked Gal4 complex profile would be different since Gal4 is no more repressed by Gal80 (in glucose conditions) and so Gal4 has the ability to interact with other complexes. As expected, the Gal4 complex profile was different under raffinose/galactose condition and additionally there were more complexes formed (Figure 3-6B).[61, 67]

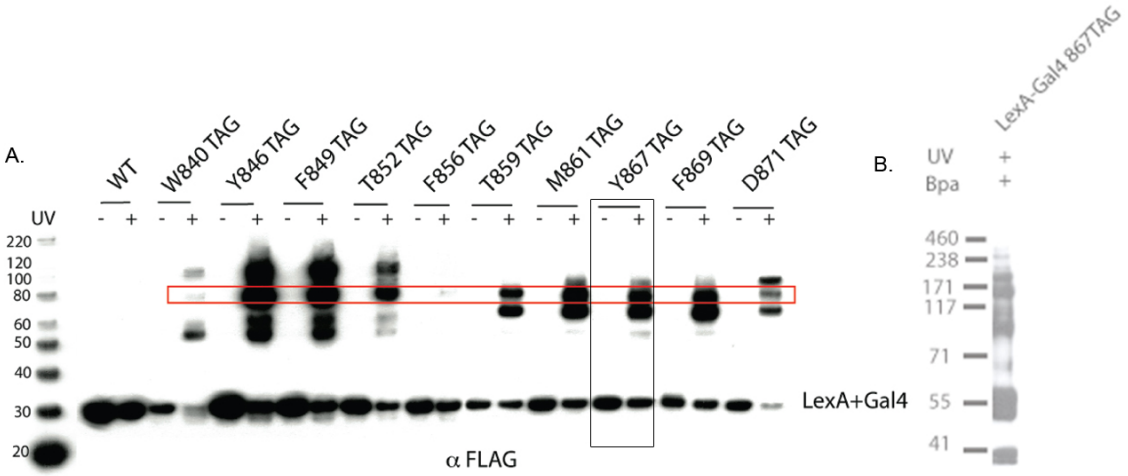
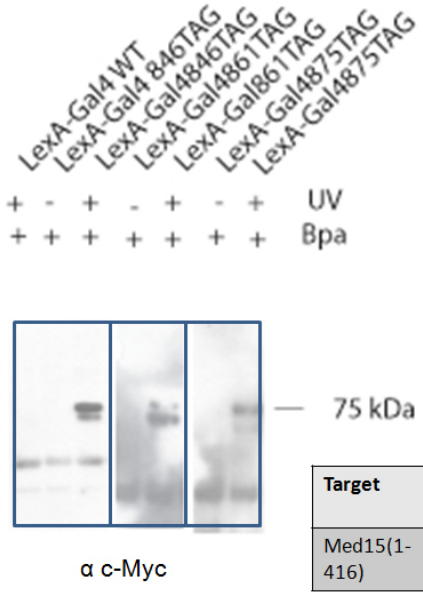


Figure 3-6: In vivo photo-crosslinking experiments with Gal4 containing pBpa mutants for different sugar conditions. Crosslinked complexes are only observed with application of UV light and pBpa. (A.) Experiments were carried out in glucose sugar conditions and (B.) experiments were carried out in raffinose/galactose sugar conditions. A comparison of both sugar conditions for Gal4867 (red box in A.) and B. show that the Gal4 complex profiles are different and more Gal4 complexes are seen in raffinose/galactose conditions.

Next we tested the ability for Gal4 TAD to crosslink with Med15 (1-416) in vivo. We carried out our experiment under raffinose and galactose conditions to increase the chances for Med15(1-416) to interact with uninhibited Gal4 TAD. In this experiment, LexA-Gal4TAG mutants were co-expressed with Myc-tagged

Med15(1-416) and subsequently crosslinking experiments were carried out. Results show that Med15 indeed interacts directly with Gal4 (Figure 3-7).

Previously in chapter 2, the Gal4 TAD was shown to crosslink to Gal80. Contrary to reported literature until this time, Gal4 in our crosslinking experiments showed an extended binding sequence beyond residues 851-871(Figure 3-6A red box).[49, 71] In similar analysis, we wanted to determine if Gal4 interacted with Med15 with an extended residue sequence. In Figure 3-7, Med15 interaction with Gal4 again extends beyond residues 851-871 on the Gal4 TAD. It is possible that the length of the Gal4 TAD sequence that participates in binding to Med15 may be important in gene up-regulation.



Target	Target MW	Target + Gal4 MW
Med15(1-416)	47	75

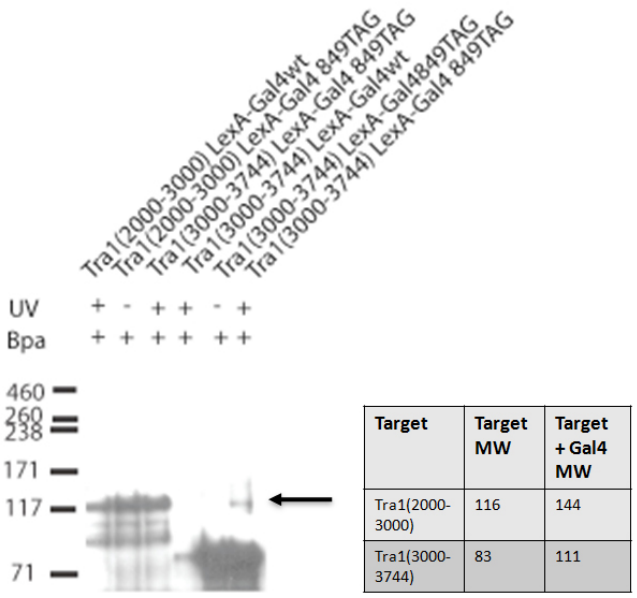
Figure 3-7: In vivo photo-crosslinking captures the moderate affinity interaction between LexA-Gal4 fusion protein and the Mediator protein, Med15. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion proteins containing pBpa at positions 846, 861 and 875, in addition to a plasmid expressing Myc-Med15(1-416) were irradiated with UV light (365 nm) for I hr. Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). For all constructs, a crosslink with Med15(1-416) was observed. The MW for a Gal4-Med15 complex is ~ 75kDa

D.2. Tra1 is a direct partner of Gal4

Based on the result of Med15, we wanted to tease out other modest-affinity and/or transient protein binding interactions. In vitro and in vivo experimental evidences exist to support that components of the SAGA complex directly contact activators.[14, 47, 48] In fact, the Hahn group reported that Tra1 (component of SAGA) and Taf12 (a shared subunit of SAGA and TFIID) crosslinked to two TADs.[47, 48] Starting with Tra1, which is 433 kDa, we wanted to investigate if Tra1 directly contacts Gal4. Previously, Mapp and coworkers explored the multiple functional activator binding sites within the Tra1 protein. In this study, they created ~50 kDa fragments of Tra1 spanning amino acids 1900-3500, which was expressed in E. coli as maltose binding protein (MBP) fusion proteins. Using fluorescence polarization, their work suggested that only Tra1 (3092-3524) had any measurable affinity for the TADs used in this study.[72] Given this experimental evidence that suggests that Tra1 (3092-3524) binds to activator TADs, we carried out crosslinking experiments with Tra1 residues 3000-3744, which spans Tra1 (3092-3524) and another Tra1 fragment Tra1(2000-3000) from the sequence region that did not show binding to TADs.

To investigate if Tra1 directly contacts the Gal4 TAD, LexA-Gal4 residue849TAG mutant was co-expressed with Myc-tagged Tra1 (2000-3000) or Tra1 (3000-3744) in yeast cells and grown in media containing 2% raffinose + 2% galactose with or without 2 mM pBpa. After crosslinking experiments were carried out we were able to visualize a crosslink between Gal4TAD and Tra1 (3000-3744) but not for Tra1 (2000-3000) construct (Figure 3-8). This result

further supports that the Tra1 (3000-3744) sequence within Tra1 is important for binding to TADs like Gal4 in vivo. Additionally, the results from this experiment signal that the activator's interaction with Tra1 may be responsible for SAGA recruitment to the promoter.[47]



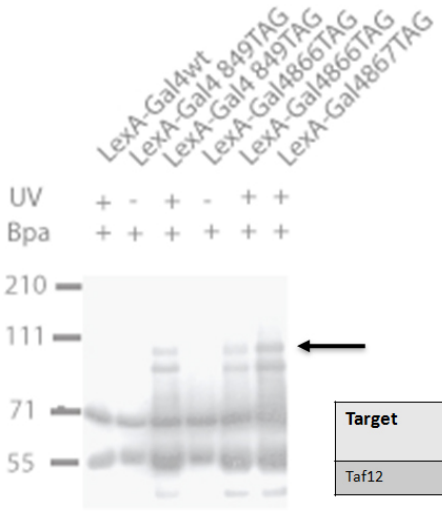
Target	Target MW	Target + Gal4 MW
Tra1(2000-3000)	116	144
Tra1(3000-3744)	83	111

Figure 3-8: In vivo photo-crosslinking captures interaction between LexA-Gal4 fusion protein and the SAGA subunit, Tra1. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion proteins containing pBpa at position 849, in addition to a plasmid expressing Myc-Tra1(2000-3000) and (3000-3744) were irradiated with UV light (365 nm) for I hr. Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). Between the two Tra1 constructs, a crosslink with Tra1(3000-3744) was observed. The MW for a Gal4-Tra1(3000-3744) complex is ~ 111kDa.

D.3. Taf12 is a direct partner of Gal4

Another protein that has been shown through in vitro experiments to crosslink to activators is the protein Taf12, which equally associates with the SAGA complex.[47, 48] In an in vitro crosslinking experiment, Taf12 was reported to interact with two prototypical activators, Gal4 and Gcn4.[47, 48] Similar to in vivo photo-crosslinking experiments carried out with Tra1, we wanted to assess if the

Taf12 protein was also a direct binding protein of Gal4. In order to do this, LexA-Gal4TAG mutants were co-expressed with full length Myc-tagged Taf12 in yeast cells and grown in media containing 2% raffinose + 2% galactose with or without 2 mM pBpa, after which crosslinking experiments were carried out.



Target	Target MW	Target + Gal4 MW
Taf12	69	97

Figure 3-9: In vivo photo-crosslinking captures the moderate affinity interaction between LexA-Gal4 fusion protein and Taf12. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion proteins containing pBpa at positions 849, 866 and 867, in addition to a plasmid expressing full length Myc-Taf12 were irradiated with UV light (365 nm) for I hr. Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). For all constructs, a crosslink with Taf12 was observed. The MW for a Gal4-Taf12 complex is ~ 97kDa.

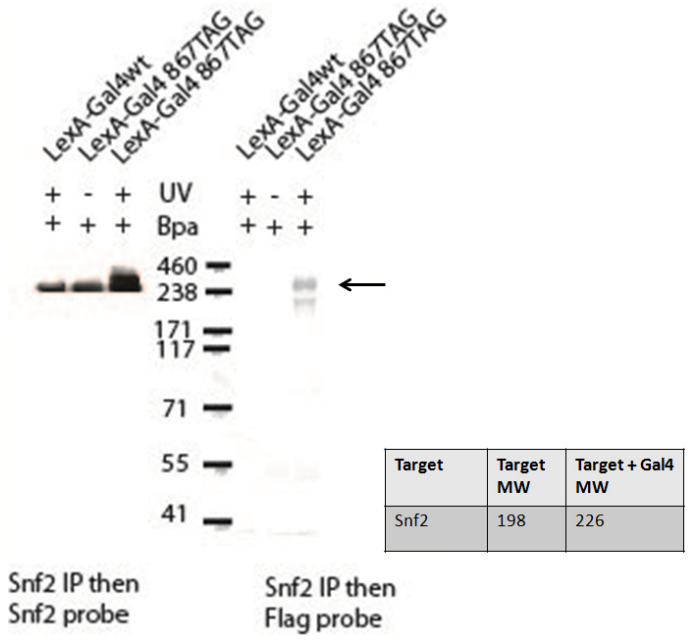
Results from our crosslinking experiments showed that Taf12 crosslinks to all the Gal4 TAD mutants tested (Figure 3-9). Figure 3-9 also shows that Taf12 interacts with the residues in the middle of the Gal4 TAD. Based on the results from this experiment, one could also propose mechanistically that apart from Tra1 or in addition to Tra1, activators utilize interactions with Taf12 to recruit the SAGA complex during transcription.

*D.4. Snf2 is a direct partner of Gal4**

Components of the Swi/Snf complex have been suggested to interact with transcriptional activators. [15, 35, 56, 57, 65, 82-85] For example, in vitro assays have suggested that Snf2, Swi1 and Snf5, subunits within the Swi/Snf complex, to be possible targets of TADs. [56, 57, 64] Therefore, we wanted to use in vivo photo-crosslinking experiments to find subunits within the Swi/Snf complex that directly contacts activators. Since, the in vivo crosslinking strategy (with the Gal4 genetically incorporated, photo-labile amino acid pBpa) was demonstrated to serve as a useful method for capturing a direct high affinity protein-protein interaction (chapter 2) and a known transient protein-protein interaction (Chapter 3, section D.1.) we hypothesized that this strategy will be applicable to the Swi/Snf complex.[73]

Given that structural studies of Swi/Snf in complex with the nucleosome suggested that the catalytic subunit Snf2 is closely positioned to the activator, first direct interaction of Snf2 and Gal4 in vivo was determined. [28, 66, 67] Similar crosslinking experiments as reported earlier in this section were carried out. In these experiments the ability for Gal4 to capture a native binding partner at endogenous levels was also tested, therefore Snf2 was not co-expressed. Briefly, live yeast cells expressing the LexA-Gal4 Phe867Bpa were photo-crosslinked, and then the Gal4-Snf2 complexes were immunopurified with a Snf2 antibody before the complex was visualized with the immuno-detectable FLAG probe. Our results (shown in Figure 3-10) revealed that Snf2 directly contacts our Gal4 TAD construct. This experiment not only shows for the first time that

Snf2 interacts with Gal4 in vivo but we also demonstrated that in vivo photo-crosslinking strategy was appropriate for capturing activator targets at their native concentrations. On the mechanistic end, it is possible that the interaction between Gal4 and the Swi/Snf complex is mediated by Snf2.



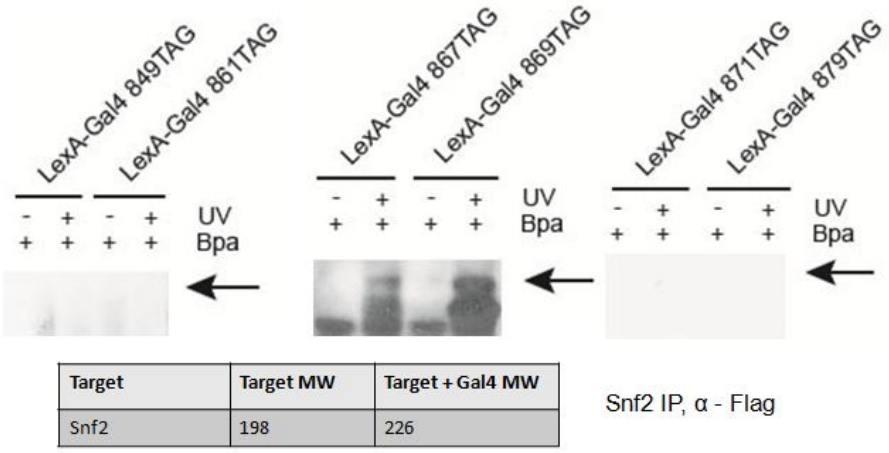
Target	Target MW	Target + Gal4 MW
Snf2	198	226

Figure 3-10: In vivo photo-crosslinking captures interaction between LexA-Gal4 fusion protein and Snf2. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion protein containing pBpa at positions 867, was irradiated with UV light (365 nm) for I hr. Subsequently the cell lysates were immunoprecipitated with an antibody to Snf2 and resolved by Western blot (α-FLAG) or (Myc). A crosslink with Snf2 was observed. The MW for a Gal4-Snf2 complex is ~ 226kDa.

Next, we wanted to determine if Swi1 and Snf5, other proposed binding partners from the Swi/Snf complex, indeed directly contact Gal4. Using a similar strategy employed for in vivo identification of Snf2, live yeast cells expressing the LexA-Gal4Phe867Bpa were photo-crosslinked, and then immuno-enriched with either a Swi1 or Snf5 antibody before being visualized with the immuno-detectable FLAG probe. Contrary to the result seen for the Snf2 experiments, no crosslinked products with Gal4 were detected. A possible interpretation for the

Swi1 and Snf5 experiments is that these two subunits do not directly contact Gal4, or the Swi1 and Snf5 antibodies used in these experiments were not capable of recognizing Gal4-Swi1 and Gal4-Snf5 complexes. In addition, these subunits have very low concentrations with about 100-500 copies per cell. Therefore, it is equally possible that the low expression levels of these coactivators subunits contributed to our results.

From our crosslinking experiments it was established that Snf2 was a direct binding partner of Gal4 and so further characterization of this interaction was carried out. First we wanted to know how much of the Gal4 TAD sequence participated in this interaction.



Target	Target MW	Target + Gal4 MW
Snf2	198	226

Figure 3-11: In vivo photo-crosslinking suggesting that Snf2 interacts with residues in the middle of the Gal4 TAD. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion protein containing pBpa at positions 849, 861,867, 869, 871 and 879 was irradiated with UV light (365 nm). Subsequently the cell lysates were immunoprecipitated with an antibody to Snf2 and resolved by Western blot (α-FLAG). Crosslinks were seen for only positions 867 and 869.

To test this, additional in vivo photo-crosslinking experiments with the Gal4 TAD containing pBpa incorporated at different positions were carried out. Figure 3-11 suggests that Gal4-Snf2 interaction is restricted to residues within the middle of

the Gal4 TAD, since this complex is not visualized when pBpa is incorporated at positions 849 and 861 (closer to the N terminus), and 871 and 879 (closer to the C terminus).

In additional efforts to further characterize Gal4-Snf2 interaction, some mutations that have been shown to affect Gal4 interaction with binding partners without negatively affecting its ability to activate transcription were made.[48, 74, 75] Again crosslinking experiments revealed that these mutations either did not affect Gal4's ability to interact with Snf2 or were not sufficient to disrupt binding to Snf2 (Figure 3-12). The latter is more consistent with other mutagenesis experiment that support that point mutations are not effective for abolishing the function of activators.[1, 75]

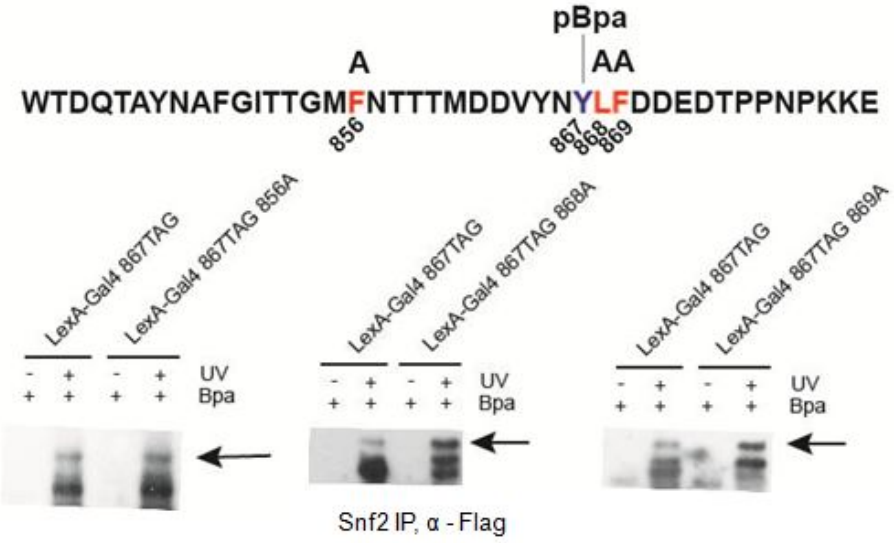


Figure 3-12: In vivo photo-crosslinking experiments, suggesting that alanine point mutations at TAD positions 856, 868 and 869 do not disrupt Gal4-Snf2 interaction. Live yeast cells bearing plasmids expressing LexA-Gal4 fusion protein containing pBpa at positions 867 and alanine mutations at positions 856 or 868 or 869 was irradiated with UV light (365 nm). Subsequently the cell lysates were immunoprecipitated with an antibody to Snf2 and resolved by Western blot (α-FLAG). Crosslinks to Snf2 was seen for the entire constructs, suggesting that the alanine mutations are not sufficient for disruption snf2 interaction with Gal4.

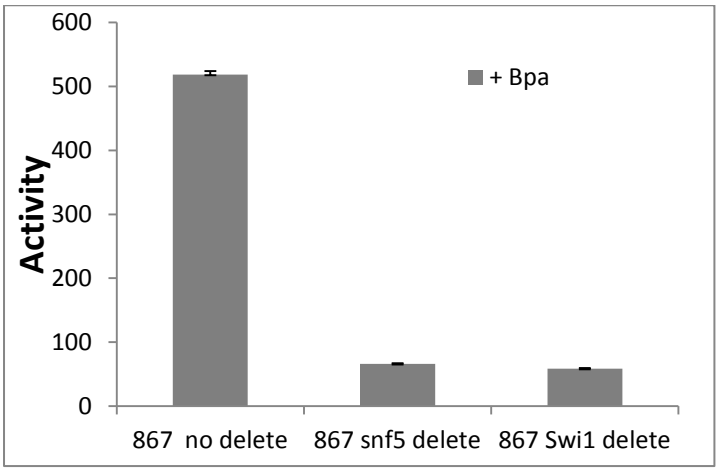


Figure 3-13: The functional impact of incorporating pBpa into LexA-Gal4867TAG in yeast non-delete, Snf5 delete and Swi1 delete LS41 strains. Liquid β-galactosidase assay was used to assess their ability to up regulate transcription in an integrated LacZ reporter gene in *S. cerevisiae.* Each activity is the average of values from at least three independent experiments with the indicated error (SDOM).

Gal4 interaction with Snf2 revealed that this Swi/Snf subunit may be the only direct binding partner and required interaction for an activator like Gal4 to engage the Swi/Snf complex. Furthermore, based on the functional results from the β-galactosidase assays of Figure 3-13, the subunits Swi1 and Snf5 may be carrying out other functions that are required for complex recruitment but not specifically for directly contacting the activator. This proposal is supported by experimental evidence from similar experiments with the VP16 TAD.[76] In this experiment, it was determined that the VP16 TAD does not interact with Snf5 but Snf5 may work cooperatively with Snf2 for recruitment of the Swi/Snf complex, with VP16 directly contacting Snf2.[76] In similar results, enrichment with either a Swi1 or Snf5 antibody did not result in any detectible crosslinked product for VP16. In addition, crosslinking experiment in yeast strains lacking either Swi1 or

Snf5, suggests that Swi1 is not a direct target of VP16.[76] The results with Med15, Tra1, Taf12 and Snf2 signify that a complete interaction map of the direct binding partners of transcriptional activators is possible and achievable.

D.5. Implications of shared and unshared activator targets

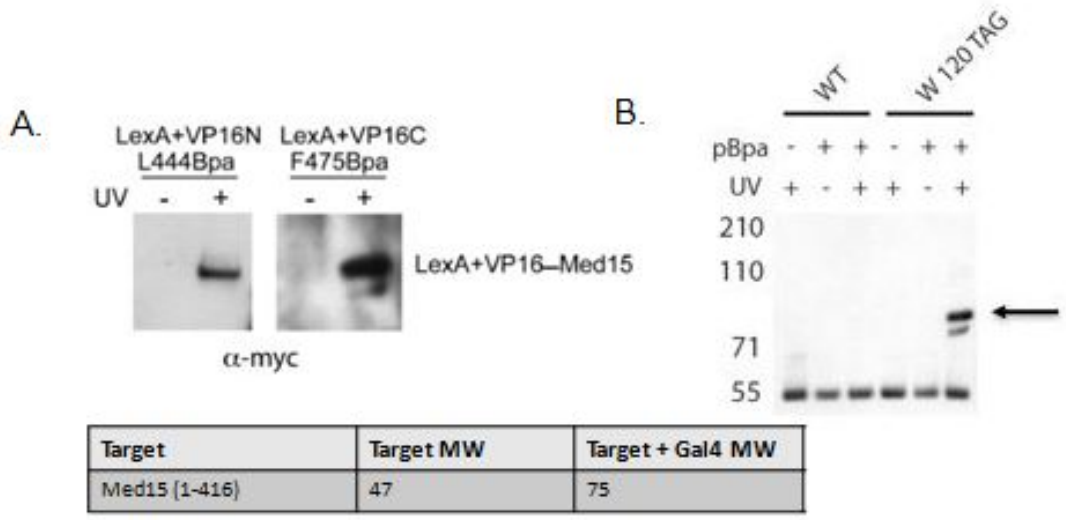
Transcriptional activators function by binding an array of transcriptional proteins, to facilitate assembly of the large (>50 proteins) complex needed for transcription. Therefore, there is a great possibility for transcriptional activators to use promiscuous binding sequences to interact with their binding partners also this style of binding may be conserved within a class of activators. [47, 74, 77] In fact, considerable evidence exists that different activators share at least a subset of co-activator binding partners in order to up-regulate transcription. As an example, the TADs of Gal4, Gcn4, VP16 and XLy have all been shown to bind to the co-activator Med15 despite considerable differences in sequence.[18, 36, 78, 79] In another example the multi-domain protein CBP, which has been implicated in several processes (including diseases) and serves to integrate signals in the nucleus,[80, 81] is highly conserved among metazoans, and it's KIX domain interacts with a set of amphipathic activators like Creb and p53.[82-84] Apart from the therapeutic advantages presented for understanding shared/unshared targets of TADs, scientists can gain detailed mechanistic understanding of how the different activator partners contribute to transcription.

Since there are in vitro experimental evidences that TADs bind the same targets, two other well studied amphipathic TADs VP16 and Gcn4, were assessed for direct contact with the targets of Gal4 that were identified through in

vivo photo-crosslinking. The Viral Protein 16 (VP16) has been proposed to interact with several targets (TFIID, SAGA amongst others).[77, 85] The VP16 TAD is unique because it comprises of two potent sub-domains, an amino terminal VP16N (residues 413-456) and a carboxyl terminal VP16C (residues 446-490) and they can function independently from one another.[86, 87] Another well studied amphipathic activator, Gcn4, has been shown to also bind to Tra1, Med15 and Taf12 in vitro.[47] Gcn4 is a member of the AP-1 transcription factor family, has been shown to be responsible for the induction of ~40 genes associated with the biosynthesis of amino acids and aminoacyl-tRNA synthetase.[88-91] VP16 and GCN4 are good activators for such comparative study because they have been shown to have overlapping targets with Gal4. A comparison of the three activators will help define the mechanistic conservation of amphipathic activator targets.

First, in vivo photo-crosslinking was used to determine whether Med15 is shared by VP16 and Gcn4 TADs. Previously, some of my colleagues, tested the ability for 3 constructs from the TAD sequences shown in Figure 3-14, LexA+VP16N444Bpa and LexA+VP16C475Bpa (for VP16) and LexA+Gcn4120Bpa (for Gcn4), to incorporate pBpa as well as maintain activator function. Their results revealed that these constructs were able to incorporate pBpa and functions as activators. Therefore, using these constructs, in vivo experiments were carried out to investigate if Med15 was a direct partner of VP16 and Gcn4 TADs as well. Briefly, the three TAD constructs (two for VP16 and one for Gcn4) were was co-expressed with Myc-tagged Med15(1-416) in

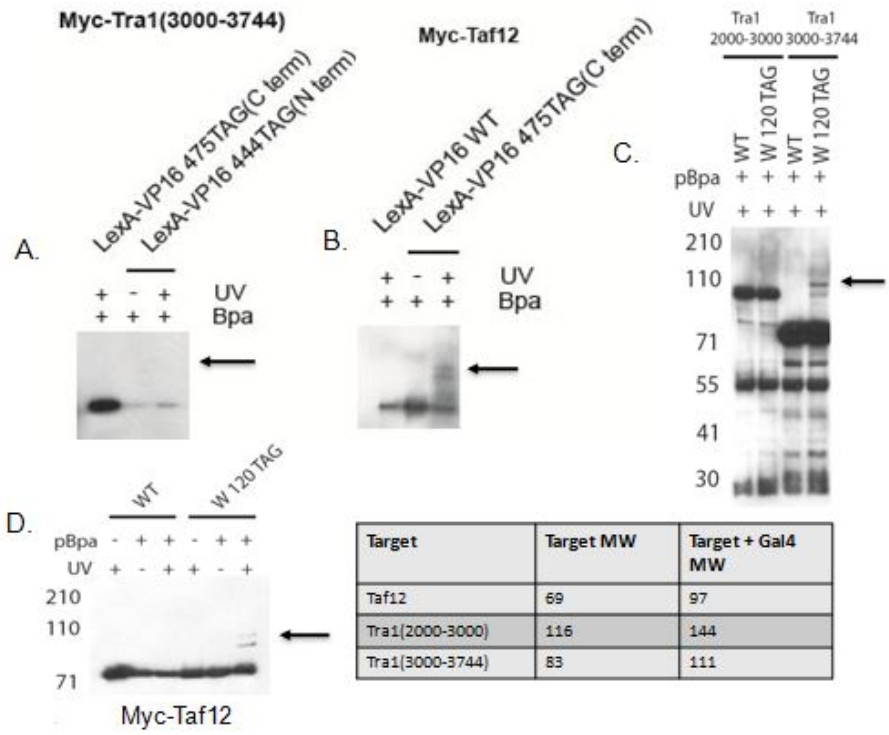
LS41 yeast strain and grown in media containing 2% raffinose + 2% galactose with or without 2 mM pBpa, after which, crosslinking experiments were carried out. Results from our crosslinking experiments showed that Med15(1-416) crosslinks to both VP16 and Gcn4 (Figure 3-14).



Target	Target MW	Target + Gal4 MW
Med15 (1-416)	47	75

Figure 3-14: In vivo photo-crosslinking captures the moderate affinity interaction between LexA-VP16N, LexA-VP16C and LexA-Gcn4 fusion proteins and the Mediator protein, Med15. Live yeast cells bearing plasmids expressing each TAD-Bpa construct, in addition to a plasmid expressing Myc-Med15(1-416) were irradiated with UV light (365 nm). Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). For all constructs, a crosslink with Med15(1-416) was observed. The MW for the LexA-TAD complexes with Med15 is ~ 75kDa.

Next we hypothesized that the in vivo photo-crossing experiment can be used to determine if Taf12, Tra1 and Snf2 are direct targets of VP16 and Gcn4.

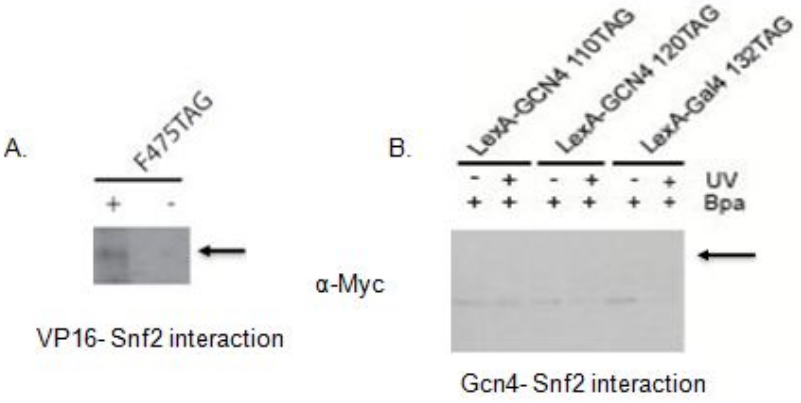


Target	Target MW	Target + Gal4 MW
Taf12	69	97
Tra1(2000-3000)	116	144
Tra1(3000-3744)	83	111

Figure 3-15: Using in vivo photo-crosslinking to determine the direct binding partners of VP16 and Gcn4 TADs with live yeast cells bearing plasmids that expressed three TAD-Bpa constructs. VP16N term and VP16C term were expressed, in addition to a plasmid expressing either (A.) Myc-Tra1(3000-3744) or (B.) full length Myc-Taf12, and Gcn4 TAD was expressed, in addition to a plasmid expressing (C.) Mcy-Tra1(2000-3000) or Mcy-Tra1(3000-3744) or (D.) full length Taf12. The cells were irradiated with UV light (365 nm) and subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). For all constructs, VP16 directly contacts Taf12 and not Tra1, while Gcn4 contacts both Tra1 and Taf12.

In the case of determining if Taf12 and Tra1 are shared targets of VP16 and Gcn4, crosslinking experiments were carried out with yeast cell expressing Myc version of full length Taf12 alongside LexA+VP16N444Bpa or LexA+VP16C475Bpa (for VP16) or LexA+Gcn4120Bpa and Tra1(3000-3744) alongside LexA+VP16N444Bpa or LexA+VP16C475Bpa (for VP16) or LexA+Gcn4120Bpa. Crosslinking experiments revealed that although both TADs directly contact Taf12, VP16 did not directly contact the Tra1 fragment used for these experiments (Figure 3-15).

Finally, crosslinking experiments were carried out to assess if our VP16 and Gcn4 constructs could directly contact the coactivator Snf2 at endogenous levels. In these experiments, the TAD constructs were expressed but Snf2 was not co-expressed. Formed TAD-Snf2 complexes were immunopurified with a Snf2 antibody before the complex was visualized with the immuno-detectable FLAG probe. Our results revealed that Snf2 directly contacts VP16 but not Gcn4 (Figure 3-16).



Target	Target MW	Target + Gal4 MW
Snf2	198	226

Figure 3-16: In vivo photo-crosslinking captures Snf2 as the direct binding partner of VP16 but not Gcn4. Live yeast cells bearing plasmids expressing VP16TAG475 or Gcn4 construct with Bpa at positions 110, 120 and 132 were irradiated with UV light (365 nm). Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-Myc). Results show that VP16 directly contacts Snf2 but Gcn4 does not, at all positions tested.

After identifying the coactivator targets that direct contact Gal4, VP16 and Gcn4 in vivo using photo-crosslinking, we can summarize from Figure 3-17 that all the TADs directly contact Med15 and Taf12 as a target, Gal4 and Gcn4 only directly contact Tra1 and Gal4 and VP16 only directly contact Snf2. Taken together, these results suggest that transcriptional activators do not always

contact the same targets and therefore do not entirely use the same mechanism for assembling coactivator complexes during transcription.

In conclusion, the in vivo experimental analysis outlined in this section gave mechanistic insight into binding interactions between transcriptional activation domains and co-activators and further, shared or overlapping direct targets found after comparing the binding partners of Gal4 with two other transcriptional activators, VP16 and Gcn4.

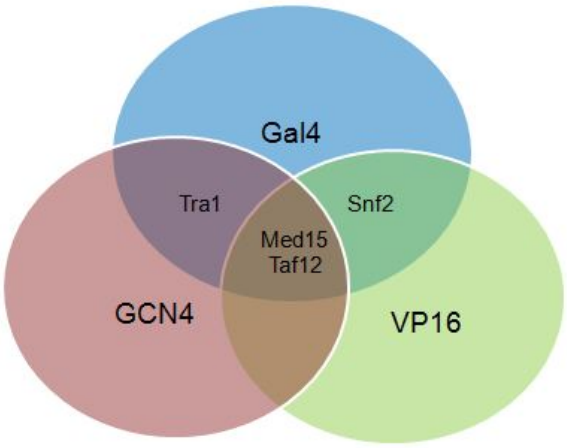


Figure 3-17: Summary of shared direct targets between Gal4, VP16 and Gcn4 based on our In vivo photo-crosslinking experiment.

In depth study of TAD interaction with Med15 will serve as an excellent model for understanding what mechanisms transcriptional activators utilize for interacting with a shared target. Such basic understanding will positively influence studies meant to screen for molecules that will serve to inhibit activators that contact the same coactivator. Similarly, the differences between the shared targets can be exploited for the development of inhibitors selective for a given activator(s). For example, Snf2, which did not interact with all 3

activators, is an excellent candidate for studying mechanistic preferences of activators that do not share same coactivators. Further, the ATPase Snf2 is highly conserved and may be a viable target for small molecule intervention in diseases with misregulated gene profile.[92-94] Finally, studies from this section serves as a test case for similar experiments that need to be accomplished to provide high level understanding for the development of efficient screens for small molecules transcriptional activator inhibitors and subsequently, therapies for diseases caused by misregulation in gene expression.

E. Methods

Yeast Strain LS41 [JPY9::pZZ41, *Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41] was used for the crosslinking experiments. Swi1 and Snf5 delete strains were made by gene disruption via PCR in LS41 and used for delete crosslinking experiments. pBpa was purchased from Chem-Impex International (Wood Dale, IL). All plasmids described below were constructed using standard molecular biology techniques. The sequences of all the isolated plasmids were verified by sequencing at the University of Michigan Core Facility (Ann Arbor, MI).

Table 3-1: Plasmids used in this study

Plasmid name	**Function**

pLexAGal4 846TAG, 849TAG, 861TAG, 866TAG, 867TAG, 869TAG, 871TAG, 875TAG, 879TAG and Wt	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid or Wt which does not contain a TAG
pLexAGal4 867TAG856A 867TAG868A, 869TAG856A	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid plus the appropriate alanine mutation
pLexAVP16N 439TAG, 442TAG, 444TAG, and Wt	Expresses LexA(1-202)+VP16 (446-490)+FLAG tag with a TAG replacing the codon of the existing amino acid or Wt which does not contain a TAG
pLexAVP16C 473TAG, 475TAG, 479TAG and Wt	Expresses LexA(1-202)+VP16 (413-456)+FLAG tag with a TAG replacing the codon of the existing amino acid or Wt which does not contain a TAG
pLexAGcn4 110TAG, pLexAGcn4 120TAG, 132TAG and Wt	Expresses LexA(1-202)+Gcn4(110-132)+FLAG tag with a TAG codon replacing the codon of the existing amino acid or Wt which does not contain a TAG

pSNRtRNA-pBpaRS	Expresses tRNA under the control of the SNR52 promoter and contains synthetase specific for pBpa ptRNA-pBpaRS
pMyc Med15(1-416)	Expresses Med15 (1-416) fused to c-Myc tag
pMyc Taf12	Expresses Taf12 fused to c-Myc tag
pMyc Tra1(2000-3000) and (3000-3744)	pMyc Tra1(2000-3000) and (3000-3744) each fused to a c-Myc tag

pLexAGal4 (840-881) TAGs and Wt

pLexA(1-202)+Gal4(840-881) TAGs was created as previously described (in Section F of Chapter 2) In the case of plasmids with alanine mutants, site-directed mutagenesis was used to replace the existing amino acid codon with GCT within the Gal4 TAD. In general, PCR primers were designed to have ~15 bases of homology on either side of the TAG mutation. Quick Change (Stratagene, La Jolla, CA) was used to incorporate the TAG mutants using manufacturer recommended conditions.

pLexAVP16 N and pLexAVP16C TAGs and Wt

A high copy plasmid expressing LexA(1-202)+VP16N (413-456)+FLAG tag and LexA(1-202)+VP16C (446-490))+FLAG tag under the control of the ADH1 promoter was created from pCLexA containing EcoRI and BamHI sites.

Primers 5'- catgaattcATGGCCCCCCCGACCGATGTC-3' and

5’catggatccTTACTTGTCATCGTCGTCCTTGTAGTCTCCCGGCCCCGGGGAATCCC-3’ were used to amplify VP16 (413-456) using pMVP16 as a template. The amplified PCR product was digested with EcoRI and BamHI and inserted into pCLexA digested with EcoRI and BamHI and calf intestinal phosphate treated to create pLexAVP16N. Primers 5’ catgaattcATGTTGGGGGACGGG- 3’ and (5’-catggatccTTACTTGTCATCGTCG -3’) were used to amplify VP16 (446-490) using pMVP16 as a template. The amplified PCR product was digested with EcoRI and BamHI and inserted into pCLexA digested with EcoRI and BamHI and calf intestinal phosphate treated to create pLexAVP16C. Plasmids containing various amber mutants in the VP16 TAD were derived from pLexAVP16N and pLexAVP16C. To create each plasmid, site-directed mutagenesis was used to replace an existing amino acid codon with TAG codon within the VP16C or VP16N TAD. In general, PCR primers were designed to have ~15 bases of homology on either side of the TAG mutation. QuikChange (Stratagene, La Jolla, CA) was used to incorporate the TAG mutants using manufacturer recommended conditions.

pLexAGcn4(110-132) TAGs and Wt

In a similar fashion to the VP16 plasmid construction, a high copy plasmid expression LexA(1-202)+Gcn4(107-144)+FLAG tag under the control of the ADH1 promoter was created from pCLexA containing EcoRI and BamHI sites.

Primers 5’-GAATTCATGTTTGAGTATGAAAACCTAGAAGACAACTC-3’ and 5’-GGATCCGGATTCA ATTGCCTTATCAGCCAATG-3’ were used to amplify

Gcn4(107-144) from yeast genomic DNA. The amplified product was digested with BamHI and EcoRI and then treated with Calf intestinal phosphatase to create pLexAGcn4. Plasmids containing various amber mutants in the Gcn4 TAD were derived from pLexAGcn4. To create each plasmid, site-directed mutagenesis was used to replace an existing amino acid codon with TAG codon within the Gcn4 TAD. In general, PCR primers were designed to have ~15 bases of homology on either side of the TAG mutation. QuikChange (Stratagene, La Jolla, CA) was used to incorporate the TAG mutants using manufacturer *recommended conditions.*

pMycMed15 (1-416)

A high copy plasmid pMycMed15(1-416) expressing Med15(1-416) under the ADH1 promoter, N-terminally tagged with the c-Myc epitope was constructed by amplifying the DNA sequence encoding Med15(1-416) from yeast genomic DNA using primers (5'-GACAGGATCCATGTCTGCTGCTCCTGTCCAAGAC-3') and (5'-CGATCATATGTCAC TGATATAATTTAGAACTTGC-3') and inserted into BamHI and NdeI digested pMyc using standard molecular biology techniques. The pMyc cloning vector was created by inserting an ADH1 driven c-myc epitope tag in pGADT7 (Clontech) followed by restriction sites for gene insertion using site-directed mutagenesis using primers (5'-AGCTATGGAACAAAAGTTGATT TCTGAAGAAGATTTGGGATCCAATGCATATGATCT-3') and (5'-AGCTTGATCA TATGCATTGGATCCCAAATCTTCTTCAGAAATCAACTTTTGTTCCAT-3').

pMycTaf12

A high copy plasmid pMycTaf12 expressing full length Taf12 under the ADH1 promoter, N-terminally tagged with the c-Myc epitope was constructed by amplifying the DNA sequence encoding Taf12 from yeast genomic DNA using primers (5'-GCCCATATGATGTCTTCCAATCCAGA-3') and (5'- CCG CCAT GGTTATTTTTTTGTATTCAA-3') and inserted into Nde1 and Nco1 digested pMyc using standard molecular biology techniques. The pMyc cloning vector was created by inserting an ADH1 driven c-myc epitope tag in pGADT7 (Clontech) followed by restriction sites for gene insertion using site-directed mutagenesis using primers (5'-AGCTATGGAACAAAAGTTGATTTCTGAAGAAGATTTGGG ATCCAATGCATATGATCT-3') and (5'-AGCTTGATCATATGCATTGGATCC CAA ATCTTCTTCAGAAATCAACTTTTGTTCCAT-3').

pMycTra1(2000-3000) and (3000-3744)

High copy plasmids pMycTra1(2000-3000) and (3000-3744) expressing Tra1(2000-3000) and Tra1(3000-3744) under the ADH1 promoter, N-terminally tagged with the c-Myc epitope was constructed by amplifying the DNA sequence encoding Tra1(2000-3000) from yeast genomic DNA using primers (5'-GCCCTCGAGAGAGATTTGTT CATATC-3') and (5'-CGGGGATCCTTATAACTC CCTAATCT-3') and Tra1(3000-3744) from yeast genomic DNA using primers (5'-GCCCTCGAAACAACGGGTCTG GATC-3') and (5'-CGGGGATCC TTAGAACCATGGCATGA-3') were inserted into Xho1 and BamHI digested pMyc using standard molecular biology techniques. The pMyc cloning vector was

created by inserting an ADH1 driven c-Myc epitope tag in pGADT7 (Clontech) followed by restriction sites for gene insertion using site-directed mutagenesis using primers (5'-AGCTATGGAACAAAAGTTGATTTCTGAAGAAGATTTGGG ATCC AATGCATATGATCT-3') and (5'-AGCTTGATCATATGCATTGGATCCCAA ATCTTCTTC AGAAATCAACTTTTGTTCCAT-3').

Construction of Snf5 and Swi1 delete strains

The yeast delete strains were made by gene disruption via PCR using a method described earlier (Longtine, M. S et al, *Yeast* 14, 953–961 (1998)). All the delete strains were derived from LS41 [JPY9::pZZ41, *Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41]. Plasmid PFa6-TRP1 (generously donated by Karbstein group, University of Michigan)(ref) was used as a template to clone out deletion inserts using target-gene-specific primer pairs as designated in Table 3-2.

Table 3-2: Primers used for PCR based gene deletion

Primer	**Purpose**	**Primer Sequence** **5'-3'**
Snf5-Fwd-1	Round 1 PCR	CATCAAGGGAACATATAGTAAAGAACTACACAAAAGCAACA *CGGATCCCCGGGTTAATTAA*
Snf5-Rev-1	Round 1 PCR	GGTTATTTACATCTCCGGTATATTTTATATATGTGTATATATTTT *GAATTCGAGCTCGTTTAAAC*

Snf5-Fwd-2	Round 2 PCR	**CATAAACACCAAAACAAAGCAT**CATCAAGGGAA CATATAGTAAAG
Snf5-Rev-2	Round 2 PCR	**GATAATACAAATTCTTCCAC**GGTTATTTACATCT CCGGTA
Swi1-Fwd	Round 1 PCR	ATGGATTTCTTTAATTTGAATAATAATAATAATAATAATAC *CGGATCCCCGGGTTAATTAA*
Swi1-Rev	Round 1 PCR	TCATTCCAAATTGGTTAGGATATCATTTTTT AAATTGTAAAG*GAATTCGAGCTCGTTTAAAC*

The underlined sequences correspond to the sequence on the pFa6-TRP1 plasmid and the sequences in italics are gene specific sequences. The sequences in bold are Snf5 gene specific sequences and are ~ 20 bp upstream and downstream of Snf5 sequence from Round 1 PCR product.

In case of Swi1 deletion, pFa6-TRP1 was used as template and PCR inserts were cloned out using primers Swi1-Fwd and Swi1-Rev. 1-5 ug of the PCR product was transformed into LS41 and spread on plates containing SC media + 2% Glucose, lacking uracil and tryptophan. After 3-4 days, the colonies grown were screened for deletion strains by lysing a small amount of cells in the colony. Briefly, a small amount of the colony (~ 0.25-0.5 uL) was taken into a PCR tube containing 20 uL of 20 mM NaOH. The tube was boiled for 20 min at 95 °C in a PCR machine and spun down. The supernatant (0.5 – 1 uL) was used as a template and using sequencing primers, the deletion was verified by gel electrophoresis and DNA sequencing. In case of the Snf5 deletion, there was no

successful deletion with one round of PCR and hence a ~ 60 bp Snf5 specific homologous sequence was cloned upstream and downstream of the Trp1 sequence by two rounds of PCR using primers described in Table 2. Screening and selection was done as described for Swi1 deletion and verified by DNA sequencing.

Incorporation of pBpa into LexA+TAD constructs

LS41 yeast was transformed with various pLexAVP16 TAG mutant plasmids and pSNRtRNA-pBpaRS. Individual colonies were grown to saturation in 5 mL SC media lacking histidine and tryptophan for selection and 2% raffinose, 30 °C, with agitation. Starter cultures were then used to inoculate 5 mL SC media lacking histidine and tryptophan, containing 2% raffinose and 2% galactose. For pBpa incorporation, 50 μL of 100 mM pBpa (dissolved in 1M NaOH) and 50 μL 1M HCl were added to the above cultures. The cultures were grown overnight at 30 °C, with agitation to an OD_{660} of ~1.0. 3 OD's of cells were harvested and lysed in 12 uL pellet lysis buffer (50 mM Tris Acetate, pH 7.9, 150 mM KOAc, 20% glycerol, 0.2% Tween-20, 2 mM MgOAc) containing complete EDTA free protease inhibitor tablets (Roche), 7 uL 1 mM DTT, and 7 uL 4X LDS NuPAGE dye (Invitrogen). Lysates were boiled at 95 °C and analyzed using Western blot with anti-FLAG (M2) antibody (Sigma).

In vivo cross-linking

To perform in vivo cross-linking, individual colonies of each pLexA-TAD TAG mutant were grown in 5 mL SC media containing 2% raffinose or but lacking

histidine and tryptophan for selection. The cultures were incubated overnight at 30 °C with agitation. Following incubation, these cultures were used to inoculate 100 mL cultures of SC media containing 2% raffinose and 2% galactose. For pBpa incorporation, 1 mL of 100 mM pBpa (dissolved in 1M NaOH) and 1 mL 1M HCl were added to the above cultures. For control cultures, 1 mL 1M NaOH and 1 mL 1M HCl were added. The cultures were incubated overnight at 30 °C with agitation to an OD_{660} of ~1.0. When cultures reached the appropriate OD_{660}, the cells were spun down by centrifuging at 3901 rcf, 4^{o}C for 5min following which the cell pellets were washed with SC media lacking histidine and tryptophan. The cell pellets were resuspended in 2mL H-W- media + 2% raffinose, 2% galactose and transferred to small cell culture dishes and subjected to UV irradiation at 365 nm light (Eurosolar 15 W UV lamp) with cooling for 0.5 h. The cells were isolated by centrifugation and stored at -80^{o}C until lysis. For crosslinking studies with co-expressed cofactors and the deletion strains, the procedure was identical except that cells were grown in SC media lacking histidine, leucine, and tryptophan and, were grown in 2% raffinose and 2% galactose. For lysis, cells were resuspended in 600 µL Lysis buffer (50 mM Hepes-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-Deoxycholate and 2X Complete Mini, EDTA Free Protease Inhibitor (Roche) and lysed using glass beads by vortexing at 4 ◦C. Subsequently, the lysate was pelleted and the supernatant incubated with 10 µL of LexA antibody (sc-1725, Santa Cruz Biotechnologies) for 2 h at 4 ◦C for immunoprecipitation. The protein bound to the antibody was isolated by incubation for 1 h with either ~50 µL of

prewashed protein G magnetic beads (Dynal Corporation, Invitrogen, Carlsbad, CA) or ~ 25 uL prewashed protein G agarose beads (Millipore) at 4 °C. After immunoprecipitation, the beads were washed 6X with 1 mL Wash Buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.1% Na-Deoxycholate and 1 mM EDTA) and stored dry at -80 °C until elution. The crosslinked sample was eluted from the beads by heating at 95 °C for 10 min in NuPAGE 4x LDS Sample buffer (Invitrogen, Carlsbad, CA) containing 250 mM DTT and probed using Western Blot analysis using anti-FLAG (M2) antibody (Sigma, St. Louis, MO) or anti-myc antibody (SC-40, Santa Cruz Biotechnology, Santa Cruz, CA).

β-*Galactosidase assays in delete strains*

To evaluate the ability of each TAG mutant to activate transcription in the presence or absence of 1 mM pBpa, saturated cultures (SC media + 2% raffinose) of each mutant were used to inoculate 5 mL SC media containing 2% raffinose/galactose but lacking histidine and tryptophan for selection. The cells were grown to an OD of 0.8-1.5 and harvested. The activity of each construct was monitored using β-galactosidase assays as previously described in chapter 2.

F. References

1. Ptashne, M., *How eukaryotic transcriptional activators work.* Nature, 1988. **335**(6192): p. 683-9.
2. Ptashne, M. and A. Gann, *Transcriptional activation by recruitment.* Nature, 1997. **386**(6625): p. 569-77.
3. Perou, C.M., T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-

Dale, P.O. Brown, and D. Botstein, *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
4. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing, the p53 network.* Nature, 2000. **408**: p. 307-310.
5. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. **2**(10): p. 740-9.
6. Lee, L.W. and A.K. Mapp, *Transcriptional switches: chemical approaches to gene regulation.* J Biol Chem, 2010. **285**(15): p. 11033-8.
7. Koehler, A.N., *A complex task? Direct modulation of transcription factors with small molecules.* Curr Opin Chem Biol, 2010. **14**(3): p. 331-40.
8. Berg, T., *Inhibition of transcription factors with small organic molecules.* Curr Opin Chem Biol, 2008. **12**(4): p. 464-71.
9. Arndt, H.D., *Small molecule modulators of transcription.* Angew Chem Int Ed Engl, 2006. **45**(28): p. 4552-60.
10. Berggard, T., S. Linse, and P. James, *Methods for the detection and analysis of protein-protein interactions.* Proteomics, 2007. **7**(16): p. 2833-42.
11. Melcher, K., *New chemical crosslinking methods for the identification of transient protein-protein interactions with multiprotein complexes.* Curr Protein Pept Sci, 2004. **5**(4): p. 287-96.
12. Perkins, J.R., I. Diboun, B.H. Dessailly, J.G. Lees, and C. Orengo, *Transient protein-protein interactions: structural, functional, and network properties.* Structure, 2010. **18**(10): p. 1233-43.
13. 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.
14. Bhaumik, S.R. and M.R. Green, *SAGA is an essential in vivo target of the yeast acidic activator Gal4p.* Genes Dev, 2001. **15**: p. 1935-1945.
15. Bhaumik, S.R., T. Raha, D.P. Aiello, and M.R. Green, *In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer.* Genes Dev, 2004. **18**(3): p. 333-43.
16. Bryant, G.O. and M. Ptashne, *Independent recruitment in vivo by Gal4 of two complexes required for transcription.* Mol Cell, 2003. **11**(5): p. 1301-9.
17. Chang, C., F. Gonzalez, B. Rothermel, L. Sun, S.A. Johnston, and T. Kodadek, *The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro.* J Biol Chem, 2001. **276**(33): p. 30956-63.
18. Jeong, C.J., S.H. Yang, Y. Xie, L. Zhang, S.A. Johnston, and T. Kodadek, *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. **40**(31): p. 9421-7.
19. Klein, J., M. Nolden, S.L. Sanders, J. Kirchner, P.A. Weil, and K. Melcher, *Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA.* J Biol Chem, 2003. **278**(9): p. 6779-86.
20. Koh, S.S., A.Z. Ansari, M. Ptashne, and R.A. Young, *An activator target in the RNA polymerase II holoenzyme.* Mol Cell, 1998. **1**(6): p. 895-904.

21. Li, X.Y., A. Virbasius, X. Zhu, and M.R. Green, *Enhancement of TBP binding by activators and general transcription factors.* Nature, 1999. **399**(6736): p. 605-9.
22. Melcher, K. and S.A. Johnston, *GAL4 interacts with TATA-binding protein and coactivators.* Mol Cell Biol, 1995. **15**(5): p. 2839-48.
23. 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.
24. Xie, Y., C. Denison, S.H. Yang, D.A. Fancy, and T. Kodadek, *Biochemical characterization of the TATA-binding protein-Gal4 activation domain complex.* J Biol Chem, 2000. **275**(41): p. 31914-20.
25. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
26. Ptashne, M., *Genes and Signal.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
27. Neely, K.E., A.H. Hassan, A.E. Wallberg, D.J. Steger, B.R. Cairns, A.P. Wright, and J.L. Workman, *Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays.* Mol Cell, 1999. **4**(4): p. 649-55.
28. Peterson, C.L., Workman, J.L., *Promoter targeting and chromatin remodeling by the SWI/SNF complex.* Curr Opin Genet Dev., 2000. **10**(2): p. 187-92.
29. Asturias, F.J., Y.W. Jiang, L.C. Myers, C.M. Gustafsson, and R.D. Kornberg, *Conserved structures of mediator and RNA polymerase II holoenzyme.* Science, 1999. **283**(5404): p. 985-7.
30. Flanagan, P.M., R.J. Kelleher, 3rd, M.H. Sayre, H. Tschochner, and R.D. Kornberg, *A mediator required for activation of RNA polymerase II transcription in vitro.* Nature, 1991. **350**(6317): p. 436-8.
31. Kim, Y.J., S. Bjorklund, Y. Li, M.H. Sayre, and R.D. Kornberg, *A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II.* Cell, 1994. **77**(4): p. 599-608.
32. Davis, J.A., Y. Takagi, R.D. Kornberg, and F.A. Asturias, *Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction.* Mol Cell, 2002. **10**(2): p. 409-15.
33. Taatjes, D.J., A.M. Naar, F. Andel, 3rd, E. Nogales, and R. Tjian, *Structure, function, and activator-induced conformations of the CRSP coactivator.* Science, 2002. **295**(5557): p. 1058-62.
34. Taatjes, D.J., T. Schneider-Poetsch, and R. Tjian, *Distinct conformational states of nuclear receptor-bound CRSP-Med complexes.* Nat Struct Mol Biol, 2004. **11**(7): p. 664-71.
35. Lee, Y.C., J.M. Park, S. Min, S.J. Han, and Y.J. Kim, *An activator binding module of yeast RNA polymerase II holoenzyme.* Mol Cell Biol, 1999. **19**(4): p. 2967-76.

36. Park, J.M., H.S. Kim, S.J. Han, M.S. Hwang, Y.C. Lee, and Y.J. Kim, *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. **20**(23): p. 8709-19.
37. Swanson, M.J., H. Qiu, L. Sumibcay, A. Krueger, S.J. Kim, K. Natarajan, S. Yoon, and A.G. Hinnebusch, *A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo.* Mol Cell Biol, 2003. **23**(8): p. 2800-20.
38. Novatchkova, M. and F. Eisenhaber, *Linking transcriptional mediators via the GACKIX domain super family.* Curr Biol, 2004. **14**(2): p. R54-5.
39. Long, R.M., L.M. Mylin, and J.E. Hopper, *GAL11 (SPT13), a transcriptional regulator of diverse yeast genes, affects the phosphorylation state of GAL4, a highly specific transcriptional activator.* Mol Cell Biol, 1991. **11**(4): p. 2311-4.
40. Gaudreau, L., M. Keaveney, J. Nevado, Z. Zaman, G.O. Bryant, K. Struhl, and M. Ptashne, *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.
41. Suzuki, Y., Y. Nogi, A. Abe, and T. Fukasawa, *GAL11 protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in Saccharomyces cerevisiae.* Mol Cell Biol, 1992. **12**(10): p. 4806.
42. Zhang, F., L. Sumibcay, A.G. Hinnebusch, and M.J. Swanson, *A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p.* Mol Cell Biol, 2004. **24**(15): p. 6871-86.
43. 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.
44. Lallet, S., H. Garreau, C. Garmendia-Torres, D. Szestakowska, E. Boy-Marcotte, S. Quevillon-Cheruel, and M. Jacquet, *Role of Gal11, a component of the RNA polymerase II mediator in stress-induced hyperphosphorylation of Msn2 in Saccharomyces cerevisiae.* Mol Microbiol, 2006. **62**(2): p. 438-52.
45. Leroy, C., L. Cormier, and L. Kuras, *Independent recruitment of mediator and SAGA by the activator Met4.* Mol Cell Biol, 2006. **26**(8): p. 3149-63.
46. Sakurai, H., Y. Hiraoka, and T. Fukasawa, *Yeast GAL11 protein is a distinctive type transcription factor that enhances basal transcription in vitro.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8382-6.
47. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. **18**(3): p. 369-78.
48. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. **25**(20): p. 9092-102.
49. Ansari, A.Z., S.S. Koh, Z. Zaman, C. Bongards, N. Lehming, R.A. Young, and M. Ptashne, *Transcriptional activating regions target a cyclin-dependent kinase.* Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14706-9.

50. Grant, P.A., D.E. Sterner, L.J. Duggan, J.L. Workman, and S.L. Berger, *The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes.* Trends Cell Biol, 1998. **8**(5): p. 193-7.
51. 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.
52. Brown, C.E., L. Howe, K. Sousa, S.C. Alley, M.J. Carrozza, S. Tan, and J.L. Workman, *Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit.* Science, 2001. **292**(5525): p. 2333-7.
53. Green, M.R., *Eukaryotic transcription activation: right on target.* Mol Cell, 2005. **18**(4): p. 399-402.
54. Garbett, K.A., M.K. Tripathi, B. Cencki, J.H. Layer, and P.A. Weil, *Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction.* Mol Cell Biol, 2007. **27**(1): p. 297-311.
55. Fuxreiter, M., P. Tompa, I. Simon, V.N. Uversky, J.C. Hansen, and F.J. Asturias, *Malleable machines take shape in eukaryotic transcriptional regulation.* Nat Chem Biol, 2008. **4**(12): p. 728-37.
56. Prochasson, P., K.E. Neely, A.H. Hassan, B. Li, and J.L. Workman, *Targeting activity is required for SWI/SNF function in vivo and is accomplished through two partially redundant activator-interaction domains.* Mol Cell, 2003. **12**(4): p. 983-90.
57. Neely, K.E., A.H. Hassan, C.E. Brown, L. Howe, and J.L. Workman, *Transcription activator interactions with multiple SWI/SNF subunits.* Mol Cell Biol, 2002. **22**(6): p. 1615-25.
58. Ferreira, M.E., P. Prochasson, K.D. Berndt, J.L. Workman, and A.P. Wright, *Activator-binding domains of the SWI/SNF chromatin remodeling complex characterized in vitro are required for its recruitment to promoters in vivo.* Febs J, 2009. **276**(9): p. 2557-65.
59. Schwabish, M.A. and K. Struhl, *The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo.* Mol Cell Biol, 2007. **27**(20): p. 6987-95.
60. Bryant, G.O., V. Prabhu, M. Floer, X. Wang, D. Spagna, D. Schreiber, and M. Ptashne, *Activator control of nucleosome occupancy in activation and repression of transcription.* PLoS Biol, 2008. **6**(12): p. 2928-39.
61. Govind, C.K., S. Yoon, H. Qiu, S. Govind, and A.G. Hinnebusch, *Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo.* Mol Cell Biol, 2005. **25**(13): p. 5626-38.
62. Memedula, S. and A.S. Belmont, *Sequential recruitment of HAT and SWI/SNF components to condensed chromatin by VP16.* Curr Biol, 2003. **13**(3): p. 241-6.
63. Yudkovsky, N., C. Logie, S. Hahn, and C.L. Peterson, *Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators.* Genes Dev, 1999. **13**(18): p. 2369-74.

64. Ferreira, M.E., S. Hermann, P. Prochasson, J.L. Workman, K.D. Berndt, and A.P. Wright, *Mechanism of transcription factor recruitment by acidic activators.* J Biol Chem, 2005. **280**(23): p. 21779-84.
65. Laurent, B.C., I. Treich, and M. Carlson, *The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation.* Genes Dev, 1993. **7**(4): p. 583-91.
66. Cote, J., J. Quinn, J.L. Workman, and C.L. Peterson, *Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex.* Science, 1994. **265**(5168): p. 53-60.
67. Lemieux, K. and L. Gaudreau, *Targeting of Swi/Snf to the yeast GAL1 UAS G requires the Mediator, TAF IIs, and RNA polymerase II.* Embo J, 2004. **23**(20): p. 4040-50.
68. Wands, A.M., N. Wang, J.K. Lum, J. Hsieh, C.A. Fierke, and A.K. Mapp, *Transient-state kinetic analysis of transcriptional activator.DNA complexes interacting with a key coactivator.* J Biol Chem, 2011. **286**(18): p. 16238-45.
69. Melcher, K., *The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins.* J Mol Biol, 2000. **301**(5): p. 1097-112.
70. Majmudar, C.Y., B. Wang, J.K. Lum, K. Hakansson, and A.K. Mapp, *A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry.* Angew Chem Int Ed Engl, 2009. **48**(38): p. 7021-4.
71. Thoden, J.B., L.A. Ryan, R.J. Reece, and H.M. Holden, *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. **283**(44): p. 30266-72.
72. Majmudar, C.Y., A.E. Labut, and A.K. Mapp, *Tra1 as a screening target for transcriptional activation domain discovery.* Bioorg Med Chem Lett, 2009. **19**(14): p. 3733-5.
73. Majmudar, C.Y., L.W. Lee, J.K. Lancia, A. Nwokoye, Q. Wang, A.M. Wands, L. Wang, and A.K. Mapp, *Impact of nonnatural amino acid mutagenesis on the in vivo function and binding modes of a transcriptional activator.* J Am Chem Soc, 2009. **131**(40): p. 14240-2.
74. Wu, Y., R.J. Reece, and M. Ptashne, *Quantitation of putative activator-target affinities predicts transcriptional activating potentials.* Embo J, 1996. **15**(15): p. 3951-63.
75. Ansari, A.Z., R.J. Reece, and M. Ptashne, *A transcriptional activating region with two contrasting modes of protein interaction.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13543-8.
76. Krishnamurthy, M., A. Dugan, A. Nwokoye, Y.H. Fung, J.K. Lancia, C.Y. Majmudar, and A.K. Mapp, *Caught in the act: covalent crosslinking captures activator-coactivator interactions in vivo.* ACS Chem Biol, 2011.
77. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo.* J Biol Chem, 2002. **277**(48): p. 46043-50.

78. Wu, Z., G. Belanger, B.B. Brennan, J.K. Lum, A.R. Minter, S.P. Rowe, A. Plachetka, C.Y. Majmudar, and A.K. Mapp, *Targeting the transcriptional machinery with unique artificial transcriptional activators.* J Am Chem Soc, 2003. **125**(41): p. 12390-1.
79. Lu, Z., S.P. Rowe, B.B. Brennan, S.E. Davis, R.E. Metzler, J.J. Nau, C.Y. Majmudar, A.K. Mapp, and A.Z. Ansari, *Unraveling the mechanism of a potent transcriptional activator.* J Biol Chem, 2005. **280**(33): p. 29689-98.
80. Janknecht, R., *The versatile functions of the transcriptional coactivators p300 and CBP and their roles in disease.* Histol Histopathol, 2002. **17**(2): p. 657-68.
81. Goodman, R.H. and S. Smolik, *CBP/p300 in cell growth, transformation, and development.* Genes Dev, 2000. **14**(13): p. 1553-77.
82. Radhakrishnan, I., G.C. Perez-Alvarado, D. Parker, H.J. Dyson, M.R. Montminy, and P.E. Wright, *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions.* Cell, 1997. **91**(6): p. 741-52.
83. Radhakrishnan, I., G.C. Perez-Alvarado, D. Parker, H.J. Dyson, M.R. Montminy, and P.E. Wright, *Structural analyses of CREB-CBP transcriptional activator-coactivator complexes by NMR spectroscopy: implications for mapping the boundaries of structural domains.* J Mol Biol, 1999. **287**(5): p. 859-65.
84. Langlois, C., C. Mas, P. Di Lello, L.M. Jenkins, P. Legault, and J.G. Omichinski, *NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53.* J Am Chem Soc, 2008. **130**(32): p. 10596-604.
85. Vignali, M., D.J. Steger, K.E. Neely, and J.L. Workman, *Distribution of acetylated histones resulting from Gal4-VP16 recruitment of SAGA and NuA4 complexes.* Embo J, 2000. **19**(11): p. 2629-40.
86. Triezenberg, S.J., R.C. Kingsbury, and S.L. McKnight, *Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression.* Genes Dev, 1988. **2**(6): p. 718-29.
87. Hayes, S. and P. O'Hare, *Mapping of a major surface-exposed site in herpes simplex virus protein Vmw65 to a region of direct interaction in a transcription complex assembly.* J Virol, 1993. **67**(2): p. 852-62.
88. Hinnebusch, A.G., *The Molecular Biology of the Yeast Saccharomyces (Jones, E. W., Pringle, J. R. and Broach, J. R. eds.).* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992: p. 319-414.
89. Myasnikov, A.N., K.V. Sasnauskas, A.A. Janulaitis, and M.N. Smirnov, *The Saccharomyces cerevisiae ADE1 gene: structure, overexpression and possible regulation by general amino acid control.* Gene, 1991. **109**(1): p. 143-7.
90. Mosch, H.U., B. Scheier, R. Lahti, P. Mantsala, and G.H. Braus, *Transcriptional activation of yeast nucleotide biosynthetic gene ADE4 by GCN4.* J Biol Chem, 1991. **266**(30): p. 20453-6.

91. Struhl, K., *Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R., eds.).* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992: p. 833-859.
92. Wilson, B.G. and C.W. Roberts, *SWI/SNF nucleosome remodellers and cancer.* Nat Rev Cancer, 2011. **11**(7): p. 481-92.
93. Hang, C.T., J. Yang, P. Han, H.L. Cheng, C. Shang, E. Ashley, B. Zhou, and C.P. Chang, *Chromatin regulation by Brg1 underlies heart muscle development and disease.* Nature, 2010. **466**(7302): p. 62-7.
94. Hargreaves, D.C. and G.R. Crabtree, *ATP-dependent chromatin remodeling: genetics, genomics and mechanisms.* Cell Res, 2011. **21**(3): p. 396-420.

CHAPTER 4
CONCLUSIONS AND FUTURE DIRECTIONS

A. Summary of dissertation

Transcriptional activators use two domains (DBD and TAD) in order to modulate gene expression, regulate the timing and extent to which mRNA levels are upregulated in response to cellular stimuli.[1] The DBD is responsible for localizing the activator to a specific set of sequences in DNA and the TAD makes contact with several protein and protein complexes to assemble the transcriptional machinery for transcription of that gene to be accomplished.[2] Many diseases ranging from cancer to diabetes arise from gene misregulation facilitated by malfunctioning transcriptional activators.[3, 4] Therefore, a major goal towards the treatment of these disease states has been to generate molecules that can influence or mimic the function of natural activator proteins.[5-8] However, this effort has been hampered by the lack of structural and mechanistic information regarding the interactions between activators and the transcriptional machinery, a direct result of insufficient methodologies available for studying these interactions. Towards overcoming this hurdle, an in vivo photo-crosslinking methodology has been presented which provides a way to covalently capture both tight and modest-affinity protein-protein interactions in living cells. To push forward towards tailored library screens for small molecule modulators of transcription, scientists need to not only know the direct binding partners of

transcriptional activators but also to characterize their binding interfaces. This will yield valuable insight that can be applied to the generation of tailored screens for artificial activators.

In the body of work presented in this dissertation, we used in vivo photo-crosslinking strategy to identify direct binding partners of transcriptional activators in live cells. In Chapter 2, nonsense suppression was achieved using a bio-orthogonal pSNR tRNA/aaRS to specifically and efficiently incorporate the unnatural amino acid pBpa at multiple sites within the TAD of the prototypical amphipathic activator Gal4. Importantly, the incorporation of pBpa within the Gal4 TAD was achieved with minimal impact on activator function. Additionally in Chapter 2, this in vivo photo-crosslinking strategy was used successfully to capture the high affinity interaction between Gal4 and its suppressor protein, Gal80 (Figure 4-1). The identification of the well-characterized Gal4-Gal80 interaction was an important first step in validating in vivo photo-crosslinking as an indispensable tool for studying transcriptional activator protein interactions.

While TAD-repressor interactions such as that between Gal4 and Gal80 can be high in affinity, the interactions between a TAD and the various coactivator complexes tend to be much more moderate in affinity and transient in nature. These characteristics of TAD-coactivator protein interactions have made the study of this interaction more difficult.

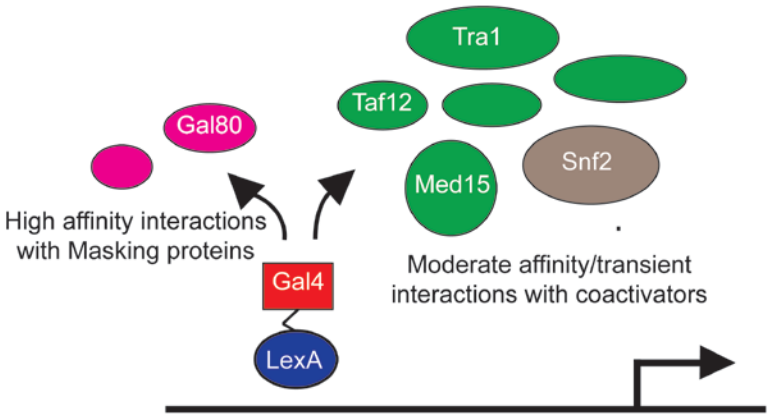


Figure 4-1: Summary of coactivator targets identified to directly contact the Gal4 TAD through in vivo photo-crosslinking experiments from chapters 2 and 3. In vivo photo-crosslinking was used to capture a high affinity masking protein interaction (Gal80), as well as moderate affinity and or transient interactions with coactivator proteins co-expressed (green circle) and native concentrations (gray circle)

In Chapter 3, I describe how this challenge was overcome and demonstrated that the in vivo photo-crosslinking strategy is powerful enough to capture modest-affinity and/or transient activator interactions. Specifically, using Gal4, we showed that Gal4 directly contacts Med15, Taf12 and Tra1 in vivo (Figure 4-1). Moreover, the identification of endogenous protein targets was achieved using this in vivo photo-crosslinking strategy. As a clear example of this, we experimentally showed that Gal4 directly binds to endogenous Snf2 (Figure 4-1). Next, we extended these studies to include two other well-characterized amphipathic activators, VP16 and Gcn4, and examined the extent of overlapping targets between Gal4, VP16 and Gcn4. These experiments suggest that, while amphipathic activators may recruit the same coactivator complex to a given promoter, they do not necessarily target the same protein subunits within that coactivator complex. These subtle differences in coactivator recruitment reveal interactions specific to a particular activator that can be explored for therapeutics.

In summary, the presented in vivo photo-crosslinking methodology is a powerful tool for teasing out both high- and modest-affinity protein interactions, in living cells. Further characterization of the TAD-coactivator interactions identified through these studies will be important in the design of small molecule transcriptional modulators.

B. Future directions

Given the importance of transcription, it is not surprising that gene misregulation is associated with almost every disease.[9, 10] Therefore, there have been numerous efforts towards the development of artificial transcriptional modulators to influence or mimic the function of natural activator proteins.[11-13] Although the work presented in this thesis has led to the identification of several of the direct, in vivo targets of transcriptional activators, detailed characterization of these interaction interfaces is still required in order to aid in the design of small molecule transcriptional modulators. Moreover, the identification of novel activator interactions may provide additional opportunities for small molecule development. Currently, the optimal methodology to carry out these characterizations is mass spectrometry, which should be employed in conjunction with in vivo photo-crosslinking for defining other unknown targets of TADs.

Mass spectrometry has been employed for studying protein-protein interactions and has been proven to be a highly sensitive and powerful method. In particular, multidimensional protein identification technology (MudPIT) has been utilized for detecting proteins in a complex mixture. The Yates group has

used this method to analyze a large-scale yeast protein mixture, resulting in the identification of 1,484 proteins.[14] In their on-line setup (Figure 4-2), trypsin-digested proteins were loaded onto a biphasic 2D microcapillary column and the peptide mixture was separated by strong cation exchange (SCX) and reverse-phase (RP-HPLC) before being eluted directly onto an ESI-MS/MS. Afterwards peptide sequences were analyzed by the SEQUEST software to identify proteins.[16]

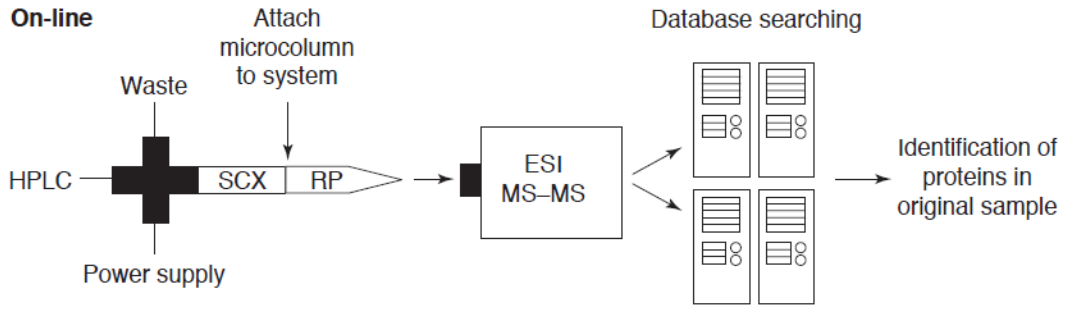


Figure 4-2: On-line mass spectrometry setup.[15] The protein sample in not handled in between protein separation and the mass spectrometer.

The yeast proteins identified from studies from the Yates group were representatives from all sub-cellular compartments and some had extreme hydrophobicity and isoelectric points.[14] For the purposes of studying endogenous transcriptional proteins, perhaps the most important observation from this experiment was the identification of several low abundance proteins. In fact, about 54% of the proteins identified from this study had CAI values that were less than 0.2, indicating that MudPIT is an effective tool for analyzing low abundance proteins, including transcriptional proteins. Among the 1, 484 proteins identified, 45 were transcription proteins including the SNF5, SWI6 and SWI4 subunits of

the Swi/Snf complex.[14] As such, the utility of MudPIT appears promising for further identification and characterization of direct binding partners of the Gal4 TAD. In fact, preliminary experiments by the Mapp lab have coupled in vivo crosslinking experiments with MudPIT analysis to identify the masking protein Gal80 as a binding partner of Gal4. Using similar experiments, the direct binding partners found from the work in this dissertation can be confirmed and further the binding sites within these targets can be teased out (further discussed in Section B.1).

B.1. Development of probes and therapeutic small molecules

MudPIT possesses remarkable sensitivity for identifying low abundant proteins and therefore can be employed for further analysis of our crosslinked protein samples. After the direct, in vivo targets of Gal4 identified from this work are confirmed by MudPIT, further, the binding sites within the targets can be characterized for the protein interface utilized for interaction. One way to identify the binding site, while utilizing mass spectrometry coupled with photo-crosslinking is to use isolated TADs from activators, containing pBpa, like Gal4 and perform in vitro crosslinking experiments with different fragments of the established target. Using western blotting, the regions on the target that crosslink with the TAD can be first identified. Furthermore, the crosslinked fragments + TAD can be submitted for a more challenging mass spectrometry analysis for exact binding site(s) determination.

From the targets reported in this dissertation as direct targets of Gal4, Snf2 is a good candidate for binding site determination not only because it is

reported here for the first time as a direct binding partner of Gal4 in vivo, but Snf2 interacted with Gal4 endogenous levels. Fragments spanning the entire sequence of Snf2 can be expressed as fusion proteins before crosslinking experiments with pBpa containing TADs. Subsequently, the physiological relevance of the fragments of Snf2 that interact with Gal4 could be determined by measuring their activity in Snf2 deletion strains. Such information will guide studies with small molecule probes that can be used to further understand TAD-partner binding interfaces. Towards the design of artificial TAD therapeutics, small molecule screens to identify molecules that disrupt an endogenous activator-coactivator interaction can be achieved easily if the binding interface between a TAD and its partner is better understood. This is a major reason for detailed structural understanding of TAD-coactivator interactions. In the long term, efforts to determine the identity of activator binding partners and the binding sites will establish screening targets for artificial activator and inhibitor design for therapeutic and other applications.

B.2. Map the entire network of PPIs needed for transcription

Transcriptional activators are known to have a multi-protein binding profile and make many contacts with proteins in the transcriptional machinery. Despite numerous studies directed at identifying the direct binding partners of activators, the direct, in vivo targets of transcriptional activators remain largely unknown. While in vivo photo-crosslinking has led to the identification of several direct, in vivo targets of Gal4, one limitation to using this strategy is that only already proposed targets can be identified. This poses a problem, as many of the

putative binding partners of transcriptional activators do not have high-quality antibodies available. Therefore, MudPIT offers the potential to map the entire network of PPIs that take place during transcription, as this technique can be utilized for identifying unknown targets. Importantly, mass spectrometry provides a complete picture of the proteins that bind to transcriptional activators when coupled with photo-crosslinking, which in vivo photo-crosslinking by itself cannot offer. In conclusion, in vivo photo-crosslinking accompanied by mass spectrometry techniques (like MudPIT) will be useful for not only identifying unknown partners of TADs but for answering questions pertaining to the primary mechanisms utilized for transcription..

C. References

1. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
2. Ptashne, M., *Genes and Signal.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
3. Chen, X., S.T. Cheung, S. So, S.T. Fan, C. Barry, J. Higgins, K.M. Lai, J. Ji, S. Dudoit, I.O. Ng, M. Van De Rijn, D. Botstein, and P.O. Brown, *Gene expression patterns in human liver cancers.* Molecular Biology of the Cell, 2002. **13**(6): p. 1929-39.
4. Perou, C.M., T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, and D. Botstein, *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
5. Mapp, A.K., A.Z. Ansari, M. Ptashne, and P.B. Dervan, *Activation of gene expression by small molecule transcription factors.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 3930-5.
6. Dervan, P.B. and B.S. Edelson, *Recognition of the DNA minor groove by pyrrole-imidazole polyamides.* Curr Opin Struct Biol, 2003. **13**(3): p. 284-99.
7. Faria, M. and C. Giovannangeli, *Triplex-forming molecules: from concepts to applications.* J Gene Med, 2001. **3**(4): p. 299-310.
8. Lum, J.K. and A.K. Mapp, *Artificial transcriptional activation domains.* Chembiochem, 2005. **6**(8): p. 1311-5.

9. Melnick, A. and J.D. Licht, *Deconstructing a disease: RARalpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia.* Blood, 1999. **93**(10): p. 3167-215.
10. Mistry, A.R., E.W. Pedersen, E. Solomon, and D. Grimwade, *The molecular pathogenesis of acute promyelocytic leukaemia: implications for the clinical management of the disease.* Blood Rev, 2003. **17**(2): p. 71-97.
11. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. **2**(10): p. 740-9.
12. Pandolfi, P.P., *Transcription therapy for cancer.* Oncogene, 2001. **20**(24): p. 3116-27.
13. Majumdar, C.Y. and A.K. Mapp, *Chemical approaches to transcriptional regulation.* Curr Opin Chem Biol, 2005. **9**: p. 467-477.
14. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nat Biotechnol, 2001. **19**(3): p. 242-7.
15. Washburn, M.P., Yates, J. R.,, *New methods for proteome analysis: Multidimentional chromatograph and mass spectromentry*. Trends in Biotechnology. 2000.
16. Link, A.J., J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, and J.R. Yates, 3rd, *Direct analysis of protein complexes using mass spectrometry.* Nat Biotechnol, 1999. **17**(7): p. 676-82.