MECHANISTIC INVESTIGATIONS OF TRANSCRIPTIONAL ACTIVATOR FUNCTION FOR THE DESIGN OF SYNTHETIC REPLACEMENTS

by

Chinmay Y. Majmudar

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2008

Doctoral Committee:

Associate Professor Anna K. Mapp, Chair
Professor Carol A. Fierke
Professor Gary D. Glick
Assistant Professor Anuj Kumar

To my parents and grandparents

Acknowledgments

I would like to thank Prof. Anna Mapp for guidance, encouragement and support throughout my time at Michigan. I have certainly become a better scientist by learning a great deal from you, from skillfully designing experiments to performing them successfully and finally writing and presenting the results in an engaging fashion; I am extremely grateful for your mentorship. Thank you for being so considerate about all our needs and especially your patience, since a lot of this work took a lot longer than we all anticipated.

Also, thanks to the rest of my thesis committee, Prof. Carol Fierke, Prof. Gary Glick and Prof. Anuj Kumar for their guidance and additionally for allowing me access to use reagents and equipment in their laboratory without which a lot of the experiments described here would have been difficult to perform.

I would particularly like to thank the members of the Mapp Lab that were present when I joined the lab as a first year graduate student: Bin Chen, Aaron Minter, Garrette Belanger, Zhiqian Wu, Amelia Fuller, Marcelle Ferguson, Zikiya Norton, Steve Rowe, Brian Brennan and Jenifer Lum. Thank you for helping me with experiments, assisting with candidacy and providing useful advice about being successful in graduate school. Specifically, I would like to thank Jenifer Lum, with whom I think that I collaborated extremely successfully on several projects and from whom I learned almost all the techniques (and borrowed reagents for experiments) that are described in this thesis.

Thank you to the more recent members of the Mapp lab: Caleb Bates, Amy Danowitz, Sara Buhrlage, Ibrahim Bori, Jonas Hojfedt, Ryan Casey, Jean-Paul Desaulniers, Chris Taylor, Lori Lee, Amberlyn Wands, Adaora Nwokoye, Jody Lancia and Malathy Krishnamurthy for assisting with several on going collaborations and making everyday in the lab an interesting and exciting day to look forward to. I would also like to thank former undergraduate and rotation students that worked on overlapping

projects with me: Lev Prasov, Stephen Winter, Myra Bertrand, Annie Labut and Marchello Cavitt.

I would like to thank our collaborator Dr. Aseem Ansari (University of Wisconsin), for all his insight, advice, reagents for past and ongoing collaborations and also allowing me to visit his laboratory and learn new techniques that ended up being particularly useful to overcome a major roadblock in one of the projects. I would also like to thank the high throughput protein lab at the Life Science Institute, in particular, Dr. William Clay Brown for his help in designing protein expression constructs for us. Also, Leila Foroughi (Matzger Lab) and Dr. Jennifer Meagher (LSI) for putting in a lot of effort towards crystallizing Med15.

I would also like to thank Manoj Cheriyan in the Fierke lab for helping me throughout my graduate career inside and outside Chemistry, starting from my rotation in the Fierke Lab. Also, the graduate students and post-docs in the Pecoraro group for helping me with peptide synthesis and members of the Glick group for assistance with mammalian cell culture and protein purification.

Finally, I would like to thank Bo Wang and Prof. Kristina Hakånsson for collaborating with us and to Bo for putting in a lot of time on a technically challenging mass spectrometry project.

And thank you to my family for everything else!

Table of Contents

List of Figures

List of Tables

List of Abbreviations

TF	Transcription Factor
ATF	Artificial Transcriptional Activator
CBP	CREB Binding Protein
DBD	DNA Binding Domain
PIC	Pre-initiation Complex
TAD	Transcriptional Activation Domain
GST	Glutathione S Transferase
MBP	Maltose Binding Protein
SAGA	Spt-Ada-Gcn5-acetyltransferse
hDM2	human double minute 2
ELISA	Enzyme Linked Immunosorbant Assay
VP16	Viral Protein 16
OD	Optical Density
TBP	TATA Binding Protein
TAF	TBP Associated Factors
ADH	Alcohol Dehydrogenase
CIP	Calf Intestinal Phostatase
PBS	Phosphate Buffered Saline
PEG	Polyethylene Glycol
mOCR	Mutated Overcome Classical Restriction protein
HRP	Horse Radish Peroxidase
SD	Standard Deviation
SDOM	Standard Deviation Of the Mean
HEK	Human Embryonic Kidney
FP	Fluorescence Polarization
CCS	Combined Consensus hydrophobicity Scale
PAGE	Polyacrylamide Gel Electrophoresis

Abstract

MECHANISTIC INVESTIGATIONS OF TRANSCRIPTIONAL ACTIVATOR FUNCTION FOR THE DESIGN OF SYNTHETIC REPLACEMENTS

by

Chinmay Y. Majmudar

Chair: Anna K. Mapp

Transcriptional activators play a critical role in regulating gene expression, precisely controlling the transcriptional response of their cognate genes in a signal-responsive fashion. Transcriptional activation occurs when the activator localizes at a specific DNA sequence and facilitates the assembly of the transcriptional machinery (RNA polymerase II and associated transcription factors) at a gene. Malfunctioning transcriptional activators are associated with a significant percentage of human diseases; greater than 50% of all cancers, for example, are associated with mutations in the transcriptional activator p53. Thus, the development of activator artificial transcriptional factors (ATFs) that functionally replicate their natural counterparts is emerging as a potential therapeutic strategy. One of the biggest hurdles to this goal is the lack of knowledge about the binding interactions utilized by natural transcriptional activators to up-regulate gene expression. The major goal of this work is to delineate the features of activator binding interactions in order to develop useful activator ATFs.

Natural transcriptional activators exhibit a multi-partner binding profile in vitro and there is evidence that this is also true in vivo, although the identities of the binding partners are unknown. To investigate the feasibility that a single binding interaction could lead to transcription function, peptide ligands for the postulated activator target Med15(Gal11) were identified through a binding screen. Satisfyingly, localizing these ligands to a promoter in *S. cerevisiae* results in transcription activation that is Med15(Gal11)-specific. Activator ATFs constructed with these ligands were not, however, as active as natural activators. Activator ATFs with enhanced function could be created by incorporating ligands that were able to interact with more than one partner. Ligands that interact with both Med15(Gal11) and the SAGA component Tra1 upregulated transcription to higher levels than those targeting Med15(Gal11) alone. In addition, the incorporation of a masking interaction into the activator ATFs led to a profoundly positive impact on function. Finally, towards identifying the functionally relevant binding partners of transcriptional activators, a nonsense suppression strategy was adapted to enable incorporation of photoactivatable, crosslinking amino acids into natural transcriptional activators in *S. cerevisiae*. Crosslinking experiments with Gal4 thus modified revealed at least three binding partners that will be the subject of further study.

CHAPTER 1

INTRODUCTION

A. Abstract

Gene transcription is a fundamental cellular process that utilizes a cascade of protein-protein, small molecule-protein and protein-DNA interactions.[1] Disruption of any number of transcription-related interactions present in healthy cells leads to aberrant gene expression and often to disease states as varied as cancer and diabetes.[2, 3] For example, mutations in the transcriptional activator and tumor suppressor p53 that alter its function are associated with 50% of all human cancers.[4, 5] Replacements of malfunctioning transcription factors with functional mimics and molecules that repair defects in these pathways will be extremely useful for biomedical applications for treatment of such disease states.[6-8] In an early advance in this therapeutic arena, scientists at Sangamo Biosciences engineered a zinc finger-based transcription factor to induce angiogenesis in a mouse model.[9] This artificial protein, when expressed in mice, was able to up regulate transcription of the VEGF gene leading to an acceleration of experimental wound healing. Approaches to develop newer generations of such therapeutic molecules that have advantageous stability, delivery and functional properties will have an immense positive impact on the feasibility of transcriptional therapies.[10-13] However, as outlined in this Chapter, there are significant gaps in our knowledge of transcriptional regulation that must be filled in order for more effective transcription-targeted therapeutic agents to be envisioned.

B. Overview of transcription

Transcription entails the creation of RNA from a DNA template by the multi-subunit RNA polymerase holoenzyme and its associated transcription factors.[1] In eukaryotes, there are 3 types of RNA polymerases. RNA pol I is responsible for the

transcription of 28S, 18S, and 5.8S rRNA genes while RNA pol II transcribes all protein encoding mRNA and snRNA. The third polmerase, RNA pol III, transcribes 5S rRNA and tRNA genes.[14-16] The studies presented here deal with regulation of RNA pol II transcription leading to mRNA production, which is eventually translated into protein by the ribosome.

Transcription of a particular gene is initiated when, in response to internal or external stimuli, proteins known as transcriptional activators localize to specific DNA binding sites upstream of the gene.[17] Subsequently, activators recruit multiple protein complexes that comprise the transcriptional machinery to the promoter in order to turn on gene expression.[18, 19]

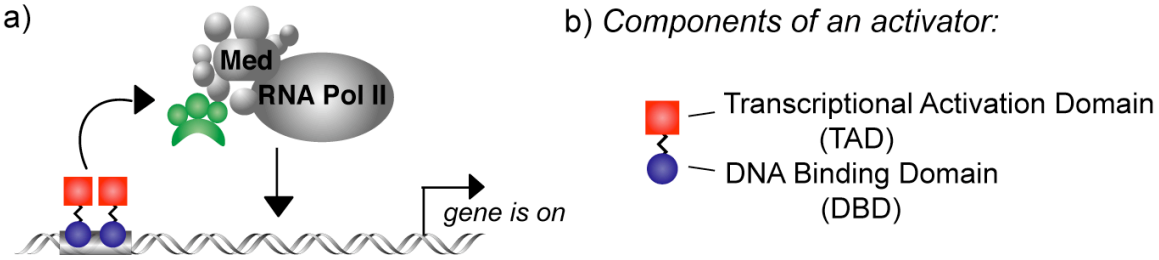


Figure 1.1 Transcriptional activators. a) Activators function by recruiting the transcriptional machinery to a gene promoter. b) An activator is minimally constructed of two modular domains, a TAD (red) and a DBD (blue) usually linked by an unstructured linker region.

B.1 Transcriptional activators

Transcriptional activators are modular proteins that are minimally constructed of a DNA binding domain (DBD) and a transcriptional activation domain (TAD) (Figure 1.1). These modules are able to function independently of each other and it is possible to swap the TADs and DBDs between two activators.[20, 21] In other words, if the TAD of activator A is attached to the DBD of activator B, a new functional activator with the gene-targeting properties of activator B will be created.

B.1.a DNA binding domain (DBD)

The DBD provides much of the gene-targeting specificity of the activator because it is responsible for localizing the protein to its cognate DNA sequence in the promoter of the gene that is to be transcribed. Natural DBDs consist of a wide variety of structural folds including zinc fingers, zinc clusters, leucine zippers and the helix-turn-helix motif

that are known to interact with DNA with high specificity and with affinities ranging from low to high nM.[22-25] These interactions are also structurally well characterized, with many solid state and solution structures illuminating the details of the protein-DNA interactions. For example, extensive conformational and thermodynamic studies with the Gcn4 DBD revealed the amino acid residues involved in DNA recognition and how the DBD is able to use water molecules to differentially interact with the AP-1 and ATF/CREB DNA binding sites (Figure 1.2).[26] However, there is no recognition code for predicting the DNA binding specificity of a natural DBD.[27] Using chromatin immunoprecipitation (ChIP)-on-chip techniques combined with in vivo footprinting, it has been possible to identify the DNA binding sites of DBDs within the genome, and in many cases these have been correlated with gene expression profiles.[28-30] From these experiments it has been possible to detect all of the genes that are regulated by a particular activator.[31] However, DBDs do not generally play a role in controlling the level to which an activator upregulates a particular gene.[17]

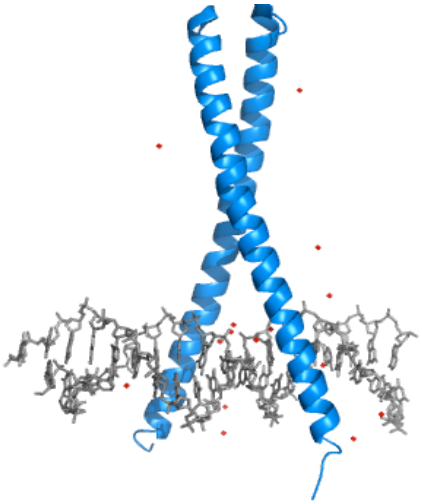


Figure 1.2 Structure of Gcn4 bound to DNA. NMR structure of Gcn4 bZIP domain (blue) that mediates dimerization and DNA binding. bZIP consists of a coiled-coil dimerization domain-leucine-zipper and an adjoining basic region which mediates binding to the AP-1 half site (shown in gray). Water molecules that mediate protein-DNA interactions are shown as red dots (PDB ID: 2DGC)

B.1.b Transcriptional activation domain (TAD)

The TAD controls the timing and level of upregulation of a particular gene to be transcribed. In contrast to the DBD, much less is known about the mechanistic details of how this domain functions. There is also little or no sequence conservation within TADs across transcriptional activators, so this has limited the ability to predict the location of a TAD based on sequence alone.[1] Thus in absence of any higher classification, TADs are often classified depending on their overall sequence composition as proline rich, glutamine/asn rich or acid rich (for examples see Figure 1.3).[32-39] More recently, using

the residues commonly found in known TADs, Piskacek and coworkers have developed an algorithm based on a 9 amino acid sequence motif that can predict the location of a TAD in a given protein sequence.[40] Although this algorithm can successfully identify TADs in natural activators such as Gal4 and VP16, it also gives rise to false positives and further refinements in the selection criteria are required for it to be a general tool to predict unknown TADs.

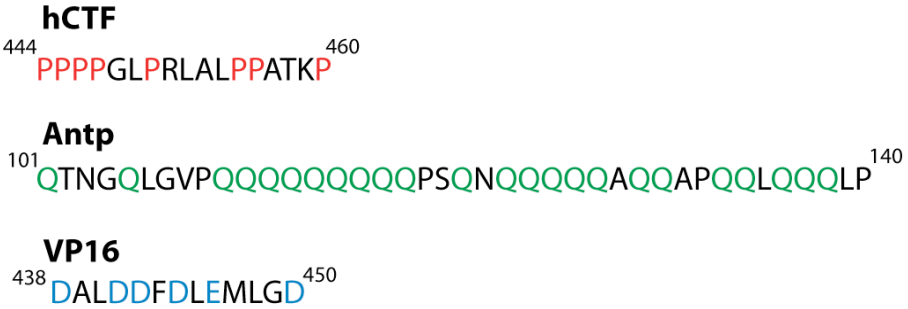


Figure 1.3 Examples of transcriptional activation domains. TADs are often classified based on their sequence composition. Human CAAT box Transcription Factor (hCTF) contains a proline (red) rich TAD. The Drosophila Antennapedia protein (Antp) TAD contains several glutamine (green) residues. The TAD from the Herpes simplex virus Viral Protein 16 (VP16) is acid rich.

The largest and most well studied class of TADs is the so called acid-rich class. As the name suggests, the name 'acid-rich' arose from the preponderance of glutamic and aspartic acids found in the domain, an unusual feature. However, mutagenesis studies revealed that it is the interspersed hydrophobic amino acids that play a more critical role in function. For example, in the acid-rich VP16 TAD it was found that the leucine and phenylalanine residues were critical for interactions with transcription proteins and cellular function, while individual acidic residues were dispensable.[41-48] This has been more recently supported by kinetic studies of acidic TAD-target complexes that suggest that the overall acidity, but not individual acidic amino acids is critical for an initial electrostatic binding event with targets followed by a structural reorganization.[49-51] In the final complex, the majority of the TAD-target contacts are mediated by hydrophobic amino acid side chains. Thus, acidic TADs have been more appropriately renamed amphipathic TADs.

One property of amphipathic TADs is that they exhibit promiscuous binding profiles, interacting at least in vitro with many protein binding partners within and

outside of the transcriptional machinery. For example, the TAD of the yeast activator Gal4 has more than ten putative binding partners identified through biochemical and/or genetic means.[52-63] As described in more detail in subsequent sections, it has been challenging to validate these interactions using traditional experimental approaches. There are thus a wealth of unresolved questions regarding the mechanism by which amphipathic TADs activate transcription. For example, it is still unclear as to what are the relevant functional targets of amphipathic TADs and how each individual interaction correlates to the overall levels of transcription obtained. To further complicate matters, TADs often target multiple binding sites on the same protein in vitro and it is unclear what the physiological relevance of each of these interactions is.[64]

TADs also participate in a number of interactions with partners outside the transcriptional machinery that regulate the timing and levels of gene expression. These include post-translational modifications such as ubiquitylation, acetylation, phosphorylation, sumoylation, neddylation, methylation and glycosylation.[65-73] Additionally, there are also masking interactions that shield a TAD from interacting with cellular components until it is required for activation.[4, 74] For example, the amphipathic yeast activator Gal4 is masked by the repressor protein Gal80.[75] Addition of galactose to the cell stimulates a conformational change that unmasks the Gal4 TAD, enabling it to activate transcription of its cognate galactose-catabolism genes.[76] Although the role of masking interactions in signaling networks is well-established, it is not clear how these interactions impact the level of gene activation by a given activator.

B.1.b.i TAD Structure

There is limited structural information on TADs, particularly in the absence of a binding partner; to date there are no reported structures of free TADs. The prevailing model is that TADs are unstructured in the absence of a binding partner but adopt a helical conformation when bound to a target.[77, 78] More specifically, the amphipathic helix is proposed to be adopted by TADs when they bind targets (Figure 1.4 and 1.6).[51, 79-81] In particular, the hydrophobic face of the helix makes the important contacts while acidic residues don't seem to be important for the primary interaction. A recent NMR study of the VP16 TAD in complex with a subunit of the general transcriptional machinery component TFIIH also showed that the TAD adopted a 9-residue α helix.[82]

However, there are also suggestions that amphipathic TADs such as Gal4 and Gcn4 may form other structures such as a β-sheet, although these investigations were performed with free TADs in absence of a binding partner.[44, 83]

The best-characterized examples of TAD structure are from complexes with masking proteins. In contrast to activator-transcriptional machinery protein complexes, these interactions are often of higher affinity and specificity which could possibly explain why such complexes have been easier to characterize.[75, 84] A well studied example is the crystal structure of the p53 TAD in complex with its masking protein mDM2, which shows the p53 TAD as an amphipathic helix with the three critical hydrophobic residues buried in a hydrophobic cleft in mDM2 (Figure 1.4).[85] Notably, the three residues important for the masking interaction have also been shown to be important for transcriptional machinery interactions and in turn function.[45] These findings illustrate that masking proteins contact an overlapping set of TAD residues that are important for function, perhaps suggesting a role for masking proteins in modulating the activity of activators.

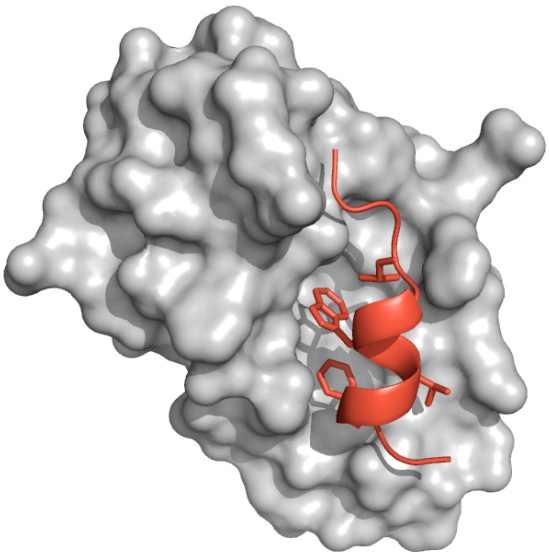


Figure 1.4 Activator-Masking protein interaction. Crystal structure of the amphipathic helical p53 TAD (red) in complex with its masking protein mDM2 (gray surface). The hydrophobic side chains are explicitly shown. The phenylalanine, leucine and tryptophan residues buried in the mDM2 cleft are important for mDM2 binding and for transcriptional activation. PDB ID: 1YCR.

Extensive structural information for other TAD-transcriptional machinery complexes is still not available because the relevant targets of most TADs have not yet been indisputably identified. Moreover, the binding sites on the postulated targets are often not known, leaving researchers to structurally characterize activators in complex with large protein fragments which further hinders the success of obtaining crystals or performing solution-based methods.

B.1.b.ii TAD targets

Over the last three decades, extensive efforts have been invested in identifying the targets of transcriptional activators within the transcriptional machinery. Although much progress has been made, the identity of the bonafide activator targets is still strongly debated. For example, a new set of targets and thus a new activation mechanism has been postulated approximately every 5 years, completely reshaping the way the transcriptional community thinks about TAD function (Figure 1.5).[11]

Initially, it was thought that activators would function by targeting components of the transcriptional machinery that were present at all promoters. Thus, researchers focused on components of the general transcriptional machinery and RNA pol II; an early candidate was the TATA binding protein (TBP).[86, 87] Bolstering this hypothesis were studies that correlated TADs' in vitro binding affinity for TBP with an activator's cellular potency.[60, 88, 89] However, it was later shown that TFIIB rather than TBP recruitment was important for transcriptional activation by an amphipathic activator.[90, 91] Following these studies, TFIIB, TFIIH, TFIIA and RNA pol II were suggested as activator targets.[92-95] Nonetheless, it was shown that supplementation of these components in an in vitro transcription assay did not stimulate activated transcription and thus other targets were probably important for TAD function.

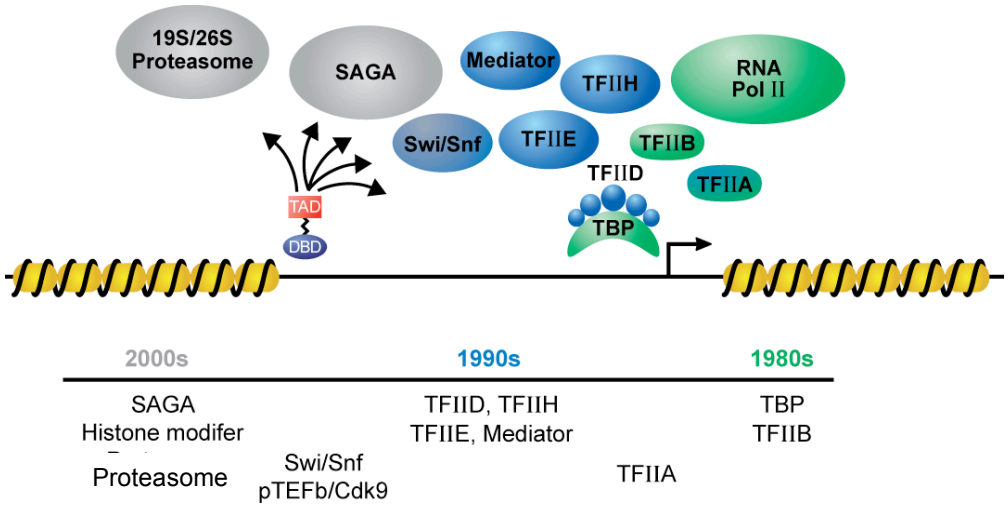


Figure 1.5 A timeline of transcriptional activator targets. During the past three decades, many transcriptional machinery complexes and proteins with these complexes have been identified as key binding partners of TADs. However, significant debate still exists about the relevance of most of these interactions.

Subsequently, coactivators were identified as activator targets; coactivators were defined as proteins or protein complexes that did not interact with DNA but served as a bridge between DNA-bound transcription factors and the components of the general transcriptional machinery. TFIID and its associated TAFs were the first candidates considered in this capacity.[96-98] It was proposed that activators could interact with a TAF and thereby recruit the entire TFIID complex to the promoter. Although it was shown that in vitro activators can stimulate activated transcription in presence of TAFs, genetic experiments in yeast that inactivated individual TAFs showed that only a small fraction of genes were dependent on TAFs.[99, 100]

Concurrently, experiments by several groups led to the identification of another complex of proteins that could "mediate" the function of activators in vitro, called the Srb/Mediator complex.[101-103] Cryo-EM studies with purified yeast Mediator showed a modular complex, consisting of a head that interacts with RNA pol II, a middle module and a tail region, which was predicted to interact with TADs.[104] A series of genetic and biochemical experiments in yeast suggested that activators interact with the Mediator components Med17(Srb4), CDK8(Srb10), Med15(Gal11), and Med2.[57, 59, 65, 105] Further, it was shown using genetic experiments that Mediator was required for the function of several amphipathic activators such as Gal4 and Gcn4.[106, 107]

Interestingly, recent work by Struhl and coworkers has indicated that Mediator may not be required by activators at all promoters in yeast.[108] They show using ChIP that TFIID and not Mediator is recruited to ribosomal and glycolytic gene promoters that utilize activators such as Rap1. Thus, it may be that some promoters are dependent on TFIID, some are dependent on Mediator and others are dependent on both coactivator complexes. Possibly the constitutively active or housekeeping ribosomal promoters operate independently of Mediator (and likewise require TFIID). In contrast, tightly regulated promoters such as the inducible Gal1 (controlled by Gal4), developmentally regulated promoters and activators that respond to environmental stress or non-optimal growth conditions (requiring a rapid increase in gene production levels) probably utilize contacts with Mediator and TFIID that assist in synergistically recruiting RNA polII and the rest of the transcriptional machinery.

Finally, to rationalize how activators and other components of the general

transcriptional machinery efficiently bound DNA in the context of chromatin, several groups investigated whether chromatin-remodeling and modifying enzymes are recruited by TADs.[109-113] In vitro experiments showed that TADs from Gcn4, Swi5, VP16, and the mammalian glucocorticoid receptor could interact with ATP-dependent nucleosome remodeling enzymes (Gcn5, SWI/SNF) and it was predicted that remodeling would greatly enhance stimulation of transcription from chromatinized templates.[114-116] In support of this, it was shown that while SWI/SNF activity was not required for the activator Gal4 to bind to and activate transcription from nucleosome-free binding sites, the complex is required for Gal4 to bind to and function at low-affinity, nucleosomal binding sites in vivo.[117] Subsequently, cell-based experiments revealed that TADs also interacted with components of the SAGA complex, which contains a subset of TAFs (but no TBP) as well as enzymes capable of covalently modifying histones.[118, 119] For example, in vivo FRET experiments from the Green laboratory showed that Tra1, in the context of the chromatin modifying complex SAGA was a target of Gal4 and required for Mediator recruitment to the upstream activating sequence (UAS) of the endogenous Gal1 gene.[53, 54] Similarly, the mammalian homolog of Tra1, TRRAP, is required for gene upregulation by the activators p53 and c-Myc.[120, 121]

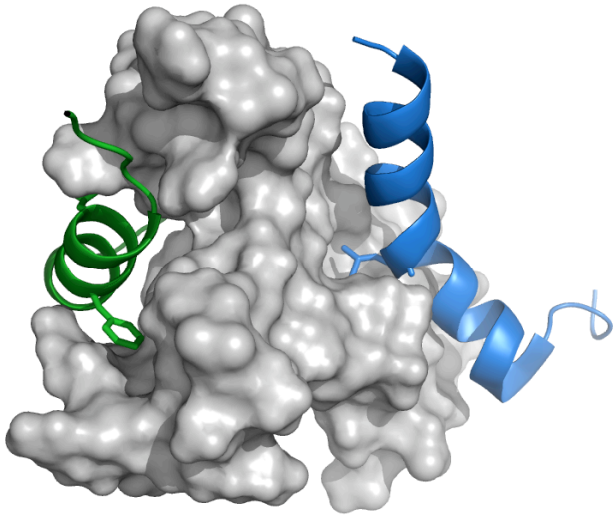
The examples listed above are only some of the prominent targets of TADs discovered in the last three decades, and there still is no general consensus among targets. One of the reasons for this disparity is the limitations of the methods currently used to discover TAD targets. Most of the approaches to date are in vitro and their cellular relevance is known. For example, the Gal4 TAD binds to the non-transcription protein lysozyme in vitro, but this interaction has no functional role in cells; thus in vitro binding is not necessarily a good test for identifying a relevant TAD target.[122] Further, the current cell-based methods involving genetic and formaldehyde-based global crosslinking also have limitations (see Chapter 4 for a more detailed discussion).[58, 123] Thus, newer methods are required to systematically determine activator targets. It is likely that strategies that determine the direct targets of activators in the native functional context will be particularly useful towards this end. The identification of such physiologically relevant protein targets of activators will have profound implications both for the general

mechanistic understanding of transcriptional regulation and for the design and discovery of molecules that can modulate transcription.

B.1.b.iii TAD binding site(s)

An additional complicating factor in the analysis of TAD-coactivator interactions, is that many coactivators appear to contain multiple binding sites for transcriptional activators.[107] For example, the mammalian coactivator CBP/p300 has multiple activator binding domains (GACKIX domain, CH1 and CH3) and within these domains, activators have been shown to target unique binding sites (Figure 1.6).[124, 125] Moreover, the activator p53 has been shown to interact with GACKIX, CH1, CH3 and IBiD domains within CBP/p300 and these individual interactions have different contributions to the overall activator function.[64] Thus, it is not only important in identifying the relevant coactivator targets of TADs, but also the individual binding sites that contribute to function.

Figure 1.6 Structure of the GACKIX domain of CBP bound to MLL and c-Myb. The amphipathic helical TADs from MLL (green) and c-Myb (blue) bind at distinct sites on the CBP (gray, surface rendering). The hydrophobic residues that make key contacts are shown. The TADs bind two distinct binding sites within the GACKIX domain. PDB ID: 2AGH



Current methods to determine coactivator binding sites involve the use of mutagenesis to delete portions of coactivators in vitro and in cells.[126, 127] However, mutagenesis experiments are often difficult to interpret due to pleitropic effects resulting from the changes in coactivator conformation and disruption of transcriptional machinery complexes. For example, a single point mutant in the coactivator Med15(Gal11) (N342V) results is sizable conformational change that enables a new series of Med15-protein interactions.[128, 129] On the other hand, solution and solid state structures of activators in

complex with coactivators are extremely useful in determining activator binding sites, however as discussed earlier, they are challenging to obtain.

Thus, more straightforward approaches to map out all the activator binding sites and correlating them to function are required. Determination of binding sites will help in further understanding activator function and will also assist in the long term goal of developing molecules to regulate transcription by mimicking or inhibiting activator-coactivator interactions.

C. Activator ATFs

The modular nature of activators enables the replacement of the TAD and DBD with natural and non-natural counterparts leading to the creation of activator artifical transcription factors (ATFs).[10] Activator ATFs are in high demand for a variety of applications in medicine, industry and for mechanistic investigation of transcription (vide infra).

Towards these applications, replacements for the DBDs have been successfully developed with protein based DBDs such as DBDs from the endogenous DNA binding proteins Gal4 and LexA. Further approaches using designed zinc fingers have also been successfully developed.[130-133] Non-protein based DBDs such the polyamides, triplex-forming oligonucleotides (TFOs) and peptide nucleic acids (PNAs) have also been designed and used in vitro and in vivo experiments.[134-140] On the other hand, there has been limited development of novel TAD replacements.[141] For instance, the most widely used TADs are short sequences derived from existing natural TADs (Table 1.1). While these TADs function well in cell culture, they will have limited use in therapeutic applications where proteolysis and cellular delivery are a major concern. Thus, one of the long term goals of our research is to develop small molecule TADs that when combined with small molecule DBDs will lead to the creation of fully synthetic activator ATFs displaying advantageous stability and delivery properties. However, it has been challenging to develop small molecules TADs that function with the high levels of activity displayed by their endogenous counterparts.

LexA-Gal4 and Gal4-VP16	Early examples of activator ATF[20, 21]
VP2	Optimized peptide sequence from VP16[142]
B42	Random E. *coli* fragments[143]
AH peptide	Designed to form amphipathic helix[144]
Retinoblastoma (RB) peptide ligands	Two hybrid to bind RB protein[145]
Poly Q and poly P peptides	Activity in vitro [36]
ATF14	Peptide sequence from VP16[146]
P201	In vivo screen of random octamers peptide attached to Gal4(1-100)[147]
LexA-GFP fusions random 15mer peptides	Activity based selection[148]
Gal80bp peptide	Phage display against Gal80[149]
6000-yeast genes fused to Gal4 DBD	[150]
Mammalian cDNA fused to Gal4 DBD	[151]
KIX peptide ligands	Phage display against KIX[152]
KIX peptide ligands	Phage display against KIX[153]
RNA	Activity based selection [154, 155]
Isoxazolidines (small molecules)	Design (active in vitro and in vivo)[156, 157]
Wrenchnolol (small molecule)	Optimized hit from Screen to inhibit ESX-Sur2 (active in vitro)
KIX peptoids	In vitro screen against KIX[158]
Hydrastine (natural product)	Computational screen based on LogP[159]
Santonin (natural product)	Computational screen based on LogP[159]

Table 1.1 List of artificial TADs developed to date

In particular, there are only a handful of small molecule TADs that function to upregulate transcription.[156, 157, 160] In absence of any other design criteria, these small molecule TADs were designed based on the general amphipathic character seen in a number of natural activators. While a great advancement in the field, these molecules moderately activate transcription in cells and are well suited for purposes that do not need extremely high levels of transcriptional activation.[159, 161] However, additional functional features are required to improve their activity to levels similar to potent natural activators like p53 or Hif1α to use them as analogous functional replacements.

The primary reason for the lack of small molecules that function with high levels of activity is the incomplete understanding of the features natural activators use to obtain their cellular activity. Methods to probe the importance and determine the identity, thermodynamic and kinetic parameters of individual activator-target interactions will be particular useful for understanding the contribution of each interaction towards potency.

In the long term, deciphering all the features utilized by natural activators for their overall function will facilitate the incorporation of these features into the design of small molecules leading to potent activator ATFs.

C.1 Applications of activator ATFs

One of the most important applications of activator ATFs is for the development of transcription-based therapeutics to restore the proper function of malfunctioning endogenous transcriptional regulators.[6] For example, overexpression of the transcriptional repressor REST/NRSF in medulloblastoma, leads to the suppression of genes critical for the proper differentiation of neuronal cells.[162, 163] Majumder and co workers have shown that up-regulation of the REST/NRSF-controlled genes by an activator ATF consisting of the REST DBD fused to the VP16 TAD, induces tumor cell apoptosis, demonstrating the utility of transcription-based therapies.[164] Further, activator ATFs could also be used in conjunction with gene therapy strategies to fine-tune the production of therapeutic proteins.[165]

Activator ATFs are also useful in biomanufacturing, molecules that up-regulate transcription to prescribed levels can be used to increase the concentrations of biosynthetic enzymes in order to boost product yields.[166, 167] Finally, in synthetic biology, transcriptional networks are key building blocks used to construct cell-based devices and networks, and molecules that function in a predictable and orthogonal fashion relative to natural regulators will be particularly useful.[168]

While activator ATFs have a number of practical applications, in the near term they are extremely useful as mechanistic tools to understand natural activator function. For example, in the 1980s, the Ptashne group demonstrated the modularity of transcriptional activators by fusing a TAD from one activator with a DBD from another and showed that the artificial activator created functioned using the unique characteristics of each individual domain.[20, 21] Subsequently, they also showed that a peptide, AH, designed to mimic an amphipathic helix was able to activate transcription robustly, indicating that perhaps an amphipathic helix is an important feature found in natural activators.[144] Further, several groups have identified random and designer peptide sequences that function as transcriptional activators when localized to DNA; these studies

have highlighted the chemical functionality that is important for activation.[146, 149, 152, 153] [143]

However, there are many questions regarding the mechanism by which the artificial molecules and natural activators function that remain unanswered. Specifically, given that natural activators interact with a variety of binding partners, the use of activator ATFs to completely characterize each individual interaction will delineate the importance of each recruitment event to the overall activity of natural activators. In the longer term, understanding the features necessary for controlling transcriptional activity will lead to newer generations of activator ATFs that closely mimic endogenous activators.

D. Overview of thesis

The work described here focuses on probing the mechanism of natural transcriptional activators towards the long-term goal of designing activator ATFs for therapeutically and biotechnologically useful purposes. We have found that specifically targeting the Mediator component Med15(Gal11) gives rise to transcriptional activation and the level of transcription seemed to vary upon the binding site on Med15(Gal11) (Chapter 2). Further, we show that incorporating additional binding interactions inside and outside the transcriptional machinery are strategies that activators use to achieve potent levels of transcriptional activation (Chapter 3). Finally, we investigate the binding partners of transcriptional activators and the nature of the activator binding sites using both in vitro and cell-based photo crosslinking. We find that the activator Gal4 targets multiple proteins in cells and we identify a small subset of conserved binding sites on the coactivator Med15(Gal11) to be important for the function of endogenous activators (Chapter 4). Incorporation of the features described here into future generations of activator ATFs (Chapter 5), leading to the design of small molecules that function in cells with similar activity as natural activators, will lead to the development of therapies to remedy misregulated transcription and methods to artificially regulate gene expression levels.

References

1. Ptashne, M.; Gann, A., *Genes & Signals*. Cold Spring Harbor Laboratory: New York, 2001.
2. Perou, C. M.; Sorlie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Rees, C. A.; Pollack, J. R.; Ross, D. T.; Johnsen, H.; Aksien, L. A.; Fluge, O.; Pergamenschikov, A.; Williams, C.; Zhu, S. X.; Lonning, P. E.; Borresen-Dale, A. L.; Brown, P. O.; Botstein, D., Molecular portraits of human breast tumours. *Nature* **2000,** 406, (6797), 747-752.
3. Nakae, J.; Biggs, W. H., 3rd; Kitamura, T.; Cavenee, W. K.; Wright, C. V.; Arden, K. C.; Accili, D., Regulation of insulin action and pancreatic beta-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1. *Nat Genet* **2002,** 32, (2), 245-53.
4. Vogelstein, B.; Lane, D.; Levine, A. J., Surfing the p53 network. *Nature* **2000,** 408, (6810), 307-310.
5. Hainaut, P.; Hollstein, M., p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res* **2000,** 77, 81-137.
6. Darnell, J. E., Transcription factors as targets for cancer therapy. *Nature Reviews Cancer* **2002,** 2, (10), 740-749.
7. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M., Arresting Cancer Proliferation by Small Molecule Gene Regulation. *Chem Biol* **2004**.
8. Pandolfi, P. P., Transcription therapy for cancer. *Oncogene* **2001,** 20, (24), 3116-3127.
9. Rebar, E. J.; Huang, Y.; Hickey, R.; Nath, A. K.; Meoli, D.; Nath, S.; Chen, B.; Xu, L.; Liang, Y.; Jamieson, A. C.; Zhang, L.; Spratt, S. K.; Case, C. C.; Wolffe, A.; Giordano, F. J., Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat Med* **2002,** 8, (12), 1427-32.
10. Ansari, A. Z.; Mapp, A. K., Modular design of artificial transcription factors. *Curr. Opin. Chem. Biol.* **2002,** 6, 765-772.
11. Mapp, A. K.; Ansari, A. Z., A TAD further: exogenous control of gene activation. *ACS Chem Biol* **2007,** 2, (1), 62-75.
12. Mapp, A. K., Regulating transcription: a chemical perspective. *Organic & Biomolecular Chemistry* **2003,** 1, (13), 2217-2220.
13. Majmudar, C. Y.; Mapp, A. K., Chemical approaches to transcriptional regulation. *Curr Opin Chem Biol* **2005,** 9, (5), 467-74.
14. Paule, M. R.; White, R. J., Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res* **2000,** 28, (6), 1283-98.
15. Geiduschek, E. P.; Kassavetis, G. A., The RNA polymerase III transcription apparatus. *J Mol Biol* **2001,** 310, (1), 1-26.
16. Cramer, P., Recent structural studies of RNA polymerases II and III. *Biochem Soc Trans* **2006,** 34, (Pt 6), 1058-61.
17. Ptashne, M., How Eukaryotic Transcriptional Activators Work. *Nature* **1988,** 335, (6192), 683-689.
18. Ptashne, M., Control of gene transcription: An outline. *Nature Medicine* **1997,** 3, (10), 1069-1072.

19. Ptashne, M.; Gann, A., Transcriptional activation by recruitment. *Nature* **1997,** 386, (6625), 569-577.
20. Brent, R.; Ptashne, M., A Eukaryotic Transcriptional Activator Bearing the DNA Specificity of a Prokaryotic Repressor. *Cell* **1985,** 43, (3), 729-736.
21. Keegan, L.; Gill, G.; Ptashne, M., Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **1986,** 231, (4739), 699-704.
22. Garvie, C. W.; Wolberger, C., Recognition of specific DNA sequences. *Molecular Cell* **2001,** 8, (5), 937-946.
23. Luscombe, N. M.; Austin, S. E.; Berman, H. M.; Thornton, J. M., An overview of the structures of protein-DNA complexes. *Genome Biol* **2000,** 1, (1), REVIEWS001.
24. Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C., DNA Recognition by Gal4 - Structure of a Protein DNA Complex. *Nature* **1992,** 356, (6368), 408-414.
25. Littlefield, O.; Nelson, H. C., A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocrystal. *Nat Struct Biol* **1999,** 6, (5), 464-70.
26. Keller, W.; Konig, P.; Richmond, T. J., Crystal structure of a bZIP/DNA complex at 2.2 A: determinants of DNA specific recognition. *J Mol Biol* **1995,** 254, (4), 657-67.
27. Pabo, C. O.; Nekludova, L., Geometric analysis and comparison of protein-DNA interfaces: Why is there no simple code for recognition? *Journal of Molecular Biology* **2000,** 301, (3), 597-624.
28. Jothi, R.; Cuddapah, S.; Barski, A.; Cui, K.; Zhao, K., Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* **2008**.
29. Molina, M. P.; Cain, C.; Bargonetti, J., In vivo footprinting and DNA affinity chromatography for analysis of p53 DNA binding ability. *Methods Mol Biol* **2003,** 234, 151-70.
30. Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B., Genome-wide mapping of in vivo protein-DNA interactions. *Science* **2007,** 316, (5830), 1497-502.
31. Smeenk, L.; van Heeringen, S. J.; Koeppel, M.; van Driel, M. A.; Bartels, S. J.; Akkers, R. C.; Denissov, S.; Stunnenberg, H. G.; Lohrum, M., Characterization of genome-wide p53-binding sites upon stress response. *Nucleic Acids Res* **2008,** 36, (11), 3639-54.
32. Almlof, T.; Wright, A. P. H.; Gustafsson, J. A., Role of Acidic and Phosphorylated Residues in Gene Activation by the Glucocorticoid Receptor. *Journal of Biological Chemistry* **1995,** 270, (29), 17535-17540.
33. Blair, W. S.; Bogerd, H. P.; Madore, S. J.; Cullen, B. R., Mutational Analysis of the Transcription Activation Domain of Rela - Identification of a Highly Synergistic Minimal Acidic Activation Module. *Molecular and Cellular Biology* **1994,** 14, (11), 7226-7234.
34. Courey, A. J.; Holtzman, D. A.; Jackson, S. P.; Tjian, R., Synergistic Activation by the Glutamine-Rich Domains of Human Transcription Factor Sp1. *Cell* **1989,** 59, (5), 827-836.
35. Courey, A. J.; Tjian, R., Analysis of Sp1 Invivo Reveals Multiple Transcriptional Domains, Including a Novel Glutamine-Rich Activation Motif. *Cell* **1988,** 55, (5), 887-898.

36. Gerber, H. P.; Seipel, K.; Georgiev, O.; Hofferer, M.; Hug, M.; Rusconi, S.; Schaffner, W., Transcriptional Activation Modulated by Homopolymeric Glutamine and Proline Stretches. *Science* **1994,** 263, (5148), 808-811.
37. Sadowski, I.; Ma, J.; Triezenberg, S.; Ptashne, M., Gal4-Vp16 Is an Unusually Potent Transcriptional Activator. *Nature* **1988,** 335, (6190), 563-564.
38. Kim, T. K.; Roeder, R. G., Transcriptional activation in yeast by the proline-rich activation domain of human CTF1. *J Biol Chem* **1993,** 268, (28), 20866-9.
39. Mermod, N.; Oneill, E. A.; Kelly, T. J.; Tjian, R., The Proline-Rich Transcriptional Activator of Ctf/Nf-I Is Distinct from the Replication and DNA-Binding Domain. *Cell* **1989,** 58, (4), 741-753.
40. Piskacek, S.; Gregor, M.; Nemethova, M.; Grabner, M.; Kovarik, P.; Piskacek, M., Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* **2007,** 89, (6), 756-68.
41. Sullivan, S. M.; Horn, P. J.; Olson, V. A.; Koop, A. H.; Niu, W.; Ebright, R. H.; Triezenberg, S. J., Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Research* **1998,** 26, (19), 4487-4496.
42. Callus, B. A.; Mathey-Prevot, B., Hydrophobic residues Phe(751) and Leu(753) are essential for STAT5 transcriptional activity. *Journal of Biological Chemistry* **2000,** 275, (22), 16954-16962.
43. Chen, R. H.; Fields, S.; Lipsick, J. S., Dissociation of Transcriptional Activation and Oncogenic Transformation by V-Myb. *Oncogene* **1995,** 11, (9), 1771-1779.
44. Leuther, K. K.; Salmeron, J. M.; Johnston, S. A., Genetic evidence that an activation domain of GAL4 does not require acidity and may form a beta sheet. *Cell* **1993,** 72, (4), 575-85.
45. Lin, J.; Chen, J.; Elenbaas, B.; Levine, A. J., Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. *Genes Dev* **1994,** 8, (10), 1235-46.
46. Regier, J. L.; Shen, F.; Triezenberg, S. J., Pattern of Aromatic and Hydrophobic Amino-Acids Critical for One of 2 Subdomains of the Vp16 Transcriptional Activator. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** 90, (3), 883-887.
47. Sainz, M. B.; Goff, S. A.; Chandler, V. L., Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo. *Molecular and Cellular Biology* **1997,** 17, (1), 115-122.
48. Drysdale, C. M.; Duenas, E.; Jackson, B. M.; Reusser, U.; Braus, G. H.; Hinnebusch, A. G., The Transcriptional Activator Gcn4 Contains Multiple Activation Domains That Are Critically Dependent on Hydrophobic Amino-Acids. *Molecular and Cellular Biology* **1995,** 15, (3), 1220-1233.
49. Hermann, S.; Berndt, K. D.; Wright, A. P., How transcriptional activators bind target proteins. *J Biol Chem* **2001,** 276, (43), 40127-32.
50. Ferreira, M. E.; Hermann, S.; Prochasson, P.; Workman, J. L.; Berndt, K. D.; Wright, A. P., Mechanism of transcription factor recruitment by acidic activators. *J Biol Chem* **2005,** 280, (23), 21779-84.
51. Jonker, H. R.; Wechselberger, R. W.; Boelens, R.; Folkers, G. E.; Kaptein, R., Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **2005,** 44, (3), 827-39.

52. Archer, C. T.; Burdine, L.; Kodadek, T., Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking. *Mol Biosyst* **2005,** 1, (5-6), 366-72.
53. Bhaumik, S. R.; Green, M. R., SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes Dev* **2001,** 15, (15), 1935-45.
54. Bhaumik, S. R.; Raha, T.; Aiello, D. P.; Green, M. R., In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev* **2004,** 18, (3), 333-43.
55. Bryant, G. O.; Ptashne, M., Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol Cell* **2003,** 11, (5), 1301-9.
56. Chang, C.; Gonzalez, F.; Rothermel, B.; Sun, L.; Johnston, S. A.; Kodadek, T., The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. *J Biol Chem* **2001,** 276, (33), 30956-63.
57. Jeong, C. J.; Yang, S. H.; Xie, Y.; Zhang, L.; Johnston, S. A.; Kodadek, T., Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **2001,** 40, (31), 9421-7.
58. Klein, J.; Nolden, M.; Sanders, S. L.; Kirchner, J.; Weil, P. A.; Melcher, K., Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA. *J Biol Chem* **2003,** 278, (9), 6779-86.
59. Koh, S. S.; Ansari, A. Z.; Ptashne, M.; Young, R. A., An activator target in the RNA polymerase II holoenzyme. *Mol Cell* **1998,** 1, (6), 895-904.
60. Li, X. Y.; Virbasius, A.; Zhu, X. C.; Green, M. R., Enhancement of TBP binding by activators and general transcription factors. *Nature* **1999,** 399, (6736), 605-609.
61. Melcher, K.; Johnston, S. A., GAL4 interacts with TATA-binding protein and coactivators. *Mol Cell Biol* **1995,** 15, (5), 2839-48.
62. Swaffield, J. C.; Melcher, K.; Johnston, S. A., A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein. *Nature* **1995,** 374, (6517), 88-91.
63. Xie, Y.; Denison, C.; Yang, S. H.; Fancy, D. A.; Kodadek, T., Biochemical characterization of the TATA-binding protein-Gal4 activation domain complex. *J Biol Chem* **2000,** 275, (41), 31914-20.
64. Teufel, D. P.; Freund, S. M.; Bycroft, M.; Fersht, A. R., Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proc Natl Acad Sci U S A* **2007,** 104, (17), 7009-14.
65. Ansari, A. Z.; Koh, S. S.; Zaman, Z.; Bongards, C.; Lehming, N.; Young, R. A.; Ptashne, M., Transcriptional activating regions target a cyclin-dependent kinase. *Proc Natl Acad Sci U S A* **2002,** 99, (23), 14706-9.
66. Bouwman, P.; Philipsen, S., Regulation of the activity of Sp1-related transcription factors. *Molecular and Cellular Endocrinology* **2002,** 195, (1-2), 27-38.
67. Chrivia, J. C.; Kwok, R. P.; Lamb, N.; Hagiwara, M.; Montminy, M. R.; Goodman, R. H., Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **1993,** 365, (6449), 855-9.
68. Jackson, S. P.; Tjian, R., O-Glycosylation of Eukaryotic Transcription Factors - Implications for Mechanisms of Transcriptional Regulation. *Cell* **1988,** 55, (1), 125-133.

69. Khidekel, N.; Hsieh-Wilson, L. C., A 'molecular switchboard'--covalent modifications to proteins and their impact on transcription. *Org Biomol Chem* **2004,** 2, (1), 1-7.
70. Lamarre-Vincent, N.; Hsieh-Wilson, L. C., Dynamic glycosylation of the transcription factor CREB: A potential role in gene regulation. *Journal of the American Chemical Society* **2003,** 125, (22), 6612-6613.
71. Lamph, W. W.; Dwarki, V. J.; Ofir, R.; Montminy, M.; Verma, I. M., Negative and Positive Regulation by Transcription Factor Camp Response Element-Binding Protein Is Modulated by Phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **1990,** 87, (11), 4320-4324.
72. Sadowski, I.; Niedbala, D.; Wood, K.; Ptashne, M., Gal4 Is Phosphorylated as a Consequence of Transcriptional Activation. *Proceedings of the National Academy of Sciences of the United States of America* **1991,** 88, (23), 10510-10514.
73. Ferdous, A.; Sikder, D.; Gillette, T.; Nalley, K.; Kodadek, T.; Johnston, S. A., The role of the proteasomal ATPases and activator monoubiquitylation in regulating Gal4 binding to promoters. *Genes Dev* **2007,** 21, (1), 112-23.
74. des Etages, S. A.; Falvey, D. A.; Reece, R. J.; Brandriss, M. C., Functional analysis of the PUT3 transcriptional activator of the proline utilization pathway in Saccharomyces cerevisiae. *Genetics* **1996,** 142, (4), 1069-82.
75. Ansari, A. Z.; Reece, R. J.; Ptashne, M., A transcriptional activating region with two contrasting modes of protein interaction. *Proc Natl Acad Sci U S A* **1998,** 95, (23), 13543-8.
76. Thoden, J. B.; Sellick, C. A.; Reece, R. J.; Holden, H. M., Understanding a transcriptional paradigm at the molecular level. The structure of yeast Gal80p. *J Biol Chem* **2007,** 282, (3), 1534-8.
77. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **2005,** 6, (3), 197-208.
78. Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G., Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A* **2000,** 97, (16), 8868-73.
79. Parker, D.; Jhala, U. S.; Radhakrishnan, I.; Yaffe, M. B.; Reyes, C.; Shulman, A. I.; Cantley, L. C.; Wright, P. E.; Montminy, M., Analysis of an activator : coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular Cell* **1998,** 2, (3), 353-359.
80. Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L., Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **1997,** 277, (5330), 1310-1313.
81. Zor, T.; De Guzman, R. N.; Dyson, H. J.; Wright, P. E., Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *Journal of Molecular Biology* **2004,** 337, (3), 521-534.
82. Langlois, C.; Mas, C.; Di Lello, P.; Jenkins, L. M.; Legault, P.; Omichinski, J. G., NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53. *J Am Chem Soc* **2008,** 130, (32), 10596-604.

83. Van Hoy, M.; Leuther, K. K.; Kodadek, T.; Johnston, S. A., The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets. *Cell* **1993,** 72, (4), 587-94.
84. Schon, O.; Friedler, A.; Bycroft, M.; Freund, S. M. V.; Fersht, A. R., Molecular mechanism of the interaction between MDM2 and p53. *J Mol Biol* **2002,** 323, (3), 491-501.
85. Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P., Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **1996,** 274, (5289), 948-953.
86. Horikoshi, M.; Hai, T.; Lin, Y. S.; Green, M. R.; Roeder, R. G., Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* **1988,** 54, (7), 1033-42.
87. Stringer, K. F.; Ingles, C. J.; Greenblatt, J., Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* **1990,** 345, (6278), 783-6.
88. Nedialkov, Y. A.; Triezenberg, S. J., Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region. *Arch Biochem Biophys* **2004,** 425, (1), 77-86.
89. Wu, Y.; Reece, R. J.; Ptashne, M., Quantitation of putative activator-target affinities predicts transcriptional activating potentials. *Embo J* **1996,** 15, (15), 3951-63.
90. Lin, Y. S.; Ha, I.; Maldonado, E.; Reinberg, D.; Green, M. R., Binding of general transcription factor TFIIB to an acidic activating region. *Nature* **1991,** 353, (6344), 569-71.
91. Choy, B.; Green, M. R., Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* **1993,** 366, (6455), 531-6.
92. Stargell, L. A.; Struhl, K., The TBP-TFIIA interaction in the response to acidic activators in vivo. *Science* **1995,** 269, (5220), 75-8.
93. Xiao, H.; Pearson, A.; Coulombe, B.; Truant, R.; Zhang, S.; Regier, J. L.; Triezenberg, S. J.; Reinberg, D.; Flores, O.; Ingles, C. J.; et al., Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. *Mol Cell Biol* **1994,** 14, (10), 7013-24.
94. Ozer, J.; Bolden, A. H.; Lieberman, P. M., Transcription factor IIA mutations show activator-specific defects and reveal a IIA function distinct from stimulation of TBP-DNA binding. *J Biol Chem* **1996,** 271, (19), 11182-90.
95. Tan, Q.; Linask, K. L.; Ebright, R. H.; Woychik, N. A., Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. *Genes Dev* **2000,** 14, (3), 339-48.
96. Burley, S. K.; Roeder, R. G., Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* **1996,** 65, 769-99.
97. Tjian, R., Molecular machines that control genes. *Sci Am* **1995,** 272, (2), 54-61.
98. Dynlacht, B. D.; Hoey, T.; Tjian, R., Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **1991,** 66, (3), 563-76.
99. Moqtaderi, Z.; Bai, Y.; Poon, D.; Weil, P. A.; Struhl, K., TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* **1996,** 383, (6596), 188-91.

100. Walker, S. S.; Reese, J. C.; Apone, L. M.; Green, M. R., Transcription activation in cells lacking TAFIIS. *Nature* **1996,** 383, (6596), 185-8.
101. Koleske, A. J.; Young, R. A., An RNA polymerase II holoenzyme responsive to activators. *Nature* **1994,** 368, (6470), 466-9.
102. Flanagan, P. M.; Kelleher, R. J., 3rd; Sayre, M. H.; Tschochner, H.; Kornberg, R. D., A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **1991,** 350, (6317), 436-8.
103. Kim, Y. J.; Bjorklund, S.; Li, Y.; Sayre, M. H.; Kornberg, R. D., A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **1994,** 77, (4), 599-608.
104. Asturias, F. J.; Jiang, Y. W.; Myers, L. C.; Gustafsson, C. M.; Kornberg, R. D., Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **1999,** 283, (5404), 985-7.
105. Myers, L. C.; Gustafsson, C. M.; Hayashibara, K. C.; Brown, P. O.; Kornberg, R. D., Mediator protein mutations that selectively abolish activated transcription. *Proc Natl Acad Sci U S A* **1999,** 96, (1), 67-72.
106. Govind, C. K.; Yoon, S.; Qiu, H.; Govind, S.; Hinnebusch, A. G., Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo. *Mol Cell Biol* **2005,** 25, (13), 5626-38.
107. Park, J. M.; Kim, H. S.; Han, S. J.; Hwang, M. S.; Lee, Y. C.; Kim, Y. J., In vivo requirement of activator-specific binding targets of mediator. *Molecular and Cellular Biology* **2000,** 20, (23), 8709-8719.
108. Fan, X.; Chou, D. M.; Struhl, K., Activator-specific recruitment of Mediator in vivo. *Nat Struct Mol Biol* **2006,** 13, (2), 117-20.
109. Fry, C. J.; Peterson, C. L., Chromatin remodeling enzymes: who's on first? *Curr Biol* **2001,** 11, (5), R185-97.
110. Kadonaga, J. T., Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **1998,** 92, (3), 307-13.
111. Workman, J. L.; Kingston, R. E., Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* **1998,** 67, 545-79.
112. Tsukiyama, T.; Wu, C., Chromatin remodeling and transcription. *Curr Opin Genet Dev* **1997,** 7, (2), 182-91.
113. Carpenter, A. E.; Memedula, S.; Plutz, M. J.; Belmont, A. S., Common effects of acidic activators on large-scale chromatin structure and transcription. *Mol Cell Biol* **2005,** 25, (3), 958-68.
114. Neely, K. E.; Hassan, A. H.; Wallberg, A. E.; Steger, D. J.; Cairns, B. R.; Wright, A. P.; Workman, J. L., Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. *Mol Cell* **1999,** 4, (4), 649-55.
115. Yudkovsky, N.; Logie, C.; Hahn, S.; Peterson, C. L., Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes Dev* **1999,** 13, (18), 2369-74.
116. Wallberg, A. E.; Neely, K. E.; Gustafsson, J. A.; Workman, J. L.; Wright, A. P.; Grant, P. A., Histone acetyltransferase complexes can mediate transcriptional activation by the major glucocorticoid receptor activation domain. *Mol Cell Biol* **1999,** 19, (9), 5952-9.

117. Burns, L. G.; Peterson, C. L., The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. *Mol Cell Biol* **1997,** 17, (8), 4811-9.
118. Grant, P. A.; Sterner, D. E.; Duggan, L. J.; Workman, J. L.; Berger, S. L., The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes. *Trends Cell Biol* **1998,** 8, (5), 193-7.
119. Kwok, R. P.; Lundblad, J. R.; Chrivia, J. C.; Richards, J. P.; Bachinger, H. P.; Brennan, R. G.; Roberts, S. G.; Green, M. R.; Goodman, R. H., Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **1994,** 370, (6486), 223-6.
120. McMahon, S. B.; Van Buskirk, H. A.; Dugan, K. A.; Copeland, T. D.; Cole, M. D., The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell* **1998,** 94, (3), 363-74.
121. Ard, P. G.; Chatterjee, C.; Kunjibettu, S.; Adside, L. R.; Gralinski, L. E.; McMahon, S. B., Transcriptional regulation of the mdm2 oncogene by p53 requires TRRAP acetyltransferase complexes. *Mol Cell Biol* **2002,** 22, (16), 5650-61.
122. Melcher, K., The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins. *Journal of Molecular Biology* **2000,** 301, (5), 1097-1112.
123. Hall, D. B.; Struhl, K., The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. *J Biol Chem* **2002,** 277, (48), 46043-50.
124. Chan, H. M.; La Thangue, N. B., p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J Cell Sci* **2001,** 114, (Pt 13), 2363-73.
125. McManus, K. J.; Hendzel, M. J., CBP, a transcriptional coactivator and acetyltransferase. *Biochem Cell Biol* **2001,** 79, (3), 253-66.
126. Kim, D. H.; Kim, G. S.; Yun, C. H.; Lee, Y. C., Functional conservation of the glutamine-rich domains of yeast Gal11 and human SRC-1 in the transactivation of glucocorticoid receptor Tau 1 in Saccharomyces cerevisiae. *Mol Cell Biol* **2008,** 28, (3), 913-25.
127. Lu, Z.; Ansari, A. Z.; Lu, X. Y.; Ogirala, A.; Ptashne, M., A target essential for the activity of a nonacidic yeast transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (13), 8591-8596.
128. Himmelfarb, H. J.; Pearlberg, J.; Last, D. H.; Ptashne, M., GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators. *Cell* **1990,** 63, (6), 1299-309.
129. Hidalgo, P.; Ansari, A. Z.; Schmidt, P.; Hare, B.; Simkovich, N.; Farrell, S.; Shin, E. J.; Ptashne, M.; Wagner, G., Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes & Development* **2001,** 15, (8), 1007-1020.
130. Beerli, R. R.; Barbas, C. F., Engineering polydactyl zinc-finger transcription factors. *Nature Biotechnology* **2002,** 20, (2), 135-141.
131. Beerli, R. R.; Dreier, B.; Barbas, C. F., Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci U S A* **2000,** 97, (4), 1495-1500.
132. Blancafort, P.; Segal, D. J.; Barbas, C. F., 3rd, Designing transcription factor architectures for drug discovery. *Mol Pharmacol* **2004,** 66, (6), 1361-71.

133. Lin, Q.; Barbas, C. F.; Schultz, P. G., Small-molecule switches for zinc finger transcription factors. *J Am Chem Soc* **2003,** 125, (3), 612-613.
134. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (8), 3930-3935.
135. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Current Opinion in Structural Biology* **2003,** 13, (3), 284-299.
136. Liu, B.; Han, Y.; Ferdous, A.; Corey, D. R.; Kodadek, T., Transcription activation by a PNA-peptide chimera in a mammalian cell extract. *Chemistry & Biology* **2003,** 10, (10), 909-916.
137. Nielsen, P. E., Peptide nucleic acid targeting of double-stranded DNA. In *Drug-Nucleic Acid Interactions*, 2001; Vol. 340, pp 329-340.
138. Faria, M.; Giovannangeli, C., Triplex-forming molecules: from concepts to applications. *Journal of Gene Medicine* **2001,** 3, (4), 299-310.
139. Kuznetsova, S.; Ait-Si-Ali, S.; Nagibneva, I.; Troalen, F.; Le Villain, J. P.; Harel-Bellan, A.; Svinarchuk, F., Gene activation by triplex-forming oligonucleotide coupled to the activating domain of protein VP16. *Nucleic Acids Research* **1999,** 27, (20), 3995-4000.
140. Uil, T. G.; Haisma, H. J.; Rots, M. G., Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucleic Acids Res* **2003,** 31, (21), 6064-78.
141. Lum, J. K.; Mapp, A. K., Artificial transcriptional activation domains. *Chembiochem* **2005,** 6, (8), 1311-5.
142. Tanaka, M., Modulation of promoter occupancy by cooperative DNA binding and activation-domain function is a major determinant of transcriptional regulation by activators in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (9), 4311-4315.
143. Ruden, D. M.; Ma, J.; Li, Y.; Wood, K.; Ptashne, M., Generating yeast transcriptional activators containing no yeast protein sequences. *Nature* **1991,** 350, (6315), 250-2.
144. Giniger, E.; Ptashne, M., Transcription in Yeast Activated by a Putative Amphipathic Alpha-Helix Linked to a DNA-Binding Unit. *Nature* **1987,** 330, (6149), 670-672.
145. Yang, M.; Wu, Z.; Fields, S., Protein-peptide interactions analyzed with the yeast two-hybrid system. *Nucleic Acids Res* **1995,** 23, (7), 1152-6.
146. Stanojevic, D.; Young, R. A., A highly potent artificial transcription factor. *Biochemistry* **2002,** 41, (23), 7209-7216.
147. Lu, X. Y.; Ansari, A. Z.; Ptashne, M., An artificial transcriptional activating region with unusual properties. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (5), 1988-1992.
148. Abedi, M.; Caponigro, G.; Shen, J.; Hansen, S.; Sandrock, T.; Kamb, A., Transcriptional transactivation by selected short random peptides attached to lexA-GFP fusion proteins. *BMC Mol Biol* **2001,** 2, 10.
149. Han, Y.; Kodadek, T., Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. *J Biol Chem* **2000,** 275, (20), 14979-84.

150. Titz, B.; Thomas, S.; Rajagopala, S. V.; Chiba, T.; Ito, T.; Uetz, P., Transcriptional activators in yeast. *Nucleic Acids Res* **2006,** 34, (3), 955-67.
151. Wiesner, C.; Hoeth, M.; Binder, B. R.; de Martin, R., A functional screening assay for the isolation of transcription factors. *Nucleic Acids Res* **2002,** 30, (16), e80.
152. Frangioni, J. V.; LaRiccia, L. M.; Cantley, L. C.; Montminy, M. R., Minimal activators that bind to the KIX domain of p300/CBP identified by phage display screening. *Nature Biotechnology* **2000,** 18, (10), 1080-1085.
153. Rutledge, S. E.; Volkman, H. M.; Schepartz, A., Molecular recognition of protein surfaces: High affinity ligands for the CBPKIX domain. *J Am Chem Soc* **2003,** 125, (47), 14336-14347.
154. Saha, S.; Ansari, A. Z.; Jarrell, K. A.; Ptashne, M., RNA sequences that work as transcriptional activating regions. *Nucleic Acids Research* **2003,** 31, 1565-1570.
155. Buskirk, A. R.; Kehayova, P. D.; Landrigan, A.; Liu, D. R., In vivo evolution of an RNA-based transcriptional activator. *Chemistry & Biology* **2003,** 10, (6), 533-540.
156. Minter, A. R.; Brennan, B. B.; Mapp, A. K., A small molecule transcriptional activation domain. *Journal of the American Chemical Society* **2004,** 126, (34), 10504-10505.
157. Buhrlage, S. J.; Brennan, B. B.; Minter, A. R.; Mapp, A. K., Stereochemical promiscuity in artificial transcriptional activators. *J Am Chem Soc* **2005,** 127, (36), 12456-7.
158. Liu, B.; Alluri, P. G.; Yu, P.; Kodadek, T., A potent transactivation domain mimic with activity in living cells. *J Am Chem Soc* **2005,** 127, (23), 8254-8255.
159. Casey, R. J.; Desaulniers, J. P.; Hojfeldt, J. W.; Mapp, A. K., Expanding the repertoire of small molecule transcriptional activation domains. *Bioorg Med Chem* **2008**.
160. Kwon, Y.; Arndt, H. D.; Mao, Q.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J Am Chem Soc* **2004,** 126, (49), 15940-1.
161. Rowe, S. P.; Casey, R. J.; Brennan, B. B.; Buhrlage, S. J.; Mapp, A. K., Transcriptional up-regulation in cells mediated by a small molecule. *J Am Chem Soc* **2007,** 129, (35), 10654-5.
162. Lawinger, P.; Venugopal, R.; Guo, Z. S.; Immaneni, A.; Sengupta, D.; Lu, W. Y.; Rastelli, L.; Carneiro, A. M. D.; Levin, V.; Fuller, G. N.; Echelard, Y.; Majumder, S., The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells. *Nature Medicine* **2000,** 6, (7), 826-+.
163. Schoenherr, C. J.; Anderson, D. J., The Neuron-Restrictive Silencer Factor (Nrsf) - a Coordinate Repressor of Multiple Neuron-Specific Genes. *Science* **1995,** 267, (5202), 1360-1363.
164. Immaneni, A.; Lawinger, P.; Zhao, Z. Y.; Lu, W. Y.; Rastelli, L.; Morris, J. H.; Majumder, S., REST-VP16 activates multiple neuronal differentiation genes in human NT2 cells. *Nucleic Acids Research* **2000,** 28, (17), 3403-3410.
165. Toniatti, C.; Bujard, H.; Cortese, R.; Ciliberto, G., Gene therapy progress and prospects: transcription regulatory systems. *Gene Ther* **2004,** 11, (8), 649-57.
166. Weikert, S.; Papac, D.; Briggs, J.; Cowfer, D.; Tom, S.; Gawlitzek, M.; Lofgren, J.; Mehta, S.; Chisholm, V.; Modi, N.; Eppler, S.; Carroll, K.; Chamow, S.; Peers, D.; Berman, P.; Krummen, L., Engineering Chinese hamster ovary cells to maximize sialic acid content of recombinant glycoproteins. *Nat Biotechnol* **1999,** 17, (11), 1116-21.

167. Fussenegger, M., The impact of mammalian gene regulation concepts on functional genomic research, metabolic engineering, and advanced gene therapies. *Biotechnol Prog* **2001,** 17, (1), 1-51.
168. McDaniel, R.; Weiss, R., Advances in synthetic biology: on the path from prototypes to applications. *Curr Opin Biotechnol* **2005,** 16, (4), 476-83.

CHAPTER 2

THE DEVELOPMENT OF ACTIVATOR ATFS[*]

A. Abstract

Malfunctioning transcriptional regulators are at the heart of many human diseases. For example, the transcriptional activator p53 is mutated in over half of human cancers preventing it from activating vital tumor suppression genes.[4] Currently, gene therapy approaches to adenovirally deliver p53 to tumor cells are undergoing clinical trials for the treatment of head and neck cancer (Advexin, Introgen Therapeutics).[5] However until further advances in gene therapy are made, these approaches have limited efficiency due to the problems associated with cellular delivery and immune response.[6, 7] Thus, the discovery of synthetic molecules that functionally replace transcriptional regulators, artificial transcription factors or ATFs, hold great promise for mechanistic studies and ultimately for therapeutic applications. Historically it has been challenging to develop ATFs that activate transcription (activator ATFs) that function with the specificity and activity levels of their natural counterparts. One difficulty has been that the binding profile of endogenous activators is not well understood; the physiologically relevant targets of activators in the transcriptional machinery (so-called coactivators) have for the most part not yet been identified. Recently, however, several genetic and biochemical studies provide strong evidence for three proteins, Med15(Gal11), Tra1 and Taf12, as legitimate binding partners of activators, although definitive in vivo evidence is still lacking (discussed more completely in Chapter 4). Further, the functional relevance or the

[*] Portions of this chapter were taken from Wu, Z; Belanger, G.; Brennan, B.B.; Lum, J.K.; Minter, A.R.; Rowe, S.P.; Plachetka, A.; Majmudar, C.Y.; Mapp, A.K. *J. Am. Chem. Soc.* **2003**, *125*, 12390-1 and Majumudar C.Y.; Lum, J.K.; Prasov, L.; Mapp, A.K. *Chem. Biol.* **2005**, *12* 313-321. Experiments in Figure 2.4 were performed by Aaron Minter and Annette Plaschetka. Experiments in Figure 2.5 were performed in collaboration with Brian Brennan. Experiments in Figures 2.7-2.10 were performed in collaboration with Jenifer Lum. I carried out the Experiments in Figures 2.11 and 2.12.

thermodynamic parameters for the interactions between activators and Med15(Gal11), Tra1 or Taf12 have not been completed elucidated.

In this Chapter, we use a synthetic peptide library to identify ligands that interact with the yeast Mediator protein Med15(Gal11). We find that ligands with micromolar dissociation constants are able to activate transcription to varying levels in yeast reporter gene assays. The difference in activity is attributed to the distinct binding sites that are targeted on Med15(Gal11) and not due to the affinity of the ligands for Med15(Gal11). In addition, we show that in contrast to other natural and artificial activators, the function of these ligands is solely dependent on the presence of Med15(Gal11), setting the stage for the development of cell-type and tissue-specific activator ATFs.

B. Introduction

B.1 The Mediator complex

Mediator was originally discovered as a stable complex of proteins that could "mediate" the function of activators in vitro.[8] It was also shown that Mediator formed a stable complex with RNA polymerase II, leading researchers in the late 1990s to hypothesize that Mediator was the bridge required for activators to interact with RNA polymerase II (Figure 2.1).[9, 10] In support of this, several genetic and biochemical experiments showed that activators interacted with a subset of Mediator components (vide infra).[11] However, the exact nature of the interaction between DNA-bound regulators and Mediator that results in gene upregulation was not completely understood.

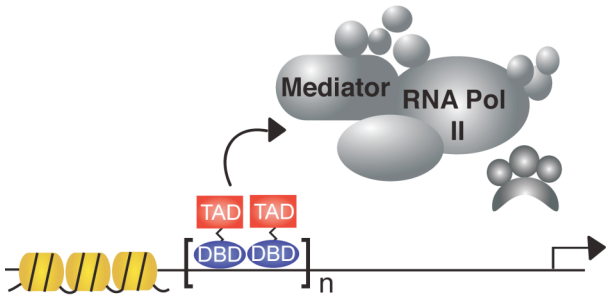


Figure 2.1. Model of transcription initiation. Activators localize on DNA (black line) and interact with a variety of protein complexes (in gray) including Mediator to recruit RNA pol II to the gene promoter.

Mediator is functionally conserved from yeast through humans, consisting of 21-25 proteins.[12] Mediator was originally identified in yeast as a complex consisting of ~20 proteins, and extensive cross-species comparisons and bioinformatics analyses have

detected metazoan counterparts for almost all yeast Mediator subunits suggesting that a universal "Mediator complex" appeared early during eukaryotic evolution (Table 2.1).[1, 13] Thus yeast has proven to be a great model system to investigate the functions of Mediator and transcription in general. It contains most of the functionally important transcriptional machinery components yet, a less complicated network of signaling found in metazoan systems. Further, due to the ease with which genes can be manipulated in yeast, *S. cerevisiae* has proven to be more tractable than human cell culture for basic research aimed at understanding fundamentally important principles for activator function.

	S. cerevisiae	***H. sapiens***	***D. melanogaster***	**Location**
Med1	Med1	Trap 220/Arc205/PBP	Trap 220	Middle
Med2	Med2			Tail
Med3	Pgd1/Hrs1/ Med3			Tail
Med4	Med4	Trap 36/Arc36/p34	Trap 36	Middle
Med5	Nut1			Middle
Med6	Med6	hMed6/Arc33/p32	Med6	Head
Med7	Med7	hMed7/Arc34/p36	Med7	Middle
Med8	Med8	Arc32	Arc32	Head
Med9	Cse2/Med9		CG5134	Middle
Med10	Nut2/Med10	hNut2/hMed10	Nut2	Middle
Med11	Med11	Hspc296	Med21	Head
Med12	Srb8	Trap 230/Arc240	Kto	CDK/Srb
Med13	Ssn2/Srb9	Trap 240/Arc250	Skd/pap/Bli	CDK/Srb
Med14	Rgr1	Trap 170/Arc150/p110	Trap 170	Tail
Med15	Gal11	Arc105/Pcqap/TIG-1	Arc105	Tail
Med16	Sin4	Trap 95/p96b	Trap 95	Tail
Med17	Srb4	Trap 80/Arc77/p78	Trap 80	Head
Med18	Srb5	p28b	p28/CG14802	Head
Med19	Rox3	Lcmr1	CG5546	Head
Med20	Srb2	hTrfp/p28a	Trfp	Head
Med21	Srb7	hSrb7 /p21	Trap 19	Middle
Med22	Srb6	Surf5	Med24	Head
Med23		Trap 150β/Arc130/hSur2	Trap 150β	N/A
Med24		Trap 100/Arc100	Trap 100	N/A
Med25		Arc92/Acid1	Arc92	N/A
Med26		Arc70	Arc70	N/A
Med27		Trap 37	Trap 37	N/A
Med28		Fksg20	Med23	Head
Med29		hIntersex	Intersex	Head
Med30			Trap 25	Head
Med31	Soh1	hSoh1	Trap 18	Middle
CDK8	Srb10/Ssn3	CDK8	CDK8	CDK/Srb
CycC	Srb11/Ssn8	CycC	CycC	CDK/Srb

Table 2.1. Components of the Mediator complex.[1] Components of the Mediator are functionally conserved from yeast through humans. The Mediator is composed of three modules, the head, middle and tail. (see Figure 2.1). Mediator is also often associated with the CDK/Srb sub-module.

Although originally hotly debated, the components of Mediator are currently well agreed upon, but the precise function of each individual protein is still unknown. Cryo-EM studies with purified yeast Mediator depicted a modular complex consisting of a head that interacted with RNA polymerase II (8 proteins), a middle module (8 proteins) and a tail region (5 proteins), which was predicted to interact with TADs (Figure 2.2).[14] The Mediator was also shown to associate with the CDK (cyclin-dependent kinase) module (4 proteins) in a few specific contexts.[15] Studies by Asturias and coworkers revealed that Mediator undergoes conformational changes when it interacted with RNA polymerase II and Tijian and coworkers showed that a similar conformational change in the human Mediator was elicited by activators.[2, 16, 17] Together these studies suggest that Mediator serves as a conformational relay system between DNA bound regulators and the transcriptional machinery. Specifically, it is in the tail module that most of the putative coactivators reside, with subunits Med14(Rgr1), Med3(Pgd1), Med2 and Med15(Gal11) identified in a number of genetic screens as effectors of positive transcriptional regulation.[18-20] Additionally, head domain component Med17(Srb4) and components of the tail module including Med15(Gal11) have been implicated as direct targets of activators using in vitro crosslinking and pull-down assays.[21-24]

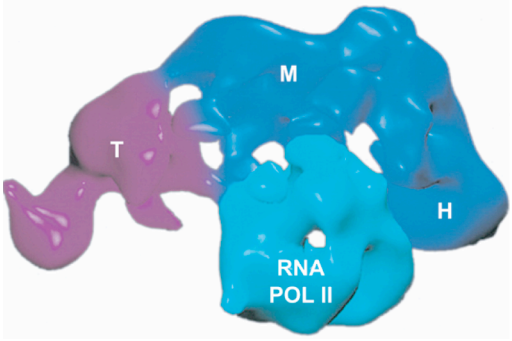


Figure 2.2
The Mediator complex.
Model of the yeast Mediator in complex with RNA Polymerase II determined from cryo-EM images.[2] Three domains have been identified, termed Head (H), Middle (M) and Tail (T).

In an elegant 'activator by-pass' study aimed at elucidating which Mediator proteins might be direct binding partners of activators, Ptashne and coworkers localized various components of Mediator to DNA by fusing them to DNA binding proteins and then evaluated their ability to activate transcription in this context.[3, 25] Med15(Gal11) localized to DNA was not only able to activate transcription two orders of magnitude higher than any other Mediator component, but it was also able to activate to similar

levels as the potent yeast activator Gal4. DBD fusions of other components Med2, Med20(Srb2) also resulted in activation, although to lesser levels (Table 2.2). These findings illustrated the possibility that artificially localizing Med15(Gal11) to DNA by creating ligands for Med15(Gal11) could potentially create activator ATFs and provided some evidence that Med15(Gal11) is a likely target of endogenous activators.

LexA fusions	Activities
LexA	0.9 ± 0.4
Med1	3.0 ± 1.0
Med2	39.0 ± 9.0
Med4	7.0 ± 2.0
Med6	4.0 ± 1.0
Med12	0.5 ± 0.3
Med15	**1230.0± 93.0**
Med17	3.0 ± 1.0
Med20	15.0 ± 3.0
Med21	0.5 ± 0.5
CDK8	8.0 ± 2.0
Spt20	29.0 ±3.0
Spt3	7.0 ± 1.0
Spt8	8.0 ± 0.5
Ahc1	5.0 ± 1.0
Tra1	30.0 ± 7.0
Taf1	3.0 ± 2.0
Taf30	1.0 ± 0.3
Taf9	18.0 ± 3.0
Taf10	23.0 ± 4.0

Table 2.2
Activator by-pass experiments using non-classical activators.[3] Various components of the yeast transcriptional machinery were fused to the LexA DBD and the level of activation elicited by each fusion was evaluated using β-galactosidase assays in yeast.

B.2 Med15(Gal11) an activator target

Med15(Gal11) is a 120 kDa protein that resides in the tail domain of the yeast Mediator complex. It was originally identified as being necessary for proper galactose utilization in yeast and postulated to be important in the regulation of galactose-inducible genes.[26-28] Subsequently a number of genetic and biochemical studies have implicated it as a likely target of DNA-bound transcriptional activators, including the activator Gal4 that regulates genes responsible for galactose metabolism.[29] However, the function of Med15(Gal11) remains unknown. The N-terminus of Med15(Gal11) has been computationally predicted to contain a GACKIX domain with sequence homology to the coactivators ARC105 and CREB binding protein (CBP), leading researchers to suggest that ARC105 is the metazoan homolog of yeast Med15(Gal11).[30] However, the rest of Med15(Gal11), apart from a glutamine rich region, is not homologous to ARC105,

raising questions about their functional relationship. More recently, the solution structure of the N-terminus (residues 2-93) of yeast Med15(Gal11) has also revealed structural similarity to GACKIX domains of CBP and ARC105.[31] The C-terminus of Med15(Gal11) has been observed to associate with Tfa1 and Tfa2, subunits of the general transcription factor TFIIE as well as stimulate phosphorylation of the carboxy-terminal domain of RNA Polymerase II by TFIIH (Figure 2.3).[32-34]

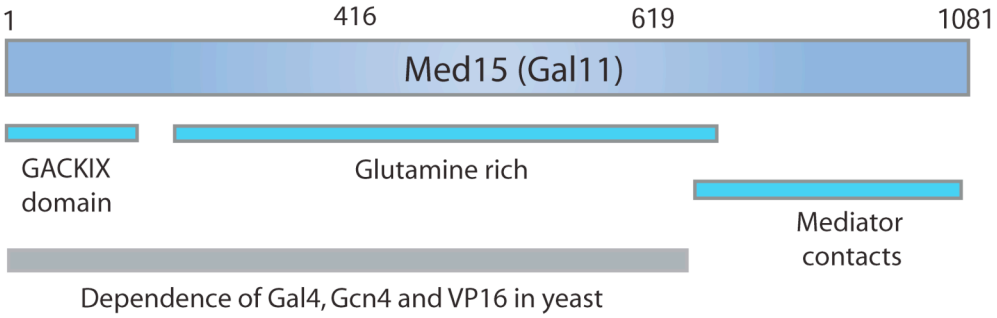


Figure 2.3 Protein interaction domains of Med15(Gal11). Med15(Gal11) is 1081 amino acid yeast protein that has been shown to interact with the Mediator and transcriptional activators. A conserved GACKIX domain similar to that found in the mammalian coactivator CBP is located at the N- terminus.

Several recent studies have implicated Med15(Gal11) as a target of both natural and artificial transcriptional activators. Genetic experiments show that Med15(Gal11) is important for transcription of *GAL* and *MATα* genes and necessary for activated transcription by the endogenous activators Gal4, Gcn4, Swi5, Msn2, VP16 and Met4.[28, 35-40] Further, direct binding experiments and integrated yeast reporter assays by the Kodadek and Kim laboratories showed that different fragments of Med15(Gal11) interacted with a number of natural TADs in vitro and deletions of these fragments in yeast caused a decrease in activation.[18, 19, 29] For example, Med15(1-351) has been shown to interact with the Gal4 TAD in vitro.[29] Med15(Gal11) has also shown to be the primary target of the artificial peptidic TAD XL_Y.[41, 42] Since we initiated our studies, in vitro crosslinking studies by the Hahn group have identified Med15(Gal11) as one of three targets of the Gal4 and Gcn4 TADs from yeast whole cell extracts.[21, 22] More recently, the Lehming group has shown using a split-ubiquitin assay in yeast that Gal4 and Gcn4 interact with Med15(Gal11).[43]

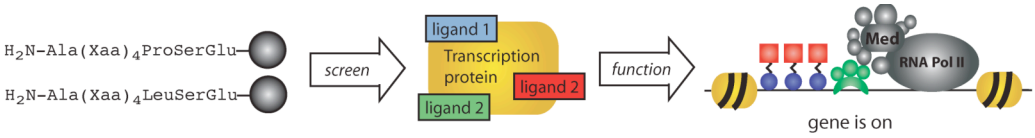
Given the evidence suggesting that Med15(Gal11) is an activator target, we hypothesized that ligands for this protein would serve as tools to investigate the features necessary to create novel transcriptional activators with robust cellular activity. In particular, we hypothesized that localizing Med15(Gal11) to DNA with synthetic ligands would lead to the creation of activator ATFs. Further, we could use these ligands to investigate the affinity and orientation requirements for recruiting Med15(Gal11) and the relationship of these factors to transcriptional activation. For these purposes, we utilized a screening strategy to identify ligands that bound different surfaces of the protein.

B.3. Synthetic peptide library screen

In order to identify ligands for Med15(Gal11) a combinatorial peptide library was screened against Med15(Gal11) in vitro. The design of the library was based on the minimal sequences from endogenous activators that activated transcription to high levels. Natural TADs often contain repeats of short sequence motifs, with each motif able to activate transcription on its own, albeit to a lesser extent than the combination. One of the shortest such motifs is found in the VP16 TAD, which contains four repeats of a consensus eight amino acid sequence.[44] It has been shown that the fusion of eight amino acids from this TAD (DFDLDMLG) to a DBD creates a strong activator. Thus, it was decided to create an octapeptide library with four residues randomized, creating a balance between library size and sequence coverage. While a peptide library was chosen for the ease of creating a library with diverse chemical functionally and detection of positive hits, a similar approach could be later extended to screen a nonbiopolymer-based library. Two libraries, AXXXXPSE and AXXXXLSE (with X=19 of the 20 natural amino acids, without lysine), were designed for the screen, giving a theoretical total of 260,642 unique sequences. In one library Leu, Ser and Glu were incorporated at the carboxy terminus. These amino acids are commonly found in the amphipathic class of endogenous transcriptional activators.[45] The inclusion of proline in the second library was used to bias the library against peptide sequences that would form helices, to obtain sequences distinct from natural activators. Former students Dr. Aaron Minter and Dr. Annette Plaschetka used an ELISA-based screen to obtain 37 ligands from these libraries that bound to bacterially expressed Med15(186-619) (Figure 2.4).

The ligands isolated fall into 3 main sequence categories (randomized region): hydrophobic, amphipathic and positively charged, indicating that they might be targeting different surfaces of Med15(Gal11). To verify this, coworkers Dr. Zhiqian Wu and Dr. Brian Brennan performed fluorescence polarization-based competition binding experiments with the ligands identified in the screen, showing that the ligands isolated were binding to different surfaces of Med15(Gal11). Further, they found that these ligands bound Med15(Gal11) in the low micromolar range similar to other TADs such as VP2 and XL_Y suggesting that some of these should activate transcription.

a)



b)

Library 1: AXXXXLSE						Library 2: AXXXXPSE			
1.	NITY	9.	RARV	17.	QRRV	22.	SFHR	30.	VLGW
2.	ETVS	10.	SHRT	18.	SLNR	23.	FWLF	31.	IWLF
3.	RRGV	11.	GRRA	19.	EDRR	24.	NQQW	32.	YFEV
4.	TWRR	12.	RVYR	20.	RWTI	25.	DLTM	33.	VHPV
5.	HTRH	13.	AHRR	21.	NLRT	26.	FYRN	34.	FNWE
6.	RDRT	14.	SRHR			27.	MHFP	35.	FTLW
7.	FARR	15.	RQRN			28.	HYYY	36.	LWFF
8.	QRRT	16.	YHRT			29.	PVLG	37.	IFFF

Figure 2.4 Synthetic peptide library screen.
a) ELISA-based screening strategy
b) Ligands that bound different surfaces of Med15(Gal11). Ligands fall into at least three categories: amphipathic (yellow), positively charged (blue) and hydrophobic (orange).

B.4 Activator ATFs

Given the ligands target Med15(Gal11) with a similar affinity as natural and unnatural TADs, we hypothesized that when fused to a DBD, leading to the creation of activator ATFs, they should activate transcription in cells to high levels. Full length endogenous activators in an optimal context can upregulate gene expression 10-5000 fold over basal levels. In order to successfully restore the function of malfunctioning endogenous activators, it is likely that activator ATFs will have to activate transcription

to natural activator levels. Additionally, since the ligands were obtained based on their ability to bind Med15(Gal11), they could potentially function by interacting with a single transcriptional machinery component. Activators ATFs that function through specific interactions would be extremely useful therapeutically to precisely target diseased cells or tissues.

C. Results

C.1 Activity in yeast

To assess the ability of the Med15(Gal11) ligands to function as transcriptional activation domains, Garrette Belanger and Dr. Jenifer Lum made yeast expression plasmids that encoded for fusions of these ligands to the N-terminus of the LexA DBD and performed qualitative X-gal plate assays in a yeast reporter assays. They identified 3 ligands, #17, #28 and #32 that activated transcription to levels similar to the positive control VP2. Quantitative evaluation of ligands #17 and #28 in yeast β-galactosidase assays revealed that #28 activated to 60% of the activity levels seen with VP2. (Figure 2.5)

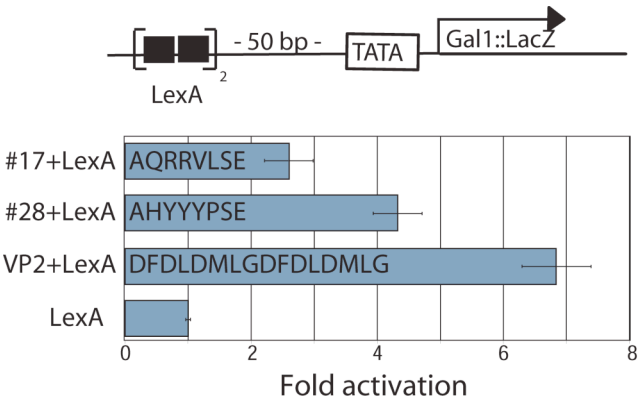


Figure 2.5 Med15(Gal11)Ligands in yeast.
Ligands+LexA fusion were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the ligand+LexA fusion by the activity of LexA alone. Error is reported as SDOM from triplicate experiments.

A surprising observation from this study was that contrary to previously held dogma, a correlation between the affinity of each ligand to its transcriptional target Med15(Gal11) and its activation potential was not observed.[46, 47] For example the artificial TAD XL_Y binds Med15(Gal11) with a K_D of 2.2 ± 0.4 μM and activates transcription >500 fold over basal, however #17 with a similar affinity for Med15(Gal11) (K_D 2.3 ± 0.4 μM) activates transcription only ~2.5 fold (Figure 2.5 and data not shown). We hypothesized that this could be due to the ligands targeting different surfaces of

Med15(Gal11). Specifically, some binding sites on Med15(Gal11) might be obscured in the context of Mediator and further some ligands might disrupt key Med15(Gal11)-Mediator contacts; hence ligands that target those sites are unable to activate transcription in the context of the cell.

Significantly, given the low micromolar Med15(Gal11) affinities observed for the different ligands, the endogenous activator VP2 and the potent artificial activator XL_Y, it may not be necessary to find high affinity coactivator ligands to activate transcription in the future. Indeed, it may be sufficient to find molecules that target a transcriptionally relevant site with moderate affinity. Subsequently, the Schepartz and Kodadek laboratories have also reported similar conclusions where they found that the affinity of ligands that bind different surfaces of the coactivator CBP does not correlate with the activity of these ligands in mammalian cell-based transcription assays.[48, 49]

C.2 Optimization of ligand activity

The cellular activity of the activator ATFs constructed from the Med15(Gal11) ligands and the LexA DBD was well below the 10-5000 fold levels seen with sequences found in endogenous activators.[50] The low activity of these molecules ultimately limits their utility as activator ATFs and we decided to investigate different strategies to improve their activation potential.

The Med15(Gal11) ligands were obtained by screening for an interaction with isolated Med15(Gal11) in vitro in the context of a resin-bound peptide. In the cell however, Med15(Gal11) is in a complex of proteins within the Mediator and certain binding sites might be obscured. Thus, it is conceivable that the orientation of the ligands is not optimal and we need to investigate different options to find favorable conditions to recruit Med15(Gal11). There are several factors such as the ability of the DBD to effectively project the ligand away from DNA and the accessibility of Med15(Gal11) from a particular position on the promoter that might impact how well Med15(Gal11) is recruited.

C.2.a The role of the DBD

Although natural transcriptional activators are modular proteins and transcriptional activation domains typically function independent of the identity of the DNA binding domain, the level of transcription elicited in each context varies.[51-53] The

ability of natural TADs to interact with multiple protein targets may account for this functional flexibility. While changing the DBD and hence the projection of the TAD might affect interactions with a subset of transcription protein targets, it is unlikely to affect all. We hypothesized that since the ELISA-based screen was done with the ligands containing a free amino-terminus, a similar orientation might be required for interaction with Med15(Gal11) in cells. For this purpose, we fused them to the N-terminus of the LexA DBD. However, only a small fraction activated transcription and to only low levels. Modeling studies of LexA by Kaptein and coworkers indicated that the amino terminus of LexA is in close proximity to the DNA, raising the possibility that amino terminal ligand-fusions may not be optimal to project the ligand away from DNA for Med15(Gal11) recruitment.[54, 55] To test this hypothesis ligand #28 was fused to the C-terminus of the LexA DBD and evaluated for activation in yeast. Unfortunately, there was negligible improvement in the ability of this ligand to activate transcription as quantitated by measuring β-galactosidase activity (data not shown). Thus, in the context of the LexA DBD, #28 was able to activate transcription only to low levels.

Next, we decided to investigate a different DBD derived from the *S. cerevisiae* protein Gal4, Gal4(1-147). The first 65 residues of Gal4 form a Cys_6-Zn cluster that binds to DNA and the subsequent 30 residues mediate dimerization through helix-helix interactions that enhance sequence-specific DNA binding.[56-58] The amino acids 97-147 serve as a linker region of undefined structure between the TAD and the DBD (Figure 2.6).[57] A plasmid was constructed encoding ligand #28 fused to the carboxyl terminus of Gal4(1-147). In this context, ligand #28 is displayed in the opposite orientation (carboxy terminus free) relative to the original LexA fusion.

Figure 2.6 Domains of Gal4. The yeast transcriptional activator, Gal4 has 3 major domains, DBD (1-65), dimerization (65-94) and two TADs (148-196 and 768-881) Regions 97-147 comprise a linker of unknown structure.

The Gal4(1-147)-based activator was more potent than the #28+LexA fusion protein, with 14.5-fold activation compared to the DBD alone (Figure 2.7). The change in orientation from the LexA to Gal4 DBD is not the only factor contributing to the increase in function. Attachment of #28 to the carboxy terminus of the DBD Gal4(1-100) produced an activator approximately 4-fold less active than Gal4(1-147)+#28. Consequently, the additional linker residues present in the Gal4(1-147)+#28 fusion protein likely play a role in the activity increase, projecting the ligand from the DNA more effectively and providing more favorable conditions for the Med15(Gal11) interaction. A similar reliance upon linkers has been noted with natural transcriptional activation domains such as ATF14 and VP2, both acid-rich TADs derived from the potent viral coactivator VP16 as well as other nonnatural TADs. [51, 52, 59]

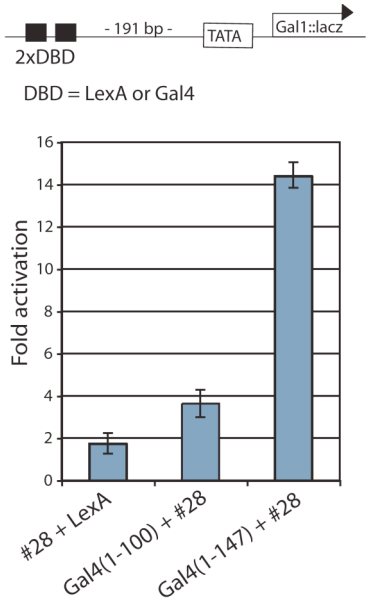


Figure 2.7
DNA binding domain dependence of #28. #28 fused to the amino terminus of LexA or the carboxy terminus of Gal4(1-00) and Gal4(1-147) was evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the DBD+#28 fusion by the activity of the DBD alone. Error is reported as SDOM from triplicate experiments.

C.2.b Binding sites in promoter

The position of the DNA binding site relative to the transcription start site has also been previously shown to affect how well an activator upregulates transcription.[25, 60-62] To probe this, activator ATF #28 was tested in yeast strains bearing 2 binding sites either 50 bp or 191 bp upstream of the TATA box. The extent to which transcription was affected by the binding site position varied with the DNA binding domain employed in the ATF. For #28+LexA, function was better at a binding site 50 bp from the TATA box, with fold activity dropping by half when moved to 191 bp. The opposite trend was observed with Gal4(1-147) as the DBD (Figure 2.8). In that case, lower levels of

activation were obtained when the binding sites were closer to the TATA box. These latter results parallel the activity of the Med15(Gal11) protein itself when it is employed as an activator ATF. The fusion protein Gal4(1-147)+Med15(Gal11) is a potent activator at the 191 bp distance; at 50 bp, however, activity drops by half.[25] The results obtained with Gal4(1-147)+#28 are suggestive of a mechanism by which #28 specifically recruits Med15(Gal11) to DNA in an orientation analogous to the positioning in the Gal4(1-147)+Med15(Gal11) fusion protein.

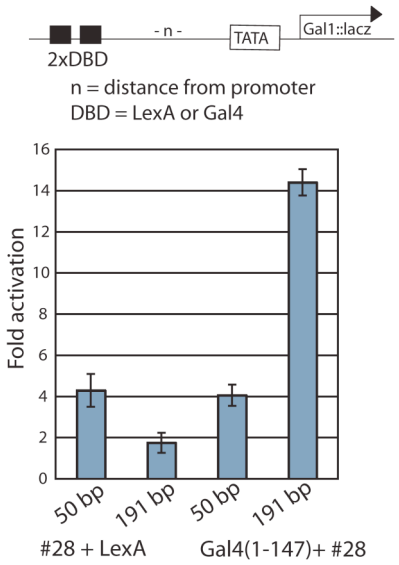


Figure 2.8 Promoter binding site dependence of #28.
#28 + LexA or Gal4(1-147) +#28 was evaluated in β-galactosidase assays in yeast at different binding sites in the promoter. Fold activation was determined by dividing the activity of the #28 DBD fusion by the activity of the DBD alone. Error is reported as SDOM from triplicate experiments.

C.2.c Increasing the activity of the other ligands

Given the increase in activity obtained for #28, we investigated if we could apply these criteria to other ligands. We attached the two other ligands previously identified to be active when fused to LexA, ligands #17 and #32, to Gal4(1-147) and evaluated their activity in yeast strains bearing 2 Gal4 sites 191 bp from the TATA box. In the case of #17, only 2-fold activity relative to the DBD alone was observed, comparable to the results obtained with the LexA DBD. Similar results were obtained when #17 was fused to Gal4(1-100); in addition, moving the binding sites closer to the transcriptional start site did not provide an increase in activity (data not shown). In contrast, ligand #32 exhibited quite modest activity when fused to the amino terminus of LexA (1.4-fold) but the activity increased to 11.5-fold when it was attached to Gal4(1-147), comparable to the activity of #28 (Figure 2.9). Overall, these experiments provided two artificial activators with improved functional profiles, Gal4(1-147)+#28 and Gal4(1-147)+#32. Interestingly,

both these ligands were from the synthetic library with proline in the third position that was designed to isolate molecules with sequences distinct from natural activators, suggesting that they function using a distinct mechanism (vide infra).

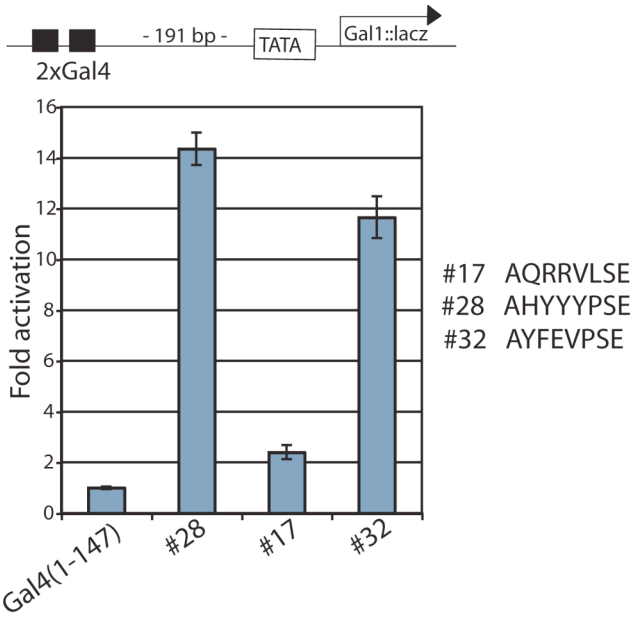


Figure 2.9 Optimized conditions for Med15(Gal11) ligands. Gal4(1-147)+ligand fusions were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the ligand Gal4 fusion by the activity of the Gal4 DBD alone. Error is reported as SDOM from triplicate experiments.

C.3 Specificity

Since the ligands were obtained based on their ability to bind Med15(Gal11) in vitro, we were eager to see if they indeed targeted Med15(Gal11) in the cellular context. Another important question we sought to answer was whether the transcriptional activity of the ligands was solely due to the interaction with Med15(Gal11). To explore this, β-galactosidase assays were carried out in a yeast strain in which Med15(Gal11) had been deleted from the genome. This experiment is possible because Med15(Gal11) is not an essential protein, although yeast bearing this alteration exhibit a slow growth phenotype. [26] Gal4(1-147)+#28 and Gal4(1-147)+#32 nearly undergo a complete loss of function in yeast strains bearing no Med15(Gal11) as compared to the strain containing Med15(Gal11) (Figure 2.10). This was in contrast to the positive control, Gal4(1-147)+ATF14, a sequence taken from the potent viral coactivator VP16, that showed only a 2-fold decrease in activation levels. Although ATF14 is known to interact with Med15, it has several additional putative targets in the transcriptional machinery, and thus it is not surprising that its function is attenuated rather than abrogated in the absence of Med15(Gal11). [63]

This point was further investigated by carrying out the same set of experiments using a yeast strain in which the central region of Med15(Gal11)(residues 186–619) had been deleted.[64] This mutation minimized the deleterious phenotype of the Med15(Gal11) delete strain and enabled us to test if the binding sites for ligands #28 and #32 were in this region because the original binding screen was carried out with this fragment. Gratifyingly, nearly identical results were obtained, with #28 and #32 showing little or no activity in this strain, while the fold activation of ATF14 was similar to the strain with Med15(Gal11) present. Taken together, these data suggest that both #28 and #32 are dependent upon a binding interaction with Med15(Gal11) for upregulating transcription.

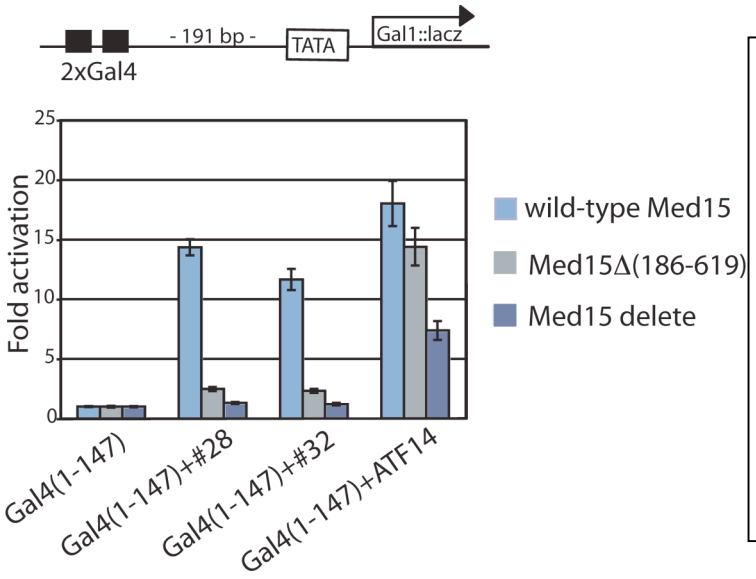


Figure 2.10 Specificity of Med15(Gal11) ligands. #28 and #32 were attached to Gal4(1-147) and tested for activity in yeast strains that contained Med15(Gal11) or Med15(Δ186-619) or without Med15(Gal11). Fold activation was calculated by dividing the activity of the particular construct by the DBD alone. Error is reported as SDOM.

We were also interested in evaluating our ligands in different cell types. In the long term, the development of activator ATFs that function in certain cell types will be therapeutically useful for specifically targeting diseased cells (for example see [65, 66]). Intriguingly, acid-rich TADs such as Gal4 or VP16, commonly used for activator ATF construction, function in all eukaryotes tested, from yeast through humans.[67-69] Despite differences in RNA polymerase II holoenzyme composition, there is evidently significant conservation across species with regard to activator targets. It has been challenging, however, to identify metazoan homologs of yeast Med15(Gal11). Recently, structural and computational evidence for homology between the amino terminus of Med15(Gal11) and the mammalian protein ARC105 was reported.[30, 31] In addition, the two proteins contain a glutamine-rich stretch of amino acids. However, the region of Med15(Gal11) used in our

original binding screen (residues 186–619) shares little sequence similarity with ARC105 or any other identified metazoan protein, and we thus anticipated that activators that function through an interaction with this region would not be able to function in mammalian cells. To investigate this, a plasmid encoding the most active of the peptides (#28) was transiently transfected into human embryonic kidney cells (HEK293 cells) along with a reporter plasmid bearing five Gal4 binding sites within an E1b promoter upstream of a SEAP reporter gene following standard protocols. As a positive control, we also examined the activity of a VP16-derived transcriptional activation domain fused to Gal4(1-147). No activation by ligand #28 was observed while the VP16-derived TAD functioned well in this context (Figure 2.11). Finally, to verify that these activators were in fact being expressed in the HEK293 cells, we performed immunofluorescence staining of the cells (Figure 2.12). Taken together, the experiments in mammalian cells reinforce the yeast results, indicating that #28 is dependent upon Med15(Gal11) for function.

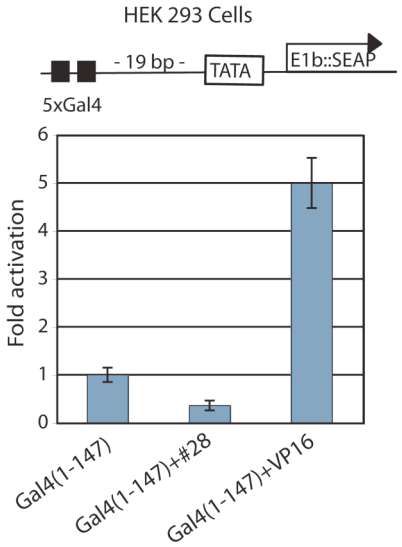


Figure 2.11 #28 in mammalian cells. AHYYYPSE (#28) was attached to Gal4(1-147) and tested for activity HEK 293 using secreted alkaline phosphatase assay. Fold activation was calculated by dividing the activity of the Gal4(1-147)+#28 by the DBD alone. Error is reported as SDOM.

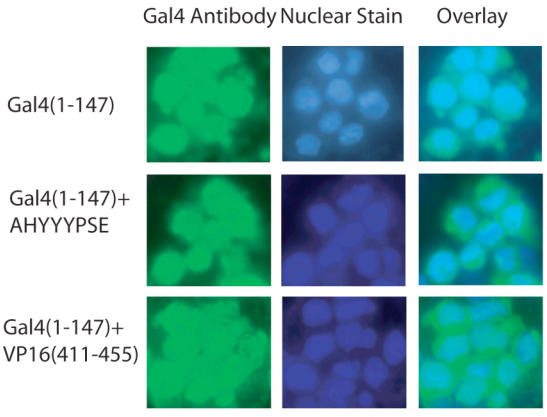


Figure 2.12 Immunofluorescence staining of HEK293 cells. HEK 293 cells expressing Gal4(1-147) constructs were probed with an α Gal4 antibody to detect expression of Gal4 (1-147) fusions. Expression of Gal4 is indicated in green. The nucleus is stained blue. The overlay show that >95% of cells expressed Gal4(1-147).

D. Conclusions and future directions

We have shown here that ligands identified using an in vitro binding screen for Med15(Gal11) serve as artificial TADs that function in yeast cell-based transcription assays. Further, low micromolar affinity for Med15(Gal11) is sufficient to recruit Med15(Gal11) and activate transcription. Contrary to the model in the literature, ligand-target binding affinity did not correlate with function. However, differences in binding sites for activators with similar affinities for Med15(Gal11) seemed to play a role in the ability of the different ligands to activate transcription in yeast.

When Med15(Gal11) itself is directly attached to LexA, localizing to it DNA, it results in >1000 fold levels of activation. However none of the ligands when attached to LexA or Gal4 DBDs displayed such high levels of activity. Perhaps, the fragment of Med15 we screened against did not contain the appropriate binding sites to recruit Med15(Gal11) in the same orientation. Additional studies aimed at discovering the binding sites of these and other Med15-targeting TADs in the future will provide insight into the relationship between binding site and potency (Chapter 4). Specifically, robust endogenous activators may have evolved to target privileged binding sites that recruit Med15(Gal11) in the optimal orientation and studies that create molecules that target these sites will lead to new generations of more potent activator ATFs.

We also show that the Med15(Gal11) ligands differ from typical natural or artificial TADs. In contrast to natural activators, as experiments described here revealed that for at least two of the artificial TADs, a Med15(Gal11) interaction is required and sufficient for transcription function. In the future, activator ATFs engineered with the Med15-specific ligands combined with ligands targeting other protein targets might provide a simple strategy to improve potency and will be outstanding tools for probing the mechanistic origins of transcriptional synergy (Chapter 3).

More practically, since the screening strategy provides TADs that function through binding interactions with individual transcriptional machinery proteins, targeting other cell-type-specific or organism-specific proteins provides a mechanism for the creation of activators whose functional specificity extends beyond that imposed by the DNA binding domain. Finally, as the screening strategy is equally applicable to

combinatorial libraries of small molecules, these results provide a framework for building tunable, uniquely specific small molecule transcriptional regulators.

E. Experimental

General methods

All techniques used for yeast manipulations were carried out in accordance with standard protocols.[70] All other general molecular biology techniques were carried out as described. The yeast strains used for testing the LexA and Gal4 fusions of ligands from the synthetic library screen include LS41 [JPY9::ZZ41, *Mat*α *his3*Δ*200 leu2*Δ*1 trp1*Δ*63 ura3-52 lys2*Δ*385 gal4 gal80* URA::pZZ41] (pZZ41 contains two LexA binding sites upstream of the TATA box), JPY52::JP185, *Mat*α *his3*Δ*200 leu2*Δ*1 trp1*Δ*63 ura3-52 lys2*Δ*385 gal4*Δ*11 med15*::LYS, URA::pJP185 (pJP185 contains two Gal4 binding sites 50bp upstream of the TATA box), ZL2, *Mat*α *his3*Δ*200 leu2*Δ*1 trp1*Δ*63 ura3-52 lys2*Δ*385 gal4*Δ*11 med15*::TRP1, URA::pJP169 (pJP169 contains two LexA binding sites upstream of the TATA box) and JPY52::JP188, *Mat*α *his3*Δ*200 leu2*Δ*1 trp1*Δ*63 ura3-52 lys2*Δ*385 gal4*Δ*11 med15*::LYS2, URA::pJP188 (pJP188 contains two Gal4 binding sites 191bp upstream of the TATA box) (Gifts from Dr. Aseem Ansari).

For the experiments with JPY52::JP185 and JPY52::JP188 Med15(Gal11)Δ yeast strains, a low copy plasmid, ycplac111 Gal11WT expressing full length Med15(Gal11) from its native promoter was transformed into yeast (Gift from Dr. Aseem Ansari).

LexA Plasmids

The Med15 ligands were expressed as fusions to the N terminus of the LexA DBD in yeast. The expression of the protein was driven by the strong, constitutively active ADH promoter and has a His marker in yeast and ampicillin selection in *E. coli*. The parent plasmid pNLexA (high copy in yeast) used for this purpose was purchased from Origene. Fusions of all 37 ligands to LexA were performed by Dr. Jenifer Lum and Garrette Belanger. C-terminal LexA fusions were accomplished by using pHyb-Zeo from Invitrogen. These plasmids were prepared by Garrette Belanger and are also high copy in

yeast and express proteins under the control of the ADH promoter. pHyb-Zeo has a Zeocin marker for yeast (high copy) and an ampicillin selection in *E. coli*.

Gal4(1-147) plasmids

Plasmids encoding Gal4(1-147)+#17, Gal4(1-147)+#28, and Gal4(1-147)+#32 were generated from pGBKT7 (Clontech) by Jenifer Lum. It expresses fusions to the C-terminus of Gal4 and is kanamycin selectable in *E. coli*, contains a TRP marker for yeast (high copy) and expresses proteins under the control of the ADH promoter.

ATF-14 was generated from pGBKT7 by first annealing oligonucleotides encoding the peptide, (5′-AA TTC tgt ggt gat gct ttg gat gat ttt gat ttg gat atg ttg TAA-3′ and 5′-TC GA TTA caa cat atc caa atc caa atc aaa atc atc caa agc atc aga acc aca G-3′) resulting in sticky ends corresponding to the restriction sites EcoRI/SalI. The duplex oligonucleotides were phosphorylated with T4 polynucleotide kinase and ligated with T4 DNA ligase into pGBKT7 predigested sequentially with EcoRI and SalI and treated with calf intestinal phosphatase. The resulting plasmids were amplified in DH5α *E. coli* (Invitrogen), selected on LB-agar plates containing 50 μg/ml kanamycin, and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

Gal4(1-100) plasmids

Plasmids encoding Gal4(1-100)+#28 and Gal4(1-100)+#17 were generated from RJR217 using homologous recombination (Gift from Dr. Aseem Ansari). These fusions were expressed under the control of the strong β-actin promoter and the plasmid has a His marker in yeast (low copy) and has an ampicillin selection in *E. coli*. Oligos (5'-GCA TTG TTA ACA GGA TTA TTT GTA CAA GAT NNN NNN NNN NNN NNN NNN NNN NNN NNN TAA AAC ATT TGA AGT TTC CAT ACT TTT GAT ACT TTT GAA G'-3) and its complement were cotransformed into the yeast strain LS41 with the plasmid RJR217 which was predigested with Sal1 and treated with Mung Bean nuclease. The resulting plasmids incorporating #28 or #17 were amplified and subsequently extracted from yeast. The plasmids were amplified in DH5α *E. coli* (Invitrogen), selected on LB-agar plates containing 50 μg/ml kanamycin, and isolated from cultures using a QIAprep

Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

Med15(Gal11)Δ(186-619) plasmid

The YCplac111+Med15Δ(186-619) plasmid was generated from the parent YCplac111 + Ga11WT plasmid using site-directed mutagenesis (Gift of Dr. Aseem Ansari). This plasmid expresses Med15(Gal11) under the control of its native promoter, is ampicillin selectable in *E. coli* and possesses a Leu marker (low copy) for yeast selection. Briefly, two sets of oligonucleotides were designed to insert XhoI restriction sites either before or after the sequence of the region to be deleted. Site-directed mutagenesis was performed using the QuikChange Kit (Stratagene) according to manufacturer's instructions. The methylated parent plasmid was then digested with DpnI and the nicked mutagenized plasmid was amplified in SMART *E. coli* cells (Genlantis), selected on LB-agar plates containing 100 μg/ml ampicillin and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). This modified plasmid was then subjected to the same mutagenesis procedure using the second set of oligonucleotides. After insertion of both of the XhoI restriction sites, the amplified plasmid was digested with XhoI and gel purified or religated without purification with T4 DNA ligase. The resulting YCplac111+Med15Δ(186-619) plasmid was amplified in SMART *E. coli* cells and selected on LB-agar plates containing 100 μg/ml ampicillin. The new YCplac111+Med15Δ(186-619) plasmid was subsequently isolated using a QIAprep Spin Miniprep Kit (Qiagen) and the sequence was verified at the University of Michigan Core Facility.

Mammalian Gal4(1-147)+#28 plasmid

For use in the human cell experiments, a plasmid encoding Gal4(1-147)+#28 was generated from pM (Clontech) by ligation of an oligonucleotide pair encoding #28 (5′-AA TTC GGT TCT GGT GGT TCT GGT GCT CAT TAT TAT TAT CCA TCT GAA TAA-3′ and 5′-TCGA TTA TTC AGA TGG ATA ATA ATA ATG AGC AGA ACC ACC AGA ACCG-3′) into pM that had been predigested with EcoRI/SalI. The resulting plasmid was amplified in DH5α *E. coli* (Invitrogen), selected on LB-agar plates

containing 0.1 mg/ml ampicillin and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequence of the isolated plasmids was verified by sequencing at the University of Michigan Core Facility.

β-galactosidase assays

The function of ligand+LexA or Gal4+ligand (peptide, artificial or natural) fusions was examined in yeast by a quantitative liquid β-galactosidase assay in accordance with established methods.[70] Briefly, the plasmids encoding the peptide fusions and the DBD plasmid (negative control) were transformed into yeast using the LiOAc method or by electroporation (Med15(Gal11) delete strains), and transformed colonies were selected by growth on synthetic complete (SC) media containing 2% raffinose and lacking the appropriate amino acid(s) for selection. Freshly transformed colonies were used to inoculate 5 ml cultures of SC media containing 2% raffinose and lacking the appropriate amino acids. The cultures were incubated overnight at 30°C with agitation. Following incubation, these cultures were used to inoculate 5 ml cultures of SC media containing 2% raffinose, 2% galactose and lacking the appropriate amino acids that were subsequently incubated overnight at 30°C with agitation to an OD_{660} of 3–4. The yeast cells were harvested and resuspended in Breaking buffer (100 mM Tris-HCl (pH 8.0), 20% glycerol) containing the Complete Protease Inhibitors cocktail (Roche). The cells were lysed by vortexing with glass beads. A portion of the cell extract was used to measure β-galactosidase activity via incubation with *o*-nitrophenyl-β-D-galactopyranoside (1 mg/ml) in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4 \cdot 7H_2O$, and 50 mM 2-mercaptoethanol [pH 7]). The reaction was stopped by adding 1 M Na_2CO_3 and the OD_{420} was measured on a Varian Cary 300 UV-vis spectrometer. The activity reported was normalized to the total protein concentration of the extract, measured using a Bradford assay kit (Bio-Rad) with BSA as the standard.

Secreted Alkaline Phosphatase Assay

The function of Gal4(1-147)+#28 was examined in human cells using a quantitative secreted alkaline phosphatase (SEAP) assay in accordance with standard protocols. For this purpose, 4 μg of plasmid encoding Gal4(1-147)+#28, the Gal4(1-147) plasmid (pM,

negative control), or the Gal4(1-147)+VP16(411-455) fusion (pM3-VP16, positive control) were transiently transfected into an equal number of human embryonic kidney 293 cells (ATCC) using the PolyFect reagent (Qiagen) according to the manufacturer's protocol. Each transfection reaction also contained 2 µg of the SEAP reporter plasmid pG5SEAP. The transfected cells were cultured at 37°C in a 5% CO_2 incubator in DMEM (Mediatech) supplemented with penicillin (100 units/ml), streptomycin (100 µg/ml), L-glutamine (290 µg/ml) and heat-inactivated FBS (Hyclone). After 72 hr, supernatant from the culture was removed and assayed for SEAP activity. Briefly, 250 µl of the supernatant was heated to 65°C to inactivate any endogenous phosphates, after which it was added to an equal volume of 2× SEAP buffer (1 mM $MgCl_2$, 20 mM L-homoarginine, and 2 M diethanolamine [pH 9.8]) and incubated at 37°C for 10 min. Finally, 20 µl of 20 mM *p*-nitrophenol phosphate in 1× SEAP buffer was added and the OD_{405} was measured at 5 min intervals for 1.5 hr using a plate reader (Molecular Devices). The activity was calculated as the change in light absorbance per minute per sample.

Immunofluorescence Staining

To verify that the Gal4(1-147) peptide ligand fusions were being expressed and transfected in approximately equal amounts in HEK293 cells, immunofluorescence staining was performed. Briefly, the transiently transfected cells were fixed on glass slides using 2% paraformaldehyde. After multiple washes using blocking buffer (0.05% saponin, 5% BSA, and PBS [pH 7.2]) anti-Gal4 antibody (Covance) was added (1:2000 dilution) and incubated for 2 hr at room temperature. After six 5 min washes, an FITC-conjugated anti-rabbit secondary antibody (Santa Cruz Biotech) was added (1:150 dilution) and incubated for 45 min at room temperature. The slides were then washed with blocking buffer 6 times for 5 min each and Hoechst (Chemicon), a nuclear stain that enables visualization of all cells, was added to the slides. The cells were visualized under a microscope (Leica DM LB connected to Spot RT slider camera, Diagnostic Instruments).

References

1. Bourbon, H. M.; Aguilera, A.; Ansari, A. Z.; Asturias, F. J.; Berk, A. J.; Bjorklund, S.; Blackwell, T. K.; Borggrefe, T.; Carey, M.; Carlson, M.; Conaway, J. W.; Conaway, R. C.; Emmons, S. W.; Fondell, J. D.; Freedman, L. P.; Fukasawa, T.; Gustafsson, C. M.; Han, M.; He, X.; Herman, P. K.; Hinnebusch, A. G.; Holmberg, S.; Holstege, F. C.; Jaehning, J. A.; Kim, Y. J.; Kuras, L.; Leutz, A.; Lis, J. T.; Meisterernest, M.; Naar, A. M.; Nasmyth, K.; Parvin, J. D.; Ptashne, M.; Reinberg, D.; Ronne, H.; Sadowski, I.; Sakurai, H.; Sipiczki, M.; Sternberg, P. W.; Stillman, D. J.; Strich, R.; Struhl, K.; Svejstrup, J. Q.; Tuck, S.; Winston, F.; Roeder, R. G.; Kornberg, R. D., A unified nomenclature for protein subunits of Mediator complexes linking transcriptional regulators to RNA polymerase II. *Molecular Cell* **2004,** 14, (5), 553-557.
2. Davis, J. A.; Takagi, Y.; Kornberg, R. D.; Asturias, F. A., Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell* **2002,** 10, (2), 409-15.
3. Cheng, J. X.; Gandolfi, M.; Ptashne, M., Activation of the Gal1 gene of yeast by pairs of 'non-classical' activators. *Curr Biol* **2004,** 14, (18), 1675-9.
4. Wang, W.; El-Deiry, W. S., Restoration of p53 to limit tumor growth. *Curr Opin Oncol* **2008,** 20, (1), 90-6.
5. Gabrilovich, D. I., INGN 201 (Advexin): adenoviral p53 gene therapy for cancer. *Expert Opin Biol Ther* **2006,** 6, (8), 823-32.
6. Sinn, P. L.; Sauter, S. L.; McCray, P. B., Jr., Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. *Gene Ther* **2005,** 12, (14), 1089-98.
7. Toniatti, C.; Bujard, H.; Cortese, R.; Ciliberto, G., Gene therapy progress and prospects: transcription regulatory systems. *Gene Ther* **2004,** 11, (8), 649-57.
8. Kelleher, R. J., 3rd; Flanagan, P. M.; Kornberg, R. D., A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **1990,** 61, (7), 1209-15.
9. Flanagan, P. M.; Kelleher, R. J., 3rd; Sayre, M. H.; Tschochner, H.; Kornberg, R. D., A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **1991,** 350, (6317), 436-8.
10. Kim, Y. J.; Bjorklund, S.; Li, Y.; Sayre, M. H.; Kornberg, R. D., A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **1994,** 77, (4), 599-608.
11. Kuras, L.; Borggrefe, T.; Kornberg, R. D., Association of the Mediator complex with enhancers of active genes. *Proc Natl Acad Sci U S A* **2003,** 100, (24), 13887-91.
12. Myers, L. C.; Kornberg, R. D., Mediator of transcriptional regulation. *Annual Review of Biochemistry* **2000,** 69, 729-749.
13. Casamassimi, A.; Napoli, C., Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie* **2007,** 89, (12), 1439-46.
14. Asturias, F. J.; Jiang, Y. W.; Myers, L. C.; Gustafsson, C. M.; Kornberg, R. D., Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **1999,** 283, (5404), 985-7.

15. Hengartner, C. J.; Thompson, C. M.; Zhang, J.; Chao, D. M.; Liao, S. M.; Koleske, A. J.; Okamura, S.; Young, R. A., Association of an activator with an RNA polymerase II holoenzyme. *Genes Dev* **1995,** 9, (8), 897-910.
16. Taatjes, D. J.; Naar, A. M.; Andel, F., 3rd; Nogales, E.; Tjian, R., Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **2002,** 295, (5557), 1058-62.
17. Taatjes, D. J.; Schneider-Poetsch, T.; Tjian, R., Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nat Struct Mol Biol* **2004,** 11, (7), 664-71.
18. Lee, Y. C.; Park, J. M.; Min, S.; Han, S. J.; Kim, Y. J., An activator binding module of yeast RNA polymerase II holoenzyme. *Molecular and Cellular Biology* **1999,** 19, (4), 2967-2976.
19. Park, J. M.; Kim, H. S.; Han, S. J.; Hwang, M. S.; Lee, Y. C.; Kim, Y. J., In vivo requirement of activator-specific binding targets of mediator. *Molecular and Cellular Biology* **2000,** 20, (23), 8709-8719.
20. Swanson, M. J.; Qiu, H.; Sumibcay, L.; Krueger, A.; Kim, S. J.; Natarajan, K.; Yoon, S.; Hinnebusch, A. G., A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol Cell Biol* **2003,** 23, (8), 2800-20.
21. Fishburn, J.; Mohibullah, N.; Hahn, S., Function of a eukaryotic transcription activator during the transcription cycle. *Mol Cell* **2005,** 18, (3), 369-78.
22. Reeves, W. M.; Hahn, S., Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* **2005,** 25, (20), 9092-102.
23. Koh, S. S.; Ansari, A. Z.; Ptashne, M.; Young, R. A., An activator target in the RNA polymerase II holoenzyme. *Mol Cell* **1998,** 1, (6), 895-904.
24. Ansari, A. Z.; Koh, S. S.; Zaman, Z.; Bongards, C.; Lehming, N.; Young, R. A.; Ptashne, M., Transcriptional activating regions target a cyclin-dependent kinase. *Proc Natl Acad Sci U S A* **2002,** 99, (23), 14706-9.
25. Gaudreau, L.; Keaveney, M.; Nevado, J.; Zaman, Z.; Bryant, G. O.; Struhl, K.; Ptashne, M., Transcriptional activation by artificial recruitment in yeast is influenced by promoter architecture and downstream sequences. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** 96, (6), 2668-2673.
26. Suzuki, Y.; Nogi, Y.; Abe, A.; Fukasawa, T., GAL11 protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in Saccharomyces cerevisiae. *Mol Cell Biol* **1988,** 8, (11), 4991-9.
27. Nishizawa, M.; Suzuki, Y.; Nogi, Y.; Matsumoto, K.; Fukasawa, T., Yeast Gal11 protein mediates the transcriptional activation signal of two different transacting factors, Gal4 and general regulatory factor I/repressor/activator site binding protein 1/translation upstream factor. *Proc Natl Acad Sci U S A* **1990,** 87, (14), 5373-7.
28. Suzuki, Y.; Nogi, Y.; Abe, A.; Fukasawa, T., GAL11 protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in Saccharomyces cerevisiae. *Mol Cell Biol* **1992,** 12, (10), 4806.
29. Jeong, C. J.; Yang, S. H.; Xie, Y.; Zhang, L.; Johnston, S. A.; Kodadek, T., Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **2001,** 40, (31), 9421-7.
30. Novatchkova, M.; Eisenhaber, F., Linking transcriptional mediators via the GACKIX domain super family. *Current biology : CB* **2004,** 14, (2), R54-5.

31. Thakur, J. K.; Arthanari, H.; Yang, F.; Pan, S. J.; Fan, X.; Breger, J.; Frueh, D. P.; Gulshan, K.; Li, D. K.; Mylonakis, E.; Struhl, K.; Moye-Rowley, W. S.; Cormack, B. P.; Wagner, G.; Naar, A. M., A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **2008,** 452, (7187), 604-9.
32. Sakurai, H.; Fukasawa, T., Functional correlation among Gal11, transcription factor (TF) IIE, and TFIIH in Saccharomyces cerevisiae. Gal11 and TFIIE cooperatively enhance TFIIH-mediated phosphorylation of RNA polymerase II carboxyl-terminal domain sequences. *J Biol Chem* **1998,** 273, (16), 9534-8.
33. Sakurai, H.; Fukasawa, T., Yeast Gal11 and transcription factor IIE function through a common pathway in transcriptional regulation. *J Biol Chem* **1997,** 272, (51), 32663-9.
34. Long, R. M.; Mylin, L. M.; Hopper, J. E., GAL11 (SPT13), a transcriptional regulator of diverse yeast genes, affects the phosphorylation state of GAL4, a highly specific transcriptional activator. *Mol Cell Biol* **1991,** 11, (4), 2311-4.
35. Zhang, F.; Sumibcay, L.; Hinnebusch, A. G.; Swanson, M. J., A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol Cell Biol* **2004,** 24, (15), 6871-86.
36. Bhoite, L. T.; Yu, Y.; Stillman, D. J., The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. *Genes Dev* **2001,** 15, (18), 2457-69.
37. Lallet, S.; Garreau, H.; Garmendia-Torres, C.; Szestakowska, D.; Boy-Marcotte, E.; Quevillon-Cheruel, S.; Jacquet, M., Role of Gal11, a component of the RNA polymerase II mediator in stress-induced hyperphosphorylation of Msn2 in Saccharomyces cerevisiae. *Mol Microbiol* **2006,** 62, (2), 438-52.
38. Leroy, C.; Cormier, L.; Kuras, L., Independent recruitment of mediator and SAGA by the activator Met4. *Mol Cell Biol* **2006,** 26, (8), 3149-63.
39. Sakurai, H.; Hiraoka, Y.; Fukasawa, T., Yeast GAL11 protein is a distinctive type transcription factor that enhances basal transcription in vitro. *Proc Natl Acad Sci U S A* **1993,** 90, (18), 8382-6.
40. Sakurai, H.; Ohishi, T.; Amakasu, H.; Fukasawa, T., Yeast GAL11 protein stimulates basal transcription in a gene-specific manner by a mechanism distinct from that by DNA-bound activators. *FEBS Lett* **1994,** 351, (2), 176-80.
41. Lu, X. Y.; Ansari, A. Z.; Ptashne, M., An artificial transcriptional activating region with unusual properties. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (5), 1988-1992.
42. Lu, Z.; Ansari, A. Z.; Lu, X. Y.; Ogirala, A.; Ptashne, M., A target essential for the activity of a nonacidic yeast transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (13), 8591-8596.
43. Lim, M. K.; Tang, V.; Le Saux, A.; Schuller, J.; Bongards, C.; Lehming, N., Gal11p dosage-compensates transcriptional activator deletions via Taf14p. *J Mol Biol* **2007,** 374, (1), 9-23.
44. Tanaka, M., Modulation of promoter occupancy by cooperative DNA binding and activation-domain function is a major determinant of transcriptional regulation by activators in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (9), 4311-4315.

45. Ptashne, M.; Gann, A., *Genes & Signals*. Cold Spring Harbor Laboratory: New York, 2001.
46. Wu, Y.; Reece, R. J.; Ptashne, M., Quantitation of putative activator-target affinities predicts transcriptional activating potentials. *Embo J* **1996,** 15, (15), 3951-63.
47. Melcher, K., The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins. *Journal of Molecular Biology* **2000,** 301, (5), 1097-1112.
48. Alluri, P.; Liu, B.; Yu, P.; Xiao, X.; Kodadek, T., Isolation and characterization of coactivator-binding peptoids from a combinatorial library. *Mol Biosyst* **2006,** 2, (11), 568-79.
49. Volkman, H. M.; Rutledge, S. E.; Schepartz, A., Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain. *J Am Chem Soc* **2005,** 127, 4649-4658.
50. Ma, J.; Ptashne, M., A New Class of Yeast Transcriptional Activators. *Cell* **1987,** 51, (1), 113-119.
51. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., Towards a minimal motif for artificial transcriptional activators. *Chemistry & Biology* **2001,** 8, (6), 583-592.
52. Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., Design of artificial transcriptional activators with rigid poly-L-proline linkers. *Journal of the American Chemical Society* **2002,** 124, (44), 13067-13071.
53. Stanojevic, D.; Young, R. A., A highly potent artificial transcription factor. *Biochemistry* **2002,** 41, (23), 7209-7216.
54. Dumoulin, P.; Ebright, R. H.; Knegtel, R.; Kaptein, R.; Granger-Schnarr, M.; Schnarr, M., Structure of the LexA repressor-DNA complex probed by affinity cleavage and affinity photo-cross-linking. *Biochemistry* **1996,** 35, (14), 4279-86.
55. Knegtel, R. M.; Fogh, R. H.; Ottleben, G.; Ruterjans, H.; Dumoulin, P.; Schnarr, M.; Boelens, R.; Kaptein, R., A model for the LexA repressor DNA complex. *Proteins* **1995,** 21, (3), 226-36.
56. Hong, M.; Fitzgerald, M. X.; Harper, S.; Luo, C.; Speicher, D. W.; Marmorstein, R., Structural basis for dimerization in DNA recognition by gal4. *Structure* **2008,** 16, (7), 1019-26.
57. Hidalgo, P.; Ansari, A. Z.; Schmidt, P.; Hare, B.; Simkovich, N.; Farrell, S.; Shin, E. J.; Ptashne, M.; Wagner, G., Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes & Development* **2001,** 15, (8), 1007-1020.
58. Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C., DNA Recognition by Gal4 - Structure of a Protein DNA Complex. *Nature* **1992,** 356, (6368), 408-414.
59. Han, Y.; Kodadek, T., Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. *J Biol Chem* **2000,** 275, (20), 14979-84.
60. Seipel, K.; Georgiev, O.; Schaffner, W., Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *The EMBO journal* **1992,** 11, (13), 4961-4968.
61. Kunzler, M.; Braus, G. H.; Georgiev, O.; Seipel, K.; Schaffner, W., Functional differences between mammalian transcription activation domains at the yeast GAL1 promoter. *The EMBO journal* **1994,** 13, (3), 641-645.

62. Ruden, D. M.; Ma, J.; Ptashne, M., No strict alignment is required between a transcriptional activator binding site and the 'TATA box' of a yeast gene. *Proceedings of the National Academy of Sciences of the United States of America* **1988,** 85, (12), 4262-4266.
63. Balciunas, D.; Hallberg, M.; Bjorklund, S.; Ronne, H., Functional interactions within yeast mediator and evidence of differential subunit modifications. *J Biol Chem* **2003,** 278, (6), 3831-9.
64. Sakurai, H.; Kim, Y. J.; Ohishi, T.; Kornberg, R. D.; Fukasawa, T., The yeast GAL11 protein binds to the transcription factor IIE through GAL11 regions essential for its in vivo function. *Proc Natl Acad Sci U S A* **1996,** 93, (18), 9488-92.
65. Hiller, M. A.; Lin, T. Y.; Wood, C.; Fuller, M. T., Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev* **2001,** 15, (8), 1021-30.
66. Kim, U.; Qin, X. F.; Gong, S.; Stevens, S.; Luo, Y.; Nussenzweig, M.; Roeder, R. G., The B-cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature* **1996,** 383, (6600), 542-7.
67. Fischer, J. A.; Giniger, E.; Maniatis, T.; Ptashne, M., GAL4 activates transcription in Drosophila. *Nature* **1988,** 332, (6167), 853-856.
68. Kakidani, H.; Ptashne, M., GAL4 activates gene expression in mammalian cells. *Cell* **1988,** 52, (2), 161-167.
69. Ma, J.; Przibilla, E.; Hu, J.; Bogorad, L.; Ptashne, M., Yeast activators stimulate plant gene expression. *Nature* **1988,** 334, (6183), 631-633.
70. Burk, D.; Dawson, D.; Stearns, T., *Methods in Yeast Genetics*. 1 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2000.

CHAPTER 3
FACTORS THAT ENHANCE THE POTENCY OF TRANSCRIPTIONAL ACTIVATORS*

A. Abstract

Transcriptional activator proteins are major players in regulating gene expression. They precisely control the timing of and levels to which their cognate genes are transcribed.[1] It is hypothesized that activators are able to achieve a desired level of control through an extensive network of interactions with partners inside and outside the transcriptional machinery.[3] For example, activators can be post-translationally modified, undergo ligand-dependent nuclear localization, interact with masking proteins and proteins in the transcriptional machinery and eventually be degraded by the proteolytic machinery.[4-14] However, it is unclear how each of these events individually contributes to the overall function of an activator.

In Chapter 2, we showed that by mimicking a single step of this cascade, specific recruitment of a single transcriptional machinery component Med15(Gal11), we were able to create activator ATFs. Unfortunately, the molecules only functioned with a fraction of the activity stimulated by natural activators. In this Chapter, we investigate the effect of incorporating additional interactions into activator ATFs in an effort to increase transcriptional output. First we create ligands for the SAGA complex component Tra1; these ligands, when attached to a DBD, are able to activate transcription to levels higher than the Med15(Gal11)-specific ligands. In contrast to the Med15 ligands described in

* Portions of this Chapter were taken from Lum, J.K.; Majmudar, C.Y.; Ansari, A.Z.; Mapp, A.K. *ACS Chem. Biol.* **2006**, *1*, 639-643. Experiments shown in Figures 3.18, 3.20, 3.21 and 3.22 were performed by Jenifer Lum. I performed experiments in Figures 3.5 - 3.15, 3.19 and Tables 3.2 - 3.4. Experiments leading to the model proposed in Figure 3.4 were performed by Steven Rowe, Brian Brennan and members of the Ansari group (U. of Wisconsin).

Chapter 2, the Tra1 ligands interact with more than one coactivator. Evidently, the TAD binding site in Tra1 is similar to binding sites within other coactivators. Further, combination of a Med15(Gal11) ligand with a Tra1 ligand leads to a synergistic increase in activation. Next, we show that engineering an interaction outside the transcriptional machinery into activator ATFs also leads to greatly enhanced function. Specifically, we show that supplementing a TAD-coactivator recruitment event with an intramolecular masking interaction leads to higher activity, presumably through protection from proteolysis and/or non-specific interactions. In the future, inclusion of these design criteria in new generations of biopolymer-based and small molecule activators will enable the development of molecules that rival the function of endogenous activators.

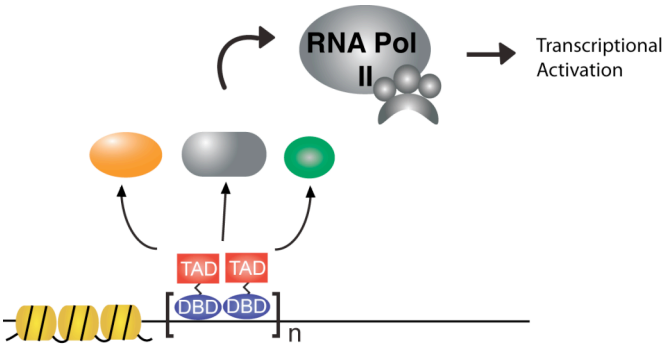


Figure 3.1 Overview of transcriptional activator interactions. Transcriptional activation domains (TADs) undergo a variety of interactions with partners inside and outside the transcriptional machinery that result in RNA polII recruitment and gene activation.

B. Introduction

B.1 Activator interactions within the transcriptional machinery

At a gene promoter transcriptional activation is initiated by the activator – dependent recruitment of chromatin remodeling enzymes and formation of the preinitiation complex (PIC) that consists of coactivators and general transcription machinery components (Figure 3.1).[1, 15] Activators are thought to recruit, stabilize and/or induce conformational changes in the preinitiation complex, signaling RNA polymerase II to initiate transcription.[16-18] However, the identity and the relevance of the exact interactions activators make with these complexes in cells are ambiguous. More recently it was shown using chromatin immunoprecipitation at the Gal1 promoter that first Spt-Ada-Gcn5-acetyltransferase (SAGA), a chromatin remodeling complex is recruited, followed by the SAGA-independent recruitment of Mediator and finally RNA

polymerase II along with the general transcriptional machinery.[19] Thus, from this study it is likely that activators such as Gal4 directly contact components of SAGA and Mediator.

Subsequently, Hahn and coworkers showed using in vitro crosslinking that Gal4 and Gcn4 interact with the SAGA component Tra1, the Mediator subunit Med15(Gal11) and Taf12, a shared subunit of SAGA and TFIID; remarkably, the crosslinking results suggest that a single peptide sequence within each activator mediates all of the binding contacts.[20, 21] These finding are consistent with the ChIP experiments mentioned above and suggest that activators utilize multiple interactions for initiating gene expression. Further, because two distinct activators interact with the same set of targets, these results also hint at the possibility of a core group of conserved activator-transcriptional machinery contacts, possibly overlapping activator binding sites on target proteins.

We hypothesized that if we could artificially create multiple interactions with the transcriptional machinery, we could probe the functional contribution of recruiting different complexes involved in transcriptional activation. In Chapter 2 we found that recruitment of Med15(Gal11) led to the activation of transcription. However, the Med15 specific ligands only moderately activated transcription and perhaps other interactions are necessary to activate transcription to levels similar to natural activators. Thus we were interested in creating ligands for other transcriptional machinery proteins and evaluating their ability to activate transcription.

B.1.a Tra1: a likely target of activators

There is considerable evidence that Tra1 is a bonafide activator target. The Workman group initially showed that the activator Hap4 specifically crosslinked with Tra1 when incubated with reconstituted (cell-free) SAGA complex.[22] Moreover, they also show using GST pulldown assays that the TADs of Hap4, Gal4, Gcn4 and VP16 interact with purified Tra1 (Figure 3.2). Subsequently, these finding were also confirmed by the Hahn laboratory through in vitro crosslinking studies using yeast nuclear extracts and the Gal4 and Gcn4 TADs, identifying Tra1 as one of three targets that these activators bound.[20, 21] Since the crosslinking studies were performed in vitro, the cellular confirmation of an activator-Tra1 interaction was lacking. Since then, Green and coworkers, using a FRET-based approach, identified Tra1 in the context of SAGA to be the sole cellular target of Gal4 in yeast.[23, 24] While this approach validates Tra1 as a

bonafide activator target, it is not known if the function of Gal4 is solely dependent on a Tra1 interaction. Further, due to the limitations of the FRET approach, Tra1 could be one of many cellular targets of Gal4 (see Chapter 4); indeed, this would be more consistent with the preponderance of evidence showing that a variety of transcription complexes are recruited to a gene promoter.[19, 25-28]

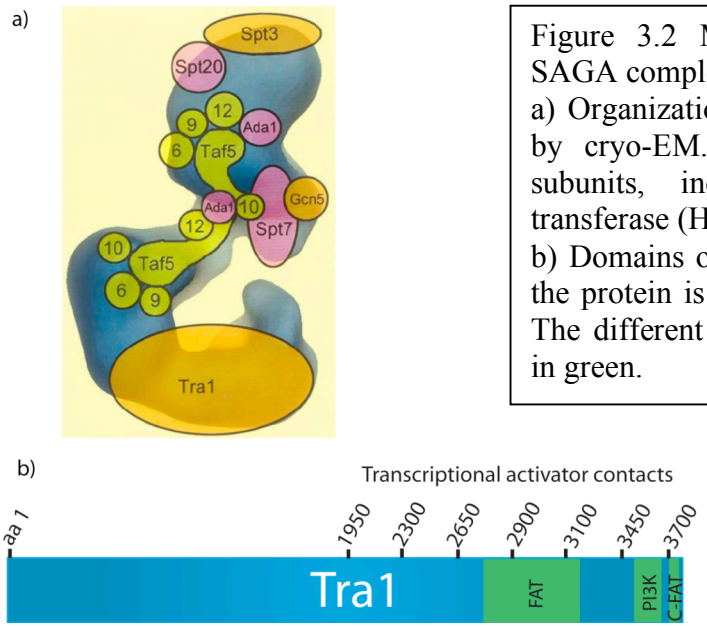


Figure 3.2 Molecular architecture of the SAGA complex and its component Tra1
a) Organization of SAGA complex revealed by cryo-EM. SAGA is composed of 14 subunits, including the histone acetyl transferase (HAT) Gcn5.[2]
b) Domains of Tra1. The C terminal half of the protein is involved in activator contacts. The different conserved domains are shown in green.

Given the evidence suggesting that Tra1 is an activator target, we wanted to investigate the interactions with Tra1 that contribute to function. Specifically, we wanted to explore if there were multiple functional activator binding sites in this protein. Furthermore, were these binding sites unique compared to other coactivator binding sites? For these purposes, we envisioned we would identify the region(s) on Tra1 that interact with natural activators. Next we could screen for ligands that bound this region to evaluate if ligands that interact with Tra1 function as transcriptional activation domains. Finally, we could combine the Tra1 ligands with Med15(Gal11) ligands to probe if multiple interactions with the transcriptional machinery would facilitate higher levels of activity.

B.2 Interactions outside the transcriptional machinery

While interactions with the transcriptional machinery are necessary for function, activators have also been shown to have interactions with proteins outside the transcriptional complex. They include interactions with enzymes that introduce covalent modifications such as phosphorylation and ubiquitylation as well as interactions with

small molecules.[4-14] These interactions primarily play a role in controlling the timing of gene expression.[10] For example the transcriptional activation domain of the mammalian activator CREB is inactive until ser133 is phosphorylated, which then results in recruitment of the coactivator CBP and upregulation of gene expression.[7]

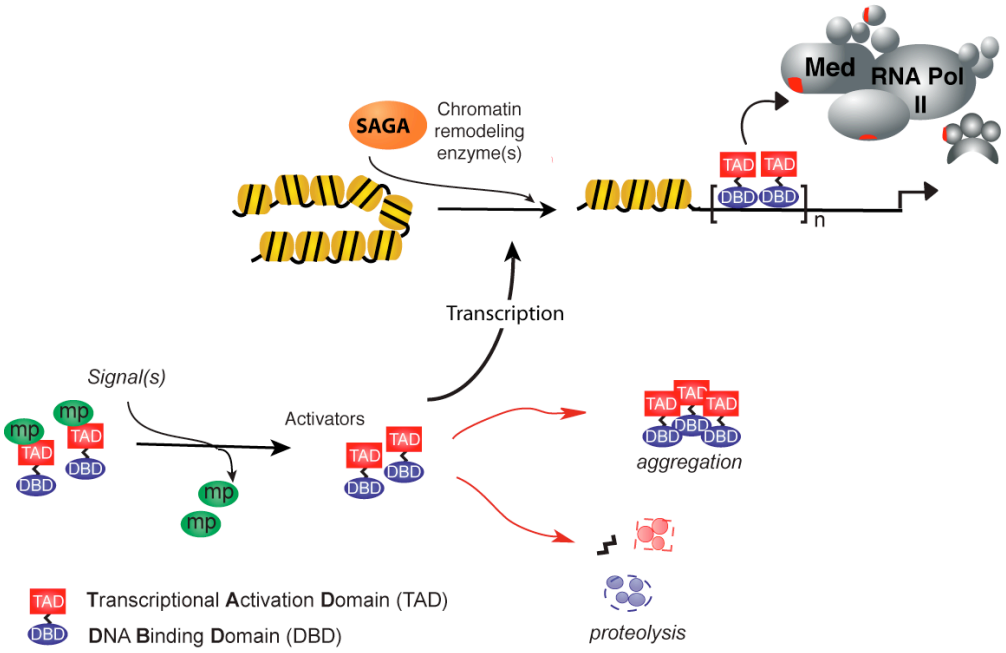


Figure 3.3 Transcriptional activator interactions inside and outside transcriptional machinery. In additional to interactions within the transcriptional machinery, the transcriptional activation domain (TAD) of an activator also interacts with masking proteins (mp) (shown in green). Masking interactions help prevent non-specific interactions such as aggregation and proteolysis that lead to non-functional activators.

Masking interactions also control the timing of gene expression. In a masking interaction overlapping residues in a TAD required for activation are concealed from the rest of the cellular environment, thus preventing activation until the residues are revealed to the transcriptional machinery. An excellent example of this is the mouse homolog of double minute 2 (hDM2), a masking protein that regulates p53 function by discriminating p53-coactivator interactions and playing a role in p53 turnover.[29-32] An instance of an intramolecular masking interaction is found in the activator Put3, which in normal cells is unable to activate transcription due to concealment of its activating region. On the addition of a specific metabolite (proline) it undergoes a conformational change that frees the transcriptional activation domain, thus enabling it to stimulate expression of its cognate genes (for additional examples see Table 3.1). [8, 33, 34]

Recent work from our group demonstrated that masking interactions also play a role in the overall activity of an activator through preventing aggregation, non-specific binding interactions and/or increasing the stability of the activator. For example, XL_Y (an artificial TAD) has its function greatly enhanced by the inclusion of a masking interaction.[35]

Activator	Masking protein	Role of interaction
p53	mDM2[32, 36]	Shields TAD from interactions, also controls degradation
E2F	RB[37]	Phosphorylation of RB releases E2F and stimulates cell proliferation
Gal4	Gal80[5, 38-41]	In the presence of glucose, Gal80 masks the TAD from interactions; addition of galactose stimulates a conformation change that exposes the TAD to the transcriptional machinery
Leu3	Intramolecular[42, 43]	Leu3 acts as a repressor protein. On the addition of a Isopropylmalate (α IPM), Leu3 undergoes a conformational change from a repressor to an activator
Put3	Intramolecular[8, 33, 34]	Addition of proline results in a conformation change that releases the TAD from interacting with an intramolecular inhibitory region and permits activation
PC4	Intramolecular[44]	Depending of the phosphorylation state of PC4, it undergoes conformational changes in a lysine rich region, which gradually effects DNA binding and transcriptional cofactor recruitment
IRF-7	IRF-5[45]	IRF-5 binds to the IRF-7 and masks its DNA binding domain, preventing activation
Hap1	Hsp70-Ydj1[46]	In absence of heme, Hap1 is bound by at least four cellular proteins including Hsp70 and Ydj1 which represses its activity. On the addition of heme these proteins dissociate from Hap1 and high levels of transcription are observed
NorR	Intramolecular[47]	Binding of Nitric oxide to the non heme iron center in the GAF domain of NorR results in the formation of a mononitrosyl iron complex and relieves intramolecular repression, enabling NorR to activate transcription
Pit-1	Pit-1β[48, 49]	Heterodimerization of Pit-1 with Pit-1β results in masking of the Pit1 TAD preventing it from interacting with its coactivator CBP
Ino2	Opi1[50]	In the presence of inositol and choline, the basic leucine zipper motif containing protein Opi1 binds the TAD of Ino2 preventing activation
Vnd	Intramolecular[51]	Vnd contains two TADs that are masked intramolecularly. When Dichaete binds Vnd, the TADs are unmasked and gene expression is initiated. Further, Vnd also acts as a repressor when it interacts the co-repressor Groucho
TAK1	TIP27[52]	TIP27 containing two zinc finger motifs represses TAK1 by interacting with the TAK1 TAD and it is suggested to affect coactivator recruitment
p65	GR[53]	Glucocorticoid receptor (GR) represses the activity of NF-kB activation domain p65 by competing with the p65-CBP interaction
Smad3	GR LBD[54]	The function of Smad3 involved in TGF-β signaling is repressed by an interaction with the GR ligand binding domain (LBD)
Mac1	Intramolecular[55]	Copper binding to two cys-rich motifs results in a conformational change that creates an intramolecular interaction in Mac1 between the DBD and the TAD, inhibiting both DNA and coactivator binding
EWS and SSAP	ZFM1[56]	Binding of the ZFM1 to the TADs of EWS and SSAP prevents them from activating transcription
ELF3 (ESX)	Intramolecular[57]	In absence of a binding partner, the ELF3 TAD interacts with the DBD preventing DNA binding. On association with a coactivator, ELF3 undergoes a conformational change that enhances DNA binding and in turn activation

Table 3.1 Activator-Masking Interactions

B.2.a XL_Y

In contrast to natural TADs, artificial TADs generally operate outside of the endogenous regulatory pathways and lack masking partners or interactions. Therefore, artificial TADs are typically unable to utilize all of the resources of the cellular machinery that contribute to functional potency and this could possibly explain why they do not activate transcription to high levels in vivo or in cell culture.[10, 29, 38, 58]

There is one exception to the trend of modestly active artificial TADs: the 16 amino acid artificial TAD XL_Y activates transcription to levels seen with full length Gal4, a potent yeast activator.[59, 60] Drs. Brian Brennan and Steven Rowe (former graduate students) in collaboration with Prof. Aseem Ansari (U. of Wisconsin) performed a series of biochemical and genetic experiments that probed the mechanism by which this activator achieved high activity levels. Using fluorescence polarization-based binding assays, they found that XL_Y interacted with low micromolar affinity with its transcriptional machinery target, Med15(Gal11). Surprisingly, they also found that XL_Y had similar affinity to part of its own DBD, Gal4(1-100). Moreover, disruption of either interaction in Med15(Gal11) mutant strains or attachment of XL_Y to a different DBD resulted in loss of activity. Further investigation identified the hydrophobic loop in the dimerization domain comprising residues (52-100) of Gal4 to be critical for XL_Y function.

From these results, it was hypothesized that the secondary contact with the Gal4 dimerization domain likely serves as a masking interaction, preventing XL_Y from participating in non-productive interactions (for example with chaperones) or proteolysis until it is exposed to its target protein Med15(Gal11), at which point XL_Y can recruit the transcriptional machinery and robustly activate transcription (Figure 3.4).[35]

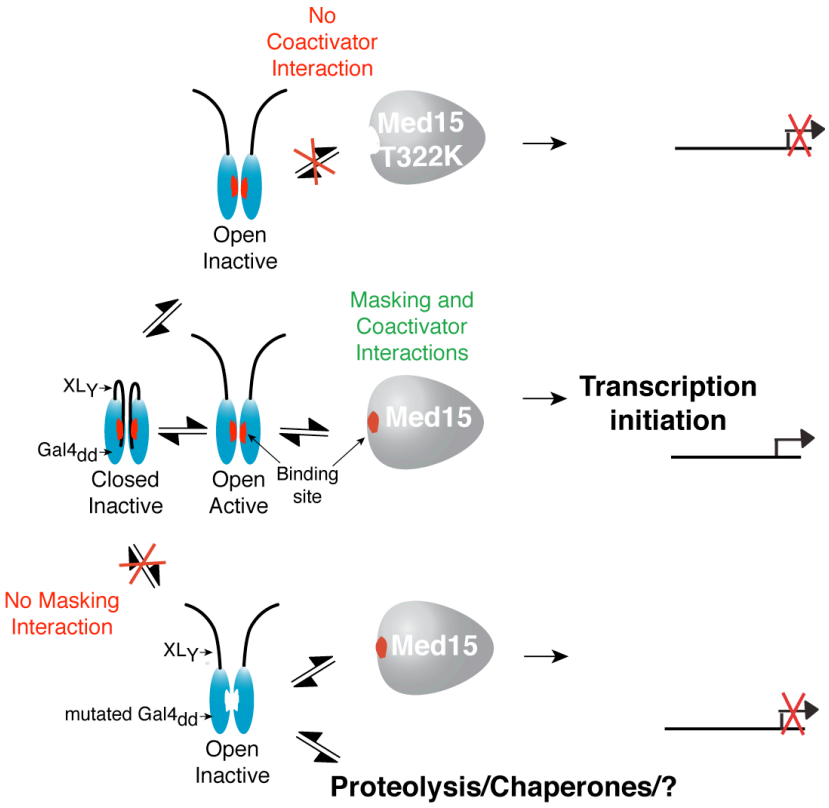


Figure 3.4 Model for XL_Y function. XL_Y can only activate transcription when it can perform two interactions, one with part of its own DNA binding domain $Gal4_{dd}$ and the other with its transcriptional machinery target Med15 (middle pathway). Disruption of either interaction leads to loss of function; top pathway does not permit coactivator binding, bottom pathway does not allow interaction with $Gal4_{dd}$. XL_Y binding sites are indicated in red and their absence is indicated by clear space.

Given that a masking interaction is essential for the function of XL_Y, it is likely that masking interactions play a general role in controlling not only the timing of endogenous activator function but also in assisting the level of upregulation. To directly probe this, we hypothesized that engineering a masking interaction into TADs that normally do not have a masking interaction should significantly augment their function.

C. Results

C.1 Creating an additional interaction within the transcriptional machinery

Given the importance of chromatin remodeling in transcriptional activation and that endogenous activators target Tra1, a component of these complexes, we hypothesized that ligands for this protein would be useful tools for investigating the importance of recruiting chromatin remodeling complexes in transcription initiation as they could be used to either specifically block or recruit such complexes. More importantly, we previously created ligands for the Mediator component Med15(Gal11) (Chapter 2) that were only able to moderately activate transcription in a Med15(Gal11)-dependent manner at a synthetic Gal1 promoter. Thus, it would be of interest to probe the functional effect upon recruitment of the SAGA complex to a gene promoter in combination with Mediator recruitment. One of the central models of transcription initiation is that multiple activators residing at a promoter work cooperatively to recruit various complexes, and our strategy would be a means to replicate this with, presumably, with greatly enhanced function.[28, 61]

C.1.a Identification of activator binding region of Tra1

In order to screen for ligands that bind Tra1, it was first necessary to identify the endogenous activator-binding module of the 433 kDa protein. We hypothesized that we could better mimic activator function by creating ligands that bound similar surfaces on Tra1 as natural activators. Presumably, endogenous activators target binding sites on Tra1 that are accessible in the cell, in context of the various proteins Tra1 associates with and ligands that target these regions should be able to interact with Tra1 in vivo. Workman and coworkers provide some insight into the region of Tra1 important for activator function; a temperature-sensitive Tra1 mutant with amino acid substitutions in the C-terminal half of Tra1 (amino acids 2226 to 3744) abolished VP16 binding in vitro and yeast containing this mutant displayed phenotypes consistent with compromised Gcn4-dependent activation.[22]

To further pinpoint a region on Tra1 targeted by the activators Gcn4, Gal4 and VP2, we created ~50 kDa fragments of Tra1 spanning amino acids 1900-3500 (Figure 3.5a). These fragments were initially expressed as GST fusions in E. *coli* to facilitate

isolation. However, an extremely low level of soluble protein was obtained, with the majority of the GST-Tra1 fusions residing in the cell pellet. Thus we decided to express and purify the Tra1 fragments as fusions to the maltose binding protein (MBP) in *E. coli* to assist in solubility and purification (Figure 3.5b), leading to ~2 mg/L of protein that was soluble up to 50-100 μM concentrations. The TADs Gcn4 (105-134), Gal4 (840-881) and VP2 were synthesized using solid phase synthesis methods and fluorescein labeled (Figure 3.5c). Using fluorescence polarization, the affinity of Gal4, Gcn4 and VP2 was assessed for each Tra1 fragment. It was found that only Tra1(3092-3524) had any measurable affinity for these TADs (Table 3.2). It is also possible that the other fragments are not fold properly when expressed in *E. coli* and thus do not bind TADs.

The binding results were quite different than those seen with the coactivator Med15(Gal11). Firstly, the TADs bind Tra1 5-10 fold weaker than they bind Med15(Gal11). Perhaps activators recruit SAGA using interactions with Tra1 and Taf12, a component of SAGA also postulated to be a target of activators.[20] Also, in the case of Tra1 only one region was found to interact with activators, while Med15 has many functional binding sites (see Chapter 4). Furthermore, consistent with the hypothesis of the existence of a subset of functionally conserved binding sites, the three endogenous TADs target the same Tra1 fragment, possibly targeting a similar binding site(s).[62, 63] However, competition experiments to verify this could not be performed conclusively since at concentrations above 500 μM the TADs showed significant aggregation.

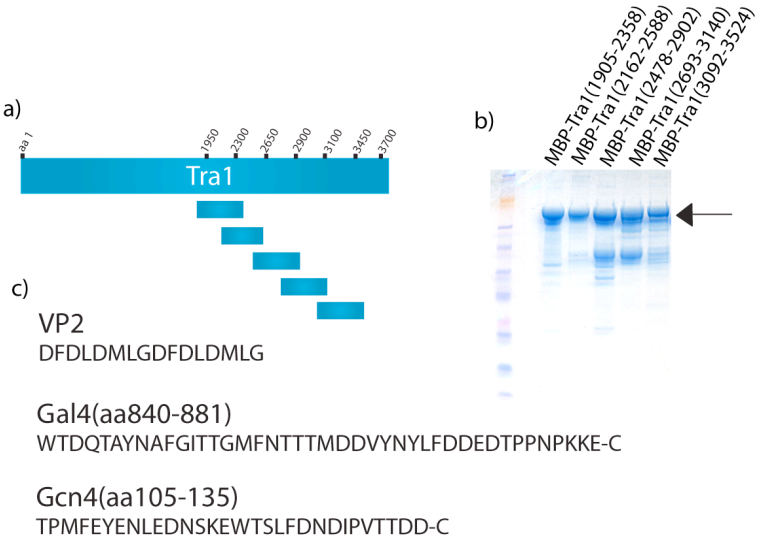


Figure 3.5 Tra1 E. *coli* expression constructs
a) Different constructs spanning the C-terminal half of Tra1 were constructed as fusions to MBP
b) Expression of MBP-Tra1 constructs in Rosetta (DE3) pLysS cells.
c) Sequences of TADs used to identify activator binding region(s) within Tra1.

	Tra1 amino acids				
	1905-2358	2162-2588	2478-2902	2693-3140	3092-3524
Gal4 (840-881)	ND	ND	ND	ND	17 ± 3 μM
Gcn4 (105-134)	ND	ND	ND	ND	>50 μM
VP2	ND	ND	ND	>50 μM	34 ± 4 μM

Table 3.2 Natural activator binding affinities for Tra1 fragments. The dissociation constant for different fragments of Tra1 was determined using FP. The data reported is an average of 3 experiments with error indicated as SD. ND = not detected.

C.1.b Screening for ligands

In order to identify ligands for Tra1, we first screened an 8 amino acid synthetic peptide library (AXXXXPSE) against Tra1(3092-3524) (Figure 3.6). The library was synthesized on solid phase on PEGA resin and had 4 variable positions giving a total possible 130,321 unique sequences. The same library was successfully used previously to yield ligands that targeted Med15(Gal11) and activated transcription (Chapter 2). Ligands for Tra1 were obtained using an ELISA-based screen and they were selected on a colorimetric basis relative to the positive control VP2, an endogenous activator that binds Tra1 and the negative control, which was the acetylated resin (Figure 3.7). The sequences of the individual peptides were obtained using Edman degradation sequencing (Figure 3.8).

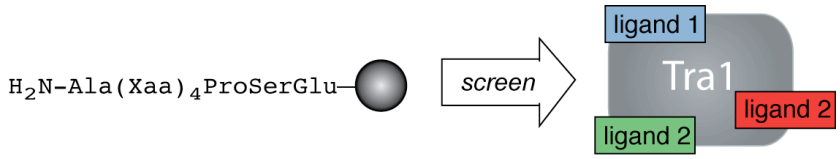


Figure 3.6 ELISA-based screen for Tra1 ligands. An 8 amino acid combinatorial library with 4 randomized positions (130,321 unique sequences) synthesized on PEGA resin was screened to obtain ligands for Tra1(3092-3534).

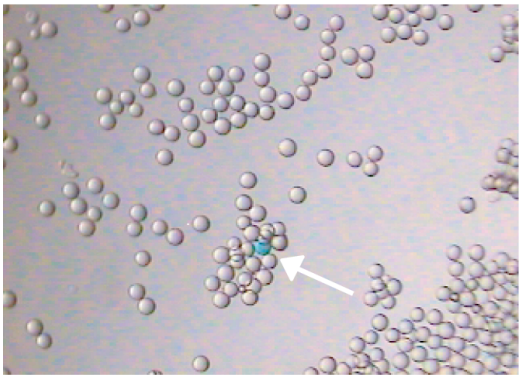


Figure 3.7 Positive hit from ELISA-based screen against Tra1. The synthetic library was incubated with MBP-Tra1 and subsequently with anti MBP-HRP antibody. The beads that bound Tra1 are indicated with the observance of blue color due to the cleavage of the TMB substrate.

In contrast to the ligands obtained that bound Med15(Gal11), the Tra1 ligands bore significant sequence homology, perhaps indicating that they target the same binding site; all the ligands were acid rich (Figure 3.8).[64] However, compared to endogenous TADs such as Gal4 that also contain a preponderance of acidic residues, a majority of these ligands did not contain interspersed hydrophobic residues. The presence of hydrophobic residues in endogenous TADs has been well documented to be important for interaction with target proteins and the dearth of hydrophobic residues was therefore somewhat surprising.[30, 62, 65-67] In order to measure the affinity of the ligands for Tra1, several ligands were synthesized and labeled with fluorescein. Using fluorescence polarization-based binding assays it was found that the ligands only weakly interacted with Tra1; a measurable dissociation constant could not be calculated due to the propensity of Tra1 to precipitate/aggregate at higher concentrations (Figure 3.9a). Nonetheless, when these ligands were fused to a DBD they activated transcription to 20-30 fold (Figure 3.9b). Given that the positive control VP2 itself binds Tra1 with a moderate affinity (30 ±4 μM), perhaps a high affinity interaction with Tra1 in vitro is not required for activator function.

ANEDDPSE AEQDDPSE ADEDTPSE
ADFDEPSE AVEDDPSE ADQDDPSE
ADMDYPSE AYEEDPSE ANFDDPSE

Figure 3.8 Sequences of ligands that bound Tra1 in the synthetic peptide library screen. Sequences of ligands obtained by Edman degradation sequencing of beads that turned blue in ELISA-based screen with Tra1. Variable positions in the ligands are indicated in gray.

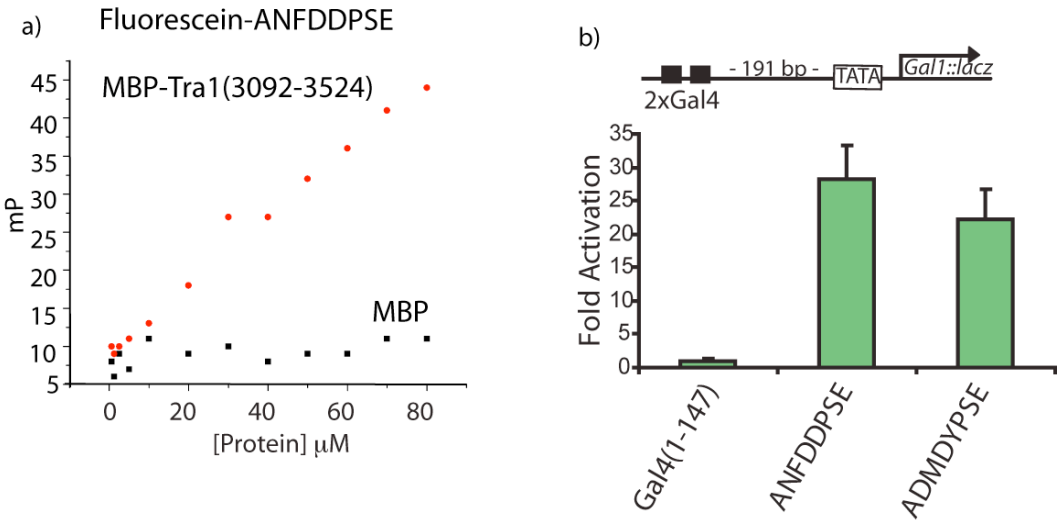


Figure 3.9 Functional evaluation of synthetic library ligands for Tra1.
a) Fluorescence polarization was used to determine the affinity of the ligands for Mbp-Tra1 or MBP. The ligands interact with Tra1 but, a dissociation constant could not be measured due to the aggregation of Tra1 at higher concentrations
b) Gal4+ligand fusions were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the Gal4+ligand fusion by the activity of the Gal4 DBD alone. Error is reported as SDOM from triplicate experiments.

C.1.c Phage display selection

As outlined above, the screen against Tra1 provided quite different results than the Med15(Gal11) screen described in Chapter 2. In the Med15(Gal11) screen, ligands that activated transcription and that did so by interacting with at least three distinct binding surfaces were identified. In contrast, the Tra1 ligands shared significant sequence similarities, suggesting that they were likely targeting a shared binding site. This was somewhat surprising since the use of a screen typically facilitates the identification of ligands that bind to different surfaces and the results may indicate that the peptide library used was not sufficiently diverse in sequence composition.[64] Further both native TADs and the Tra1-selected TADs bound to Tra1 more weakly than the Med15(Gal11) ligands interacted with their target protein. Ultimately the poor binding affinity of the Tra1 ligands limits their utility as mechanistic tools.

To address the limitations of the previous experiment, we chose to carry out a phage display selection with a commercially available phage-displayed dodecamer

library. By using a longer peptide (twelve versus eight residues) and carrying out multiple rounds of selection, affinity enhancement is often realized.[68-70] Each phage displays dodecamer peptides on the surface of the coat protein PIII and it contains five copies of the peptide. In addition, the phage library has ~10^9 different sequences, several orders of magnitude more than the octamer synthetic peptide library used for the ELISA-based screen.

No VP2	VP2 in binding buffer
ENSPLWWPQPLA	YHSQVSPPWPHM
EWALMQVPPNSK	TWIPSYPLVSRD
TTPWLPKPWYTP	NYPSWPLQLTMP
TKFTDTARYPPL	IQHEKHLHLQPR
TERLRDLPTLIR	NLSWHSFDTGTV
GLQQLRPPAQHP	SLASLSGEPVSF
ALQQTPFPSFPS	AQEQQAIMGGKL
NPCSHPTAWCTP	STNFSLESPHSW
QTPLNTLYTASS	GWPFHFTSDFAT
TYHTSYTENAGQ	

Table 3.3 Ligands for Tra1 determined by phage display. A subset of ligands isolated after 2 and 3 rounds of positive selection against Tra1. 500 μM VP2 was used in the binding buffer to isolate ligands that target different binding sites than those targeted by VP2.

Since the ELISA screen against Tra1 yielded ligands with similar sequence compositions, suggesting the ligands target a similar binding site on Tra1, we attempted to obtain ligands that targeted different binding sites on Tra1. For this purpose, the phage display was performed using two different conditions: phage incubated with Tra1, and phage incubated with Tra1 in a buffer containing the endogenous TAD VP2 (500 μM). We hypothesized that inclusion of VP2 would bias the selection *against* ligands that interact with the same surface(s) as VP2 and would enable investigation of the significance of endogenous activator binding site(s). Three rounds of positive selection were performed against Tra1 and one round of negative selection against maltose binding protein (MBP) after the first positive selection, since Tra1 was fused to MBP. DNA was extracted from the phage that were isolated under the various conditions and sequenced to obtain the encoding peptides (Table 3.3). In order to verify that the sequences obtained do in fact target Tra1, a phage-based ELISA was performed and the sequences that preferentially bound Tra1 were carried forward for activity assays in yeast (Figure 3.10). To test the ability of the ligands to activate transcription, they were fused to the Gal4 (1-147) DBD and evaluated in yeast reporter gene assays (Figure 3.11).

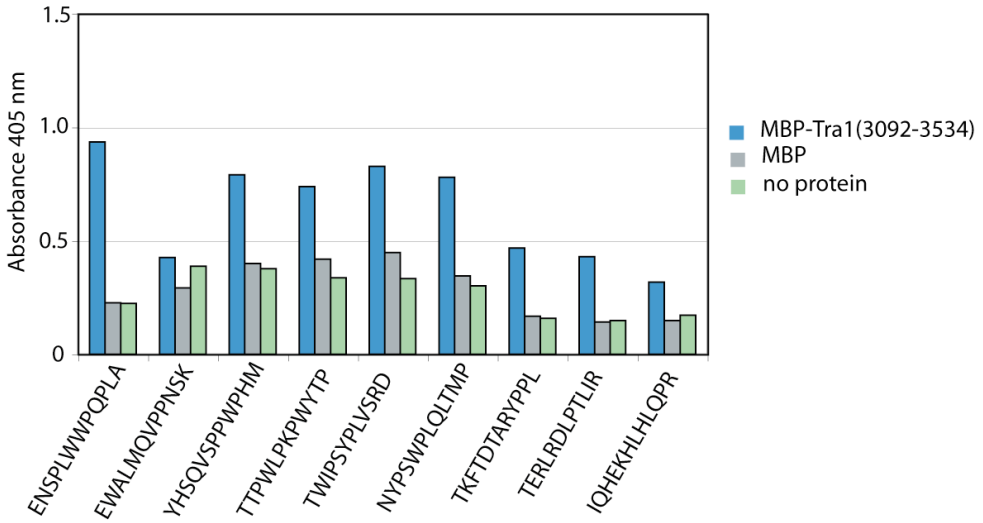


Figure 3.10 ELISA of ligands isolated from phage display. 10^{10} phage expressing select ligands isolated after 3 rounds of positive phage display selection against Tra1 were probed in an ELISA to determine if they bound Tra1 selectively over maltose binding protein (MBP) or buffer without protein. The bound phage were detected with an anti-phage antibody and visualized using colorimetric detection.

One ligand (ENSPLWWPQPLA) in particular was able to activate transcription to high levels. This ligand was obtained under the condition where no competing VP2 was used when the phage were selected to bind Tra1. Another active ligand, ALQQTPFPSFPS, was also obtained under these conditions. However, none of the ligands tested that were biased to bind sites other than VP2 activated transcription to high levels (Figure 3.11). The K_D for ENSPLWWPQPLA was, in contrast to the earlier Tra1 ligands, measurable (98 ± 8 μM) but is still >3-fold higher than that of VP2. Competitive ELISA experiments with phage-ENSPLWWPQPLA and VP2 showed that increasing concentrations of VP2 inhibited the binding of phage-ENSPLWWPQPLA to Tra1, suggesting that VP2 and ENSPLWWPQPLA target an overlapping site on Tra1 (Figure 3.13).

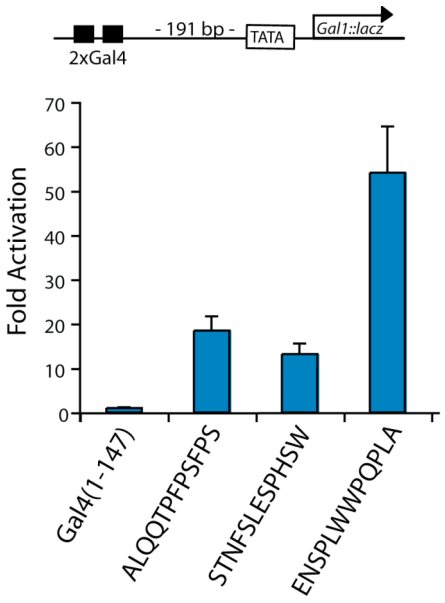


Figure 3.11
Evaluation of Tra1 ligands from phage display. Gal4+ligand fusions were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the Gal4+ligand fusion by the activity of the Gal4 DBD alone. Error is reported as SDOM from at least triplicate experiments.

A distinctive feature of molecules that function using a mechanism similar to endogenous TADs is that increasing their local concentration at the promoter, for example, by oligomerization, leads to transcriptional levels greater than the sum of those observed with the individual activators.[68, 71] Given the amphipathic nature of ENSPLWWPQPLA, it seemed possible that, similar to the natural amphipathic TADs that it resembles, it would function in a similar manner. To investigate this possibility, a dimer of ENSPLWWPQPLA was synthesized and found to bind Tra1 three-fold better (28 ± 1 μM) and correspondingly was able to activate transcription synergistically with greater than 95% confidence (four fold increase relative to the monomer) (Table 3.4 and Figure 3.12). In contrast the Med15(Gal11) ligands described in Chapter 2 did not display synergistic levels of transcription when multimerized (data not shown).[72] We attributed the lack of synergy with the Med15(Gal11) ligands to their Med15(Gal11)-specific mechanism of function. Thus, these finding are indicative of ENSPLWWPQPLA functioning in a manner similar to natural activators but distinct from the Med15(Gal11)-specific ligands.

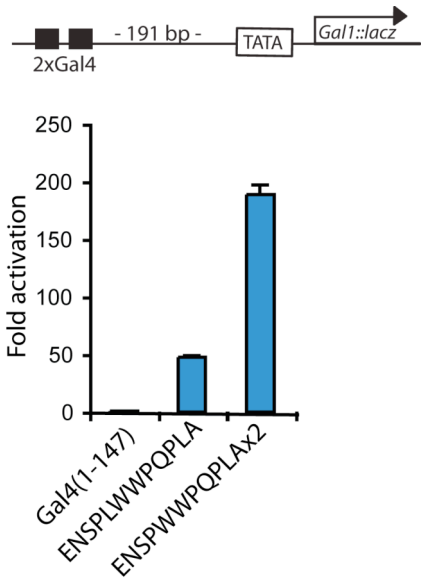


Figure 3.12
Dimerization of Tra1 ligand. Gal4+ligand fusions were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the Gal4+ligand fusion by the activity of the Gal4 DBD alone. Error is reported as SDOM from at least triplicate experiments.

The natural activator VP2, whose function the Tra1 ligand resembles, has been shown to interact with several coactivators in vitro. To investigate if the Tra1 ligand would interact with coactivators in addition to Tra1, the in vitro binding affinity for Med15(Gal11) (1-357), a protein fragment that interacts with several natural and unnatural TADs, was measured. It was found that ENSPLWWPQPLA bound Med15(Gal11) with a 81 ± 3 μM dissociation constant. The endogenous activator VP2 also targets this region of Med15(Gal11), although with a stronger affinity (3.8 ± 0.5 μM) (Table 3.4). Satisfyingly, competitive ELISA experiments with Med15(Gal11), phage-ENSPLWWPQPLA and VP2 indicated that VP2 and ENSPLWWPQPLA compete for the same binding site on Med15(Gal11) (Figure 3.13), although a VP2-induced conformational change in Med15(Gal11) that prevents ENSPLWWPQPLA from binding cannot be ruled out.

Ligand	MBP-Tra1(3092-5324)	Med15(1-357)
ENSPLWWPQPLA	98 ± 8 μM	81 ± 3 μM
ENSPLWWPQPLAgsggsgENSPLWWPQPLA	28 ± 1 μM	32 ± 2 μM

Table 3.4. Dissociation constants for Tra1 ligands. Fluorescence polarization was used to measure the dissociation constants for the fluorescein labeled ligands and Med15(1-357) or MBP-Tra1(3092-3524). The error reported is the SD from triplicate measurements.

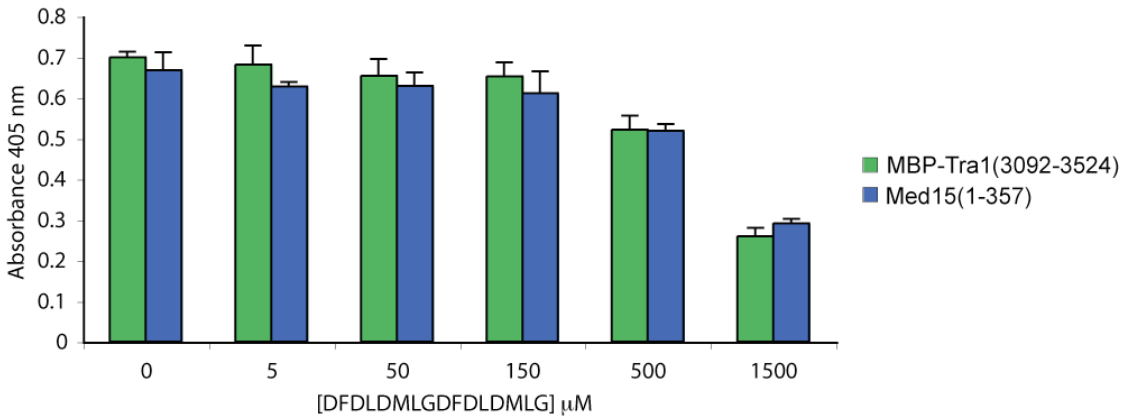


Figure 3.13 Competitive ELISA experiments with VP2 (DFDDMLGDFDLDMLG) and ENSPLWWPQPLA.
a) 10^{10} phage expressing ENSPLWWPQPLA with varying amount of VP2 were incubated with 96 well plates pre-bound with Med15(Gal11). The bound phage were detected using an anti-phage antibody and visualized by colorimetric detection. The values reported are an average of triplicate experiments and error is reported as SD
b) similar experiment as a) with Tra1 pre-bound on the plate.

To investigate the functional importance of the Med15(Gal11) interaction for the ability of ENSPLWWPQPLA to activate transcription in yeast, transcriptional assays were performed in yeast deleted for parts of Med15(Gal11). It was found that deletion of Med15(2-345) results in a reduction in the activity of ENSPLWWPQPLA, but as seen with the partially retained activity of the dimer, other interactions between the ligand and the transcriptional machinery, most likely with Tra1, are important for function (Figure 3.14). A similar functional dependence for Tra1 could not be performed in a straightforward manner due to the importance of Tra1 for yeast viability. Further, the functional reliance of ENSPLWWPQPLA on Med15(Gal11) is consistent with the Med15(Gal11) dependence of VP2, whose activity is also reduced in yeast deleted for Med15(2-345) (Chapter 4).

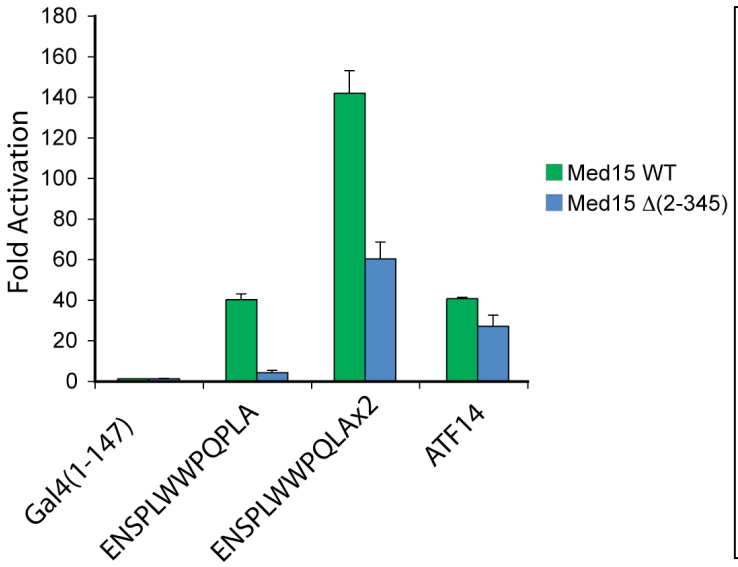


Figure 3.14 Med15(Gal11) dependence of Tra1 ligands. Tra1 ligands were evaluated for function in yeast strains with Med15Δ(2-345) or Med15 WT using quantitative β-galactosidase assays. Fold activation was determined by dividing the activity of the Gal4+ligand fusion by the activity of the Gal4 DBD alone. Error is reported as SDOM from at least triplicate experiments.

C.1.d Combinations with Med15(Gal11) ligands

In order to investigate the functional consequence of combining a Tra1 ligand with a Med15(Gal11)-specific ligand we decided to simultaneously introduce activator ATFs constructed from the Med15(Gal11) ligand AHYYYPSE and the Tra1 ligand ENSPLWWPQPLA into yeast. For this purpose, we used a yeast strain that had binding sites for the LexA and Gal4 DBDs upstream of the LacZ reporter gene. We attached the Med15(Gal11) specific ligand AHYYYPSE to Gal4(1-147) and ENSPLWWPQPLA to the LexA DBD (Figure 3.15a). To see a more pronounced effect, we also performed a similar experiment where we combined the AHYYYPSE dimer with the ENSPLWWPQPLA dimer. (Figure 3.15b) Satisfyingly, we found that the combination of these two ligands gave a statistically significant synergistic increase in transcription in yeast compared to either construct by itself.

While the mechanistic origin of synergy is still debated, two limiting models postulate that synergy could appear by simultaneously targeting multiple binding sites on the same transcriptional machinery target or due to the recruitment of multiple transcription proteins.[61, 73-77] Although, ENSPLWWPQPLA and AHYYYPSE target Med15(Gal11), they target distinct binding sites on Med15(Gal11) (data not shown). Previously, Lev Prasov in our group has shown that combination of ligands that target distinct surfaces of Med15(Gal11) leads to only additive levels of transcription.[72] Thus, it is likely that the synergy between AHYYYPSE and ENSPLWWPQPLA appears due to

the recruitment of Med15(Gal11) and Tra1. However, additional experiments will be needed to substantiate this model.

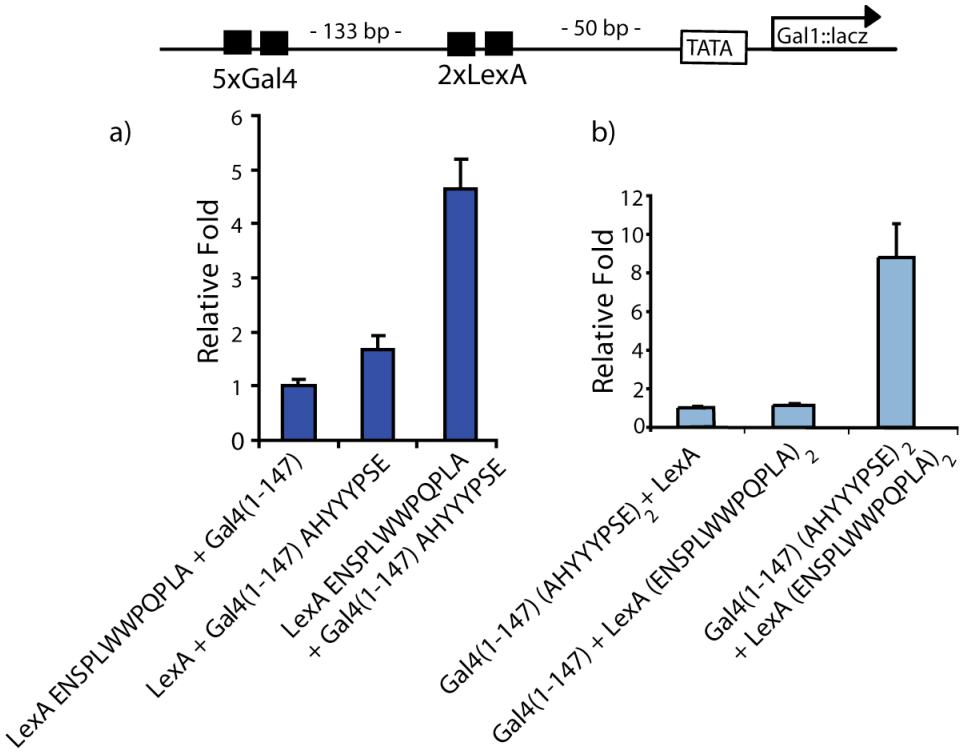


Figure 3.15 Synergistic activation of Med15(Gal11) and Tra1 Ligands. a) Monomers of Med15(Gal11) ligand (AHYYYPSE) fused to Gal4(1-147) and the Tra1 ligand (ENSPLWWPQPLA) fused to LexA were evaluated β-galactosidase assays in yeast. Relative fold is calculated by normalizing the activity of the first bar in each graph to 1. Error is reported as SDOM from at least triplicate experiments. b) same as a) expect dimers of ligands were used.

C.2 Engineering interactions outside the transcriptional machinery: incorporation of masking

The screen for Med15(Gal11) described in Chapter 2 resulted in peptides with several different sequence compositions that bound Med15(Gal11) with similar affinities. In terms of their ability to activate transcription, however, we found that only ligands with a moderately hydrophobic or amphipathic sequence were able to function as TADs. Interestingly, there were several hydrophobic ligands that bound Med15(Gal11) with low micromolar affinities, but were unable to activate transcription when fused to the LexA

DBD. Given the probable unstructured nature of these ligands in absence of a binding partner, coupled with their hydrophobicity, it appeared that proteolysis and non-specific interactions might be compromising their ability to activate transcription (Figure 3.16).[58, 78-80] We hypothesized that in a manner analogous to the artificial TAD XL_Y, incorporation of an intramolecular interaction into these ligands would provide structure to and/or mask the hydrophobic surfaces, converting them from inactivate peptides into potent transcriptional activators.[35]

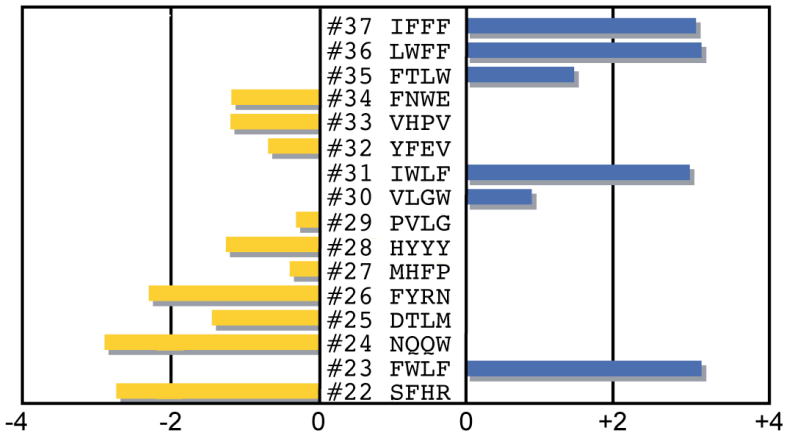


Figure 3.16 Mean hydrophobicity of Med15(Gal11) ligands. Mean hydrophobicity was determined for each peptide in library 2 (see Chapter 2 for screening details) using a hydrophobicity calculator (CCS scale). http://www.bbcm.univ.trieste.it/~tossi/HydroCalc/HydroMCalc.html

To test this hypothesis, we chose Gal4(1-100), a DBD that could provide an intramolecular binding surface(s) for at least a subset of the ligands. This region comprises a binuclear zinc cluster DNA binding domain (1-40) as well as a dimerization domain (40-100) that contains at least three binding surfaces that are known to interact with hydrophobic partners (Figure 3.17).[35, 81, 82] Residues in helices 1 and 2 are critical for dimerization of Gal4 and for DNA binding. The two other surfaces facilitate contacts that mediate transcriptional activation. Specifically, in the presence of a point mutation in the transcription protein Med15(Gal11)(N342V) termed Gal11P, residues in loop 1 and helix 3 of Gal4 are able to recruit Gal11P, converting this surface into a potent transcriptional activator.[83] The third surface primarily consisting of residues in loop 1 was found to be critical for the interaction with XL_Y. Of the three binding partners, the ligands bear most similarity to XL_Y in terms of hydrophobic content

(LTGFVQDYLLPTCIP, calculated mean hydrophobicity 1.82) and size (8 amino acid residues versus 16). Thus, Gal4(1-100) seemed an ideal choice to implement our strategy of incorporating a masking interaction to probe its importance in transcriptional activation.

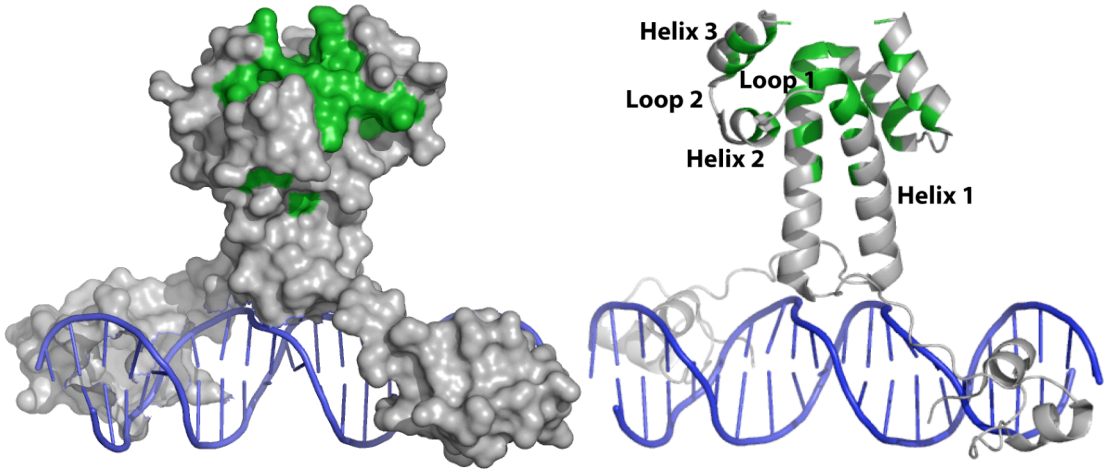


Figure 3.17 Crystal Structure of the Gal4 DNA binding and dimerization domains. Gal4(1-96) bound to its cognate DNA (blue). Surface rendering and ribbon cartoons of the structure showing the hydrophobic residues in green. The different structural features are indicated. PDB ID: 3COQ

Dr. Jenifer Lum, a former graduate student in the laboratory, made yeast expression plasmids for all 37 Med15(Gal11) ligands described in Chapter 2, fused to the Gal4(1-100) DBD. She initially screened the Gal4(1-100)+ligand fusions in an X-gal filter plate assay versus the positive control Gal4(1-100)+VP2, a reiteration of a transcriptionally active segment of the viral coactivator VP16 that is commonly used in activator ATFs and the negative control, Gal4(1-100). Qualitatively, ligands #23, #31, #34, #35, #36 and #37 displayed significant levels of activity. A more quantitative *o*-nitrophenol-galactoside (ONPG) liquid assay revealed a dramatic increase in activity for several of the ligands from library 2 (AXXXXPSE) that appeared active in the plate assay. Ligand #36, for example, only showed 1.2-fold levels of activity when attached to the LexA DBD, but upon attachment to Gal4(1-100) displayed a 600-fold increase in activity relative to the Gal4 DBD alone. The enhancement in activity correlates well with the mean hydrophobicity of the variable region within the ligands, where the most active ligands also have the highest mean hydrophobicity (Figures 3.16 and 3.18).

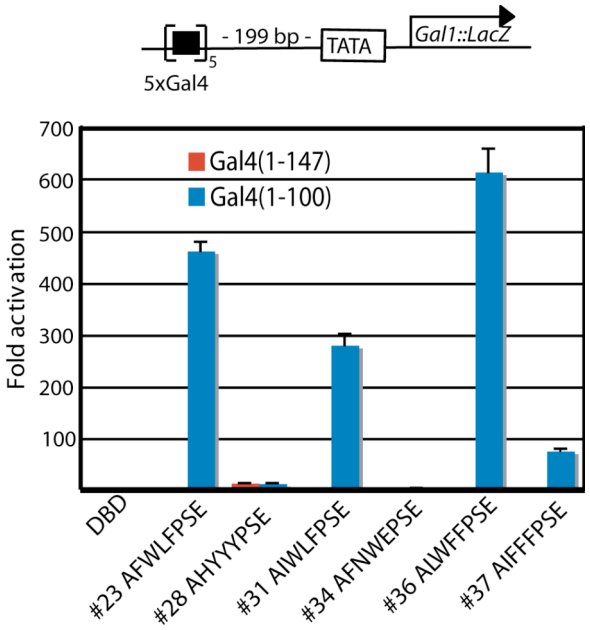


Figure 3.18 Activity of Med15(Gal11) ligands on Gal4(1-100) vs Gal4(1-147). Ligand+Gal4 fusion were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the ligand+Gal4 fusion by the activity of Gal4 alone. Error is reported as SDOM from triplicate experiments.

Not surprisingly, it was found that the dimerization domain of Gal4 played a role in the activity of the ligands; close proximity of the ligands to this domain is important to maintain the increased activity observed with #23, #31, #36 and #37 (Figure 3.18). We made fusions of these ligands to the Gal4(1–147) DBD and this negates the activity enhancement seen when they are attached to Gal4(1-100). Probably the structure of Gal4(1–147) differs from Gal4(1–100), with a portion of the hydrophobic binding surfaces in the Gal4(52–100) region obscured.[84]

C.2.a Unmasking the masking interaction

The enhancement in activity of the ligands when attached to Gal4(1-100) and its correlation to ligand hydrophobicity was suggestive of an interaction between the hydrophobic dimerization domain of Gal4 and the ligands. To directly probe this, we used fluorescence polarization to assess the affinities of the different ligands for Gal4(1–100). Initially, we decided to focus the binding studies within the hydrophobic dimerization domain of Gal4. Towards this end, the ligands were synthesized using standard FMOC solid phase synthesis protocols and labeled with fluorescein at the amino-terminus in solution. The affinity for bacterially expressed GST-Gal4(52-100) was measured and observed to correlate with the activity enhancement (data not shown). Further, the ligands were labeled at the C-terminus via a thiol linkage from a cysteine residue. The trends in binding affinity were similar to the amino-terminal labeled ligands,

with the ligands that bound the best showing the highest activity enhancement. Nonetheless, the affinities were slightly weaker, indicating the importance of a more accessible carboxy-terminus. To our dismay, control experiments with the fluorescein labeled ligands and GST alone showed similar binding affinities as the GST-Gal4(52-100) fusion.

To eliminate the interaction with GST, GST-Gal4(1-100) was bacterially expressed, and the GST tag was cleaved off using precision protease. The resulting free Gal4(1-100) was then purified by gel filtration chromatography. The binding affinity of the ligands for Gal4(1-100) was measured and gratifyingly, the ligands bound with low micromolar affinities, with a trend that followed the activity enhancement (Figure 3.19). Ligands #36 and #23 which show the greatest activity enhancement when attached to Gal4(1–100) also interact most tightly (K_Ds of 2.6 ± 0.1 and 3.1 ± 0.1 μM, respectively), ligand #31 exhibits a 2- to 3-fold higher K_D corresponding to its smaller activity enhancement (7.3 ± 0.2 μM). Ligand #37, which shows the smallest activity enhancement of the four, also binds, but as a result of the propensity of Gal4(1–100) to aggregate at concentrations above 15–20 μM, a complete binding curve could not be obtained. Interestingly, ligand #28 and VP2, TADs that activate transcription but do not show enhancement when attached to Gal4(1-100) do not detectably bind. Another noteworthy result showed that the ligands bound Gal4(1-100) with a similar affinity irrespective of the presence of its cognate DNA, indicating that the ligands can bind Gal4(1-100) in its DNA-bound and unbound conformations (data not shown). These results indicate that the interaction with Gal4(1-100) can potentially take place when the activator is free in the cellular environment, when in absence of a masking interaction it is most likely to undergo non-specific interactions and/or proteolysis.[58, 78, 85]

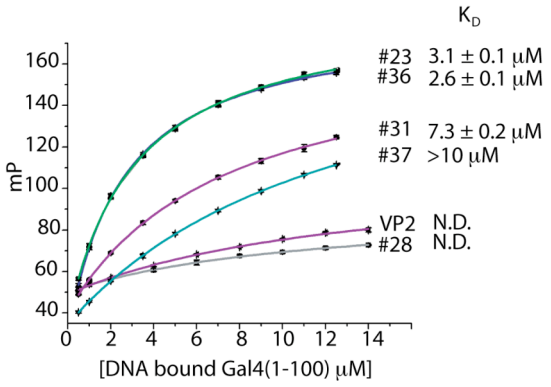


Figure 3.19 Dissociation constants of Med15(Gal11) ligands with Gal4(1-100). Each ligand was incubated with a range of concentrations of DNA-bound Gal4(1-100) and fluorescence polarization was monitored. mP is millipolarization units and errors are reported as SD. N.D. = not detected.

The in vitro binding studies between Gal4(1-100) and the ligands are consistent with a masking interaction between the ligands and the DBD. It is only when the ligands are in close proximity to Gal4(1-100), that they are able to activate transcription to high levels. To confirm an interaction in a cellular context, Jenifer Lum carried out alanine scanning mutagenesis of the dimerization domain Gal4(52-100) and measured the activity of #36, the most potent ligand, using quantitative β-galactosidase assays in each case. Alanine substitution at three residues (76, 77, and 84) resulted in a dramatic decrease (≥80%) in activity. Other residues (81, 82, and 90) had a more moderate effect on function (Figure 3.20). Residues 76, 77 and 84 define a binding interface at the second helix and first loop of the dimerization domain, a conformationally mobile and solvent-exposed region of the structure. Notably, these mutations had no effect on the activity of a control TAD, VP2, that shows no enhanced activity when attached to Gal4(1–100). Substitution at residues 69, 70, and 75 increased activity 3- to 9-fold; given their position relative to the binding site, these alanine substitutions likely increase the accessibility of the binding site. Substitution of residues 97 and 89 similarly produced an increase in activity. In this instance, attenuated dimerization of helix 3 may facilitate interaction of the activation peptide with the loop 1 binding site. In addition, the NMR structure of $Gal4_{dd}$ reveals that residues 97–100 are largely unstructured and thus may provide needed flexibility such that key residues of the activation peptide can reach the binding site.[82] Based on the NMR and crystal structures, Leu77 is ~15-25 Å from Phe97, and when extended, Gal4(97–100)+SS+ALWFFPSE is estimated to be ~35 Å long, within range of the proposed binding site. The mutagenesis studies together with the binding experiments reinforce the importance of Gal4(1-100) in the activity enhancement of these ligands and is consistent with the hypothesis that it mimics a masking interaction.

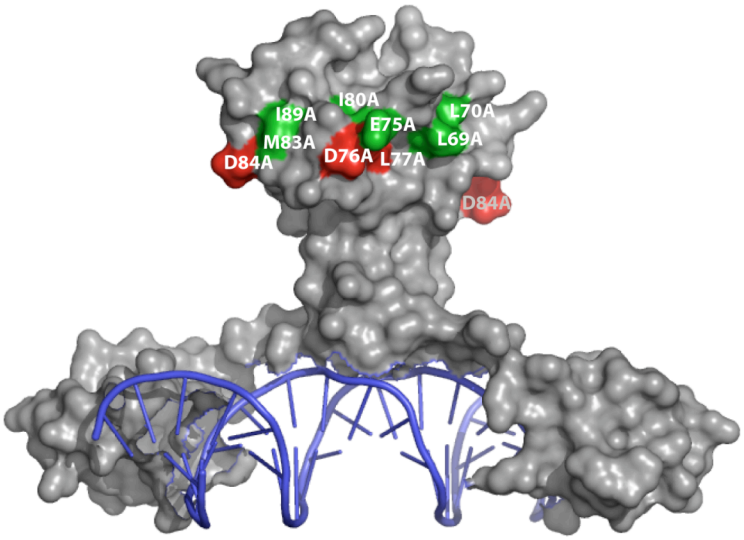


Figure 3.20 $Gal4_{dd}$ residues that alter activation. Surface rendering of the crystal structure of the Gal4(1–96) dimer. The residues in Gal4 that upon replacement with alanine increase the activity of Lex(1-202)+Gal4(40-100)+ALWFFPSE (#36) are shaded in green, while the residues that significantly decrease the activity are indicated in red.

C.2.b Expanding the scope

Hydrophobicity of the Med15(Gal11) ligands is an important factor in the enhancement of function when attached to Gal4(1-100) (Figure 3.16 and 3.18). Natural and artificial TADs also possess a preponderance of hydrophobic residues critical for function and are often unstructured in the absence of a protein binding partner. Thus, it might be possible to incorporate the masking interaction into other TADs to enhance their function. To test this idea, Jenifer Lum made yeast expression plasmids for several relatively short, non-natural peptide TADs fused to Gal4(1-100) or Gal4(1-147) and evaluated their ability to activate transcription (Figure 3.21a). KBP1.66 and KBP2.20 were isolated from a phage display screen against the mammalian coactivator CREB binding protein, CBP.[70] AH was designed to generally mimic the amphipathic helix (AH) motif found in most natural activators.[86] These ligands show no direct sequence overlap with natural yeast activators and thus are unlikely to interact with endogenous masking proteins that regulate stability and activity. G80bpA arose from a phage display screen from binding to the yeast repressor Gal80, nonetheless, it is not known to be regulated by Gal80.[68, 87] KBP1.66, KBP2.20 and Gal80bpA display a significant increase in activity upon attachment to Gal4(1-100) compared with Gal4(1-147) (49-, 49-, and 79-fold enhancement, respectively). In contrast, AH did not display a significant difference when attached to either Gal4(1-100) or Gal4(1-147). It is possible that the hydrophobic residues in AH are not well placed for an interaction with the DBD, perhaps as a result of a significant secondary structure assumed by the TAD (Figure 3.21b).

Jenifer Lum also evaluated the natural activation domains of VP16, Gal4, Gcn4, and p53. These TADs also contain an excess of hydrophobic residues and the importance of these residues in mediating up-regulation has been well documented.[5, 30, 66, 67, 88, 89] Interestingly, ATF11, Gcn4 and p53 showed an activity enhancement. In contrast, VP2 and Gal4 did not increase in activity upon attachment to Gal4(1–100) (Figure 3.20c). Correspondingly, we did not see a significant interaction between VP2 and Gal4(1-100) in vitro (Figure 3.19).

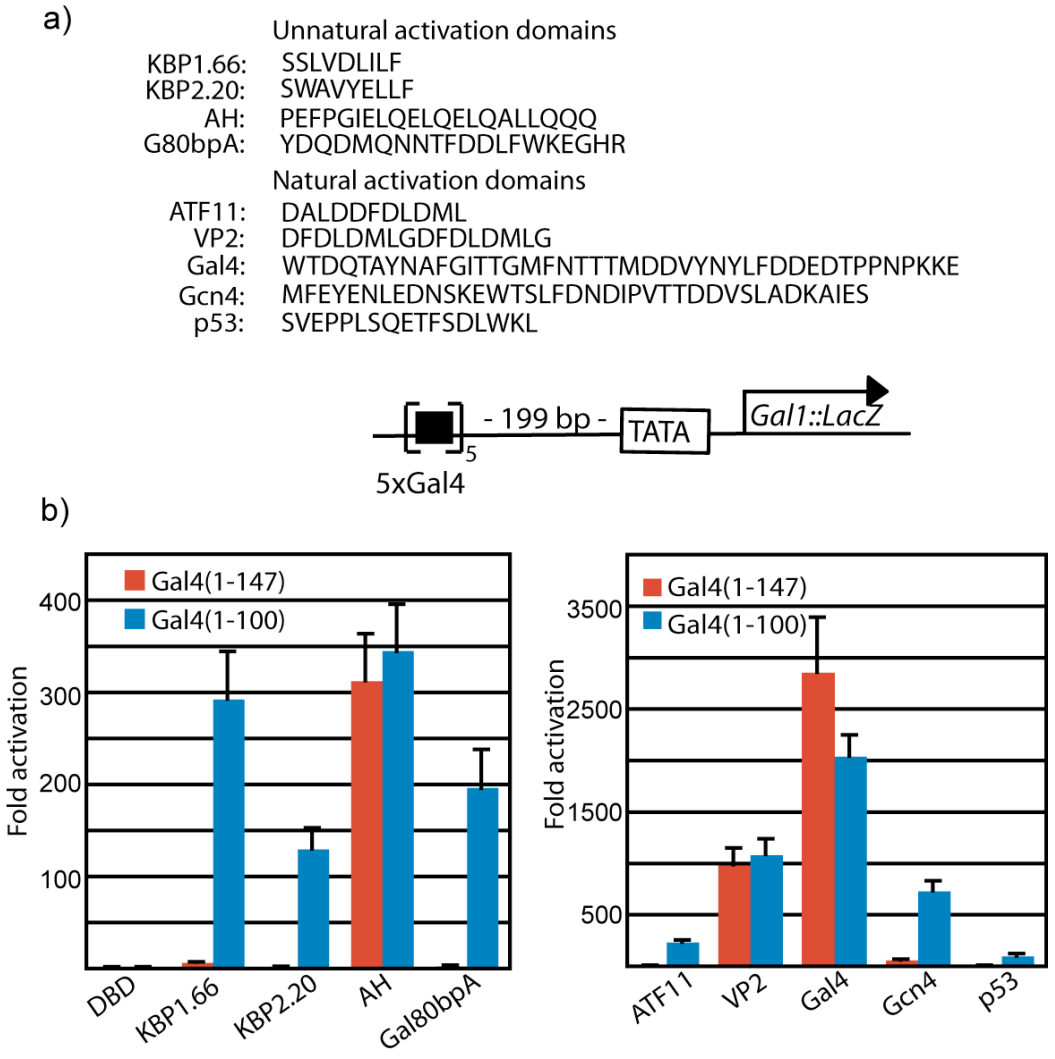


Figure 3.21 Enhancing the function of natural and unnatural TADs.
a) Sequences of unnatural and natural TADs investigated. b) Unnatural TADs fused to Gal4(1-100) or Gal4(1-147). c) Natural TADs fused to Gal4(1-100) or Gal4(1-147).Ligand+Gal4 fusions were evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the ligand+Gal4 fusion by the activity of Gal4 alone. Error is reported as SDOM from triplicate experiments.

The Tra1 ligand ENSPLWWPQPLA also contains a number of hydrophobic residues. To probe if attachment to Gal4(1-100) would also assist its activity, we compared its activity on Gal4(1-100) versus Gal4(1-147) (Figure 3.22). In this instance also we observed an activity increase on attachment to Gal4(1-100), suggesting that the function of a wide variety of hydrophobic TADs can be enhanced in this context.

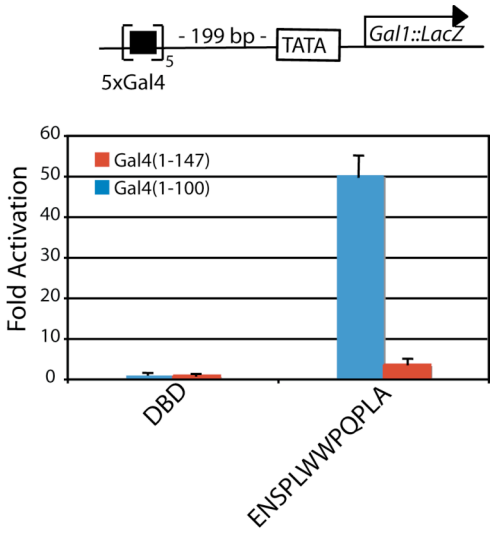


Figure 3.22 Enhancement of the function of the Tra1 ligand.
ENSPLWWPQPLA was fused to Gal4(1-100) or Gal4(1-147) and evaluated in β-galactosidase assays in yeast. Fold activation was determined by dividing the activity of the ligand+Gal4 fusion by the activity of Gal4 alone. Error is reported as SDOM from triplicate experiments.

D. Conclusions and future directions

We have shown that natural activators interact with the C-terminus of Tra1 and ligands that target this region when used fused to a DBD activated transcription to levels higher than the Med15(Gal11)-specific ligands. In particular, we found that a ligand that targets an overlapping binding site with the endogenous activator VP2 activated transcription to the highest level. Further this ligand was found to interact with another coactivator Med15(Gal11) at a shared binding site with VP2. These findings suggest that activator binding site(s) on different coactivators are similar. Thus, in the future, strategies aimed at deciphering endogenous activator binding sites will be extremely useful for activator ATF development; possibly targeting one such site would lead to molecules capable of activating transcription to high levels, as they potentially will be able to interact with similar sites on multiple coactivators (Chapter 4).

We were also able to enhance the function of the Tra1 ligands. Combination of the Tra1 and Med15(Gal11) ligands creating multiple interactions with the transcriptional machinery at the activator-relevant targets Med15(Gal11) and Tra1 led to synergistic levels of gene expression. Thus, it seems likely that endogenous activators attain high

levels of activity by interacting with more than one protein in the transcriptional machinery. Approaches to identify all the functionally relevant targets of natural activators will assist in creating robust synthetic activators by incorporating all these interactions (Chapter 4).

Interaction(s) outside of the transcriptional machinery also have a profoundly positive effect on transcription function. Micromolar ligands for the transcriptional machinery, for example, only function as robust transcription activators when such a secondary binding interaction is available. Given the hydrophobic nature of these ligands, it is likely that the secondary interaction alters the binding and stability profile of the ligands, decreasing non-specific binding and premature proteolysis that would reduce the effective functional concentration of the ligands available for transcriptional activation. Consistent with this notion, we have found using western blots that the stability of the ligands is altered when fused to Gal4(1-100) versus Gal4(1-147).[90] Moreover, this strategy of incorporating a secondary "masking" interaction appears general; several natural and artificial TADs had their activity significantly enhanced in vivo via this mechanism. Inclusion of such a strategy in future generations for small molecule and non-biopolymer-based TADs could assist in limiting non-productive interactions, facilitating their delivery to their transcriptional machinery targets and greatly elevating their function.[91, 92]

E. Experimental

Peptide synthesis of TADs

FMOC-based solid phase peptide synthesis was used to synthesize the VP2, Gal4(840-881) and Gcn4(105-134) peptides. For this purpose, an ABI 433A peptide synthesizer was used and the peptides were prepared using clear amide resin (0.40 mmol/g, Peptides International) at 0.1 mmole scale using manufacturer recommended conditions, except for forced double coupling of all residues in the synthesis of Gal4 and Gcn4. The Med15 and Tra1 ligands were synthesized using a ATT 90 peptide synthesizer (Advanced ChemTech). A cysteine was added to the C-terminus of Gal4, Gcn4 to facilitate labeling with fluorescein.

After completion of automated synthesis, all the peptides were fully deprotected and cleaved off the resin using 95% TFA, 2.5% Triisopropylsilane (TIS), and 2.5% water (5mL, 2 h). Subsequently, the solution was evaporated by blowing nitrogen to <1 mL and precipitated with cold ether (3x10 mL). Without further purification, the Tra1, Med15 ligands and the VP2 peptide were labeled at the N terminus using fluorescein 5,6 succinimidyl ester (Pierce). For this purpose, ~1 mg of crude peptide was dissolved in 20μL of DMF and 50μL CH_3CN followed by 500μL of 100mM $NaHCO_3$ pH 8.6. Approximately 10 fold excess of fluorescein 5,6 succinimidyl ester dissolved in 10μL DMF and 50μL CH_3CN was added to the peptide solution and the coupling was carried out at 37 ºC overnight. The product was characterized by electrospray mass spectrometry and purified to homogeneity using reversed-phase HPLC on a C18 column with a gradient solvent system (Buffer A: 0.1% TFA, Buffer B: CH_3CN) and stored dry at -20 ºC.

To facilitate the labeling of the Gal4, Gcn4 TADs, ~1 mg of TAD was dissolved in 50μL DMF and 50μL CH3CN followed by 500μL 0.1% TFA. To this ~10 fold excess of fluorescein maleimide (Pierce) dissolved in 50μL DMF was added and the reaction was carried out at 37 ºC overnight. The products were characterized and purified as described above.

For competition binding studies, unlabeled peptides were purified to homogeneity using reversed-phase HPLC as described above.

Synthesis of peptide library

The AXXXXPSE peptide library was synthesized using split-pool synthesis. 0.5 mmole of PEGA-NH2 resin (Novabiochem, 0.4 mmol/g) was manually coupled with 5 equivalents of HOBT, HBTU, diisopropyl ethyl amine and each amino acid. The PEGA resin was chosen due to its hydrophilicity relative to more commonly used
PEG-grafted polystyrene resins and because it has a long PEG chain tethering the peptide to the solid support, facilitating interaction with large proteins. After the addition of the residues PSE, the resin was split into 19 reaction tubes and one of 19 endogenous amino acids was added to each tube (no lysine). After the coupling proceeded for ~12 h, the resin from all the reactions was pooled, FMOC deprotected (3x10mL 20% piperidine in

DMF, 10 min) and again split into 19 reactions to repeat the same process, 3 more times. Finally, the resin was pooled and alanine was added to complete the synthesis of the library. After removal of all protecting group using 95% TFA, 2.5% TIS, 2.5% water for 4 h, the library was washed extensively with MeOH and stored dry at -20 °C.

Plasmid Construction

DNA Oligos for plasmid construction were purchased from Invitrogen or Integrated DNA Technologies (IDT).

Tra1 expression plasmids

Plasmids for bacterially expressing Tra1 were generated from pMal-c2g (New England Biolabs) or pGEX6p-1 (GE healthcare). This plasmids express proteins as fusions to MBP or GST respectively and are ampilcillin selectable in *E. coli*.

Primers 1905 Fwd (5' - CTG GTC GTG AAT GGA TCC GCG TAC CTG GTT ACA TCA TAT - 3') and 2358 rev (5' - GTA GGC TTG GAC CTC GAG CTA GCT TCT GGA CAT ATT GAC ATT - 3'); 2162 Fwd (5' - CTG GTC GTG AGC GGA TCC ATG AAT GCT TTG GAT GTC G - 3') and 2588 rev (5' - GTA TGC ATG TAA CTC GAG CTA AGA AAG CAA TGT AAT GAT AGA TCG - 3'); 2478 Fwd 5' - CTG GTC GTG AAT GGA TCC GGT TCT TTT AAT AGA GAG AGA - 3' and 2902 rev (5' - GTA GGC TTG TCA CTC GAG CTA TAC TAG GTC ATT CCA CAT ATT - 3'); 2693 Fwd (5' - CTG GTC GTG AAT GGA TCC TCA TAT GAA CAA ATT GGC CTT - 3') and 3140 Rev (5' - GTA GGC CTG CCA CTC GAG CTA TAG TTG AGG AAT GAA AGT AAT - 3'); 3092 Fwd (5' - CTG GTC GTG AAT GGA TCC TAT AAG AAC TCG AAG ATT AGG - 3') and 3524 rev (5' - GTA GGC CTG CCA CTC GAG CTA TGA TGG TAC AAA CAT TGT TTG - 3') were used to amplify Tra1 fragments from yeast genomic DNA for incorporation in pGEX6p-1. Expand high fidelity DNA polymerase (Roche) was used to amplify the products with 400 nM of each primer and 4.5mM $MgCl_2$ and 1μL of yeast genome DNA template per 50 μL reaction. The amplification conditions are [94 °C 2', 9x (94 °C 15", 52 °C 30", 68 °C 2'), 24x (94 °C 15", 59 °C 30", 68 ° C 2'), 68 °C 7', 4 °C hold]. A total of 4, 50 μL PCRs were performed for each Tra1 fragment. The PCR products were purified using a Qiagen PCR purification kit and

subsequently double digested with BamHI and EcoRI. The digestion reaction was purified using the Qiagen PCR purification kit and ligated with T4 DNA ligase into pGEX6p-1 predigested sequentially with BamHI/EcoRI and treated with calf intestinal phosphatase. The resulting plasmids were amplified in Smart *E. coli* cells (Genlantis), selected on LB-agar plates containing 100 μg/ml ampicillin, and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

MBP-Tra1 constructs were created using the same fwd primers as used for GST-Tra1. The reverse primers used are 2358 Rev (5'- GTA CGC ATG GAC GTC GAC CTA GCT TCT GGA CAT ATT GAC AAT - 3') 2588 Rev (5' - GCA CGC AGG GAC GTC GAC CTA AGA AAG CAA TGT AAT GAT AGA - 3') 2902 rev (5' - GTA CGC AGG GAC GTC GAC CTA TAC TAG GTC ATT CCA CAT ATT - 3') 3140 rev (5' - GTA CGC AGG GAC GTC GAC CTA TAG TTG AGG AAT GAA AGT AAT - 3') 3524 rev (5' - GCA CGC AGG GAC GTC GAC CTA TGA TGG TAC AAA CAT TGT TTG - 3'). The GST Tra1 fusions were used as a template to create the MBP fusions. Pfu turbo (Stratagene) was used in these PCRs to improve fidelity. 10 ng of template, 3.5mM $MgCl_2$ and 200 nM of each primer was used in each 50 μL amplification reaction using the same PCR protocol as described above. The PCR products were purified using a Qiagen PCR purification kit and subsequently double digested with BamHI and SalI. The digestion reaction was purified using the Qiagen PCR purification kit and ligated with T4 DNA ligase into pMal-c2g sequentially predigested with BamHI/SalI and treated with calf intestinal phosphatase. The resulting plasmids were amplified in Smart *E. coli* cells (Genlantis), selected on LB-agar plates containing 100 μg/ml ampicillin, and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

MBP plasmid

To create MBP-stop, a plasmid that only expressed MBP, a stop codon was inserted after MBP in pMal-c2g. Oligos (5'- GAT CCT AGT GAT GGC CAG -3')and (5' TCG ACT GGC CAT CAC TAG -3') were annealed, phosphorylated and ligated into BamHI/XbaI sequentially digested pMal-C2g and CIP treated as described in Chapter 2.

Expression Plasmid for GST-Gal4(52-100) and Gal4(1-100)

The plasmid expressing GST-Gal4(52-100) was a gift from Dr. Aseem Ansari.

Gal4(1-100) was amplified from yeast genomic DNA using primers (5'- cat gga tcc atg aag cta ctg tct tc -3') and (5'- cat gaa ttc tta atc ttg tac aaa taa tcc tg - 3') containing BamH1 and EcoR1 restriction sites. The plasmid was constructed as described above by inserting into pGEX6p-1.

Med15(Gal11)(2-345) delete

Ycplac111 Med15Δ(2-345) expressing Med15(Gal11) on a low copy plasmid under the control of the native Med15(Gal11) promoter in yeast was generated by Jenifer Lum by performing site-directed mutagenesis to introduce XhoI restriction sites at amino acids 2 and 345 in Ycplac11 Med15 WT. Subsequently, the plasmid was digested with XhoI and re-ligated to form ycplac111 Med15Δ(2-345).

Tra1 ligand plasmids

Plasmids expressing the Tra1 ligands on Gal4(1-147) were constructed by insertion of DNA oligos encoding the ligands into pGBKT7 as described in Chapter 2. pGBKT7 is high copy in yeast, contains a tryptophan marker and expresses Gal4 fusions under the control of the ADH1 promoter.

Plasmids used for masking experiments

Plasmids expressing the Med15 ligands or the Tra1 ligand on Gal4(1-100) or Gal4(1-147) were constructed by insertion of DNA oligos encoding the ligands into XbaI/SalI digested NYC317 (Gift from Dr. Aseem Ansari) or NYC317-147 (constructed by Jenifer Lum) as described in Chapter 2. These plasmids are low copy in yeast, contains a histadine marker and expresses Gal4 fusions under the control of the β-actin promoter.

Expression of Tra1 constructs

Plasmids for the expression of each Tra1 fragment fused to GST or MBP are described above. The appropriate plasmid was transformed into chemically competent

Rosetta2(DE3) pLysS *E. coli* (Novagen) and cells were plated onto LB-agar plates supplemented with ampicillin (100 μg/mL) and chloramphenicol (34μg/mL). Cultures (50 mL) from single colonies were grown overnight at 37 °C (300 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100μg/mL) and chloramphenicol (34μg/mL) before addition to 1L of Select APS Super Broth supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL). After an OD_{600} of 0.3 was reached (37 °C, 300 rpm), the cultures were cooled to 16 °C for 1 h (100 rpm), and expression was induced with IPTG (final concentration 0.1 mM) for 12 h (250 rpm). The frozen cell pellet was resuspended in Buffer A (100 mM PBS pH 7.4, 10% glycerol (v/v), 1 mM DTT, 0.2% NP-40 (w/v) and Roche Complete Protease Inhibitor Cocktail) lysed using sonication, and the GST-tagged protein was isolated using glutathione-sepharose beads (GE Healthcare) The MBP-tagged protein was isolated using amylose resin (New England Biolabs). The protein was eluted from the beads overnight at 4 ºC using Elution buffer (50mM Tris pH 8.0 0.1%, NP-40 and 15mM Glutathione) for GST tagged proteins or (100 mM PBS pH 7.4, 10% glycerol (v/v), 1 mM DTT, 0.2% NP-40 (w/v) and 10 mM maltose) for MBP fusions. The protein solution was then buffer exchanged to Storage buffer (10 mM PBS, pH 7.4, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40) using a PD-10 column (GE Healthcare). Finally, the protein was concentrated to 20-200 μM using a Vivascience 30K centrifugal filter device (polyethersulfone membrance). The protein concentration was measured using a Bradford assay (Bio-Rad) with BSA as the standard. The identity and purity of the fusion protein was verified by reducing SDS-PAGE with appropriate molecular weight standards. The protein solution was stored at –80 °C until needed.

For use in the ELISA screen and phage display selection, MBP-Tra1(3092-3524) was further purified using size exclusion chromatography. For this purpose, after elution from the amylose resin, the protein was concentrated to ~3 mL using a Vivascience 30K centrifugal filter device (polyethersulfone membrane) and then loaded onto a gel filtration column (16/60 Superdex 200, GE Healthcare) pre-equilibrated with (10 mM PBS, pH 7.0, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40). Fractions containing MBP-Tra1(3092-3524) were pooled and concentrated using a Vivascience 30K centrifugal filter device. MBP was expressed and purified using the MBP-stop plasmid

described above and using identical conditions for MBP-Tra1 except that it was only induced with IPTG for 3 h at 16 °C.

Expression of Gal4(1-100)

The GST-Gal4(1-100) or GST-Gal4(52-100) was transformed into chemically competent Rosetta2(DE3) pLysS *E. coli* (Novagen) and cells were plated onto LB-agar plates supplemented with ampicillin (100 μg/mL) and chloramphenicol (34μg/mL). Cultures (50 mL) from single colonies were grown overnight at 37 °C (300 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100μg/mL) and chloramphenicol (34μg/mL) before addition to 1L of Select APS Super Broth supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL). After an OD_{600} of 0.3 was reached (37 °C, 300 rpm), the cultures were cooled to 16 °C for 1 h (100 rpm), and expression was induced with IPTG (final concentration 0.1 mM) and 20 μM zinc sulfate (only for Gal4(1-100)) for 1.5 h (250 rpm). The frozen cell pellet was resuspended in Buffer A (100 mM PBS pH 7.4, 10% glycerol (v/v), 1 mM DTT, 0.2% NP-40 (w/v) and Roche Complete Protease Inhibitor Cocktail) lysed using sonication, and the GST-tagged protein was isolated using glutathione-sepharose beads (GE Healthcare). The protein was eluted from the beads overnight at 4 ºC using Elution buffer (50mM Tris pH 8.0 0.1%, NP-40 and 15mM Glutathione). The protein solution was then buffer exchanged to Storage buffer (10 mM PBS, pH 7.4, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40) using a PD-10 column (GE Healthcare). Finally, the protein was concentrated to 10-50 μM using a Vivascience 10K centrifugal filter device (polyethersulfone membrane). The protein solution was stored at –80 °C until needed. Cleavage of the GST tag was accomplished overnight at 4 °C using PreScission Protease (GE Healthcare) in Cleavage Buffer (50 mM Tris pH 7.0, 0.2% NP-40 (w/v), 1 mM DTT). The Gal4(1-100) solution thus obtained was concentrated using a Vivascience 10K centrifugal filter device and then loaded onto a gel filtration column (Superdex 75, GE Healthcare) to remove residual GST and to exchange the buffer to Storage Buffer. Fractions containing Gal4 (1-100) were pooled and concentrated using a Vivascience 10K centrifugal filter device. The protein concentration was measured using a Bradford assay (Bio-Rad) with BSA as the

standard. The identity and purity of the fusion protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

Expression of GST using the parental pGEX6p-1 plasmid was accomplished using an identical procedure as described above, except only 2 x 250 mL superbroth was used instead of 4 x 1L for GST-Gal4(1-100).

Measurement of Dissociation Constants

Dissociation constant measurements were carried out on a TECAN Genios Pro Plate reader (for Tra1 studies) and a Panvera Beacon 2000 (for Gal4 dimerization domain studies) at room temperature.

Prior to each Tra1 binding experiment, fluorescein-labeled ligand was resuspended in 1 mL of Storage buffer (10mM PBS pH 7.0, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40) and the concentration of the solution was determined by UV-Vis using the absorbance of fluorescein, λ=495nm (ε 66,000 M^{-1} cm^{-1} for fluorescein NHS ester and ε 72,000 M^{-1} cm^{-1} for fluorescein maleimide). Two stock solutions were prepared, 50 nM labeled peptide solution in Storage buffer, and ~100 μM Tra1 in Storage buffer containing a final labeled peptide concentration of 50 nM. 20μL of ~100 μM Tra1+50 nM labeled peptide was added to the first well of a low volume 384-well plate (Corning). To each successive well, a certain amount of the two stock solutions were mixed to obtain a range of Tra1 concentrations from ~100 μM - 0.1 μM, keeping concentrations of the labeled ligand constant throughout the experiment. In each case, 20 μL of mixture was added to the well. The plate was incubated for 10 min at room temperature and the fluorescence polarization at each Tra1 concentration was measured. The raw data was imported into Origin and fit to equation 1 using the Marquardt-Levenberg nonlinear least squares regression:

$$y = \left(\frac{a \times x}{b + x}\right) + c \qquad \textbf{Equation 1}$$

where a is maximum anisotropy – minimum anisotropy, x is protein concentration, b is the dissociation constant, c is the polarization of the free labeled peptide (no protein), and

y is response. For each experiment, the $R^2 > 0.98$ and each K_D represents the average of three separate experiments.

For binding experiments with Gal4(1-100) or Gal4(52-100), fluorescein-labeled ligand was resuspended in 1 mL of Storage buffer (10 mM PBS pH 7.4, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40) as described above. A 100 μL solution of the highest Gal4 concentration was prepared with a final concentration of 50 nM fluorescein labeled peptide and added to a glass test tube. After a 10 minute incubation, the fluorescence polarization of this solution was measured. For each successive measurement, calculated amounts of Storage buffer + 50 nM labeled peptide were added to the test tube, keeping the concentrations of the labeled ligand constant throughout the experiment while the protein concentration was varied. The raw data was analyzed as described above.

Competition experiments using fluorescence polarization

Competition experiments were performed on the Beacon 2000. Prior to each experiment, ~1 mg of purified unlabeled competitor peptide was weighed out and resuspended in DMSO (final concentration <10%) followed by ~1mL of Storage buffer to create a stock solution of unlabeled peptide (~500 μM - 2 mM). Peptide concentrations were determined using a spectrophotometer, λ = 280 or determined solely by weight in cases where their the extinction coefficient could not be determined due to the lack of tryptophans and tyrosines . A control solution with Storage buffer containing the exact amount of DMSO but no unlabeled competitor peptide was also prepared. Next, using the binding isotherm of the labeled peptide and protein, the concentration of protein that gave 50% saturation of labeled peptide was determined. This concentration of protein was kept constant throughout the experiment. For each data point, solutions were prepared that contained the predetermined protein concentration, 25 nM fluorescein labeled peptide, variable amounts of competitor peptide and Storage buffer (with appropriate percentage of DMSO) bringing the solution to a total volume of 100 μL. All solutions were incubated for 30 min at rt before measurements were taken. The raw data was imported into Origin.

ELISA

The ELISAs were carried out according to previously described protocols.[64] The peptide library was synthesized on 0.5 mmole scale as described above. It was split into 4 Alltech tubes (12 mL). Each tube was washed with MeOH (10 x 10 mL) using a vacuum manifold. The resin was then briefly dried under vacuum and allowed to swell in H_2O (5 mL) followed by washes with H_2O (5 x 5 mL). The resin was then blocked with Blocking buffer (10 mM PBS pH 7.0, 0.1% gelatin, 0.2% Tween-20, 0.01% NP-40, 10% glycerol) for 1.5 h. After blocking, the resin was washed with Wash buffer (10 mM PBS, 0.2% tween-20) (2 x 10 mL). Subsequently, 1 μM MBP was added in Blocking buffer with 0.1 mg/mL BSA (4 mL per tube) and incubated for 1 h followed by 5 x 5 mL washes with Wash buffer. To detect bound MBP, Anti-MBP-HRP (1:3000, New England Biolabs) was added in blocking buffer for 1 h. Unbound antibody was removed by washing with Wash buffer (3 x 10 mL), 100 mM PBS (3 x 10 mL) and water (1 x 5 mL). To visualize beads that bound MBP, 4 mL TMB substrate (Sigma) was added to each tube and allowed to incubate for 1 min, after which it was drained and the resin was resuspended into water and spread out into petri dishes to identify blue beads. No blue beads were isolated after 1 min incubation of the substrate. After 2 min incubation of the substrate, 2 blue beads were isolated (the time was optimized based on control ELISA experiments with the positive control ligand VP2 and MBP-Tra1). The remaining resin was transferred back into the Alltech tubes and treated with 6M guanidine for 30 min followed by DMF overnight.

To identify peptides that bind Tra1, the library was washed, and blocked as described above. After blocking, the resin was incubated for 1 h with MBP-Tra1(3092-3524) at a final concentration of 0.25 μM in binding buffer followed by washing (5 x 5 mL) with Wash buffer. The anti-MBP-HRP antibody was added for 1 hr as described above and after thoroughly washing, the TMB substrate was added. TMB was added to each alltech tube and incubated for 5 min after which the resin was resuspended in water and transferred into a petri dish to isolate the blue beads. Several beads (~1%) turned blue on the addition of TMB, however they lost their color when water was added. Only the beads that remained blue after washing with water were transferred into 1.5 mL microfuge tubes using a pipet tip. 500 μL of 6M guanidine was added to each tube containing a single bead and the beads were submitted for sequencing. Edman

degradation sequencing of the selected peptides was carried out at Michigan State University.

Phage display

Phage display was performed according to the Ph.D.-12 Phage Display Peptide Library Kit manual (New England Biolabs). Selections against MBP-Tra1(3092-3524) were done in 96 well protein binding plates (Nunc). (Control experiments were initially performed using anti-MBP antibody to verify that MBP-Tra1 and MBP do indeed bind to the plate and the most stringent Wash buffer used does not strip them off the plate). 200μL 1μM Tra1 in Storage buffer (10mM PBS pH 7.0, 10% glycerol, 0.0% NP-40) was incubated in 2 wells of the plate at 4 °C. Following an overnight incubation, the protein solution was poured off the plate. The plate was subsequently incubated with Blocking Buffer (10mM PBS pH 7.0, 10% glycerol, 0.01% NP-40, 0.1% gelatin, 0.2% tween-20, 10 mg/mL BSA and 0.02% NaN_3) for 2 h at 4 °C with shaking. After blocking, the wells were washed with Wash buffer (6 x 300μL) (10mM PBS, 0.1% Tween-20) and then 10 μL of the phage library in 100 μL of Binding buffer (10 mM PBS pH 7.0, 10% glycerol, 0.01% NP-49, 0.1% gelatin, 0.2% Tween-20, 100 μg/ml BSA) was added to each well. The phage were allowed to bind for 1 h at room temperature with shaking after which, the wells were washed (10 x 300μL) with Wash buffer. The phage were eluted by the addition of 150 mL Elution buffer (0.2M glycine-HCl pH 2.2). After a 10 min incubation, 22.5 μL of Neutralizing buffer (1M Tris pH 9.0) was added. The eluted phage were amplified in ER2728 *E. coli* and precipitated using PEG/NaCl and used for subsequent pannings of selection.

The next round of panning consisting of a negative selection, was carried out to remove phage that bound MBP and/or the plate. For this purpose, 1μM MBP in Storage buffer was added to 2 wells of a plate as described above. 2 x 10^{11} amplified phage from the previous round were added to each well and in this case, after a 1 h incubation, phage that did not bind MBP or the plate were isolated and added to 2 wells that contained Tra1 for the 2nd round of positive selection. The conditions for the 2nd panning were similar as described above, except in one well the Binding buffer used contained 500 μM VP2.

The third round of positive panning was performed similar to the 2nd panning with the two different binding buffers except the stringency of the Wash buffer was increased to contain 0.5% Tween-20.

To identify the peptide sequences encoded by each bound phage, after each panning, 15 individual phage were amplified to extract their DNA, which was sequenced at the University of Michigan Sequencing Core. Phage ELISAs were also performed as per the manual to identify phage that bound MBP-Tra1 over MBP or just the plate. Briefly, 10^9, 10^{10} and 10^{11} phage were incubated with a 96-well plate prebound with 0.1 μM MBP-tra1 or MBP or a blank well. The bound phage were detected using a HRP-conjuagted anti-phage antibody using the ABTS substrate and quantitated at 405 nm using an absorbance plate reader (Tecan Genios Pro).

Competitive ELISA

Phage expressing ENSPLWWPQPLA were amplified according to the New England Biolabs Phage display kit manual. Briefly, 10 μL of phage glycerol stock was used to infect 20 mL of LB (with tetracycline) inoculated with 200 μL of saturated culture of ER2728 and grown for 5 h at 37 °C. The cells were pelleted and the supernatant containing phage were precipitated with PEG/NaCl overnight at 4 °C. Subsequently, the phage was resuspended in PBS and re-precipitated with PEG/NaCl and finally dissolved in 10 mM PBS obtaining ~50 μL of 10^{10} phage/μL.

For the ELISA, 200 μL of 1 μM Tra1 or Med15 was added to each well of a NUNC maxisorp protein-binding 96 well plate and incubated overnight at 4 °C. The plate was blocked with Blocking buffer (10 mg/mL BSA in 10 mM PBS, 0.5% Tween-20) for 1 hr. Subsequently, 200 μL of 10^{10} phage was added to each well in 10 mM PBS, 0.5% Tween-20 with varying amounts of VP2 and incubated for 1 h. After 6 washes with 10 mM PBS, 0.5% Tween-20, 200 μL of anti-M13-hrp antibody (1:5000 dilution) was added in Blocking buffer to each well for 1 h. After 6 washes with 10 mM PBS, 0.5% Tween-20 the phage were detected using 200 μL of ABTS solution in sodium citrate with hydrogen peroxide as per the NEB phage display manual. The substrate was incubated for 1 h and the absorbance at 405 nm measured using an absorbance plate reader (Tecan Genios Pro).

β-galactosidase assays

The function of DBD+TAD fusions was examined in yeast by a quantitative liquid β-galactosidase assay as described in Chapter 2. Briefly, the plasmids encoding the peptide fusions and the DBD plasmid (negative control) were transformed into yeast using the LiOAc method and transformed colonies were selected by growth on synthetic complete (SC) media containing 2% raffinose and lacking the appropriate amino acid(s) for selection. Freshly transformed colonies were used to inoculate 5 ml cultures of SC media containing 2% raffinose and lacking the appropriate amino acids. The cultures were incubated overnight at 30°C with agitation. Following incubation, these cultures were used to inoculate 5 ml cultures of SC media containing 2% raffinose, 2% galactose and lacking the appropriate amino acids that were subsequently incubated overnight at 30°C with agitation to an OD_{660} of 3–4. The yeast cells were harvested and resuspended in Breaking buffer (100 mM Tris-HCl (pH 8.0), 20% glycerol) containing the Complete Protease Inhibitors cocktail (Roche). The cells were lysed by vortexing with glass beads. A portion of the cell extract was used to measure β-galactosidase activity via incubation with *o*-nitrophenyl-β-D-galactopyranoside (1 mg/ml) in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4 \cdot 7H_2O$, and 50 mM 2-mercaptoethanol [pH 7]). The reaction was stopped by adding 1 M Na_2CO_3 and the OD_{420} was measured on a Varian Cary 300 UV-vis spectrometer. The activity reported was normalized to the total protein concentration of the extract, measured using a Bradford assay kit (Bio-Rad) with BSA as the standard.

References

1. Ptashne, M.; Gann, A., *Genes & Signals*. Cold Spring Harbor Laboratory: New York, 2001.
2. Wu, P. Y.; Ruhlmann, C.; Winston, F.; Schultz, P., Molecular architecture of the S. cerevisiae SAGA complex. *Mol Cell* **2004,** 15, (2), 199-208.
3. Mapp, A. K.; Ansari, A. Z., A TAD further: exogenous control of gene activation. *ACS Chem Biol* **2007,** 2, (1), 62-75.
4. Ansari, A. Z.; Koh, S. S.; Zaman, Z.; Bongards, C.; Lehming, N.; Young, R. A.; Ptashne, M., Transcriptional activating regions target a cyclin-dependent kinase. *Proc Natl Acad Sci U S A* **2002,** 99, (23), 14706-9.

5. Ansari, A. Z.; Reece, R. J.; Ptashne, M., A transcriptional activating region with two contrasting modes of protein interaction. *Proc Natl Acad Sci U S A* **1998,** 95, (23), 13543-8.
6. Bouwman, P.; Philipsen, S., Regulation of the activity of Sp1-related transcription factors. *Molecular and Cellular Endocrinology* **2002,** 195, (1-2), 27-38.
7. Chrivia, J. C.; Kwok, R. P.; Lamb, N.; Hagiwara, M.; Montminy, M. R.; Goodman, R. H., Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **1993,** 365, (6449), 855-9.
8. des Etages, S. A.; Falvey, D. A.; Reece, R. J.; Brandriss, M. C., Functional analysis of the PUT3 transcriptional activator of the proline utilization pathway in Saccharomyces cerevisiae. *Genetics* **1996,** 142, (4), 1069-82.
9. Jackson, S. P.; Tjian, R., O-Glycosylation of Eukaryotic Transcription Factors - Implications for Mechanisms of Transcriptional Regulation. *Cell* **1988,** 55, (1), 125-133.
10. Khidekel, N.; Hsieh-Wilson, L. C., A 'molecular switchboard'--covalent modifications to proteins and their impact on transcription. *Org Biomol Chem* **2004,** 2, (1), 1-7.
11. Lamarre-Vincent, N.; Hsieh-Wilson, L. C., Dynamic glycosylation of the transcription factor CREB: A potential role in gene regulation. *Journal of the American Chemical Society* **2003,** 125, (22), 6612-6613.
12. Lamph, W. W.; Dwarki, V. J.; Ofir, R.; Montminy, M.; Verma, I. M., Negative and Positive Regulation by Transcription Factor Camp Response Element-Binding Protein Is Modulated by Phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **1990,** 87, (11), 4320-4324.
13. Sadowski, I.; Niedbala, D.; Wood, K.; Ptashne, M., Gal4 Is Phosphorylated as a Consequence of Transcriptional Activation. *Proceedings of the National Academy of Sciences of the United States of America* **1991,** 88, (23), 10510-10514.
14. Ferdous, A.; Sikder, D.; Gillette, T.; Nalley, K.; Kodadek, T.; Johnston, S. A., The role of the proteasomal ATPases and activator monoubiquitylation in regulating Gal4 binding to promoters. *Genes Dev* **2007,** 21, (1), 112-23.
15. Ptashne, M.; Gann, A., Transcriptional activation by recruitment. *Nature* **1997,** 386, (6625), 569-577.
16. Taatjes, D. J.; Naar, A. M.; Andel, F., 3rd; Nogales, E.; Tjian, R., Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **2002,** 295, (5557), 1058-62.
17. Taatjes, D. J.; Schneider-Poetsch, T.; Tjian, R., Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nat Struct Mol Biol* **2004,** 11, (7), 664-71.
18. Davis, J. A.; Takagi, Y.; Kornberg, R. D.; Asturias, F. A., Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell* **2002,** 10, (2), 409-15.
19. Bryant, G. O.; Ptashne, M., Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol Cell* **2003,** 11, (5), 1301-9.
20. Fishburn, J.; Mohibullah, N.; Hahn, S., Function of a eukaryotic transcription activator during the transcription cycle. *Mol Cell* **2005,** 18, (3), 369-78.
21. Reeves, W. M.; Hahn, S., Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* **2005,** 25, (20), 9092-102.

22. Brown, C. E.; Howe, L.; Sousa, K.; Alley, S. C.; Carrozza, M. J.; Tan, S.; Workman, J. L., Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* **2001,** 292, (5525), 2333-7.
23. Bhaumik, S. R.; Raha, T.; Aiello, D. P.; Green, M. R., In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev* **2004,** 18, (3), 333-43.
24. Green, M. R., Eukaryotic transcription activation: right on target. *Mol Cell* **2005,** 18, (4), 399-402.
25. Archer, C. T.; Burdine, L.; Kodadek, T., Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking. *Mol Biosyst* **2005,** 1, (5-6), 366-72.
26. Chang, C.; Gonzalez, F.; Rothermel, B.; Sun, L.; Johnston, S. A.; Kodadek, T., The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. *J Biol Chem* **2001,** 276, (33), 30956-63.
27. Zhang, F.; Sumibcay, L.; Hinnebusch, A. G.; Swanson, M. J., A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol Cell Biol* **2004,** 24, (15), 6871-86.
28. Swanson, M. J.; Qiu, H.; Sumibcay, L.; Krueger, A.; Kim, S. J.; Natarajan, K.; Yoon, S.; Hinnebusch, A. G., A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol Cell Biol* **2003,** 23, (8), 2800-20.
29. Brooks, C. L.; Gu, W., p53 ubiquitination: Mdm2 and beyond. *Mol Cell* **2006,** 21, (3), 307-15.
30. Lin, J.; Chen, J.; Elenbaas, B.; Levine, A. J., Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. *Genes Dev* **1994,** 8, (10), 1235-46.
31. Oliner, J. D.; Pietenpol, J. A.; Thiagalingam, S.; Gyuris, J.; Kinzler, K. W.; Vogelstein, B., Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature* **1993,** 362, (6423), 857-60.
32. Schon, O.; Friedler, A.; Bycroft, M.; Freund, S. M. V.; Fersht, A. R., Molecular mechanism of the interaction between MDM2 and p53. *J Mol Biol* **2002,** 323, (3), 491-501.
33. Des Etages, S. A.; Saxena, D.; Huang, H. L.; Falvey, D. A.; Barber, D.; Brandriss, M. C., Conformational changes play a role in regulating the activity of the proline utilization pathway-specific regulator in Saccharomyces cerevisiae. *Mol Microbiol* **2001,** 40, (4), 890-9.
34. Sellick, C. A.; Reece, R. J., Modulation of transcription factor function by an amino acid: activation of Put3p by proline. *Embo J* **2003,** 22, (19), 5147-53.
35. Lu, Z.; Rowe, S. P.; Brennan, B. B.; Davis, S. E.; Metzler, R. E.; Nau, J. J.; Majmudar, C. Y.; Mapp, A. K.; Ansari, A. Z., Unraveling the mechanism of a potent transcriptional activator. *J Biol Chem* **2005,** 280, (33), 29689-98.
36. Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P., Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **1996,** 274, (5289), 948-953.
37. Lee, C.; Chang, J. H.; Lee, H. S.; Cho, Y., Structural basis for the recognition of the E2F transactivation domain by the retinoblastoma tumor suppressor. *Genes Dev* **2002,** 16, (24), 3199-212.

38. Lohr, D.; Venkov, P.; Zlatanova, J., Transcriptional regulation in the yeast GAL gene family: a complex genetic network. *FASEB J* **1995,** 9, (9), 777-87.
39. Platt, A.; Reece, R. J., The yeast galactose genetic switch is mediated by the formation of a Gal4p-Gal80p-Gal3p complex. *Embo J* **1998,** 17, (14), 4086-91.
40. Sil, A. K.; Alam, S.; Xin, P.; Ma, L.; Morgan, M.; Lebo, C. M.; Woods, M. P.; Hopper, J. E., The Gal3p-Gal80p-Gal4p transcription switch of yeast: Gal3p destabilizes the Gal80p-Gal4p complex in response to galactose and ATP. *Mol Cell Biol* **1999,** 19, (11), 7828-40.
41. Zenke, F. T.; Engles, R.; Vollenbroich, V.; Meyer, J.; Hollenberg, C. P.; Breunig, K. D., Activation of Gal4p by galactose-dependent interaction of galactokinase and Gal80p. *Science* **1996,** 272, (5268), 1662-5.
42. Sze, J. Y.; Woontner, M.; Jaehning, J. A.; Kohlhaw, G. B., In vitro transcriptional activation by a metabolic intermediate: activation by Leu3 depends on alpha-isopropylmalate. *Science* **1992,** 258, (5085), 1143-5.
43. Wang, D.; Zheng, F.; Holmberg, S.; Kohlhaw, G. B., Yeast transcriptional regulator Leu3p. Self-masking, specificity of masking, and evidence for regulation by the intracellular level of Leu3p. *J Biol Chem* **1999,** 274, (27), 19017-24.
44. Jonker, H. R.; Wechselberger, R. W.; Pinkse, M.; Kaptein, R.; Folkers, G. E., Gradual phosphorylation regulates PC4 coactivator function. *FEBS J* **2006,** 273, (7), 1430-44.
45. Barnes, B. J.; Field, A. E.; Pitha-Rowe, P. M., Virus-induced heterodimer formation between IRF-5 and IRF-7 modulates assembly of the IFNA enhanceosome in vivo and transcriptional activity of IFNA genes. *J Biol Chem* **2003,** 278, (19), 16630-41.
46. Hon, T.; Lee, H. C.; Hach, A.; Johnson, J. L.; Craig, E. A.; Erdjument-Bromage, H.; Tempst, P.; Zhang, L., The Hsp70-Ydj1 molecular chaperone represses the activity of the heme activator protein Hap1 in the absence of heme. *Mol Cell Biol* **2001,** 21, (23), 7923-32.
47. D'Autreaux, B.; Tucker, N.; Spiro, S.; Dixon, R., Characterization of the nitric oxide-reactive transcriptional activator NorR. *Methods Enzymol* **2008,** 437, 235-51.
48. Ferry, A. L.; Locasto, D. M.; Meszaros, L. B.; Bailey, J. C.; Jonsen, M. D.; Brodsky, K.; Hoon, C. J.; Gutierrez-Hartmann, A.; Diamond, S. E., Pit-1beta reduces transcription and CREB-binding protein recruitment in a DNA context-dependent manner. *J Endocrinol* **2005,** 185, (1), 173-85.
49. Sporici, R. A.; Hodskins, J. S.; Locasto, D. M.; Meszaros, L. B.; Ferry, A. L.; Weidner, A. M.; Rinehart, C. A.; Bailey, J. C.; Mains, I. M.; Diamond, S. E., Repression of the prolactin promoter: a functional consequence of the heterodimerization between Pit-1 and Pit-1 beta. *J Mol Endocrinol* **2005,** 35, (2), 317-31.
50. Heyken, W. T.; Repenning, A.; Kumme, J.; Schuller, H. J., Constitutive expression of yeast phospholipid biosynthetic genes by variants of Ino2 activator defective for interaction with Opi1 repressor. *Mol Microbiol* **2005,** 56, (3), 696-707.
51. Yu, Z.; Syu, L. J.; Mellerick, D. M., Contextual interactions determine whether the Drosophila homeodomain protein, Vnd, acts as a repressor or activator. *Nucleic Acids Res* **2005,** 33, (1), 1-12.
52. Nakajima, T.; Fujino, S.; Nakanishi, G.; Kim, Y. S.; Jetten, A. M., TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. *Nucleic Acids Res* **2004,** 32, (14), 4194-204.

53. Wu, J.; Li, Y.; Dietz, J.; Lala, D. S., Repression of p65 transcriptional activation by the glucocorticoid receptor in the absence of receptor-coactivator interactions. *Mol Endocrinol* **2004,** 18, (1), 53-62.
54. Li, G.; Wang, S.; Gelehrter, T. D., Identification of glucocorticoid receptor domains involved in transrepression of transforming growth factor-beta action. *J Biol Chem* **2003,** 278, (43), 41779-88.
55. Jensen, L. T.; Winge, D. R., Identification of a copper-induced intramolecular interaction in the transcription factor Mac1 from Saccharomyces cerevisiae. *EMBO J* **1998,** 17, (18), 5400-8.
56. Zhang, D.; Paley, A. J.; Childs, G., The transcriptional repressor ZFM1 interacts with and modulates the ability of EWS to activate transcription. *J Biol Chem* **1998,** 273, (29), 18086-91.
57. Kopp, J. L.; Wilder, P. J.; Desler, M.; Kinarsky, L.; Rizzino, A., Different domains of the transcription factor ELF3 are required in a promoter-specific manner and multiple domains control its binding to DNA. *J Biol Chem* **2007,** 282, (5), 3027-41.
58. Muratani, M.; Tansey, W. P., How the ubiquitin-proteasome system controls transcription. *Nat Rev Mol Cell Biol* **2003,** 4, (3), 192-201.
59. Lu, X. Y.; Ansari, A. Z.; Ptashne, M., An artificial transcriptional activating region with unusual properties. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (5), 1988-1992.
60. Lu, Z.; Ansari, A. Z.; Lu, X. Y.; Ogirala, A.; Ptashne, M., A target essential for the activity of a nonacidic yeast transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (13), 8591-8596.
61. Carey, M.; Lin, Y. S.; Green, M. R.; Ptashne, M., A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature* **1990,** 345, (6273), 361-4.
62. Radhakrishnan, I.; Perez-Alvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E., Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* **1997,** 91, (6), 741-52.
63. Zor, T.; De Guzman, R. N.; Dyson, H. J.; Wright, P. E., Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *Journal of Molecular Biology* **2004,** 337, (3), 521-534.
64. Wu, Z.; Belanger, G.; Brennan, B. B.; Lum, J. K.; Minter, A. R.; Rowe, S. P.; Plachetka, A.; Majmudar, C. Y.; Mapp, A. K., Targeting the transcriptional machinery with unique artificial transcriptional activators. *J Am Chem Soc* **2003,** 125, (41), 12390-1.
65. Ma, B.; Pan, Y.; Gunasekaran, K.; Keskin, O.; Venkataraghavan, R. B.; Levine, A. J.; Nussinov, R., The contribution of the Trp/Met/Phe residues to physical interactions of p53 with cellular proteins. *Phys Biol* **2005,** 2, (1-2), S56-66.
66. Jackson, B. M.; Drysdale, C. M.; Natarajan, K.; Hinnebusch, A. G., Identification of seven hydrophobic clusters in GCN4 making redundant contributions to transcriptional activation. *Mol Cell Biol* **1996,** 16, (10), 5557-71.
67. Sainz, M. B.; Goff, S. A.; Chandler, V. L., Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids important for activation in vivo. *Molecular and Cellular Biology* **1997,** 17, (1), 115-122.
68. Han, Y.; Kodadek, T., Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. *J Biol Chem* **2000,** 275, (20), 14979-84.

69. Rutledge, S. E.; Volkman, H. M.; Schepartz, A., Molecular recognition of protein surfaces: High affinity ligands for the CBPKIX domain. *J Am Chem Soc* **2003,** 125, (47), 14336-14347.
70. Frangioni, J. V.; LaRiccia, L. M.; Cantley, L. C.; Montminy, M. R., Minimal activators that bind to the KIX domain of p300/CBP identified by phage display screening. *Nature Biotechnology* **2000,** 18, (10), 1080-1085.
71. Tanaka, M., Modulation of promoter occupancy by cooperative DNA binding and activation-domain function is a major determinant of transcriptional regulation by activators in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (9), 4311-4315.
72. Majmudar, C. Y.; Lum, J. K.; Prasov, L.; Mapp, A. K., Functional specificity of artificial transcriptional activators. *Chem Biol* **2005,** 12, (3), 313-21.
73. Tanaka, M.; Herr, W., Reconstitution of transcriptional activation domains by reiteration of short peptide segments reveals the modular organization of a glutamine-rich activation domain. *Molecular and cellular biology* **1994,** 14, (9), 6056-6067.
74. Vashee, S.; Willie, J.; Kodadek, T., Synergistic activation of transcription by physiologically unrelated transcription factors through cooperative DNA-binding. *Biochem Biophys Res Commun* **1998,** 247, (2), 530-5.
75. Lin, Y. S.; Carey, M. F.; Ptashne, M.; Green, M. R., Gal4 Derivatives Function Alone and Synergistically with Mammalian Activators Invitro. *Cell* **1988,** 54, (5), 659-664.
76. Oliviero, S.; Struhl, K., Synergistic Transcriptional Enhancement Does Not Depend on the Number of Acidic Activation Domains Bound to the Promoter. *Proc Natl Acad Sci USA* **1991,** 88, (1), 224-228.
77. Herschlag, D.; Johnson, F. B., Synergism in Transcriptional Activation - a Kinetic View. *Genes & Development* **1993,** 7, (2), 173-179.
78. Mayer, M. P.; Bukau, B., Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci* **2005,** 62, (6), 670-84.
79. Horwich, A. L.; Weber-Ban, E. U.; Finley, D., Chaperone rings in protein folding and degradation. *Proc Natl Acad Sci U S A* **1999,** 96, (20), 11033-40.
80. Tossi, A.; Sandri, L., Molecular diversity in gene-encoded, cationic antimicrobial polypeptides. *Curr Pharm Des* **2002,** 8, (9), 743-61.
81. Marmorstein, R.; Carey, M.; Ptashne, M.; Harrison, S. C., DNA Recognition by Gal4 - Structure of a Protein DNA Complex. *Nature* **1992,** 356, (6368), 408-414.
82. Hidalgo, P.; Ansari, A. Z.; Schmidt, P.; Hare, B.; Simkovich, N.; Farrell, S.; Shin, E. J.; Ptashne, M.; Wagner, G., Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes & Development* **2001,** 15, (8), 1007-1020.
83. Himmelfarb, H. J.; Pearlberg, J.; Last, D. H.; Ptashne, M., GAL11P: a yeast mutation that potentiates the effect of weak GAL4-derived activators. *Cell* **1990,** 63, (6), 1299-309.
84. Chakshusmathi, G.; Mondal, K.; Lakshmi, G. S.; Singh, G.; Roy, A.; Ch, R. B.; Madhusudhanan, S.; Varadarajan, R., Design of temperature-sensitive mutants solely from amino acid sequence. *Proc Natl Acad Sci U S A* **2004,** 101, (21), 7925-30.
85. Kodadek, T.; Sikder, D.; Nalley, K., Keeping transcriptional activators under control. *Cell* **2006,** 127, (2), 261-4.

86. Giniger, E.; Ptashne, M., Transcription in Yeast Activated by a Putative Amphipathic Alpha-Helix Linked to a DNA-Binding Unit. *Nature* **1987,** 330, (6149), 670-672.
87. Liu, B.; Han, Y.; Ferdous, A.; Corey, D. R.; Kodadek, T., Transcription activation by a PNA-peptide chimera in a mammalian cell extract. *Chemistry & Biology* **2003,** 10, (10), 909-916.
88. Drysdale, C. M.; Duenas, E.; Jackson, B. M.; Reusser, U.; Braus, G. H.; Hinnebusch, A. G., The Transcriptional Activator Gcn4 Contains Multiple Activation Domains That Are Critically Dependent on Hydrophobic Amino-Acids. *Molecular and Cellular Biology* **1995,** 15, (3), 1220-1233.
89. Sullivan, S. M.; Horn, P. J.; Olson, V. A.; Koop, A. H.; Niu, W.; Ebright, R. H.; Triezenberg, S. J., Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Research* **1998,** 26, (19), 4487-4496.
90. Lum, J. K.; Majmudar, C. Y.; Ansari, A. Z.; Mapp, A. K., Converting inactive peptides into potent transcriptional activators. *ACS Chem Biol* **2006,** 1, (10), 639-43.
91. Kwon, Y.; Arndt, H. D.; Mao, Q.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J Am Chem Soc* **2004,** 126, (49), 15940-1.
92. Alluri, P.; Liu, B.; Yu, P.; Xiao, X.; Kodadek, T., Isolation and characterization of coactivator-binding peptoids from a combinatorial library. *Mol Biosyst* **2006,** 2, (11), 568-79.

CHAPTER 4
CHARACTERIZING ACTIVATOR-COACTIVATOR INTERACTIONS*

A. Abstract

Because of the central role that transcriptional activators play in gene regulation, significant effort has been devoted to developing a molecular-level picture of how they function in eukaryotes. It is generally accepted that activators are involved in two major types of binding events: DNA binding events and protein-protein interactions.[3] Experimental techniques such as ChIP-on-chip and in vivo footprinting have successfully been used to identify the DNA binding sites of activators within the genome, and in many cases these have been correlated with gene expression profiles.[4-6] However, protein binding partners of activators, and more specifically, the transcriptional machinery binding partners (coactivators) have not conclusively been identified.[7] For example, a number of genetic and biochemical studies have identified coactivator proteins that interact with activators directly or indirectly.[8-13] Two recent crosslinking studies in reconstituted (cell-free) systems have implicated a small subset of these proteins as likely direct activator targets.[2, 14] Although an important step forward, there are still many questions about the exact nature of activator-coactivator interactions and the relevance of these results in the context of the cellular environment. More specifically, given the importance of activator binding site on function (Chapter 2), not only is the identification of the in vivo activator targets important, but also determining the binding site(s) within a given coactivators.[1, 2, 15, 16] In the long term, efforts to determine the identity of activator binding partners and the binding sites will establish screening targets for artificial activator and inhibitor design for therapeutic and biotechnology applications.

---

* I performed experiments in Figures 4.4 – 4.7, 4.12 – 4.14. Mass spectrometry data for Figures 4.5 and 4.8 was obtained by Bo wang (Hakånsson group) and the Michigan Proteome Consortium facility. Jenifer Lum performed experiments in Figure 4.10. Experiments in Figure 4.15 were performed in collaboration with Anna Mapp.

In this Chapter, we use crosslinking coupled with mass spectrometry for the characterization of activator binding sites and, ultimately, for the identification of binding partners. In vitro crosslinking was used to identify the binding sites of the activators Gal4, Gcn4 and VP2 within a key coactivator Med15(Gal11) and coupled with genetic experiments, these studies suggest that the TADs function through distinct but overlapping binding profiles. To characterize these interactions in vivo, in parallel we developed a strategy for reliable incorporation of the photo-crosslinkable amino acid, p-benzoylphenylalanine (pBpa), into Gal4 in yeast using nonsense suppression. Finally, using crosslinker-containing Gal4, we have performed cellular crosslinking experiments, thus setting the stage to identify its functionally relevant targets in vivo.

B. Introduction

The protein-protein interactions between the amphipathic class of transcriptional activators and their interacting partners have been extensively studied. Much of what is known about these interactions comes from studies of the amphipathic activators Gal4, Gcn4 and VP16.[1, 8, 10, 12, 17-23] Early on, chromatin immunoprecipitation (ChIP) studies showed that subsequent to activator-promoter binding, the chromatin-remodeling complex SAGA, the Mediator complex and the general transcriptional machinery are sequentially localized to the promoter.[24] However, this study did not reveal if the recruitment process was initiated by direct binding interactions of activators with components of these complexes. In vitro crosslinking experiments with DNA bound Gcn4 and Gal4 subsequently revealed that both of these activators interact with Tra1 and Taf12 in the SAGA complex and Med15(Gal11) in Mediator.[2, 14] However, whether or not these TADs target identical binding sites within these proteins was not elucidated. In addition, there was some question regarding the in vivo relevance of these results since the studies used non-native promoters and the DNA templates used were not chromatinized. There is, however additional data validating Tra1 as a target, in vivo FRET demonstrated an interaction between the Gal4 TAD and Tra1.[25-27] This study examined eighteen other potential targets, including Med15(Gal11), and only Tra1 and the masking protein Gal80 gave a positive FRET signal.[27] Although the negative results here are difficult to interpret given the reliance of FRET on the relative positioning of the

partner, they underscore the need for alternative approaches to define both the in vivo binding partners of activators and the binding sites with these partners. More recently, the Lehming group also used a similar approach where they attached half of the ubiquitin protein to Gal4, Gcn4 and probed for complementation with the other half of ubiquitin attached to different transcriptional machinery proteins in yeast.[28] This study resulted in the identification of >10 proteins, including Med15(Gal11), that associate with Gal4 and it raises the possibility that some of these proteins are not direct targets of Gal4.[29] Thus as outlined in this Chapter, we have focused on implementation of alternative in vitro and in vivo crosslinking strategies to address this unmet need.

Strategy	Description	Advantages	Disadvantages
Structural studies	NMR spectroscopic and Crystallographic studies of free activators or activators in complex with target proteins	Provides important information about the structure of activators and the residues of activators important for target binding. Can also provide information about the target binding site	In vitro; free activators are often unstructured in solution and it is difficult to obtain structural information in these cases; high protein concentrations required and activators and their targets are often insoluble/aggregate at these concentrations
Pull down experiments	Activators or their targets are immobilized on a solid support often by tagging one of the partners with GST. The immobilized partner is incubated with extract containing a potential bait in solution and subsequently the solid support is probed for retainment of the bait proteins	Convenient to perform experimentally; several examples of this techniques used to identify potential activator binding partners	In vitro; often results in false positives since the experiments are performed at non-physiological concentrations of the interacting pair; GST dimerizes in solution and activator fusion to GST can create an artificially high affinity for a target due to cooperative binding
Genetic experiments	Endogenous copies of potential coactivators are mutated or deleted and the effect of this perturbation is probed on the ability of an activator to function	In vivo, can provide functional information about the relevance for the potential coactivator; low resolution information on binding sites by determining regions on coactivators that are important for activator function	Often hard to interpret since they result in pleiotropic effects due to large-scale alteration of transcription protein complexes instead of disrupting specific individual activator-coactivator interactions
In vitro crosslinking	Incorporation of relatively small crosslinkers into a TAD results in a covalent bond between the crosslinker and the activator target when triggered by irradiation with light in the case of photo-crosslinkers or by the addition of other triggering molecules	Identifies direct activator targets; creates a small perturbation in activator structure; can probe the role of different residues on an activator for target binding; can identify activator binding sites on targets	In vitro, no direct physical evidence for these activator-target interactions *in vivo,* does not take in consideration the effects of chromatin, and covalent modifications on activators such as ubiquitylation, and ubiquitin-mediated proteolysis.
ChIP and formaldehyde-based *in vivo* crosslinking	Use of formaldehyde as a crosslinking agent in live cells enables the rapid generation of protein-protein and protein-DNA crosslinks, creating a snapshot of cellular interactions at a given time point. Immunoprecipitation with an antibody for a particular target can indicate the association between the target and a DNA bound activator.	In vivo, permits the ability to probe activator-target interactions in a time dependent manner; Using qPCR can quantitate the association between two interacting partners; formaldehyde freezes all cellular interaction preventing degradation of activator-target complexes by proteases and the proteasome that often results during cellular lysis	Formaldehyde is a non-specific crosslinker that crosslinks two lysine residues on proteins that are in close proximity, an apparent protein-protein interaction in vivo could either represent a direct crosslink between the two proteins or it might be due to multiple crosslinks that indirectly connect the two proteins; the amount of formaldehyde crosslinking depends on the number and physical location of lysines within the interacting surfaces, parameters that vary among protein-protein interactions. For this reason, the failure to observe a crosslink between two proteins does not necessarily mean that the proteins are not in contact
FRET	Two potential interacting partners are fused to different fluorophores that when in close contact display a FRET signal indicating an interaction between the two proteins	In vitro and in vivo, validates interactions between an activator and targets in cells.	All fluorophore fusions are not optimal to observe a FRET signal, there is a maximum separation distance that the fluorophores can be apart, the signal also depends on the orientation of the fluorophores; cannot be used to discover unknown interactions, require a priori knowledge of potential activator targets, so they can be fused to fluorophores to probe a potential interaction

Table 4.1 Strategies to characterize activator-target interactions

B.1 Unnatural amino acids as protein interaction probes in cells

Until recently, the state of the art in identifying protein-protein interactions and binding sites involved the use of in vitro approaches and the few cellular strategies that have been developed also have limitations (summarized in Table 4.1). Methods to incorporate a site-specific crosslinker in a protein in living cells would be extremely advantageous as any crosslinked protein discovered could be confidently assumed to be a direct binding partner.[30-32] Historically the challenge to using this method has been technical; the selective incorporation of a photoactive group at a single site within a single protein in cells has been until recently an elusive goal.

Current methods of incorporating crosslinking agents in cells involve the use of photo-Met and photo-Leu amino acids which substitute the endogenous Leu and Met in proteins.[33-35] This strategy has been successfully used in mammalian cells to identify protein-protein interactions. We tried photo-Met and photo-Leu in yeast and found that yeast do not grow under these conditions, either due to problems incorporating them or poor transport cannot incorporate them and hence do not grow. Another approach involves the in vitro labeling of proteins with a crosslinker and assisting the cellular entry of the proteins by fusion to a cell-penetrating peptide such as TAT.[36-40] However, this approach has seen only modest success in yeast.[41-44] Moreover, the expression of well-folded activators in bacteria is a longstanding challenge as these proteins typically aggregate quite extensively.

B.1.a Nonsense suppression to incorporate site-specific unnatural amino acids into proteins

In a recent breakthrough, nonsense suppression methods have made it possible to incorporate unnatural amino acids conferring advantageous chemical functionality into proteins in cells. Using this strategy, the Schultz group has developed methods to incorporate unnatural amino acids in bacteria, yeast and mammalian cells at the amber stop codon.[45-54] In order to incorporate unnatural amino acids into proteins in yeast, they used the orthogonal pair of the *E. coli* amber suppressor tyrosyl–tRNA synthetase and $tRNA_{CUA}$. The *E. coli* tyrosyl–tRNA synthetase (TyrRS) has previously been shown to efficiently aminoacylate *E. coli* $tRNA_{CUA}$ when both are genetically encoded in *S. cerevisiae* (Figure 4.1c).[55-58] To alter the amino acid specificity of the orthogonal TyrRS

so that it aminoacylates $tRNA_{CUA}$ with a desired unnatural amino acid and none of the endogenous amino acids, they used a growth-based selection strategy in yeast with a library of TyrRS containing 5 random mutants in the active site. The selection was based on the ability of TyrRS to incorporate the unnatural amino acid into 2 amber positions producing an active Gal4, which enabled the yeast to survive. They also performed a number of controls to verify that the isolated TyrRS mutants were in fact functioning by incorporating a particular amino acid. Finally the unnatural amino acids of interest were inserted into the control protein superoxide dismutase (hSOD) and incorporation was verified by mass spectrometry.[52, 53]

Using the synthetase-tRNA pairs developed by this approach, the Schultz and Wang groups have shown the incorporation of a variety of amino acids (>30) including the fluorescent amino acid dansyl alanine and amino acids that contain groups for photocrosslinking, p-Benzoyl phenyl alanine (pBpa) and p-Azido phenyl alanine (pAzpa) (Figure 4.1a).[50, 59-63] Incorporation of pBpa and pAzpa into transcriptional activators in cells could lead to new methods for probing activator-coactivator interactions in a cellular context and to determine the binding sites of activators on coactivators using mass spectrometry.[64-67]

Specifically, pBpa is well suited for this purpose because it is more stable than pAzpa and can be iradiated at 365 nm wavelengths that are less damaging to proteins and other biomolecules. Further, the diradicals formed when pBpa is excited to its triplet state reversibly relax back down to the ground state if there are no H-bond donors in close proximity.[68-70] This feature enables the continuous irradiation of this molecule to increase crosslinking yields without compromising on the reactivity of the crosslinker (Figure 4.1b).

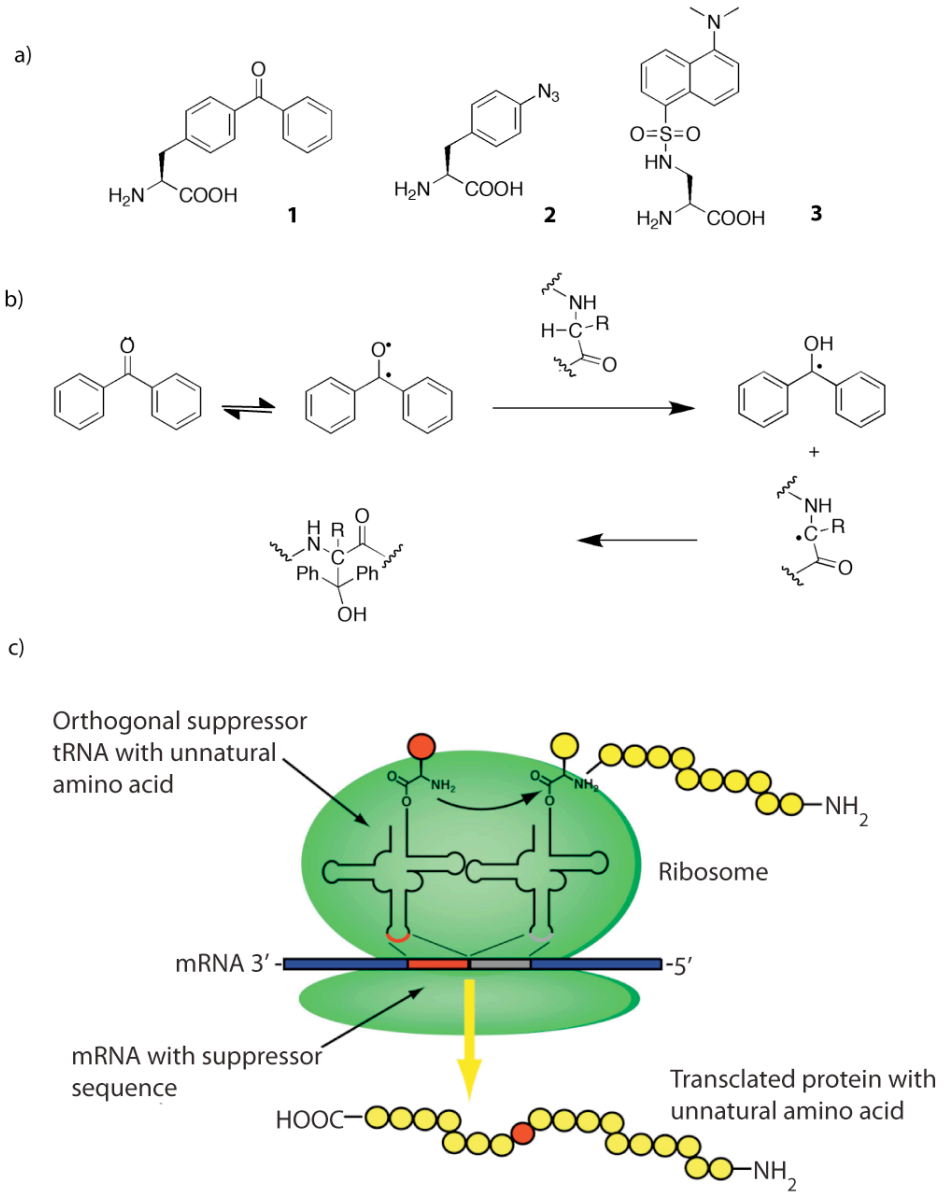


Figure 4.1 Structure of unnatural amino acids. a) p-Benzoyl phenyl alanine (pBpa), p-Azido phenyl alanine (pAzpa) and dansyl alanine. pBpa and pAzpa contain functional groups that enable the formation of protein-protein crosslinks when irradiated with UV light. Dansyl alanine is a fluorescent amino acid and can be used for FRET studies. b) Mechanism of benzophenone crosslinking. Upon exposure to UV light (365 nm), benzophenone forms a diradical that abstracts a proton from adjacent backbone C-H bonds of proteins in close proximity followed by an insertion event to form a covalent C-C linkage between itself and the protein. c) Incorporation of unnatural amino acids in cells. Orthogonal suppressor tRNA is charged with an unnatural amino acid by its cognate synthetase and then it recognizes the suppressor codon on the mRNA.

C. Results

C.1 Determination of binding site(s) of activators within Med15(Gal11)

As outlined in Chapter 2, many lines of evidence suggest that the yeast Mediator component Med15(Gal11) is important for activated transcription, specifically as a binding target of endogenous and artificial TADs.[2, 10, 13, 14, 71-74] For example, ligands that were screened to bind Med15(Gal11)(186-619) were able to specifically activate transcription in yeast (Chapter 2 and 3). It has been shown that the C-terminus of the protein is involved in transcriptional machinery interactions, with TFIIE.[75-77] The N-terminus and the middle portion of the protein have been shown to be important for activator contacts (residues 1-619).[10, 17, 18, 71, 78] However, it is not known if there is a single, generally utilized TAD binding surface in this protein or a variety of sites (as suggested by experiments in Chapter 2) used by endogenous activators.

As discussed in the final section of this Chapter, in vivo crosslinking will provide the most convincing evidence for physiologically relevant activator targets and binding sites on these targets. As we were adapting that technologically challenging strategy for our purposes, we simultaneously pursued in vitro crosslinking that would enable us to identify likely TAD binding sites to search for in vivo.

To identify the binding site, we utilized mass spectrometry coupled with photo-crosslinking methods resulting in a powerful strategy to determine activator-coactivator binding sites.[79, 80] Using this approach, the identity and location of site-specific photo-crosslinks generated can be identified by mass spectrometry, providing high resolution binding site information. More importantly, with the improvements in mass spectrometer sensitivity they require only small amounts of protein compared to structural studies that can also provide similar information but are less feasible due to the stability and solubility of several transcription proteins at high concentration.[81-88]

We decided to identify the binding sites of the Gal4, Gcn4 and VP16 TADs within Med15(Gal11). These activators upregulate transcription to high levels in cells and several lines of evidence suggest that Med15(Gal11) is a target of these activators. [2, 10, 13, 14, 71-74] Given the overall amphipathic nature of these TADs, we wanted to investigate if these TADs target the same binding site(s) on Med15(Gal11). For our initial studies, we envisioned that we would use isolated TADs from these activators and perform

crosslinking experiments with different fragments of Med15(Gal11) in vitro. Using western blots we could then narrow in on the regions that seemed to interact with these TADs. Once we identified the fragments of Med15(Gal11) that crosslink with TADs, we would repeat similar crosslinking experiments with those fragments and submit them for the more challenging mass spectrometry analysis for exact binding site(s) determination (Figure 4.2). This would provide, for the first time high resolution information about the location of the binding site. Subsequently, to investigate the physiological relevance of these results we could measure the activity of the various activators in Med15(Gal11) delete strains, since deletion of Med15(Gal11) does not result in large-scale reorganization of the transcriptional machinery.[89] Future experiments with TADs fused to DBDs would probe the binding profile in the context of a DNA bound activator and finally crosslinking experiments in yeast with crosslinkers incorporated into activators would enable the identification of binding sites in a native context.

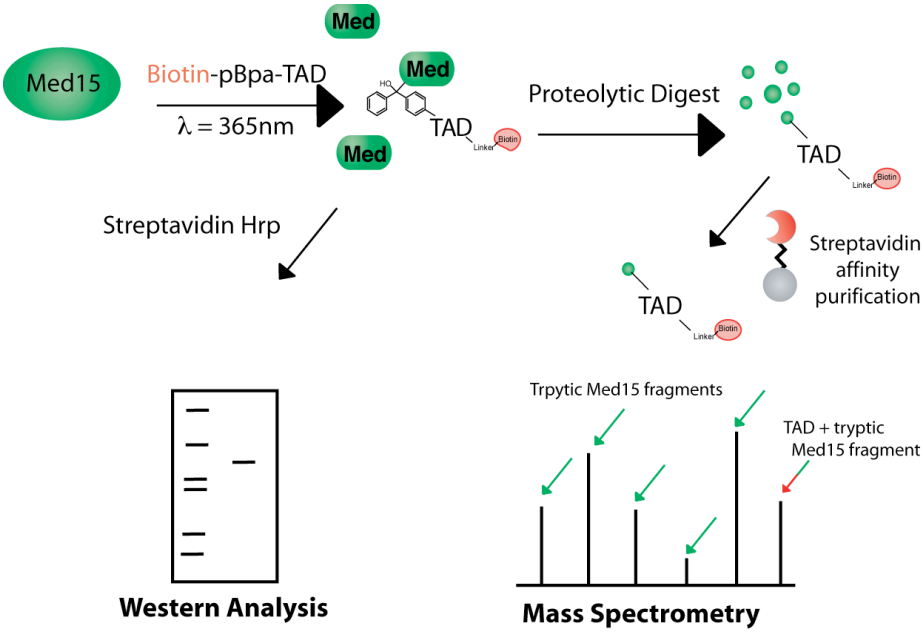


Figure 4.2 Strategy to identify binding sites of activators using in vitro crosslinking. Irradiation of Med15(Gal11) fragments combined with TADs containing pBpa and biotin will form TAD-Med15 crosslinks. Western blots using a streptavidin-HRP conjugate can be used to qualitatively probe the fragments of Med15(Gal11) that are crosslinked. To identify binding sites, the crosslinked reaction is proteolytically digested (e.g. trypsin) and probed by mass spectrometry directly or it can be enriched specifically for crosslinked fragments (red+green arrow) over tryptic Med15 fragments (green arrows) before being probed by mass spectrometry.

C.1.a Construction of TADs and expression of Med15(Gal11) fragments

Solid phase peptide synthesis was used to synthesize TADs from Gal4, Gcn4, and VP16 in which pBpa replaced particular phe residues. The sequences for each endogenous activator chosen have previously been shown to display a significant portion of the activity of the full-length proteins in yeast. Further, the importance of the particular Phe residue in interactions with targets and function has been well documented.[1, 90-92] Moreover, substitution of the particular Phe with other hydrophobic residues has shown to retain function and affinity for target proteins.[1, 2, 14] Thus, the photactivatable amino acid p-Benzoyl phenyl alanine (pBpa) was used to replace positions that contained existing Phe residues, with the attempt of maintaining the hydrophobic character of the peptide (Figure 4.3a). The peptides were also labeled with a biotin handle at the N or C terminus for purification and qualitative detection of crosslinking by western blots.

Since full-length Med15(Gal11) cannot be bacterially expressed well and with high enough purity due to its large size, in collaboration with the high throughput protein expression facility (LSI), we developed bacterial expression constructs that span the entire sequence of the protein. These constructs were designed on the basis of secondary structure predictions and produced Med15(Gal11) as fusions to the solubility and purification tags his-MBP and his-mOCR. Using this approach we were able to obtain constructs that span >90% the sequence of Med15(Gal11) (Figure 4.3b).

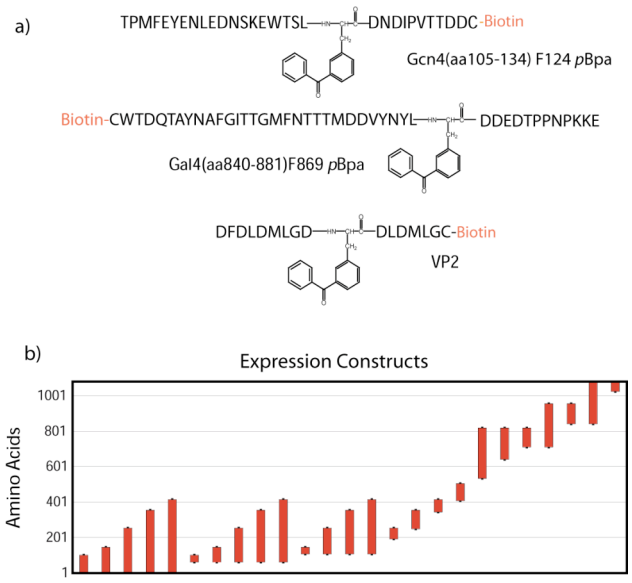


Figure 4.3 Constructs for in vitro crosslinking. a) Structure of TADs used for in vitro crosslinking. pBpa is incorporated into the TADs in place of phe residues maintaining the hydrophobic character. A biotin handle is incorporated into each TAD to facilitate detection and purification. b) Med15 *E. coli* constructs. Bacterial expression constructs spanning the entire sequence of Med15 fused to MBP or mOCR were developed at the high throughput protein expression facility at LSI. Expression constructs that gave soluble protein >75% purity after a single affinity chromatography step are shown.

C.1.b Locating the TAD binding region

To localize the fragment of Med15(Gal11) that interacted with activators, we picked a series of fragments that spanned the entire sequence of Med15(Gal11) and performed crosslinking reactions with isolated TADs. For each crosslinking reaction, the pBpa containing-TAD was incubated with Med15(Gal11) and irradiated with 365 nm light for 30 min using a 6W handheld lamp at 4 °C. The presence of the crosslinked product was detected by western blots using streptavidin-horse raddish peroxidase (HRP) following resolution by PAGE. Satisfyingly, based on previous studies highlighting the importance of the N-terminus and the middle of the protein for activator contacts, we found that all the TADs interacted with Med15(Gal11)(1-357), and a smaller fragment within that region, Med15(Gal11)(107-255), also showed a strong signal when probed with streptavidin-hrp in a western blot (Figure 4.4).[10, 17, 18, 71] This suggested that at least one binding site was contained within this region. Consistent with these results, fluorescence polarization binding experiments with this region revealed that the Gal4, Gcn4 and VP2 TADs interacted with a low micromolar dissociation constant (Figure 4.4c).

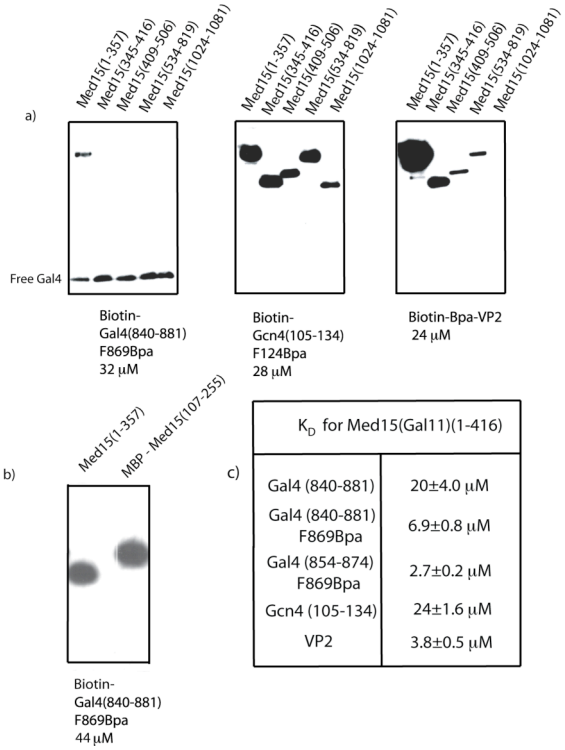


Figure 4.4 Isolation of TAD binding region of Med15(Gal11).
a) 1 mg/mL of MBP fusions spanning the sequence of Med15 (obtained from high throughput protein lab, LSI) were incubated with 24-32 μM TADs of each TAD individually and irradiated for 30 min. The detection of crosslinked products was performed using westerns blot probed with streptavidin-hrp. b) 20 μM Med15(1-357) with no tag or MBP-Med15(107-255) were incubated with 44 μM Gal4 TAD and irradiated for 30 min. c) Dissociation constants for TADs and Med15(1-416) determined by fluorescence polarization. Error is shown as SD from 3 independent experiments.

C.1.c Mass spectrometric determination of binding sites

To obtain binding site information we decided to use Med15(Gal11)(1-416) as Med15(Gal11)(345-416) also interacted with Gcn4 and VP2 (Figure 4.4b). The mass spectrometry studies were performed by our collaborator Bo Wang in the Håkansson group. The determination of the binding site is achieved by performing a proteolytic digest of the TAD-Med15(Gal11) crosslinked complex and detection of a set of peaks that correspond to molecular weights of the TAD + a proteolytic Med15(Gal11) fragment. To eliminate false positives, a candidate crosslinked product was also confirmed using MS-MS to verify it contained residues of the TAD and of Med15(Gal11).

C.1.c.i Determination of sequence coverage for Med15(Gal11)(1-357)

Since the TAD could potentially crosslink at any region within Med15(Gal11), it was necessary to ensure that >80% sequence coverage of Med15(Gal11) itself could be obtained in terms of proteolytic peptides. Using FT-ICR mass spectrometry and optimizing various instrumental parameters, our collaborator Bo Wang was able to obtain >95% sequence coverage of the protein by using trypsin for proteolytic digestions (Figure 4.5). Similar experiments with chymotrypsin showed that it provided ~80% sequence coverage.

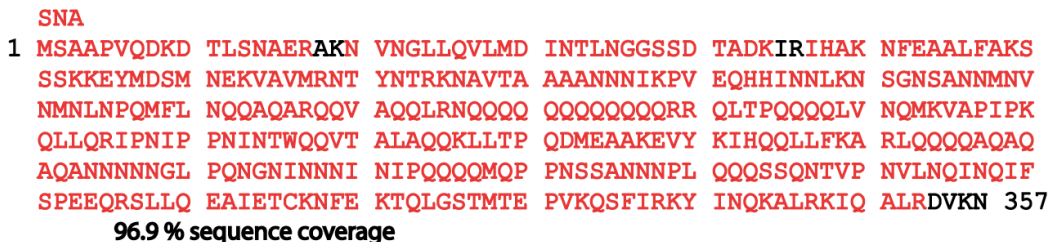


Figure 4.5 Sequence coverage of Med15(1-357) using mass spectrometry. Med15(1-357) was digested with trypsin and the proteolytic fragments were analyzed by ESI mass spectrometry. The residues that were contained in the proteolytic fragments are indicated in red. The amino acids SNA are not part of Med15(Gal11), but are left over after cleavage with TEV protease.

C.1.c.ii TAD-Med15(Gal11) crosslinking

For each crosslinking reaction, the pBpa-containing TAD was incubated with Med15(Gal11) and irradiated with 365 nm light using a 6W handheld lamp at 4 °C. Evaluation of the time dependence of UV irradiation for the Gcn4 TAD revealed that

only 5 min of 365 nm light was sufficient to produce a crosslinked product detectable via Western blot analysis. At longer times (>2 h) we noticed that the TAD crosslinked to the minor impurities present in the Med15(Gal11) sample. A 2 h time point seemed to balance TAD-Med15(Gal11) crosslinked product yields with lower levels of crosslinking to the impurity (Figure 4.6). These were the conditions used for subsequent experiments. Initially, the reaction mixture containing uncrosslinked TAD, free Med15(Gal11) and crosslinked TAD-Med15(Gal11) was subjected to trypsin digest followed by mass spectrometric analysis directly without separating the individual components. Using this approach, however, only peaks corresponding to the free TAD and tryptic Med15(Gal11) fragments were observed. Thus the crosslinked products were likely of lower abundance relative to the other components and/or were less ionizable by the ion source perhaps due to ion suppression.

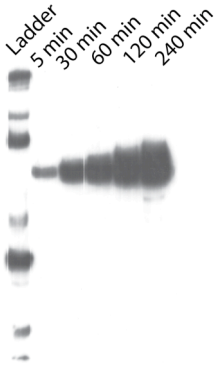


Figure 4.6 Time dependence of crosslinking. 45 μM Med15(1-416) was incubated with 70 μM Gcn4 TAD for the times indicated. The crosslinked products were visualized using a western blot with streptavidin-HRP.

C.1.c.iii Enriching for crosslinked products

In order to specifically isolate the crosslinked products for mass spectrometric analysis, we incubated the tryptic digest reaction mixture with neutravidin beads. Theoretically, only the biotinylated compounds would bind to neutravidin, which would include the free TAD and any potential crosslinked products (the TADs were labeled at the N or C terminus depending on which position prevents the loss of biotin and pBpa after trypsin digest). After several washes of the neutravidin beads, the bound material was eluted using various elution conditions. It was found that the presence of detergent was required to elute the biotinylated material from the beads; heating the beads to 100 °C, addition of organic solvents such as acetonitrile, DMF and DMSO and lowering the pH did not elute detectable amounts of material. Detergents are not compatible with mass spectrometry since they form clusters and prevent the ionization of the analyte. In an

effort to remove the detergent we moved to using LC/MS using a C18 column, which would resolve the tryptic peptides from the detergent. However, we had limited success with this approach, with the detergent clusters still being observed and no crosslinked fragment was found.

C.1.c.iv Use of cleavable disulfide linked biotin-TAD conjugates

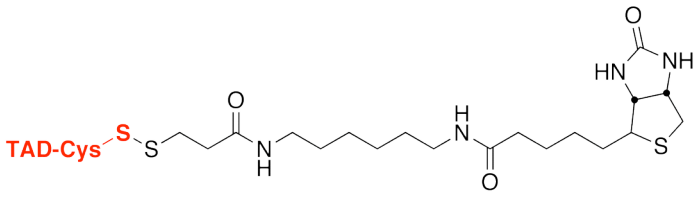


Figure 4.7 Structure of TAD linked to cleavable biotin. TADs containing pBpa and a Cys (in red) were linked to a disulfide cleavable biotin moiety.

In an effort to elute the biotinylated material from neutravidin using mass spectrometry compatible reagents, Gal4, Gcn4, and VP2 were linked to biotin via a disulfide bond, which on the addition of DTT results in the cleavage of biotin from the rest of the sequence, enabling its elution from neutravidin beads (Figure 4.7). Disulfide biotin-linked Gal4, Gcn4 or VP2 were combined with Med15(Gal11)(1-416) for 2 h and irradiated with a 150W 365 nm light source at 4 °C. At the end of this time, a 30K MW cut-off concentrator was used to remove any free TAD from the solution and the semi-purified mixture was digested with trypsin. After the trypsin was heat inactivated, the material was purified using neutravidin agarose and eluted with 100 mM DTT at 65 °C and subsequently analyzed by mass spectrometry.

C.1.c.v Location of binding sites

The TADs were observed to crosslink to Med15(Gal11) (207-217) and to Med15(Gal11) (161-174) by MS and the result was confirmed by obtaining MS-MS spectra. This was not unexpected since in Figure 4.4, Med15(Gal11)(107-255) also crosslinks with these TADs. Interestingly, recent work by Kim et. al identified homology between residues 116-255 of Med15(Gal11) and the Qr of human steroid receptor coactivator (SRC) proteins.[93] Specifically, they align two characteristic nuclear receptor binding motifs, an A and B box, to lie within this region. They also found that the mammalian glucocorticoid receptor (GR) TAD (known as tau1, GR(77-262)) was able to interact with the 116-255 fragment. In particular, they found that mutations in the B box (residues 196-202) abolished Med15(Gal11) binding and the cellular activity of the GR

TAD in yeast. Given the close proximity of the potential B box to the Med15(Gal11)(207-217) crosslinked fragment, it is likely that this region has evolved to interact with multiple activators (Figure 4.8).

Crosslinking experiments with Gal4, Gcn4, VP2 and Med15(Gal11)(1-416) also resulted in another crosslinked fragment, Med15(Gal11)(74-84). This binding site lies in helix 3 of the Med15(Gal11) GACKIX domain, an evolutionary conserved activator binding motif.[78, 94-98] The GACKIX domain has been found in mammalian coactivators CBP, ARC105 and SREBP. Comparison of the different activators binding sites across species reveals an overlapping region in helix 3 of the GACKIX domain (Figure 4.9). While these results highlight the Med15(Gal11) binding sites that are targeted by VP2 and Gcn4 in vitro, it is conceivable and indeed likely that in vivo, at physiological conditions, all or a subset of the interactions control the function of these TADs.

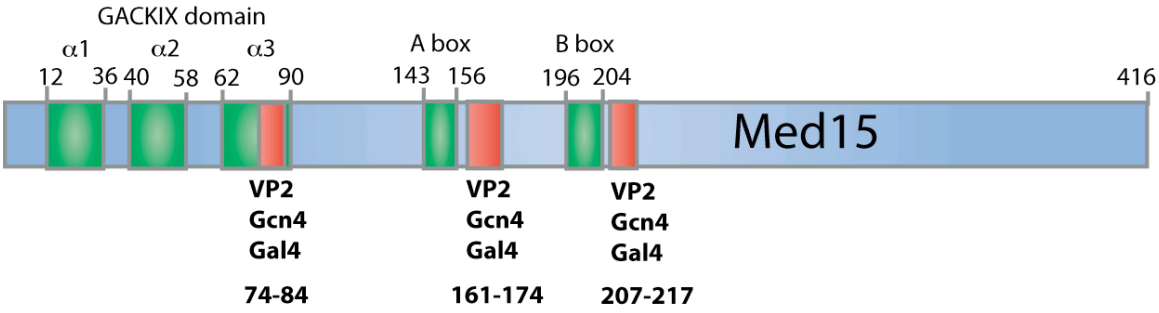


Figure 4.8 Summary of binding sites on Med15(1-416). Binding sites on Med15(1-416) as determined by mass spectrometry are indicated in red. Conserved domains are indicated in green. The Gal4 TAD used in these experiments consisted of residues 854-874 instead of the larger 840-881. This shorter peptide sequence still retains the ability to bind Med15 as shown in Figure 4.4 and it was found to be easier to detect by mass spectrometry.

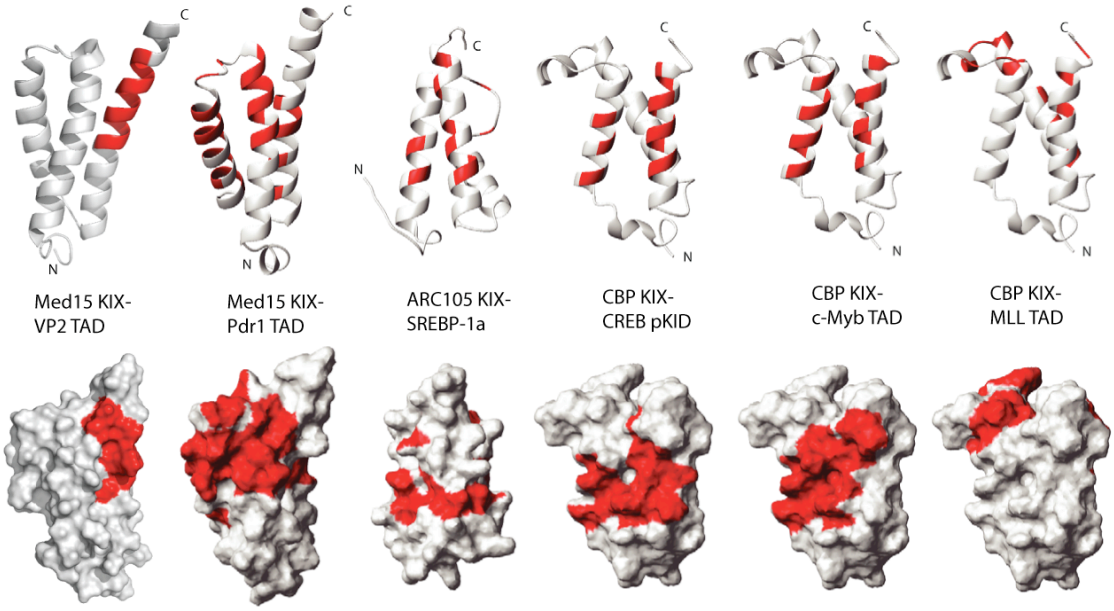


Figure 4.9 Comparison of binding of sites of activators in the GACKIX domain. The binding sites of different activators in homologous GACKIX domains are indicated in red. The solution structures of the GACKIX domains are taken from the PDB ,2KON (Med15 KIX), 2GUT (ARC105 KIX), 1KDX (CBP KIX-CREB), 1SBO (CBP KIX-c-Myb) and 2AGH (CBP KIX-MLL).

C.1.c.vi.Independent confirmation of mass spectrometry results

The binding sites within Med15(Gal11) determined by photo-crosslinking in the previous section were performed in vitro and it is conceivable that only a subset of these are physiologically relevant. To probe the significance of the interaction in cells, functional assays were performed in yeast in collaboration with Dr. Jenifer Lum. The ability of these TADs to activate transcription in reporter gene assays as Gal4 DBD fusions in various Med15(Gal11) mutant yeast strains was investigated. Interestingly, Jenifer Lum found that deletion of Med15(Gal11)(1-186) resulted only in the reduction of activity for Gal4-VP2. On the other hand, the Med15(Gal11)(1-345) deletion mutant resulted in reduction of activity of all the TADs. These studies show that in cells, VP2 most likely targets all three binding sites, while Gcn4 and Gcn4 only utilize the Med15(Gal11)(207-217) binding site for function (Figure 4.10). Additional studies with other Med15(Gal11) deletions are in progress to completely delineate the functional role of each individual binding site.

These results show that multiple activators seem to target the same binding sites on coactivators. Further these binding sites seem to be conserved between different coactivators, with the GACKIX, A and B boxes being found in many coactivators. However, the three activators investigated in this study had a different functional dependence, highlighting that in cells only a subset of binding sites are functionally relevant. Furthermore from the studies with VP2 we observe that multiple binding interactions to Med15(Gal11) are required to obtain maximal activation.

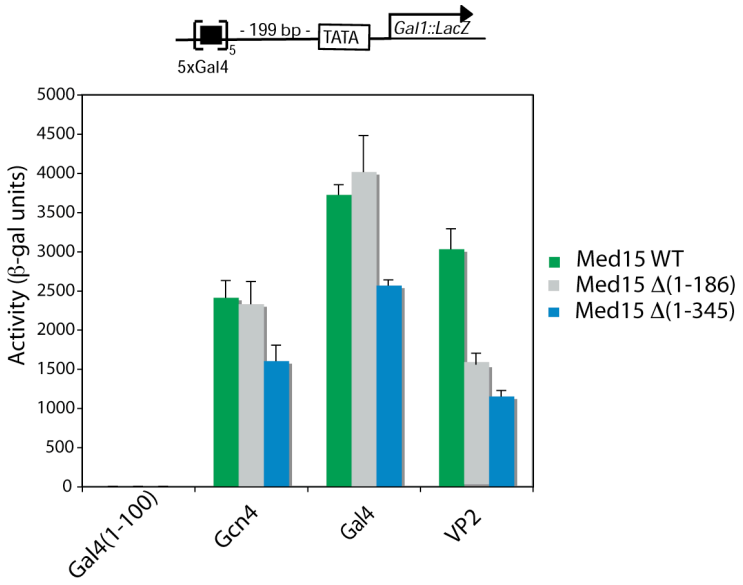


Figure 4.10 Functional evaluation of TADs in yeast. β-galactosidase assays were performed in Med15(Gal11) mutant strains containing deletions of 2 or 3 of the TAD binding sites. The TADs were fused to the Gal4(1-100) DBD. The experiments were performed in triplicate and the error is reported as SDOM.

In the future, the crosslinking combined with proteomics approaches can be extended to study other activator-coactivator interactions, to map out the critical residues important for their interactions. For example, elucidating all the physiologically relevant binding sites of the mammalian transcriptional activator (Elf1)ESX on Med23(Sur2), will assist in developing inhibitors of this interaction to treat breast cancer that results from an over expression of (Elf1)ESX regulated genes.[99]

C.2 Incorporation of unnatural amino acids in cells for in vivo crosslinking

Extension of the in vitro crosslinking strategy of determining activator binding sites to cellular approaches will not only confirm the binding sites of TADs on Med15(Gal11) but also lead to the identification of other bonafide coactivators supplementing those that have been determined using previous in vivo approaches.[27] In particular, in contrast to previous in vitro crosslinking approaches, incorporation of crosslinking amino acids pBpa and pAzpa into the TAD of activators in yeast will

facilitate the identification of cellular activator-coactivator interactions. When irradiated with UV light, pBpa and pAzpa form radicals that insert into adjacent C-H bonds of proteins, forming stable covalent crosslinks with potential direct-binding partners. For our initial experiments, we decided to incorporate pBpa and pAzpa amino acids into the well-characterized minimal Gal4 TAD comprising residues (840-881).[1] We envisioned that using western blot analysis of a Gal4 TAD-containing activator, we could visualize the incorporation of unnatural amino acids in the Gal4 TAD. Subsequently, we could use western blots and mass spectrometry to determine the identity and location of the covalent crosslinks generated by the crosslinker-incorporated activators when yeast cells are irradiated by UV light.

Several hydrophobic and acidic residues in the Gal4 TAD have been shown to be important in vitro for contact with potential coactivators and also for activity (Figure 4.11).[1, 2, 22, 100] However, replacement of a subset of these residues with other hydrophobic groups has shown to have limited impact on coactivator binding and for Gal4 function. Thus, we hypothesized that replacing residues F849, F856, F869 and D873 with pBpa or pAzpa will have minimal structural perturbation and should retain the function of the Gal4 TAD in yeast.

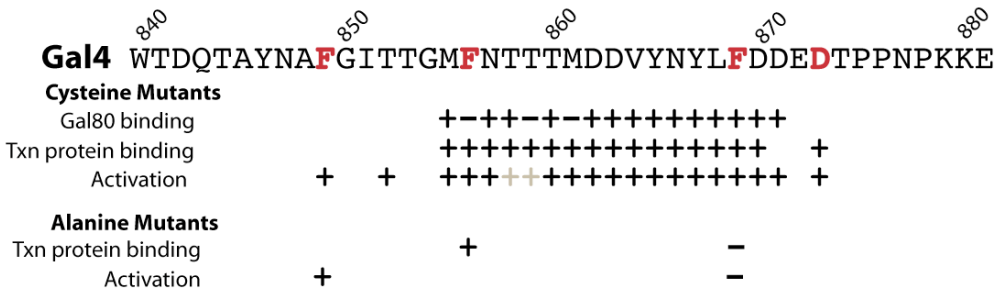


Figure 4.11 Residues in the Gal4 TAD important for function. Cys or ala mutants of the Gal4 TAD have been previously evaluated for in vitro binding and activity in yeast. + indicates the residue when mutated retains >75% function + indicates ~50-70% of wild type function and – indicates loss of function. Residues in red will be replaced with pBpa

C.2.a Incorporating pBpa and pAzpa in yeast

Using the non-sense suppression methodology, incorporation of unnatural amino acids is controlled by the amber codon.[57] Yeast expression plasmids were constructed with the Gal4 TAD (840-881) fused to the Gal4 DBD (1-147) or the LexA DBD (1-202). Site-directed mutagenesis was used to encode the amber stop codon at a specific residue. Thus, when used in cells with the appropriate tRNA/synthetase we should either see a full length product with incorporated unnatural amino acid, or we would see truncated products.

To investigate these possibilities, plasmids encoding the LexA+Gal4(840-881)F849TAG were introduced into yeast strain DSY-5 together with the cognate tRNA/synthetase (tRNAx1) encoding for pBpa or pAzpa (obtained from Dr. Peter Schultz).[52, 53] Five mL yeast cultures with and without 1 mM unnatural amino acid were grown in selective media to late-log phase, pelleted, and the contents resolved by SDS-PAGE followed by probing with a LexA antibody. Unfortunately, it was found that majority of the protein terminated at position of the TAG codon instead of incorporating any unnatural amino acid (Figure 4.12a).

C.2.a.i Optimization of incorporation efficiency

In order to increase the yield of protein that contained pBpa we decided to investigate the tRNA and synthetase expression. Previous studies by Schimmel and coworkers have demonstrated that the *E. coli* $tRNA_{CUA}$ is not expressed to high levels in yeast and the use of multiple tRNAs resulted in greater yields of full length protein product with the desired amino acid incorporated.[56, 57] Towards this end, we made expression plasmids that encoded for 2 and 3 copies of the *E. coli* $tRNA_{CUA}$ (tRNAx2 and tRNAx3) and investigated their ability to produce full length Gal4 TAD containing the unnatural amino acid. We only saw marginal improvements of yield when these tRNAs were used in combination with LexA+Gal4(840-881)F849TAG (Figure 4.12a). Another feature that affects the yield of the tRNA is the promoter it is expressed from. The original promoter used by Schultz and coworkers driving the tRNA expression has been reported to give poor yields of the tRNA as it lacks a yeast polIII promoter. Hahn and coworkers tested strong, medium and weak promoters to investigate the affect on tRNA yield.[101] They found that the medium-strength N(GTT)PR promoter allowed expression

of nearly normal levels of several proteins with pBpa incorporated. We created a plasmid based on their results, with the N(GTT)PR driving expression of a single copy of *E. coli* $tRNA_{CUA}$. Contrary to what they observed, we found that this promoter did not provide us with any higher yield (N(GTT) promoter tRNAx1). In their studies they integrated the plasmids encoding the tRNA/synthetase and their gene of interest using plasmid shuffling, perhaps improving their yields.

Recently, Chen et. al performed a similar study, where they increased the number of tRNAs to 3 and introduced a strong Pol III promoter, *pPGK1*, normally used for mRNA transcription to drive expression of these tRNAs.[102] They found that this construct gave >50 fold increase in tRNA production compared to the original plasmid. Correspondingly this translated to increased yield of protein containing the unnatural amino acid. Unfortunately, the increased yield comes at a price as the fidelity of the system seems to decrease. They show that incorporation of pBpa at a particular amber codon is only 61% while ~30% Trp is incorporated at this position. In the case of pAzpa, 90% is incorporated, but ~6% and 2% of Trp and Leu respectively are also incorporated.

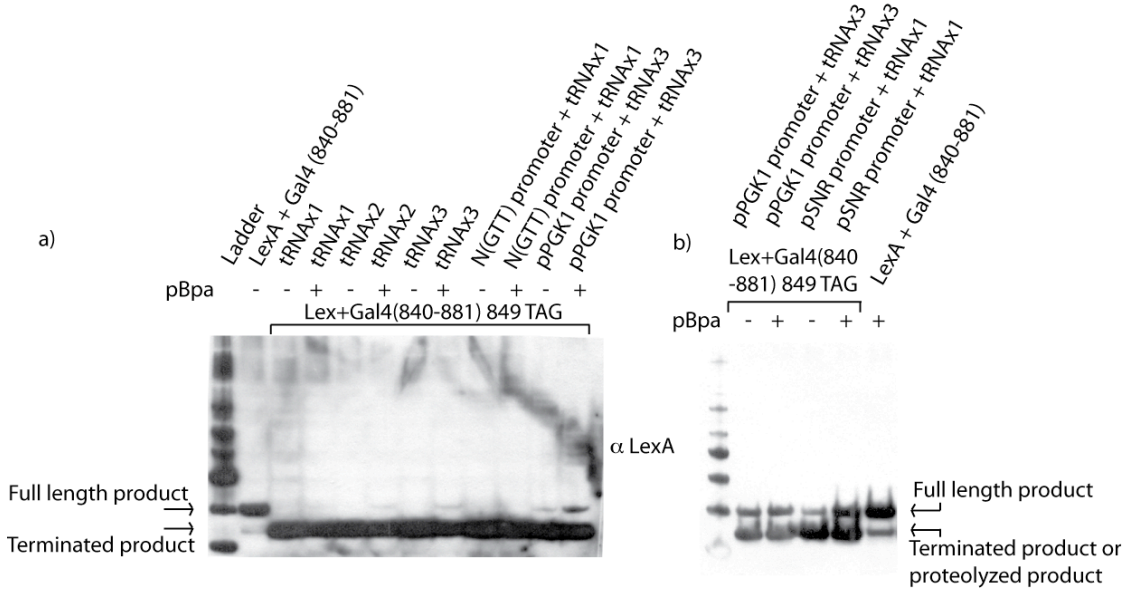


Figure 4.12 Incorporation of photo-crosslinking amino acids in the Gal4 TAD in yeast. a) Dependence of incorporation on tRNA. DSY-5 yeast expressing LexA(1-202)+Gal4(840-881)849TAG with the indicated tRNA plasmid were grown in media with and without 1 mM pBpa. 3 ODs of late log phase growing cells were probed using a LexA antibody in a western blot. b) LS41 yeast expressing LexA(1-202)+Gal4(840-881)849TAG with the indicated tRNA plasmid were probed as described in a).

We obtained the plasmids encoding three tRNAs driven by the PGK1 promoter (PGK1 promoter + tRNAx3) and evaluated their ability to incorporate pAzpa and pBpa in LexA+Gal4(840-881)F849TAG. Satisfyingly, we found that these incorporating conditions gave higher yield of full length product (Figure 4.12a). However, there was a significant amount of misincorporation (~50%), possibly Trp and/or Leu leading to full length product when no unnatural amino acid was added. However, for our purposes, these levels of misincorporated products most likely will not interfere with the crosslinking reaction by effectively competing for coactivator binding; based on related in vitro competition binding experiments it would require >10 fold excess of misincorporated protein compared to incorporated protein to completely prevent crosslinking. More recently, alleviating this concern, we have determined conditions where we observe no full-length misincorporated products in the absence of unnatural amino acids (vide infra).

Wang and coworkers also developed an optimized system for tRNA and synthetase expression.[103] They have shown that using the Pol III SNR promoter (derived from the snoRNA encoding gene SNR52) to drive tRNA expression and GDH promoter to express the synthetase, yielded ~9 fold higher levels of a control protein incorporated with tyrosine at the amber codon compared to the first generation Schultz plasmids (tRNAx1). Interestingly, they show using northern blots that the tRNA produced by the SNR promoter is in fact 100 fold lower than the tRNA produced by tRNAx1. These results suggest that the tRNA produced by the tRNAx1 plasmid is not correctly processed or modified. Based on these promising results, we developed plasmids using the SNR and GDH promoters for the incorporation of pBpa. We found that these conditions (pSNR promoter + tRNAx1) yielded the highest amount of incorporated LexA-Gal4 product, ~2 fold higher than the PGK1 promoter driving the expression of 3 tRNAs (Figure 4.12b). Further, we have also found that the incorporation levels of pBpa and pAzpa into the LexA-Gal4 constructs are comparable (Figure 4.13).

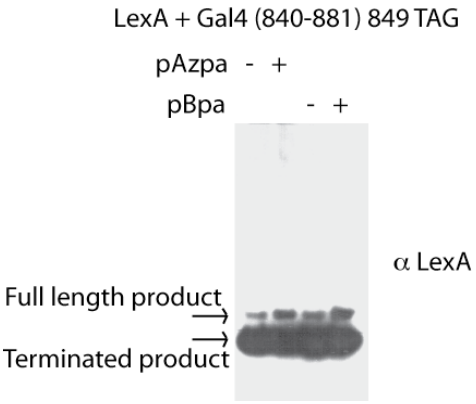


Figure 4.13
Dependence of incorporation on unnatural amino acid. DSY-5 yeast expressing LexA(1-202) + Gal4(840-881)849TAG and a plasmid expressing 3 tRNAs under the control of the PGK1 promoter and the appropriate synthetase were grown in media with and without 1 mM Azpa or pBpa. 3 OD of late log phase growing cells were probed using a LexA antibody in a western blot.

C.2.b Crosslinking in live yeast

Using the optimized conditions for expression of LexA-Gal4 constructs incorporating pBpa and pAzpa, we exposed the yeast cells to UV irradiation in order to form crosslinks between the Gal4 TAD and its binding partners. Satisfyingly, we find additional higher molecular weight protein bands in a western blot only for the samples that contained unnatural amino acid and were UV treated (Figure 4.14). Consistent with previous work showing the importance of multiple protein complexes involved in transcription we see at least three crosslinked products. One crosslinked product is at ~150 kD, potentially corresponding to a crosslink to Med15(Gal11), while another product potentially matches the a Gal80 crosslink (~80 kD). These results are indicative of functionally relevant crosslinked products and we are currently in the process of identifying them using proteomic approaches at the Michigan Proteome Consortium facility. Moreover, we will use the binding site information determined from the in vitro crosslinking studies in the previous section to assist in determining the binding site(s) of Gal4 within Med15(Gal11) in cells.

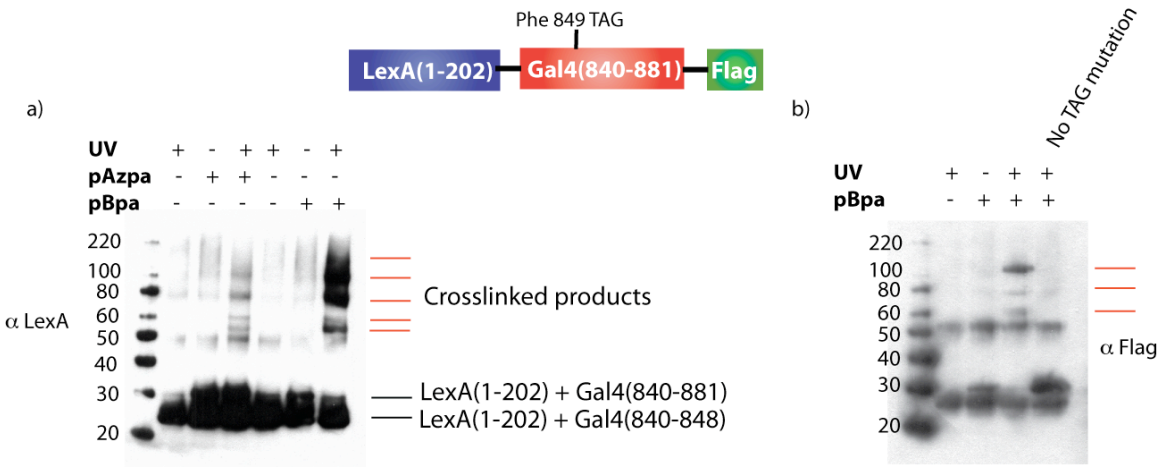
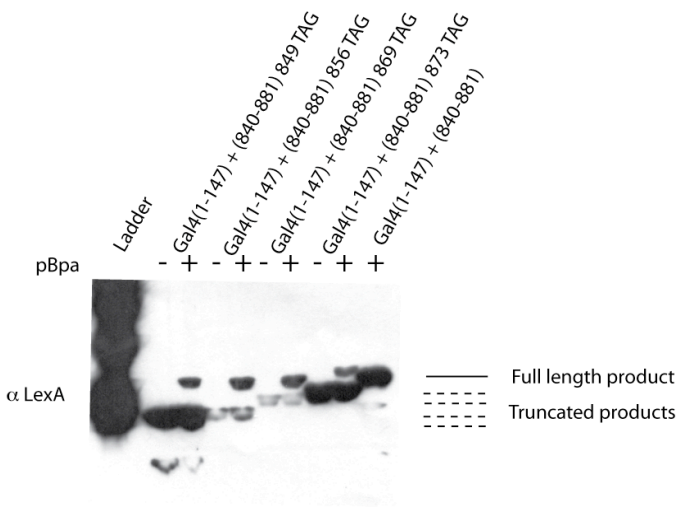


Figure 4.14 In vivo crosslinking with LexA+Gal4 constructs. LS41 yeast expressing LexA(1-202)+Gal4(840-881)849TAG and pSNR pBpa tRNA plasmids were grown in media with and without pBpa and differentially irradiated by a 365 nm light source. The activator constructs also contain a C-terminal Flag epitope tag to facilitate detection. 60 ODs of late log phase yeast were harvested and immunoprecipitated using a LexA antibody followed by western blots in which they were probed with a) a LexA antibody or b) Flag antibody. Crosslinked products are indicated in red.

More recently, we have also extended this approach to express pBpa incorporated at a variety of positions in constructs that contain the Gal4 DBD and the Gal4 TAD. In contrast to the Lex DBD+Gal4 TAD incorporation results, we find that the fidelity of incorporation is extremely high for the Gal4 DBD+TAD constructs; there is no visible misincorporated full length product when pBpa is absent from the media (Figure 4.15). One of the differences between the two DBDs used, apart from sequence composition, is that the LexA constructs were expressed on a high copy plasmid driven by the ADH promoter, while the Gal4 constructs were expressed on a low copy plasmid driven by the β-actin promoter. It is, however, not clear why these differences so significantly affect incorporation fidelity.



4.15 Incorporation of pBpa in Gal4 DBD+TAD constructs. LS41 yeast expressing Gal4(1-147+840-881) with select amber mutations and pSNR pBpa tRNA plasmid were grown in media with and without pBpa. 3 ODs of mid log phase cells were harvested and probed in a western blot with a Gal4 antibody. The lower molecular weight band in each lane represents the amber truncated product.

The Gal4 constructs developed here will be valuable tools to determine the coactivator targets in a cellular context. For example, the yeast can be grown in different sugars and the binding profile of Gal4 can be investigated in each individual case. The effect of the location of DNA binding site in the promoter on the binding profile can also be probed. Further, the effect of chromatinized DNA templates versus non-chromatinized templates can be investigated and the presence of activator DNA binding sites in general can also be investigated. Finally, the ability to incorporate the crosslinker at different positions in the TAD, facilitates the investigation of activator-specific residues important for coactivator binding.

D. Conclusions and future directions

Using photo-crosslinking and mass spectrometry we were able to develop methods to identify activator binding sites on Med15(Gal11). Verification of these results using binding and activity assays validates this technique for the future discovery of activator binding sites. Supplementary ChIP-based studies in yeast can be used to verify the recruitment of Med15(Gal11) by these activators at endogenous promoters in various Med15(Gal11) mutant strains. Somewhat unexpectedly, we found multiple overlapping activator binding sites on Med15(Gal11) for the three different TADs VP2 and Gcn4. Further, we have found the artificial TAD XL_Y to also interact with Med15(Gal11)(1-416) and binding site determination for this peptide is also underway. Moreover, several other activators have been found to interact with structurally conserved/similar binding

regions using NMR spectroscopy and mutational analysis. These findings indicate that activator binding surfaces are evolutionary conserved and predict the presence of certain binding sites on coactivator surfaces that are probably important for activation. Thus, we can rationalize our earlier studies (Chapter 2), where only a subset of ligands that targeted Med15(Gal11) activated transcription, although they had similar binding affinities. Further, the existence of conserved binding sites on different coactivators can explain how activators are able to activate transcription across species.

In the long term, the identification of activator binding sites in combination with structural studies on Med15(Gal11) and similar approaches with other coactivators, will lead to excellent screening targets for novel artificial activator discovery and also for therapeutically useful inhibitor design.

We have also shown here that unnatural amino acids can be incorporated in a TAD in yeast. More importantly, we have found that irradiation of live yeast expressing crosslinker containing activators leads to crosslinked products between activators and their binding partners. Although, the identification of the binding partners is currently underway, we find that at least three crosslinked products indicating that Gal4 most likely interacts with more than one protein in a cellular context.

In the future, studies with different activators such as Gcn4, VP2, Rap1 and Pdr1 can assess the conservation of coactivators among different classes of activators. Further incorporation of crosslinkers in the TAD and the DBD can circumvent the need for formaldehyde in ChIP to study coactivator promoter localization. In contrast to formaldehyde, the use of pBpa or pAzpa will investigate the direct binding of transcription proteins to DNA bound activators. Finally, given the recent success in using non-natural amino acids in mammalian cells, analogous crosslinking experiments with Gal4, Gcn4, VP2 in combination with mammalian activators such p53, Hif1α, ESX and p65 should be feasible in metazoan systems and would provide an explanation as to how activators from diverse species are able to activate transcription in different contexts.[45, 104]

Lastly, activator-coactivator interactions are one of first interaction pathways in transcriptional networks and determining these are a great stride forward towards mapping the network of protein-protein interactions involved in transcription. Once coactivators are identified, using unnatural amino acid mutagenesis it should be

straightforward to extend a similar approach to incorporate crosslinkers into coactivators and subsequently other transcriptional machinery proteins in efforts to map out the entire transcriptional network. A complete understanding of the interactions necessary for transcription will be useful for not only therapeutic purposes and biomanufacturing, but also in synthetic biology to reprogram these networks in efforts to build cell-based devices.

E. Experimental

Peptide Synthesis

Fmoc-based solid phase peptide synthesis was used to synthesize
VP2 DFDLDMLGD(pBpa)DLDMLGC,
Gal4(840-881)
CWTDQTAYNAFGITTGMFNTTTMDDVYNYL(pBpa)DDEDTPPNPKKE or Gal4(854-874) GMFNTTTMDDVYNYL(pBpa)DDEDT and Gcn4 TPMFEYENLEDNSKEWTSL(pBpa)DNDIPVTTDDC. FMOC-pBpa was purchased from Bachem and Chem-Impex International. The peptides were synthesized using an ABI 433A peptide synthesizer on clear amide resin (0.4 mmol/g, Peptides International) at 0.1 mmole scale using manufacturer recommended conditions, except for forced double coupling of all residues in the synthesis of Gal4 and Gcn4 after the first proline in the sequence was encountered.

To facilate labeling of the peptides with biotin, ~10 mg of peptides were dissolved in ~500 μL DMF and 100 μL CH_3CN followed by 500μL 0.1% TFA. To this a ~10 fold excess of Biotin-HPDP (*N*-(6-(Biotinamido)hexyl)-3'-(2'-pyridyldithio)-propionamide (Pierce) dissolved in 100μL DMF, 100 μL CH_3CN was added and the reaction was carried out at 37 ºC overnight. The products were characterized and purified to homogeneity using reversed-phase HPLC on a C18 column with a gradient solvent system (Buffer A: 0.1% TFA, buffer B: CH_3CN) and stored at -80 ºC.

Med15(Gal11) bacterial expression plasmids

Bacterial expression constructs expressing different fragments of Med15(Gal11) fused to the mutated OCR protein (for increasing solubility) and a His tag (for purification) were created by the High-Throughput protein lab (LSI). The plasmids also contained a cleavable TEV protease site to separate the tags from Med15(Gal11) after expression/purification. Ligation independent cloning (LIC) was used to insert the various Med15 fragments into the His-mOCR parent plasmid (IPTG inducible and ampicillin selectable).

Med15(Gal11) Expression

To identify the Med15(Gal11), fragment(s) that interact with TADs, all the His-MBP-Med15(Gal11) constructs generated by ligation independent cloning were expressed and purified at the LSI using high-throughput methods using a Biomek FX liquid handling system. The analysis of the concentration and purity of each sample was determined using a Caliper Labchip90 instrument. Samples of each protein construct were obtained in 100 mM PBS pH 7.2 with 250 mM imidazole and used without further purification for crosslinking with Gal4, Gcn4 and VP2 followed by detection by streptavidin-HRP in a western blot.

For larger scale isolations, the His-mOCR-Med15(1-416), His-mOCR-Med15(1-357) and His-MBP-Med15(107-255) plasmids were grown individually. For these purposes, plasmids encoding each construct were transformed into chemically competent Rosetta2(DE3) pLysS *E. coli* (Novagen) and cells were plated onto LB-agar plates supplemented with ampicillin (100 μg/mL) and chloramphenicol (34μg/mL). Cultures (50 mL) from single colonies were grown overnight at 37 °C (300 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100μg/mL) and chloramphenicol (34μg/mL) before addition to 1L of Select APS Super Broth supplemented with ampicillin (100μg/mL). After an OD_{600} of 0.3 was reached (300 rpm, 37 °C), the cultures were cooled to 20 °C for 1 h (100 rpm), and expression was induced with IPTG (final concentration 0.1 mM) for ~12 h (250 rpm). The cell pellet was resuspended in Buffer A (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME and Roche Complete Protease Inhibitor Cocktail) lysed using sonication, pelleted, and the His-tagged protein was isolated from the supernatant using Ni NTA-Agarose (Qiagen).

The Ni beads were washed 6 times with Wash buffer (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME, 30 mM Imidazole). The protein was eluted from the beads at 4 ºC 3 times using Elution buffer (100 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME, 300 mM Imidazole).

For His-mOCR-Med15(1-416) and His-mOCR-Med15(1-357), the His-mOCR tag was cleaved using TEV protease by incubating 0.25 mg TEV for every liter of Med15(Gal11) lysed. The eluted protein and TEV were combined in dialysis tubing (Pierce) and the buffer exchange to Storage Buffer (50 mM Tris pH 7.5 at 4 °C, 10% glycerol (v/v), 150 mM NaCl, 10 mM β-ME) overnight at 4 °C. The Med15(1-416) solution thus obtained was concentrated using a Vivascience 30K centrifugal filter device and then loaded onto a gel filtration column (Superdex 75, GE Healthcare), pre-equilibrated with Storage buffer, to remove His-mOCR. Fractions containing Med15(1-416) were pooled and concentrated using a Vivascience 10K centrifugal filter device (Vivaspin 15). The protein concentration was measured using absorbance at 280 nm. The identity and purity of the fusion protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

His-MBP-Med15(107-255) were purified similarly using gel filtration except the His-MBP tag was not cleaved as it helped stabilize the protein.

For mass spectrometry experiments the single cysteine in Med15(Gal11)(1-416) was acetylated to prevent it from reacting with the biotinylated TAD. To acetylate cysteine 306 in Med15(Gal11), Med15(1-416) was treated with 10 mM DTT for 1 h. After reducing any disufides, 40 mM iodoacetamide (1M stock in 100 mM NaHCO3) was added and the protein solution was incubated in the dark for 45 min. Excess iodoacetamide and any reducing agent were removed using a PD-10 column (GE Healthcare) pre-equilibrated with Storage buffer that did not contain any reducing agents.

In vitro Crosslinking

Biotinylated TADs were resuspended in DMSO followed by 100 mM PBS ensuring the final DMSO concentration was <10%. The concentration of the peptides were determined by measuring absorbance at 280 nm. The extinction coefficient for pBpa

was calculated to be 16,000 M^{-1} cm^{-1} (based on a titration curve with pBpa as a standard) and used to calculate the overall extinction coefficients of the peptides.

100 μL of biotinylated TAD (~100-300 μM) was mixed with 100 μL Med15 (~100-300 μM) and irradiated with a 150W 365 nm UV lamp (BIB-150P, Spectroline) at a distance of 20 cm at 4 °C for 2 h. The irradiated sample was then buffer exchanged with a 1 mL 30K MW cutoff concentrator (Millipore) with 100 mM PBS to remove free TAD (4 x 1mL). This solution was subsequently incubated with 10 μg sequencing grade trypsin (Promega) for 5 h at 37 °C. After digesting the sample, trypsin was inactivated by vortexing and incubating at 65 °C for 10 min. Next the sample was added to 100 μL of neutravidin beads (Pierce) prewashed with 100 mM PBS and incubated for 1 hr at 4 °C with shaking. To remove bound unbound tryptic Med15 fragments, the neutravidin beads were washed (6 x 1 mL) with 100 mM PBS. To elute the biotinylated products the beads were incubated 2x with 100 μL Elution buffer (100 mM PBS, 100mM DTT) for 1 h at 65 °C. The eluted samples were stored at -80 °C until mass spectrometry analysis was performed.

For mass spectrometry analysis, the samples were purified using a ZIP tip (Millipore) and exchanged to 0.1% formic acid, 70% acetonitrile or 70% methanol and a combination of MALDI and FT-ICR was performed. MALDI analysis was performed at the Michigan Proteome Consortium facility. FT-ICR electro-spray analysis was performed in the Håkansson Laboratory.

Plasmid construction

Gal4(1-147+840-881) Constructs

A low copy plasmid expressing Gal4(1-147+840-881) containing a C-terminal cysteine driven by the strong β-actin promoter was obtained from Jenifer Lum. This plasmid has a His marker for growth in yeast and an ampicillin selection in *E. coli*. TAG mutations at amino acids 849, 856, 869, and 873 were inserted as described below.

LexA(1-202)+Gal4(840-881) Constructs

A high copy plasmid expressing LexA(1-202)+Gal4(840-881)+flag peptide under the control of the ADH promoter was created from pNlexA (Origene). Amberlyn Wands

mutated the existing EcoRI and BamHI at the N terminus of LexA. Subsequently, she inserted these sites at the C-terminus of LexA to enable the creation of C terminal LexA fusions. Primers (5'-TTAC gaattc TGGACGGACCAAACTG -3') and (5'- AGT GGA TCC TTA TTT GTC GTC GTC GTC TTT ATA GTC CTC TTT TTT TGG G -3') were used to amplify Gal4 (840-881) from a plasmid expressing Gal4(1-147)+(840-881). Pfu turbo (Stratagene) was used to amplify the product with 200 nM of each primer and 4.5 mM $MgCl_2$ and 5 ng of template plasmid per 50 μL reaction. The amplification conditions are [94 °C 2', 9x (94 °C 15", 52 °C 30", 68 °C 2'), 24x (94 °C 15", 59 °C 30", 68 ° C 2'), 68 °C 7', 4 °C hold]. A total of 4, 50 μL PCRs were performed. The PCR products were purified using a Qiagen PCR purification kit and subsequently double digested with BamHI and EcoRI. The digestion reaction was purified using the Qiagen PCR purification kit and ligated with T4 DNA ligase into pCLexA (predigested with BamHI/EcoRI and treated with calf intestinal phosphatase). The resulting plasmids were amplified in Smart *E. coli* cells (Genlantis), selected on LB-agar plates containing 100 μg/ml ampicillin, and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

TAG mutations

Site-directed mutagenesis was used to insert the TAG codon at various positions in Gal4. In general primers were designed to have ~15 bases of homology on either side of the TAG mutation. Pfu turbo polymerase (Stratagene) was used to incorporate the TAG mutants with 200 nM of each primer and 4.5mM $MgCl_2$ and 5ng of template plasmid per 50 mL reaction. The amplification conditions were [95 °C 30'', 17x (95 °C 30", 55 °C 1, 68 °C 11'), 68 °C 7', 4 °C hold]. The PCR products were digested with 2 μL Dpn1 at 37 °C for 1.5 h. 5 μL of the digest was transformed in Smart *E. coli* cells (Genlantis), selected on LB-agar plates containing 100 μg/ml ampicillin, and isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). The sequences of the isolated plasmids were verified by sequencing at the University of Michigan Core Facility.

tRNA plasmids

tRNAx1 plasmid

pESC bp-2 encodes a single amber tRNA and the synthetase incorporating pBpa. pESC Az-3 encodes a single amber tRNA and the synthetase incorporating pAzpa. pESC WT encodes a single amber tRNA and the WT synthetase that incorporated tyrosine. These plasmids were obtained from Dr. Peter Schultz. They are ampicillin selectable in *E. coli* and have a Trp marker for yeast and are high copy.

tRNAx2 and tRNAx3 plasmids

Additional amber tRNAs were added by inserting NcoI and NdeI sites after the first tRNA. For this purpose, a new tRNA fragment was amplified from pESC bp-2 using primers (5' – G GG TCG ACC GGT AAG CTT CCC GAT AAG GGA GCA G -3') and (5' - CAA AAG TCC CTG AAC TTC CCC CAT GGC CTA CAT ATG TGC TGC TAG CGC CG – 3') that contained NdeI and NcoI sites after the tRNA. The PCR conditions were similar to the ones above with Pfu Turbo. The PCR product was digested with AgeI and NheI and ligated into pESC bp-2 predigested with AgeI/NheI and treated with CIP. The resulting plasmid, ptRNAx1-bp was amplified and characterized as described above.

Primers (5' - GGG TCG CCA TGG AAG CTT CCC GAT AAG GGA GCA G – 3') and (5' - CAA AAG TCC CTG AAC TTC CCC ATA TGG CCG – 3') were used to amplify the second tRNA using PCR. The PCR product was digested with NcoI and NdeI and ligated in ptRNAx1-bp predigested with NcoI/NdeI and treated with CIP. The resulting plasmid, ptRNAx2-bp was amplified and characterized as described above.

Primers (5' - GGG TCG CAT ATG AAG CTT CCC GAT AAG GGA GCA G – 3') and (5' - CAA AAG TCC CTG AAC TTC CCG CTA GCC GAC CC – 3') were used to amplify the third tRNA using PCR. The PCR product was digested with NdeI and NheI and ligated in ptRNAx2-bp predigested with NdeI/NheI and treated with CIP. The resulting plasmid, ptRNAx3-bp was amplified and characterized as described above.

N(GTT)PR promoter plasmid

To change the promoter 5' of the tRNA, DNA oligos were designed that contained the sequence of the N(GTT)PR followed by the amber tRNA and terminated

with the 3' tRNA flanking sequence. Complementary oligo set 1 (5' - CCG GTA ATA AAT ACC GGA GAT ATG ATT CAG ATG TGG AAG CGG TTT TAA AGT CCC AAT TCA TCA TTA TTT GCG TGG GGT TCC CGA G -3'), (5' - CCC TTT GGC CGC TCG GGA ACC CCA CGC AAA TAA TGA TGA ATT GGG ACT TTA AAA CCG CTT CCA CAT CTG AAT CAT ATC TCC GGT ATT TAT TA -3') and set 2 (5' - CGG CCA AAG GGA GCA GAC TCT AAA TCT GCC GTC ATC GAC TTC GAA GGT TCG AAT CCT TCC CCC ACC ACC ATT TTT TTC AAA AGT CCC TGA ACT TCC CG – 3'), (5' - CTA GCG GGA AGT TCA GGG ACT TTT GAA AAA AAT GGT GGT GGG GGA AGG ATT CGA ACC TTC GAA GTC GAT GAC GGC AGA TTT AGA GTC TGC T -3') were annealed and phosphorylated using T4 polynucleotide kinase. Both sets of oligos were then ligated simultaneously into Age/NheI digested and CIP treated pESC bp-2. The resulting plasmid, pN(GTT)PR-bp was amplified and characterized as described above.

pPR1-pGK1+3SUP4-tRNA$_{CUA}$ plasmids

pPR1-pGK1+3SUP4-tRNA$_{CUA}$-pBpa and pPR1-pGK1+3SUP4-tRNA$_{CUA}$-pAzpa expressing 3 tRNAs driven by the pGK1 promoter were obtained from Dr. Peter Schultz. They are ampicillin selectable in *E. coli* and have a Trp marker for yeast.

pSNR plasmids

pSNR WT plasmid incorporating tyrosine at the amber position was obtained from Dr. Lei Wang. pSNR pBpa and pSNR pAzpa were generated by insertion of the appropriate *E. coli* tyrosyl synthetase (amplified from pPR1-pGK1+3SUP4-tRNA$_{CUA}$-pBpa or pPR1-pGK1+3SUP4-tRNA$_{CUA}$-pAzpa) using primers 5'- AGT TCA ACT AGT ATG GCA AGC AGT AAC TTG ATT -3' and 5'-TCG ATC TCG AGT TAT TTC CAG CAA ATC AGA CA-3' followed by a SpeI/XhoI digest) into SpeI and XhoI sequentially digested and calf intestinal phosphate treated pSNR WT.

Binding experiments

The dissociation constants of fluorescence labeled TADs to Med15(Gal11) were determined using a Tecan plate reader as described in Chapter 3.

Western blots

Western blots were performed to verify the incorporation of unnatural amino acids in the Gal4(1-147+840-881) or LexA(1-202)+Gal4(840-881) amber mutant constructs. Freshly transformed yeast colonies were used to inoculate 5 ml cultures of SC media containing 2% glucose but lacking the appropriate amino acids for selection. The cultures were incubated overnight at 30°C and agitated at 250 rpm. Following incubation, these cultures were used to inoculate 5 mL cultures of SC media containing 2% glucose with or without 1 mM unnatural amino acid (dissolved in 1M NaOH) and were subsequently incubated overnight at 30 °C with agitation to an OD_{660} of 3-4. To each 5 mL culture containing unnatural amino acids, ~1 mg of pBpa or pAzpa dissolved in 50 μL 1 M NaOH was added followed by an equal volume of 1 M HCl. 3 ODs of cells were harvested, washed with water and quick-frozen in liquid nitrogen. The pellets were lysed by heating in SDS lysis buffer (50 mM TrisOAc pH 7.9, 150 mM KOAc, 20% glycerol (v/v), 0.02% Tween, 1 mM β-mercaptoethanol, 2 mM $Mg(OAc)_2$) at 90 ^oC for 10 minutes. Crude lysates were centrifuged at 14000 rpm for 2 minutes and the supernatant was resolved on NuPAGE 12% Bis-Tris gels using MES electrophoresis buffer (Invitrogen). Gal4 fusions were detected using an antibody for Gal4 (1:2000, SC-577, Santa Cruz Biotechnologies). LexA fusions were detected using an antibody for LexA (1: 1000, SC-1725, Santa Cruz Biotechnologies). Blots were developed using a horseradish peroxidase-conjugated secondary antibody (1: 5000, Santa Cruz Biotechnologies) and visualized by chemiluminescence (ECL, ECL plus or ECL advance, GE Healthcare).

In vivo crosslinking

Freshly transformed LS41 yeast colonies expressing LexA(1-202)+Gal4(840-881) 849 TAG + flag and the pSNR tRNA plasmid for pBpa or pAzpa were used to inoculate 5 ml cultures of SC media containing 2% glucose but lacking histadine and tryptophan for selection. The cultures were incubated overnight at 30°C and agitated at 250 rpm. Following incubation, these cultures were used to inoculate 50 mL cultures of SC media containing 2% glucose, with or without 1 mM unnatural amino acid (dissolved in 1M NaOH) and were subsequently incubated overnight at 30 °C with agitation to an OD_{660} of

3-4. 60 ODs of cells were isolated, washed with water and resuspended in 2 mL water and irradiated for 1 h with 365 nm light at 4 °C (Rayonet photo-reactor). Following irradiation, the cells were pelleted and stored at -80 °C until lysis. Cells were resuspended in 700 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 protease inhibitor (Roche)) and lysed using glass beads by vortexing at 4 °C using the following cycle: 30 min vortex, 10 min break, 15 min vortex. The lysate was isolated by piercing a hole in the microfuge tube and collection by centrifugation. Subsequently, the lysate was pelleted and the supernatant incubated with 10 μL of LexA antibody (D-19 or N-19, Santa Cruz Biotechnologies) for 5 h at 4 °C for immunoprecipitation. The protein bound to the antibody was isolated by incubation for 1 h with ~100 μL of prewashed protein G magnetic beads (Dynal Corporation) at 4 °C. The beads were washed 4X with 1 mL Wash Buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5% NP-40, 1% Na-Deoxycholate and 1 mM EDTA) and stored dry at -20 °C. The protein was eluted from the beads by heating at 90 °C for 10 min in NuPAGE SDS gel loading buffer (Invitrogen) and probed using a western blot as described above.

References

1. Ansari, A. Z.; Reece, R. J.; Ptashne, M., A transcriptional activating region with two contrasting modes of protein interaction. *Proc Natl Acad Sci U S A* **1998,** 95, (23), 13543-8.
2. Reeves, W. M.; Hahn, S., Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* **2005,** 25, (20), 9092-102.
3. Ptashne, M.; Gann, A., *Genes & Signals*. Cold Spring Harbor Laboratory: New York, 2001.
4. Molina, M. P.; Cain, C.; Bargonetti, J., In vivo footprinting and DNA affinity chromatography for analysis of p53 DNA binding ability. *Methods Mol Biol* **2003,** 234, 151-70.
5. Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B., Genome-wide mapping of in vivo protein-DNA interactions. *Science* **2007,** 316, (5830), 1497-502.
6. Sikder, D.; Kodadek, T., Genomic studies of transcription factor-DNA interactions. *Curr Opin Chem Biol* **2005,** 9, (1), 38-45.
7. Mapp, A. K.; Ansari, A. Z., A TAD further: exogenous control of gene activation. *ACS Chem Biol* **2007,** 2, (1), 62-75.

8. Ansari, A. Z.; Koh, S. S.; Zaman, Z.; Bongards, C.; Lehming, N.; Young, R. A.; Ptashne, M., Transcriptional activating regions target a cyclin-dependent kinase. *Proc Natl Acad Sci U S A* **2002,** 99, (23), 14706-9.
9. Chang, C.; Gonzalez, F.; Rothermel, B.; Sun, L.; Johnston, S. A.; Kodadek, T., The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. *J Biol Chem* **2001,** 276, (33), 30956-63.
10. Jeong, C. J.; Yang, S. H.; Xie, Y.; Zhang, L.; Johnston, S. A.; Kodadek, T., Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **2001,** 40, (31), 9421-7.
11. Myers, L. C.; Gustafsson, C. M.; Hayashibara, K. C.; Brown, P. O.; Kornberg, R. D., Mediator protein mutations that selectively abolish activated transcription. *Proc Natl Acad Sci U S A* **1999,** 96, (1), 67-72.
12. Swanson, M. J.; Qiu, H.; Sumibcay, L.; Krueger, A.; Kim, S. J.; Natarajan, K.; Yoon, S.; Hinnebusch, A. G., A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol Cell Biol* **2003,** 23, (8), 2800-20.
13. Zhang, F.; Sumibcay, L.; Hinnebusch, A. G.; Swanson, M. J., A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol Cell Biol* **2004,** 24, (15), 6871-86.
14. Fishburn, J.; Mohibullah, N.; Hahn, S., Function of a eukaryotic transcription activator during the transcription cycle. *Mol Cell* **2005,** 18, (3), 369-78.
15. Volkman, H. M.; Rutledge, S. E.; Schepartz, A., Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain. *J Am Chem Soc* **2005,** 127, 4649-4658.
16. Wu, Z.; Belanger, G.; Brennan, B. B.; Lum, J. K.; Minter, A. R.; Rowe, S. P.; Plachetka, A.; Majmudar, C. Y.; Mapp, A. K., Targeting the transcriptional machinery with unique artificial transcriptional activators. *J Am Chem Soc* **2003,** 125, (41), 12390-1.
17. Lee, Y. C.; Park, J. M.; Min, S.; Han, S. J.; Kim, Y. J., An activator binding module of yeast RNA polymerase II holoenzyme. *Molecular and Cellular Biology* **1999,** 19, (4), 2967-2976.
18. Park, J. M.; Kim, H. S.; Han, S. J.; Hwang, M. S.; Lee, Y. C.; Kim, Y. J., In vivo requirement of activator-specific binding targets of mediator. *Molecular and Cellular Biology* **2000,** 20, (23), 8709-8719.
19. Archer, C. T.; Burdine, L.; Kodadek, T., Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking. *Mol Biosyst* **2005,** 1, (5-6), 366-72.
20. Koh, S. S.; Ansari, A. Z.; Ptashne, M.; Young, R. A., An activator target in the RNA polymerase II holoenzyme. *Mol Cell* **1998,** 1, (6), 895-904.
21. Hall, D. B.; Struhl, K., The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. *J Biol Chem* **2002,** 277, (48), 46043-50.
22. Melcher, K.; Johnston, S. A., GAL4 interacts with TATA-binding protein and coactivators. *Mol Cell Biol* **1995,** 15, (5), 2839-48.
23. Stringer, K. F.; Ingles, C. J.; Greenblatt, J., Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* **1990,** 345, (6278), 783-6.

24. Bryant, G. O.; Ptashne, M., Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol Cell* **2003,** 11, (5), 1301-9.
25. Bhaumik, S. R.; Green, M. R., SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes Dev* **2001,** 15, (15), 1935-45.
26. Bhaumik, S. R.; Green, M. R., Interaction of Gal4p with components of transcription machinery in vivo. *Methods Enzymol* **2003,** 370, 445-54.
27. Bhaumik, S. R.; Raha, T.; Aiello, D. P.; Green, M. R., In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev* **2004,** 18, (3), 333-43.
28. Lehming, N., Analysis of protein-protein proximities using the split-ubiquitin system. *Brief Funct Genomic Proteomic* **2002,** 1, (3), 230-8.
29. Lim, M. K.; Tang, V.; Le Saux, A.; Schuller, J.; Bongards, C.; Lehming, N., Gal11p dosage-compensates transcriptional activator deletions via Taf14p. *J Mol Biol* **2007,** 374, (1), 9-23.
30. Tanaka, Y.; Bond, M. R.; Kohler, J. J., Photocrosslinkers illuminate interactions in living cells. *Mol Biosyst* **2008,** 4, (6), 473-80.
31. Melcher, K.; Chen, H. T., Identification and analysis of multiprotein complexes through chemical crosslinking. *Curr Protoc Cell Biol* **2007,** Chapter 17, Unit 17 10.
32. Sadakane, Y.; Hatanaka, Y., Photochemical fishing approaches for identifying target proteins and elucidating the structure of a ligand-binding region using carbene-generating photoreactive probes. *Anal Sci* **2006,** 22, (2), 209-18.
33. Suchanek, M.; Radzikowska, A.; Thiele, C., Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat Methods* **2005,** 2, (4), 261-7.
34. MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J., Photo-leucine incorporation reveals the target of a cyclodepsipeptide inhibitor of cotranslational translocation. *J Am Chem Soc* **2007,** 129, (47), 14560-1.
35. Vila-Perello, M.; Pratt, M. R.; Tulin, F.; Muir, T. W., Covalent capture of phospho-dependent protein oligomerization by site-specific incorporation of a diazirine photo-cross-linker. *J Am Chem Soc* **2007,** 129, (26), 8068-9.
36. Torchilin, V. P., Tat peptide-mediated intracellular delivery of pharmaceutical nanocarriers. *Adv Drug Deliv Rev* **2008,** 60, (4-5), 548-58.
37. Gump, J. M.; Dowdy, S. F., TAT transduction: the molecular mechanism and therapeutic prospects. *Trends Mol Med* **2007,** 13, (10), 443-8.
38. Gustafsson, A. B.; Gottlieb, R. A.; Granville, D. J., TAT-mediated protein transduction: delivering biologically active proteins to the heart. *Methods Mol Med* **2005,** 112, 81-90.
39. Wadia, J. S.; Dowdy, S. F., Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer. *Adv Drug Deliv Rev* **2005,** 57, (4), 579-96.
40. Brooks, H.; Lebleu, B.; Vives, E., Tat peptide-mediated cellular delivery: back to basics. *Adv Drug Deliv Rev* **2005,** 57, (4), 559-77.
41. Palm, C.; Netzereab, S.; Hallbrink, M., Quantitatively determined uptake of cell-penetrating peptides in non-mammalian cells with an evaluation of degradation and antimicrobial effects. *Peptides* **2006,** 27, (7), 1710-6.

42. Parenteau, J.; Klinck, R.; Good, L.; Langel, U.; Wellinger, R. J.; Elela, S. A., Free uptake of cell-penetrating peptides by fission yeast. *FEBS Lett* **2005,** 579, (21), 4873-8.
43. Rajarao, G. K.; Nekhotiaeva, N.; Good, L., Peptide-mediated delivery of green fluorescent protein into yeasts and bacteria. *FEMS Microbiol Lett* **2002,** 215, (2), 267-72.
44. Holm, T.; Netzereab, S.; Hansen, M.; Langel, U.; Hallbrink, M., Uptake of cell-penetrating peptides in yeasts. *FEBS Lett* **2005,** 579, (23), 5217-22.
45. Liu, W.; Brock, A.; Chen, S.; Schultz, P. G., Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* **2007,** 4, (3), 239-44.
46. Xie, J.; Schultz, P. G., A chemical toolkit for proteins--an expanded genetic code. *Nat Rev Mol Cell Biol* **2006,** 7, (10), 775-82.
47. Wang, L.; Xie, J.; Schultz, P. G., Expanding the genetic code. *Annu Rev Biophys Biomol Struct* **2006,** 35, 225-49.
48. Ryu, Y.; Schultz, P. G., Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat Methods* **2006,** 3, (4), 263-5.
49. Xie, J.; Schultz, P. G., An expanding genetic code. *Methods* **2005,** 36, (3), 227-38.
50. Xie, J.; Schultz, P. G., Adding amino acids to the genetic repertoire. *Curr Opin Chem Biol* **2005,** 9, (6), 548-54.
51. Zhang, Z.; Smith, B. A.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G., A new strategy for the site-specific modification of proteins in vivo. *Biochemistry* **2003,** 42, (22), 6735-46.
52. Chin, J. W.; Cropp, T. A.; Chu, S.; Meggers, E.; Schultz, P. G., Progress toward an expanded eukaryotic genetic code. *Chem Biol* **2003,** 10, (6), 511-9.
53. Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G., An expanded eukaryotic genetic code. *Science* **2003,** 301, (5635), 964-7.
54. Wang, L.; Schultz, P. G., Expanding the genetic code. *Chem Commun (Camb)* **2002**, (1), 1-11.
55. Wakasugi, K.; Quinn, C. L.; Tao, N.; Schimmel, P., Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. *EMBO J* **1998,** 17, (1), 297-305.
56. Edwards, H.; Trezeguet, V.; Schimmel, P., An Escherichia coli tyrosine transfer RNA is a leucine-specific transfer RNA in the yeast Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **1991,** 88, (4), 1153-6.
57. Edwards, H.; Schimmel, P., A bacterial amber suppressor in Saccharomyces cerevisiae is selectively recognized by a bacterial aminoacyl-tRNA synthetase. *Mol Cell Biol* **1990,** 10, (4), 1633-41.
58. Edwards, H.; Schimmel, P., An E. coli aminoacyl-tRNA synthetase can substitute for yeast mitochondrial enzyme function in vivo. *Cell* **1987,** 51, (4), 643-9.
59. Cellitti, S. E.; Jones, D. H.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu, H.; Brittain, S. M.; Brinker, A.; Caldwell, J.; Bursulaya, B.; Spraggon, G.; Brock, A.; Ryu, Y.; Uno, T.; Schultz, P. G.; Geierstanger, B. H., In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J Am Chem Soc* **2008,** 130, (29), 9268-81.
60. Tsao, M. L.; Summerer, D.; Ryu, Y.; Schultz, P. G., The genetic incorporation of a distance probe into proteins in Escherichia coli. *J Am Chem Soc* **2006,** 128, (14), 4572-3.

61. Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G., The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat Biotechnol* **2004,** 22, (10), 1297-301.
62. Chin, J. W.; Schultz, P. G., In vivo photocrosslinking with unnatural amino Acid mutagenesis. *Chembiochem* **2002,** 3, (11), 1135-7.
63. Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G., Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **2002,** 99, (17), 11020-4.
64. Wittelsberger, A.; Mierke, D. F.; Rosenblatt, M., Mapping ligand-receptor interfaces: approaching the resolution limit of benzophenone-based photoaffinity scanning. *Chem Biol Drug Des* **2008,** 71, (4), 380-3.
65. Ahrends, R.; Kosinski, J.; Kirsch, D.; Manelyte, L.; Giron-Monzon, L.; Hummerich, L.; Schulz, O.; Spengler, B.; Friedhoff, P., Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry. *Nucleic Acids Res* **2006,** 34, (10), 3169-80.
66. Leszyk, J.; Tao, T.; Nuwaysir, L. M.; Gergely, J., Identification of the photocrosslinking sites in troponin-I with 4-maleimidobenzophenone labelled mutant troponin-Cs having single cysteines at positions 158 and 21. *J Muscle Res Cell Motil* **1998,** 19, (5), 479-90.
67. Weber, P. J.; Beck-Sickinger, A. G., Comparison of the photochemical behavior of four different photoactivatable probes. *J Pept Res* **1997,** 49, (5), 375-83.
68. Dorman, G.; Prestwich, G. D., Benzophenone photophores in biochemistry. *Biochemistry* **1994,** 33, (19), 5661-73.
69. Kauer, J. C.; Erickson-Viitanen, S.; Wolfe, H. R., Jr.; DeGrado, W. F., p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. *J Biol Chem* **1986,** 261, (23), 10695-700.
70. Farrell, I. S.; Toroney, R.; Hazen, J. L.; Mehl, R. A.; Chin, J. W., Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat Methods* **2005,** 2, (5), 377-84.
71. Lu, Z.; Ansari, A. Z.; Lu, X. Y.; Ogirala, A.; Ptashne, M., A target essential for the activity of a nonacidic yeast transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (13), 8591-8596.
72. Majmudar, C. Y.; Lum, J. K.; Prasov, L.; Mapp, A. K., Functional specificity of artificial transcriptional activators. *Chem Biol* **2005,** 12, (3), 313-21.
73. Sakurai, H.; Hiraoka, Y.; Fukasawa, T., Yeast GAL11 protein is a distinctive type transcription factor that enhances basal transcription in vitro. *Proc Natl Acad Sci U S A* **1993,** 90, (18), 8382-6.
74. Sakurai, H.; Ohishi, T.; Amakasu, H.; Fukasawa, T., Yeast GAL11 protein stimulates basal transcription in a gene-specific manner by a mechanism distinct from that by DNA-bound activators. *FEBS Lett* **1994,** 351, (2), 176-80.
75. Sakurai, H.; Fukasawa, T., Yeast Gal11 and transcription factor IIE function through a common pathway in transcriptional regulation. *J Biol Chem* **1997,** 272, (51), 32663-9.
76. Sakurai, H.; Fukasawa, T., Functional correlation among Gal11, transcription factor (TF) IIE, and TFIIH in Saccharomyces cerevisiae. Gal11 and TFIIE cooperatively

enhance TFIIH-mediated phosphorylation of RNA polymerase II carboxyl-terminal domain sequences. *J Biol Chem* **1998,** 273, (16), 9534-8.
77. Sakurai, H.; Fukasawa, T., Functional connections between mediator components and general transcription factors of Saccharomyces cerevisiae. *J Biol Chem* **2000,** 275, (47), 37251-6.
78. Novatchkova, M.; Eisenhaber, F., Linking transcriptional mediators via the GACKIX domain super family. *Current biology : CB* **2004,** 14, (2), R54-5.
79. Robinette, D.; Neamati, N.; Tomer, K. B.; Borchers, C. H., Photoaffinity labeling combined with mass spectrometric approaches as a tool for structural proteomics. *Expert Rev Proteomics* **2006,** 3, (4), 399-408.
80. Trakselis, M. A.; Alley, S. C.; Ishmael, F. T., Identification and mapping of protein-protein interactions by a combination of cross-linking, cleavage, and proteomics. *Bioconjug Chem* **2005,** 16, (4), 741-50.
81. Son, C. D.; Sargsyan, H.; Hurst, G. B.; Naider, F.; Becker, J. M., Analysis of ligand-receptor cross-linked fragments by mass spectrometry. *J Pept Res* **2005,** 65, (3), 418-26.
82. Girault, S.; Sagan, S.; Bolbach, G.; Lavielle, S.; Chassaing, G., The use of photolabelled peptides to localize the substance-P-binding site in the human neurokinin-1 tachykinin receptor. *Eur J Biochem* **1996,** 240, (1), 215-22.
83. Girault, S.; Chassaing, G.; Blais, J. C.; Brunot, A.; Bolbach, G., Coupling of MALDI-TOF mass analysis to the separation of biotinylated peptides by magnetic streptavidin beads. *Anal Chem* **1996,** 68, (13), 2122-6.
84. Jahn, O.; Eckart, K.; Tezval, H.; Spiess, J., Characterization of peptide-protein interactions using photoaffinity labeling and LC/MS. *Anal Bioanal Chem* **2004,** 378, (4), 1031-6.
85. Hatanaka, Y.; Sadakane, Y., Photoaffinity labeling in drug discovery and developments: chemical gateway for entering proteomic frontier. *Curr Top Med Chem* **2002,** 2, (3), 271-88.
86. Cusick, M. E.; Klitgord, N.; Vidal, M.; Hill, D. E., Interactome: gateway into systems biology. *Hum Mol Genet* **2005,** 14 Spec No. 2, R171-81.
87. Jahn, O.; Eckart, K.; Brauns, O.; Tezval, H.; Spiess, J., The binding protein of corticotropin-releasing factor: ligand-binding site and subunit structure. *Proc Natl Acad Sci U S A* **2002,** 99, (19), 12055-60.
88. Sinz, A., Investigation of protein-ligand interactions by mass spectrometry. *ChemMedChem* **2007,** 2, (4), 425-31.
89. Suzuki, Y.; Nogi, Y.; Abe, A.; Fukasawa, T., GAL11 protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in Saccharomyces cerevisiae. *Mol Cell Biol* **1988,** 8, (11), 4991-9.
90. Sullivan, S. M.; Horn, P. J.; Olson, V. A.; Koop, A. H.; Niu, W.; Ebright, R. H.; Triezenberg, S. J., Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Research* **1998,** 26, (19), 4487-4496.
91. Drysdale, C. M.; Duenas, E.; Jackson, B. M.; Reusser, U.; Braus, G. H.; Hinnebusch, A. G., The Transcriptional Activator Gcn4 Contains Multiple Activation Domains That Are Critically Dependent on Hydrophobic Amino-Acids. *Molecular and Cellular Biology* **1995,** 15, (3), 1220-1233.

92. Jackson, B. M.; Drysdale, C. M.; Natarajan, K.; Hinnebusch, A. G., Identification of seven hydrophobic clusters in GCN4 making redundant contributions to transcriptional activation. *Mol Cell Biol* **1996,** 16, (10), 5557-71.
93. Kim, D. H.; Kim, G. S.; Yun, C. H.; Lee, Y. C., Functional conservation of the glutamine-rich domains of yeast Gal11 and human SRC-1 in the transactivation of glucocorticoid receptor Tau 1 in Saccharomyces cerevisiae. *Mol Cell Biol* **2008,** 28, (3), 913-25.
94. De Guzman, R. N.; Goto, N. K.; Dyson, H. J.; Wright, P. E., Structural basis for cooperative transcription factor binding to the CBP coactivator. *J Mol Biol* **2006,** 355, (5), 1005-13.
95. Radhakrishnan, I.; Perez-Alvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E., Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* **1997,** 91, (6), 741-52.
96. Thakur, J. K.; Arthanari, H.; Yang, F.; Pan, S. J.; Fan, X.; Breger, J.; Frueh, D. P.; Gulshan, K.; Li, D. K.; Mylonakis, E.; Struhl, K.; Moye-Rowley, W. S.; Cormack, B. P.; Wagner, G.; Naar, A. M., A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **2008,** 452, (7187), 604-9.
97. Yang, F.; Vought, B. W.; Satterlee, J. S.; Walker, A. K.; Jim Sun, Z. Y.; Watts, J. L.; DeBeaumont, R.; Saito, R. M.; Hyberts, S. G.; Yang, S.; Macol, C.; Iyer, L.; Tjian, R.; van den Heuvel, S.; Hart, A. C.; Wagner, G.; Naar, A. M., An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **2006,** 442, (7103), 700-4.
98. Zor, T.; De Guzman, R. N.; Dyson, H. J.; Wright, P. E., Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *Journal of Molecular Biology* **2004,** 337, (3), 521-534.
99. Asada, S.; Choi, Y.; Yamada, M.; Wang, S. C.; Hung, M. C.; Qin, J.; Uesugi, M., External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc Natl Acad Sci U S A* **2002,** 99, (20), 12747-52.
100. Zenke, F. T.; Engles, R.; Vollenbroich, V.; Meyer, J.; Hollenberg, C. P.; Breunig, K. D., Activation of Gal4p by galactose-dependent interaction of galactokinase and Gal80p. *Science* **1996,** 272, (5268), 1662-5.
101. Chen, H. T.; Warfield, L.; Hahn, S., The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol* **2007,** 14, (8), 696-703.
102. Chen, S.; Schultz, P. G.; Brock, A., An improved system for the generation and analysis of mutant proteins containing unnatural amino acids in Saccharomyces cerevisiae. *J Mol Biol* **2007,** 371, (1), 112-22.
103. Wang, Q.; Wang, L., New methods enabling efficient incorporation of unnatural amino acids in yeast. *J Am Chem Soc* **2008,** 130, (19), 6066-7.
104. Baruah, H.; Puthenveetil, S.; Choi, Y. A.; Shah, S.; Ting, A. Y., An Engineered Aryl Azide Ligase for Site-Specific Mapping of Protein-Protein Interactions through Photo-Cross-Linking. *Angew Chem Int Ed Engl* **2008,** 47, (37), 7018-7021.

CHAPTER 5
CONCLUSIONS AND FUTURE DIRECTIONS

A. Introduction

Regulated gene expression is essential for the appropriate development and the continued existence of all organisms. Transcriptional activators play a critical role in this process, precisely controlling the transcriptional response of their cognate genes based upon environmental needs. Given their importance in transcription, it is not surprising that many diseases arise due to malfunctioning transcriptional activators. Thus, activator artificial transcription factors (activator ATFs), molecules that mimic the function of natural activator proteins, have been proposed as effective replacements for malfunctioning activators to treat diseases ranging from cancer to diabetes.[1, 2] One of the major long-term goals of our research is to develop more drug-like small molecule activator ATFs for these purposes.[3-5] Activator ATFs that precisely control gene expression are also in demand for synthetic biology applications to modulate biological circuits to develop cell-based devices.[6] One of the biggest hurdles, however, is that the features endogenous activators use for function are not well defined, making it difficult to design activator ATFs that work in an analogous manner.[7]

As outlined in this dissertation, we have investigated the mechanism of endogenous transcriptional activators to determine the features they utilize to activate transcription to high levels. In the process, we have also developed new generations of activator ATFs with advantageous specificity and stability properties. Incorporation of all these factors into future generations of non-biopolymer and small molecule activator ATFs will serve to develop molecules with enhanced utility for a variety of applications.

B. Multiple interactions are required for potent function

A key observation arising from this work is that we have found that multiple interactions inside and outside the transcriptional machinery are required for enhanced cellular function of an activator. In Chapter 2, we screened a synthetic peptide library and isolated ligands that target the yeast Mediator component Med15(Gal11). We found that the ligands that functioned by targeting a single transcriptional machinery component, specifically recruiting the coactivator Med15(Gal11), were only able to activate transcription to moderate levels. While activator ATFs developed using these ligands can be useful to specifically target diseased cells and tissues, they do not function with levels of activity that resemble natural activators and other features are required for these activator ATFs to serve as effective replacements.

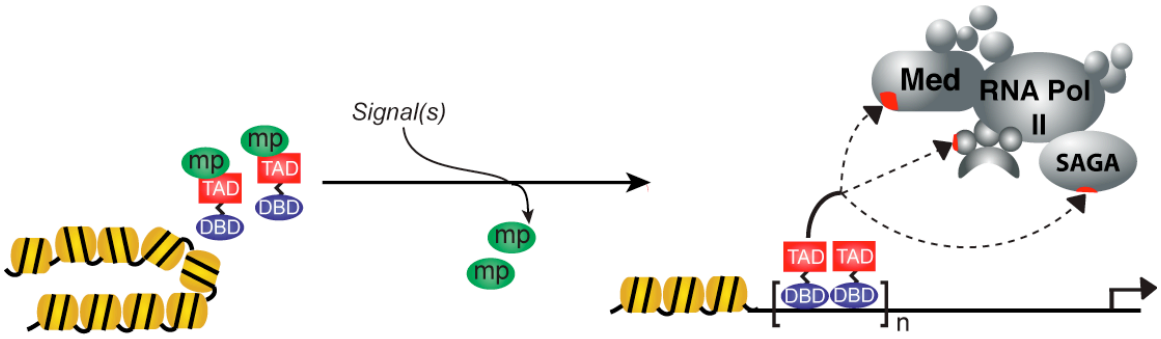


Figure 5.1 Multiple TAD interactions inside and outside the transcriptional machinery. Masking interactions with masking proteins (mp) greatly enhance activator function, presumably by increasing the amount of activator available for productive transcriptional machinery interactions. Recruiting multiple coactivator proteins, involved in different processes required for transcriptional activation, results in a synergistic increase in activation.

In Chapter 3 we isolated ligands that target the coactivator Tra1 and we found that the ligand that functioned with the highest activity did so by targeting multiple coactivators. These findings were further substantiated when we showed that combination of the Med15(Gal11)-specific ligands with Tra1 ligands resulted in a synergistic increase in activity.

Interactions outside the transcriptional machinery also seem to enhance activator function. Also in Chapter 3, we found that incorporating a masking interaction leads to a large increase in transcriptional activation compared to activators that do not incorporate a masking interaction. Although in this case the enhanced function is most likely not due to interactions that directly involve contacts with the transcriptional machinery, we hypothesized that the masking interactions positively affecting the stability of an activator, leading to effective coactivator recruitment.

Finally, in Chapter 4, we investigated the binding partners of natural activators using in vitro and in vivo crosslinking experiments. Consistent with our findings in Chapters 2 and 3 that multiple interactions enhance the function of activator ATFs, we find that the potent yeast activator Gal4 interacts with at least three different proteins, suggesting that natural activators use multiple interactions to activate transcription to high levels.

C. Existence of privileged activator binding sites on coactivators

We have found that interactions with certain activator binding sites on coactivators are more amenable to high levels of transcriptional activation than others. In Chapter 2, we found that ligands that in vitro target different surfaces of the coactivator Med15(Gal11) with similar affinities to activate transcription to varying levels in yeast. The differences in activity levels was contributed to the different binding sites on Med15(Gal11) that the ligands targeted, illustrating for the first time that binding site(s) and not solely binding affinity was also responsible for controlling activator potency.

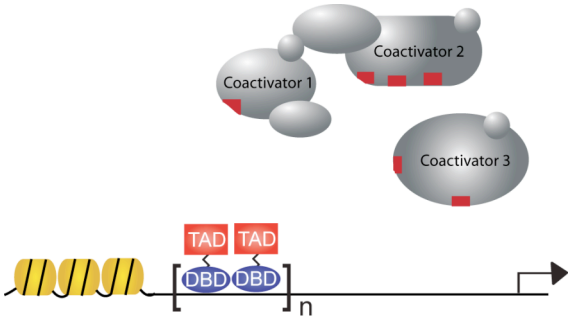


Figure 5.2 Interaction with privileged binding sites. Certain binding sites (shown in red) within coactivators are more accessible to transcriptional activators and they are conserved across different coactivators. Targeting these binding sites enables the recruiting of multiple coactivators using the same TAD surface, resulting in high levels of transcriptional activation.

In Chapter 3, we found that the Tra1 ligand that activated transcription to the highest levels targeted a binding site shared with the natural activator VP2. Moreover, the

ligand was also able to target another coactivator Med15(Gal11), at an overlapping binging site with VP2. These findings suggest that natural activators have evolved to target certain privileged binding sites on coactivators that are conserved between coactivators, despite presumed differences in function.

To further strength the idea of the existence of a set of privileged binding sites, in Chapter 4, we found using in vitro crosslinking followed by functional evaluation in cells that three different activators with varying sequence compositions, target one overlapping binding site on the coactivator Med15(Gal11). Further, BLAST searches of these binding sites indicates homology to a variety of transcription proteins across species, although the full-length proteins are not significantly homologous. These finding illustrate that natural activators target similar binding sites that have evolved to exist on multiple coactivators. Finally, the results also highlight how natural activators are able to target multiple, different proteins using the same small set of amino acid residues and activate across species.

D. Future directions

D.1 Conformation

Activators are predicted to be unstructured in solution and adopt a particular conformation upon binding. Thus investigating the structure of activators in complex with different target proteins will assist in elucidating the role of structural changes in transcription. Further structures of activator-coactivator complexes will assist in the design of small molecules to functionally mimic activators.

In Chapter 4, we determined that Med15(Gal11)(1-357) contains binding sites for the yeast activator Gal4. Several studies have shown that amphipathic activators adopt a helical conformation on target binding. However, a study by Johnston and coworkers on the Gal4 TAD has predicted it forms a β-sheet in solution.[8, 9] The Ptashne group has shown that certain proline mutations in the Gal4 TAD prevent an interaction with its masking protein Gal80, but permit an interaction with the transcriptional machinery.[10] Taken together, these results suggest that the interaction with Gal80 induces formation of some ordered structure incompatible with the presence of proline residues; in contrast, to

interact with the transcriptional machinery such a structure is not required. Thus it is likely that Gal4 may obtain a different structure depending on the target it is bound to.

To investigate the conformational preference of Gal4, in collaboration with the Ansari and Markely groups (U. of Wisconsin) we are using NMR spectroscopy to investigate the conformational changes in the solution structure of Gal4 when complexed with the inhibitor Gal80 or the coactivator Med15(Gal11)(1-357). Thus far we have found that the Gal4 TAD is unstructured (based on ^{13}C chemical shifts obtained for $^{13}C,^{15}N$ labeled Gal4 TAD) under the experimental conditions investigated. Currently, we are in the process of performing NMR experiments with Gal4 complexed to Med15(Gal11).

D.2 Solid state characterization

The identification of three binding sites on Med15(Gal11) for VP2 raises the question if these binding sites are structurally similar and whether VP2 uses a similar interaction to target these sites. Further, the exact function of Med15(Gal11) is unknown and it has long poly-glutamine stretches suggesting it might have unique structural features compared to other coactivators. To get further insight into Med15(Gal11)-activator interactions, in collaboration with the LSI protein crystallography facility and the Matzger group we are investigating the use of polymer-induced crystals to crystallize Med15(Gal11)(1-357) in complex with VP2 and other TADs. Currently, we are in the process of optimizing screening conditions to grow larger, diffractable crystals of Med15(Gal11)(1-357) complexed to VP2.

D.3 Kinetics

We have found three factors that contribute to the high levels of natural activator function: targeting specific binding sites, recruiting multiple transcriptional machinery proteins and incorporation of masking lead to increased activity levels. However, other factors may also contribute to function. For example, the kinetics of formation and dissociation of TAD-coactivator complexes are likely to be significant contributors to transcriptional output. [11-15]. Since initiation of transcription includes a series of coupled binding events including a number of activator-target interactions that result in PIC formation and transcription, the rates of association and dissociation could potentially control activator function. FRET studies can be used to probe these association and

dissociation kinetics. Using the Med15(Gal11) binding sites for VP2, Gcn4 and Gal4 TADs determined in Chapter 4, we can site specifically incorporate one FRET pair in close proximity to the activator binding site(s) on Med15(Gal11) and another in the TAD to measure the kinetics. These studies can be performed in vitro or in cells using unnatural amino acid incorporation or by labeling Med15(Gal11) with a fluorophore.

D.4 Long term goal of developing therapeutic small molecule ATFs

In the efforts described here we have primarily focused on developing design principles to create synthetic transcriptional activators that upregulate gene expression. We have found that targeting a single protein in the transcriptional machinery, in absence of any other interactions, does not yield high levels of activity. On the other hand, multiple interactions with the transcriptional machinery and at transcriptionally relevant binding sites yields high levels of activation. Further, incorporation of a masking interaction that shields the TAD from aggregation and/or premature degradation can have a profound impact on transcriptional output.

Thus, in the future the molecules that display a promiscuous binding profile in the transcriptional machinery, but also incorporate a masking interaction that controls the promiscuity to target particular privileged binding sites on coactivators will lead to molecules with greatly enhanced function. For instance, one might consider performing a small molecule screen to identify molecules that disrupt an endogenous activator-coactivator interaction thereby targeting the same binding site as the endogenous activator. Another approach could involve the design of molecules to display functional groups important for endogenous activator function on an effective scaffold. Subsequently, fusion of the prospective TAD via a flexible linker to an appropriate DBD that also incorporates a masking feature should result in molecules that rival the cell-based activity of natural activators. Such molecules will be of great benefit for biomanufacturing fermentations where, for instance, the expression of an enzyme needs to be increased to a desired level to increase product yields.

However, in order for activator ATFs to be therapeutically useful, several other criteria need to be fulfilled. For example, activator ATFs will need to integrate with cellular signals and respond to a number of physiological cues for temporal control of their function. Not only would they have to respond to external stimuli, but also traffic to

the nucleus and once they have finished activating there needs to be a mechanism to turn them off. Further complicating matters, when administered in humans, there are also non-functional considerations such as toxicity and immunogenicity that will have to be addressed. There has been some recent progress on these front, for example, small molecule responsive activators have been developed and peptidic sequences corresponding to nuclear localization signals have also been identified.[16-18]

Despite all the challenges to develop a therapeutically useful activator ATF, a biopolymer-based activator ATFs is currently undergoing clinical trials for the treatment of diabetic neuropathy (Sangamo Biosciences). Thus, the first generation of transcription-based therapeutics are at the horizon and the tools described here will assist in fulfilling the biggest limitations of small molecule activator ATFs, robust cellular activity, generating a new class of therapeutics that will be proteolytically more stable, less immunogenetic and likely be more cell permeable.

References

1. Melnick, A.; Licht, J. D., Deconstructing a disease: RAR alpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia. *Blood* **1999,** 93, (10), 3167-3215.
2. Mistry, A. R.; Pedersen, E. W.; Solomon, E.; Grimwade, D., The molecular pathogenesis of acute promyelocytic leukaemia: implications for the clinical management of the disease. *Blood Reviews* **2003,** 17, (2), 71-97.
3. Darnell, J. E., Transcription factors as targets for cancer therapy. *Nature Reviews Cancer* **2002,** 2, (10), 740-749.
4. Pandolfi, P. P., Transcription therapy for cancer. *Oncogene* **2001,** 20, (24), 3116-3127.
5. Majmudar, C. Y.; Mapp, A. K., Chemical approaches to transcriptional regulation. *Curr Opin Chem Biol* **2005,** 9, (5), 467-74.
6. McDaniel, R.; Weiss, R., Advances in synthetic biology: on the path from prototypes to applications. *Curr Opin Biotechnol* **2005,** 16, (4), 476-83.
7. Mapp, A. K.; Ansari, A. Z., A TAD further: exogenous control of gene activation. *ACS Chem Biol* **2007,** 2, (1), 62-75.
8. Leuther, K. K.; Salmeron, J. M.; Johnston, S. A., Genetic evidence that an activation domain of GAL4 does not require acidity and may form a beta sheet. *Cell* **1993,** 72, (4), 575-85.
9. Van Hoy, M.; Leuther, K. K.; Kodadek, T.; Johnston, S. A., The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets. *Cell* **1993,** 72, (4), 587-94.

10. Ansari, A. Z.; Reece, R. J.; Ptashne, M., A transcriptional activating region with two contrasting modes of protein interaction. *Proc Natl Acad Sci U S A* **1998,** 95, (23), 13543-8.
11. Paal, K.; Baeuerle, P. A.; Schmitz, M. L., Basal transcription factors TBP and TFIIB and the viral coactivator E1A 13S bind with distinct affinities and kinetics to the transactivation domain of NF-kappaB p65. *Nucleic Acids Res* **1997,** 25, (5), 1050-5.
12. Ferreira, M. E.; Hermann, S.; Prochasson, P.; Workman, J. L.; Berndt, K. D.; Wright, A. P., Mechanism of transcription factor recruitment by acidic activators. *J Biol Chem* **2005,** 280, (23), 21779-84.
13. Hermann, S.; Berndt, K. D.; Wright, A. P., How transcriptional activators bind target proteins. *J Biol Chem* **2001,** 276, (43), 40127-32.
14. Wu, Y.; Reece, R. J.; Ptashne, M., Quantitation of putative activator-target affinities predicts transcriptional activating potentials. *Embo J* **1996,** 15, (15), 3951-63.
15. Nedialkov, Y. A.; Triezenberg, S. J., Quantitative assessment of in vitro interactions implicates TATA-binding protein as a target of the VP16C transcriptional activation region. *Arch Biochem Biophys* **2004,** 425, (1), 77-86.
16. Buskirk, A. R.; Landrigan, A.; Liu, D. R., Engineering a ligand-dependent RNA transcriptional activator. *Chem Biol* **2004,** 11, (8), 1157-1163.
17. Penolazzi, L.; Borgatti, M.; Lambertini, E.; Mischiati, C.; Finotti, A.; Romanelli, A.; Saviano, M.; Pedone, C.; Piva, R.; Gambari, R., Peptide nucleic acid-DNA decoy chimeras targeting NF-kappaB transcription factors: Induction of apoptosis in human primary osteoclasts. *Int J Mol Med* **2004,** 14, (2), 145-52.
18. Holm, T.; Netzereab, S.; Hansen, M.; Langel, U.; Hallbrink, M., Uptake of cell-penetrating peptides in yeasts. *FEBS Lett* **2005,** 579, (23), 5217-22.