Covalent Chemical Capture of Transcriptional Protein-Protein Interactions Using Genetically Incorporated Photo-crosslinking Amino Acids

by

Amanda Elizabeth Dugan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2013

Doctoral Committee:

Professor Anna K. Mapp, Chair
Associate Professor Patrick J. O'Brien
Professor Janet L. Smith
Associate Professor Raymond C. Trievel

Acknowledgements

The path to graduation was not an easy one to navigate and I am indebted to many people who have helped me along the way. First and foremost is my advisor, Anna, who I somehow managed to convince to let me work in her lab. The early years of graduate school were full of uncertainty and self-doubt - I stumbled a lot and made a lot of mistakes. Anna always showed extended patience in dealing with my concerns and pushed me to accomplish bigger goals than I thought I was capable of doing. As my graduate career comes to an end, I am more confident in myself and in my future as a researcher and I know that her support and mentorship got me to this point. The woman, quite frankly, deserves a medal. I am also eternally grateful for the support and encouragement I received from the outstanding scientists that sat on my thesis committee. I will always appreciate the honesty with which Janet, Pat, and Ray evaluated my work and I know their detailed critiques have made me a better scientist and public speaker. Many thanks also to the current and previous people behind the Program in Chemical Biology, especially Laura Howe, Amy McGovern, and Justine Altman, for keeping me on track over the last five years.

My lab family members, new and old, have played an instrumental role in my scientific and personal development. It has been such a great experience to have worked every day with co-workers who I also consider friends. I owe much of my success to Dr. Malathy Krishnamurthy, who took me under her wing and waited patiently as I made mistakes, relentlessly asked her questions, and hounded her for help in figuring out why my project wasn't working. Malathy selflessly shared her project with me, taught me how to evaluate my data in the context of a bigger picture, and worked with me until I was technically proficient in lab. She gets a medal too. To Ningkun, thank you for your friendship, for innumerable coffee breaks, and always leading the way. I could not have made it this far without you. Chinmay, you have been a great collaborator and friend

who has taught me the value of innovation and perseverance and that there's nothing a beer and kati roll can't fix. It has truly been a pleasure working with someone who is equally stubborn as I. To my cubbymate, Adaora, thank you for the years of moral support and friendship. Thanks also to other members of the early crosslinking team including Chinmay, Lori, Hugo, and Jody for the years of collaboration. I am also grateful to have worked with the outstanding young people who make up the current crosslinking group – Rachel, Cassie, and Micah – you guys are going to do so many great things and I wish you the very best. Rachel, I'm so happy you're a part of the lab and that we've been able to work as friends for the past year. To Chris, Amy, Jonas, Paul, Steve, JP, Vik, James, Jason, Lindsey, Aaron and the rest of the Mapp Lab – thank you for the years of conversation and the countless times my face has hurt from laughing with you all.

The friends I have made in graduate school have been a true blessing to me. My amazing cohort in the Program in Chemical Biology, particularly Rachel, Ningkun, Elin, and Eli, thank you for letting me be a part of your lives. You guys are the best and I can't put into words how much your friendship has meant to me. To my friends from the greatest state ever (Ohio), Anne and Asia, you are amazing and I love you both.

Finally, I owe so much to my family, both immediate and extended. Thank you to my parents for teaching me the importance of hard work, to respect myself and my goals, and to never give up on something that is important to me. I love you more than words can describe. To my brothers, Andy, Justin, Ethan, and Bart and my sister-in-law Kim – you are my best friends in the world and your support means the world to me. To my grandparents– thank you for the years of encouragement and unconditional love. My extended family, particularly my in-laws Tom, Chris, Jim, Wendy and Grandpa and Gram, thank you for welcoming me into your family with open arms and for giving me my husband, Michael, who has been my best friend, partner and ally for the last eight years. Michael, thanks for proposing to me in the middle of graduate school because, you know, I

didn’t have enough things to work on. I still love you, though. We have shared so much together and I can’t wait to see what life has in store for us. In honor of you and our first child, I am dedicating this thesis.

Table of Contents

List of Figures

List of Tables

Abstract

COVALENT CHEMICAL CAPTURE OF TRANSCRIPTION PROTEIN-PROTEIN INTERACTIONS USING GENETICALLY INCORPORATED PHOTOCROSSLINKING AMINO ACIDS

by

Amanda E. Dugan

Chair: Anna K. Mapp

Critical to the regulation of nearly every cellular process are protein-protein interaction (PPI) networks whose constituent contacts differ dramatically in affinities, lifetimes, and interaction interface area. For example, PPIs are essential to transcriptional regulation, a process in which transcriptional activators dynamically engage in a diverse series of binding interactions to recruit numerous multi-protein coactivator complexes to the promoter of a gene to be transcribed. Activators make specific, high affinity contacts with masking proteins but they also engage in weaker affinity, more transient interactions with the transcriptional machinery. These weaker affinity, transient PPIs are important as they are believed to play a significant role in modulating the function, localization, and specificity of key complexes that are central to PPI networks. Given the fundamental role that PPIs play in executing cellular processes such as transcription initiation, a longstanding goal has been to design small molecules that will transiently modulate these contacts and reveal key insights into how they function in the context of their associated networks. However, this goal has proven to be quite challenging as currently just a small fraction of PPIs are targeted with small molecules. Of the thousands of contacts that remain, many are considered impossible to target, particularly those that are more transient in nature. In many cases, characterizing these PPIs using commonly

employed affinity based methods is challenging. Thus, methods capable of capturing a broad spectrum of PPIs, particularly in their native cellular environment, should prove incredibly useful.

The major goal of this thesis is to demonstrate the utility of *in vivo* photocrosslinking with a genetically incorporated photo-labile amino acid p-benzoyl-L-phenylalanine (Bpa) in capturing the direct targets of transcriptional activators in living cells. Using this methodology, the interaction with the Mediator protein Med15 was captured, indicating the power of this approach in capturing weaker affinity interactions that have historically been challenging to study. Furthermore, the direct targets within the Swi/Snf chromatin remodeling complex were identified, indicating the compatibility of this approach in studying transcriptional systems that often utilize low-abundance complexes. In combination with formaldehyde crosslinking, *in vivo* crosslinking with Bpa demonstrates for the first time the ability to capture the direct targets of DNA bound activators. And finally, using mass spectrometry to identify binding partners of transcriptional activators reveals several novel targets, thus contributing to a more complete interaction map for transcriptional networks.

Chapter 1 Introduction[1]

1.A. Overview of protein-protein interaction networks

Protein-protein interactions (PPIs) form the backbone of nearly every facet of cellular function. Remarkably complex, PPI networks are composed of thousands of contacts that range from very stable and high affinity to interactions that are much weaker and more transient in nature. Exemplifying complex PPI profiles are transcriptional activators, a family of signal responsive proteins that localize to gene promoters and recruit the necessary transcriptional complexes required to upregulate gene expression. In the cell, activators are believed to dynamically interact with numerous protein complexes, and the contacts include high-affinity interactions with masking proteins in addition to weaker, transient interactions with transcriptional coactivator complexes. As with other functionally important transient PPIs, visualization of direct activator-coactivator interactions has proven challenging given the dynamics of the association and structural plasticity of the interaction interfaces. As a result, much remains unclear regarding the mechanism by which activators recruit coactivator complexes *in vivo* in order to upregulate transcription. Thus, a longstanding goal has been develop probes and tools that can be used to define these interactions in a cellular setting.

Although the goal of modulating PPIs has been recognized for some time, these interactions have been historically challenging to target with small molecules.[2-4] Indeed, the majority of PPIs have been broadly classified as "undruggable" and among the estimated 650,000 PPIs in the cell, far less than

[1] Portions of this Chapter are from from published work regarding the challenges associated with targeting protein-protein interactions: (1) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. *ACS Chem. Biol.* **2012**, *7*, 1311.

0.01% have been targeted with inhibitors.[2,4-6] Upon closer examination of the PPIs utilized in transcriptional systems, a sharp decline is observed in the amount of high-resolution structural and mechanistic data required for the rational design of small molecule modulators intended to perturb these interactions. This lack of information can be attributed in part to the limitations of the methods currently available for studying the broad spectrum of PPIs that exist and the even smaller number of methods able to examine PPIs in their native cellular environment.[7-9] As such, new methodologies with this application in mind are essential to advancing progress toward creating complete functional maps of transcription PPI networks.

1.A.1. Diversity of PPIs and associated challenges in targeting with small molecules

PPIs differ greatly in binding affinity and contact area, making a one-size-fits-all approach in targeting them difficult to achieve. In contrast to protein-ligand interfaces (PLIs) such as those between an enzyme and its substrate, PPIs tend to be larger and flatter, with an average surface area of 1940 +/- 760 Å^2.[3,10,11] Consistent with their wide distribution of contact areas, PPIs exhibit a wide range of affinity values, with examples of pM dissociation constants in more stable complexes and mM values in transient complexes.[8,12,13] Several examples of PPIs with a diverse range of affinities and contact areas are highlighted in Figure 1-1 and Table 1-1.[14-18] The scope of this diversity is postulated to be even greater than currently appreciated given the underrepresentation in the Protein Data Bank (PDB) of transient, low affinity and membrane associated PPIs.[8,19] Structure-based drug design approaches have been particularly successful in targeting tight affinity and small surface area PPIs (e.g. most similar to PLIs) as they have been more amenable to characterization using standard structural and biochemical approaches. In fact, a 2012 query of the 2P2IDB and TIMBAL databases, which analyze PPI interfaces and the small molecules available to target PPIs, respectively, revealed an overwhelming number of PPI inhibitors for

contacts with smaller surface areas (less than 1,800 Å^2) and relatively tight affinities (nM dissociation constants). In fact, this class of PPIs was found to be targeted by approximately 68% of reported small molecules.[1,20,21] In contrast, only about 10% of known inhibitors target PPIs that are of moderate to weak affinity (≥1 μM). We expect that this disparity in targeting ability will diminish with the advancement of methods better-suited to study this challenging class of PPIs, thus revealing features necessary to guide drug discovery efforts.

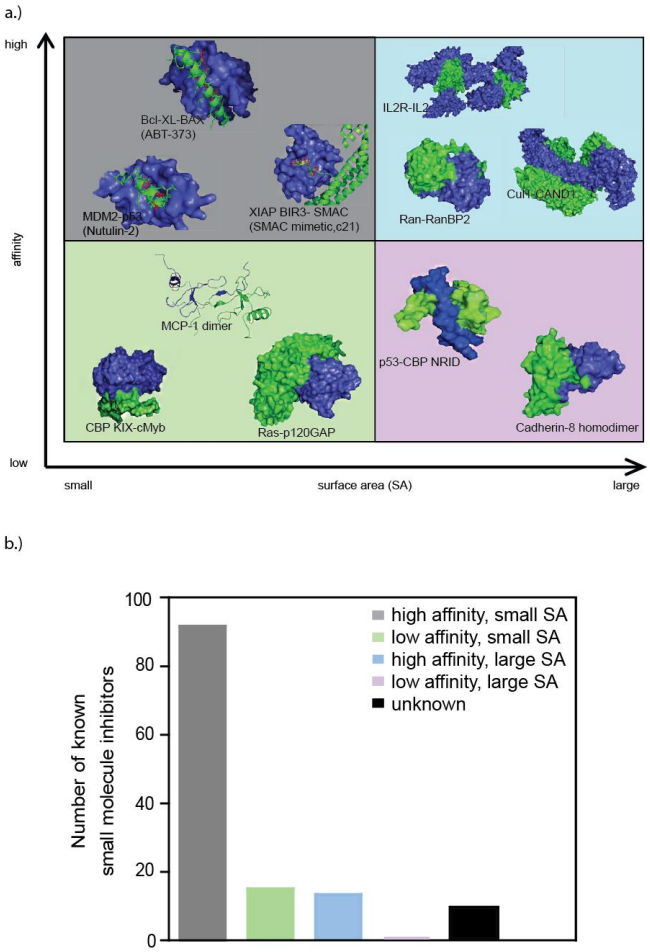

Figure 1-1 (Top) PPIs from a variety of interaction networks cover a broad spectrum of affinities and contact areas. Analysis of PPIs using these two variables leads to four major quadrants to describe these interactions: strong and concise (grey, upper left), strong and broad (blue, upper right), weak and concise (green, lower left), and weak and broad (pink, lower right). Examples of each type of interaction from the Protein Data Bank (PDB) are shown. Strong and concise: p53-MDM2 (1YCQ), Bcl-xL-BAX (2XA0), XIAP-SMAC (2JK7, 1G73); strong and broad: IL-2-IL2R (2ERJ), Ran-RanBP2 (1RRP), Cul1-CAND1 (1U6G);

weak and concise: CBP KIX-cMyb complex (1SB0), MCP-1 dimer (1DOM), and Ras-p120GAP (1WQ1); weak and broad: cadherin-8-homodimer (1ZXK) and p53-CBP NRID (2L14). (Bottom) PPIs with known inhibitors were gathered from the 2P2IDB and TIMBAL databases and categorized based on the PPI binding affinities (as reported in the PDBbind) and the surface area of the interaction (as measured by InterProSurf.) Analysis revealed that an overwhelming majority of known PPI inhibitors target PPIs with smaller surface areas (<1800 Å^2) and relatively tight affinities (<1 mM). Colors on graph correspond to the quadrant they target. Table with values corresponding to PPI placement in Figure 1-1 is shown below. Analysis performed by Dr. Andrea Thompson, UM.

PPI	K_D (μM)	Interface Area (Å^2)
Mdm2-p53[22]	0.06	1067-1481
XIAP-Bir3-SMAC[23]	0.42	650.6
BcL-XL-BAX[24]	0.015	502-828.6
IL2R-IL2[25]	0.01	1950
Cul1-CAND1[26]	High affinity	4898.6
Ran-RanBP2[27]	0.001	4928.6
MCP-1 dimer	0.5	1720
CBP KIX-cMyb[28]	15	1320
Ras-p120GAP[29]	17	3145
p53-CBP NRID[30]	10	2288
cadherin 8 homodimer[31]	80	2663.1

Table 1-1 – Associated values for Figure 1-1. Binding affinities and interface areas for PPIs within each category in Figure 1-1 are listed. Interface surface areas were calculated using 2P2IDB database for the PDB IDs listed in Figure 1-1 and binding affinities were pulled from the literature.

1.B. Transcriptional activator PPIs: a case study in the difficulties in studying and targeting transient interactions.

Transient and weaker affinity contacts are of particular interest in biological systems as they enable individual proteins to serve multiple functions, often by re-using the same interfaces. This strategy is heavily featured in transcriptional up-regulation, a process in which transcriptional activators engage in a multitude of interactions using a single transcriptional activation domain (TAD) to assemble the transcriptional machinery at the promoter of a gene. This assembly requires a series of dynamic binding interactions between the TAD and a variety of coactivator complexes, including chromatin-modifying enzymes, helicases, and scaffolding complexes (Figure 1-2).[32]

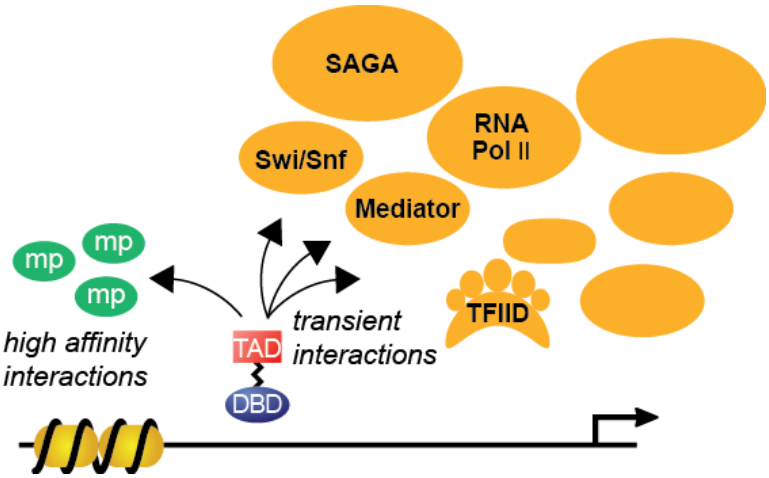

Figure 1-2 The transcriptional activation domain (TAD) of amphipathic activators can engage in high-affinity interactions, such as those with masking proteins (mp), but the interactions between the TAD and coactivator complexes are more moderate in affinity and transient in nature.

Activators can additionally make high affinity (low nM) contacts with masking proteins that repress their activity. These activator-masking protein interactions fall solidly in the area of Figure 1-1 that has proven to be the most amenable to targeting with small molecules. Indeed, a crystal structure of the activator p53 bound to its masking protein MDM2 shows an interaction more reminiscent of a PLI, with a concise alpha-helix of the p53 TAD binding tightly to a well-defined cleft on MDM2.[33] Not surprisingly, the p53-Mdm2 interaction is one of the most successfully targeted by small molecules due to its well-defined binding pocket, the structure of which was solved thanks to an advantageous compatibility of this class of PPIs with crystallization conditions.[34] Whereas interactions with masking proteins can be specific and high affinity, activator-coactivator interactions are often mediated through lower affinity, transient contacts.[32,35-39] As shown in Figure 1-3, the PPIs that p53 maintains with coactivators during recruitment are far more diverse in contact area (~800-2500 $Å^2$) and exhibit dissociation constants that fall in the high nM to low mM range.[33,40-42] Thus, the interactions most critical to understanding how p53 upregulates gene expression belong to the class of PPIs that have historically been the most challenging to study.

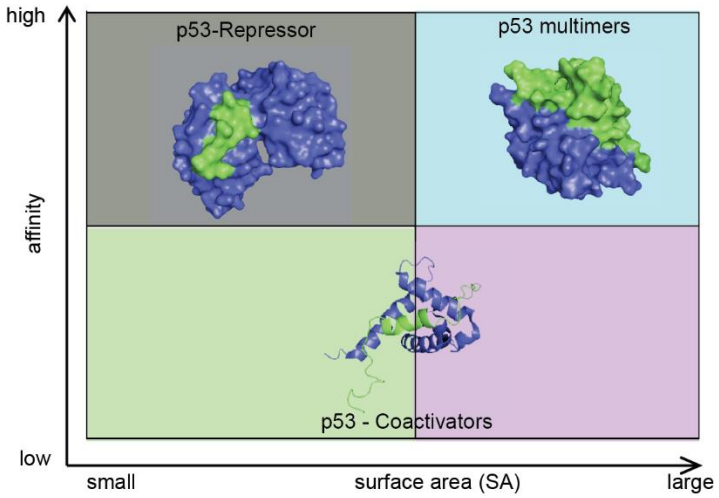

Figure 1-3 The field of transcriptional regulation is a prime example of a single system that engages in a diverse range of interactions. The transcriptional activator and tumor suppressor p53 engages in interactions that include the four major types of PPIs. The interaction between p53 and its repressor protein Mdm2 is high affinity with a concise surface area (PDB:1YCQ, grey, 70 nM, 1067 $Å^2$)[43] whereas the p53 tetramer (PDB: 1HS5, blue, 50 nM, 2502 $Å^2$)[44] also maintains a high affinity interaction but with a broader PPI interface. The interactions of p53 with the transcriptional machinery and coactivators such as the Taz2 domain of p300 (17 mM) [40] are often lower in affinity and range in interaction area (PDB:1 2GS0, 795 $Å^2$ to PDB: 2L14, 2288.6 $Å^2$). Thus, studying the interactions of a single system can pose significant challenges if the methods used are not suitable for covering this wide range of PPIs

1.B.1. Multi-protein complexes in transcriptional PPI networks

Further utilization of weaker affinity, transient PPIs in transcriptional systems can be seen in the coactivator complexes themselves, which are typically composed of at least one enzyme, such as an ATPase, and a series of non-enzymatic factors, such as scaffolding and regulatory proteins, that can associate either stably or transiently with the complex and help fine-tune activity, subcellular location, and/or selectivity. For example, the combinatorial assembly of BAF-type chromatin remodeling complexes involved in transcriptional regulation permits subunit exchange during the transition from a pluripotent stem cell into a neuron progenitor cell and finally into a fully differentiated neuron (Figure 1-4).[45,46] In mouse embryonic stem cells, an esBAF complex containing the BRG1 ATPase (but not BRM) and two copies of BAF 155 (but not BAF 170) is responsible for regulating transcriptional networks controlling pluripotency in ES

cells.[47,48] Interestingly, complexes containing BAF 155 and not BAF170 have yet to be identified in other cell types, suggesting the importance of this subunit specifically in regulating pluripotency.[49] Further use of subunit exchange and combinatorial assembly is seen in the formation of complex isoforms specific to the cell types in which they exist, including npBAF in neuron progenitor cells and nBAF in neuronal cells (Figure 1-4). These isoforms are believed to differentially interact with cellular complexes known to be critical to cellular differentiation and dendrite formation.[49] Thus, subunit exchange of multi-protein complexes regulates critical stages of cellular development. As another example, the Snf1 kinase/AMPK complex exchanges a regulatory subunit to dictate subcellular localization of the complex. In this case, the Gal83 complex isoform is believed to play the most significant role in regulating transcription of galactose inducible genes, whereas the Sip2 and Sip1 isoforms are believed to play important roles in the cytoplasm and vacuole, respectively. Thus, transient interactions play a central role in regulating transcriptional PPI networks. Further, in the case of multi-protein complexes, using small molecules to target the assembly and/or disassembly of the constituent contacts has the potential to uncover important insights into how different complex isoforms function.

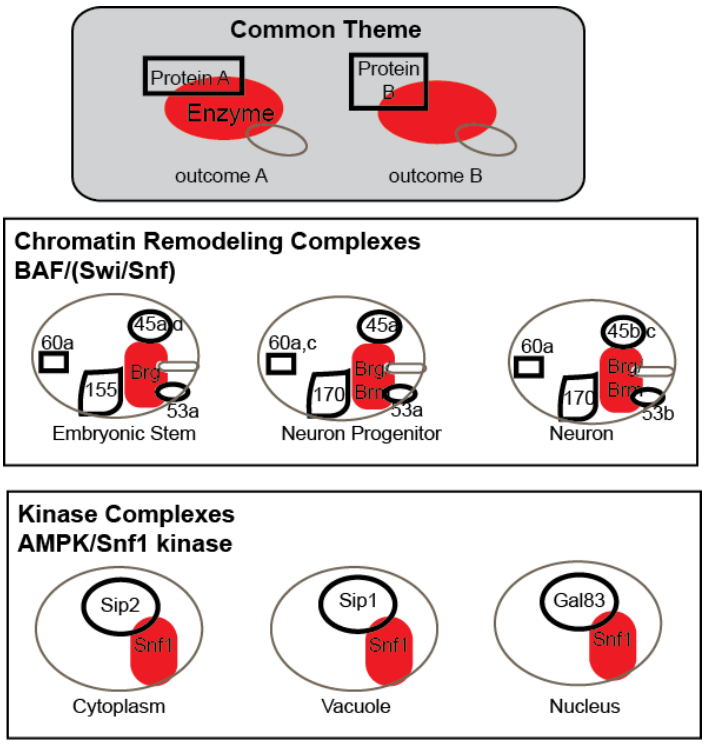

Figure 1-4 (Top) At the core of transcriptional PPI networks are multi-protein complexes that are often composed of at least one enzyme component (red) and several scaffolding and regulatory proteins that can associate either stably or transiently with the core subunit. In many cases, exchange of the non-enzymatic subunits can fine-tune complex localization and activity. (middle) As an example, the BAF-type remodeling complex uses subunit exchange to modulate the progression from an embryonic stem cell to a fully differentiated neuron.[49] (bottom) Subunit exchange dictates the subcellularl localization of the Snf1/AMPK complex.

Traditional routes for molecular probe discovery have developed compounds that inhibit the enzyme components of multi-protein complexes; however, in dynamically exchanging systems, targeting only the enzyme component may not be the most informative approach.[4] Rather, inhibiting or even promoting specific PPIs could be of even greater value for understanding transcriptional regulation and a wide range of essential cellular processes, a prospect that is supported by a growing appreciation of the potential of PPIs as drug targets.[50-53] However, achieving this goal has been fraught with difficulty, in part due to the challenges associated with characterizing transient, moderate affinity PPIs. The direct result of these difficulties is a shortage of the information necessary to guide molecular probe design.[7,54,55]

1.C. Challenges in studying activator-coactivator PPIs

To date, advances in understanding the underlying mechanism of activated transcription have largely been from studies that have focused on amphipathic activators, the largest and most well-characterized class in eukaryotes.[56] This family of activators includes the aforementioned p53 as well as such classic model systems as the viral activator VP16 and the yeast activators Gcn4 and Gal4. One key feature of amphipathic activators is their ability to activate transcription across eukaryotes, regardless of their species of origin within this domain of life.[57-60] As such, when fused to an appropriate DNA binding domain, the TADs of mammalian activators such as p53 are able to upregulate expression of protein encoding genes in *S. cerevisiae*.[61-63] This not only implies a high conservation of the fundamental components of this process across kingdoms, but it also suggests that studies of activators in less complex model organisms, such as yeast, can be used to form hypotheses about the mechanism of activated transcription in more complex systems. For this reason, *S. cerevisiae* has long been a critically important model organism for defining the mechanism of eukaryotic gene transcription.

1.C.1. Structure of activator PPIs

Amphipathic activators are so named due to the amino acid composition of their transcriptional activation domain (TAD), which is used to maintain contacts with inhibitory masking proteins, transcriptional complexes, and other contacts outside of the transcriptional machinery. The primary structure of amphipathic TADs is generally defined by stretches of acidic and polar residues interspersed with hydrophobic amino acids that are believed to play a substantial role in activator binding and function.[32,64-67] Like other intrinsically disordered proteins, the secondary structure of TADs is ill-defined in the absence of a binding partner, although it has been suggested that certain regions of the VP16 TAD adopt a partial α-helical conformation in solution.[68] Upon binding to a protein partner,

TADs have been shown to fold into an α-helix, with the hydrophobic face of the helix reported to contain many hotspots involved in maintaining binding interactions.[68-71] For example, the crystal structures of p53-MDM2 and Gal4-Gal80 beautifully complement biochemical data supporting the importance of hydrophobic residues in activator binding and function, with each structure showing hotspot residues in each TAD buried deep within the hydrophobic binding cleft on their respective masking proteins (Figure 1-5).[42,72,73] In fact, small molecules that inhibit the p53-Mdm2 interaction do so by mimicking just three hotspot amino acids in the p53 TAD (Phe19, Trp23, and Leu26) that contribute the most binding energy to the interaction.[74]

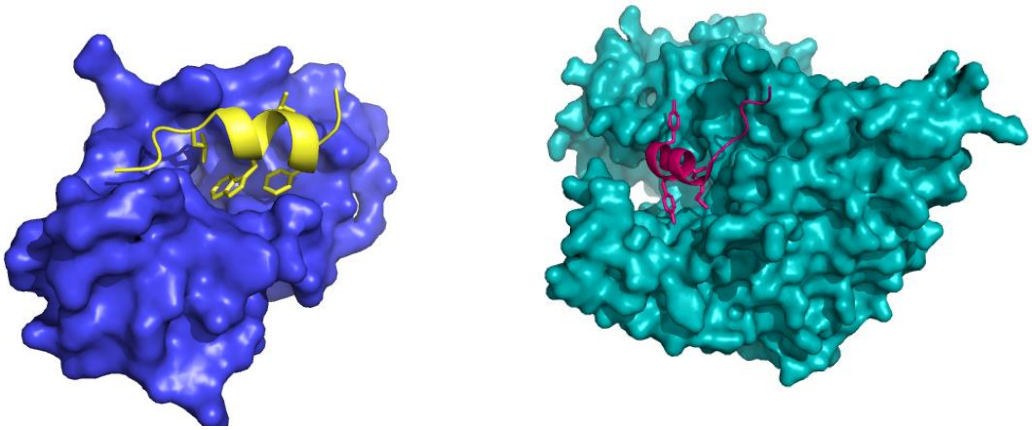

Figure 1-5 Activator-masking protein interactions are often defined by high affinity and small interaction areas. Both the p53-Mdm2(blue/yellow) and Gal4-Gal80 (cyan/magenta) interaction entail a small helix of the TAD of each activator binding a well defined groove on their respective masking proteins. The side chains of hydrophobic residues in each amphipathic TAD are shown as these are often considered to play a substantial role in these binding interactions. For example, Phe19, Trp23, and Leu26 in the p53 TAD are considered hotspots as they contributed the majority of binding energy in the p53-Mdm2 interaction.

In contrast, limited structural data is available for activator-coactivator interactions because transient, moderate-affinity PPIs are often not amenable to crystallization by standard techniques. As with most intrinsically disordered proteins, NMR spectroscopy has been the preferred method for examining the structure of unbound activators and weaker affinity activator-coactivator

interactions. Of the few high-resolution structures that do exist for activator-coactivator PPIs, most are limited to truncated segments of the TAD; thus, a complete picture of these binding events has proven difficult to obtain. Toward this goal, a recent approach by the Fersht group using residual dipolar coupling NMR combined with small angle X-ray scattering (SAXS) and molecular dynamics modeling was used to solve the structure of a modified p53 tetramer bound to the TAZ2 domain of p300.[75] The production of high-resolution structures of activator PPIs may be an increasingly realistic goal as more sensitive methods are developed and refined. However, as discussed in the next section, structural characterization of activator PPIs is contingent on the identification of the direct cellular targets of transcriptional activators, a significant roadblock that has hindered efforts toward this goal.

1.C.2 The elusive targets of amphipathic activators

The identification of the direct, cellular targets of amphipathic activators has proven to be very challenging. The absence of well-defined activator interaction profiles has stymied the discovery of small molecules that can modulate or mimic these PPIs.[32,76] For example, the few compounds capable of mimicking activator function were designed to contain hydrophobic and polar substituents as a means to simulate the amphipathic nature of activators.[77-79] However, these molecules only modestly upregulate gene expression and thus require significant optimization in order to achieve activation levels on par with natural activators. Toward this goal, elucidating the identities of activator targets *in vivo* and the subsequent characterization of these interactions should reveal key mechanistic information that can be used to tailor the activity of small molecules that can mimic or modulate activator function *in vivo.* Critical to this objective are methods capable of capturing the wide range of PPIs that activators maintain, especially the more transient interactions that have proven most challenging to study. Furthermore, because many of these PPIs occur in the context of multi-protein complexes that often employ dynamically exchanging subunits, it is

particularly useful to have a method that can capture these interactions in their native cellular environment.

1.C.2.a ChIP and genetic approaches to examine activator PPIs

Chromatin immunoprecipitation (ChIP) methods have been critical in establishing the fundamental role of activators in recruiting coactivator complexes and assembling the pre-initiation complex (PIC), which includes RNA polymerase II as well as other essential transcription factors. This approach uses formaldehyde to rapidly and non-specifically crosslink protein-DNA and protein-protein interactions in cells or on DNA templates, immobilizing the complexes that activators recruit to a particular promoter in order to regulate gene expression.[80] Thus, the advantage of ChIP is that it allows for the examination of activator recruitment on DNA, providing the most relevant information regarding initiation events in cells. Genetic studies have also contributed to this effort by introducing point mutations or deleting full proteins or protein domains and examining its effect on cell phenotype, complex recruitment, and functional output.[81-83] ChIP studies have helped lay the groundwork in identifying the proteins and complexes involved in transcriptional PPI networks. However, the nature of ChIP experiments does not allow one to easily distinguish a direct versus an indirect interaction between two proteins. Thus, the subunits within these recruited complexes that serve as activator targets *in vivo*, in addition to the location of activator binding sites within these targets, remain unclear via this route. Furthermore, genetic studies can be difficult to interpret owing to pleiotropic effects that result from deleting proteins that are critical to a number of PPI networks.[84] As such, a single deletion has the potential to affect more than just the transcriptional system under investigation. In addition, deletion of a protein required for the structural integrity of a complex can cause dissociation of the constituent subunits. For example, Gal4 has been shown to recruit the SAGA chromatin-modifying complex to the Gal1 promoter and genetic experiments were performed to determine which SAGA subunits participate in this recruitment.

While deletion of Spt8 resulted in a significant loss of SAGA recruitment by Gal4, the same result was achieved upon deletion of Spt7, Spt20 and Ada1, three subunits shown to be important for structural integrity of the complex.[85]. Thus, these approaches have left a significant knowledge gap regarding the specific coactivator targets of activators in cells.

1.C.*2.b Biochemical approaches to identify activator-coactivator PPIs*

Based on the results of genetic and ChIP studies, it is expected that activators target multiple protein partners throughout the course of transcription initiation. Indeed, biochemical approaches such as GST-pulldowns and *in vitro* crosslinking have highlighted the binding promiscuity of amphipathic activators, as demonstrated in Table 1-2 by their association with multiple targets. In fact, these analyses have generated over fifteen putative targets for VP16 (Table 1-2). *In vitro* binding studies have also identified interactions with little to no cellular relevance, including a reported interaction between Gal4 and bacterial lysozyme as well as binding of VP16 to Gal80. [86,87] This comes as a result of examining activator binding in the context of purified complexes or individual purified subunits, conditions that hardly resemble the complex and dynamic environment of the cell.[88] Providing additional support for studying activator PPIs in cells are studies that examined the recruitment of the Mediator complex by Gal4, leading to the identification of the Med15 subunit as a target of this activator during complex assembly at the promoter. Early characterization of the interaction between Gal4 and Med15 was carried out with purified components in solution. However, a recent examination of a ternary DNA-Gal4-Med15 complex suggests that the affinity of this interaction is enhanced by the presence of DNA, a finding that highlights the importance of examining activator-coactivator interactions in a cellular, DNA bound context.[35,89] Moreover, experiments with different classes of activators have shown that changing the location of DNA binding sites relative to the transcription start site can have vastly different outcomes on the transcriptional activity of the activator.[90] As such, *in vitro* experiments are likely

to overlook variables that may play a critical role in transcription *in vivo*, thus providing an incomplete or inaccurate representation of cellular events. Finally, many *in vitro* approaches, including GST-pulldowns and co-immunoprecipitations, rely on an affinity-based capture of activator PPIs and are therefore not ideally suited for the study of more transient, moderate affinity activator PPIs. Additionally, further analysis of these putative interactions reveals that activators can bind multiple sites within a single coactivator subunit. For example, p53 has been found to bind at least four different domains on the coactivator p300.[91] However, identifying which of these sites are functionally relevent in a cellular context has been difficult to establish.

Complex	Target	Activa	Method
Mediator	Med6	VP16	genetically introduced hexahistidine tags [92]
	Med16	Gcn4	in vitro crosslinking [38]
	Srb7	Gcn4	GST pulldown [93]
	Gal11/ Med15	VP16, Gal4, Gcn4	Far Western, VP16 affinity columns, SPR on polII holoenzymes purified from WT and mediator subunit null strains, in vitro crosslinking [94]
SAGA	Spt3	VP16	formaldehyde xlinking [95]
	Spt7	VP16	formaldehyde xlinking [95]
	Spt20	VP16	formaldehyde xlinking [95]
	Ada1	VP16	formaldehyde xlinking, genetically introduced hexahistidine tags [92,95]
	Tra1	VP16, Gal4	formaldehyde xlinking, Co-IP, in vitro crosslinking [95-98]
	Ada2	VP16	Co-IP [99]
	Ada3	VP16	formaldehyde xlinking [95]
	Gcn5	VP16	formaldehyde xlinking [95]
Swi/Snf	Snf2	VP16, Gcn4	genetically introduced hexahistidine tags,photo-crosslinking affinity label transfer method, GST-pulldown [92,100,101]
	Swi1	VP16, Gcn4	SPR, photo-crosslinking affinity label transfer method [36,100]
	Arp7	VP16, Gcn4	GST-pulldown [101]
	Swp73	Gcn4	GST-pulldown [93]

	Snf5	VP16, Gcn4	SPR, photo-crosslinking affinity label transfer method [36,100]
TFIIB	TFIIB	VP16	formaldehyde xlinking, affinity pulldown [95,102]
TFIID	Taf11	VP16	formaldehyde xlinking [95]
	Taf14	VP16	formaldehyde xlinking [95]
	Taf4	VP16	genetically introduced hexahistidine tags [92]
	TBP	VP16	genetically introduced hexahistidine tags, VP16 pulldown, SPR, formaldehyde crosslinking [36,92,95,103-106]
TFIID/ SAGA	Taf5	VP16	formaldehyde xlinking [95]
	Taf6	VP16	formaldehyde xlinking [95]
	Taf9	VP16	formaldehyde xlinking [95]
	Taf10	VP16	formaldehyde xlinking [95]
	Taf12	VP16, Gal4	formaldehyde xlinking, genetically introduced hexahistidine tags, in vitro crosslinking [92,95,98]
	Taf6	VP16	genetically introduced hexahistidine tags [92]
TFIIH	Tfb1	VP16	GST pulldown, NMR binding assay [107-109]
Proteasome	Sug1	Gal4, VP16	GST-VP16 pulldown using met labeled partner proteins [86]
RSC	Sth1	Gcn4	GST-pulldown [93]

Table 1-2 A variety of methods have been used to identify putative binding partners of transcriptional activators. This list is not meant to be inclusive, but rather highlight the promiscuous binding profile of transcriptional activators, particularly in an *in vitro* context.

C.2.c Investigating activator-coactivator interactions in vivo

In vivo approaches designed to address the limitations of the aforementioned *in vitro* methods have been used to identify the putative binding partners of activators in their native cellular environment. For example, yeast two hybrid (Y2H) studies have been heavily employed to look at PPIs in cells. This method is particularly advantageous for activators as it is sensitive enough to detect weakly interacting proteins; however, this sensitivity is also its downfall as Y2H screens are notoriously susceptible to high rates of false positives.[110] Another sensitive method, *in vivo* Forster Resonance Energy Transfer (FRET), has produced some convincing results regarding the interactions between Gal4 and

individual subunits of the SAGA chromatin modifying complex in yeast. Gal4 was tagged with a cyan fluorescent protein (CFP) and then tested for binding to enhanced yellow fluorescent protein (eYFP) tagged SAGA subunits. In the context of these studies, binding is implied upon enhancement of FRET as it indicates the two proteins under investigation are in close proximity (less than 50 Å). Of the coactivator subunits examined, only Tra1 showed FRET with Gal4, suggesting that this component was particularly critical for recruitment of SAGA.[111] While an important study, the absence of FRET for the remaining SAGA subunits (as well as several non-SAGA proteins identified in vitro to be Gal4 targets, including TBP) is difficult to interpret as a real result, especially given the sensitivity of FRET experiments to proper positioning of the fluorophores in use.[112] Furthermore, for proteins that primarily exist in the context of multi-protein complexes, the fluorophore can often be obscured or may disrupt subunit interactions within the complex.

C.2.d Use of crosslinking reagents to capture activator-coactivator PPIs

Perhaps the most convincing data in the identification of the direct interactions of transcriptional activators comes from studies using crosslinking moieties which capture and stabilize transient PPIs, affording the ability to study them under conditions that would otherwise prove disruptive. To date, numerous crosslinking reagents have been introduced to stabilize activator-coactivator interactions. Early experiments with chemical crosslinkers generally focused on incorporating the crosslinking moiety in a purified activator and analyzing the crosslinked products formed when introduced to either purified complexes or in whole cell lysates.[38,92,97,100] For example, in a series of elegant studies, the Hahn group has demonstrated the utility of crosslinking in capturing activator PPIs in vitro by using a PEAS crosslinker to capture the direct targets of Gal4 and Gcn4 in yeast cell lysate and later using a genetically incorporated photocrosslinking amino acid, p-benzoyl-L-phenylalanine (Bpa) to capture the interactions of TATA binding protein (TBP) both in vivo and in live yeast.[38,113] Their studies in yeast

demonstrated that TBP directly contacts the chromatin modifying SAGA complex, allowing them to map the molecular positions of protein components within the pre-initiation complex (PIC) in yeast.

The added advantage of using genetically incorporated photocrosslinking amino acids such as Bpa is that they can be site-specifically incorporated within a protein domain in cells, thus removing the requirement of introducing the crosslinking moiety post-translationally on a purified protein, such as that required when using popular PEAS, Sulfo-SBED and many NHS-ester crosslinking reagents. Site specific incorporation is achieved through the use of an evolved Bpa-specific tRNA synthetase (BpaRS) that selectively charges a tRNA variant whose anti-codon loop matches that of an amber stop codon (UAG) in the mRNA being translated.[114] Thus, mutation of specific codons in a gene of interest to an amber stop codon allows for Bpa to be site-selectively incorporated in the protein of interest *in vivo.* Furthermore, the tRNABpa and BpaRS react bio-orthogonally; the BpaRS does not acylate natural tRNAs with Bpa and the tRNA Bpa cannot be charged with natural amino acids by either endogenous synthetases or the BpaRS. The Schultz group first developed the use of nonsense suppression to incorporate unnatural amino acids such as Bpa into proteins in bacterial cells. Since their initial studies in *E. coli*, Bpa has subsequently been introduced in a similar fashion in eukaryotes including yeast and mammalian cells, allowing for the covalent capture of PPIs in the respective cell types.[115-118] Additional photo-crosslinking moieties have also been introduced into proteins via nonsense suppression including p-azidophenylalanine (pAzpa) and aliphatic diazirine-containing amino acids, thus providing several options with which to carry out *in vivo* photocrosslinking studies. While this technology can be used to validate the many putative partners that have already been identified, a more powerful application of this methodology comes with the use of mass spectrometry to identify novel activator interactions in a high-throughput manner. Finally, because activators can bind coactivators

both on and off DNA, the development of a method that allows for the capture of DNA bound activator interactions will be of great use in the mapping the contacts required specifically at the promoter of genes.

1.D. Overview of Thesis

In the remaining chapters of this thesis, I demonstrate the utility of Bpa crosslinking in live yeast in capturing the direct binding partners of transcriptional activators whose PPIs vary greatly in affinity, interaction area, and lifetimes. In Chapter 2, we show that Bpa crosslinking in yeast is capable of capturing moderate affinity activator interactions, specifically that between VP16 and the Mediator subunit Med15. We additionally demonstrate the capture of low abundance, endogenous targets by capturing the interactions between Gal4 and VP16 and components of the Swi/Snf chromatin-remodeling complex. Chapter 3 demonstrates the power of a novel method, Tandem Reversible and Irreversible Crosslinking (TRIC), in capturing the direct targets of DNA-bound activators, a goal that has not been accomplished until the creation of this approach. In Chapter 4, we use mass spectrometry to identify novel targets of the transcriptional activator Gal4, thus bringing us one step closer to a complete interaction profile for transcriptional activators. And finally, in Chapter 5, we discuss the future applications of these approaches in guiding small molecule discovery and the use of nonsense suppression to incorporate new chemistries into transcriptional activators, with the intent of yielding key insights into the mechanisms surround coactivator recruitment.

1.E. References for Chapter 1

(1) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. *ACS Chem. Biol.* **2012**, *7*, 1311.
(2) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. *Nat Rev Drug Discov* **2006**, *5*, 993.
(3) Surade, S.; Blundell, T. L. *Chem Biol* **2012**, *19*, 42.

(4) Hopkins, A. L.; Groom, C. R. *Nat Rev Drug Discov* **2002**, *1*, 727.
(5) Drews, J. *Science* **2000**, *287*, 1960.
(6) Stumpf, M. P.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. *Proc Natl Acad Sci U S A* **2008**, *105*, 6959.
(7) Berggard, T.; Linse, S.; James, P. *Proteomics* **2007**, *7*, 2833.
(8) Nooren, I. M.; Thornton, J. M. *J Mol Biol* **2003**, *325*, 991.
(9) Melcher, K. *Curr Protein Pept Sci* **2004**, *5*, 287.
(10) Perot, S.; Sperandio, O.; Miteva, M. A.; Camproux, A. C.; Villoutreix, B. O. *Drug Discov Today* **2010**, *15*, 656.
(11) Fuller, J. C.; Burgoyne, N. J.; Jackson, R. M. *Drug Discov Today* **2009**, *14*, 155.
(12) Lo Conte, L.; Chothia, C.; Janin, J. *J Mol Biol* **1999**, *285*, 2177.
(13) Nooren, I. M.; Thornton, J. M. *EMBO J* **2003**, *22*, 3486.
(14) Lee, C. W.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2010**, *49*, 9964.
(15) Patel, S. D.; Ciatto, C.; Chen, C. P.; Bahna, F.; Rajebhosale, M.; Arkus, N.; Schieren, I.; Jessell, T. M.; Honig, B.; Price, S. R.; Shapiro, L. *Cell* **2006**, *124*, 1255.
(16) Stauber, D. J.; Debler, E. W.; Horton, P. A.; Smith, K. A.; Wilson, I. A. *Proc Natl Acad Sci U S A* **2006**, *103*, 2788.
(17) Sun, H.; Stuckey, J. A.; Nikolovska-Coleska, Z.; Qin, D.; Meagher, J. L.; Qiu, S.; Lu, J.; Yang, C. Y.; Saito, N. G.; Wang, S. *J Med Chem* **2008**, *51*, 7169.
(18) Wu, G.; Chai, J.; Suber, T. L.; Wu, J. W.; Du, C.; Wang, X.; Shi, Y. *Nature* **2000**, *408*, 1008.
(19) Vaynberg, J.; Fukuda, T.; Chen, K.; Vinogradova, O.; Velyvis, A.; Tu, Y.; Ng, L.; Wu, C.; Qin, J. *Mol Cell* **2005**, *17*, 513.
(20) Bourgeas, R.; Basse, M. J.; Morelli, X.; Roche, P. *PLoS One* **2010**, *5*, e9598.
(21) Higueruelo, A. P.; Schreyer, A.; Bickerton, G. R.; Pitt, W. R.; Groom, C. R.; Blundell, T. L. *Chem Biol Drug Des* **2009**, *74*, 457.
(22) Lai, Z.; Auger, K. R.; Manubay, C. M.; Copeland, R. A. *Arch Biochem Biophys* **2000**, *381*, 278.
(23) Liu, Z.; Sun, C.; Olejniczak, E. T.; Meadows, R. P.; Betz, S. F.; Oost, T.; Herrmann, J.; Wu, J. C.; Fesik, S. W. *Nature* **2000**, *408*, 1004.
(24) Ku, B.; Liang, C.; Jung, J. U.; Oh, B.-H. *Cell Res* **2011**, *21*, 627.
(25) Boyman, O.; Sprent, J. *Nat Rev Immunol* **2012**, *12*, 180.
(26) Feng, S.; Shen, Y.; Sullivan, J. A.; Rubio, V.; Xiong, Y.; Sun, T.-P.; Deng, X. W. *The Plant Cell* **2004** , *16*, 1870.
(27) Seewald, M. J.; Kraemer, A.; Farkasovsky, M.; Korner, C.; Wittinghofer, A.; Vetter, I. R. *Mol Cell Biol* **2003**, *23*, 8124.
(28) Zor, T.; Mayr, B. M.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. *J Biol Chem* **2002**, *277*, 42241.
(29) Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmuller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A. *Science* **1997**, *18*, 333.

(30) Ferreon, J. C.; Lee, C. W.; Arai, M.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Proc Natl Acad Sci U S A* **2009**, *106*, 6591.
(31) Vendome, J.; Posy, S.; Jin, X.; Bahna, F.; Ahlsen, G.; Shapiro, L.; Honig, B. *Nat Struct Mol Biol* **2011**, *18*, 693.
(32) Mapp, A. K.; Ansari, A. Z. *ACS Chem Biol* **2007**, *2*, 62.
(33) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948.
(34) Wang, W.; Hu, Y. *Med Res Rev* **2012**, *32*, 1159.
(35) Wands, A. M.; Wang, N.; Lum, J. K.; Hsieh, J.; Fierke, C. A.; Mapp, A. K. *J Biol Chem* **2010**, *286*, 16238.
(36) Ferreira, M. E.; Hermann, S.; Prochasson, P.; Workman, J. L.; Berndt, K. D.; Wright, A. P. H. *J Biol Chem* **2005**, *280*, 21779.
(37) Fuxreiter, M.; Tompa, P.; Simon, I. *Bioinformatics* **2007**, *23*, 950.
(38) Herbig, E.; Warfield, L.; Fish, L.; Fishburn, J.; Knutson, B. a.; Moorefield, B.; Pacheco, D.; Hahn, S. *Mol Cell Biol* **2010**, *30*, 2376.
(39) Melcher, K. *J Mol Biol* **2000**, *301*, 1097.
(40) Feng, H.; Jenkins, L. M.; Durell, S. R.; Hayashi, R.; Mazur, S. J.; Cherry, S.; Tropea, J. E.; Miller, M.; Wlodawer, A.; Apella, E.; Bai, Y. *Structure* **2009**, *17*, 202.
(41) Lee, W.; Harvey, T. S.; Yin, Y.; Yau, P.; Litchfield, D.; Arrowsmith, C. H. *Nat Struct Biol* **1994**, *1*, 877.
(42) Popowicz, G. M.; Czarna, A.; Holak, T. A. *Cell Cycle* **2008**, *7*, 2441.
(43) Sakaguchi, K.; Saito, S.; Higashimoto, Y.; Roy, S.; Anderson, C. W.; Appella, E. *J Biol Chem* **2000**, *275*, 9278.
(44) Rajagopalan, S.; Huang, F.; Fersht, A. R. *Nuc Acids Res* **2010**, 1.
(45) Lessard, J.; Wu, J. I.; Ranish, J. A.; Wan, M.; Winslow, M. M.; Staahl, B. T.; Wu, H.; Aebersold, R.; Graef, I. A.; Crabtree, G. R. *Neuron* **2007**, *55*, 201.
(46) Yoo, A. S.; Crabtree, G. R. *Curr Opin Neurobiol* **2009**, *19*, 120.
(47) Ho, L.; Ronan, J. L.; Wu, J.; Staahl, B. T.; Chen, L.; Kuo, A.; Lessard, J.; Nesvizhskii, A. I.; Ranish, J. A.; Crabtree, G. R. *Proc Natl Acad Sci U S A* **2009**, *106*, 5181.
(48) Yan, Z.; Wang, Z.; Sharova, L.; Sharov, A. A.; Ling, C.; Piao, Y.; Aiba, K.; Matoba, R.; Wang, W.; Ko, M. S. H. *Stem Cells* **2008** , *26* , 1155.
(49) Ho, L.; Crabtree, G. R. *Nature* **2010**, *463*, 474.
(50) Cravatt, B. F.; Simon, G. M.; Yates, J. R., 3rd *Nature* **2007**, *450*.
(51) Powers, E. T.; Morimoto, R. I.; Dillin, A.; Kelly, J. W.; Balch, W. E. *Annu Rev Biochem* **2009**, *78*, 959.
(52) Makley, L. N., Gestwicki, J.E. *Chem Biol Drug Des* **2013**, *81*, 22.
(53) Thiel, P.; Kaiser, M.; Ottmann, C. *Angew Chem Int Ed Engl* **2012**, *51*, 2012.
(54) Perkins, J. R.; Diboun, I.; Desailly, B. H.; Lees, J. G.; Orengo, C. *Structure* **2010**, *18*, 1233.
(55) A., S.; Schaft, D.; Roguev, A.; Pijnappel, W. W. M. P.; Stewart, A. F.; Shevchenko, A. *Mol Cell Proteomics* **2002**, *1*, 204.

(56) Ptashne, M.; Gann, A. *Genes & Signals*; Cold Spring Harbor Laboratory: New York, 2001.
(57) Fischer, J. A.; Giniger, E.; Maniatis, T.; Ptashne, M. *Nature* **1988**, *332*, 853.
(58) Ma, J.; Przibilla, E.; Hu, J.; Bogorad, L.; Ptashne, M. *Nature* **1988**, *334*, 631.
(59) Kakidani, H.; Ptashne, M. *Cell* **1988**, *52*, 161.
(60) Webster, N.; Jin, J. R.; Green, S.; Hollis, M.; Chambon, P. *Cell* **1988**, *52*, 169.
(61) Smardova, J.; Smarda, J.; Koptikova, J. *Differentiation* **2005**, *73*, 261.
(62) Scharer, E.; Iggo, R. *Nuc Acids Res* **1992**, *20*, 1539.
(63) Tokino, T.; Thiagalingam, S.; el-Deiry, W. S.; Waldman, T.; Kinzler, K. W.; Vogelstein, B. *Hum Mol Genet* **1994**, *3*, 1537.
(64) Regier, J. L., Shen, F., Triezenberg, S.J. *Proc Natl Acad Sci USA* **1993**, *90*, 883.
(65) Cress, W. D.; Triezenberg, S. J. *Science* **1991**, *251*, 87.
(66) Giniger, E.; Ptashne, M. *Nature* **1987**, *330*, 670.
(67) Ruden, D. M. *Chromosoma* **1992**, *101*, 342.
(68) Jonker, H. R.; Wechselberger, R. W.; Boelens, R.; Folkers, G. E.; Kaptein, R. *Biochemistry* **2005**, *44*, 827.
(69) Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L. *Science* **1997**, *277*, 1310.
(70) Shen, F.; Triezenberg, S. J.; Hensley, P.; Porter, D.; Knutson, J. R. *J Biol Chem* **1996**, *271*, 4827.
(71) Langlois, C.; Mas, C.; Di Lello, P.; Jenkins, L. M. M.; Legault, P.; Omichinski, J. G. *J Am Chem Soc* **2008**, *130*, 10596.
(72) Thoden, J. B.; Ryan, L. A.; Reece, R. J.; Holden, H. M. *J Biol Chem* **2008**, *283*, 30266.
(73) Kumar, P. R. Y., Y.; Sternglanz, R.; Johnston, S.A.; Joshua-Tor, L. *Science* **2008**, *319*, 1090.
(74) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. *Science* **2004**, *303*, 844.
(75) Wells, M.; Tidow, H.; Rutherford, T. J.; Markwick, P.; Jenson, M. R.; Mylonas, E.; Svergun, D. I.; Blackledge, M.; Fersht, A. R. *Proc Nat Acad Sci, USA* **2008**, *105*, 5762.
(76) Lee, L. W.; Mapp, A. K. *J Biol Chem* **2010**, *285*, 11033.
(77) Minter, A. R.; Brennan, B. B.; Mapp, A. K. *J Am Chem Soc* **2004**, *126*, 10504.
(78) Xiao, X.; Yu, P.; Lim, H.-S.; Sikder, D.; Kodadek, T. *Angew Chem Int Ed Engl* **2007**, *46*, 2865.
(79) Dongju, J.; Shimogawa, H.; Kwon, Y.; Mao, Q.; Sato, S.-i.; Kamisuki, S.; Kigoshi, H.; Uesugi, M. *J Am Chem Soc* **2009**, *131*, 4774.
(80) O., A.; Geisberg, J. V.; Sekinger, E.; Yang, A.; Mogtaderi, Z.; Struhl, K. *Curr Protoc Mol Biol* **2005**, *Ch. 21*.

(81) White, M.; Riles, L.; Cohen, B. A. *Genetics* **2009**, *181*, 435.
(82) Yousef, A. F.; Xu, G. W.; Mendez, M.; Brandl, C. J.; Mymryk, J. S. *Int. J. Cancer* **2008**, *122*, 942.
(83) Hu, Z.; Killion, P. J.; Iyer, V. R. *Nat Genet* **2007**, *39*, 683.
(84) Dudley, A. M.; Janse, D. M.; Tanay, A.; Shamir, R.; G.M., C. *Mol Syst Biol* **2005**, *1*, 2005.0001.
(85) Bhaumik, S. R.; Green, M. R. *Mol Cell Biol* **2002**, *22*, 7365.
(86) Swaffield, J. C.; Melcher, K.; Johnston, S. A. *Nature* **1995**, *374*.
(87) Vashee, S.; Kodadek, T. *Proc Nat Acad Sci, USA* **1995**, *92*, 10683.
(88) Barberis, A.; Petrascheck, M. *Encylopedia of Life Sciences* **2003**, 1.
(89) Majmudar, C. Y.; Wang, B.; Lum, J. K.; Hakansson, K.; Mapp, A. K. *Angew Chem Int Ed Engl* **2009**, *48*, 7021.
(90) Remacle, J. E.; Albrecht, G.; Brys, R.; Braus, G. H.; Hyulebroeck, D. *EMBO J* **1997**, *16*, 5722.
(91) Teufel, D. P.; Freund, S. M.; Bycroft, M.; Fersht, A. R. *Proc Nat Acad Sci, USA* **2007**, *104*, 7009.
(92) Klein, J.; Nolden, M.; Sanders, S. L.; Kirchner, J.; Weil, P. A.; Melcher, K. *J Biol Chem* **2003**, *278*, 6779.
(93) Swanson, M. J., Qiu, H., Sumibcay, L., Krueger, A., Kim, S., Natarajan, K., Yoon, S., and Hinnebusch, A.G. *Mol Cell Biol* **2003**, *23*, 2800.
(94) Lee, Y. C.; Park, J. M.; Min, S.; Ii, P.; Mol, H.; Biol, C.; Kim, Y.-j. *Mol Cell Biol* **1999**, *19*, 2967.
(95) Hall, D. B.; Struhl, K. *J Biol Chem* **2002**, *277*, 46043.
(96) Brown, C. E.; Howe, L.; Sousa, K.; Alley, S. C.; Carrozza, M. J.; Tan, S.; Workman, J. L. *Science* **2001**, *292*, 2333.
(97) Reeves, W. M.; Hahn, S. *Mol Cell Biol* **2005**, *25*, 9092.
(98) Fishburn, J.; Mohibullah, N.; Hahn, S. *Mol cell* **2005**, *18*, 369.
(99) Silverman, N.; Agapite, J.; Guarente, L. *Proc Nat Acad Sci, USA* **1994**, *91*, 11665.
(100) Neely, K. E.; Hassan, A. H.; Brown, C. E.; Howe, L.; Workman, J. L. *Mol Cell Biol* **2002**, *22*, 1615.
(101) Neely, K. E.; Hassan, A. H.; Wallberg, A. E.; Steger, D. J.; Cairns, B. R.; Wright, A. P.; Workman, J. L. *Mol Cell* **1999**, *4*, 649.
(102) Lin, Y. S.; Ha, I.; Maldonado, E.; Reinberg, D.; Green, M. R. *Nature* **1991**, *353*, 569.
(103) Shooltz, D. D.; Alberts, G. L.; Triezenberg, S. J. *Protein Expr Purif* **2008**, *59*, 297.
(104) Stringer, K. F.; Ingles, C. J.; Greenblatt, J. *Nature* **1990**, *345*, 783.
(105) Nedialkov, Y. A.; Triezenberg, S. J. *Arch Biochem Biophys* **2004**, *425*, 77.
(106) Ingles, C. J.; Shales, M.; Cress, W. D.; Triezenberg, S. J.; Greenblatt, J. *Nature* **1991**, *351*, 588.
(107) Xiao, H.; Pearson, A.; Coulombe, B.; Truant, R.; Zhang, S.; Regier, J. L.; Triezenberg, S. J.; Reinberg, D.; Flores, O.; Ingles, C. J.; Greenblatt, J. *Mol Cell Biol* **1994**, *14*, 7013.

(108) Di Lello, P.; Nguyen, B. D.; Jones, T. N.; Potempa, K.; Kobor, M. S.; Legault, P.; Omichinski, J. G. *Biochemistry* **2005**, *44*, 7678.
(109) Langlois, C.; Mas, C.; Di Lello, P.; Jenkins, L. M. M.; Legault, P.; Omichinski, J. G. *J Am Chem Soc* **2008**, *130*, 10596.
(110) Bruckner, A.; Polge, C.; Lentze, N.; Auerbach, D.; Schlattner, U. *Int J Mol Sci* **2009**, *10*, 2763.
(111) Bhaumik, S. R.; Raha, T.; Aiello, D. P.; Green, M. R. *Genes Dev* **2004**, *18*, 333.
(112) Piston, D. W.; Kremers, G.-J. *Trends in biochemical sciences* **2007**, *32*, 407.
(113) Mohibullah, N.; Hahn, S. *Genes Dev* **2008**, *22*, 2994.
(114) Wang, Q.; Parrish, A. R.; Wang, L. *Chem Biol* **2009**, *16*, 323.
(115) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964.
(116) Hino, N.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. *Nat protocols* **2006**, *1*, 2957.
(117) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A. *Nat Methods* **2005**, *2*, 3.
(118) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J Am Chem Soc* **2009**, *131*, 14240.

Chapter 2 In vivo photocrosslinking captures transient, moderate affinity activator-coactivator interactions in living cells

2.A. Summary[2]

Transient and moderate affinity protein-protein interactions (PPIs) play a critical role in the regulation of essential cellular processes including protein folding, ubiquitylation, and transcription. A number of disease states are believed to be the result of aberrations within these protein networks. Therefore, a longstanding therapeutic goal has been to design small molecules that can tunably modulate the constituent interactions.[3-11] However, as discussed in Chapter 1, the discovery of small molecule modulators has been hindered by a lack of structural and mechanistic information due to limitations of the approaches currently available for studying transient PPIs in their native environment. For example, *in vitro* co-crystallization and co-purification and *in vivo* two-hybrid studies are best suited for probing stably associated proteins, but are less ideal for studying proteins that engage in modest affinity and/or transient multi-protein binding interactions.[12-16] Thus, the limitations of the current methods demonstrate a

[2] Portions of this chapter are from two published manuscripts. (1) Krishnamurthy, M.; Dugan, A.; Nwokoye, A.; Fung, Y.-H.; Lancia, J. K.; Majmudar, C. Y.; Mapp, A. K. *ACS Chem Biol* **2011**, *6*, 1321. and (2) Lancia, J. K.; Nwokoye, A.; Dugan, A.; Pricer, R.; Joiner, C.; Mapp, A. K. *Biopolymers* **2013**, *manuscript accepted*.

The individual contributions to the data presented in this chapter are as follows: Dr. Malathy Krishnamurthy was responsible for the creation and testing of LexA+VP16C constructs as well as the design and creation of the Swi1 and Snf5 deletion strains. Amanda Dugan created and tested the LexA+VP16N constructs. Hugo Fung and Adaora Nwokoye tested crosslinking of Gcn4 and Gal4, respectively, to Snf2. Cassandra Joiner and Amanda Dugan designed and created the LexA+Gal4 F856Bpa/Azpa methionine mutants and, along with Rachel Pricer, tested the incorporation, activity, and crosslinking of Bpa and Azpa mutants. Dr. Chinmay Majmudar provided intellectual contributions.

need for approaches that can probe these interactions in their native cellular environment.

In this chapter, we address this issue by employing an *in vivo* photocrosslinking method that uses a genetically incorporated unnatural amino acid, p-benzoyl-L-phenylalanine (Bpa). We demonstrate the utility of this method in capturing the moderate affinity binding partners of the prototypical activator VP16 in live yeast, with particular focus on a previously identified putative interaction with Med15. Furthermore, *in vivo* photocrosslinking is used to examine the shared targets of amphipathic activators in yeast, including low abundance complexes such as the chromatin-modifying coactivator complex Swi/Snf. Specifically, we capture subunits within the Swi/Snf complex that may serve as a common handle for activators during recruitment of this complex in cells. Furthermore, using the interaction between the yeast activator Gal4 and its suppressor Gal80 as a model, we address some factors that should be considered when interpreting negative results of crosslinking experiments (i.e. when crosslinking to a target is not observed), such as the positional context of the crosslinker as well as differential reactivities of the crosslinking moieties. Taken together, the data presented in this chapter demonstrates the utility of genetically encoded photo-activatable amino acids in characterizing activator-coactivator complexes in vivo and further suggests that this strategy can be implemented more broadly for the capture and discovery of transient protein-protein interactions in their native environment.

2.B. Background : Activator-coactivator complex formation

Transcriptional activators are signal responsive regulatory proteins that assemble the transcriptional machinery at the promoter of a gene through dynamic binding interactions with a variety of coactivator complexes, including chromatin-modifying, helicase, and scaffolding complexes.[11,17,18] Activators are modular in architecture and are minimally composed of a DNA binding domain (DBD) that localizes the activator to its cognate DNA binding site and a transcriptional

activation domain (TAD) that mediates the majority of contacts with transcriptional complexes. Based on biochemical data, we know that the interactions between activators and suppressor proteins tend to be high affinity and specific in nature whereas activator-coactivator interactions appear to be mediated through lower affinity transient contacts (Figure 2-1).[11,16,18-21] As mentioned in Chapter 1, this characteristic of activators has made the study and characterization of activator-coactivator interactions a tremendously challenging endeavor.

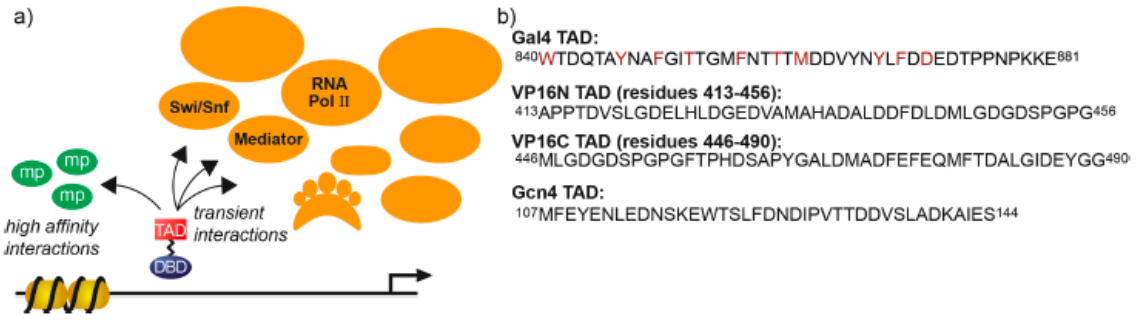

Figure 2-1 (a) The transcriptional activation domain (TAD) of amphipathic activators can engage in high affinity interactions, such as those with masking proteins (mp), but the interactions between the TAD and coactivator complexes are more moderate in affinity and transient in nature.[11,16 18-21] (b) Amphipathic activators share little sequence homology but do share binding targets, at least *in vitro*. The photocrosslinking amino acid, Bpa, has been incorporated within the Gal4 TAD (positions of incorporation highlighted in red) with little impact on the function and binding profile of the TAD.[17]

In vivo co-localization studies have broadly defined the complexes that are recruited by amphipathic activators during transcription but have not readily provided information on the specific, direct coactivator targets within these complexes.[22-24] For example, the well-characterized amphipathic activator VP16 has been shown to recruit the Swi/Snf chromatin-remodeling complex early in transcription initiation, as evidenced by both *in vivo* and *in vitro* co-localization studies.[25-30] *In vitro* assays have identified several subunits within this complex as possible targets of VP16 but *in vivo* interaction studies have not distinguished with of the components are the relevant binding partner(s) in cells.[19,31,32] Additionally, the results of *in vitro* binding studies do not always guarantee a true interaction *in vivo*. The intrinsically disordered nature of amphipathic TADs,

combined with their hydrophobic amino acid composition, affords them a promiscuous binding profile, particularly *in vitro*. *In vitro* binding experiments with the activator Gal4 identified several subunits within the transcriptional machinery as putative targets, but this approach also identified bacterial lysozyme as a binding partner, an unnatural interaction whose relevance in cells is non-existent.[33] As such, *in vivo* methodologies that can capture transient activator-coactivator interactions in their native environment would be extremely beneficial in identifying the relevant cellular binding partners of transcriptional activators.

2.C. Covalent chemical capture in cells

Covalent chemical capture using genetically incorporated photocrosslinking amino acids such as Bpa is a promising tool in the identification of directly interacting PPIs in cells. This approach requires the development of a nonsense suppression strategy to incorporate Bpa site-specifically into proteins in living cells and, upon irradiation of these cells, capture the direct targets of Bpa-containing proteins. Development of this strategy was initiated by the Schultz lab and then further optimized by the Mapp group, totaling over two decades of development efforts to make this a feasible strategy for the in vivo capture of activator targets in yeast.[17,34] Briefly, incorporation of Bpa into proteins in yeast is accomplished using a Bpa-specific tRNA synthetase (BpaRS), derived from wild type E. coli Tyrosyl RS, to selectively charge a tRNA variant whose anti-codon loop matches that of an amber stop codon (UAG) in the mRNA being translated. Thus, mutation of specific codons in the open reading frame of a gene to an amber stop codon allows for Bpa to be site-selectively incorporated in the protein of interest. Furthermore, the tRNABpa and BpaRS react bio-orthogonally, therefore the BpaRS does not acylate natural tRNAs with Bpa and the tRNA Bpa should not be charged with natural amino acids by either endogenous synthetases or the BpaRS (Figure 2-2).[35]

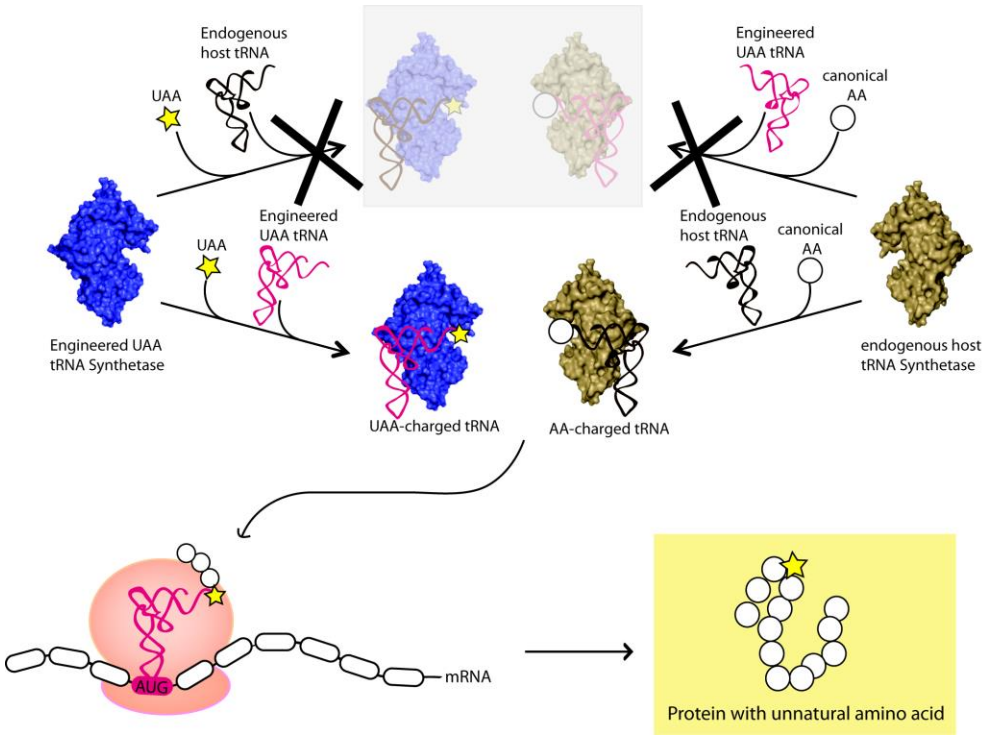

Figure 2-2 Nonsense suppression allows genetic incorporation of unnatural amino acids (UAAs) into proteins and peptides. Successful expression of a protein (or peptide) with a UAA incorporated site specifically requires a tRNA/aminoacyl tRNA synthetase pair (red and blue, respectively) that recognizes a rare stop codon such as the AUG amber stop codon during the translation process. In order for incorporation to be selective, the tRNA/aminoacyl tRNA synthetase pair must be orthogonal to the tRNA/aminoacyl synthetase pairs that install the canonical amino acids.

2.C.1. In vivo photocrosslinking in Saccharomyces cerevisiae

In early work pioneered by the Schultz lab, Bpa incorporation in yeast was first demonstrated to occur in high yields and with high fidelity as measured by expression of the highly abundant and well-expressed protein, superoxide dismutase.[34] However, as documented by our lab, transfer of this technology to more challenging protein systems required extensive optimization to increase overall mutant protein expression and ultimately support a more global application of this technology in yeast.[2,17] These studies found that expression of the Bpa tRNA/RS pair could be increased by placing it under the control of a Pol III pSNR52 promoter, thus enhancing incorporation of Bpa in the activator Gal4 as a result of increased Bpa tRNA/RS availability. Furthermore, unlike other

promoters tested to achieve this result, the pSNR52 promoter was able to drive expression of the tRNA/RS without impacting the fidelity of the system. Using this optimized system, Dr. Chinmay Majmudar and colleagues in our lab found that Bpa could be incorporated into TADs with high yields and with little impact on transcriptional activator binding and function.

Once the optimal incorporation conditions were determined, our group next performed crosslinking experiments to evaluate if the *in vivo* binding partners of Gal4 could be captured using this approach. Bpa is advantageous for *in vivo* studies as it can be activated at longer UV wavelengths (~365 nm), thus reducing the amount of protein and DNA damage that can occur as a result of irradiation at higher energy wavelengths. Photoactivation of Bpa at 365 nm leads to the covalent capture of the direct binding partners of the activators which can then be isolated from yeast lysates after breaking irradiated cells open and analyzed by immunological methods such as Western blotting (as discussed in Chapters 2 and 3) or by mass spectrometric methods (as discussed in Chapter 4) (Figure 2-3).

Crosslinking studies our group and others went on to demonstrate that in vivo crosslinking with Bpa is a useful method for capturing direct, high-affinity PPIs.[17,36-38] More specifically, Bpa placement within the TAD of the activator Gal4 did not impair function of the protein and photoactivation led to covalent capture of its high affinity (low nanomolar K_D) suppressor protein Gal80.[17] However, while successful in the case of a very tight PPI, this method has not been employed in the case of moderate-affinity, transient interactions such as those between activators and coactivators. In the following sections, we test the utility of *in vivo* Bpa crosslinking in capturing VP16-coactivator interactions and for resolving the identity of the Swi/Snf components targeted by this activator.

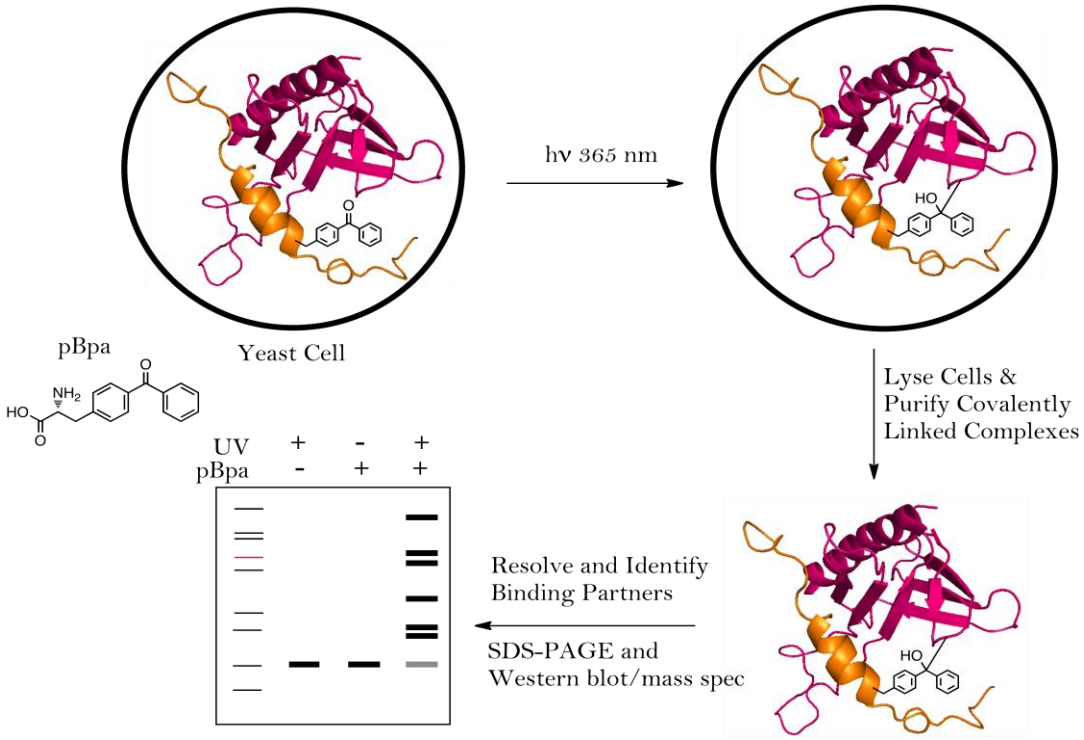

Figure 2-3 . Schematic of in vivo photocrosslinking in yeast with Bpa. Bpa is incorporated into the TAD of an activator in live yeast using nonsense suppression. These yeast cells are then irradiated with UV light (365 nm) to activate the Bpa, leading to covalent capture of direct binding partners through C-H insertion chemistry. The cells are then lysed and the covalent adducts are isolated using a variety of purification methods. The identity of the crosslinked partner can then be identified through traditional immunological detection methods such as Western blotting or resolved by mass spectrometric methods to identify novel binding partners of transcriptional activators.

2.D. The viral activator VP16

During lytic infection of the herpes simplex virus (HSV-1), the amphipathic activator VP16 upregulates the expression of immediate-early genes required for the replication and survival of the virus.[39] As expected for such a critical viral protein, VP16 is a powerful activator, able to initiate transcription rapidly and efficiently.[40] In contrast to many other amphipathic activators, VP16 lacks a covalently bound DBD, instead using its core domain to bind two host proteins, Oct1 and Hcf1, to largely mediate its contact with DNA.[41] Functional experiments on the VP16 TAD reveal that the full TAD (413-490) is actually comprised of two sub-domains that can function independently from one another, an amino

terminal VP16N (residues 413-456) and a carboxy terminal VP16C (residues 446-490) (Figure 2-1,b).[42,43] Further analysis of VP16N identified a minimal peptide sequence (DFDLDMLG) that was sufficient to activate transcription when tethered to a DNA binding moiety, with transcriptional activity increasing as multiple copies of the peptide were introduced.[44-48] As such, experiments focusing on VP2, an activation domain bearing two copies of the DFDLDMLG sequence, yielded significant insights on how natural activators function. Biochemical studies with VP2 and each subdomain of VP16 have identified a multitude of putative binding partners, such as general transcription factors including TBP and its associated TAFs as well as components of the Mediator, SAGA, and Swi/Snf complexes, among others.[24,26,27,29,32,49-67] However, given the largely *in vitro* context of these experiments, the role of important factors including coactivator dynamics and localization, chromatin organization, and proteasomal regulation remains untested under these conditions. Therefore, while numerous putative binding partners of VP16 have been identified, the direct targets of VP16 in the native cellular context remains unclear.

2.D.1. Incorporation of Bpa in the VP16 TAD

To discover proteins that interact with different regions of the VP16 TAD *in vivo*, we incorporated Bpa within regions of each subdomain shown to be involved in forming protein interactions (VP16N: L439, F442, L444; VP16C: F473, F475, F479).[43,68,69] Each Bpa-containing construct was expressed in *Saccharomyces cerevisiae* as a fusion protein bearing the bacterial LexA DBD and a carboxy-terminal Flag tag for detection (Figure 2-4,a). All six Bpa mutants were assayed for Bpa incorporation and activation potential in a yeast strain with an integrated LacZ reporter gene under the control of a GAL1 promoter bearing two LexA binding sites (Figure 2-4 b, c). Therefore, the advantage of using a heterologous activator construct is that a single promoter containing binding sites for LexA exists within the yeast strain used in our studies, thus simplifying functional analyses of all mutant proteins. As in previous studies, we found that efficient

Bpa crosslinking was affected by positioning of Bpa within the protein.[17,70] As such, the LexA+VP16C F479Bpa mutant was removed from further testing as it displayed poor incorporation. LexA+VP16N L444Bpa and LexA+VP16C F475Bpa, on the other hand, showed good incorporation and activity (Figure 2-4 b, c), demonstrating that these two constructs were the most suitable for use in crosslinking experiments. We subsequently examined the crosslinking profiles of each of these constructs and found that, as expected, each maintains a multi-protein binding profile, with several covalent crosslinked complexes detected by Western blot (Figure 2-5). Additionally, these crosslinking profiles were consistently repeatable; as a result, these two constructs were selected to be the focus of additional crosslinking investigations.

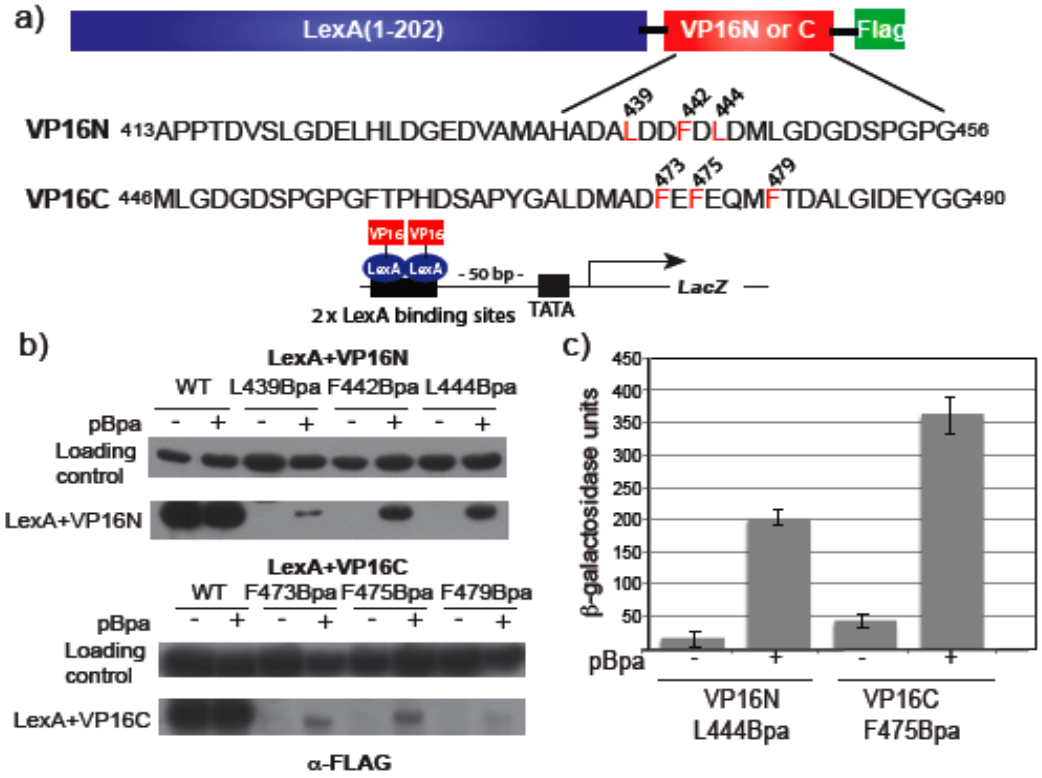

Figure 2-4 Incorporation of Bpa within the VP16 TAD. (a) Plasmids encoding the DNA binding domain (DBD) of LexA fused to either the N- or C-terminal VP16 TAD as well as a FLAG tag were constructed. The LexA DBD was utilized to exclusively examine transcriptional activation at the 2 unique LexA binding sites upstream of the LacZ reporter in *S. cerevisiae.* Positions at which Bpa mutagenesis was carried out are within regions of the VP16N or VP16C subdomains known to participate in coactivator binding (sites of

incorporation highlighted in red). (b) Yeast cells bearing plasmids encoding the various LexA+VP16 constructs and the Bpa specific tRNA/synthetase pair expressed by pSNRtRNA-pBpaRS were grown in the presence or absence of 1 mM Bpa and analyzed by Western blot with an α-Flag antibody (c) LexA+VP16N L444Bpa and LexA+VP16C F475Bpa were assessed for their ability to upregulate transcription of an integrated LacZ reporter gene in *S. cerevisiae* as measured by liquid b-galactosidase assays. Each activity is the average of values from at least three independent experiments with the indicated error (SDOM). Dr. Malathy Krishnamurthy performed experiments with LexA+VP16C and Amanda Dugan performed experiments with LexA+VP16N.

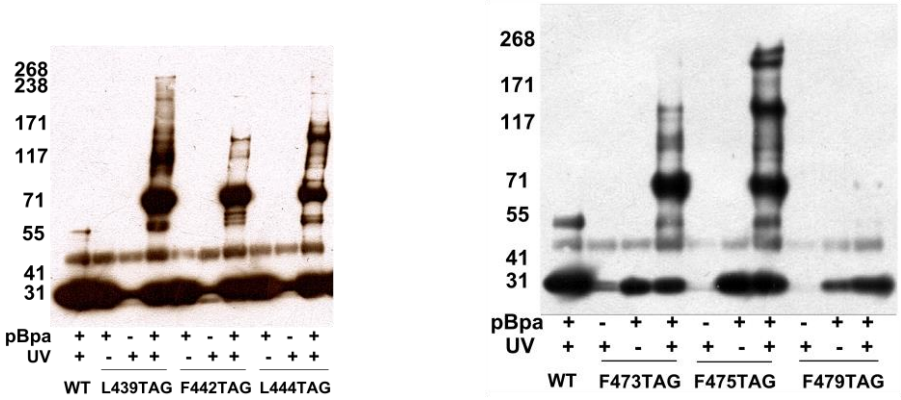

Figure 2-5 Photocrosslinking profiles of LexA+VP16N and LexA+VP16C Bpa containing mutants. Bpa was incorporated in LexA+VP16N (left) at positions 439, 442, and 444 and in LexA+VP16C (right) at positions 473, 475, and 479 in yeast. Yeast cells were irradiated with UV light for 30 minutes on ice to activate Bpa and covalently capture proteins directly binding to each subdomain of Bpa. Western blot is visualized with an α-FLAG antibody. Dr. Malathy Krishnamurthy carried out experiments with LexA+VP16C and Amanda Dugan carried out experiments with LexA+VP16N.

2.D.2. Capturing VP16-Med15 in live yeast: a model moderate affinity interaction

A strong body of evidence exists to support the Mediator subunit and coactivator Med15 as a target of VP16 and this model that is supported by *in vivo* deletion and mutagenesis experiments from our own lab.[18,70] The interaction between Med15 and activators such as VP16 is moderate in affinity (high nM to low μM) with dissociation constants ≥2 orders of magnitude weaker than the Gal4-Gal80 interaction (Figure 2-6 a).[17,18,70,71] Thus, this interaction appeared to be an excellent test case of the effectiveness of the *in vivo* crosslinking strategy for capturing moderate affinity binding interactions. We first tested the ability of each VP16 subdomain to crosslink to the coactivator Med15 *in vivo* by co-expressing myc-tagged Med15 alongside either LexA+VP16N L444Bpa or

LexA+VP16C F475Bpa and irradiating live yeast with 365 nm UV light. The covalent adducts were isolated from yeast lysate and analyzed by Western blot (Figure 2-6 b). A direct contact between each subdomain of VP16 and Med15 was observed and this was dependent upon irradiation, thus validating the utility of this method in capturing a moderate affinity in vivo interaction of a transcriptional activator.

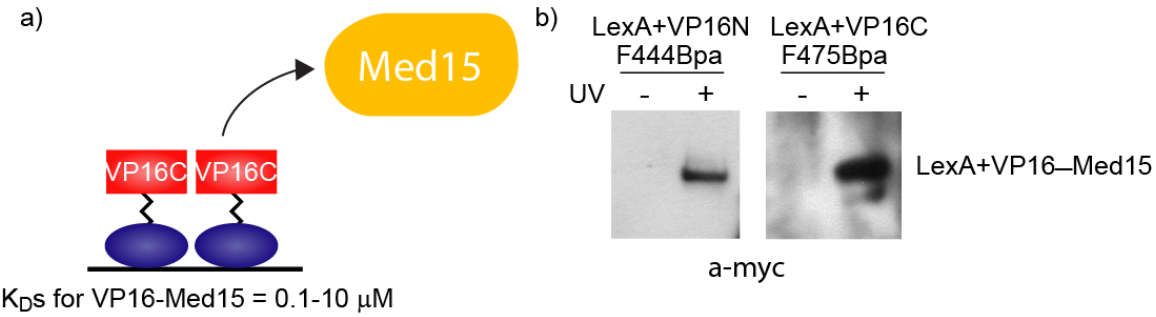

Figure 2-6 In vivo photocrosslinking captures the moderate affinity interaction between LexA+VP16 and the Mediator protein, Med15. a) VP16 has been shown to interact transiently with the coactivator Med15, with equilibrium binding measurements placing the affinity of the TAD for Med15 in the moderate category. DNA-bound homodimers exhibit the highest affinity (0.1 μM) for Med15 and isolated TADs display affinities for Med15 that falls in the low to mid-micromolar range (*16, 44*). (b) Live yeast cells bearing plasmids expressing LexA+VP16N L444Bpa or LexA+VP16C F475Bpa fusion proteins, in addition to a plasmid expressing myc-Med15(1-416) were irradiated with UV light (365 nm) for 30 minutes. Subsequently, cell lysates were immunoprecipitated with α-LexA and analyzed by Western blot (α-myc). For both constructs, a crosslink with Med15 is observed. Dr. Malathy Krishnamurthy performed experiments with LexA+VP16C and Amanda Dugan performed experiments with LexA+VP16N.

2.D.3. Capturing the direct interactions of VP16 within the Swi/Snf chromatin remodeling complex

As outlined earlier, the Swi/Snf chromatin-remodeling complex has also been proposed to be a direct binding partner of activators such as VP16 but there is conflicting evidence as to which subunit(s) serves as the activator-binding motif *in vivo*. In yeast, Swi/Snf is a twelve subunit multi-protein complex that, when recruited to promoters such as GAL1, enhances transcription by repositioning nucleosomes and increasing transcription factor accessibility at the promoter. As discussed in Chapter 1, the activity and localization of many transcriptional complexes including Swi/Snf are modulated by the presence of subunits that can associate either stably or transiently with the complex. For example, the human

homolog of the Swi2/Snf2 ATPase, Brg1/BRM, has been shown to function independently but its activity is enhanced significantly upon addition of three other subunits (INI1 (ySnf5), BAF155 (ySwi3), and BAF170 (ySwi3)). The direct effects of the exchange of these subunits can be seen during neuronal development, where modulation of the complex results in the differentiation of an embryonic stem cell into a fully developed neuron. In yeast, VP16 enhances Swi/Snf recruitment to promoters such as GAL1 used in our studies and structural studies of Swi/Snf in complex with the nucleosome suggest that the catalytic subunit Snf2 is positioned close to the activator.[25-30,72-74] However, *in vitro* binding studies have shown that several additional subunits can serve as activator binding partners (Swi1, Snf5).[19,31,32] We hypothesized that the *in vivo* crosslinking strategy could be used to test if the Swi/Snf complex is directy bound by VP16 in the cell and, if so, to identify the Swi/Snf subunits that are directly bound by VP16 in the native complex environment. We secondly wished to identify which subdomain was important for binding these subunits, thus revealing a more complete picture of recruitment by VP16 *in vivo*.

2.D.4. VP16 and the Snf2 ATPase

In the case of both the LexA+VP16N L444Bpa and the LexA+VP16C F475Bpa activators, irradiation of live yeast cells expressing the activators followed by visualization of all crosslinked products via immunodetection of the FLAG tag revealed several bands in the 130-220 kDa molecular weight range, consistent with the size range expected for covalent complexes with the Snf2, Swi1, and Snf5 subunits (Figure 2-7 c,d). To test this, immunoprecipitation of whole-cell extracts from irradiated cells with an antibody to Snf2 was carried out. In these experiments, no detectable LexA+VP16N-Snf2 product was observed, even when additional Bpa incorporation positions at L439 and F442 were tested (data not shown). However, as seen in Figure 2-7b, the LexA+VP16C F475Bpa mutant crosslinks directly to endogenous Snf2. Consistent with this result, point mutations (F479A and F479P) known to decrease VP16 coactivator binding in

vitro were introduced and, in line with these earlier biochemical experiments, abrogation of VP16 crosslinking to Snf2 in vivo is observed.[55,59,62] These results are consistent with the recent structural model proposed by Dechassa et al that places Snf2 proximal to the transcriptional activator in the context of a Swi/Snf-nucleosome-activator complex.[30] Further, the data suggests that it is the C-terminus of the VP16 TAD that is responsible for the bulk of the Snf2 recruitment activity.

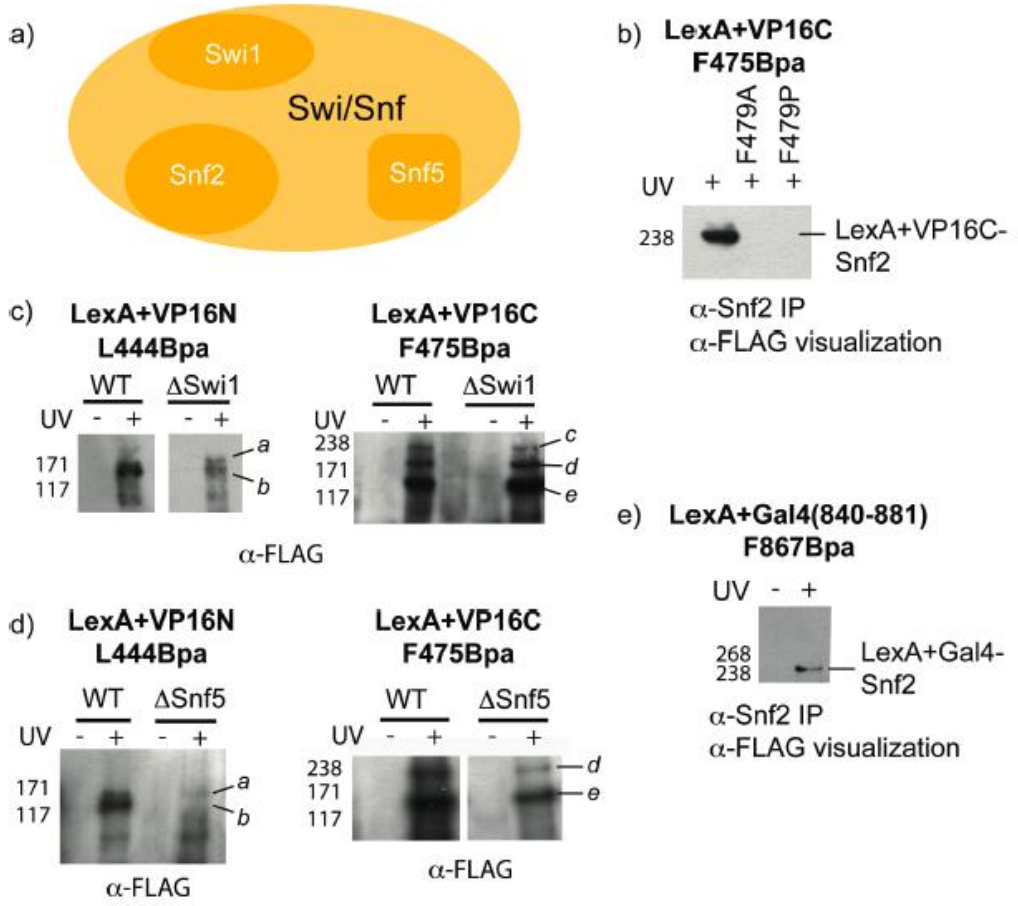

Figure 2-7 Analysis of VP16 crosslinking to the Swi/Snf coactivators, Snf2, Swi1 and Snf5. (a) The recruitment of the Swi/Snf chromatin remodeling complex by VP16 has been proposed to occur through interactions with the Snf2, Swi1 and Snf5 subunits although the direct binding partners *in vivo* have not been determined (*17, 33, 34*). (b) Live yeast cells expressing LexA+VP16C F475Bpa were irradiated with 365 nm light (30 minutes) and subsequently the cell lysates were immunoprecipitated with an antibody to Snf2 and resolved by Western blot (α-FLAG), revealing a direct interaction between VP16C and endogenous Snf2. In line with previous biochemical experiments, when phenylalanine 479 in VP16C was mutated to either alanine or proline, crosslinking to Snf2 was abolished. (c,d) LexA+VP16C F475Bpa and LexA+VP16N

L444Bpa were expressed in yeast strains lacking either Swi1 or Snf5 and the live yeast cells were irradiated with 365 nm light. Subsequently, cell lysates were immunoprecipitated (α-LexA) and resolved by Western blot (α–FLAG). In the individual blots for LexA+VP16N, the marks a and b denote crosslinked protein bands at the appropriate size for Swi1 and Snf5, respectively. In the individual blots for LexA+VP16C, the marks c, d, and e indicate bands at the appropriate size for Snf2, Swi1 and Snf5, respectively. (e) To test if Gal4 also contacts Snf2, crosslinking experiments were carried out with live yeast cells expressing LexA+Gal4 F867Bpa as in (b). Dr. Malathy Krishnamurthy created the Swi/Snf deletion strains and performed crosslinking with LexA+VP16C, Amanda Dugan performed experiments with LexA+VP16N, Hugo Fung carried out experiments with LexA+Gcn4, and Adaora Nwokoye performed experiments with LexA+Gal4.

2.D.5. VP16 and Snf5, Swi1 and Snf6

In contrast to the Snf2 immunoprecipitation experiments, enrichment with either a Swi1 or Snf5 antibody did not result in any detectable crosslinked product (data not shown). We realize that this could be one effect of crosslinking experiments in that, depending on the location of the crosslinked bond formed, epitope recognition by the antibody could be obscured in a similar fashion to what is seen in formaldehyde crosslinking experiments. To probe these interactions further, we generated yeast strains lacking either Swi1 or Snf5 and carried out crosslinking experiments. No differences in crosslinked product formation between the WT strain and the Swi1 delete strain were observed with either VP16-derived activator, suggesting that Swi1 is not a direct target of VP16 (Figure 2-7c). In contrast, deletion of Snf5 disrupts the normal binding pattern of LexA+VP16N L444Bpa, consistent with Snf5 interacting with VP16N (Figure 2-7d). However, upon deletion of Snf5, LexA+VP16C F475Bpa displays no change in binding pattern, implying that the VP16C TAD does not interact with Snf5.

In vitro data does not strongly support an interaction between Snf6 and VP16 but the structural model of a VP16-Swi/Snf complex from EM studies suggests that the Snf6 subunit is the closest in proximity to the activator.[30,32] As such, we decided to test if this subunit was involved in Swi/Snf recruitment. In the absence of a commercially available antibody, we created a myc-Snf6 construct and co-expressed this alongside our VP16 Bpa mutants. Following irradiation of cells with UV light, we lysed the yeast and immunoprecipitated the covalent activator adducts with a LexA antibody and probed the subsequent Western blot

with a myc-antibody to detect covalently bound Snf6. To our surprise, a clean band corresponding to the molecular weight of a VP16N-Snf6 complex was detected, indicating that VP16N does in fact directly contact Snf6 in cells, despite earlier *in vitro* studies testing and missing the capture of this interaction (Figure 2-8). Together with the results of Figure 2-7b, these data support a model in which the subdomains of VP16 work cooperatively to recruit the Swi/Snf complex, with VP16C directly contacting Snf2 and VP16N depending on Snf6 and Snf5 during transcription initiation.

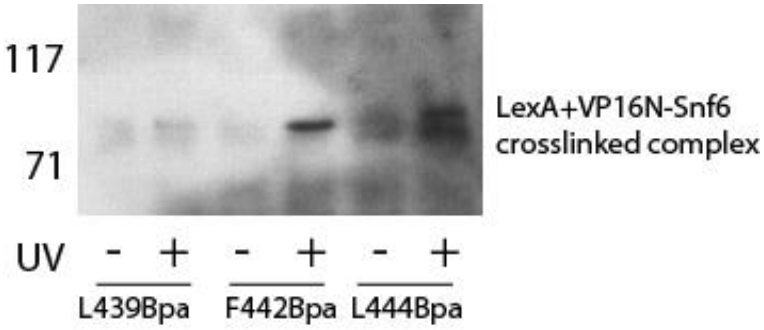

Figure 2-8 *In vivo* photocrosslinking captures the Swi/Snf subunit Snf6, an unlikely target according to *in vitro* studies. LexA+VP16N L444Bpa and LexA+VP16C F475Bpa were co-transformed alongside myc-Snf6 in yeast and subjected to in vivo photocrosslinking experiments as described earlier. Immunoprecipitation of resulting cell lysates with a LexA antibody followed by Western blot detection with an α-myc antibody reveals a direct interaction between each subdomain of VP16 and Snf6. LexA+VP16C was not tested for crosslinking to Snf6. The appearance of two bands in these experiments has been observed with other myc-tagged proteins including Med15. Given the importance of ubiquitin mediated degradation in regulation activators, this lower band could be a partially degraded crosslinked product. However, reports from other labs suggest that proteins that are heavily phosphorylated often display several bands on a blot. Thus, the two bands shown could indicate a crosslinked VP16N-Snf6 product, one that is more phosphorylated than the other.

2.E. Examining shared targets of amphipathic activators in the Swi/Snf complex

Snf2 is an ATPase that is essential for Swi/Snf function and is highly conserved among eukaryotes, making it a likely shared target among other transcriptional activators.[75] In fact, in addition to VP16, the amphipathic activators Gal4 and Gcn4 have been shown to recruit Swi/Snf to a variety of promoters *in vivo* and *in vitro*, suggesting that these activators contact a conserved set of targets within this complex.[19,29,31,76-80] To determine if Snf2 is a shared target of these activators, Gal4 and Gcn4 were modified to contain Bpa within regions of each

TAD implicated in coactivator binding and then tested for their ability to crosslink to Snf2. As shown in Figure 2-7e, Gal4 makes a direct contact with Snf2, whereas Gcn4 does not for any position tested (Figure 2-9). These data suggest that Snf2 (Brg1/Brm in metazoans) could be a key target for small molecule probe development in order to characterize the role of the conserved Swi/Snf complexes that are associated with pathophysiological processes.[81-84] However, further studies will be needed to dissect if VP16 and Gal4, as well as other activators, interact with the same binding site within Snf2. Together, this data demonstrates the powerful advantage that *in vivo* photocrosslinking methods provide in capturing the direct cellular binding partners of transcriptional activators, as demonstrated by the identification of Swi/Snf targets that were missed by *in vitro* studies.

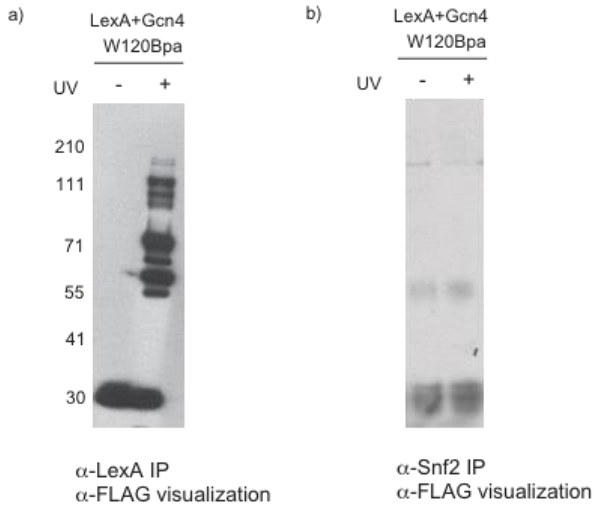

Figure 2-9 LexA-Gcn4 does not appear to crosslink to Snf2 in vivo. Live yeast cells expressing LexA+Gcn4 W120Bpa were irradiated with 365 nm light for 30 minutes. (a) Cell lysates were immunoprecipitated with an antibody to LexA and resolved by Western blot (α-FLAG), indicating multiple crosslinked products. (b) Cell lysates were immunoprecipitated with an antibody to Snf2 and resolved by Western blot (α-FLAG). Similar results were obtained with LexA+Gcn4 with Bpa incorporated at positions F108, Y110, D115, K118, T121, L123, F124, N126, T132 or K140.

2.F. Considerations when interpreting a negative crosslinking result

Through these studies, we have found that it is difficult to interpret a negative result from *in vivo* photocrosslinking studies. A negative result, as in the case of

the Swi1 and others, could arise from a variety of factors including lack of a binding interaction, low unnatural amino acid (UAA) incorporation yield and/or fidelity, poor positioning of the UAA, and the poor reactivity of the activated UAA with the amino acids in the binding partner. As a case in point, the first example of in cell photocrosslinking experiments from our group led to the *in situ* characterization of the binding interface between Gal4 and its masking protein Gal80. In these experiments, Majmudar et al incorporated Bpa at positions spanning the Gal4 TAD and then carried out crosslinking experiments with these mutants to capture the Gal4-Gal80 interaction in yeast. From the results, they noticed that for position 856, significant functional and structural data exists to support the involvement of this residue in a direct Gal4-Gal80 interaction yet little to no crosslinking was observed.[17] However, the LexA+Gal4 F856Bpa mutant remained repressed in the presence of Gal80, indicating that incorporation of Bpa at this position did not negatively impact the Gal80 binding interaction. Thus we launched an examination of the remaining facets of the *in vivo* crosslinking experiment and, importantly demonstrate the substantial role that the crosslinking mechanism and sequence context play in the ability to capture a PPI. This case study of a PPI provides an additional framework for designing successful in vivo crosslinking experiments.

2.F.1. The Gal4-Gal80 interaction

In response to carbon source availability, the well-characterized yeast activator Gal4 regulates genes responsible for galactose catabolism and its function is highly regulated by its masking protein Gal80.[85-88] In the presence of glucose, Gal80 binds Gal4 with low nanomolar affinity and prevents Gal4 from recruiting the necessary transcriptional complexes to upregulate gene expression.[71] Conversely, in the absence of glucose and presence of galactose, inhibition of Gal4 by Gal80 is lifted, allowing transcription to occur. Functional and structural data have mapped the residues in Gal4 involved in contacting Gal80 and thus which residues were most likely to yield crosslinks in our photocrosslinking

studies.[71,88-91] As in earlier experiments, a heterologous construct, LexA+Gal4-flag was used in these studies.

As described earlier, *in vivo* photocrosslinking with LexA+Gal4 F856Bpa yielded no complex between Gal4 and Gal80 despite several lines of evidence supporting an interaction. Bpa incorporation at this position was evaluated and subsequent Western blot analysis and functional data indicated that LexA+Gal4 F856Bpa was not only being expressed, but that it was also fully functional and sensitive to Gal80 inhibition under glucose conditions.

2.F.2. Differences in Bpa and Azpa crosslinking mechanism and reactivity

On explanation for the failure to capture Gal80 is intramolecular quenching of Bpa. Upon irradiation with 350-365 nm UV light, Bpa forms a diradical that proceeds to form a covalent bond with nearby protein backbones and amino acid side chains through C-H insertion chemistry.[13] Although Bpa can react with most C-H bonds, it has experimentally been shown to react preferentially with methionine (Met) where it reacts at distances beyond the 3.1 A reactive radius.[92] Specifically, the apparent preference of Bpa for methionine suggests that Bpa efficiency can be altered dramatically when placed in close proximity to methionine's thioether side chain.[92-94] Further analysis of position 856 in the Gal4 TAD reveals two methionines in close proximity to the Bpa side chain, thus we hypothesized that the methionine residues at positions 855 and 861 are internally "quenching" Bpa, thereby preventing it from crosslinking to Gal80. Consistent with this hypothesis, when Met855 and Met861 are mutated either individually or collectively to alanine, we see that the resulting mutants are functional and, importantly, that crosslinking of Gal4 to Gal80 is restored (Figure 2-11). These data are consistent with a model in which an intramolecular crosslink was competing with the intermolecular reaction in the LexA+Gal4 F856Bpa mutant and led to a false negative in the original crosslinking experiments.

Thus a different photoactivatable amino acid, Azpa, was used as this can be incorporated into proteins in yeast using the optimized nonsense suppression system. The amino acid preference for Azpa crosslinking is less clear because it has a more complex crosslinking mechanism compared to Bpa.[13] During excitation at ~254 nm of light it forms a nitrene that can then insert into C-H or heteroatom-H bonds. However, if the nitrene does not react within the ~10^{-4} s excitation, it will rearrange into a more stable ketenimine which can then react with nucleophiles, including surrounding solvent.[13,95] Although the differences in reactive mechanism between Bpa and Azpa are known, a direct comparison of the effect of these reactivities on the experimental outcome of crosslinking studies has yet to be established.

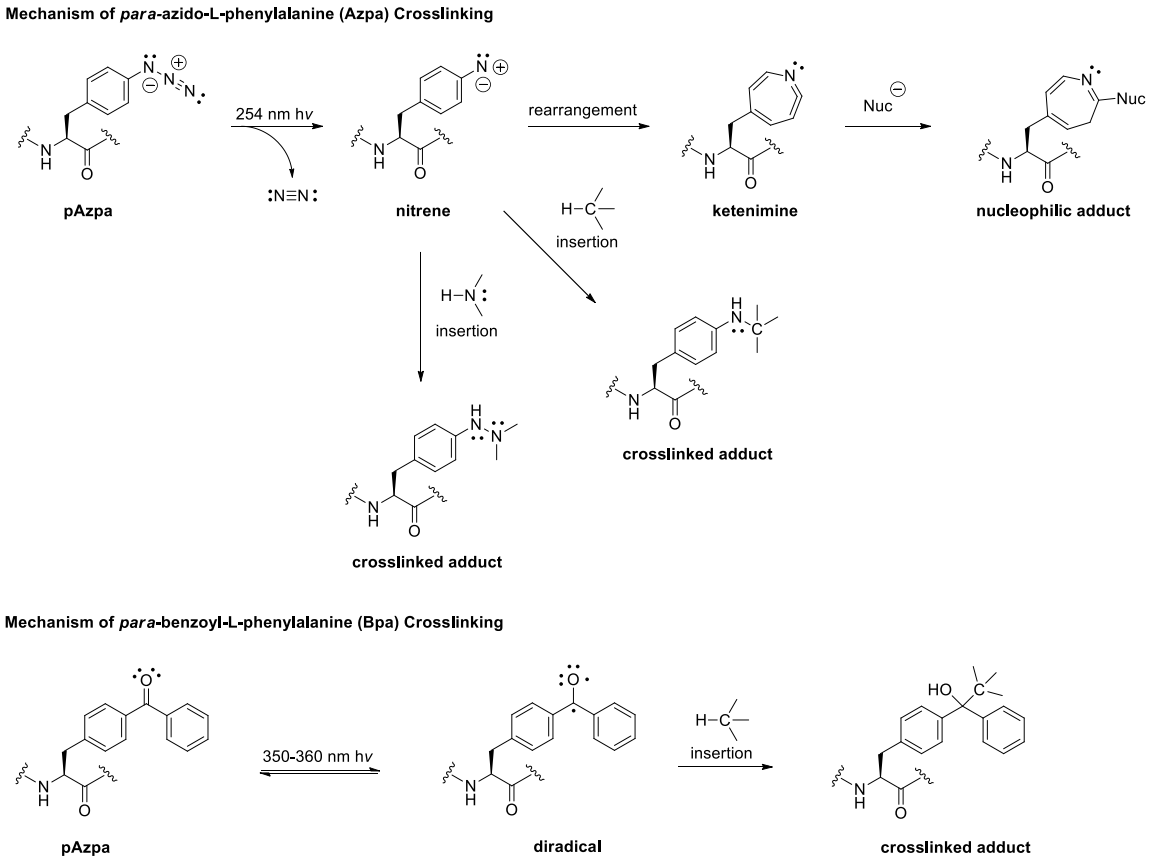

Figure 2-10 (top) Azpa crosslinking mechanism. Upon irradiation with ~254 nm UV light, Azpa forms a nitrene that can form covalent crosslinking adducts with nearby C-H and N-H bonds. In the absence of a crosslinking partner, the nitrene relaxes into a ketenimine that then reacts with nucleophiles such as primary amines to form a stable end product. Once activated, Azpa cannot relax back to its ground state and be reactivated for crosslinking investigations. (b). Bpa crosslinking mechanism. Upon irradiation with 350-365

nm UV light, Bpa forms a diradical that can relax back down to its ground state (reversible excitation). However, in the presence of a binding partner, the diradical can also form a covalent bond through a C-H insertion reaction.

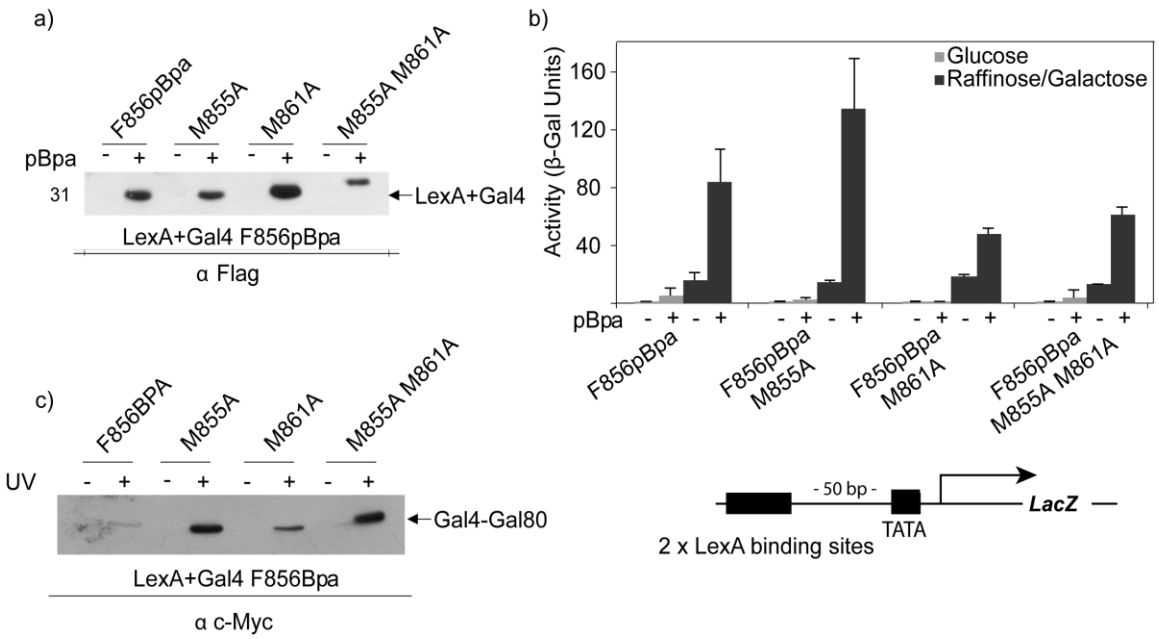

Figure 2-11 Evaluating intramolecular quenching of Bpa in LexA+Gal4 F856Bpa by neighboring methionines at residues 855 and 861. (a). M855 and M861 in the Gal4 TAD were individually and collectively mutated to alanine to remove the presence of the thioether side chain that has been shown to have a "magnet" effect on Bpa crosslinking. Incorporation of Bpa in LexA+Gal4 F856Bpa and the methionine mutants was tested as described earlier, showing the production of full-length protein for each construct. (b). All mutants were tested for functional output as measured by liquid β-galactosidase assays. Activator activity was measured as a function of expression of an integrated LacZ reporter gene under the control of a GAL1 promoter bearing two LexA binding sites. All constructs maintained sensitivity to glucose repression and galactose induction. (c). Mutants were assayed for ability to crosslink to myc-Gal80 in live yeast. As discussed in the text, LexA+Gal4 F856Bpa is unable to form crosslinks with myc-Gal80 despite functional and structural evidence implicating this residue in the interaction face. Upon mutation of neighboring methionines to alanine, crosslinking to Gal80 is restored, indicating a probable intramolecular quenching by positions 855 and 861 in the Gal4 TAD. Experiments performed alongside Rachel Pricer and Cassandra Joiner in the Mapp Lab.

2.F.2.a. Impact on UAA reactivity on the outcome of crosslinking studies

Here we incorporate Azpa into the Gal4 TAD using the same optimized system developed by our lab for Bpa incorporation. As before, Azpa incorporation is driven by an Azpa specific tRNA/RS whose expression is controlled by the pSNR52 promoter. Utilization of the expression conditions outlined earlier leads to the incorporation of Azpa at position 856, with the resulting mutant proving to

be fully expressed and functional (Figure 2-12a,b). We next performed a direct comparison of Azpa and Bpa at postion 856, revealing that LexA+Gal4 F856Azpa readily crosslinks to Gal80 whereas LexA+Gal4 F856Bpa does not. As expected, switching methionine for alanine at positions 855 and 861 yielded no changes in Azpa crosslinking, consistent with the reactivity profile of this amino acid (Figure 2-12c). Thus, we find that to avoid false negatives, one should carry out crosslinking experiments with more than one UAA mutant since a small change in position can have a dramatic effect on crosslinking outcome. The results shown with Gal4-Gal80 illustrate that careful consideration of the innate reactivity of the UAA utilized is a key factor in the successful application of the *in vivo* photocrosslinking strategy.

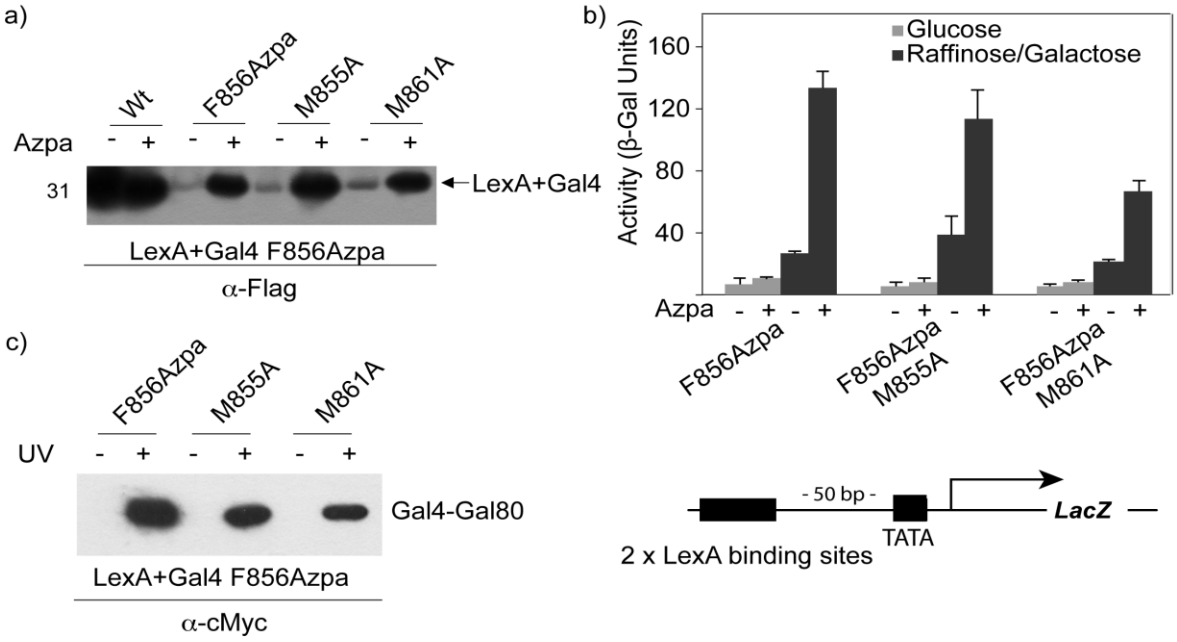

Figure 2-12 Azpa crosslinks to myc-Gal80 regardless of the presence of neighboring methionines. (a). M855 and M861 in the Gal4 TAD were individually and collectively mutated to alanine to remove the presence of the thioether side chain that has been shown to have a "magnet" effect on Bpa crosslinking. Incorporation of Azpa in LexA+Gal4 F856Azpa and the methionine mutants was tested as described earlier, showing the production of full-length protein for each construct. (b). All mutants were tested for functional output as measured by liquid β-galactosidase assays. Activator activity was measured as a function of expression of an integrated LacZ reporter gene under the control of a GAL1 promoter bearing two LexA binding sites. All constructs maintained sensitivity to glucose repression and galactose induction. (c). Mutants were assayed for ability to crosslink to myc-Gal80 in live yeast. In contrast to studies with Bpa, LexA+Gal4 F856Azpa readily forms crosslinks with myc-Gal80. Mutation of neighboring methionines to alanine does not appear to effect LexA+Gal4 F856Azpa crosslinking to Gal80, consistent with the crosslinking mechanism of Azpa. Experiments performed alongside Rachel Pricer and Cassandra Joiner in the Mapp Lab.

Taken together, these data demonstrate that the genetically encoded photocrosslinkers Bpa and Azpa are a viable and perhaps indispensible tool for capturing moderate affinity and transient protein-protein interactions in cells. Employing the *in vivo* photocrosslinking strategy revealed an interaction model for the cooperative recruitment of the chromatin-remodeling Swi/Snf complex to gene promoters and further identified the Snf2 coactivator and ATPase to be a direct target of the prototypical activators VP16 and Gal4. These data represent a significant step toward the development of a complete interaction map of the direct binding partners of transcriptional activators, long an elusive goal. Furthermore, applying the considerations presented in this chapter will facilitate the successful implementation of the *in vivo* crosslinking methodology for this class of moderate affinity, transient interactions and further sets the stage for the dissection of the complex interactions of the many other cellular mechanisms that function through similar PPI networks.

2.G. Experimental

Yeast Strain LS41 [JPY9::pZZ41, *Mata his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41] was used for the crosslinking experiments. Swi1 and Snf5 deletion strains were made by gene disruption via PCR in LS41 and used for deletion 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 of Plasmids used in this Chapter

Plasmid name	**Function**

pLexAVP16N	Expresses LexA(1-202)+VP16 (413-456)+FLAG tag
pLexAVP16C	Expresses LexA(1-202)+VP16 (446-490)+FLAG tag
pLexAVP16N 439TAG, pLexAVP16N 442TAG, pLexAVP16N 444TAG,	Expresses LexA(1-202)+VP16 (446-490)+FLAG tag with a TAG replacing the codon of the existing amino acid
pLexAVP16C 473TAG, pLexAVP16C 475TAG, pLexAVP16C 479TAG	Expresses LexA(1-202)+VP16 (413-456)+FLAG tag with a TAG replacing the codon of the existing amino acid
pLexAGal4 849TAG, pLexAGal4 867TAG, pLexAGal4 869TAG	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid
pLexAGcn4 108TAG, pLexAGcn4 110TAG, pLexAGcn4 115TAG, pLexAGcn4 118TAG, pLexAGcn4 120TAG, pLexAGcn4 121TAG, pLexAGcn4 123TAG, pLexAGcn4 124TAG, pLexAGcn4 126TAG, pLexAGcn4 132TAG,	Expresses LexA(1-202)+Gcn4(107-144)+FLAG tag with a TAG codon replacing the codon of the existing amino acid

pLexAGcn4 140TAG	
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
pLexAGal4	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag
pLexAGal4 849TAG, pLexAGal4 856TAG	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid
pSNRtRNA-pAzpaRS	Expresses tRNA under the control of the SNR52 promoter and contains synthetase specific for pAzpa
c-Myc-Gal80	Expresses full-length Gal80+c-Myc tag

Construction of plasmids

1. pLexAVP16 N and pLexAVP16C

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 treated with calf intestinal phosphate 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.

2. pLexAVP16N 439TAG, pLexAVP16N 442TAG, pLexAVP16N 444TAG, pLexAVP16C 473TAG, pLexAVP16C 475TAG, pLexAVP16C 479TAG

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.

3. pLexAGcn4(107-144)

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.

4. pLexAGal4 (840-881)

pLexA(1-202)+Gal4(840-881) was created as previously described (Majmudar, CY et al, 2009)

5. 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(1-415) from yeast genomic DNA using primers (5'-GACAGGATCCATGTCT GCTGCTCCTGTCCAAGAC-3') and (5'- CGATCATATGTCACTGATATAATTTAGAACTTGC-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'-AGCTATGGAACAAAAGTTGATTTCTGAAGAAGATTTGGGATCCAATGCATATGATCT-3') and (5'-AGCTTGATCATATGCATTGGATCCCAAATCTTCTTCAGAAATCAACTTTTGTTCCAT-3').

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**CATCAAGGGAACAAAAG
Snf5-Rev-2	Round 2 PCR	**GATAATACAAATTCTTCCAC**GGTTATTTACATCTCCGGTA
Swi1-Fwd	Round 1 PCR	ATGGATTTCTTTAATTTGAATAATAATAATAATAATAATAATAC *CGGATCCCCGGGTTAATTAA*

Swi1-Rev	Round 1 PCR	TCATTCCAAATTGGTTAGGATATCATTTTTTAAATTGTAAAG *GAATTCGAGCTCGTTTAAAC*

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, *Mata his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41]. Plasmid PFa6-TRP1 (generously donated by Karbstein group, University of Michigan) was used as a template to clone out deletion inserts using target-gene-specific primer pairs as designated in Table 2.

Table 2: Primers used for PCR based gene deletion

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 the colony using 20 mM NaOH. 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(1-202)+VP16N and LexA(1-202)+VP16C and expression of myc-Med15 in S. cerevisiae pellet

LS41 yeast was transformed with various pLexAVP16 TAG mutant plasmids and pSNRtRNA-pBpaRS plasmid. Individual colonies were grown to saturation in 5 mL SC media lacking histidine and tryptophan for selection and 2% Raffinose at 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 the cell pellets were 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). To test expression of the myc-Med15(1-416) construct, the same protocol was followed except that LS41 were additionally transformed with the pMyc-Med15(1-416) plasmid and grown in SC media lacking histidine, tryptophan and leucine. Lysates were analyzed using Western blot with anti-myc antibody (Santa Cruz Biotech, sc-40).

In vivo cross-linking

To perform in vivo cross-linking, individual colonies of each pLexAVP16 TAG mutant were grown in 5 mL SC media containing 2% Raffinose 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 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 OD660 of ~1.0. When cultures reached the appropriate OD_{660}, the cells were spun down by centrifuging at 3901 rcf at 4^{o}C for 5 min. following which the cell pellets were washed with SC media lacking histidine and tryptophan. The cell pellets were resuspended in 2mL SC media lacking histidine and tryptophan + 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 mycMed15(1-416), myc-Snf6, myc-Gal80 and the deletion strains, the procedure was identical except that cells were grown in SC media lacking histidine, leucine, and tryptophan and, due to poor growth in raffinose and galactose, deletion cultures were grown in 2% glucose. 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 ^{o}C. After immunoprecipitation, the beads were washed 6 times 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 ^{o}C until elution. The crosslinked sample was eluted from the beads by heating at 95 ^{o}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 HRP, Santa Cruz Biotechnology, Santa Cruz, CA).

β-*Galactosidase assays*

To evaluate the ability of each LexA+VP16 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 + 2% Galactose but lacking histidine and tryptophan for selection. The cells were grown to an OD of 1.5-2.0 and harvested. The activity of each construct was monitored using β-galactosidase assays as previously described.

2.H. References-Chapter 2

(1) Krishnamurthy, M.; Dugan, A.; Nwokoye, A.; Fung, Y.-H.; Lancia, J. K.; Majmudar, C. Y.; Mapp, A. K. *ACS Chem Biol* **2011**, *6*, 1321.
(2) Lancia, J. K.; Nwokoye, A.; Dugan, A.; Pricer, R.; Joiner, C.; Mapp, A. K. *Biopolymers* **2013**, *manuscript accepted*.
(3) Lee, L. W.; Mapp, A. K. *J Biol Chem* **2010**, *285*, 11033.
(4) Koehler, A. N. *Curr Opin Chem Biol* **2010**, *14*, 331.
(5) Arkin, M. R.; Whitty, A. *Curr Opin Chem Biol* **2009**, *13*, 284.
(6) Evans, C. G.; Chang, L.; Gestwicki, J. E. *J Med Chem* **2010**, *53*, 4585.
(7) Ong, D. S.; Kelly, J. W. *Curr Opin Cell Biol* **2011**, *23*, 231.
(8) Berg, T. *Curr Opin Chem Biol* **2008**, *12*, 464.
(9) Arndt, H. D. *Angew Chem Int Ed Engl* **2006**, *45*, 4552.
(10) Blazer, L. L., Neubig, R.R. *Neuropsychopharmacology* **2009**, *34*, 126.
(11) Mapp, A. K.; Ansari, A. Z. *ACS Chem Biol* **2007**, *2*, 62.
(12) Berggård, T.; Linse, S.; James, P. *Proteomics* **2007**, *7*, 2833.
(13) Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol Biosyst* **2008**, *4*, 473.
(14) Melcher, K. *Curr Protein Pept Sci* **2004**, *5*, 287.
(15) Perkins, J. R.; Diboun, I.; Dessailly, B. H.; Lees, J. G.; Orengo, C. *Structure* **2010**, *18*, 1233.
(16) Fuxreiter, M.; Tompa, P.; Simon, I. *Bioinformatics* **2007**, *23*, 950.
(17) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J Am Chem Soc* **2009**, *131*, 14240.
(18) Wands, A. M.; Wang, N.; Lum, J. K.; Hsieh, J.; Fierke, C. A.; Mapp, A. K. *J Biol Chem* **2010**, *286*, 16238.

(19) Ferreira, M. E.; Hermann, S.; Prochasson, P.; Workman, J. L.; Berndt, K. D.; Wright, A. P. H. *J Biol Chem* **2005**, *280*, 21779.
(20) Herbig, E.; Warfield, L.; Fish, L.; Fishburn, J.; Knutson, B. A.; Moorefield, B.; Pacheco, D.; Hahn, S. *Mol Cell Biol* **2010**, *30*, 2376.
(21) Melcher, K. *J Mol Biol* **2000**, *301*, 1097.
(22) Bryant, G. O.; Ptashne, M. *In Vivo* **2003**, *11*, 1301.
(23) Cosma, M. P.; Tanaka, T.; Nasmyth, K. *Cell* **1999**, *97*, 299.
(24) Hall, D. B.; Struhl, K. *J Biol Chem* **2002**, *277*, 46043.
(25) Memedula, S.; Belmont, A. S. *Curr Biol* **2003**, *13*, 241.
(26) Yudkovsky, N.; Logie, C.; Hahn, S.; Peterson, C. L. *Genes Dev* **1999**, *13*, 2369.
(27) Herrera, F. J.; Triezenberg, S. J. *J Virol* **2004**, *78*, 9689.
(28) Hassan, A. H.; Neely, K. E.; Workman, J. L. *Cell* **2001**, *104*, 817.
(29) Neely, K. E.; Hassan, A. H.; Wallberg, A. E.; Steger, D. J.; Cairns, B. R.; Wright, A. P.; Workman, J. L. *Mol Cell* **1999**, *4*, 649.
(30) Dechassa, M. L.; Zhang, B.; Horowitz-Scherer, R.; Persinger, J.; Woodcock, C. L.; Peterson, C. L.; Bartholomew, B. *Mol Cell Biol* **2008**, *28*, 6010.
(31) Prochasson, P.; Neely, K. E.; Hassan, A. H.; Li, B.; Workman, J. L. *Mol Cell* **2003**, *12*, 983.
(32) Neely, K. E.; Hassan, A. H.; Brown, C. E.; Howe, L.; Workman, J. L. *Mol Cell Biol* **2002**, *22*, 1615.
(33) Vashee, S.; Kodadek, T. *Proc Nat Acad Sci, USA* **1995**, *92*, 10683.
(34) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964.
(35) Xie, J.; Schultz, P. G. *Methods* **2005**, *36*, 227.
(36) Chin, J. W.; Schultz, P. G. *Chembiochem* **2002**, *3*, 1135.
(37) Mohibullah, N.; Hahn, S. *Genes Dev* **2008**, *22*, 2994.
(38) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A. *Nat Methods* **2005**, *2*, 3.
(39) Mossman, K. L.; Sherburne, R.; Lavery, C.; Duncan, J.; Smiley, J. R. *J Virol* **2000**, *74*, 6287.
(40) Sadowski, I.; Ma, J.; Triezenberg, S. J.; Ptashne, M. *Nature* **1988**, *335*, 563.
(41) Lai, J. S.; Herr, W. *Mol Cell Biol* **1997**, *17*, 3937.
(42) Triezenberg, S. J.; Kingsbury, R. C.; McKnight, S. L. *Genes Dev* **1988**, *2*, 718.
(43) Regier, J. L., Shen, F., Triezenberg, S.J. *Proc Natl Acad Sci USA* **1993**, *90*, 883.
(44) Ansari, a.; Mapp, a.; Nguyen, D.; Dervan, P.; Ptashne, M. *Chemistry & Biology* **2001**, *8*, 583.
(45) Chen, J.; Peterson, K. R.; Iancu-Rubin, C.; Bieker, J. J. *Proc Nat Acad Sci, USA* **2010**, *107*, 16846.
(46) Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; P.B., D. *J Am Chem Soc* **2002**, *124*, 13067.
(47) Seipel, K.; Georgiev, O.; Schaffner, W. *Biol Chem* **1994**, *375*, 463.

(48) Cress, W. D.; Triezenberg, S. J. *Science* **1991**, *251*, 87.
(49) Kobayashi, N.; Boyer, T. G.; Berk, a. J. *Mol Cell Biol* **1995**, *15*, 6465.
(50) Xiao, H.; Pearson, A.; Coulombe, B.; Truant, R.; Zhang, S.; Regier, J. L.; Triezenberg, S. J.; Reinberg, D.; Flores, O.; Ingles, C. J.; Greenblatt, J. *Mol Cell Biol* **1994**, *14*, 7013.
(51) Lin, Y. S.; Ha, I.; Maldonado, E.; Reinberg, D.; Green, M. R. *Nature* **1991**, *353*, 569.
(52) Kobayashi, N.; Horn, P. J.; Sullivan, S. M.; Triezenberg, S. J.; Boyer, T. G.; Berk, a. J. *Mol Cell Biol* **1998**, *18*, 4023.
(53) Stringer, K. F.; Ingles, C. J.; Greenblatt, J. *Nature* **1990**, *345*, 783.
(54) Hayashi, F.; Ishima, R.; Liu, D.; Tong, K. I.; Kim, S.; Reinberg, D.; Bagby, S.; Ikura, M. *Biochemistry* **1998**, *37*, 7941.
(55) Jonker, H. R.; Wechselberger, R. W.; Boelens, R.; Folkers, G. E.; Kaptein, R. *Biochemistry* **2005**, *44*, 827.
(56) Klein, J.; Nolden, M.; Sanders, S. L.; Kirchner, J.; Weil, P. A.; Melcher, K. *J Biol Chem* **2003**, *278*, 6779.
(57) Shooltz, D. D.; Alberts, G. L.; Triezenberg, S. J. *Protein Expr Purif* **2008**, *59*, 297.
(58) Nedialkov, Y. A.; Triezenberg, S. J. *Arch Biochem Biophys* **2004**, *425*, 77.
(59) Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L. *Science* **1997**, *277*, 1310.
(60) Hori, R. T.; Xu, S.; Hu, X.; Pyo, S. *Nuc Acids Res* **2004**, *32*, 3856.
(61) Mittler, G.; Stuhler, T.; Santolin, L.; Uhlmann, T.; Kremmer, E.; Lottspeich, F.; Berti, L.; Meisterernst, M. *EMBO J* **2003**, *22*, 6494.
(62) Yang, F.; DeBeaumont, R.; Zhou, S.; Naar, A. M. *Proc Natl Acad Sci U S A* **2004**, *101*, 2339.
(63) Uhlmann, T.; Boeing, S.; Lehmbacher, M.; Meisterernst, M. *J Biol Chem* **2007**, *282*, 2163.
(64) Utley, R. T.; Ikeda, K.; Grant, P. A.; Cote, J.; Steger, D. J.; Eberharter, A.; John, S.; Workman, J. L. *Nature* **1998**, *394*, 498.
(65) Silverman, N.; Agapite, J.; Guarente, L. *Proc Nat Acad Sci, USA* **1994**, *91*, 11665.
(66) Ikeda, K.; Stuehler, T.; Meisterernst, M. *Genes cells* **2002**, *7*, 49.
(67) Langlois, C.; Mas, C.; Di Lello, P.; Jenkins, L. M. M.; Legault, P.; Omichinski, J. G. *J Am Chem Soc* **2008**, *130*, 10596.
(68) Sullivan, S. M.; Horn, P. J.; Olsen, V. A.; Triezenberg, S. J. *Nucl Acids Res* **1998**, *26*, 4487.
(69) Shen, F.; Triezenberg, S. J.; Hensley, P.; Porter, D.; Knutson, J. R. *J Biol Chem* **1996**, *271*, 4827.
(70) Majmudar, C. Y.; Wang, B.; Lum, J. K.; Hakansson, K.; Mapp, A. K. *Angew Chem Int Ed Engl* **2009**, *48*, 7021.
(71) Thoden, J. B.; Ryan, L. A.; Reece, R. J.; Holden, H. M. *J Biol Chem* **2008**, *283*, 30266.
(72) Côté, J.; Quinn, J.; Workman, J. L.; Peterson, C. L. *Science* **1994**, *265*, 53.

(73) Lemieux, K.; Gaudreau, L. *The EMBO journal* **2004**, *23*, 4040.
(74) Peterson, C. L.; Workman, J. L. *Curr Opin Genet Dev* **2000**, *10*, 187.
(75) Laurent, B. C.; Treich, I.; Carlson, M. *Genes Dev* **1993**, *7*, 583.
(76) Ferreira, M. E.; Prochasson, P.; Berndt, K. D.; Workman, J. L.; Wright, A. P. H. *FEBS J* **2009**, *276*, 2557.
(77) Schwabish, M. A.; Struhl, K. *Mol Cell Biol* **2007**, *27*, 6987.
(78) Bryant, G. O.; Prabhu, V.; Floer, M.; Wang, X.; Spagna, D.; Schreiber, D.; Ptashne, M. *PLoS Biology* **2008**, *6*.
(79) Govind, C. K.; Yoon, S.; Qiu, H.; Govind, S.; Hinnebusch, A. G. *Mol Cell Biol* **2005**, *25*, 5626.
(80) Swanson, M. J., Qiu, H., Sumibcay, L., Krueger, A., Kim, S., Natarajan, K., Yoon, S., and Hinnebusch, A.G. *Mol Cell Biol* **2003**, *23*, 2800.
(81) Wilson, B. G.; Roberts, C. W. *Nat Rev Cancer* **2011**, *11*, 481.
(82) Hang, C. T.; Yang, J.; Han, P.; Cheng, H. L.; Shang, C.; Ashley, E.; Zhou, B.; Chang, C. P. *Nature* **2010**, *466*, 62.
(83) Hargreaves, D. C.; Crabtree, G. R. *Cell Res* **2011**, *21*, 396.
(84) Reisman, D., Glaros, S., and Thompson, E.A. *Oncogene* **2009**, *28*, 1653.
(85) Ansari, A. Z.; Reece, R. J.; Ptashne, M. *Proc Natl Acad Sci U S A* **1998**, *95*, 13543.
(86) Ma, J.; Ptashne, M. *Cell* **1987**, *50*, 137.
(87) Lohr, D.; Venkov, P.; Zlatanova, J. *FASEB J* **1995**, *9*, 777.
(88) Jiang, F.; Frey, B. R.; Evans, M. L.; Friel, J. C.; Hopper, J. E. *Society* **2009**, *29*, 5604.
(89) Kumar, P. R. Y., Y.; Sternglanz, R.; Johnston, S.A.; Joshua-Tor, L. *Science* **2008**, *319*, 1090.
(90) Gill, G. P., M. *Cell* **1987**, *51*, 121.
(91) Dickson, R. C. G., C.J.; Martin, A.K. *Nuc Acids Res* **1990**, *18*, 5213.
(92) Wittelsberger, A.; Thomas, B. E.; Mierke, D. F.; Rosenblatt, M. *FEBS Lett* **2006**, *580*, 1872.
(93) Wittelsberger, A. M., D.F.; Rosenblatt, M. *Chem Biol Drug Des* **2008**, *71*, 380.
(94) Rihakova, L. D., M.; Auger-Messier, M.; Pérodin, J.; Boucard, A.A.; Guillemette, G.; Leduc, R.; Lavigne, P.; Escher, E. *J Recept Signal Transduct Res* **2002**, *22*, 297.
(95) Knowles, J. R. *Acc. Chem. Res.* **1972**, *5*, 155.

Chapter 3 Tandem reversible and irreversible crosslinking (TRIC) using genetically incorporated Bpa captures the direct interactions of DNA bound activators in S. cerevisiae.

3.A.Summary[3]

Methods for capturing protein-protein interactions (PPIs) in their native environment are critical for the construction of complete interaction maps of protein networks involved in such cellular processes as protein folding, signaling, and transcription. Often these networks rely on transient and moderate affinity interactions to execute their core function, historically the most difficult class of PPIs to isolate using traditional biochemical methods. Described herein is an approach for capturing PPIs in *S. cerevisiae* that employs the genetically incorporated photo-crosslinking amino acid *p*-benzoyl-L-phenylalanine (Bpa). In these studies, Bpa is site-specifically incorporated in the transcriptional activation domain (TAD) of the activator VP16 using nonsense suppression technology optimized for use in yeast. *In vivo* photocrosslinking with Bpa reveals a direct contact between VP16 and the general transcription factor TATA-binding protein (TBP). To our knowledge, this is the first time a combinatorial crosslinking approach using tandem reversible formaldehyde crosslinking and irreversible

[3] This chapter is primarily comprised of data that is unpublished. Contributions to the work in this chapter is as follows: Amanda Dugan was responsible for the experimental design and execution of experiments testing activator crosslinking to myc-TBP, endogenous TBP and initial TRIC data. Rachel Pricer was responsible for obtaining a high-quality figure of the VP16-TBP TRIC interaction and she and Cassandra Joiner were additionally responsible for the execution of PCR experiments examining localization to the Gal1 promoter.

Bpa photo-crosslinking (TRIC) has been used to capture the direct interactions of DNA bound transcriptional activators in live yeast. This methodology is further used to identify TBP as a direct, cellular target of promoter bound VP16, thus resolving a nearly decade long debate over the relevance of this interaction. As such, TRIC can be used to capture direct interactions in a variety of PPI networks.

3.B. Background

Nearly every physiological process requires an intricately woven network of protein-protein interactions (PPIs) to dutifully carry out the executive orders of the cell. Often these networks are replete with PPIs that vary greatly in interaction interface areas, affinities, and lifetimes.[1] Illustrative of this complexity are transcriptional PPI networks whose key players, transcriptional activators, must recruit numerous multi-protein coactivator complexes in order to upregulate gene expression.[2] *In vivo* co-localization and chromatin immunoprecipitation (ChIP) studies have been instrumental in identifying the complexes that are recruited by activators to the promoter, but the inherent limitations of these approaches have rendered them incapable of distinguishing the individual subunits within these complexes that serve as activator targets *in vivo*. For example, the formaldehyde used in ChIP experiments functions by non-specifically crosslinking protein-DNA and protein-protein interactions to covalently stabilize the complexes; however, because formaldehyde forms crosslinks with any nucleophilic side chains, this approach lacks the resolution to distinguish directly interacting subunits of complexes. Furthermore, affinity-based co-purification methods have proven to be poorly suited for studying transient PPIs, resulting in a plethora of conflicting reports that have hindered our understanding of the mechanisms surrounding coactivator recruitment at a given promoter.

A relevant example of this struggle is the mechanism of recruitment of TBP to promoters by the activator VP16. As with other activators, VP16 uses its transcriptional activation domain (TAD) to initiate co-localization of numerous multi-protein complexes, including the Swi/Snf and SAGA chromatin remodeling

and modifying complexes, respectively, as well as Pol II and the associated holoenzyme[3,4] One essential protein within the yeast holoenzyme, TBP, is recruited early on in transcription initiation and has been shown to localize to promoters in a VP16-dependent fashion.[3,5,6] Data from *in vitro* biochemical studies suggest that this recruitment occurs through a direct interaction with TBP while *in vivo* localization data from *Saccharomyces cerevisiae* supports an indirect mechanism whereby TBP recruitment is mediated through the SAGA chromatin modifying complex.[3,7-15] To resolve this and the many other outstanding questions surrounding the direct binding network of transcriptional activators, an alternative experimental strategy is needed.

Studies in Chapter 2 demonstrated that *in vivo* photocrosslinking with the genetically incorporated unnatural amino acid *p*-benzyol-L-phenylalanine (Bpa) is a powerful tool to covalently capture the direct targets of transcriptional activators in live yeast, including transient and moderate affinity interactions that have historically proven the most difficult to study. In this Chapter, implementation of the optimized incorporation system in yeast is used to site-specifically incorporate Bpa into the TAD of the activator VP16. Using *in vivo* photo-crosslinking, the mechanism of recruitment of TBP by VP16, just one of many longstanding conflicts in the literature, is resolved. Furthermore, the first example of the dual application of irreversible *in vivo* photocrosslinking and reversible formaldehyde crosslinking to isolate the direct interactions of promoter-bound activators is described. This approach reveals important details regarding the actions of activators occurring at a specific promoter.

3.C. The VP16-TBP interaction: the SAGA continues

In yeast, TATA-binding protein (TBP) is a core component of the general transcription factor TFIID complex and it is essential for yeast viability and RNA Pol II-mediated transcription of protein-encoding genes.[16,17] TFIID is one of the first general transcription factors to be recruited during transcription initiation.

During this process, the saddle-shaped TBP subunit of TFIID binds the minor groove of DNA, introducing a severe bend at the promoter and unwinding the double helix (Figure 3-1).[18,19] DNA recognition by TBP is required for formation of the pre-initiation complex, which includes RNA polymerase II itself. The C-terminal core domain of TBP is highly conserved among eukaryotes and in humans, and abnormal polyglutamine rich expansions in mutated TBP results in deregulated transcription and the onset of Spinocerebellar Ataxia-17, a Huntington-like neurological disorder.[20] Given the essential nature of this protein, it is not surprising that the mechanism by which it is recruited to gene promoters has been the focus of innumerable studies over the last several decades (Figure 3-1).

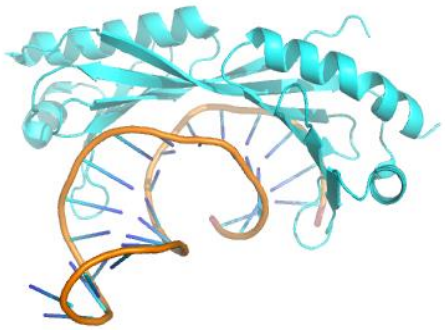

Figure 3-1 TBP binding to DNA (PDB: 1YTB). Binding of Tata binding protein (TBP) to DNA is believed to be the rate-limiting first step in assembly of the pre-initiation complex (PIC) during transcription initiation. As depicted here, binding of TBP to the minor groove of DNA induces a bend in the double helix, promoting the unwinding of DNA for RNA Polymerase II function.

Studies with *in vivo* formaldehyde crosslinking have demonstrated TBP recruitment by VP16, but were unable to clarify whether this interaction occurs through a direct or indirect mechanism.[3] One model has been proposed in which VP16 directly targets the concave DNA binding surface of TBP during

recruitment. This model is substantiated by demonstrations that point mutations within the DNA binding surface of TBP abrogate VP16 binding in addition to evidence demonstrating the ability of VP16 to correctly orient TBP on the promoter.[7,8] Because an activator-TBP complex has the potential to interfere with TBP binding to the TATA box, the nature of this particular interaction has been postulated to be transient in that it does not persist during PIC formation and function. In addition to these studies, significant *in vitro* data exists to support a direct interaction between VP16 and TBP.[9-12] Moreover, in line with previous biochemical data, *in vitro* crosslinking indicates that TBP has more than one binding site for VP16 as several different VP16-TBP crosslinked species were observed by Western blot.[13]

Although there is strong support for a direct interaction model, several studies have proposed that the recruitment of TBP by VP16 occurs through a more indirect mechanism. For example, one study suggests TBP recruitment occurs through activator interactions with TFIIB as VP16 is able to interact with a TFIID/A/B complex but unable to interact with a TFIID/A complex alone.[15] Another study suggests that the TBP/VP16 interaction is dependent on the SAGA subunit ADA2 after observing that GST-VP16 was unable to capture TBP from lysates lacking ADA2.[21] Finally, a study from the Hahn lab demonstrated that deletion of the SAGA subunit Spt3 results in loss of TBP recruitment by VP16, also supporting an indirect model of recruitment.[14]

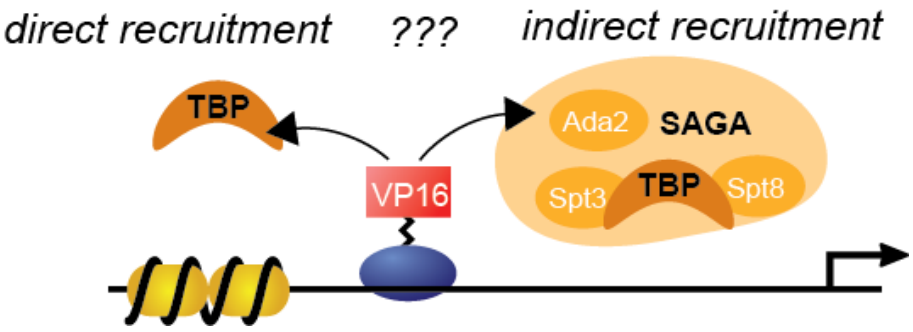

Figure 3-2 The mechanism by which VP16 recruits TBP to the Gal1 promoter in yeast remains unclear. Several in vitro binding studies indicate that TBP is a putative target of VP16 in cells, but *in vivo* co-

localization studies support a more indirect model of recruitment, one that is mediated through the SAGA chromatin modifying complex.

3.C.1. Examining crosslinking of VP16 to myc-TBP in yeast

Like many other transcriptional complexes, both SAGA and TFIID have been postulated to be combinatorially assembled in cells; therefore it is critical to examine the TBP-VP16 interaction in the native context where the full influence of complex subunit exchange can be considered.[22,23] To determine if TBP is recruited through a direct interaction with VP16, *in vivo* photocrosslinking was carried out with a VP16 construct described in Chapter 2. Briefly, a LexA+VP16-FLAG plasmid was created by fusing the DNA binding domain of the bacterial repressor protein LexA to either the amino terminal subdomain of the VP16 TAD (412-456, VP16N) or carboxy terminal VP16 TAD (446-490, VP16C). These plasmids were then transformed into a yeast strain bearing an integrated LacZ reporter gene under the control of a GAL1 promoter containing two LexA binding sites. Co-transformed along with the pLexA-VP16-flag plasmids was a plasmid for the Tyr tRNA/RS Bpa required for incorporation of Bpa at position 444 in VP16N and position 475 in VP16C. Incorporation of Bpa at both positions tested showed an excellent incorporation and activity profile. Therefore, these constructs were selected for use in crosslinking experiments (Figure 3-3).

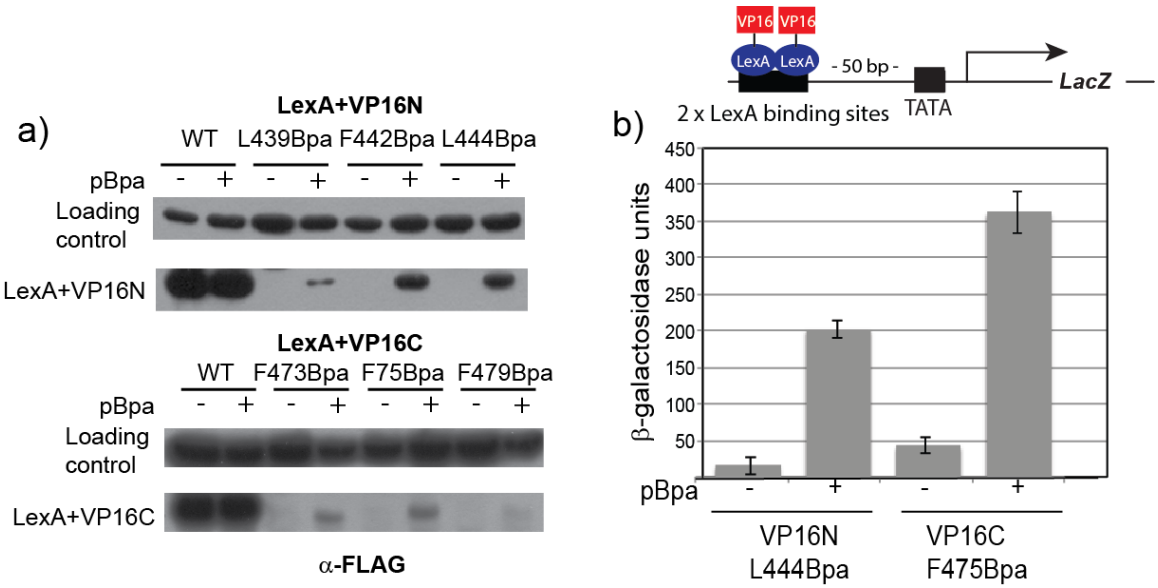

Figure 3-3 LexA+VP16-FLAG plasmids were created by fusing the DNA binding domain of the bacterial repressor protein LexA to either the amino terminal subdomain of the VP16 TAD (412-456, VP16N) or carboxy terminal VP16 TAD (446-490, VP16C). These plasmids were transformed into a yeast strain bearing an integrated LacZ reporter gene under the control of a GAL1 promoter containing two LexA binding sites. Co-transformed along the pLexA-VP16-flag plasmids was a plasmid for the Tyr tRNA/RS Bpa required for incorporation of Bpa. A) Three positions in each subdomain of VP16 were tested for Bpa incorporation and transcriptional activity as measured by expression of the integrated LacZ reporter gene. Given the excellent incorporation and activity of the LexA+VP16N L444Bpa and LexA+VP16C F475Bpa constructs were chosen for crosslinking experiments.

Initial experiments were carried out with a myc-tagged TBP construct and LexA+VP16 L444Bpa or LexA+VP16C F475Bpa. Because C-terminal tagging of TBP impairs its function in yeast, the myc tag was placed on the amino terminus of the protein.[24] Yeast cells expressing either VP16 construct and myc-TBP were then grown to mid-log phase and irradiated for 30 minutes with 365 nm UV light while kept on ice. Post-irradiation, the yeast were lysed and their lysates immunoprecipitated with LexA antibody to precipitate all activator-containing complexes. This mixture was then resolved and analyzed by Western blot with a myc-HRP antibody to detect the presence of a VP16-TBP interaction. For both VP16 constructs tested, a covalent adduct with TBP was detected (Figure 3-4, top panel). Thus, in this experimental context, a direct interaction between VP16 and TBP in living cells is occurring. Further supporting these data, a point mutation was introduced into myc-TBP at a position shown in vitro to be critical

for binding to VP16 (L114K) and this mutation abrogated crosslinking of VP16 to TBP, indicating that the presence of Bpa in VP16 was not responsible for mediating the interaction.[5] Furthermore, myc-TBP and myc-TBP L114K expressed at the same levels in yeast, thus the lack of crosslinking was not due to poor mutant expression issues (Figure 3-4, bottom panel).

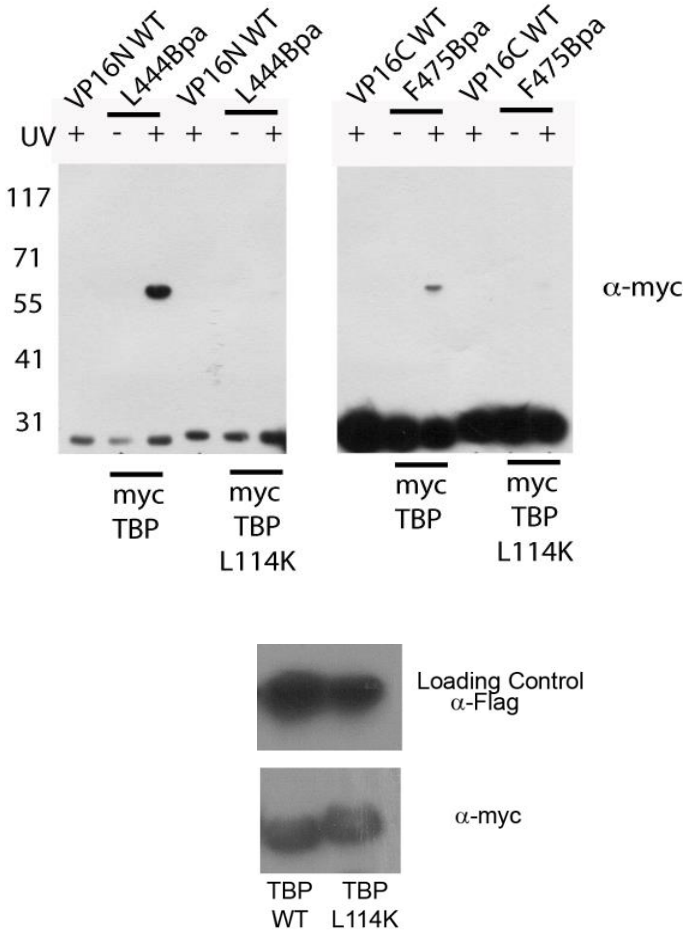

Figure 3-4 (top) VP16 directly contacts myc-TBP in live yeast. Yeast co-expressing LexA+VP16N L444Bpa or LexA+VP16C F475Bpa and myc-TBP were irradiated with UV light to activate Bpa and covalently capture interacting proteins. After lysis, lysates were IP'd with a LexA antibody and the resulting Western blot was probed with an α-myc antibody to detect a VP16-TBP interaction. Introduction of a point mutation (L114K) in myc-TBP known to be important for TBP recruitment and binding by VP16 also abrogates crosslinking in vivo. (bottom) Abrogation of crosslinking with TBP L114K is not due to poor mutant expression compared to WT TBP.

To further define the interaction, Bpa was incorporated at residue 114 in myc-TBP, a position previously shown to be critical in maintaining an interaction with VP16 *in vitro*, and crosslinking was carried out in cells co-expressing LexA+VP16N WT. Consistent with earlier data, a covalently bound VP16-TBP complex was isolated from irradiated yeast (Figure 3-5). Thus, regardless of whether Bpa was incorporated in the activator or the recruited transcription factor, the VP16-TBP interaction was captured in living yeast.

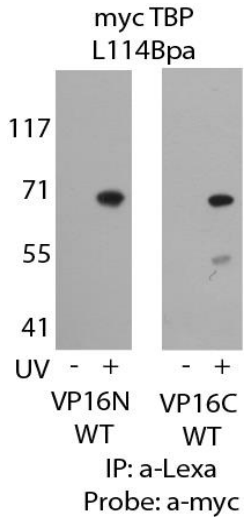

Figure 3-5 TBP crosslinks to LexA+VP16N WT and LexA+VP16C WT. "Reverse" crosslinking was performed with Bpa containing TBP and wild-type constructs of VP16N and VP16C. Yeast co-expressing myc-TBP L114Bpa and LexA+VP16N WT or LexA+VP16C WT were subjected to UV crosslinking and subsequently lysed. Lysates were IP'd with an a-LexA antibody and the resulting Western was probed with an α-myc antibody. Placing the Bpa on the partner protein TBP results in the same outcome, with a covalent VP16-TBP adduct observed by Western blot.

3.C.2. Crosslinking of VP16 to endogenous TBP

Given the sensitivity of PPI networks to fluctuations in protein levels, specifically the overexpression of TBP in yeast, crosslinking of VP16 to endogenous TBP was next pursued. First, a set of four antibodies were tested for their ability to immunoprecipitate free TBP from yeast lysate, and one antibody performed most consistently in these experiments. With this in hand, live yeast co-expressing LexA+VP16N L444Bpa or LexA+VP16C F475Bpa and the Bpa tRNA/RS were irradiated with UV light and subsequently lysed. Yeast lysates were

immunoprecipitated with TBP antibody followed by Western blot identification of TBP-VP16 covalent products using an α-FLAG antibody. Whereas a crosslinked product was readily observable for VP16 and myc-TBP, no observable crosslinked product was detected in these studies, even after several rounds of optimization. As TBP is required for transcription by all three RNA polymerases, it is present in fairly high abundance in yeast cells (about 20,000 copies/cell).[16] However, the covalently bound VP16-TBP was at a low enough concentration to be outside the limits of the current detection capabilities. According to Sigma, the manufacturer of the α-Flag antibody used in these studies, increasing the number of Flag tags on the protein being investigated results in a 10- to 20- fold increase in detection capability of the antibody, thus allowing for the detection of low femtomolar quantities of protein. After altering the Flag tag on VP16 to a 5x Flag epitope, crosslinking experiments were repeated and a direct contact between endogenous TBP and VP16 was observed (Figure 3-6, left panel, 60 kDa). Furthermore, when a double point mutation (L439P, F442P) was introduced into the VP16N TAD, crosslinking to TBP was abolished, indicating that this interaction is specific and not mediated by the presence of Bpa or the enhanced Flag tag (Figure 3-6, right panel). Thus, *in vivo* photocrosslinking was able to capture a direct interaction with an endogenously expressed transcription factor in live yeast.

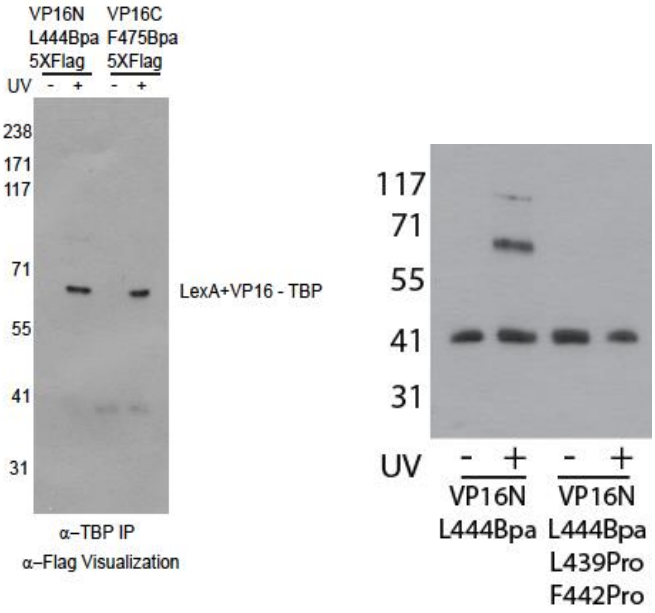

Figure 3-6 VP16 crosslinks to endogenous TBP. LexA+VP16N L444Bpa-5xflag and LexA+VP16C F475Bpa were transformed into yeast and cells were subjected to crosslinking under UV light. After lysis, cell lysates were immunoprecipitated with an α-TBP antibody and the subsequent Western blot was probed with an α-Flag antibody to detect a covalently bound TBP-VP16 complex. (left blot) We found that both subdomains directly contacted TBP in cells, as determined by the presence of a band corresponding to the molecular weight of a TBP-VP16 product (60 kDa). (right blot) Introduction of a double point mutation in the VP16N TAD known to impair VP16 activity and binding resulted in the abrogation of crosslinking to TBP in yeast. This data indicates that the addition of a 5xflag tag and the presence of Bpa within the TAD are not responsible for the formation of this interaction in yeast.

3.D. TBP is a shared target of amphipathic activators in yeast

As articulated in Chapters 1 and 2, amphipathic activators share little sequence homology yet have been suggested to share a common set of targets in the transcriptional machinery.[25,26] In addition to VP16, several lines of evidence exist to support the recruitment of TBP to the GAL1 promoter by the yeast activator Gal4. In the case of Gal4, the available literature suggests that TBP recruitment occurs through an indirect mechanism in which Gal4 recruits the SAGA complex and SAGA recruits TBP. Several SAGA subunits have been shown to be required for both SAGA and TBP recruitment by Gal4.[27] Of particular interest is the Spt3 subunit which has been shown to have no role in SAGA complex integrity but is believed to play a significant role in TBP recruitment.[28-30] Genetic and biochemical studies have supported a Spt3p-TBP interaction and *in vitro* and

in vivo crosslinking studies were able to finally capture a direct interaction between Spt3 and TBP.[14,31-33] Furthermore, deletion of Spt3 interferes with the recruitment of TBP but does not affect the promoter occupancy of SAGA or Gal4.[30,34] Thus, other amphipathic activators were examined for their targeting of TBP *in vivo*, including the yeast activators Gal4 and Gcn4. LexA+Gal4-5xflag and LexA+Gcn4-5xflag fusion proteins were expressed separately in yeast and subjected to *in vivo* crosslinking as described earlier. Immunoprecipitation of the resulting cell lysates with TBP antibody followed by Western blot detection of the covalent complex with a α-Flag antibody indicates that endogenous TBP is indeed a shared target of these three amphipathic activators, suggesting a common direct mechanism of recruitment of this essential yeast protein (Figure 3-7).

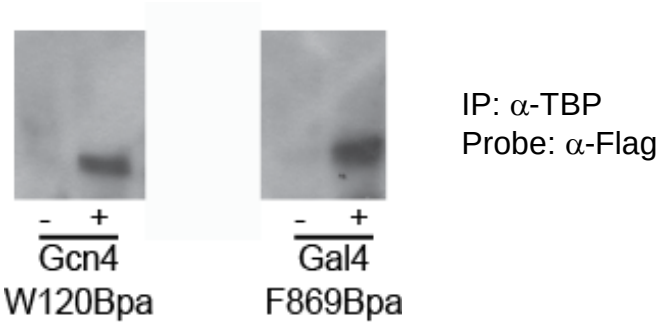

Figure 3-7 LexA+Gcn4 W120Bpa and LexA+Gal4 F869Bpa both crosslink to endogenous TBP in live yeast. After UV irradiation, yeast were harvested and the resulting lysates were immunoprecipitated with an α-TBP antibody. The following Western blot was probed with α-Flag antibody to detect the presence of a covalently bound Gcn4 or Gal4. In both activators, we were able to detect a direct interaction with endogenous TBP.

3.E. Tandem reversible and irreversible crosslinking (TRIC)

While *in vivo* photocrosslinking with Bpa is a powerful approach for capturing PPIs in their native environment, it is limited in its ability to differentiate the location and timing of the interaction, information that is critical to advancing our understanding of the mechanism of gene expression *in vivo*. Thus, a method that affords the ability to examine direct activator interactions with these variables in mind would be tremendously advantageous. So far, chromatin

immunoprecipitation (ChIP) has come the closest to this goal, using formaldehyde chemical crosslinking to rapidly initiate protein-protein and protein-DNA crosslinks in cells. However, although researchers have used this approach for years to examine the complexes recruited to specific regions of the yeast genome, ChIP lacks the resolution to define the specific proteins within these recruited complexes that are engaging in direct interactions with transcriptional activators and other coactivator complexes during upregulation of gene expression. Given the utility of Bpa crosslinking in achieving this, the complementary approaches of reversible formaldehyde crosslinking and irreversible Bpa photo-crosslinking (TRIC) were combined to examine the direct interactions of DNA bound proteins. In TRIC, cells are first crosslinked with formaldehyde which rapidly permeates the cell and reacts non-specifically with primary amines to covalently bind protein-DNA and protein-protein interactions. Formaldehyde crosslinked cells are then UV irradiated to activate the Bpa, thus forming site-specific covalent crosslinks with the direct binding partners of the activator within the immobilized complexes. The chromatin of the cells is then isolated and washed to remove non-covalently bound protein and, following sonication, the DNA bound complexes can be immunoprecipitated. Upon reversal of the formaldehyde crosslinks, only the interactions that were covalently captured with Bpa remain intact and these irreversibly crosslinked PPIs can be resolved on a Western blot, as before (Figure 3-8). The work described herein demonstrates the utility of this methodology in building a comprehensive map of PPIs critical to the regulation of eukaryotic gene expression, further allowing for the examination of direct interactions in a promoter-specific context while also setting a foundation for the application of this approach in examining other complex PPI networks in yeast.

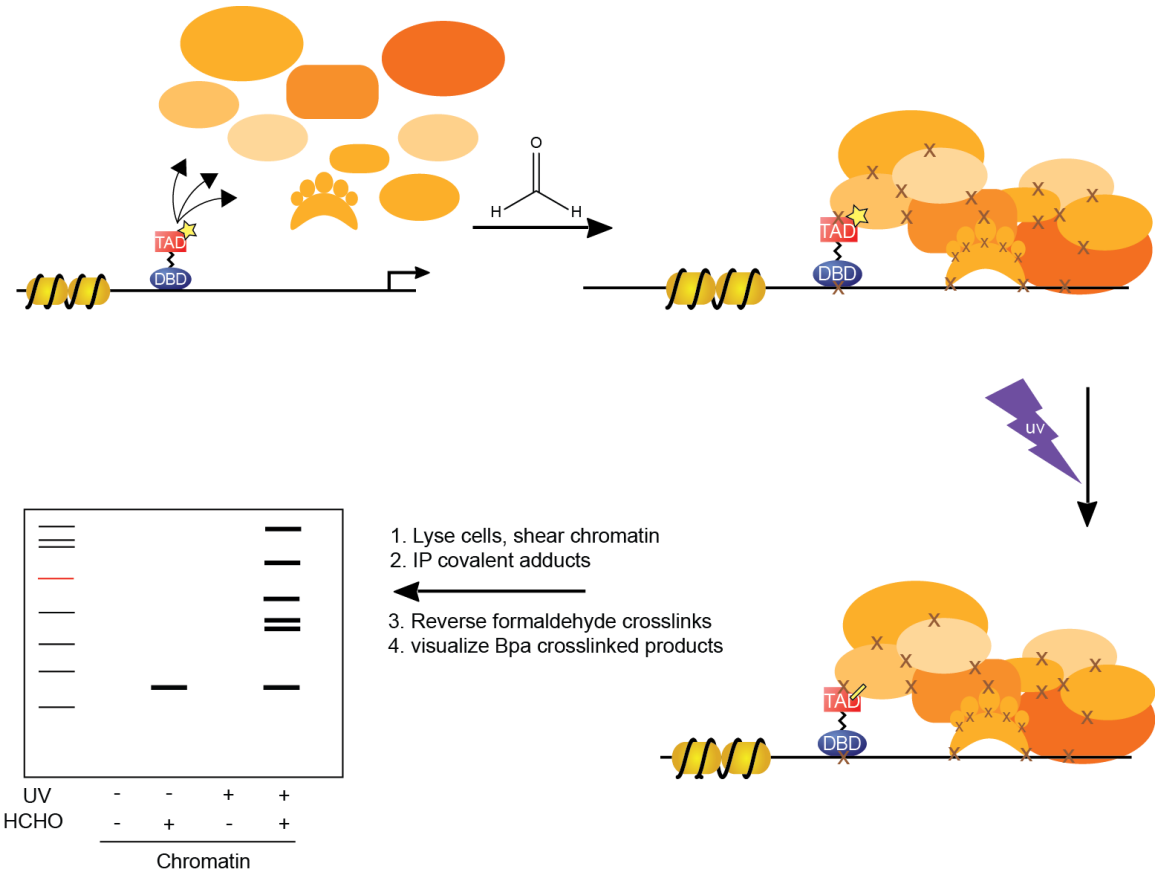

Figure 3-8 Overview of TRIC. Briefly, cells are treated with formaldehyde to rapidly initiate protein-protein and protein-DNA crosslinks. The formaldehyde reaction is then quenched and the cells are irradiated with UV light to activate the Bpa and capture directly interacting proteins within the formaldehyde crosslinked complex. The chromatin fraction of the cells is then isolated, washed, and sonicated to shear and solubilize the chromatin. Following immunoprecipitation, the formaldehyde crosslinks are reversed, leaving only the Bpa covalent irreversible products. These covalent products can then be detected by Western blot, thus revealing the direct interactions of DNA bound activators.

3.E.1. Capturing direct targets of DNA bound VP16

Using the specific interaction between TBP and VP16 to test the TRIC method, we first transformed yeast with LexA+VP16N L444Bpa and the Tyr tRNA/RS required for Bpa incorporation. Cells were harvested at mid-log phase OD660 and then formaldehyde crosslinked for 20 minutes to establish protein-protein and protein-DNA crosslinks, essentially rendering these interactions immobilized in yeast. Formaldehyde crosslinked cells were then irradiated under UV light for 30 minutes on ice to activate Bpa and form irreversible covalent crosslinks between VP16 and TBP. Upon cell lysis, the chromatin fraction of the cell was

isolated, washed to remove non-covalently bound proteins, and then sonicated to solubilize the chromatin. This soluble fraction was then immunoprecipitated with a TBP antibody and immobilized to a magnetic bead. The formaldehyde crosslinks were reversed in elution buffer with boiling at 95 °C, leaving only the irreversible Bpa crosslinks intact, and the immunoprecipitated proteins were then separated using SDS-PAGE. Any detectible VP16-TBP complex from the DNA fraction was then visualized on Western blot with an α-FLAG antibody. As expected, immunoprecipitation of yeast lysates yields a covalent adduct between VP16 and TBP in the presence of UV light as well as treatment with UV and formaldehyde. However, when the chromatin fraction of the cells is isolated, a TBP-VP16 crosslink is only observed when formaldehyde was added to cells prior to UV irradiation (Figure 3-9). Thus, this method confers a unique advantage in visualizing PPIs of DNA-bound proteins, exemplified here with a DNA-bound activator and an essential transcription factor.

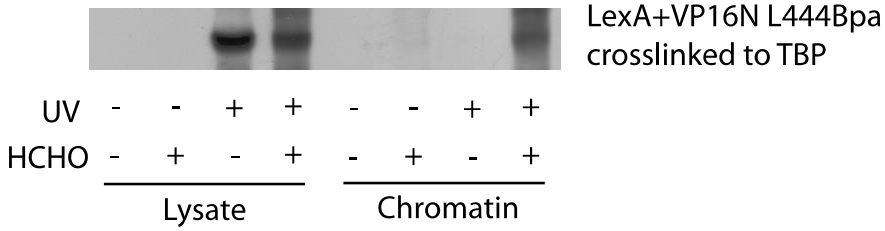

Figure 3-9 Tandem reversible and irreversible crosslinking captures the direct targets of DNA bound transcriptional activators. Identical cultures of yeast expressing LexA+VP16N L444Bpa were either crosslinked under UV light for 30 minutes, formaldehyde crosslinked with 1% formaldehyde for 20 minutes, or treated with a combination of both procedures (formaldehyde followed by UV crosslinking). Cell lysates were then immunoprecipitated per the standard protocol with an α-TBP antibody and the samples were boiled for 20 minutes at 95*C to reverse the formaldehyde crosslinks, leaving only the irreversible Bpa crosslinked products intact. The chromatin fractions of these cultures were washed to remove non-covalently bound protein and then the chromatin was sheared and solubilized using sonication. Soluble chromatin was then immunoprecipitated with an α-TBP antibody and the formaldehyde crosslinks were reversed by boiling for 20 minutes, leaving only the irreversible Bpa crosslinked products intact. The resulting Western blot of lysates and chromatin were probed with an α-Flag antibody. As expected from earlier experiments (Figure 6), we see a covalent VP16-TBP complex present in cell lysates from yeast irradiated with UV light. For the chromatin fraction, the formaldehyde is required to stabilize DNA-protein interactions, so only in these samples do we immunoprecipitate protein from the chromatin fraction. However, upon reversal of the formaldehyde, only in the lane where Bpa was activated to form an

irreversible covalent bond do we see the VP16-TBP complex, indicating that this interaction occurs on DNA. Experiments done alongside Rachel Pricer.

As mentioned previously, the yeast strain used in these studies carries an integrated LacZ reporter gene that is under the control of a Gal1 promoter that contains two LexA binding sites for our activator constructs. Thus, TRIC experiments should report on interactions at this modified promoter. To verify this method, the TRIC protocol was carried out again but this time the DNA was examined. As a control, an aliquot of soluble chromatin was retained following sonication. Using primers designed to amplify the Gal1 promoter, PCR was carried out on the input as well as the DNA retained post-IP. The PCR products were separated on an agarose gel and stained with ethidium bromide. As shown in Figure 3-10, the input lanes all show the presence of the Gal1 promoter, as does a LS41 genomic DNA control. However, only in conditions where formaldehyde is added to the culture is Gal1 amplified from the immunoprecipitated material (Figure 3-10). Thus, the VP16-TBP interaction seen in Figure 3-9 is indeed bound to Gal1, as predicted.

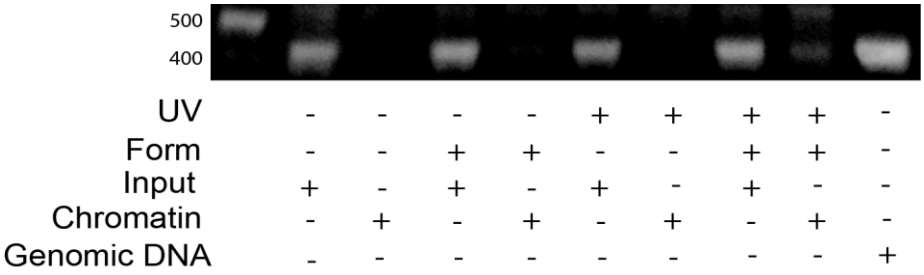

Figure 3-10 VP16-TBP interaction occurs at the Gal1 promoter. The yeast strain used in these studies has an integrated LacZ reporter gene under the control of a Gal1 promoter that has two LexA binding sites. As such, the DNA bound interactions we observe for LexA+VP16N L444Bpa should only be occurring at this spot in the yeast genome. To verify that this is the case, we carried out TRIC experiments in yeast, this time focusing only on the chromatin fraction. Like before, chromatin was washed to remove non-covalently bound protein and then sonicated to shear and solubilize the chromatin. For each condition tested, we saved an aliquot of solubilized chromatin as input for this experiment (pre-immunoprecipitation). We then IP'd the chromatin with a TBP antibody followed by reversal of the formaldehyde crosslinks and digestion of protein and RNA with proteinase K and RNase I, respectively. The remaining DNA and the input DNA was then subjected to PCR amplification with Gal1 promoter primers. The Gal1 product is about 440 bps long, approximately the size of the product we see on the DNA gel above. As expected, all input lanes and our

genomic DNA control show a large amplified band for Gal1, indicating its presence in the chromatin fraction. After immunoprecipitation, we only see the Gal1 DNA in conditions where formaldehyde was used to stabilize TBP and VP16 to the DNA, thus allowing us to IP it out (+F, -UV; +F+UV). The leftmost lane contains a 1 kilobase ladder.

3.F. Conclusion

TBP is just one of many essential transcriptional proteins whose recruitment by activators is documented in a series of conflicting reports in the literature. Using the long-contested VP16-TBP interaction as a model, we demonstrate that *in vivo* photocrosslinking is a valuable tool in clarifying the nature of the interactions that exist in transcriptional PPI networks, a feat that traditional methods were unable to accomplish. This work led to the validation of TBP as a direct cellular binding partner of VP16 and additionally identified that this protein is a shared target with the activators Gal4 and Gcn4. Further work will need to be carried out to determine if all three activators target the same region on TBP or if they differentially recruit TBP by contacting different surfaces, information that can be exploited when developing small molecules to specifically target this critical interaction.

We furthermore demonstrate the utility of tandem reversible and irreversible crosslinking (TRIC) in capturing the *in vivo* direct interactions of DNA bound activators, an accomplishment that, to the best of our knowledge, has not been achieved up until this point. This approach should reveal critical information about the mechanism of coactivator recruitment to DNA, thus yielding key insights to guide drug discovery efforts. More broadly, this method should be useful in the capture and identification of other DNA bound PPIs, including histone modifying and remodeling enzymes and proteins involved in DNA repair pathways. Additionally, this approach could even be used to look at the direct interactions within PPI networks localized to other areas of the cell including membrane PPIs as well as those occurring in the mitochondria, peroxisomes, and endoplasmic reticulum. As the TRIC method becomes more refined, the inclusion of rapid mixing protocols that reduce formaldehyde crosslinking to a

shorter timescale should allow for the addition of a time scale in these experiments. This would allow for changes in the activator-interaction profile during recruitment to be observed in a time-dependent fashion after induction, thus building a more complete picture of transcriptional interaction maps.

3.G. Experimental

Yeast Strain LS41 [JPY9::pZZ41, *Mata his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41] was used for the crosslinking experiments. Bpa 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).

Plasmid name	**Function**
pLexAVP16N WT	Expresses LexA(1-202)+VP16 (413-456)+FLAG tag
pLexAVP16C WT	Expresses LexA(1-202)+VP16 (446-490)+FLAG tag
pLexAVP16N 444TAG	Expresses LexA(1-202)+VP16 (413-456)+FLAG tag with a TAG replacing the codon of the existing amino acid

pLexAVP16C 475TAG	Expresses LexA(1-202)+VP16 (446-490)+FLAG tag with a TAG replacing the codon of the existing amino acid
pLexAGcn4 120TAG 5xflag	Expresses LexA(1-202)+Gcn4(107-144)+5xFLAG tag with a TAG codon 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
pMyc TBP, pMyc TBP L114K, pMyc TBP L114Bpa	Expresses full length TBP fused to c-myc tag on N terminus of protein;

	replacing position L114 with either a Lysine or TAG codon for Bpa incorporation
pLexAVP16N 444TAG 5xFlag	Expresses LexA(1-202)+VP16(413-456)+5xFLAG tag
pLexAGal4 849TAG	Expresses LexA(1-202)+Gal4(840-881)+FLAG tag with a TAG replacing the codon of the existing amino acid
pLexAVP16C 475TAG 5xFlag	Expresses LexA(1-202)+VP16(446-490)+5xFLAG tag

3.G.1. Construction of plasmids

1. pLexAVP16 N and pLexAVP16C

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 treated with calf intestinal phosphate 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.

2. pLexAVP16N 444TAG, pLexAVP16C 475TAG

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.

3. pLexAGcn4(107-144)

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.

4. pLexAGal4 (840-881)

pLexA(1-202)+Gal4(840-881) was created as previously described (Majmudar, CY et al, 2009)

5. pGADT7 myc-TBP

A high copy plasmid pGADT7-myc expressing TBP under the ADH1 promoter, N-terminally tagged with the c-Myc epitope was constructed by amplifying the DNA sequence encoding Med15(1-415) from yeast genomic DNA using primers (5'-catCATATGATGGCCGATGAGGAACGTTTAAAGG-3') and (5'-atgCTCGAGTCACATTTTTCTAAATTCACTTAGC -3') and inserted into NdeI and XhoI 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'-AGCTATGGAACAAAAGTTGATTTCTGAAGAAGATTTGGGATCCAATGCATATGATCT-3') and (5'-AGCTTGATCATATGCATTGGATCCCAAATCTTCTTCAGAAATCAACTTTTGTTCCAT-3').

3.G.2. Incorporation of pBpa into LexA(1-202)+VP16N and LexA(1-202)+VP16C and expression of myc-TBP in S. cerevisiae pellet

LS41 yeast was transformed with various pLexAVP16 TAG mutant plasmids and pSNRtRNA-pBpaRS plasmid. Individual colonies were grown to saturation in 5 mL SC media lacking histidine and tryptophan for selection and 2% Raffinose at 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 the cell pellets were 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). To test expression of the myc-TBP construct, the same protocol was followed except that LS41 were additionally transformed with the pGADT7-mycTBP plasmid and grown in SC media lacking histidine, tryptophan and leucine. Lysates were analyzed using Western blot with anti-myc antibody (Santa Cruz Biotech, sc-40).

3.G.3. In vivo cross-linking

To perform in vivo cross-linking, individual colonies of each pLexAVP16 TAG mutant were grown in 5 mL SC media containing 2% Raffinose 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 100 mM pBpa dissolved in 1M NaOH and 1 mL 1M HCl were added to the above cultures. The cultures were incubated overnight at 30 °C with agitation to an OD660 of ~1.0. When cultures reached the appropriate OD_{660}, the cells were spun down by centrifuging at 3901 rcf at 4^{o}C for 5 min. following which the cell pellets were washed with SC media lacking histidine and tryptophan. The cell pellets were resuspended in 2mL SC media lacking histidine and tryptophan + 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 mycTBP, the procedure was identical except that cells were grown in SC media lacking histidine, leucine, and tryptophan. 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 ~ 8 uL prewashed protein G agarose beads (Millipore) at 4 ^{o}C. After immunoprecipitation, the beads were washed 6 times 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 ^{o}C until elution. The crosslinked sample was eluted from the beads by heating at 95 ^{o}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 HRP, Santa Cruz Biotechnology, Santa Cruz, CA).

3.G.4. β-Galactosidase assays

To evaluate the ability of each LexA+VP16 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 + 2% Galactose but lacking histidine and tryptophan for selection. The cells were grown to an OD of 1.5-2.0 and harvested. The activity of each construct was monitored using β-galactosidase assays as previously described (Majmudar, C.Y. et al 2009).

3.G.5. Tandem reversible and irreversible crosslinking

To perform TRIC, individual colonies of each pLexAVP16 TAG mutant were grown in 5 mL SC media containing 2% Raffinose 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 100 mM pBpa dissolved in 1M NaOH and 1 mL 1M HCl were added to the above cultures. The cultures were incubated overnight at 30 °C with agitation to an OD660 of ~1.0. When cultures reached the appropriate OD_{660}, the

cells were spun down by centrifuging at 3901 rcf at 4^{o}C for 5 min. following which the cell pellets were washed with SC media lacking histidine and tryptophan.

The cell pellets only receiving UV treatment were resuspended in 2mL SC media lacking histidine and tryptophan + 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 cell pellets receiving formaldehyde treatment were resuspended in 1 mL DI water and added to a 100 mL solution of 1% formaldehyde in water (2.7 mL of 37% formaldehyde solution into 97.3 mL water). Cells were crosslinked with formaldehyde for 20 minutes before being quenched with 30 mL of 2M Glycine. Cells were then centrifuged and washed with 50 mL DI water. Samples intended to additionally receive UV crosslinking were resuspended in 2 mL SC media lacking histidine and tryptophan +2% Raffinose, 2% Galactose and transferred to a small cell culture dish and subjected to UV irradiation at 365 nm UV light (Eurosolar 15 W UV lamp with cooling for 0.5h.

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 for 45 minutes at 4 °C. We found in these studies that complete cellular lysis is necessary to eliminate background signal caused by cell lysis during sonication. Subsequent lysates were immunoprecipitated with 8 uL TBP antibody (santa cruz, sc-33736) and incubated for 2 hours at 4 deg C. The remaining pellet is then washed 4x with “Harsh” ChIP buffer (50 mM HEPES-KOH pH 7.5, 1M NaCl, 1 mM EDTA, 1% Triton X-100, 1% Na-Deoxycholate) followed by 2 washes with regular ChIP 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)). Pellets were resuspended in 600 µL ChIP lysis buffer containing protease inhibitor and sonicate at a setting of 10% (3 on Marsh lab sonicator, Fisher Scientific) for 2

minutes with 30 sec pulse on/off. Samples were then centrifuged in the cold room for 20 minutes at max speed. Soluble chromatin was then removed from the pellet and immunoprecipitated with TBP antibody (santa cruz, sc-33736) for 2 hours, 4 deg C. 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 ~ 8 uL prewashed protein G agarose beads (Millipore) at 4 ^{o}C. After immunoprecipitation, the beads were washed 6 times 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 ^{o}C until elution. The crosslinked sample was eluted from the beads and formaldehyde crosslinks reversed by heating at 95 ^{o}C for 20 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).

For studies examining the DNA IP'd during TRIC, the TRIC protocol was followed with the exception of 50 uL of solubilized chromatin being saved prior to immunoprecipitation. Additionally, lysates were discarded in these experiments.

To examine the size of the sheared chromatin, 50 µL TE/SDS was added to the Input samples and incubated overnight at 65°C to reverse crosslinks. 2.5 µL proteinase K (20 mg/mL stock) was then added and incubated at 50°C for 3 hours to digest proteins, followed by a PCR cleanup. 0.5 µL RNAse A (1 mg/mL stock) was added and then incubated at 37°C for 30 min. samples were visualized on 1% agarose gel. Smear should show between 300-600 bps.

For PCR on TRIC samples, 90 µL TE/SDS was added to 50 uL input and incubated overnight at 65°C followed by PCR Cleanup and elution in 58 µL EB buffer. DNA is then measured and PCR set up (See table). For Chromatin, beads were washed 2x with lysis buffer, 1 time with 500 mM NaCl lysis buffer, 1

time with wash buffer, and 1 time with TE buffer (10 mM Tris-HCl, 1 mM EDTA, 0.01% SDS (5 g in 500 mL ex)

. 50 µL elution buffer(50 mM Tris-HCl, 10 mM EDTA, 1% SDS) was then added to the beads and vortexed briefly before incubating at 65°C for 30 minutes, with vortexing every 5 minutes to resuspend the beads. Beads were centrifuged for 30 sec at 3000 rpm and the eluent transfered to a new tube. 120 µL TE/SDS was added and incubated overnight at 65°C followed by a PCR cleanup, elution in 58 µL EB buffer. Measure DNA on nanodrop and set up PCR (See table). Products were visualized on a 1% agarose gel stained with ethidium bromide.

For 25 µL Final PCR Volume
2.5 µL 10x Pfu buffer
0.5 µL 10 mM DTNPs
0.5 µL each primer
0.5 µL Pfu Turbo
2.0 µL template DNA
17.5 µL water
1 µL 50 mM $MgCl_2$

3.H. References – Chapter 3.

(1) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. *ACS Chem. Biol.* **2012**, *7*, 1311.
(2) Mapp, A. K.; Ansari, A. Z. *ACS Chem Biol* **2007**, *2*, 62.
(3) Hall, D. B.; Struhl, K. *J Biol Chem* **2002**, *277*, 46043.
(4) Chen, X.-F.; Lehmann, L.; Lin, J. J.; Vashisht, A.; Schmidt, R.; Ferrari, R.; Huang, C.; McKee, R.; Mosley, A.; Plath, K.; Kurdistani, S. K.; Wohlschlegel, J.; Carey, M. *Cell Rep* **2012**, *2*, 1061.
(5) Kim, T.; Hashimoto, S.; Kelleher, R.; Flanagan, P. *Nature* **1994**, *369*, 252.
(6) Herrera, F. J.; Triezenberg, S. J. *J Virol* **2004**, *78*, 9689.
(7) Nishikawa, J.-I.; Kokubo, T.; Horikoshi, M.; Roeder, R. G.; Nakatani, Y. *Proc Nat Acad Sci, USA* **1997**, *94*, 85.
(8) Kays, A. R.; Schepartz, A. *Biochemistry* **2002**, *41*, 3147.
(9) Nedialkov, Y. A.; Triezenberg, S. J. *Arch Biochem Biophys* **2004**, *425*, 77.
(10) Stringer, K. F.; Ingles, C. J.; Greenblatt, J. *Nature* **1990**, *345*, 783.
(11) Lin, Y. S.; Ha, I.; Maldonado, E.; Reinberg, D.; Green, M. R. *Nature* **1991**, *353*, 569.
(12) Shen, F.; Triezenberg, S. J.; Hensley, P.; Porter, D.; Knutson, J. R. *J Biol Chem* **1996**, *271*, 4827.
(13) Klein, J.; Nolden, M.; Sanders, S. L.; Kirchner, J.; Weil, P. A.; Melcher, K. *J Biol Chem* **2003**, *278*, 6779.
(14) Mohibullah, N.; Hahn, S. *Genes Dev* **2008**, *22*, 2994.
(15) Hori, R. T.; Xu, S.; Hu, X.; Pyo, S. *Nuc Acid Res* **2004**, *32*, 3856.
(16) Cormack, B. P.; Struhl, K. *Cell* **1992**, *69*, 685.
(17) Giaever, G.; Chu, A. M.; Ni, L.; Connelly, C.; Riles, L.; Veronneau, S.; Dow, S.; Lucau-Danila, A.; Anderson, K.; Andre, B.; Arkin, A. P.; Astromoff, A.; El-Bakkoury, M.; Bangham, R.; Benito, R.; Brachat, S.; Campanero, S.; Curtiss, M.; Davis, K.; Deutschbauer, A.; Entian, K. D.; Flaherty, P.; Foury, F.; Garfinkel, D. J.; Gerstein, M.; Gotte, D.; Guldener, U.; Hegemann, J. H.; Hempel, S.; Herman, Z.; Jaramillo, D. F.; Kelly, D. E.; Kelly, S. L.; Kotter, P.; LaBonte, D.; Lamb, D. C.; Lan, N.; Liang, H.; Liao, H.; Liu, L.; Luo, C.; Lussier, M.; Mao, R.; Menard, P.; Ooi, S. L.; Revuelta, J. L.; Roberts, C. J.; Rose, M.; Ross-Macdonald, P.; Scherens, B.; Schimmack, G.; Shafer, B.; Shoemaker, D. D.; Sookahi-Mahadeo, S.; Storms, R. K.; Strathern, J. N.; Valle, G.; Voet, M.; Volckaert, G.; Wang, C. Y.; Ward, T. R.; WIlhelmy, J.; Winzeler, E.; Yang, Y.; Yen, G.; Youngman, E.; Yu, K.; Bussey, H.; Boeke, J. D.; Snyder, M.; Phillipson, P.; Davis, R. W.; Johnston, M. *Nature* **2002**, *418*, 387.
(18) Wu, J.; Parkhurst, K. M.; Powell, R. M.; Brenowitz, M.; Parkhurst, L. J. *J Biol Chem* **2001**, *276*, 14614.
(19) Gaston, K.; Bell, A.; Busby, S.; Fried, M. *Nuc Acids Res* **1992**, *20*, 3391.
(20) Nakamura, K.; Jeong, S. Y.; Uchihara, T.; Anno, M.; Nagashima, K.; Nagashima, T.; Ikeda, S.; Tsuji, S.; Kanazawa, I. *Hum Mol Genet* **2001**, *10*, 1441.
(21) Barlev, N. A.; Candau, R.; Wang, L.; Darpino, P.; Silverman, N.; Berger, S. L. *J Biol Chem* **1995**, *270*, 19337.

(22) Belotserkovskaya, R.; Sterner, D. E.; Deng, M.; Sayre, M. H.; Lieberman, P. M.; Berger, S. L. *Mol Cell Biol* **2000**, *20*, 634.
(23) Sander, S. L.; Garbett, K. A.; Weil, P. A. *Mol Cell Biol* **2002**, *22*, 6000.
(24) Govind, C. K.; Yoon, S.; Qiu, H.; Govind, S.; Hinnebusch, A. G. *Mol Cell Biol* **2005**, *25*, 5626.
(25) Johnson, S. A. S.; Dubeau, L.; Kawalek, M.; Dervan, A.; Schonthal, A. H.; Dang, C. V.; Johnson, D. L. *Mol Cell Biol* **2003**, *23*, 3043.
(26) Hsu, J.-Y.; Juven-Gershon, T.; Marr, M. T.; Wright, K. J.; Tjian, R.; Kadonaga, J. T. *Genes Dev* **2008**, *22*, 2353.
(27) Bhaumik, S. R.; Green, M. R. *Mol Cell Biol* **2002**, *22*, 7365.
(28) Grant, P. A.; Duggan, L.; Cote, J.; Roberts, S. M.; Brownell, J. E.; Candau, R.; Ohba, R.; Owen-Hughes, T.; Allis, C. D.; WInston, F.; Berger, S. L.; Workman, J. L. *Genes Dev* **1997**, *11*, 1640.
(29) Sterner, D. E.; Grant, P. A.; Roberts, S. M.; Duggan, L. J.; Belotserkovskaya, R.; Pacella, L. A.; Winston, F.; Workman, J. L. *Mol Cell Biol* **1999**, *19*, 86.
(30) Larschan, E.; Winston, F. *Genes Dev* **2001**, *15*, 1946.
(31) Dudley, A. M.; Rougeulle, C.; Winston, F. *Genes Dev* **1999**, *13*, 2940.
(32) Eisenmann, D. M.; Arndt, K. M.; Ricupero, S. L.; Rooney, J. W.; Winston, F. *Genes Dev* **1992**, *6*, 1319.
(33) Lee, T.-I.; Young, R. A. *Genes Dev* **1998**, *12*, 1398.
(34) Bhaumik, S. R.; Green, M. R. *Genes Dev* **2001**, *15*, 1935.

Chapter 4 Identification of novel targets of transcriptional activators using in vivo photocrosslinking and mass spectrometry

4.A.Background[4]

Fundamental to the regulation of eukaryotic gene expression are transcriptional activators, a family of signal responsive proteins that recruit numerous coactivator complexes to the promoter of a gene and assemble the requisite pre-initiation complex (PIC). Regarding these complexes, much effort has been put forth over the past several decades to construct an interaction map detailing the individual protein subunits that serve as the direct binding partners of activators *in vivo*. Biochemical studies have been able to contribute a modest amount of useful information toward this goal; however, in many cases the approaches used have been poorly-suited to accommodate the discovery of activator targets in a high-throughput fashion. Recently, mass spectrometric methods for the analysis of complex protein mixtures have been hailed as a superior route for the discovery of activator interactions. Indeed, several publications in recent years have employed this strategy to examine activator recruitment *in vitro*, with

[4] The work in this Chapter is mostly unpublished data. Contributions to the work in this Chapter are as follows: Dr. Chinmay Majmudar, Dr. Jody Lancia and Yik-Hong Fung were responsible for the extensive optimization of the Bpa-crosslinking/MS protocol (XLMS), an effort that spanned 5 years of work. Summary of many of these early optimizations can be found in the thesis of Dr. J.K. Lancia. Y.H. Fung and Dr. C.Y. Majmudar were additionally responsible for the implementation of cryo-lysis methods to reduce proteolytic degradation of low-abundance transcription factors. All MudPIT analyses were performed by our collaborator, Dr. Sherry Niessen, from the Cravatt Lab at The Scripps Research Institute in La Jolla, California. Dr. Chinmay Majmudar and Amanda Dugan ran several XLMS experiments, most of which focused on using isotopically labeled proteins to enhance signal of crosslinked peptides during MS analysis. Validation of Snf1/AMPK targets from MS experiments were planned and performed by Dr. C.Y. Majmudar, Amanda Dugan, and Rachel Pricer.

particular focus on well-characterized Pol II holoenzyme and coactivator complexes.[1,2] While this work has garnered a deepened appreciation of the complexities surrounding pre-initiation complex assembly, such a powerful platform has yet to be employed to investigate activator recruitment beyond the core set of classical targets. Moving forward, we chose to combine mass spectrometry with *in vivo* photocrosslinking to discover new, direct targets of the prototypical yeast transcriptional activator Gal4. These studies resulted in the capture of several previously unidentified targets of Gal4, a subset of them being in the Snf1 kinase complex, which shares significant homology with the human AMPK complex, a leading drug target candidate for the regulation of type II diabetes.

Transcriptional activators are modular proteins, possessing a DNA binding domain (DBD) that recognizes cognate sequences within the promoter and a transcriptional activation domain (TAD) that mediates the contacts required for recruitment of transcriptional complexes. The associations that an activator makes with a single TAD are numerous and can vary greatly in affinity, ranging from low nanomolar interactions with masking proteins to more moderate affinity interactions with coactivator complexes. One particularly well-characterized system includes the yeast activator Gal4, whose activity is regulated by changes in carbon source availability. In the presence of glucose, Gal4 is tightly repressed by its masking protein Gal80; however, in the presence of galactose, an inducing sugar, repression of Gal4 is relieved, leaving Gal4 free to interact with complexes involved in galactose sensing and catabolism.[3-5] Historically, investigations of Gal4 have utilized this switch-like activation to examine the mechanism of transcriptional regulation by Gal4, focusing primarily on the interactions with Gal80 and classical transcriptional complexes including SAGA, Swi/Snf, Mediator and other components of the Pol II holoenzyme.[3-18] However, little has been done in the way of establishing a map of Gal4 PPIs outside the previously

identified associations with the core transcriptional machinery. This is primarily due to the dearth of methods available to capture a broad range of PPIs in the native cellular environment combined with a limited number of approaches available to analyze complex protein mixtures, particularly those of low-abundance. Methods that additionally allow for protein identification in a high-throughput fashion would facilitate activator PPI discovery. In order to achieve the longstanding goal of building a complete interaction map for transcriptional networks, new approaches, or a combination of approaches, must be utilized.

4.B. Toward activator interaction discovery

As demonstrated in Chapters 2 and 3, photocrosslinking with genetically incorporated amino acids affords the ability to capture both stable, high affinity as well as transient and moderate affinity PPIs in their native cellular environment.[8,10] Briefly, this method relies on an evolved tRNA/tRNA synthetase (RS) pair that work to site-selectively incorporate the photocrosslinking unnatural amino acid (UAA) in response to an amber (UAG) stop codon in the mRNA being translated.[19,20] Work from our group and others has implemented this technology in yeast to incorporate the photo-labile UAA p-benzoyl-L-phenylalanine (Bpa) to capture transcriptional PPIs, notably those between coactivator complexes as well as activator-coactivator interactions.[8,10,21] Investigations with the yeast activator Gal4 demonstrated that Bpa incorporation along the Gal4 TAD occurs with minimal impact on activator binding and function. Furthermore, subsequent crosslinking of these Gal4 constructs in yeast indicated that, as expected, each position tested yielded a multi-protein binding profile, with position 849 appearing to capture protein partners with high efficiency relative to the other positions tested. (Figure 4-1). Since these initial experiments, the identities of these crosslinked complexes have been examined in detail using classical immunodetection techniques, finding important targets within the transcriptional machinery including the Mediator protein Med15, the TATA-associated factor Taf12, the Swi/Snf ATPase Snf2, and the SAGA subunit Tra1,

among others.[5][10] However, while these studies were important in filling in pieces of the activator-coactivator puzzle, this hypothesis driven approach is neither suitable nor practical for discovery mode identification of activator binding partners. Therefore, there is a clear need to access a more high-throughput and unbiased route to discover the identity of the many crosslinked partners that exist.

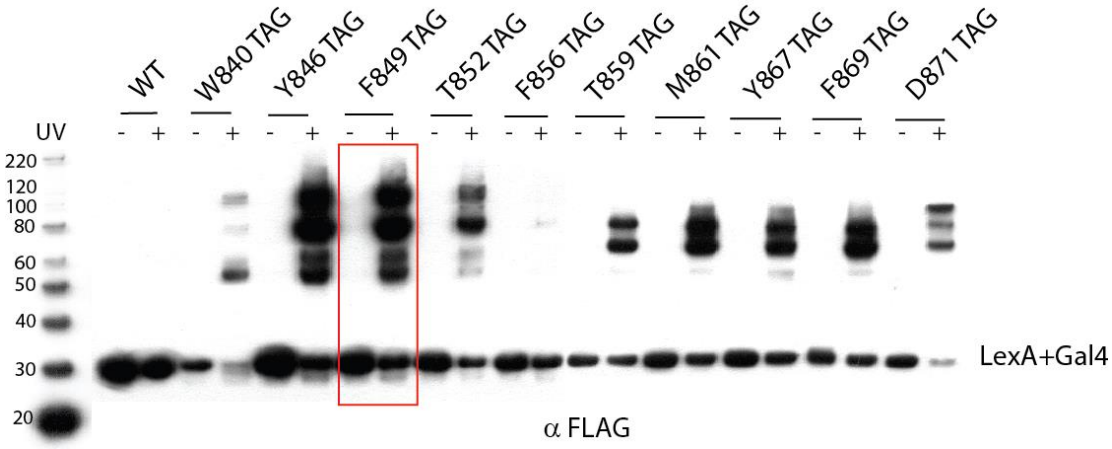

Figure 4-1 . Gal4 exhibits a multiprotein binding profile for each position within the Gal4 TAD incorporating Bpa. Yeast expressing the various Gal4 Bpa mutants were crosslinked under UV light to capture the direct binding partners of Gal4 in live cells. The covalent adducts were immunoprecipitated with an α-LexA antibody and probed with an α-Flag antibody[8]. LexA+Gal4 F849Bpa (highlighted by red square) was selected for MS analysis given its robust crosslinking profile.

4.B.1. Multi-dimensional protein identification technology (MuDPIT) to analyze complex protein mixtures

The mass spectrometry based approach multi-dimensional protein identification technology (MuDPIT), offers several advantages for the discovery of novel interaction partners. First, mass spectrometry is the most sensitive technique that enables discovery of novel targets with no prior information required for target identification. Further, the MuDPIT strategy is designed to effectively separate and analyze complex protein mixtures, thus providing added resolution and sensitivity over other traditional in gel digestion MS approaches, and it has already been used successfully in proteomics studies in *S. cerevisiae*. MudPIT

[5] Gal4 crosslinking in vivo to Taf12 and Tra1 is unpublished data, A.Nwokoye, Y-H. Fung, and C.Y. Majmudar

was first described by the Yates lab where it was used to examine the proteome of yeast, proving sensitive enough to detect membrane proteins and low-abundance transcription factors.[22,23] Whereas traditional proteomics approaches require protein mixtures to first be separated by 2-dimensional electrophoresis and then isolated manually by excising and digesting specific bands from the gel, MudPIT eliminates the need for cumbersome in-gel techniques by digesting the protein mixture directly and then separating the resulting peptides using a high resolution 2D liquid chromatography approach. This entails sample peptides first separated on a strong cation exchange (SCX) column followed by reverse phase (RPLC) chromatographic separation to achieve maximum resolution of peptide species. As the peptides are eluted off the chromatography column, they are automatically injected into the mass spectrophotometer for tandem MS (MS/MS) analysis. During MS/MS runs, the masses of the digested peptides are first measured before being fragmented further using collision-induced dissociation, followed by measurement of the masses of the fragmentation products.[24]

Due to the enormous amount of data this method generates, powerful computational programs are used to assign amino acid sequences and relative abundance of each peptide. Historically, MS-based approaches have demonstrated excellent sequence coverage for only the most abundant peptides in a given sample. However, advocates of MudPIT suggest that this method is sensitive enough to identify even a single peptide in a complex mixture, thereby conferring an advantage over in-gel techniques that would fail to detect such a small quantity of protein. Demonstrating the power of a variation of this approach, Shen and colleagues were able to use an enhanced RPLC-MS/MS technique to identify over 2000 proteins in human plasma whose abundance varied over six orders of magnitude.[25] Given this sensitivity, MudPIT appears to offer the most straightforward route for identifying the complex mixture of binding partners of Gal4 captured in our crosslinking experiments and additionally be the most

powerful approach for analyzing transcriptional complexes that are notoriously low in abundance and often difficult to detect.

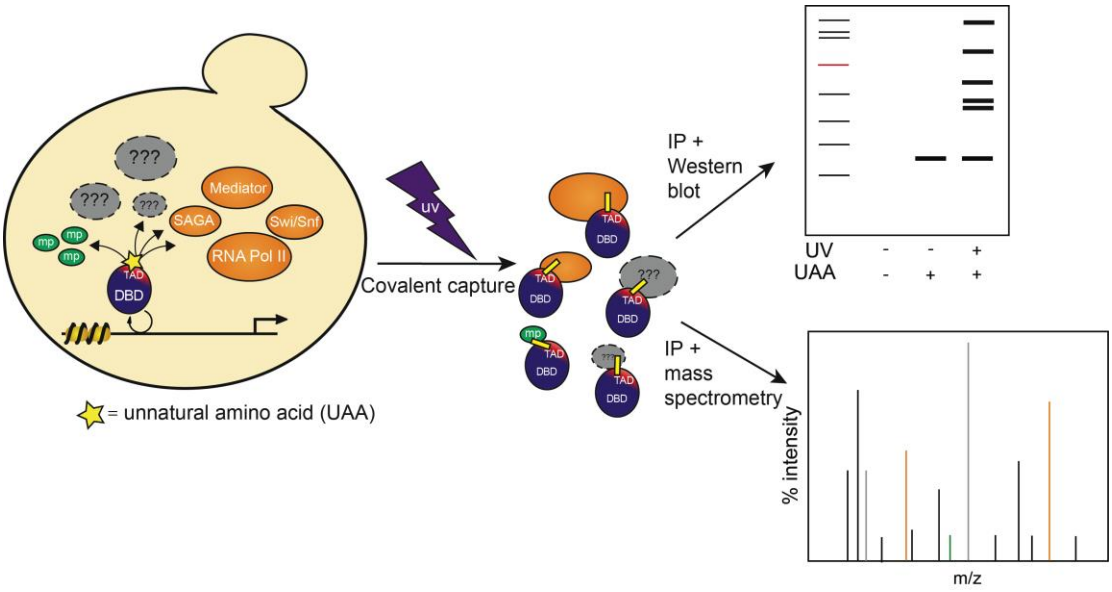

Figure 4-2 Approaches for the identification of activator targets from *in vivo* photocrosslinking studies. Live yeast expressing Bpa containing transcriptional activators (blue and red) are irradiated with UV light to covalently capture the direct interactions of the activator with masking proteins (green), transcriptional machinery proteins (orange) and novel targets (grey). Cells are then lysed and the covalent complexes can be immuno-purified out of cell lysates. These purified products can then be identified via traditional immunodetection techniques such as Western blotting (top) or using mass spectrometric based methods which concurrently allow for the identification of novel binding partners.

4.B.2. Inherent challenges of crosslinking-MS studies

While the challenges facing a combinatorial crosslinking-MS approach have been well-documented, significant advancements have been made since the initial report of MudPIT, making this goal more feasible as more enhancements are introduced. [26-29] One particular challenge associated with studying transcriptional systems, for example, is the relatively low abundance of these proteins in cells, with some proteins as low as 50-100 copies/cell (compare with >200,000 copies/cell for some ribosomal proteins).[30] Although mass spectrometric approaches are often biased toward higher abundance proteins, researchers have found several ways to improve the sensitivity and signal of low abundance peptides, such as using longer, thinner capillaries to more efficiently separate

peptides during RPLC as well as using isotopically labeled proteins to increase the signal to noise ratio of these experiments (SILAC).[31,32] As an additional measure, cryo-lysis procedure was incorporated to prevent the degradation of already low abundance transcriptional complexes. In this approach, crosslinked cells are kept frozen with liquid nitrogen and lysed in a planetary ball mill at around -200 °C, a temperature low enough to severely reduce the activity of proteases that could interfere with these studies. Importantly, a collaboration with Dr. Ben Cravatt, a leading expert in MS-based proteomics research and co-director of the Center for Physiological Proteomics at the Scripps Research Institute, was established.

4.C. Combining in vivo photo-crosslinking and MuDPIT to identify novel binding partners of Gal4 in yeast

To date, the analysis of activator interactions using MudPIT has been mostly limited to experiments examining the *in vitro* assembly of well-studied transcriptional components on naked DNA and chromatinized templates. As discussed in previous chapters, this *in vitro* examination of transcriptional assembly does not necessarily correlate to transcriptional events that occur at the promoter *in vivo*. Therefore, in this chapter, we chose to examine activator interactions in live yeast using *in vivo* photocrosslinking and use the full power of MudPIT to identify novel targets of Gal4 rather than focusing on previously identified transcriptional partners. The advantage of this combined approach is that the activator PPI network is kept in its native environment, allowing for the effects of factors such as cellular localization, nuclear membrane proteins, post-translational modifying enzymes, ubiquitin ligases, and proteosome components to be considered, as all of these have been shown to be important for regulating activator function.[33-40] Thus, the use of Bpa crosslinking in live yeast combined with MS analysis should yield significantly more information regarding the full

spectrum of interactions that activators directly engage in during transcription (Figure 4-2).

4.C.1. Proof of principle: identification of crosslinked Gal4-Gal80 using MudPIT

As an initial experiment to test the feasibility of this approach, Drs. Majmudar and Lancia first examined if the protocol was robust enough to capture the well-characterized Gal4-Gal80 interaction. These studies utilized a LexA+Gal4 F849Bpa-1xflag-6His construct whose two C-terminal tags could be used as purification handles to obtain a clean sample for MS analysis. The sensitivity of this construct to Gal80 repression under glucose conditions was examined as well as its ability to upregulate transcription of the integrated LacZ reporter gene in our yeast strain under inducing galactose conditions, as described in Chapter 2. As shown in Figure 4-3 lower left panel, LexA+Gal4 F849Bpa-1xflag-6His was indeed activated in galactose and repressed in glucose, indicating that this construct is interacting with Gal80 in the conditions of the experiment.

Next, to examine the Gal4-Gal80 interaction with MuDPIT, experiments were scaled up to 1 L cultures of yeast transformed with a LexA+Gal4 F849Bpa-1xflag-6His construct and the tRNA/RS required for Bpa incorporation. Yeast cells were grown in glucose (i.e. conditions that promote repression of Gal4 by Gal80) and then UV-irradiated for 30 minutes on ice. As a control, a batch of cells was also grown in galactose which should relieve repression by Gal80 and theoretically result in a reduced observation of this interaction via MS studies. Each batch of cells was lysed using the specialized cryo-lysis procedure and the crosslinked Gal4 products were purified from yeast lysate using an optimized tandem purification strategy designed to reduce false positives. This purification protocol takes advantage of the fact that the activator is covalently crosslinked to its binding partners, thus allowing stringent purification conditions to be used without any concern of disrupting the activator interactions. Briefly, covalent products were bound to nickel agarose to capture the 6xHis tag on Gal4 and then stringently washed under denaturing conditions to eliminate non-specific

hydrophobic interactions. Following elution off the nickel agarose, the partially purified adducts were added to agarose conjugated with α-Flag antibody to capture the Flag tag on the covalently crosslinked activator products. Subsequently, the bound Gal4 products are eluted off the Flag-agarose through competition with a flag peptide, a mild elution option that leaves unwanted protein bound to the α-flag agarose. Finally, the purified products are buffer exchanged into a MS compatible ammonium bicarbonate buffer, allowing for the simultaneous removal of excess flag peptide, before being trypsin digested, frozen, and sent to Dr. Sherry Niessen at the Center for Physiological Proteomics for MuDPIT analysis (Figure 4-3 top panel). As shown in Figure 4-3 lower right panel, capture of the Gal4-Gal80 interaction was indeed observable by MuDPIT analysis and, furthermore, this interaction was observed to be five times more abundant under glucose conditions than in galactose. The presence of Gal80 in the galactose sample is not completely surprising as it has been suggested that Gal80 does not necessarily dissociate from Gal4 during derepression of galactose inducible genes.[41,42] Together, this data indicates that our experimental setup is transcriptionally responsive in yeast and that Bpa crosslinking and MuDPIT should serve as a reliable platform upon which to examine activator PPIs *in vivo*.

4.C.2. Using in vivo photocrosslinking and MuDPIT to capture and identify Gal4 PPIs

Original experiments with the Gal4 TAD (Figure 4-1) tested several sites of incorporation for Bpa, resulting in the finding that Bpa incorporation and crosslinking efficiency is significantly influenced by the positional context of the crosslinking moiety within the protein domain.[8] Furthermore, given the low abundance of transcription factors, it was critical to use a Gal4 construct that showed a robust and repeatable crosslinking profile as this will enhance the capture of these minimally available proteins in cells. As such, LexA+Gal4

F849Bpa was chosen as the focus of future experiments as this position displays an intense crosslinking profile and should provide a better sample signal during MS analysis. While additional efforts to increase sample signal were attempted by isotopically labeling yeast cultures with either ^{15}N-labeled ammonium sulfate or ^{13}C, ^{15}N labeled Lysine and Leucine (SILAC), this optimization effort proved unsuccessful in our hands. As a result of low levels of incorporation of the isotopically labeled amino acids, this strategy was not pursued further.

Because the initial MuDPIT analysis of Gal4 crosslinked products from 1L cultures showed low signal to noise for many transcription factors, the reaction conditions of these experiments was scaled up by expressing 10-15 L of yeast in an effort to boost the signal of crosslinked peptides. Yeast were harvested at mid-log phase growth and irradiated with UV light for 30 minutes with cooling before freezing the cells and subjecting them to lysis using the cryogenic planetary ball mill. Following purification, samples were sent for analysis. A selected summary of results are summarized in Table 4-1. The proteins listed were selected based on a minimum number of 5 spectral counts, at least 5 fold enrichment in +UV conditions over –UV conditions, and the cellular abundance was factored in.

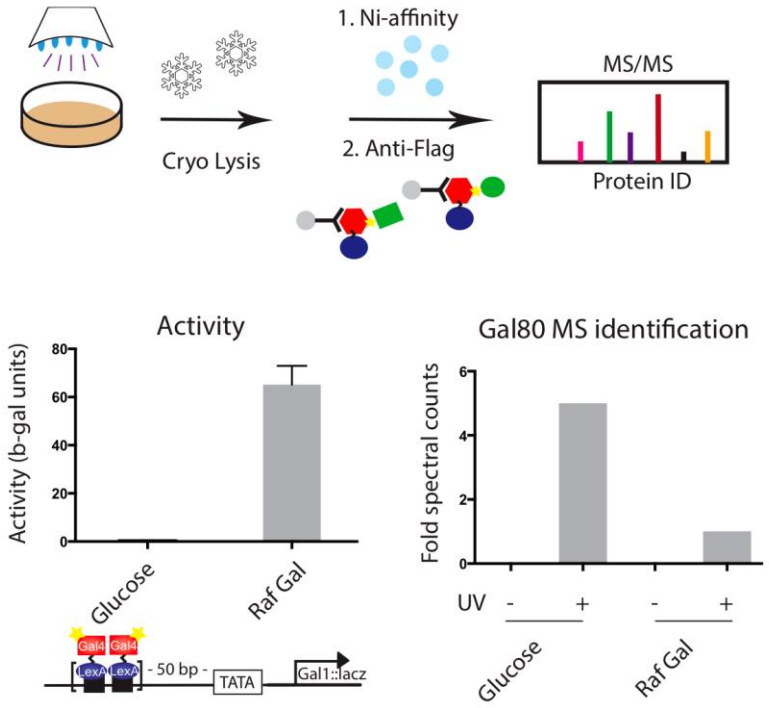

Figure 4-3. Gal4 directly targets Gal80 as identified by a combined *in vivo* photocrossking and MuDPIT approach.(top panel) A 1L yeast culture grown in glucoses was irradiated with UV light to form covalent adducts between LexA+Gal4 F849Bpa-6HIS-flag and its binding partners under repressive conditions. Cells were then lysed using a procedure that uses liquid nitrogen to keep cells frozen during lysis, thus reducing the activity of proteases. Using a tandem purification protocol, the lysates are then incubated with nickel agarose to affinity purify the 6HIS tag on Gal4, followed by incubation with Flag-agarose to immunoprecipitated out the flag tagged activator-adducts. Purified complexed proteins were then digested and submitted to Dr. Sherry Niessen for MuDPIT analysis. (Bottom Left) β-galactosidase activity assays indicate that LexA+Gal4 F849Bpa 6HIS-flag is sensitive to transcriptional repression in the presence of Glucose and can be activated in the presence of the inducing sugar Galactose. (Bottom Right) MuDPIT analysis correlates with activity data in that Gal80 binding is enriched under repressive glucose conditions compared to activating galactose conditions.

Table 4-1 Novel targets of Gal4 captured through *in vivo* photocrosslinking and MuDPIT identification.

Target	Spectral Counts (- UV)	Spectral Counts (+ UV)
Blm10	0	6
Kc11	0	7
Xpo1	0	6
Rrp5	1	9
Mam3	3	17
Gal83	4	21
Trm82	1	5
Far8	1	5
Snf1	696	704

Crosslinking-MS studies revealed several previously unidentified targets belonging to complexes that have some precedence in influencing transcription including proteasome proteins (Blm10), membrane-bound kinases (Kc11), exportin proteins (XpoI), and cell cycle arrest proteins (Far8) as well as protein targets whose functional relevance is more abstract, including tRNA methytransferases (Trm82) and mitochondrial proteins(Mam3). Of the targets identified in this analysis, two of particular note belong to the Snf1 kinase complex, a key regulatory complex involved in glucose/galactose sensing and catabolism. Although both Gal4 and Snf1 have been implicated in regulating galactose inducible genes, little evidence exists to support an interaction *in vitro* and, to our knowledge, no data exists to support this interaction in cells.[43] These data suggest that a key player in galactose catabolism has been left largely overlooked in studies involving Gal4; thus, subsequent validation efforts were focused upon these interactions.

4.D. The Snf1 complex in galactose catabolism

In the response to stressors such as glucose depletion, the Snf1 kinase complex plays a critical role in regulating metabolic response where it is essential for

transcription of genes involved in gluconeogenesis and galactose catabolism. This complex is heterotrimeric in composition, containing a catalytic alpha subunit (Snf1), a regulatory gamma subunit (Snf4), and a third beta subunit that exchanges between Sip1, Sip2, and Gal83 to regulate subcellular localization of the complex.[44] The Snf1 complex is believed to function in glucose repression and galactose activation pathways by acting on the Mig1 repressor, a DNA binding protein shown to localize to genes regulated by Gal4. Under conditions of high glucose, Snf1 is inactivated and remains largely localized in the cytoplasm, leaving Mig1 to localize to the nucleus and recruit the Cyc8-Tup1 complex to DNA to inhibit transcription. In contrast, when glucose is depleted and galactose is introduced, Snf1 becomes activated and localizes to the nucleus where it is believed to phosphorylate Mig1 and induce its export to the cytoplasm, thus allowing transcription to occur (Figure 4-4).[45-48] Additional evidence suggests that the Snf1 complex interacts with components of the Mediator complex to increase activity of the Pol II holoenzyme.[49] Thus, significant evidence exists to support an important role of the Snf1 kinase complex in upregulating transcription of galactose inducible genes. However, while several studies have examined the interaction of this complex with other activators, no studies have been performed to examine how this complex is recruited to promoters controlled by Gal4, such as the GAL1 promoter used in these investigations. Thus, a significant aspect regarding how Gal4 regulates expression of galactose inducible genes remains unexplored.

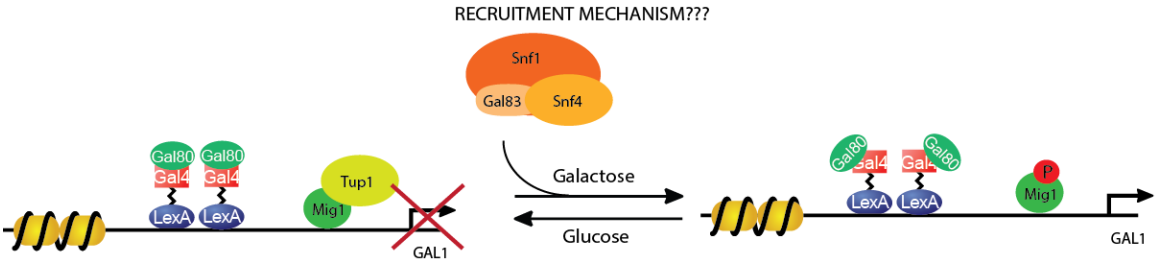

Figure 4-4 The Snf1 kinase complex has been shown to be critical during stress response to low glucose. Under conditions of high glucose, the complex is inactivated and cytoplasmic but shuttles to the nucleus under activating conditions of low glucose. Snf1 complex has implicated as important to transcription of

galactose inducible promoters including the GAL1 promoter used in our studies and data supports a role of this complex in phosphorylating the transcriptional repressor Mig1, thereby shuttling it to the cytoplasm and relieving transcriptional repression. However, the mechanism by which the Snf1 complex is recruited to the GAL1 promoter has never been examined.

During Snf1 activation, the Gal83 isoform has been shown to be the complex that predominantly localizes to the nucleus, with the Sip2 isoform staying largely cytoplasmic and the Sip1 isoform maintaining a vacuolar subcellular localization.[44] Consistent with this model, the MuDPIT data shows a nearly five-fold increase over background for spectral counts supporting a nuclear Gal4-Gal83 interaction when cells are grown in galactose. Furthermore, the remaining two beta subunits, Sip1 and Sip2, are present in the data but not significantly enriched in either condition tested. As shown in Table 1, a significant number of spectral counts is also observed for a Gal4-Snf1 interaction, but the +UV sample is hardly enriched over background signal for Snf1. The ample quantity of Snf1 in the control non-UV sample can be explained by the reactivity of Snf1 with the nickel purification columns, given the presence of a poly-histidine stretch in the amino-terminal portion of Snf1 that binds the Ni-agarose extremely well. Thus, as a result of the tandem purification method, any available form of Snf1, either free or crosslinked with Gal4, was purified from solution. As such, it was critical to test crosslinking of Gal4 to the individual subunits of the Snf1 complex using Western blots and validate the hits from the MS data.

4.D.1. Crosslinking Gal4 to myc-tagged Snf1 complex components

In the absence of antibodies for all subunits of the Snf1 complex, 6x myc-tagged Snf1, Snf4, Sip1, Sip2 and Gal83 constructs were created to allow for immunological detection with an α-myc HRP antibody. These constructs were co-expressed alongside the LexA+Gal4 F849Bpa-1xflag-6His activator and the tRNA/RS required for Bpa incorporation. Live yeast were then irradiated with UV light for 30 minutes and following cell lysis, the lysates were immunoprecipitated with a LexA antibody to isolate the Gal4 crosslinked products. The resulting

Western blot was then probed with an α-myc antibody to detect the presence of a covalently bound myc-tagged Snf1 complex subunit. As expected based on previous localization data, no crosslinked adduct between Gal4 and Sip1, the subunit that is primarily localized to the vacuole in yeast, was observed. Additionally, no crosslinking between Snf4 and Gal4 was observed, suggesting that this subunit does not serve as a target during recruitment of Snf1. However, as discussed in Chapter 2, further tests will have to be conducted in order to verify this negative result. In contrast, a clean crosslinking product for Gal83 and Snf1 was detected (Figure 4-5a, b), indicating a direct interaction between these subunits and Gal4 during transcriptional regulation. A surprising result, Sip2 was also found to be a direct target of Gal4 in live yeast, contrary to localization data demonstrating a largely cytoplasmic residency of this protein (Figure 4-5c). Additionally, because Gal4 and the Snf1 complex have been demonstrated to be inactivated under conditions of high glucose, we hypothesized that this interaction should not occur in yeast grown in this sugar. Indeed, as expected, crosslinking between Gal4 and the Snf1 kinase subunits is significantly diminished in the presence of glucose, further supporting the interdependent role of these proteins in regulating stress response pathways in response to changes in carbon source.

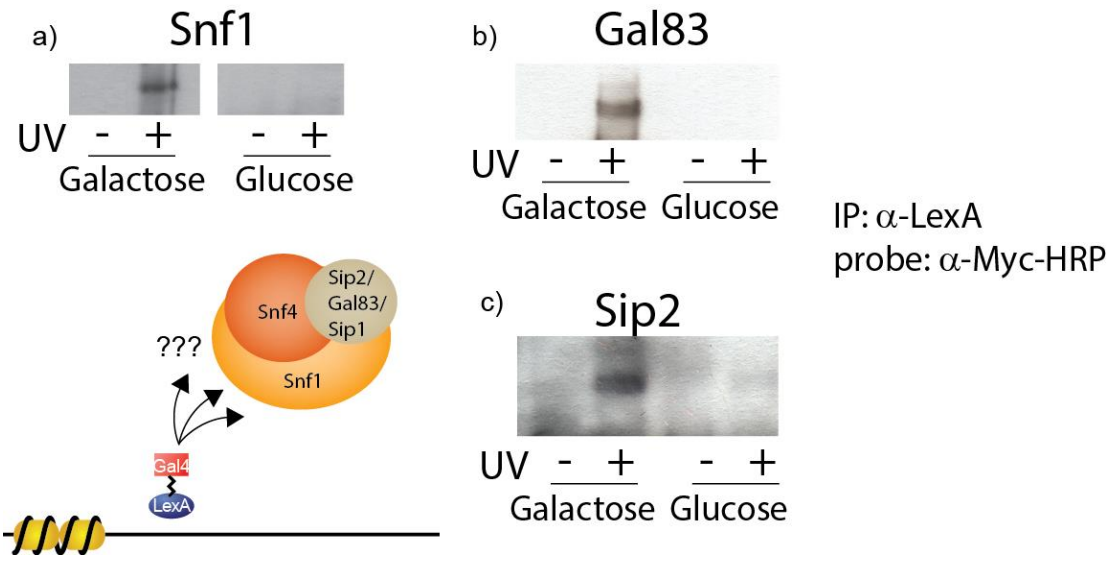

Figure 4-5 *In vivo* photocrosslinking with Bpa captures direct targets of Gal4 within the Snf1 kinase complex. 6xmyc tagged version of each subunit of the Snf1 kinase complex were co-expressed alongside LexA+Gal4 F849Bpa 6HIS-flag and the tRNA/RS required for Bpa incorporation in yeast. Yeast cells were irradiated with UV light and then lysed. The covalent complexes were purified from yeast lysate through immunoprecipitation with a LexA antibody followed by subsequent visualization of crosslinked products via Western blotting with a myc antibody. Under repressive glucose conditions, crosslinking to Snf1 subunits was not observed. However, when cells were grown under glucose limiting conditions, Gal4 crosslinked readily to Snf1, Gal83 and Sip2. Thus in this experimental context, Gal4 makes direct contacts with three Snf1 subunits in live yeast.

4.D.2. Snf1 homology in other eukaryotes

The Snf1 complex is highly conserved among eukaryotes including plant and mammalian systems where it functions in stress-response pathways to maintain energy homeostasis. The mammalian counterpart of the yeast Snf1 kinase complex, AMPK, plays a significant role in maintaining cellular homeostasis by functioning in some cases as a tumor suppressor and additionally as a regulator of energy response. [50,51] Given its important role in the cell, AMPK is currently emerging as a relevant target in the treatment of diseases that exhibit abnormal metabolic profiles including certain cancers as well as diabetes.[52-55] Additionally, AMPK has been suggested to be an important regulator of activator function, involved in the activation of p53 mediated apoptosis under conditions of glucose depletion and additionally responsible for the phosphorylation of numerous

transcriptional targets including critical coactivators such as p300 and histone deacetylases as well as activators such as p53 and FOXO3. [56,57] Thus, the Snf1/AMPK complex likely plays an important role in transcriptional PPI networks and regulating expression of genes required for cell survival.

4.D.3. Examining crosslinking of VP16 and Snf1 kinase

Studies in yeast with the activator Gal4 indicated that recruitment of the Snf1 complex occurs through targeting of the catalytic Snf1 subunit and through an additional contact with either of two regulatory subunits, Gal83 and Sip2. Given the high homology between yeast Snf1 and mammalian AMPK, the interactions of VP16, a viral activator that functions in the HSV-1 infection of mammalian cells, with the Snf1 complex was examined to determine if recruitment also occurs through targeting of the essential Snf1 subunit. LexA+VP16N L444Bpa was transformed into yeast alongside the tRNA/RS required for Bpa incorporation. Following irradiation of live yeast, the cells were lysed and the lysates immunoprecipitated with a Snf1 antibody to pull out endogenous Snf1 and the subsequent Western blot was probed with an α-Flag antibody to detect the presence of a covalent Snf1-VP16 complex. As with Gal4, a direct contact between VP16N and Snf1 was observed, suggesting that this subunit is a common target among these two activators (Figure 4-6). However, further experiments are required to examine if Gal4 and VP16 target overlapping domains on Snf1.

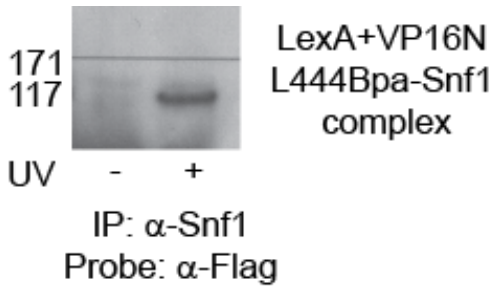

Figure 4-6 VP16 directly contacts endogenous Snf1 in yeast. Yeast expressing LexA+VP16N L444Bpa and the tRNA/RS required for Bpa incorporation were irradiated with UV light to activate Bpa crosslinking.

Following lysis, yeast lysates were immunopurified with a Snf1 antibody and subsequent Western blots were probed with a Flag antibody to detect the presence of a covalently bound VP16. The observation of a VP16-Snf1 crosslinked product indicates that the Snf1 subunit is a shared target between the amphipathic activators Gal4 and VP16.

4.D.4. Future investigations of Gal4-Snf1 kinase interactions

In vivo photocrosslinking combined with MuDPIT analysis revealed that the Snf1 kinase complex is a target of Gal4 in yeast, with further analysis indicating that the Snf1, Gal83 and Sip2 subunits specifically act as handles for recruitment of this complex *in vivo*. The Snf1 complex has been suggested to phosphorylate transcriptional activators and additionally associate with the Mediator complex thus establishing its importance in transcriptional pathways. However, it remains unclear whether the Gal4-Snf1 complex interactions identified in this study occur in the context of a free or DNA-bound activator. As such, investigations examining these interactions in the context of a DNA bound activator using tandem reversible and irreversible crosslinking (TRIC) are currently underway. Furthermore, although some transcriptional activators have been identified as substrates of Snf1/AMPK, it is unknown whether this is also true for Gal4. Snf1 is a serine/threonine kinase known to phosphorylate substrates containing a consensus sequence of ϕXRXX**S**/**T**XXXϕ.[58] Interestingly, a portion of the Gal4 TAD contains a loose variation of this sequence (MFNTT**T**MDDV), suggesting that Gal4 may also be a substrate of the Snf1 complex in addition to the other transcriptional complexes with which it has been shown to associate. As such, the nature of the Gal4-Snf1 interaction will be investigated further, focusing in particular on variations in the phosphorylation state of Gal4 as a function of Snf1 activity. Finally, we are poised to investigate additional interactions identified through MuDPIT analysis of Gal4 crosslinked products, including the interaction with the proteasomal component Blm10, a complex whose regulation of activators is believed to have a significant impact on activator function. In these future experiments, it will likely be beneficial to

evaluate the utility of other affinity tags, such as streptavidin or GST, to avoid contamination problems that were observed during Ni-affinity purification of the 6xHis tag used in these studies.

4.E. Lessons learned from in vivo crosslinking and MuDPIT analysis

Although the combined *in vivo* photocrosslinking and MuDPIT approach was successful in identifying several novel, direct interactions of the yeast activator Gal4, a significant limitation to this approach became increasingly evident with each batch of sample analyzed. Specifically, the bias of mass spec based approaches such as MuDPIT in identifying high-abundance peptides proved to be a major hindrance in obtaining a well-defined interaction profile of Gal4. For example, the most prominent results from these studies (i.e. the proteins with the highest spectral counts) included highly abundant cellular proteins such as molecular chaperones and ribosomal proteins. In accordance with this bias, the more abundant Gal83 (3,500 copies/cell) was enriched in the MS data but Sip2 (300 copies/cell) was overlooked, despite both being direct binding partners of Gal4. This limitation has been present since the invention of MuDPIT where, of the 1,484 proteins observed from the yeast proteome, only 19% of the proteins present at <5000 copies/cell were observed, compared to a nearly 90% sequence coverage for proteins present at >50,000 copies/cell.[22] Furthermore, similar to our findings, isotopic labeling does not alleviate this bias, as follow up studies examining changes in the abundance of 688 yeast proteins again demonstrated a poor sequence coverage for those present at <5000 copies/cell.[59] As such, it was not surprising that MuDPIT analysis of crosslinked Gal4 products failed to identify transcriptional proteins that have been identified as direct partners of Gal4 through previous *in vivo* crosslinking investigations including Med15 (606 copies/cell), Taf12 (930 copies/cell), and Snf2 (217 copies/cell). Thus, while the work in this chapter demonstrates that *in vivo* photocrosslinking and mass spec analysis was critical for the identification of the Snf1 kinase complex as a novel target of Gal4, the inability of this method to

identify low-abundance, validated targets of Gal4 indicates that significant advances must be made in order to use this approach to its fullest potential. Furthermore, it demonstrates that follow-up *in vivo* crosslinking experiments are absolutely essential to validate hits from MuDPIT analysis.

4.F. Conclusions

In sum, the data in this chapter represents a significant step toward the development of complete and accurate interaction maps of transcriptional activators, long an elusive goal. A combination of *in vivo* photocrosslinking with the unnatural amino acid Bpa followed by covalent adduct analysis using the powerful mass spectrometric method MudPIT resulted in the identification of several subunits of the Snf1 kinase complex as novel, direct targets of the yeast activator Gal4. Although several attempts to enhance the signal of low abundance peptides were made, further optimizations will be required to reduce the bias of MuDPIT toward the identification of higher abundance peptides and thus use this method to its fullest potential. In the future, we believe this approach will be effective in establishing a more complete picture of the interactions that comprise transcriptional networks as well as other PPI systems that utilize a broad range of contact types.

4.G. Experimental

Yeast Strain LS41 [JPY9::pZZ41, *Mata his3*Δ*200 leu2*Δ*1 trp1*Δ*63 ura3-52 lys2*Δ*385 gal4* URA::pZZ41] was used for the crosslinking experiments. Bpa 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).

Plasmid name	Function
pLexAVP16N 444TAG	Expresses LexA(1-202)+VP16 (413-456)+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
pMyc Gal80	Expresses full length Gal80 fused to c-myc tag on N terminus of protein;
pLexA+Gal4 F849Bpa 6His-Flag	Expresses LexA(1-202)+Gal4(840-881)+6xHIS and 1xFLAG tag with a TAG replacing the codon of the

	existing amino acid
p6XMyc-Snf1	Expresses full length Snf1 fused to a 6x c-myc tag on C terminus of protein;
p6XMyc-Gal83	Expresses full length Gal83 fused to a 6x c-myc tag on C terminus of protein;
p6XMyc-Sip2	Expresses full length Sip2 fused to a 6x c-myc tag on C terminus of protein;
p6XMyc-Snf4	Expresses full length Snf4 fused to a 6x c-myc tag on C terminus of protein;
p6XMyc-Sip1	Expresses full length Sip1 fused to a 6x c-myc tag on C terminus of protein;

4.G.1. In vivo cross-linking to capture the Gal4-Gal80 interaction

To perform in vivo cross-linking, an individual colony of pLexA+Gal4 F849TAG-6HIS-flag was grown in 10 mL SC media containing 2% Raffinose 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 1L cultures of SC media containing 2% Glucose. Control samples were grown in 2% Raffinose and 2% Galactose. For pBpa incorporation, 10 mL 100 mM pBpa dissolved in 1M NaOH and 10 mL 1M HCl were added to the above cultures. The cultures were incubated overnight at 30 °C with agitation to an OD660 of ~0.8. When cultures reached the appropriate OD_{660}, the cells were spun down by centrifuging at 6000 rcf at 4^{o}C for 15min. following which the cell pellets were washed with SC media lacking histidine and tryptophan. The cell pellets were resuspended in 2mL SC media lacking histidine and tryptophan + either 2% Raffinose and 2% Galactose or 2% Glucose and transferred to 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.

Cryolysis procedure

Cryolysis was performed using the Retsch planetary ball mill PM 200 in the Skiniotis Lab (LSI, University of Michigan). Chambers and balls were submerged in liquid nitrogen to cool equipment down. Frozen yeast cells were crushed in a chilled mortar and pestle (chilled in liquid nitrogen) and were then added to the chilled chambers. 3 balls/chamber were then added and then the lid was clamped in place. Liquid nitrogen was ladled over the closed chambers repeatedly to keep cells and equipment chilled. Balanced chambers were loaded into the cryomill machine and clamped into the machine. Machine was run for two minutes at 500 r.p.m. When done, the chambers were removed and again chilled with liquid nitrogen and then run for another cycle in the PM200 machine. This procedure was repeated for a total of 3 cycles. Chambers were removed from machine after final spin and the frozen lysed yeast were recovered from the chambers and placed in a 50 mL Falcon tube on dry ice.

Affinity purification

Lysed yeast were allowed to thaw in a cold water bath and then centrifuged for 30 minutes at 4 deg C, 9299 rcf (Beckman Coulter Allegra X-22R, fixed rotor). While cells were spinning, 300-500 μL Ni-agarose bead slurry was washed three times with 1 mL chilled PBS. After centrifugation is complete, supernatant from lysed yeast is transferred to a clean 50 mL Falcon tube and the Ni-agarose beads (resuspended in 1 mL chilled PBS) are added to the supernatant. Beads and supernatant are incubated for 1 hour on rotating carousel in 4 deg C cold room. After incubation, beads are spun down (low rpm) and the supernatant is pipetted off. The Ni-agarose is resuspended in the remaining supernatant, 5 mL chilled Nickel Wash Buffer (100 mM PBS pH 7.0, 150 mM NaCl, 10% glycerol, 0.5% NP-40 + 2 protease inhibitor tablets (Roche Complete minis, EDTA free) and 10 mM betamercaptoethanol) was added to beads and then transferred to a 15 mL falcon tube. Beads in buffer were centrifuged at 2500 rpm, 2 minutes, 4 deg C (Beckman Coulter Allegra X-22R, bucket rotor) and the supernatant decantant. This procedure was repeated 4 more times for a total of five 5mL washes. The beads were then washed twice with 5 mL cold 100 mM PBS. PBS was decanted and the beads were resuspended in ~500 μL 100 mM PBS and transferred to a chilled 1.75 Eppendorf flip cap tube. Residual beads on the 15 mL tube were washed with about 500 μL 100 mM PBS and transferred to Eppendorf tube. Beads were spun briefly at low rpm (Eppendorf 5417C microfuge). The PBS was pipetted off and discarded.

Beads were resuspended in 500 μL chilled Nickel Elution Buffer (100 mM PBS, pH 7.0, 500 mM Imidazole, 0.5% Tween-20) and mixed gently for 1 minute by inversion and then recentrifuged. The supernatant (eluant) was pipetted off and collected in a separate, clean 1.75 mL Eppendorf tube. Another 500 μL chilled Nickel Elution Buffer was added to the beads and mixed by inversion for 10 minutes, 4 deg C in a cold room. During the second elution, a 10K concentrator

was washed with 100 mM PBS. The first elution was centrifuged to remove any lingering Ni-agarose and the eluant was pipetted out and added to the concentrator. Concentrator was spun at 15,000xG for 5 minutes at 4 deg C. Elutions were repeated for a total of 3 elutions, each elution added to the concentrator after centrifugation to remove residual Ni-agarose. After concentrating the three elutions, the volume should be close to 100 μL. 500 μL of 100 mM chilled PBS was added to the protein and then run through the concentrator. This was repeated for a total of 3 PBS buffer exchanges to get a final volume of 100-150 μL of protein.

400 μL of Flag-agarose was washed three times with 1 mL Flag Wash Buffer (100 mM PBS, 150 mM NaCl, 0.5% NP-40). Concentrated sample was added to the Flag-agarose (Sigma, M2). The concentrator was washed with a small volume of 100 mM PBS and added to the Flag-agarose incubation (2 hours, 4 deg C, with rotation on carousel). Flag-agarose was then washed ten times with 1 mL/wash Flag Wash Buffer, spinning between washes (5000 rpm, 30 sec, Eppendorf 5417C). Agarose was washed twice with 100 mM PBS to dilute out the NP-40. 1 mL of Flag Elution Buffer (100 mL PBS, pH 7.0, 0.1% NP-40, 350 mM NaCl) containing 10 mg/mL flag peptide was thawed from -80 deg C. 333 μL Flag elution buffer was added to washed Flag-agarose and incubated for 30 minutes, 4 deg C, with rotation. Beads were briefly centrifuged and the elution transferred to a new 1.75 mL Eppendorf tube. Elution was quick frozen in liquid nitrogen and stored in -80 deg C freezer. Second elution was performed in the same fashion and the final 333 μL was added to beads for a third elution that was allowed to run overnight. The next morning, the two frozen elutions were thawed and spun down to remove residual Flag-agarose. Elutions were transferred to new tubes and then centrifuged again to remove any chance of residual beads remaining. These first two elutions, cleared of all beads, were then concentrated in a 10K concentrator at 15,000 xG until down to 100 μL. The overnight elution is then added and centrifuged down to 100 μL. 500 μL of Ammonium bicarbonate

(50 mM) was then added to the concentrator for a total of 4 buffer exchange cycles with the protein. Concentrated sample was then transferred to a clean Eppendorf tube. Concentrator was washed with a little ammonium bicarbonate and then added to the sample. Sample was quick frozen in liquid nitrogen, packed in dry ice, and shipped to the Cravatt Lab overnight.

4.G.2. MuDPIT analysis

All MuDPIT analyses were performed by Dr. Sherry Niessen at the Center for Physiological Proteomics in La Jolla, California. Crosslinked products were denatured by resuspending in an equal volume of 8M urea, 50 mM Tris (pH 8.0). 10-20 mg of protein was then reduced with 10 mM Tris(2-carboxyethyl) phosphine HCl (TCEP, Sigma) for 30 minutes at room temperature. The proteins were then alkylated with fresh 12.5 M iodoacetamide (IAA; Sigma) and the concentration of urea was reduced to 2 M by adding 50 mM Tris (pH 8.0). Crosslinked proteins were then digested overnight with 1 mM $CaCl_2$ and 0.2 mg Trypsin and incubated at 38 deg C. Digested peptides were then acidified in 5% formic acid and centrifuged at 17,000 xg for 15 minutes. Half of the digested mixture was pressure loaded onto a biphasic strong cation exchange/reverse phase capillary column and separated by 2D liquid chromatography and tandem MS using an 11 step gradient on an LTQ-Orbitrap LX hybrid mass spectrometer. MS spectra were acquired in centroid mode, with a mass range of 400-1,800 in the Orbitrap analyzer with resolution set at 30,000 followed by 7 MS/MS scans in the ion trap. All MS/MS spectra were collected using a normalized collision energy of 35% and an isolation window of 2 Da. One microscan was applied for all experiments in the Orbitrap or LTQ. Spray voltage was set to 2.50 kV, and the flow rate through the column was 0.20 uL/min.

4.G.3. Crosslinking with myc-Snf1 subunits.

For crosslinking studies with mycSnf1 subunits, the procedure was identical except that cells were scaled down to 100 mL cultures in SC media lacking

histidine, leucine, and tryptophan. 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) (or in the case of VP16-Snf1, lysates were immunoprecipitated with a Snf1 antibody – Santa Cruz, sc-15621) 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 ~ 8 uL prewashed protein G agarose beads (Millipore) at 4 ^{o}C. After immunoprecipitation, the beads were washed 6 times 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 ^{o}C until elution. The crosslinked sample was eluted from the beads by heating at 95 ^{o}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 HRP, Santa Cruz Biotechnology, Santa Cruz, CA).

4.G.4. β.Galactosidase assays

To evaluate the ability of each LexA+Gal4 F849Bpa-6HIS-flag to activate transcription in the presence or absence of glucose, saturated cultures (SC media + 2% Raffinose) of each mutant were used to inoculate 5 mL SC media containing either 2%Glucose or 2% Raffinose & 2% Galactose but lacking histidine and tryptophan for selection. The cells were grown to an OD of 0.8-1 and harvested. The activity of each construct was monitored using β-galactosidase assays as previously described.[60]

4.F. References- Chapter 4

(1) Chen, X.-F.; Lehmann, L.; Lin, J. J.; Vashisht, A.; Schmidt, R.; Ferrari, R.; Huang, C.; McKee, R.; Mosley, A.; Plath, K.; Kurdistani, S. K.; Wohlschlegel, J.; Carey, M. *Cell Rep* **2012**, *2*, 1061.
(2) Sikorski, T. W.; Joo, Y. J.; Ficarro, S. B.; Askenazi, M.; Buratowski, S.; Marto, J. A. *J Biol Chem* **2012** *287*, 35397.
(3) Ma, J.; Ptashne, M. *Cell* **1987**, *50*, 137.
(4) Jiang, F.; Frey, B. R.; Evans, M. L.; Friel, J. C.; Hopper, J. E. *Society* **2009**, *29*, 5604.
(5) Ansari, A. Z.; Reece, R. J.; Ptashne, M. *Proc Natl Acad Sci U S A* **1998**, *95*, 13543.
(6) Bhaumik, S. R.; Green, M. R. *Genes Dev* **2001**, *15*, 1935.
(7) Bhaumik, S. R.; Green, M. R. *Methods in enzymology* **2003**, *370*, 445.
(8) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J Am Chem Soc* **2009**, *131*, 14240.
(9) Majmudar, C. Y.; Wang, B.; Lum, J. K.; Hakansson, K.; Mapp, A. K. *Angew Chem Int Ed Engl* **2009**, *48*, 7021.
(10) Krishnamurthy, M.; Dugan, A.; Nwokoye, A.; Fung, Y.-H.; Lancia, J. K.; Majmudar, C. Y.; Mapp, A. K. *ACS Chem Biol* **2011**, *6*, 1321.
(11) Bryant, G. O.; Prabhu, V.; Floer, M.; Wang, X.; Spagna, D.; Schreiber, D.; Ptashne, M. *PLoS Biology* **2008**, *6*.
(12) Melcher, K.; Johnston, S. A. *Mol Cell Biol* **1995**, *15*, 2839.
(13) Wu, Y.; Reece, R. J.; Ptashne, M. *EMBO J* **1996**, *15*, 3951.
(14) Jeong, C. J.; Yang, S. H.; Xie, Y.; Zhang, L.; Johnston, S. a.; Kodadek, T. *Biochemistry* **2001**, *40*, 9421.
(15) Yudkovsky, N.; Logie, C.; Hahn, S.; Peterson, C. L. *Genes Dev* **1999**, *13*, 2369.
(16) Brown, C. E.; Howe, L.; Sousa, K.; Alley, S. C.; Carrozza, M. J.; Tan, S.; Workman, J. L. *Science* **2001**, *292*, 2333.
(17) Reeves, W. M.; Hahn, S. *Mol Cell Biol* **2005**, *25*, 9092.
(18) Fishburn, J.; Mohibullah, N.; Hahn, S. *Mol cell* **2005**, *18*, 369.
(19) Wang, Q.; Parrish, A. R.; Wang, L. *Chem Biol* **2009**, *16*, 323.
(20) Wang, Q.; Wang, L. *Methods Mol Biol* **2012**, *794*, 199.
(21) Mohibullah, N.; Hahn, S. *Genes Dev* **2008**, *22*, 2994.
(22) Washburn, M. P.; Wolters, D.; Yates, J. R. r. *Nat Biotechnol* **2001**, *19*, 242.
(23) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. r. *Nat Biotechnol* **1999**, *17*, 676.
(24) Delahunty, C. M.; Yates, J. R.R. *Biotechniques* **2007**, *43*, 563.
(25) Shen, Y.; Zhang, R.; Moore, R. J.; Kim, J.; Metz, T. O.; Hixson, K. K.; Zhao, R.; Livesay, E. A.; Udseth, H. R.; Smith, R. D. *Anal Chem* **2005**, *77*, 3090.
(26) Sinz, A. *Mass Spec Rev* **2006**, *25*, 663.
(27) Petrotchenko, E. V.; Borchers, C. H. *Mass Spectrom Rev* **2010**, *29*, 862.
(28) Tabb, D. L. *Nat Methods* **2012**, *9*, 879.
(29) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. *Mol Cell Proteomics* **2010**, *9*, 1634.

(30) Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. *Nature* **2003**, *425*, 737.
(31) Guzzetta, A. W.; Chien, A. S. *Proteome Res* **2005**, *4*, 2412.
(32) Yates, J. R. r.; Ruse, C. I.; Nakorchevsky, A. *Annu. Rev. Biomed. Eng.* **2009**, *11*, 49.
(33) Igarishi, D.; Ishida, S.; Fukazawa, J.; Takahashi, Y. *Plant Cell* **2001**, *13*, 2483.
(34) Colomina, N.; Liu, Y.; Aldea, M.; Gari, E. *Mol Cell Biol* **2003**, *23*, 7415.
(35) Brickner, D. G.; Cajigas, I.; Fondufe-Mittendorf, Y.; Ahmed, S.; Lee, P.-C.; Widom, J.; Brickner, J. H. *PLoS Biol* **2007**, *5*, e81.
(36) Menon, B. B.; Sarma, N. J.; Pasula, S.; Deminoff, S. J.; Willis, K. A.; Barbara, K. E.; Andrews, B.; Santangelo, G. M. *Proc Nat Acad Sci, USA* **2005**, *102*, 5749.
(37) Brown, C. R.; Kennedy, C. J.; Delmar, V. A.; Forbes, D. J.; Silver, P. A. *Genes & Dev.* **2008**, *22*, 627.
(38) Geng, F.; Wenzel, S.; Tansey, W. P. *Annu. Rev. Biochem.* **2012**, *27*, 177.
(39) Musselman, C. A.; Lalonde, M.-E.; Cote, J.; Kutateladze, T. G. *Nat Struct Mol Biol* **2012**, *19*, 1218.
(40) Appella, E.; Anderson, C. W. *Eur. J. Biochem.* **2001**, *268*, 2764.
(41) Leuther, K. K.; Johnston, S. A. *Science* **1992**, *256*, 1333.
(42) Bhaumik, S. R.; Raha, T.; Aiello, D. P.; Green, M. R. *Genes Dev* **2004**, *18*, 333.
(43) Lo, W.-S.; Gamache, E. R.; Henry, K. W.; Yang, D.; Pillus, L.; Berger, S. L. *EMBO J* **2005**, *24*, 997.
(44) Vincent, O.; Townley, R.; Kuchin, S.; Carlson, M. *Genes Dev.* **2001**, *15*, 1104.
(45) Gancedo, J. M. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 334.
(46) DeVit, M. J.; Waddle, J. A.; Johnston, M. *Mol Biol Cell* **1997**, *8*, 1603.
(47) Ostling, J.; Ronne, H. *Eur. J. Biochem.* **1998**, *252*, 162.
(48) Treitel, M. A.; Kuchin, S.; Carlson, M. *Mol Cell Biol* **1998**, *18*, 6273.
(49) Kuchin, S.; Treich, I.; Carlson, M. *Proc Nat Acad Sci, USA* **2000**, *97*, 7916.
(50) Faubert, B.; Boily, G.; Izreig, S.; Griss, T.; Samborska, B.; Dong, Z.; Dupuy, F.; Chambers, C.; Fuerth, B. J.; Viollet, B.; Mamer, O. A.; Avizonis, D.; DeBerardinis, R. J.; Siegel, P. M.; Jones, R. G. *Cell Metab* **2013**, *17*, 113.
(51) Hardie, D. G. *Int J Obes* **2008**, *32*, S7.
(52) Wang, W.; Guan, K. L. *Acta Physiol* **2009**, *196*, 55.
(53) Liang, J.; Mills, G. B. *Cancer Res* **2013**, *73*, 2929.
(54) Viollet, B.; Lantier, L.; Devin-Leclerc, J.; Hebrard, S.; Amouyal, C.; Mounier, R.; Foretz, M.; Andreelli, F. *Front Biosci* **2009**, *14*, 3380.
(55) Zhang, B. B.; Zhou, G.; Li, C. *Cell Metab* **2009**, *9*, 407.
(56) Okoshi, R.; Ozaki, T.; Yamamoto, H.; Ando, K.; Koida, N.; Ono, S.; Koda, T.; Kamijo, T.; Nakagawara, A.; Kizaki, H. *J Biol Chem* **2008**, *283*, 3979.
(57) Mihaylova, M. M.; Shaw, R. J. *Nat Cell Biol* **2011**, *13*, 1016.

(58) Dale, S.; Wilson, W. A.; Edelman, A. M.; Hardie, D. G. *FEBS Lett* **1995**, *361*, 191.
(59) Washburn, M. P.; Koller, A.; Oshiro, G.; Ulaszek, R. R.; Plouffe, D.; Deciu, C.; Winzeler, E.; Yates, J. R. r. *Proc Nat Acad Sci, USA* **2003**, *100*, 3107.
(60) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J Am Chem Soc* **2009**, *131*, 14240.

Chapter 5 Conclusions and Future Directions

5.A. Introduction

Protein-protein interaction (PPI) networks govern nearly every essential regulatory process in the cell including protein synthesis and folding, cell signaling, and transcriptional regulation. These networks are complex, comprised of thousands of contacts that vary in affinities, interaction area, and lifetimes. Of particular importance to biological systems are networks that heavily utilize transient PPIs of moderate to low affinity, as these are essential to the proper function of regulatory cellular processes. Transient interactions often enable a protein to engage in multiple contacts, and this can occur through re-use of the same interface. Aberrations within PPI networks are a hallmark of various disease states; therefore, there is great interest in designing compounds to target PPIs in order to learn more about how they function and, in the long term, restore normal cell phenotype. However, this effort has been hindered by a lack of structural and mechanistic data, particularly with regards to transient PPIs that have historically been the most challenging to study.

The challenges associated with studying transient PPI networks are particularly apparent in the process of transcriptional initiation. At the heart of this process are transcriptional activators, a family of signal responsive proteins that localize to gene promoters and use a single transcriptional activation domain (TAD) to recruit numerous multi-protein coactivator complexes required for up regulation of gene expression. Activators are dynamically engaged with the transcriptional machinery through a coordinated series of contacts that range greatly in affinities

and interaction areas. Although co-localization studies have identified the complexes that are recruited by activators during transcription, the direct targets within these complexes in cells remains largely unknown. Furthermore, due to the limitations of methods such as co-immunoprecipitation in detecting transient, lower affinity interactions, many binding partners of activators may remain unidentified. As a result, the interaction map for transcriptional PPI networks is currently incomplete and is lacking critical information necessary to aid the rational design of small molecules to transiently perturb this system reveal key insights into how it functions.

As demonstrated in this dissertation, *in vivo* photocrosslinking with a genetically incorporated photo-labile amino acid is a powerful approach to capture the direct binding partners of transcriptional activators in living cells. This includes transient, moderate affinity interactions that are typically challenging to study. Furthermore, the use of a combination of *in vivo* chemical capture and mass spectrometry revealed several novel binding partners of a yeast activator, indicating the power of this combined approach in building a more complete interaction map for transcriptional PPI networks. Finally, a novel method was developed to examine the direct interactions of DNA bound transcriptional activators *in vivo*, thus affording the ability to examine the most functionally relevant interactions in a cellular context. In all, the work performed in this dissertation not only contributes to advancing the longstanding goal of creating complete interaction maps for activators, but it also indicates that these approaches can be used more broadly in the examination of other regulatory PPI networks that rely on transient interactions to function.

5.B. In vivo photocrosslinking is a powerful method for the capture of a broad range of activator PPIs in yeast

A key finding from this work is that covalent chemical capture using genetically encoded p-benzoyl-l-phenylalanine (Bpa) is a powerful tool to capture the direct, cellular binding partners of transcriptional activators. The utility of this

methodology is demonstrated in Chapter 2 by the *in vivo* covalent chemical capture of the moderate affinity interaction between VP16 and Med15 (K_D~ 0.1-10 μM) as well as capture of a higher affinity contact with TBP (K_D~33 nM), an interaction whose occurrence in cells has been contested prior to this study.[1-3] Furthermore, capture of the direct targets of VP16 within the low abundance Swi/Snf complex demonstrates the power of this approach in studying transcriptional PPI networks whose constituent interactions are often mediated by low-abundance proteins. The data in Chapters 2 and 3 also demonstrate that a negative result must be interpreted cautiously, as the lack of a crosslinking product could arise from a variety of factors including poor positioning of the crosslinking amino acid or an inherent reactivity that is poorly suited for a given PPI interface. Furthermore, the use of a detection system that lacks the sensitivity to observe a crosslinked product may be a significant problem when examining very small quantities of crosslinked proteins.[4,5] Finally, when combined with mass spectrometric based methods designed to identify complex protein mixtures, *in vivo* photocrosslinking is tremendously useful in identifying the novel targets of transcriptional activators, as demonstrated in Chapter 4 by the discovery of Snf1, Gal83, and Sip2 as direct targets of the transcriptional activator Gal4.

5.C.Recruitment of multi-protein complexes involves the targeting of multiple subunits

At the core of transcriptional PPI networks are multi-protein complexes that are composed of at least one enzymatic subunit, several scaffolding proteins, and any number of regulatory subunits that associate either stably or transiently with the complex to modulate complex function and localization. The experiments of Chapter 2 examined the recruitment of the Swi/Snf chromatin remodeling complex whose core enzymatic subunit, the Snf2 ATPase, associates with at least 9-11 other proteins in the cell. *In vivo* photocrosslinking with VP16 showed that recruitment of this complex occurs through a direct interaction with the

enzymatic Snf2 subunit and additional targeting of Snf5 and Snf6, three proteins that have been shown to be functionally interdependent.[6] Furthermore, in mammalian systems, INI1/Snf5 has been demonstrated to enhance the activity of BRG1/BRM (Snf2), thus indicating the role of this subunit in modulating complex function.[7] In Chapter 4, we examined the interactions of Gal4 with the Snf1 kinase complex and found that Gal4 targets the catalytic subunit Snf1 in addition to two regulatory proteins, Gal83 and Sip2, that associate transiently with the complex to fine tune subcellular localization.[8] Thus, it appears that activators target multiple subunits of coactivator complexes and furthermore that the enzymatic subunit is often, but not always, a shared target of amphipathic activators. As such, the results of the studies in this thesis support a model wherein activators target the minimal functional component of the complex while additionally targeting associated factors, perhaps as a means to recruit specific complex isoforms whose subunit association confers differential activity profiles.

5.D. Future Directions

5.D.1. Moving into an endogenous system

The work outlined in this thesis contributed several novel findings including in Chapter 2 the first demonstration of the utility of *in vivo* photocrosslinking in capturing a transient, moderate-affinity interaction. In Chapter 3, the use of TRIC to examine the VP16-TBP interaction was the first demonstration of the capture and visualization of the direct interactions of DNA bound activators. And finally, the work in Chapter 4 was the first study of its kind in using *in vivo* photocrosslinking combined with MuDPIT to identify novel interaction partners of the activator Gal4.

While the studies with model systems in this thesis (i.e. LexA-VP16 in yeast) yielded a significant amount of data, the long-term goal is to examine activator function in a context that more closely resembles its natural state. Specifically, the future goal would be to site specifically introduce an amber stop codon into

the endogenous gene, for example Gal4, so that Bpa containing activator is expressed at levels more comparable to that found in nature. However, due to relatively poor incorporation rates of Bpa in activators expressed from plasmids (about 20%), significant optimization in Bpa mutant expression will be required before this can be achieved. If this can be achieved, however, it will represent a significant step toward understanding how natural activators function to upregulate transcription across the genome.

Thinking more broadly in the context of entire transcriptional PPI networks, this technology can also be employed to examine the inter- and intra-molecular interactions of other transcriptionally relevant complexes. For example, combinatorial complex assembly is a popular strategy employed in transcriptional systems to control complex localization and function. Perhaps *in vivo* photocrosslinking can be used to investigate the identity and positioning of exchangeable subunits within these complexes and, furthermore, identify changes in the interactions with other complexes that come as a result of this exchange. Different isoforms of the Pol II holoenzyme, general transcription factor complexes, and chromatin modifying complexes, for example, have all been reported; however, the contact fluctuations within the PPI network that come as a result of these differentially composed complexes remains to be seen. To address this, a combined approach such as that used in Chapter 4 would be useful given the ability of proteomics approaches to examine a spectrum of interactions in a relatively unbiased manner.

Furthermore, one could easily imagine the utility of the *in vivo* crosslinking strategy in examining the PPIs that govern not just transcriptional initiation, but also events surrounding elongation and termination. For example, the interactions that RNA polymerase II makes with elongation and termination factors and the interactions of transcriptional complexes with the nuclear pore complex could all be examined using Bpa crosslinking. Furthermore, the changes in interactions of Pol II on DNA could theoretically be examined using

TRIC. In particular, by decreasing the crosslinking time of the formaldehyde using rapid mixing techniques, work in this area may be able to reveal a detailed portrait of the PPIs regulating Pol II function over the entire course of transcription. Not only would this be of great benefit to the field, but it has the potential to identify PPIs that could be selectively targeted to activate or inhibit transcription at each identifiable stage.

5.D.2. Incorporation of other unnatural amino acids to examine the mechanism of activator recruitment

Although the work in this thesis primarily focused on the use of Bpa to capture activator-coactivator interactions *in vivo*, the incorporation of other amino acids could provide complementary information to these studies. For example, p-azido-phenylalanine not only provides a crosslinking moiety for complementary *in vivo* crosslinking studies, but the azide functional group also provides a handle with which new functionalities can be added through the use of Click chemistry or a Staudinger ligation. In theory, an Azpa containing activator could be labeled with an alkynl-fluorophore, either through copper catalyzed reaction in cell lysates or through a strain-promoted reaction that removes the requirement of toxic copper. In this way, the labeling of proteins in living cells would allow for the visualization of protein movement and complex dynamics, adding yet another dimension of information to the initial crosslinking studies.

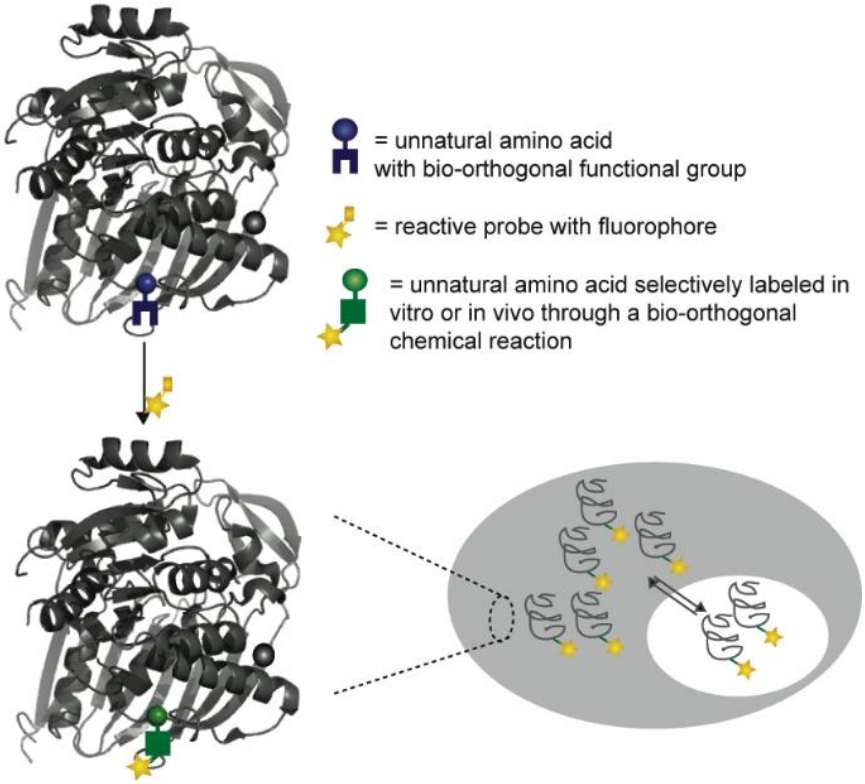

Figure 5-1 Schematic of fluorophore conjugation to proteins labeled with Azpa. Azpa containing proteins are expressed in cells that have taken up a fluorophore with a bio-orthogonal reactive handle. This leads to fluorescent labeling of the protein in cells and allows protein localization and dynamics to be monitored. Alternatively, fluorescent unnatural amino acids can be directly incorporated into proteins provided the tRNA/RS have been engineered for this function.

In light of issues detailed in Chapter 4 surrounding purification strategies, a particularly attractive strategy is to incorporate Bpa derivatives that possess an alkynyl handle, thus removing the need for exogenous tags such as the 6xHis tag used in our earlier purification strategy. This not only provides another handle for purification, but as the alkyne is not naturally available, it will be specific to the protein in which it is incorporated. Thus, the Bpa-alkyne derivatives should react bioorthogonally with a biotin-azide handle, for example, and eliminate cross reactivity issues associated with the use of Ni-agarose purification strategies. Preliminary work done with Cassandra Joiner and Dr. James Clayton demonstrates that Bpa containing an installed alkyne handle (Bpyne), is able to be incorporated into Gal4 in yeast and, furthermore, that this modification does

not impair crosslinking ability (Figure 5-2). In the future, we foresee Bpyne being tremendously useful in crosslinking experiments, as following cell lysis, we should be able to click on a biotin-azide purification handle to pull out covalent activator complexes, thus removing the need for antibodies that can be costly and ineffective.

Bpa-alkyne derivatives and p-Azido-phenylalanine can also be useful in conjugating fluorophores to proteins for *in situ* or *in vivo* visualization studies. Because both of these amino acids can currently be incorporated into proteins in yeast, one could reasonably begin labeling these proteins using click chemistry to covalently attach a fluorescein or Cy5 fluorophore. Labeled proteins could then be visualized using a number of available microscopy techniques. For example, the shuttling of transcriptional complexes into and out of the nucleus could be observed via this route, adding complementary data to what may be observed in TRIC studies. Future work in this area may also focus on the direct incorporation of fluorophores using the nonsense suppression strategy. However, in order for this to be realized, a tRNA/RS pair would have to be selectively evolved to incorporate the fluorescent amino acid. While this appears more plausible for smaller fluorescent groups such as coumarin, the quantum yield of such fluorophores is unfortunately often too low to overcome background fluorescence in cells. Fluorophores like Cy3, Cy5, fluorescein and the Alexa series are far better for *in situ* visualization, but are much larger in size and may be more difficult to fit into the active site of a tRNA synthethetase, even after extensive engineering. Regardless, labeling of proteins for visualization studies appears promising via nonsense suppression routes.

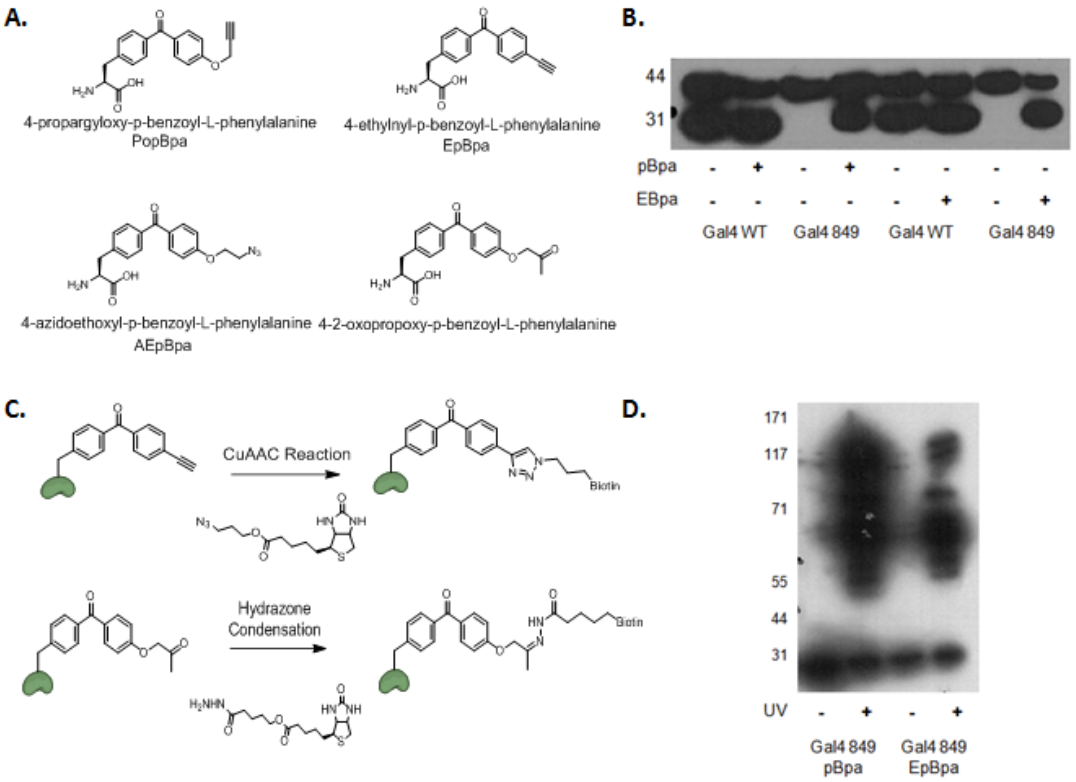

Figure 5-2 A. Several Bpa derivatives are being pursued in the Mapp group, each with an orthogonally reactive handle including terminal alkynes, azides, and ketones. B. One amino acid, Bpyne/Epyne, has already been incorporated into Gal4 in our lab and with yields comparable to Bpa. D. When tested for crosslinking potential, bot Bpa and Bpyne displayed a similar crosslinking profile. C. Examples of how high efficiency purification handles can be conjugated to Bpa derivatives. Work in these figures executed by Dr. James Clayton and Cassandra Joiner.

In sum, the future of unnatural amino acids in the study of PPI networks is promising. In this dissertation, two amino acids in particular, p-benzoyl-L-phenylalanine (Bpa) and p-azidophenylalanine (Azpa) were used to covalently capture the direct interactions of transcriptional activators in live yeast. Moving forward, studies with these amino acids and others with new chemistries can be applied to other PPI networks whose constituent interactions have proven challenging to study.

5.E. References-Chapter 5

(1) Shen, F.; Triezenberg, S. J.; Hensley, P.; Porter, D.; Knutson, J. R. *J Biol Chem* **1996**, *271*, 4827.

(2) Wands, A. M.; Wang, N.; Lum, J. K.; Hsieh, J.; Fierke, C. A.; Mapp, A. K. *J Biol Chem* **2010**, *286*, 16238.
(3) Majmudar, C. Y.; Wang, B.; Lum, J. K.; Hakansson, K.; Mapp, A. K. *Angew Chem Int Ed Engl* **2009**, *48*, 7021.
(4) Lancia, J. K.; Nwokoye, A.; Dugan, A.; Pricer, R.; Joiner, C.; Mapp, A. K. *Biopolymers* **2013**, *manuscript accepted*.
(5) Majmudar, C. Y.; Lee, L. W.; Lancia, J. K.; Nwokoye, A.; Wang, Q.; Wands, A. M.; Wang, L.; Mapp, A. K. *J Am Chem Soc* **2009**, *131*, 14240.
(6) Laurent, B. C.; Carlson, M. *Genes Dev* **1992**, *6*, 1707.
(7) Phelan, M. L.; Sif, S.; Narlikar, G.; Kingston, R. E. *Mol Cell* **1999**, *3*, 247.
(8) Vincent, O.; Townley, R.; Kuchin, S.; Carlson, M. *Genes Dev.* **2001**, *15*, 1104.