CHARACTERIZATION OF THE DYNAMIC INTERACTIONS OF TRANSCRIPTIONAL ACTIVATORS

by

Amberlyn M. Wands

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2010

Doctoral Committee:

Associate Professor Anna K. Mapp, Chair
Professor Hashim M. Al-Hashimi
Professor E Neil G. Marsh
Associate Professor Jorge A. Iñiguez-Lluhí

Acknowledgements

I have so many people to thank for helping me throughout my graduate school career. First, I would like to thank my advisor Dr. Anna Mapp for all of the guidance you have given me, as well as allowing me the freedom to express myself as a scientist. Your patience and confidence in my abilities means a lot to me, and I promise to keep working on presenting myself to others in a positive yet assertive manner. I would also like to thank you for taking the time to instill in your students the importance of thinking and writing critically about scientific concepts, which I know we will carry with us into our future careers. Next I would like to thank my committee members for their time, and for always asking me challenging questions that made me look at my projects from a different perspective. I would also like to give a special thanks to Dr. Carol Fierke and Dr. John Hsieh for their willingness to work on a collaboration with people starting with a minimal background in the field of transient kinetics. Their love of solving kinetic problems is inspiring, and I appreciate being given the opportunity to work with them.

I was fortunate from the beginning when I was a rotation student in Anna's lab to be mentored by Jenifer Lum and Chinmay Majmudar. I want to thank Jen for not only inspiring me to be a better scientist but also a better person, and for really caring about my well being. I want to thank Chinmay for teaching me that you have to be resourceful to get ahead in life, and for pushing me to be the best that I can be and never letting me give up. I was also fortunate to be able to be a mentor to Ningkun Wang, and I am looking forward to hearing about all of the great work that she will produce in her graduate career. To the rest of my lab mates, I am really going to miss seeing you every day. In particular, to Amy Danowitz, your friendship has been a blessing to me. I want to thank you for always being true to yourself and for being an oyster. To Caleb Bates, Jonas Højfeldt, Lori Lee (the "new" post-doc), Chris Taylor, and Ryan Casey, I'm glad I was given the opportunity to get to know you better over the years, and I wish you all the best in your future endeavors.

I was also fortunate to make friends within the Chemistry Department as well. The first real friends that I made in Michigan were with my fellow chemical biology students Li Wang and Jingjie Mo; our first two years of graduate school were filled with many late nights studying for our classes. In addition, I later met my good friend and roommate June Yang during my rotation in Anna Mapp's lab, with whom I shared a fondness for reading fiction books and watching movies in our spare time. I am also very grateful that I met my dear friends Kendra Reid and Charles Evans, whom I will miss. To Kendra, I especially want to thank you for sharing my same sense of humor; you have made me laugh through some pretty hard times, and I hope I was able to do the same for you. I would also like to acknowledge the good times that I had with my friends Tasneem Patwa, Sarah Nehm (who was technically a Pharmaceutical Sciences student), and Gayle Gawlik.

Finally, I would like to thank my family and friends back home for their unwavering support. To my mother and my brother, who I know are proud of me for working hard and pursuing my career. To my father and Amy who were able to come to my defense. To my Aunt Cinda and Uncle Bruce, as well as my cousins Brian, David, and Heather, as well as their families, who I always looked forward to seeing when I came home for the holidays. And finally to my Grandma Freeland for her patience when I kept saying it might take me a little longer to graduate than I thought. And last but not least, I would like to thank my friends Dona Williams, Julie Widener, and Jackie Franceschetti for believing in me.

Table of Contents

List of Figures

(continued)

List of Tables

List of Abbreviations

3-AT	3-Amino-1,2,4-triazole
ABTS	2,2'-Azino-bis(3-Ethylbenzthiazoline-6-Sulfonic Acid)
αCTD	alpha C-terminal domain
AFP	Autofluorescent protein
ATF	Artificial transcription factor
ATP	Adenosine-5'-triphosphate
bp	base pair
BSA	Bovine serum albumin
CBP	CREB binding protein
CCPGCC	Tetracysteine motif for biarsenical labeling
ChIP	Chromatin immunoprecipitation
CPMG RD	Carr-Purcell-Meiboom-Gill relaxation dispersion
CPRG	Chlorophenol red-β-D-galactopyranoside
CRE	cAMP response elements
CREB	cAMP response element-binding
cryo-EM	cryo-electron microscopy
Cys	Cysteine
DBD	DNA-binding domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EC50	Half maximal effective concentration
ECFP	Enhanced cyan fluorescent protein
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immunosorbent assay
ERα	Estrogen receptor alpha
EYFP	Enhanced yellow fluorescent protein
FlAsH	4,5-bis(1,3,2-dithiarsolan-2-yl)-fluorescein
FP	Fluorescence polarization
FRET	Förster resonance energy transfer
Gal4-dd	Gal4 dimerization domain
GST	Glutathione S-transferase
GTF	General transcription factor
H_2O_2	Hydrogen peroxide
His	Histidine
His_6	Hexahistidine
HRP	Horseradish peroxidase
IDP	Intrinsically disordered protein
Im	Imidazole
IPTG	Isopropyl β-D-1-thiogalactopyranoside

IUP	Intrinsically unstructured protein
$K_{d,app}$	Apparent dissociation constant
KID	Kinase-inducible domain of CREB
KIX	Kinase-inducible domain interacting domain of CBP
k_{obs}	observed rate
LiOAc	Lithium acetate
MBP	Maltose binding protein
MDM2	Murine double minute 2
min	Minutes
mRNA	Messenger RNA
NA	NeutrAvidin
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NRSE	Neuron-restrictive silencer element
NRSF	Neuron restrictive silencer factor
ONPG	*ortho*-Nitrophenyl-β-galactoside
P1	Rb-binding peptide with sequence YGLWILWCDEEGLDLG
P201	Peptide with sequence YLLPTCIP
pBpa	*p*-benzoyl-L-phenylalanine
PBS	Phosphate Buffered Saline
PDB	Protein data bank
PIC	Pre-initiation complex
Py	Pyrrole
Rb	Retinoblastoma tumor suppressor protein
RDCs	Residual dipolar couplings
RE1	Repressor element 1
ReAsH	4,5-bis(1,3,2-dithiarsolan-2-yl)-resorufin
REST	Repressor element 1 silencing transcription factor
RNA	Ribonucleic acid
RNA Pol II	RNA Polymerase II
rpm	Revolutions per minute
SAGA	Spt-Ada-Gcn5-Acetyltransferase
SDOM	Standard deviation of the mean
SPR	Surface plasmon resonance
TAD	Transcriptional activation domain
TAF	TBP associated factor
TBP	TATA-binding protein
TFIID	Transcription factor II D
TFO	Triplex-forming oligonucleotide
XL_Y	Peptide with sequence LTGLFVQDYLLPTCIP
Y2H	Yeast two-hybrid
ZF	Zinc finger
Zn	Zinc

Abstract

CHARACTERIZATION OF THE DYNAMIC INTERACTIONS OF TRANSCRIPTIONAL ACTIVATORS

by

Amberlyn M. Wands

Chair: Anna K. Mapp

Transcription is initiated through a series of coupled binding equilibria between transcriptional activators and their array of protein targets within the transcriptional machinery. However, previous efforts to kinetically characterize these interactions have produced conflicting models for the mechanism of complex formation, which is hampering the discovery of non-natural mimics of their transcriptional activation domains (TADs). Using fluorescence stopped-flow techniques, we determined that the activators Gal4, Gcn4, and VP16 interact with the same coactivator, Med15, via a two-step binding mechanism comprised of a bimolecular association step and a conformational change. We further hypothesized that the life-times of these interactions

should be more revealing of differences in activator potency (i.e., transcriptional output); thus, we analyzed the microscopic rate and equilibrium constants defining the individual steps within our mechanism, in order to identify key trends that can differentiate the activators from one another in terms of their ability to recruit the transcriptional machinery to a gene promoter. We determined that it is the favorability of the conformational change step and its partition ratio that correlates with the ability of an activator to stimulate transcription. Future studies will focus on determining how the different structural propensities of the TAD sequences contribute to the stability of the intermediate that they form.

Furthermore, another significant challenge in the development of artificial transcription factors (ATFs) is a lack of small molecules that can be used to localize them to a gene promoter in a cellular context. We propose a novel approach to accomplish this task which relies on the interaction of a ligand with an endogenous DNA-bound protein. To this end, we have successfully used an *in vitro* phage display selection with a random 12 amino acid peptide library to isolate ligands that are capable of interacting with DNA-binding proteins, such as the DNA-binding domain of the yeast transcriptional activator Gal4 (residues 1-100) and the bacterial repressor LexA (residues 1-202). Future studies will entail the implementation of a selection with a conformation-constrained peptide library to obtain ligands that may possess increased stability and specificity within the cellular milieu. In particular, protein scaffolds that promote helix stabilization would aid in the future identification of peptidomimetic or small molecule replacements.

CHAPTER 1

INTRODUCTION

A. Introduction

Nature has developed an intricate system to control the differential expression of protein-encoding genes from a common genome, thus giving rise to various cell and tissue types within the human body [1]. In order to turn on a desired set of genes in response to signaling cues, proteins known as transcriptional activators interact with and recruit the transcriptional machinery to the appropriate promoter regions (Figure 1.1) [2]. The dynamic protein-DNA and protein-protein interactions that activators form throughout this process regulate the timing and extent of mRNA levels produced [3]. It is therefore not surprising that malfunctioning transcription factors can lead to aberrant gene expression profiles and disease states, including developmental defects, cancer, and diabetes [4-7]. My thesis research efforts have focused on studying the mechanism by which natural activators carry out these functions to aid in the development of artificial replacements that can be used as mechanistic tools and therapeutic agents.

B. Activated transcription

RNA polymerase II (RNA Pol II) is responsible for synthesizing messenger RNA transcripts (mRNA) from a genomic DNA template [8]. RNA Pol II activity requires the assembly of general transcription factors (GTFs) onto core promoter elements of a gene

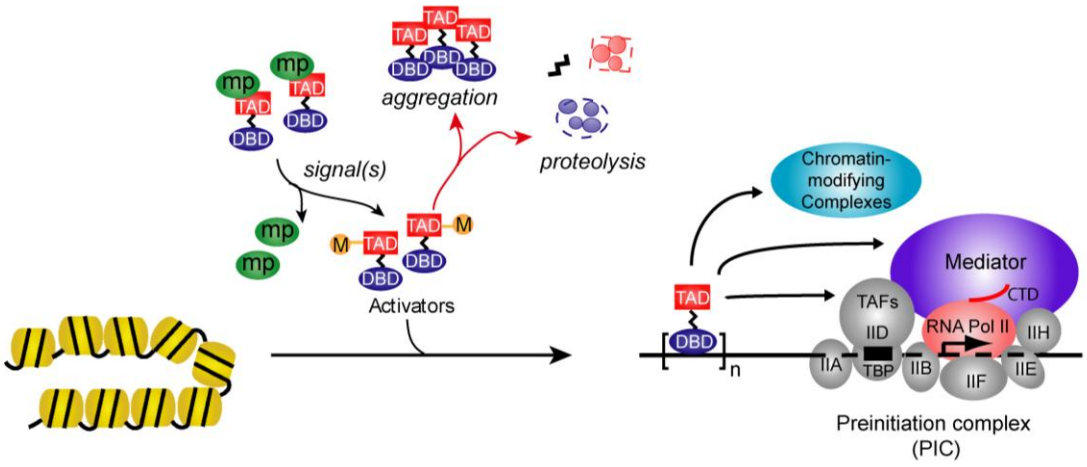


Figure 1.1: Function of a transcriptional activator. Transcriptional activators contain two essential domains, a DNA-binding domain (DBD, blue circle) and a transcriptional activation domain (TAD, red square). The cell uses a number of mechanisms to regulate activator function, including masking proteins (mp) that preclude exposure of the TAD until signaling pathways (such as post-translational modifications) render it available. During transcriptional activation, activators bind to specific sites on DNA and recruit transcriptional coactivators to gene promoters, thus facilitating assembly of the pre-initiation complex (PIC).

which serve to specifically position the polymerase to recognize the transcription start site and aid in the initiation of transcription; this is known as the pre-initiation complex (PIC), and its assembly is sufficient to support basal transcription levels [3, 9]. Activated transcription occurs when sequence-specific DNA-binding proteins called transcriptional activators facilitate assembly of the PIC at the promoter in a timely manner by the recruitment of proteins termed coactivators [10]. One class of coactivators achieves this by serving as adapters to components of the general transcriptional machinery [11]. For example, transcriptional activators are unable to stimulate transcription in a reconstituted yeast system with pure RNA Pol II and GTFs (regardless of the fact that they have been shown to bind many of these components *in vitro*) until addition of a 20-25 polypeptide

complex termed Mediator [2, 9, 12-17]. Consistent with the role of Mediator as an intermediary factor, a cryo-electron microscopy (EM) image of the yeast Mediator-RNA Pol II complex reveals the existence of 'head' and 'middle' domains which make multiple contacts on the back face of the polymerase, leaving ~75% of the RNA Pol II surface available for interactions with other components of the PIC, while a 'tail' domain of Mediator extends away from RNA Pol II and contains proteins that have been shown to be targets of transcriptional activators (Figure 1.2) [14, 18]. Another class of coactivators functions to overcome the repressive effects of chromatin packaging at the promoter [19-23]. This is achieved by their modification of the structure of chromatin via ATP-dependent remodeling (e.g., Swi/Snf complex) or histone-modification (e.g., SAGA complex) [23-25]. However, although recruitment is the general mechanism of transcriptional activation in eukaryotes, the identity and timing of the complexes that are recruited are gene-specific events [26].

C. Natural transcriptional activators

C.1. Modular architecture

In order to carry out their function, activators must participate in many protein–DNA and protein–protein interactions, yet this can be accomplished with a fairly simple architecture. Transcriptional activators utilize minimally two key domains: a DNA-binding domain (DBD) that localizes the activator to the appropriate site within the genome, and a transcriptional activation domain (TAD) that interacts with targets within the transcriptional machinery [2, 10]. This modularity was originally demonstrated by Brent and Ptashne in a 'domain-swapping' experiment, in which the DBD of the bacterial

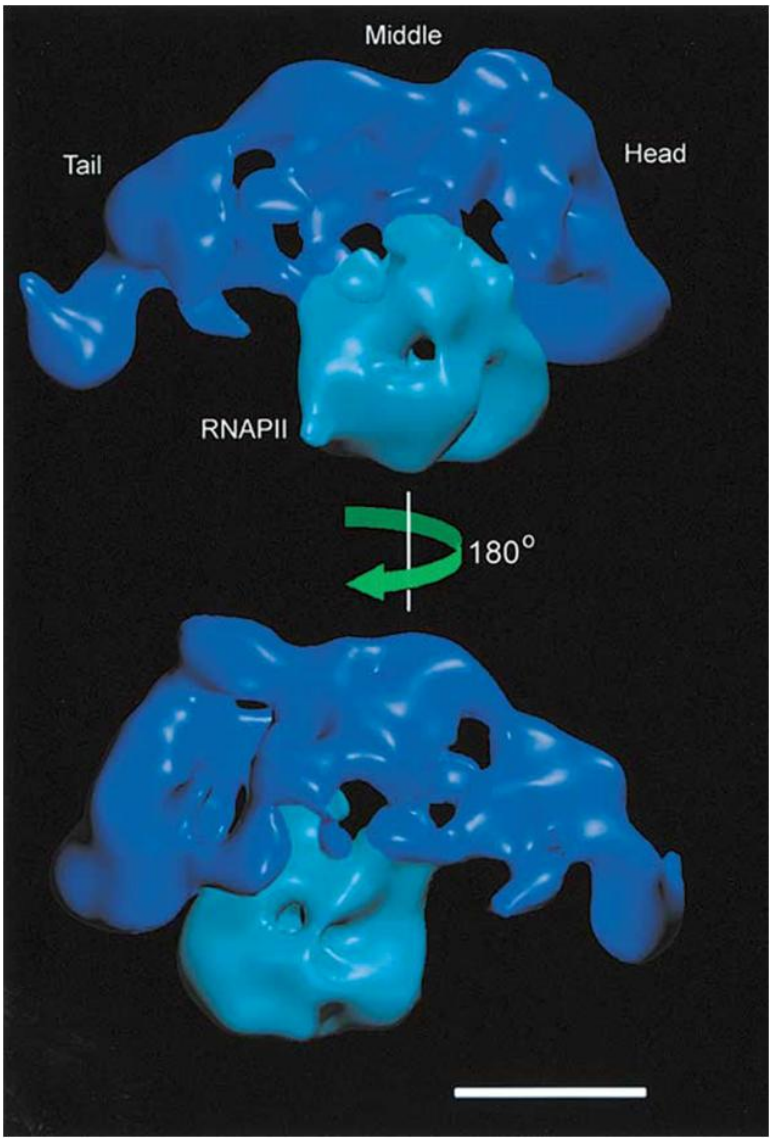


Figure 1.2: Mediator complex in yeast. The Mediator complex is composed of 20-25 proteins and is thought to serve as a conduit between transcriptional activators and RNA polymerase II. The cryo-EM image of the yeast Mediator-RNA Pol II complex reveals the three domains of Mediator termed Head (H), Middle (M) and Tail (T). Scale bar = 100 Å. Reprinted from Molecular Cell, v10, Davis, J.A.; Takagi, Y.; Kornberg, R.D.; Asturias, F.J., Structure of the yeast RNA Polymerase II Holoenzyme: Mediator Conformation and Polymerase Interaction, 409-415, Copyright (2002), with permission from Elsevier.

transcriptional repressor LexA was fused to the TAD of the yeast transcriptional activator Gal4, resulting in a hybrid protein that was able to function as a transcriptional activator at LexA binding sites [27]. Indeed, this revelation has allowed for the implementation of a modular replacement strategy for the development of artificial transcription factors (ATFs), in which the functional domains of endogenous transcriptional regulators are substituted with artificial counterparts [28]. Therefore, in this section the important structural and binding characteristics of DBDs and TADs from natural activators that are required to reconstitute activator function will be further discussed.

C.1.A. DNA-binding domain (DBD)

DNA-binding domains (DBDs) utilized by transcriptional activators are typically composed of globular yet diverse protein folds, as determined by their extensive structural characterization by nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography [29]. DBDs bind to a specific DNA sequence by possessing a complementary surface of favorable electrostatic and/or van der Waals interactions with the base pairs and sugar-phosphate backbone, most commonly through the use of side chains protruding from recognition motifs that can interact with the major and minor grooves of B-form DNA [30-31]. Moreover, multimerization of these polypeptide chains can increase the specificity of the sequences that activators are capable of targeting, as well as their binding affinity. For example, due to the ability of the yeast activator Gcn4 to form a coiled-coil homodimer via its leucine zipper, the basic region of its continuous alpha helices are positioned to interact with adjacent pseudo-palindromic 4 bp half sites (Figure 1.3a) [30, 32]. Similarly, the $Zn(II)_2Cys_6$ binuclear clusters of a homodimer of the

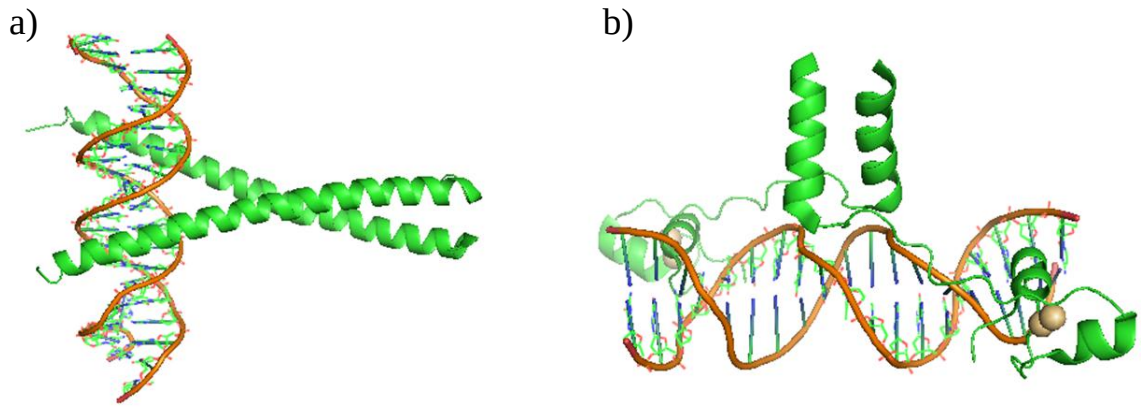


Figure 1.3: DBD-DNA interactions. a) Ribbon diagram of the crystal structure of the basic region leucine zipper (bZIP) DNA-binding motif of Gcn4 bound to DNA as a dimer (PDB ID: 1YSA). b) Ribbon diagram of the crystal structure of the zinc cluster DNA-binding motif of Gal4 bound to DNA as a dimer (PDB ID: 1D66).

yeast activator Gal4 can make base-pair specific contacts to highly conserved CGG triplets located at the ends of a 17 bp recognition sequence (Figure 1.3b) [33-34]. Even further, the viral activator VP16 utilizes protein-protein interactions with two host-cell factors, HCF-1 and Oct-1, to assemble a multiprotein complex that enhances its affinity for a "TAATGARAT" recognition element [35]. To this end, Chapter 3 of this dissertation discusses a new strategy to utilize peptide-protein interactions to hijack these DNA-binding properties.

C.1.B. Transcriptional activation domain (TAD)

C.1.B.1. Structural studies

Transcriptional activation domains (TADs) have been classified based on their amino acid composition as amphipathic (containing an interspersion of polar and hydrophobic residues), proline-rich, or glutamine-rich [4, 28]. The largest and best-studied class of

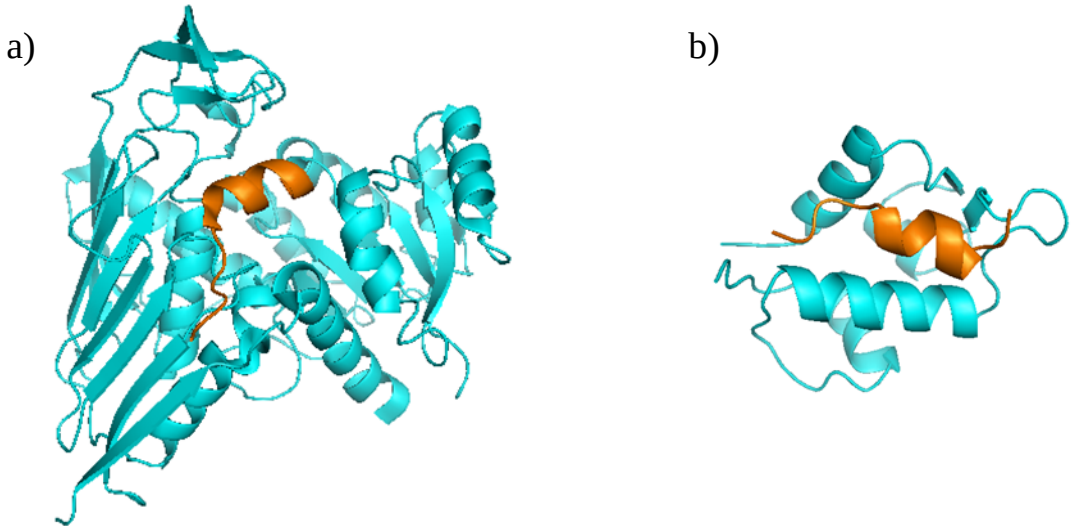


Figure 1.4: TAD-target interactions. a) Ribbon diagram of the crystal structure of the Gal4 TAD (orange) bound as an α-helix to its masking protein Gal80 (cyan) (PDB ID: 3E1K). b) Ribbon diagram of the crystal structure of the p53 TAD (orange) bound as an α-helix to its masking protein MDM2 (cyan) (PDB ID: 1YCR).

activators is the amphipathic class [2]. They are classified as intrinsically unstructured or disordered proteins (IUPs or IDPs), defined as proteins that lack intrinsic globular structure or contain long disordered segments in their normal, functional state [26, 36-39]. This common structural feature was initially established from circular dichroism (CD) and/or nuclear magnetic resonance (NMR) experiments performed on isolated TADs (e.g., TADs from p53, Gal4, Gcn4, VP16, c-Myc, NF-κB, GR, AR, ER), which illustrated their largely unstructured character in aqueous solution and neutral pH [25-26, 37-38, 40-48]. However, under specific conditions, such as hydrophobic solvents or lowered pH, these TADs have a propensity for α-helix or β-sheet formation; conditions that are fulfilled upon interaction with target proteins as well [43-44, 49]. For example, the crystal structures of the TAD•masking protein complexes of Gal4•Gal80 and p53•MDM2 show the TADs bound as α helices (Figure 1.4) [50-52]. However, with

these limited structural studies it remains unclear if there is a common structural motif, such as an α-helix, that is utilized by all activators of this class.

Recent studies have shown that the coactivator-binding motifs located within activators are primed for target recognition by the presence of transiently-structured motifs. For example, the N-terminus of the tumor suppressor p53 (residues 1-93) consists of a transcriptional activation domain (TAD, residues 1-67) followed by a proline-rich region (PRR, residues 67-93), both of which are intrinsically disordered [53-56]. Residual dipolar couplings (RDCs) measured to characterize the local structure and dynamics of its conformational ensembles have shown that *isolated* p53(1-93) is intrinsically disordered, but with two regions of nascent secondary structure: a single helix turn populated at a level of ≈30% at residues 22-25 within the TAD (compared with the 2.5 turn helix that residues 18-26 form in complex with MDM2), and an elevated degree of stiffness in the PRR due to a raised population of polyproline II (PPII) conformations [40]. Furthermore, these conformational preferences were also observed in the context of the DNA•full-length p53 complex. A model generated of the DNA•full-length p53 complex suggests that the rigid PRR plays a structural role in projecting the TAD away from the surface of the protein, most likely to be positioned for interaction with coactivator targets. Similarly, the molecular shape of the full-length viral coactivator VP16 (residues 1-490) has been elucidated by Luisi and coworkers using small-angle x-ray scattering (SAXS), which revealed a globular structural core from which the essentially disordered C-terminal TAD extends (residues 412-490) [36]. NMR analysis by Han and coworkers of this *isolated* VP16 TAD indicates that although the polypeptide is completely devoid of globular structure in aqueous solution, it does contain four regions with transient structural order

in 5% of the conformers; in particular, residues 472-479 are able to form a relatively well-defined helix [57]. However, it is still unknown what the impact these pre-structured motifs have on the binding mechanism of VP16 with its target proteins, as well as if this feature is shared by all TADs of this class; these questions will be discussed further in Chapter 2 of this dissertation.

C.1.B.2. Target binding specificity

In general, amphipathic TADs are capable of interacting with and recruiting an overlapping subset of coactivator protein complexes to the promoters of the genes they are regulating [24, 48, 58-60]. This idea was first illustrated by research conducted within the model system of *Saccharomyces cerevisiae* using chromatin immunoprecipitation (ChIP) techniques, in which proteins that are *indirectly* associated with DNA were identified via formaldehyde crosslinking [19-20, 26, 61]. For example, Bryant and Ptashne used ChIP to measure the Gal4-dependent appearance of the SAGA histone-acetyltransferase (HAT) and Mediator complexes to the promoters of the coordinately induced *GAL* genes of *GAL1* and *GAL7* following galactose induction, while work by Lemieux and Gaudreau detected the recruitment of the Swi/Snf complex as well [62]. Similarly, ChIP analysis of the Gcn4-regulated *ARG1* gene upon induction of amino acid starvation revealed a greater recruitment of the SAGA, Mediator, and Swi/Snf complexes over that of the Mbf1 protein and the CCR4-NOT, Paf1, and RSC complexes which have been shown to be required at other Gcn4-dependent promoters [24, 60, 63-64].

Researchers have since been developing an array of chemical crosslinking techniques that can identify the *direct* TAD binding partners located within these protein complexes,

in order to characterize the binding interactions that are responsible for recruitment [65-68]. To achieve this, a nondiffusible crosslinking probe is placed in proximity to the TAD, which forms a covalent linkage nonspecifically to proteins separated by very short distances (on the order of angstroms) [24, 67, 69-77]. One such *in vitro* experiment performed by Hahn and coworkers on the TADs of Gal4 and Gcn4 in their DNA-bound forms identified their direct binding partners within preinitiation complexes (PICs) that had been assembled on a *HIS4* promoter containing one upstream Gcn4 binding site. Both Gal4 and Gcn4 TADs crosslinked with Tra1 (located within both the SAGA and NuA4 complexes), Taf12 (located within both the SAGA and TFIID complexes), and Med15(Gal11) (located within the tail module of the Mediator complex) [76-77]. Significantly, these types of studies illustrate that TADs of this class are not only capable of recruiting shared complexes, but do so by interacting with shared coactivator targets (Table 1.1). Furthermore, ongoing work by Mapp and coworkers entails using non-natural amino acid incorporation methods to perform these crosslinking studies in an *in vivo* setting. Using this method, they have successfully characterized the Gal4-Gal80 binding interface, and are currently optimizing methods to identify additional partners within the crosslinked adducts using mass spectrometry [78-79].

C.1.B.3. Potency

Historically, it has been found that the affinity by which a TAD interacts with transcriptional machinery targets correlates with its transcriptional output (i.e., its potency). For instance, this phenomenon was observed by Ptashne and coworkers when they analyzed deletion mutants of the TAD Gal4(840-881) to determine if there existed a

Table 1.1: Comparison of ***representative*** targets of Gal4, Gcn4, and VP16 within the *Saccharomyces cerevisiae* transcriptional machinery.

Complex	Protein	TAD		
		Gal4	Gcn4	VP16
GTFs	TBP	SPR [80] Pull-down [81]† Cross-linking (direct) [73]†	NA	Pull-down [81]† Cross-linking (indirect) [59]† Cross-linking (direct) [73]†
	TFIIB	SPR [80]	NA	Cross-linking (indirect) [59]†
TFIID/SAGA	Taf12(yTaf61/68)	Cross-linking (direct) [73, 77]†	Pull-down [82]† Cross-linking (direct) [76]†	Cross-linking (indirect) [59]† Cross-linking (direct) [73]†
	Taf9(yTaf17/20)	NA	Pull-down [83]†	Cross-linking (indirect) [59]†
	Taf6(yTaf60)	Cross-linking (direct) [73]†	Pull-down [83]†	Pull-down [83]† Cross-linking (indirect) [59]† Cross-linking (direct) [73]†
	Taf5(yTaf90)	NA	Pull-down [83]†	Pull-down [83]† Cross-linking (indirect) [59]†
SAGA	Tra1	Pull-down [84] Cross-linking (direct) [77]† FRET [85]	Pull-down [84] Cross-linking (direct) [76]†	Pull-down [84]
Swi/Snf	Swi1	SPR [25]	Cross-linking (direct) [24]†	SPR [25] Pull-down [24]
	Snf5	SPR [25]	Cross-linking (direct) [24]†	SPR [25] Pull-down [24]
	Swi2/Snf2	NA	Cross-linking (direct) [24]†	Pull-down [24]
Mediator	Cdk8(Srb10)	Cross-linking (direct) [86]†	Pull-down [86]	Pull-down [86]
	Med15(Gal11)	Pull-down [87]† Cross-linking (direct) [77]†	Pull-down [87]† Cross-linking (direct) [76]†	Pull-down [87]†

† Targets are located within a multi-protein complex.

correlation between their ability to upregulate expression of a reporter gene in yeast with their ability to bind target proteins *in vitro* [80]. The authors observed that the activators Gal4(DBD)-(840-854), (840-857), (840-869), and (840-881) elicit transcriptional activity approximately proportional to the length of the TAD, and that this trend is additionally reflected in their binding affinities ($K_{d,apparent}$) for the general transcription factors TBP and TFIIB. In addition, Sung and coworkers observed similar results with the p53 TAD when various hydrophobic amino acids (positionally conserved to those within the VP16 TAD that are critical for activity) were replaced with hydrophilic ones and investigated

for their ability to interact with TBP [88]. In order to determine whether this hypothesis holds upon comparing TADs derived from different activators, Melcher determined the binding affinities of TADs from both strong (Gal4, VP16, p53) and weak (AH, Gcn4, TAT) activators for TBP, in which he found that their activating potentials correlated with their ability to bind to this transcriptional machinery target [89].

More recent studies have provided evidence that in fact $K_{d,app}$ is not a uniformly good predictor of activator function. For instance, Mapp and coworkers isolated ligands from eight-residue synthetic peptide libraries (#28: AHYYYPSE and #17: AQRRVLSE) that were capable of binding to the coactivator Med15(Gal11) with binding affinities comparable to that of the extremely potent P201 peptide (YLLPTCIP) [90]. In addition, Schepartz and coworkers isolated a pair of ligands, PPKID4^{P} (4^{P}) and PPKID6^{U} (6^{U}), that possess similar binding affinities for the KIX domain of the coactivator CBP, yet 6^{U} is a less potent transcriptional activator than 4^{P} [91]. Rather, because transcription is initiated through a series of coupled binding equilibria between activators and their array of coactivator targets within the transcriptional machinery, we hypothesized that a kinetic analysis of the life-times of these interactions should be more revealing of activator potency. Even further, such an analysis could provide a mechanistic explanation for the ability of activators to interact differently with the same coactivator target (Chapter 2).

D. Activator artificial transcription factors (activator ATFs)

Considerable interest exists in the discovery of molecules that can mimic the function of a transcriptional activator for use as mechanistic tools and therapeutic agents [92]. In the following section, progress towards the development of non-natural DNA-binding

domain (DBD) and transcriptional activation domain (TAD) replacements for use in activator artificial transcription factors (activator ATFs) will be described, with a focus on optimizing fundamental properties including DNA binding specificity, stability from degradation, delivery into cells, and potency.

D.1.Artificial DNA-binding domains (DBDs)

Much work has been done to develop artificial DBDs that can bind with high specificity and affinity to predetermined DNA sequences [28]. One way in which this binding has been achieved is by mutating amino acids on natural protein scaffolds to recognize novel sequences [93-95]. The Cys_2His_2 zinc finger (ZF) fold has proven to be enormously versatile as a DNA-targeting entity, and is composed of ~30 amino acids folded into a ββα structure that is stabilized by hydrophobic interactions and the coordination of a zinc ion. In terms of sequence recognition, each ZF makes its primary sidechain–base interactions to three adjacent nucleotides in the sense strand of the DNA duplex by inserting its α helix into the major groove (Figure 1.5a) [29, 96]. One simple method for creating a ZF protein capable of binding to a predetermined DNA sequence is through the "modular assembly" approach, in which pre-existing, single finger 'modules' with known specificities are assembled into a multifinger array [96-97]. In practice, ZF proteins composed of 3–6 fingers with apparent dissociation constants in the picomolar to nanomolar range have been attached to proteinacious transcriptional activation domains and used successfully to upregulate endogenous genes in both mammalian cell culture and animal models [93, 97-98]. Despite this success, stable delivery of these proteins

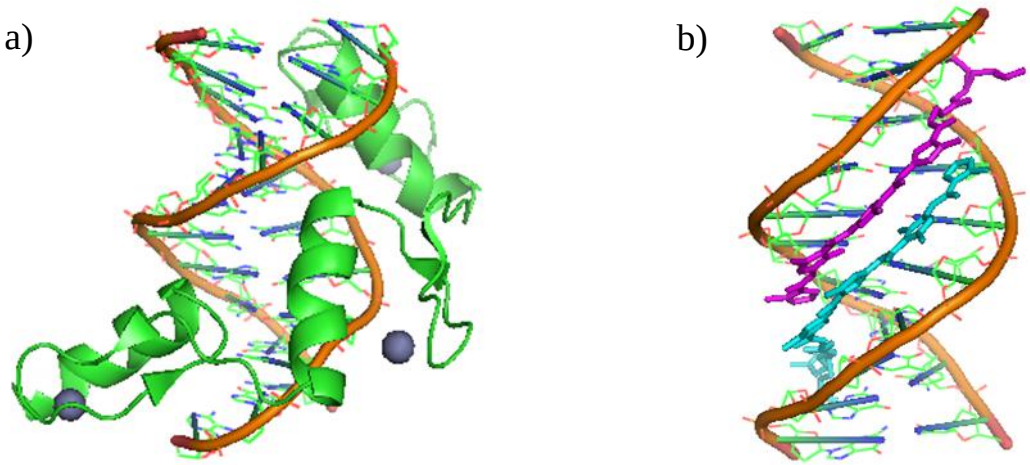


Figure 1.5: Artificial DNA-binding domains (artificial DBDs). a) Ribbon diagram of the crystal structure of three zinc fingers from Zif268 bound to the major groove of DNA (PDB ID: 1AAY). b) Crystal structure of polyamides bound as a dimer to the minor groove of DNA (PDB ID: 407D).

remains a challenge because they must be administered by viral vectors [99].

Another approach is using the hydrogen bonding properties of nucleic acid-like molecules as sequence-specific gene targeting agents. Triplex-forming oligonucleotides (TFOs), for example, bind to the major groove of DNA by forming Hoogsteen or reverse Hoogsteen hydrogen bonds with purine bases of Watson-Crick base pairs, thus utilizing a third strand binding code [100-101]. For instance, in purine TFOs Gs bind the G of a G•C pair and As bind the A of an A•T pair, while in pyrimidine TFOs protonated Cs bind the G of a G•C pair and Ts bind the A of an A•T pair [101]. Although triplex binding is restricted to polypurine stretches of DNA, recent bioinformatic studies have found that the largest relative concentrations of these sequences are found in regulatory regions (especially in promoter zones) [102]. In addition, modifications to the sugar-phosphate backbone and nucleoside bases can be made in order to improve their binding properties

and stability under physiological conditions; although limitations still exist on their efficient delivery into cells [103]. For example, attachment of a VP16-derived peptide to a TFO by Young and coworkers created a transcriptional activator that upregulated expression of a reporter gene up to 30-fold upon delivery into mammalian cell culture via liposome transfection techniques [104].

Finally, tailoring the DNA-binding properties of natural products is another attractive choice for the development of non-natural DBDs. In particular, considerable progress has occurred in developing programmable polyamides that can readily target specific DNA sequences [105]. Polyamides are inspired from the natural products distamycin and netropsin; they consist of pyrrole (Py), imidazole (Im), and hydroxypyrrole (Hp) amino acids linked through amide bonds that enable recognition of A•T, T•A, C•G, and G•C base pairs through the formation of specific hydrogen bonds with minor groove functionality (Figure 1.5b), although overall specificity varies with sequence context [95, 106-108]. Although several different polyamide structural motifs exist, the hairpin polyamide in which a flexible amino acid tether connects two polyamide arms is used most commonly [105, 107, 109]. Hairpin polyamides exhibit DNA binding affinities with dissociation constants in the picomolar to nanomolar range, and have been shown in several applications to traffic to the nucleus and to interact with their cognate DNA sites [110-112]. They have also been used as the basis for several different activator ATF constructs that function in cell-free and in cellular systems [113-115]. In contrast to protein DBDs, polyamides are synthesized on solid-phase and can be conjugated to either peptidic or nonpeptidic TADs; however, since activator ATFs that contain such non-protein DBDs typically use much smaller TADs to minimize the overall size of the

construct, this can impose limitations on their eventual potency [116-117]. In addition, in cellular systems these constructs often require special modifications to enhance permeability and typically target shorter DNA sequences (6-8 base pairs for an eight-ring hairpin) relative to proteins. Thus, an alternative approach to obtain cell-permeable molecules that can achieve the DNA-binding specificity equal to that of a protein would be an important contribution to the field (Chapter 3).

D.2.Artificial transcriptional activation domains (TADs)

The most common TADs used in the construction of activator ATFs are derived from the activation domains of natural proteins [4, 28]. Activating sequences from the viral protein VP16, the yeast activator Gal4, and the p65 subunit of the human activator NF-κB have all been attached to ZF proteins and function as activator ATFs in mammalian cell culture, while smaller peptides have even been used on non-proteinacious DBDs [100, 118-119]. For instance, a monomeric or dimeric repeat of eight residues of VP16 (VP1 and VP2) could upregulate a reporter gene *in vitro* using yeast nuclear extracts when attached to a polyamide, and a monomeric or dimeric repeat of an 11 residue sequence taken from VP16 (ATF14 and ATF29, respectively) could upregulate a reporter gene in mammalian cell culture when attached to a TFO [104, 113-114]. However, a screen in *Saccharomyces cerevisiae* of random 8-residue peptides attached to the DBD of Gal4 (residues 1-100) identified P201, a peptidic TAD that is as potent as full-length Gal4 [120]. This enhanced transcriptional activity of a small peptide was later attributed to a second binding interaction with the dimerization domain of Gal4 that protects P201 from

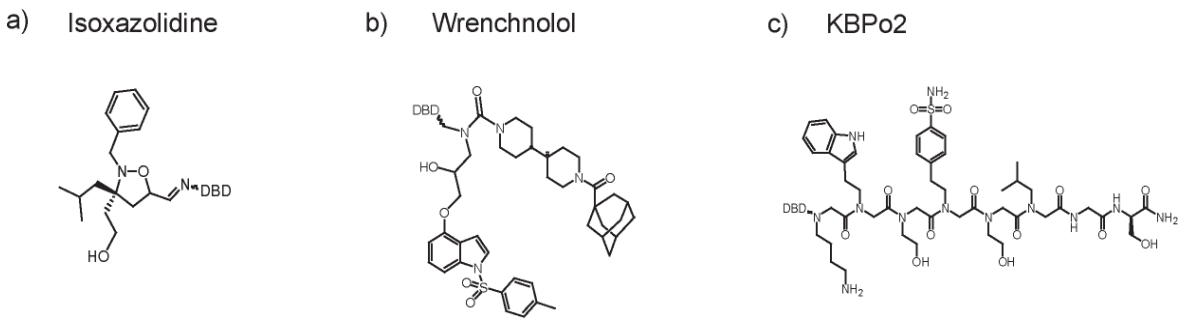


Figure 1.6: Small molecule/peptidomimetic-based artificial transcriptional activation domains (artificial TADs).

proteolytic degradation [121]. Therefore, small molecules and peptidomimetics that can mimic TAD function and bypass the degradation pathway of the cell are desirable alternatives.

Small molecules and peptidomimetics are often considerably smaller than protein-derived TADs and thus may exhibit advantageous cell permeability properties [92]. The first small molecule TAD to be reported was an amphipathic isoxazolidine (Figure 1.6a), and its ability to upregulate transcription was demonstrated initially in a cell-free assay where it proved to be as active as a larger peptidic activator (MW 290 vs 1674) [122]. This amphipathic isoxazolidine also functions in mammalian cell culture in a dose-dependent manner with up to 80-fold activation at 1 μM [123]. In addition, a hydrophobic molecule named wrenchnolol (Figure 1.6b) designed originally as an inhibitor of the ESX-Sur2 interaction is also capable of functioning as an activation domain upon localization to DNA [115]. When conjugated to a polyamide, this molecule upregulated transcription in a cell-free system 3.5-fold over background [124]. This synthetic activator ATF was, however, inactive in mammalian cell culture, which was possibly caused by limited nuclear localization. Finally, a peptoid (oligo-N-substituted

glycines) named KBPo2, which was identified from a library screen against the KIX domain of CBP (Figure 1.6c), is capable of activating robustly (up to 1000-fold) transcription of a reporter gene in mammalian cells in a 2-hybrid assay with an EC_{50} of 10μM [125]. Furthermore, a polyamide-peptoid conjugate is capable of activating transcription of 45 endogenous genes that contain multiple binding sites for the polyamide within their promoters by at least 3-fold in HeLa cells [117]. These examples illustrate the types of functional groups that are conducive to mediating interactions with components of the transcriptional machinery. Yet, it is still not known if these types of molecules are also able to utilize the same binding mechanism as natural TADs, be it simple colocalization or the involvement of a conformational change step, and what effect this might have on their potency (Chapter 2).

E. Thesis summary

Although to date it has been established that purely synthetic activator ATFs are indeed capable of perturbing the expression profiles of cellular systems, there is a need to increase activation potency and gene-targeting specificity before their full potential as therapeutic agents can be assessed [5, 126-127].

In order to examine the factors that influence natural TAD potency, a further understanding of the mechanism of activator•target complex formation is therefore required. To this end, Chapter 2 discusses the utilization of fluorescence stopped flow techniques to kinetically characterize the ability of three prototypical eukaryotic activators with differential potency to interact with the key coactivator Med15. These studies have revealed that the ability of TADs to undergo a conformational change to

form a final stable complex with its target, relative to dissociation, correlates with the extent to which they can upregulate transcription. This partition ratio appears to be dictated by the amount of inherent structural stability of the TAD within the intermediate complex, and is an important consideration for the development of future generations of nonnatural TAD mimics.

Furthermore, due to the difficulty of developing cell-permeable molecules that can bind to unique sites within the genome, another approach to localizing a transcriptional regulatory domain to a gene of interest would be to hijack the DNA-binding specificity of an endogenous protein. To this end, Chapter 3 focuses on the applicability of various screening/selection techniques of random peptide libraries to identify novel binding surfaces on DNA-binding proteins, with future implications to the discovery of their small molecule or peptidomimetic replacements.

F. References

1. Lee, T.I. and R.A. Young, *Transcription of eukaryotic protein-coding genes.* Annu Rev Genet, 2000. 34: p. 77-137.
2. Ptashne, M. and A. Gann, *Genes & Signals*. 2001, New York: Cold Spring Harbor Laboratory.
3. Ptashne, M., *Control of gene transcription: An outline.* Nature Medicine, 1997. 3(10): p. 1069-1072.
4. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. 2(1): p. 62-75.
5. Pandolfi, P.P., *Transcription therapy for cancer.* Oncogene, 2001. 20(24): p. 3116-27.
6. Chen, X., et al., *Gene expression patterns in human liver cancers.* Mol Biol Cell, 2002. 13(6): p. 1929-39.
7. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. 406(6797): p. 747-52.
8. Koleske, A.J. and R.A. Young, *The RNA polymerase II holoenzyme and its implications for gene regulation.* Trends Biochem Sci, 1995. 20(3): p. 113-6.
9. Kornberg, R.D., *Mediator and the mechanism of transcriptional activation.* Trends Biochem Sci, 2005. 30(5): p. 235-9.
10. Ptashne, M., *How Eukaryotic Transcriptional Activators Work.* Nature, 1988. 335(6192): p. 683-689.
11. Ptashne, M. and A. Gann, *Transcriptional activation by recruitment.* Nature, 1997. 386(6625): p. 569-577.
12. Kim, Y.J., et al., *A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II.* Cell, 1994. 77(4): p. 599-608.
13. Myers, L.C., et al., *Mediator protein mutations that selectively abolish activated transcription.* Proc Natl Acad Sci U S A, 1999. 96(1): p. 67-72.
14. Myers, L.C. and R.D. Kornberg, *Mediator of transcriptional regulation.* Annual Review of Biochemistry, 2000. 69: p. 729-749.
15. Bourbon, H.M., et al., *A unified nomenclature for protein subunits of Mediator complexes linking transcriptional regulators to RNA polymerase II.* Molecular Cell, 2004. 14(5): p. 553-557.
16. Chadick, J.Z. and F.J. Asturias, *Structure of eukaryotic Mediator complexes.* Trends Biochem Sci, 2005. 30(5): p. 264-71.
17. Fan, X., D.M. Chou, and K. Struhl, *Activator-specific recruitment of Mediator in vivo.* Nat Struct Mol Biol, 2006. 13(2): p. 117-20.
18. Asturias, F.J., et al., *Conserved structures of mediator and RNA polymerase II holoenzyme.* Science, 1999. 283(5404): p. 985-7.
19. Cosma, M.P., T. Tanaka, and K. Nasmyth, *Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter.* Cell, 1999. 97(3): p. 299-311.
20. Agalioti, T., et al., *Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter.* Cell, 2000. 103: p. 667-78.

21. Chan, H.M. and N.B. La Thangue, *p300/CBP proteins: HATs for transcriptional bridges and scaffolds.* J Cell Sci, 2001. 114(Pt 13): p. 2363-73.
22. Bhaumik, S.R. and M.R. Green, *Interaction of Gal4p with components of transcription machinery in vivo.* Methods Enzymol, 2003. 370: p. 445-54.
23. Ferreira, M.E., et al., *Activator-binding domains of the SWI/SNF chromatin remodeling complex characterized in vitro are required for its recruitment to promoters in vivo.* FEBS J, 2009. 276(9): p. 2557-65.
24. Neely, K.E., et al., *Transcription activator interactions with multiple SWI/SNF subunits.* Mol Cell Biol, 2002. 22(6): p. 1615-25.
25. Ferreira, M.E., et al., *Mechanism of transcription factor recruitment by acidic activators.* J Biol Chem, 2005. 280(23): p. 21779-84.
26. Biddick, R. and E.T. Young, *The disorderly study of ordered recruitment.* Yeast, 2009. 26(4): p. 205-20.
27. Brent, R. and M. Ptashne, *A Eukaryotic Transcriptional Activator Bearing the DNA Specificity of a Prokaryotic Repressor.* Cell, 1985. 43(3): p. 729-736.
28. Ansari, A.Z. and A.K. Mapp, *Modular design of artificial transcription factors.* Curr Opin Chem Biol, 2002. 6(6): p. 765-72.
29. Garvie, C.W. and C. Wolberger, *Recognition of specific DNA sequences.* Mol Cell, 2001. 8(5): p. 937-46.
30. Keller, W., P. Konig, and T.J. Richmond, *Crystal structure of a bZIP/DNA complex at 2.2 A: determinants of DNA specific recognition.* J Mol Biol, 1995. 254(4): p. 657-67.
31. Luscombe, N.M., et al., *An overview of the structures of protein-DNA complexes.* Genome Biol, 2000. 1(1): p. REVIEWS001.
32. Ellenberger, T.E., et al., *The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex.* Cell, 1992. 71(7): p. 1223-37.
33. Marmorstein, R., et al., *DNA Recognition by Gal4 - Structure of a Protein DNA Complex.* Nature, 1992. 356(6368): p. 408-414.
34. Hong, M., et al., *Structural basis for dimerization in DNA recognition by Gal4.* Structure, 2008. 16(7): p. 1019-26.
35. Wysocka, J. and W. Herr, *The herpes simplex virus VP16-induced complex: the makings of a regulatory switch.* Trends Biochem Sci, 2003. 28(6): p. 294-304.
36. Grossmann, J.G., et al., *Molecular shapes of transcription factors TFIIB and VP16 in solution: implications for recognition.* Biochemistry, 2001. 40(21): p. 6267-74.
37. Dyson, H.J. and P.E. Wright, *Coupling of folding and binding for unstructured proteins.* Curr Opin Struct Biol, 2002. 12(1): p. 54-60.
38. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. 6(3): p. 197-208.
39. Sugase, K., H.J. Dyson, and P.E. Wright, *Mechanism of coupled folding and binding of an intrinsically disordered protein.* Nature, 2007. 447(7147): p. 1021-5.
40. Wells, M., et al., *Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain.* Proc Natl Acad Sci U S A, 2008. 105(15): p. 5762-7.

41. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Curr Opin Struct Biol, 2009. 19(1): p. 31-8.
42. Leuther, K.K., J.M. Salmeron, and S.A. Johnston, *Genetic evidence that an activation domain of GAL4 does not require acidity and may form a beta sheet.* Cell, 1993. 72(4): p. 575-85.
43. Van Hoy, M., et al., *The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets.* Cell, 1993. 72(4): p. 587-94.
44. Schmitz, M.L., et al., *Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation.* J Biol Chem, 1994. 269(41): p. 25613-20.
45. Dahlman-Wright, K., et al., *Structural characterization of a minimal functional transactivation domain from the human glucocorticoid receptor.* Proc Natl Acad Sci U S A, 1995. 92(5): p. 1699-703.
46. Shen, F., et al., *Critical amino acids in the transcriptional activation domain of the herpesvirus protein VP16 are solvent-exposed in highly mobile protein segments. An intrinsic fluorescence study.* J Biol Chem, 1996. 271(9): p. 4819-26.
47. Warnmark, A., et al., *The N-terminal regions of estrogen receptor alpha and beta are unstructured in vitro and show different TBP binding properties.* J Biol Chem, 2001. 276(49): p. 45939-44.
48. Jonker, H.R., et al., *Structural properties of the promiscuous VP16 activation domain.* Biochemistry, 2005. 44(3): p. 827-39.
49. Cress, W.D. and S.J. Triezenberg, *Critical Structural Elements of the Vp16 Transcriptional Activation Domain.* Science, 1991. 251(4989): p. 87-90.
50. Thoden, J.B., et al., *Understanding a transcriptional paradigm at the molecular level. The structure of yeast Gal80p.* J Biol Chem, 2007. 282(3): p. 1534-8.
51. Thoden, J.B., et al., *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. 283(44): p. 30266-72.
52. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. 274(5289): p. 948-53.
53. Fields, S. and S.K. Jang, *Presence of a potent transcription activating sequence in the p53 protein.* Science, 1990. 249(4972): p. 1046-9.
54. Lin, J., et al., *Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein.* Genes Dev, 1994. 8(10): p. 1235-46.
55. Dawson, R., et al., *The N-terminal domain of p53 is natively unfolded.* J Mol Biol, 2003. 332(5): p. 1131-41.
56. Teufel, D.P., et al., *Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53.* Proc Natl Acad Sci U S A, 2007. 104(17): p. 7009-14.
57. Kim, D.H., et al., *Multiple hTAF(II)31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16.* BMB Rep, 2009. 42(7): p. 411-7.
58. Sheppard, K.A., et al., *Transcriptional activation by NF-kappaB requires multiple coactivators.* Mol Cell Biol, 1999. 19(9): p. 6367-78.

59. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo.* J Biol Chem, 2002. 277(48): p. 46043-50.
60. Govind, C.K., et al., *Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo.* Mol Cell Biol, 2005. 25(13): p. 5626-38.
61. Cosma, M.P., *Ordered recruitment: gene-specific mechanism of transcription activation.* Mol Cell, 2002. 10(2): p. 227-36.
62. Bryant, G.O. and M. Ptashne, *Independent recruitment in vivo by Gal4 of two complexes required for transcription.* Mol Cell, 2003. 11(5): p. 1301-9.
63. Neely, K.E., et al., *Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays.* Mol Cell, 1999. 4(4): p. 649-55.
64. Swanson, M.J., et al., *A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo.* Mol Cell Biol, 2003. 23(8): p. 2800-20.
65. Trakselis, M.A., S.C. Alley, and F.T. Ishmael, *Identification and mapping of protein-protein interactions by a combination of cross-linking, cleavage, and proteomics.* Bioconjug Chem, 2005. 16(4): p. 741-50.
66. Sadakane, Y. and Y. Hatanaka, *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): p. 209-18.
67. Melcher, K. and H.T. Chen, *Identification and analysis of multiprotein complexes through chemical crosslinking.* Curr Protoc Cell Biol, 2007. Chapter 17: p. Unit 17 10.
68. Tanaka, Y., M.R. Bond, and J.J. Kohler, *Photocrosslinkers illuminate interactions in living cells.* Mol Biosyst, 2008. 4(6): p. 473-80.
69. Ansari, A.Z., R.J. Reece, and M. Ptashne, *A transcriptional activating region with two contrasting modes of protein interaction.* Proc Natl Acad Sci U S A, 1998. 95(23): p. 13543-8.
70. Koh, S.S., et al., *An activator target in the RNA polymerase II holoenzyme.* Mol Cell, 1998. 1(6): p. 895-904.
71. Jackson, V., *Formaldehyde cross-linking for studying nucleosomal dynamics.* Methods, 1999. 17(2): p. 125-39.
72. Bhaumik, S.R. and M.R. Green, *SAGA is an essential in vivo target of the yeast acidic activator Gal4p.* Genes Dev, 2001. 15(15): p. 1935-45.
73. Klein, J., et al., *Use of a genetically introduced cross-linker to identify interaction sites of acidic activators within native transcription factor IID and SAGA.* J Biol Chem, 2003. 278(9): p. 6779-86.
74. Herrera, F.J. and S.J. Triezenberg, *VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection.* J Virol, 2004. 78(18): p. 9689-96.
75. Archer, C.T., L. Burdine, and T. Kodadek, *Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking.* Mol Biosyst, 2005. 1(5-6): p. 366-72.

76. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. 18(3): p. 369-78.
77. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. 25(20): p. 9092-102.
78. Majmudar, C.Y., et al., *Impact of nonnatural amino acid mutagenesis on the in vivo function and binding modes of a transcriptional activator.* J Am Chem Soc, 2009. 131(40): p. 14240-2.
79. Majmudar, C.Y., et al., *A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry.* Angew Chem Int Ed Engl, 2009. 48(38): p. 7021-4.
80. Wu, Y., R.J. Reece, and M. Ptashne, *Quantitation of putative activator-target affinities predicts transcriptional activating potentials.* EMBO J, 1996. 15(15): p. 3951-63.
81. Melcher, K. and S.A. Johnston, *GAL4 interacts with TATA-binding protein and coactivators.* Mol Cell Biol, 1995. 15(5): p. 2839-48.
82. Natarajan, K., et al., *yTAFII61 has a general role in RNA polymerase II transcription and is required by Gcn4p to recruit the SAGA coactivator complex.* Mol Cell, 1998. 2(5): p. 683-92.
83. Drysdale, C.M., et al., *The Gcn4p activation domain interacts specifically in vitro with RNA polymerase II holoenzyme, TFIID, and the Adap-Gcn5p coactivator complex.* Mol Cell Biol, 1998. 18(3): p. 1711-24.
84. Brown, C.E., et al., *Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit.* Science, 2001. 292(5525): p. 2333-7.
85. Bhaumik, S.R., et al., *In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer.* Genes Dev, 2004. 18(3): p. 333-43.
86. Ansari, A.Z., et al., *Transcriptional activating regions target a cyclin-dependent kinase.* Proc Natl Acad Sci U S A, 2002. 99(23): p. 14706-9.
87. Park, J.M., et al., *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. 20(23): p. 8709-19.
88. Chang, J., et al., *Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein.* J Biol Chem, 1995. 270(42): p. 25014-9.
89. Melcher, K., *The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins.* J Mol Biol, 2000. 301(5): p. 1097-112.
90. Wu, Z., et al., *Targeting the transcriptional machinery with unique artificial transcriptional activators.* J Am Chem Soc, 2003. 125(41): p. 12390-1.
91. Volkman, H.M., S.E. Rutledge, and A. Schepartz, *Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain.* J Am Chem Soc, 2005. 127(13): p. 4649-58.
92. Majmudar, C.Y. and A.K. Mapp, *Chemical approaches to transcriptional regulation.* Curr Opin Chem Biol, 2005. 9(5): p. 467-74.
93. Blancafort, P., D.J. Segal, and C.F. Barbas, 3rd, *Designing transcription factor architectures for drug discovery.* Mol Pharmacol, 2004. 66(6): p. 1361-71.

94. Falke, D. and R.L. Juliano, *Selective gene regulation with designed transcription factors: implications for therapy.* Curr Opin Mol Ther, 2003. 5(2): p. 161-6.
95. Uil, T.G., H.J. Haisma, and M.G. Rots, *Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities.* Nucleic Acids Res, 2003. 31(21): p. 6064-78.
96. Pabo, C.O., E. Peisach, and R.A. Grant, *Design and selection of novel Cys(2)His(2) zinc finger proteins.* Annual Review of Biochemistry, 2001. 70: p. 313-340.
97. Beerli, R.R. and C.F. Barbas, *Engineering polydactyl zinc-finger transcription factors.* Nature Biotechnology, 2002. 20(2): p. 135-141.
98. Ansari, A.Z., *Fingers reach for the genome.* Nat Biotechnol, 2003. 21(3): p. 242-3.
99. Sinn, P.L., S.L. Sauter, and P.B. McCray, Jr., *Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production.* Gene Ther, 2005. 12(14): p. 1089-98.
100. Kuznetsova, S., et al., *Gene activation by triplex-forming oligonucleotide coupled to the activating domain of protein VP16.* Nucleic Acids Research, 1999. 27(20): p. 3995-4000.
101. Faria, M. and C. Giovannangeli, *Triplex-forming molecules: from concepts to applications.* Journal of Gene Medicine, 2001. 3(4): p. 299-310.
102. Goni, J.R., X. de la Cruz, and M. Orozco, *Triplex-forming oligonucleotide target sequences in the human genome.* Nucleic Acids Res, 2004. 32(1): p. 354-60.
103. Knauert, M.P. and P.M. Glazer, *Triplex forming oligonucleotides: sequence-specific tools for gene targeting.* Hum Mol Genet, 2001. 10(20): p. 2243-51.
104. Stanojevic, D. and R.A. Young, *A highly potent artificial transcription factor.* Biochemistry, 2002. 41(23): p. 7209-7216.
105. Gottesfeld, J.M., et al., *Regulation of gene expression by small molecules.* Nature, 1997. 387(6629): p. 202-205.
106. O'Hare, C.C., et al., *DNA sequence recognition in the minor groove by crosslinked polyamides: The effect of N-terminal head group and linker length on binding affinity and specificity.* Proc Natl Acad Sci USA, 2002. 99(1): p. 72-77.
107. Dervan, P.B. and B.S. Edelson, *Recognition of the DNA minor groove by pyrrole-imidazole polyamides.* Current Opinion in Structural Biology, 2003. 13(3): p. 284-299.
108. Nickols, N.G., et al., *Improved nuclear localization of DNA-binding polyamides.* Nucleic Acids Res, 2007. 35(2): p. 363-70.
109. Dervan, P.B., *Molecular recognition of DNA by small molecules.* Bioorganic & Medicinal Chemistry, 2001. 9(9): p. 2215-2235.
110. Wurtz, N.R., et al., *Inhibition of DNA binding by NF-kappa B with pyrrole-imidazole polyamides.* Biochemistry, 2002. 41(24): p. 7604-9.
111. Dudouet, B., et al., *Accessibility of nuclear chromatin by DNA binding polyamides.* Chemistry & Biology, 2003. 10(9): p. 859-867.
112. Nickols, N.G., et al., *Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface.* ACS Chem Biol, 2007. 2(8): p. 561-71.

113. Mapp, A.K., et al., *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): p. 3930-3935.
114. Ansari, A.Z., et al., *Towards a minimal motif for artificial transcriptional activators.* Chemistry & Biology, 2001. 8(6): p. 583-592.
115. Shimogawa, H., et al., *A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction.* J Am Chem Soc, 2004. 126(11): p. 3461-71.
116. Liu, B. and T. Kodadek, *Investigation of the relative cellular permeability of DNA-binding pyrrole-imidazole polyamides.* J Med Chem, 2009. 52(15): p. 4604-12.
117. Xiao, X., et al., *A cell-permeable synthetic transcription factor mimic.* Angew Chem Int Ed Engl, 2007. 46(16): p. 2865-8.
118. Tanaka, M. and W. Herr, *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): p. 6056-6067.
119. Sadowski, I., et al., *Gal4-Vp16 Is an Unusually Potent Transcriptional Activator.* Nature, 1988. 335(6190): p. 563-564.
120. Lu, X.Y., A.Z. Ansari, and M. Ptashne, *An artificial transcriptional activating region with unusual properties.* Proceedings of the National Academy of Sciences of the United States of America, 2000. 97(5): p. 1988-1992.
121. Lu, Z., et al., *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): p. 8591-8596.
122. Minter, A.R., B.B. Brennan, and A.K. Mapp, *A small molecule transcriptional activation domain.* Journal of the American Chemical Society, 2004. 126(34): p. 10504-10505.
123. Rowe, S.P., et al., *Transcriptional up-regulation in cells mediated by a small molecule.* J Am Chem Soc, 2007. 129(35): p. 10654-5.
124. Kwon, Y., et al., *Small molecule transcription factor mimic.* J Am Chem Soc, 2004. 126(49): p. 15940-1.
125. Liu, B., et al., *A potent transactivation domain mimic with activity in living cells.* J Am Chem Soc, 2005. 127(23): p. 8254-8255.
126. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. 2(10): p. 740-9.
127. Toniatti, C., et al., *Gene therapy progress and prospects: transcription regulatory systems.* Gene Ther, 2004. 11(8): p. 649-57.

CHAPTER 2

TRANSIENT KINETIC ANALYSIS OF TRANSCRIPTIONAL ACTIVATOR•DNA COMPLEXES INTERACTING WITH THE KEY COACTIVATOR MED15[1]

A. Introduction

Transcription is initiated through a series of coupled binding equilibria between transcriptional activators and their array of protein targets within the transcriptional machinery [2]. However, previous efforts to kinetically characterize these interactions have produced conflicting models for the mechanism of complex formation [3-5], which is hampering the discovery of non-natural mimics of TAD function for use in activator ATFs. To address this issue, we employed fluorescence stopped-flow techniques to obtain the minimal mechanism by which three different activators are capable of interacting with the same coactivator, Med15.

We further hypothesized that the life-times of these interactions should be more revealing of activator potency. Thus, we analyzed the microscopic rate and equilibrium constants defining the individual steps within our mechanism, in order to identify key trends that can differentiate the activators from one another in terms of their ability to recruit the transcriptional machinery to a gene promoter (Figure 2.1).

[1] This project was a collaborative effort with Dr. Carol Fierke and Dr. John Hsieh, for which a manuscript has been submitted for publication.

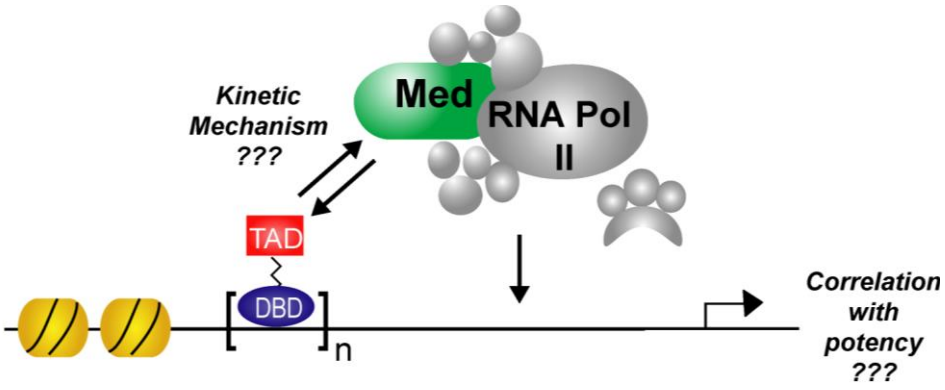


Figure 2.1: Incomplete mechanism of transcriptional activation. Schematic of a transcriptional activator localized upstream of a gene promoter through a DNA-binding domain (DBD) (blue circle) which is poised for recruitment of the transcriptional machinery through interaction with the transcriptional activation domain (TAD) (red square). Analysis of the life-times of TAD-target interactions should be more revealing of activator potency (i.e., transcriptional output).

B. Background: TAD•Target complex formation

The largest class of activators, the amphipathic class, is characterized by interspersed polar and hydrophobic amino acid residues within their TADs despite little sequence homology within the class (Figure 2.2a). TADs of this class are *largely* devoid of secondary structural elements in aqueous solution [6-8], and are classified as intrinsically disordered [9-11]. However, the TAD segments involved in the binding interaction with a target protein are found *predominantly* as amphipathic helices within the complex [12-14]. In particular, it is the hydrophobic face of the amphipathic helix that participates in the majority of the contacts, which have also been shown through mutational analysis to be important for transcriptional activity [15]. For example, three hydrophobic residues (Phe^{475}, Met^{478}, and Phe^{479}) within the 9-residue helix (Asp^{472}-Thr^{480}) of the $VP16_{456-490}$ TAD make crucial contacts with Tfb1/TFIIH (Figure 2.2b) [16].

a)

TAD	Sequence
Gal4	^{840}WTDQTAYNAFGITTGMFNTTTMDDVYNYLFDDEDTPPNPKKE^{881}
Gcn4	^{107}MFEYENLEDNSKEWTSLFDNDIPVTTDDVSLADKAIES^{144}
VP16	^{456}GFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG^{490}
p53	^{1}MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAM^{40}
KID	^{116}DSVTDSQKRREILSRRPSYRKILNDLSSDAPG^{147}

The side chains of the amino acids shown in green are hydrophobic; those in red are polar.

b)

Figure 2.2: Natural TADs. (a) Sequences from the TADs of Gal4, Gcn4, and VP16 used in this study for fluorescence stopped-flow experiments and for transcriptional activation assays in yeast. Sequences from the TADs of p53 and KID have been previously characterized by NMR methods. (b) Ribbon diagram of the lowest-energy structure of the TAD of $VP16_{456\text{-}490}$ (orange) bound to $Tfb1_{1\text{-}115}$ (cyan) (PDB ID: 2K2U). Residues 472-480 of VP16 form an α-helix.

Due to the inherent structural flexibility of the TADs, they are composed of a vast ensemble of conformational isomers that are in dynamic equilibrium, with the population time of each conformer following statistical thermodynamic distributions dictated by the heights of the energy barriers separating them [17]. Thus, the mechanism of molecular recognition that occurs during the process of TAD•target complex formation has been *proposed* to occur via three potential models (Figure 2.3) [18]. In the first model, that of conformational selection, the TAD conformers responsible for molecular recognition are not necessarily present at the highest populations in solution, but rather are selected for binding by the transcriptional machinery target because they more closely resemble the conformer present in the bound state [19]. Furthermore, as these binding conformers are depleted from the solution, there is a shift in equilibrium toward these conformers that subsequently propagates the binding event [17, 19]. In the second model, that of induced-fit, folding of the intrinsically disordered TAD is induced by association with the target protein (a.k.a., coupled binding and folding) [18, 20]. According to this model, an ensemble of unstructured TAD conformers bind to the target and evolve into a final high-affinity complex as the TAD forms intermolecular interactions with the binding site residues. In the third model, that of conformational adaption, a combination of conformational selection and induced-fit occurs. In the remainder of this section, results from previous studies using nuclear magnetic resonance (NMR) and surface plasmon resonance (SPR) spectroscopy to characterize TAD-target interactions will be presented in support of these potential models. In addition, the limitations of using SPR to perform such a transient kinetic analysis will be discussed, which we hypothesize can be addressed upon implementation of fluorescence stopped-flow techniques.

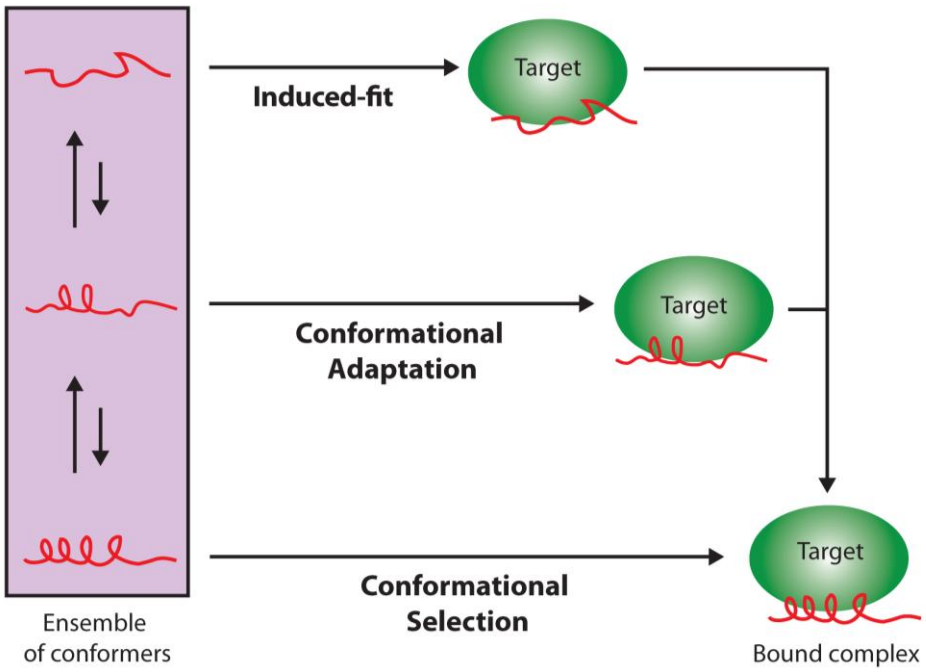


Figure 2.3: Schematic representation of potential two-step binding models for TAD•target complex formation.

B.1. Nuclear magnetic resonance (NMR) spectroscopic analysis of TAD structure

NMR spectroscopy is commonly used to obtain structures of the low-energy (or ground-state) conformers of a protein that are present at equilibrium using experimentally generated distance, angle, and orientation restraints [21-22]; however, it is the high-energy (or excited-state) conformers, which are virtually invisible to conventional biophysical tools due to their sparse population, that are hypothesized to provide a more complete understanding of protein function (e.g., molecular recognition) [23]. Nevertheless, because the kinetic barriers separating the ground and excited-state conformers typically results in molecular rearrangements that occur on the micro- to millisecond timescale or longer, detailed information on the exchange event between

these two states can be gathered through the use of Carr-Purcell-Meiboom-Gill relaxation dispersion (CPMG RD) NMR methods [23-24].

Briefly, CPMG RD NMR methods take advantage of the fact that a single NMR spin (probe) evolves at different frequencies (the chemical shift) when in states A and B. If the spin exists in a single state for time *T* and a 'refocusing' pulse is applied exactly in the center of this interval, then the overall frequency evolution of the spin for the total *T* period is reduced to zero. By contrast, if the pulse is applied 'off center' such that the probe is not in a single state during *T*, then refocusing is not complete, so that by the end of the total time period the spin will have evolved at an effective frequency that depends on the duration of the probe in each state and the number of jumps that were executed. If the rate of application of the pulses increases, the effects of the jumps are smaller in the sense that refocusing is improved. By measuring peak intensities as a function of the number of pulses (CPMG frequency (ν_{CPMG})), expressed as a relaxation rate, R_2^{eff} (s^{-1}), a 'relaxation-dispersion' curve is obtained (Figure 2.4). From these data, values for the chemical shift of the excited state, as well as the rate constants of exchange between states A and B (k_{AB} and k_{BA}, and thus the populations of interconverting conformers), can be obtained.

To this end, these NMR methods have recently facilitated the detection of preformed secondary structure elements within the excited-state conformers of isolated TADs. For example, Han and coworkers examined the unbound full-length TAD of VP16 (residues 412-490) and observed four short regions with a propensity for α-helicity in 5% of the conformers, the most stable of which occurs at residues 472-479 [25]; this region encompasses the $hTAF_{II}31$-binding motif (residues 472-483) identified by Verdine and

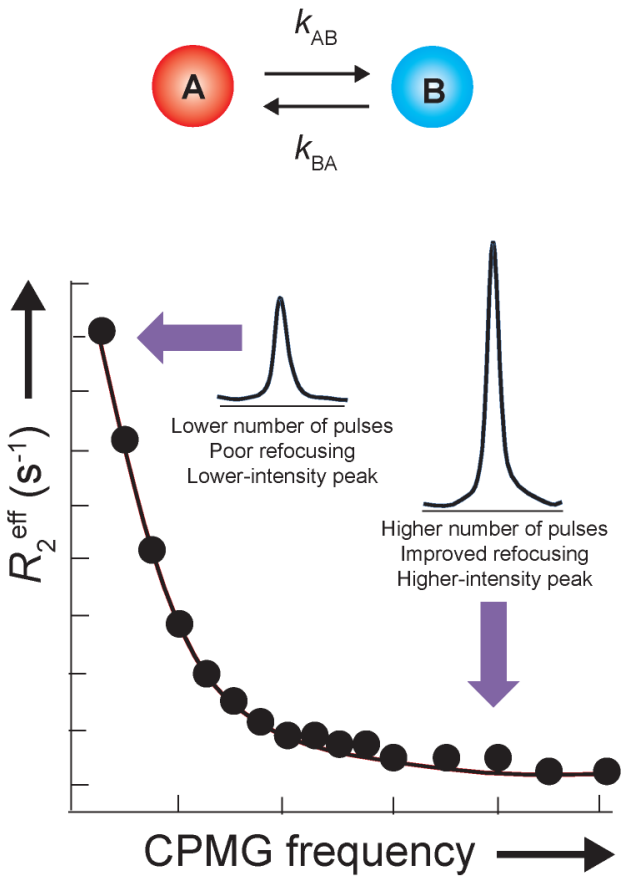


Figure 2.4: Relaxation dispersion NMR spectroscopy. A single NMR spin (probe) interconverts between two conformational states, A and B, as a function of time with its own trajectory (i.e., with regards to the time it spends in each state and the number of times it jumps between states). During the trajectory, pulses are applied that lead to a modulation of the relaxation rate of the probe. By measuring the peak intensity as a function of the number of pulses (CPMG frequency), expressed as a relaxation rate, R_2^{eff}, a relaxation-dispersion curve is obtained, from which details of the exchange reaction can be obtained.

coworkers [12], as well as the Tfb1/TFIIH-binding motif (residues 472-480) identified by Omichinski and coworkers [16]. Similarly, Fersht and coworkers were able to analyze the unbound full-length TAD of p53 (residues 1-93) to identify a 4-amino acid polypeptide region of helical structure populated to a level of ~30% for residues 22-25 [8], which is encompassed by the region found fully helical when bound to MDM2 (residues 18-25)

[14]. Yet, the primary role that these low-populated conformers with transiently formed structures serve during TAD•target complex formation has yet to be determined. For instance, if proceeding via the conformational selection model, they would be the conformers which are solely responsible for complex formation; alternatively, their presence may simply illustrate which polypeptide regions are most primed for induced folding upon interaction with a binding partner [26]. In support of the latter, NMR studies by Wright and coworkers on the TAD of KID (residues 116-147) demonstrated that it binds to the KIX domain of the coactivator CBP via an induced-fit mechanism, in which an ensemble of non-specifically associated encounter complexes evolve to the final bound state via a low-affinity intermediate (in which the α_B region (residues 133 - 138 and 141) of KID possesses up to 70% of their helical character) [5, 27]. However, the actual molecular recognition pathway most likely lies between these two extremes; due to the fact that the short segments of pre-structured motifs observed within the isolated TADs do not encompass the whole region found structured in the final complexes, these conformers probably serve as primary sites of contact which undergo further rearrangements to optimize their intermolecular interactions within the binding site [19, 26]. Overall, these structural studies have provided valuable information about the dynamic equilibrium that exists within the TADs.

B.2. Surface plasmon resonance (SPR) spectroscopic analysis of TAD-Target interactions

In principle, the time-course of a reaction from the moment when the species are mixed (t = 0) until equilibrium is established (and its concentration dependence) contains

all the information for the evaluation of the individual steps along the reaction pathway, including the formation/decay of intermediates [28]. Thus, transient state kinetic methods provide a means by which to elucidate the binding mechanism of protein complex assembly. To this end, surface plasmon resonance (SPR) spectroscopy coupled with a flow system creates a label-free method with which to analyze the interactions of biomolecules in real-time [29-30].

In a typical SPR experiment, a ligand is first immobilized onto the surface of a sensor chip, whereby an analyte is then passed over the chip in a continuous flow (Figure 2.5). Interaction of the analyte with the ligand increases the molecular mass near the chip surface, causing a change in the refractive index that can be monitored over time in order to obtain information about the association rate of the complex. Upon reaching equilibrium, sample buffer is then injected over the surface in order to obtain information about the dissociation rate of the complex. The experiment is subsequently repeated at various analyte concentrations, and the resulting sensograms (commonly described by either a single or double exponential) can then be analyzed by global analysis. Experimental conditions required to obtain *accurate* kinetic information are different from those used for affinity determination [30]. For instance, high flow rate of analyte (>30 μL/min) and low surface capacity of ligand (<100 RU) can minimize the underestimation of the association rate (k_{on}) caused by limitations in transferring the analyte from bulk solution to the chip surface (mass transport effects). These conditions can also minimize re-binding of the analyte during the dissociation phase, which would lead to an underestimation of k_{off}.

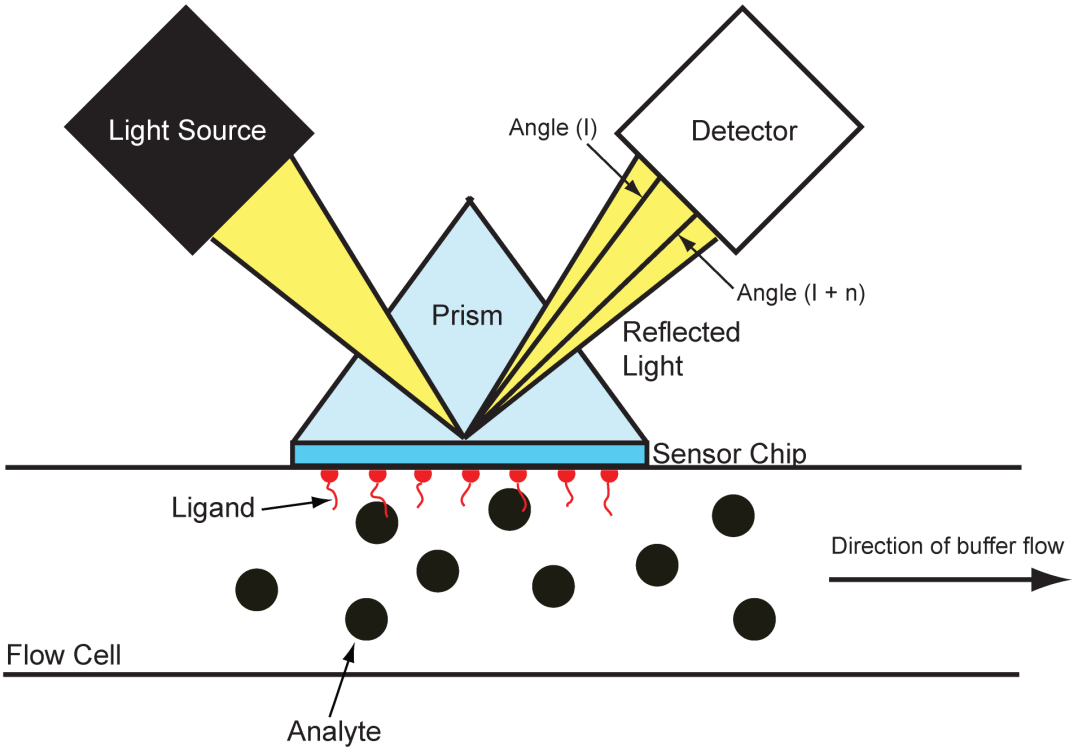


Figure 2.5: Surface plasmon resonance (SPR). SPR occurs when light is reflected from a conducting film at the interface between two media with different refractive indices. A change in the intensity of reflected light is dependent on the refractive index close to chip surface. When the refractive index changes due to protein binding, it changes the resonance signal monitored by the instrument and can be used to extract thermodynamic and kinetic information.

To date, analysis of TAD-target interactions using SPR has not yielded a unified binding mechanism. For instance, the simplest possible mechanism by which activators could interact with targets within the transcriptional machinery is by a single bimolecular collision step. Indeed, in support of this model, measurement of the kinetic rate constants for the association and dissociation of TAD-target interactions using SPR has resulted in sensograms that fit well to a one-step (Langmuir) binding isotherm ($A + B \leftrightarrow AB$), which predicts that both the association and dissociation phases are described by a single exponential term [31]. For example, Schmitz and coworkers reported one-step binding

when the NF-κB p65 TAD (residues 471-551) attached to the Gal4 DBD was immobilized to a sensor chip using amine coupling and the general transcription factors TBP, TFIIB, as well as the viral coactivator E1A 13S were injected over the surface [32]. In addition, Ptashne and coworkers also reported one-step binding when Gal4(1-100)-Gal4(TAD) constructs pre-bound to an immobilized DNA duplex containing two Gal4 binding sites interacted with either TBP and TFIIB (which can occur simultaneously) or its masking protein Gal80 (which blocks the interaction of TBP and TFIIB) [33]. One interpretation of this data supports a recruitment model in which assembly of the pre-initiation complex *in vivo* proceeds via the simultaneous association of multiple promoter-bound activators to multiple binding partners within the transcriptional machinery, and therefore a conformational change in the unstructured TAD is not required for stable complex formation.

Alternatively, if the TAD exits in an equilibrium between populations of "*totally* unstructured" low affinity conformations and "*mostly* unstructured" high-affinity conformations (i.e., conformations with regions of transient pre-formed secondary structure), at low enough target (analyte) concentration only the high-affinity conformation of the TAD (ligand) will bind using SPR and the sensograms will be monophasic; this model is referred to as heterogenous ligand binding [30, 34-35]. For example, Han and coworkers examined the binding of Helix 2 (H2) of the VP16 TAD (residues 469-485) to immobilized $hTAF_{II}31$ using SPR and analyzed the data according to a one-step binding isotherm [25]. Rather than binding via an induced-fit "random coil → α helix" transition of the VP16 TAD *after* binding, as hypothesized previously by Verdine and coworkers [12], the authors argue for binding via conformational selection

Table 2.1: Summary of the analysis of TAD-target interactions via SPR.

Activator	Target	Binding model	Reference
Gal4(DBD)-NF-κB p65	TBP TFIIB E1A 13S	1:1 Langmuir	[32]
Gal4(DBD)-Gal4	TBP TFIIB Gal80	1:1 Langmuir	[33]
TAD	**Target**	**Binding model**	**Reference**
VP16	$hTAF_{II}31$	1:1 Langmuir	[25]
	TBP Swi1 Snf5	Rapid equilibrium followed by a slow conversion step	[4]
Gal4	TBP Swi1 Snf5	Rapid equilibrium followed by a slow conversion step	[4]
c-*Myc*	TBP	Rapid equilibrium followed by a slow conversion step	[36]
AR-AF1	RAP74/TFIIF SRC-1a	1:1 Langmuir	[37]

of a "pre-formed helix" (residues 472-479). Similarly, SPR sensograms for the binding of immobilized AR-AF1 (residues 142-485) to both RAP74/TFIIF and the coactivator SRC-1a were analyzed according to a one-step binding mechanism [37], possibly due to the presence of regions with secondary structure [38].

Finally, there is also evidence that activators could interact with transcriptional machinery targets through a more complex, two-step mechanism involving a conformational change step, as evidenced by the observation of biphasic sensograms by SPR [30, 34]; this is consistent with a model in which the TAD undergoes a conformational change *after* binding to the target, regardless of the amount of initial secondary structure it possesses. Indeed, Wright and coworkers observed biphasic sensograms between the interaction of the *c-Myc* TAD (residues 1-143) and TBP, which involved a fast (rapid-equilibrium) phase that is less favored at elevated salt

concentrations, as well as a slow phase that is more affected by temperature, thus supporting a model in which a rapidly forming complex based on electrostatic interactions converts slowly through hydrophobic interactions to a more stable form [36]. They observed analogous results for the binding of the additional amphipathic TADs of Gal4 (residues 769-881) and VP16 (residues 413-490) to TBP, as well as the targets Swi1 (residues 329-547) and Snf5 (residues 1-334), while the control construct VP16Δ456 (devoid of the 472-479 helix previous shown to be important for binding to multiple targets) showed no binding to any of the three targets [4]. It is also important to note that at high enough target (analyte) concentrations, both high and low-affinity conformations of the TAD (ligands) are capable of binding, which can also result in the observation of two phases [30, 34-35]; however, through additional experiments the authors determined that their data was most consistent with a model involving a conformational change and not the heterogeneous ligand binding model.

B.2.A. Limitations

From these data, it is evident that SPR is not the optimal method for analysis of the transient kinetics of TAD-target interactions, as exemplified by the observation of both mono and biphasic sensograms by the TADs of Gal4 and VP16. One possible explanation for such inconsistencies is that because the response measured by SPR for the conformational change step is caused by a change in the hydrodynamic radius of the complex [39], the technique of SPR may not be sensitive enough to reproducibly detect a slight "conformational tightening" of pre-structured motifs during the stabilization of the TAD•target complexes. In order to enhance its sensitivity, researchers have started to

perform SPR experiments simultaneously with surface plasmon fluorescence spectroscopy (SPFS) [40]. Furthermore, diffusion-limited association reactions occur on a time-scale that is too fast for the SPR instrument to accurately measure [41]. Thus, Wright and coworkers treated the fast phase in their sensograms as a rapid equilibrium, whereby they lost information pertaining to the bimolecular association and dissociation rate constants [4]. Rather, as will be presented in the next section, fluorescence stopped-flow techniques are better suited to address these issues for the accurate study of the process of TAD•target complex formation.

C. Experimental Design: Characterization of the interactions of DNA-bound activators with the coactivator Med15 via fluorescence techniques

Due to the limitations presented above, we have implemented a strategy to investigate TAD-target interactions using fluorescence stopped-flow techniques in order to unravel the minimal mechanism by which the TADs of the prototypical activators Gal4 (residues 840-881), Gcn4 (residues 107-144), and VP16 (residues 456-490) interact with the key coactivator Med15(Gal11). In addition, once obtained, the thermodynamic and kinetic rate constants for the individual steps within this mechanism will be used to identify key trends that provide insight into their differential potency.

C.1. Fluorescence stopped-flow spectroscopy

Fluorescence stopped-flow spectroscopy is a rapid-mixing method that has been implemented in the elucidation of enzymatic mechanisms, as well as changes in protein structure [42-43]. In the stopped-flow apparatus, the binding reaction is initiated by

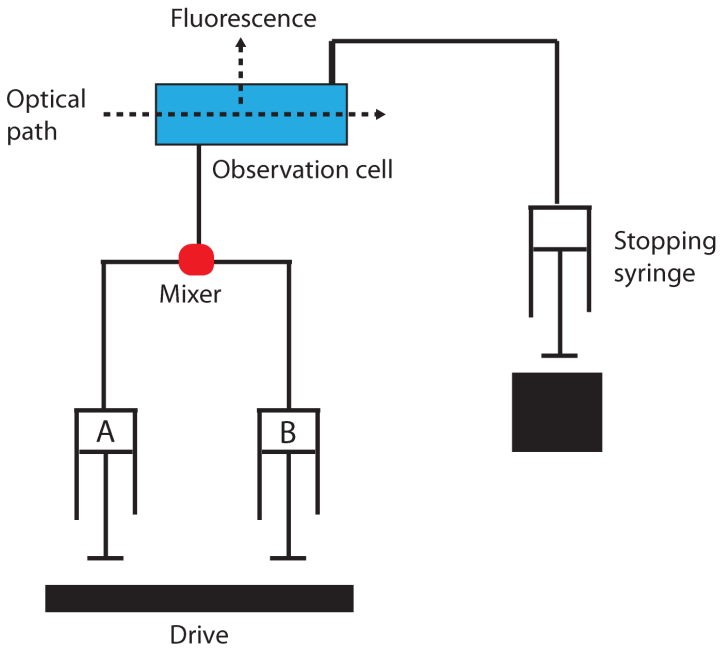


Figure 2.6: Fluorescence stopped-flow spectroscopy. Small volumes of solutions in two separate syringes (A and B) are first driven into a rapid mixing chamber. This freshly mixed solution then flows through the observation cell, displacing its original contents and filling the stopping syringe. Once the plunger of the stopping syringe hits a block, the steady-state flow is stopped instantaneously, and data collection is triggered in the observation cell. The fluorophore is excited and its emission is recorded over a fixed number of time points within a designated time-domain in order to generate a curve.

pushing two solutions maintained in separate syringes through a mixing chamber and into an observation cell (Figure 2.6) [42, 44]. This steady state flow is only ~1-2 milliseconds old when it reaches the observation cell (known as the dead time of the instrument), which in turn empties the previous contents of the observation cell into the stopping syringe until it is filled and the plunger hits a block ("stopped-flow"). Once this occurs, data collection at the observation cell is triggered, in which the change in the emission property of the fluorophore is recorded over time. The resulting curve can then be best fit to either a single or double exponential, and the concentration dependence of the observed rate constant(s) can be determined. We therefore wanted to apply this technique

to monitor the formation of TAD•target complexes over time with a higher time-domain resolution than can be achieved with SPR. However, in order to achieve this, a fluorescent probe had to be installed within our system, which was achieved by the utilization of TADs within the context of a preformed DNA•activator complex.

C.2. DNA-bound activators

In our experimental set-up each TAD is characterized in the context of a DNA-bound intact activator. Not only does this approach provide a facile way to install a fluorophore into the system for signal detection purposes (i.e., at the 5' end of the DNA duplex), but it also better mimics their presentation from a gene promoter. To achieve this, the TADs were attached to the first 100 amino acids of Gal4, which encompasses both the $Zn(II)_2Cys_6$ binuclear cluster DNA-binding domain (DBD) and the dimerization domain (dd), for which a crystal structure of it complexed to a single DNA binding site (consisting of two half sites) has been solved (Figure 2.7) [45]. From this structure, Gal4(1-100) appears to be predominantly well-folded, and is predicted to project the TADs for unhindered access if attached to the C-terminus. However, once the TADs were attached, the activator required an additional N-terminal MBP solubility tag to prevent aggregation due to their intrinsically disordered properties. Furthermore, as will be seen, interaction of preformed DNA•activator complexes with the coactivator Med15(Gal11) resulted in a change in fluorescence (due to either direct or indirect perturbation of the fluorophore on the DNA) that can be monitored under either equilibrium conditions or in real-time.

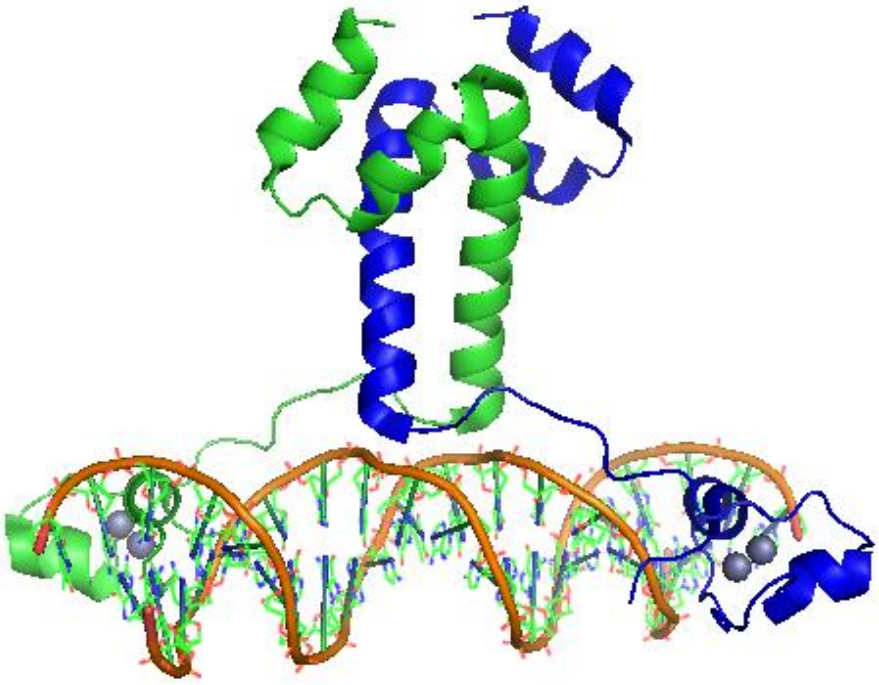


Figure 2.7: DNA-bound Gal4(1-100). Ribbon diagram of the crystal structure of a Gal4(1-100) dimer bound to DNA (PDB ID: 3COQ). The two subunits of the dimer are colored in blue and green, respectively, with zinc ions as gray spheres.

C.3. Coactivator target:Med15(Gal11)

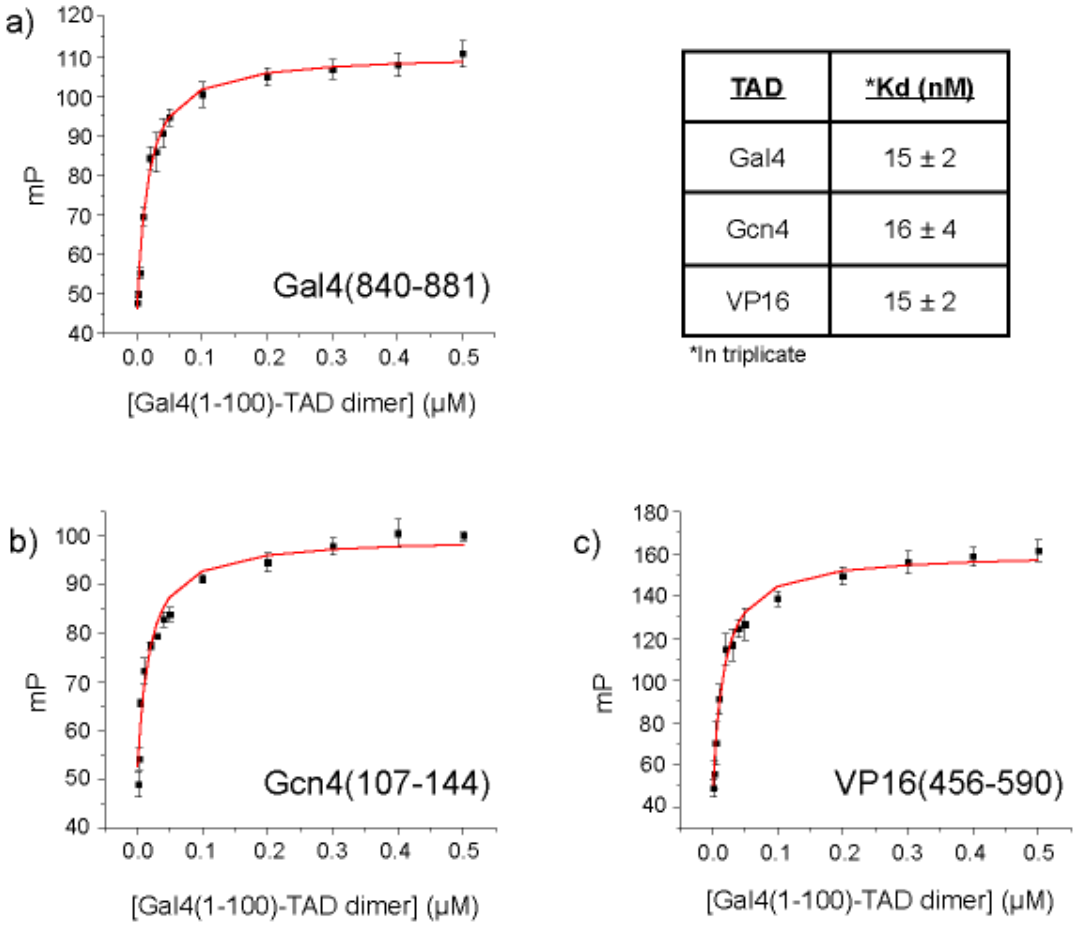
The tail module of the yeast Mediator complex is believed to be an important target for transcriptional activators, and it is within this module that the coactivator Med15(Gal11) resides via an interaction with its C-terminal domain [46-48]. In fact, there is considerable evidence that the TADs of Gal4, Gcn4, and VP16 share a functionally important binding interaction with the N-terminus of Med15 [49-51]. For instance, i*n vitro* crosslinking experiments performed by Mapp and coworkers revealed contacts between the N-terminus of Med15 (residues 1-416) and an incorporated photoactive p-benzoyl phenylalanine (pBpa) unnatural amino acid within the isolated TADs of Gal4 and Gcn4; these experiments identified an overlapping functional binding

site at around residue 200 of Med15 [51]. In addition, similar *S. cerevisiae* crosslinking experiments performed *in vivo* using nonsense suppression methods to incorporate the amino acid pBpa within the TAD of VP16 revealed its direct interaction with this fragment of Med15 as well (unpublished data). Based on these studies, we hypothesize that determining the mechanism by which the protoypical activators Gal4, Gcn4, and VP16 interact with the N-terminal domain of this relevant target Med15(Gal11) would be an excellent approach to analyze the binding properties that correlate with favorable activity. As described in the following section, we expressed and purified the N-terminal 345 residues of Med15 with a GST solubility tag, and monitored its interaction with preformed fluorescently-labeled DNA•activator complexes by following the increase in fluorescence that occurs upon binding.

D. Results: Analysis of the DNA•activator•Med15 assembly pathway

D.1. Overall affinity ($K_{d,apparent}$) of DNA-activator interactions

Upon analysis of DNA•activator complex formation through fluorescence polarization (FP), the three activator constructs exhibited an identical binding affinity (15 ± 5 nM) for 1 nM of a fluorescein-labeled consensus DNA binding site (composed of two half sites) (Figure 2.8); these K_ds are consistent with those previously reported for Gal4(1-100) and this DNA sequence [45], indicating that the presence of the MBP solubility tag was not hindering the ability of the activators to interact with DNA. Thus, the DNA binding function of the activators is independent of the TAD, consistent with the modular architecture of most transcriptional activators [2, 9]. Additionally, in order to ensure that the activator is binding sequence-specifically to the DNA, we then determined



TAD	*Kd (nM)
Gal4	15 ± 2
Gcn4	16 ± 4
VP16	15 ± 2

*In triplicate

Figure 2.8: Dissociation constants of activators for DNA. A constant 1 nM concentration of the 5'-Fluorescein-labeled duplex DNA was incubated with varying concentrations of Gal4(1-100)-TAD for 30 min at room temperature and the resultant polarization values at each protein concentration were obtained on a Beacon 2000 instrument (PanVera Corp). Each value is the average of three independent experiments with the indicated error (standard deviation). The solid line is a curve fit to these data according to a 1:1 binding interaction of duplex DNA to dimeric activator [1].

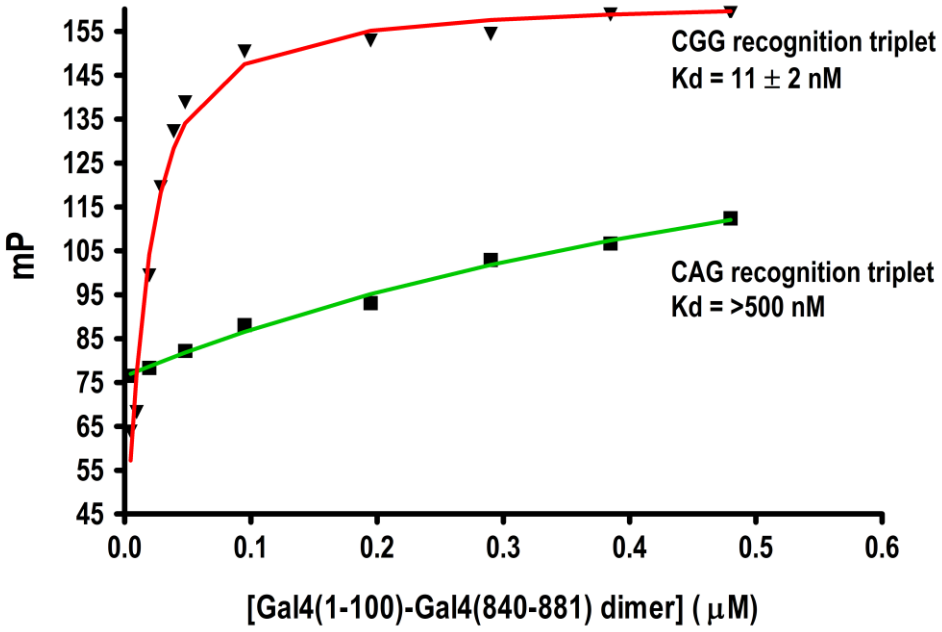


Figure 2.9: DNA-binding specificity of the activators. A constant 1 nM concentration of the 5'-Fluorescein-labeled duplex DNA was incubated with varying concentrations of Gal4(1-100)-Gal4(840-881) for 30 min at room temperature and the resultant polarization values at each protein concentration were obtained on a Beacon 2000 instrument (PanVera Corp). Mutation of the terminal CGG recognition triplet (red line) to a CAG (green line) abrogates the ability of the activator to make specific contacts with the DNA.

that the Gal4 activator is unable to bind to a duplex DNA with a doubly symmetric mutation within the palindromic CGG half sites under analogous conditions (Figure 2.9).

Next, the optimal activator:DNA ratio for utilization in the Med15 binding experiments was determined. From the data presented in Figure 2.8, it is apparent that the percent of the 1 nM DNA that is in the bound form begins to decrease significantly at a concentration of ≤100 nM of activator dimer (corresponding to a threshold of ~87% of the DNA in the bound state according to simulation studies, Figure 2.10). When the DNA concentration was increased to 25 nM in order to increase the signal (while maintaining

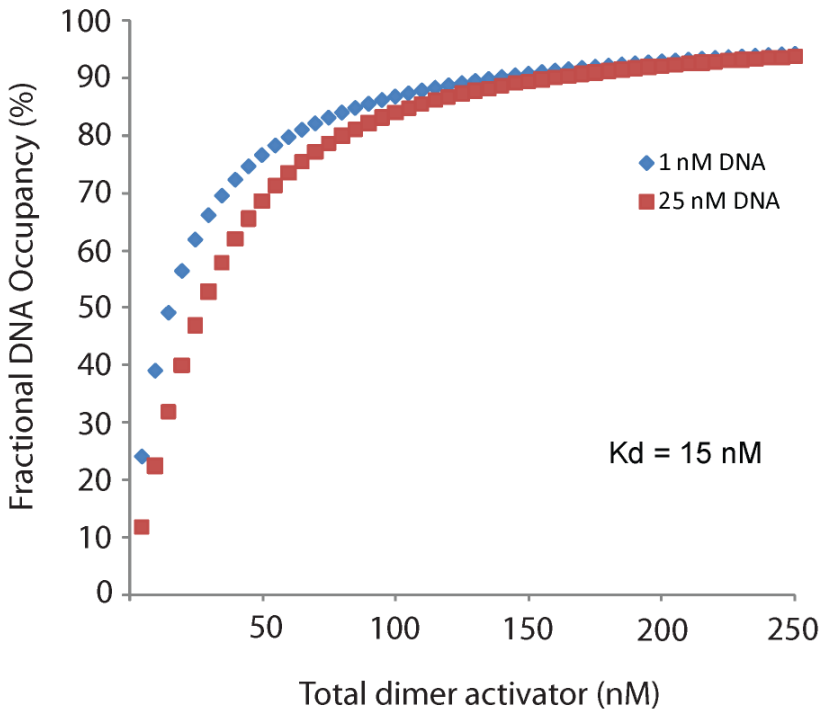


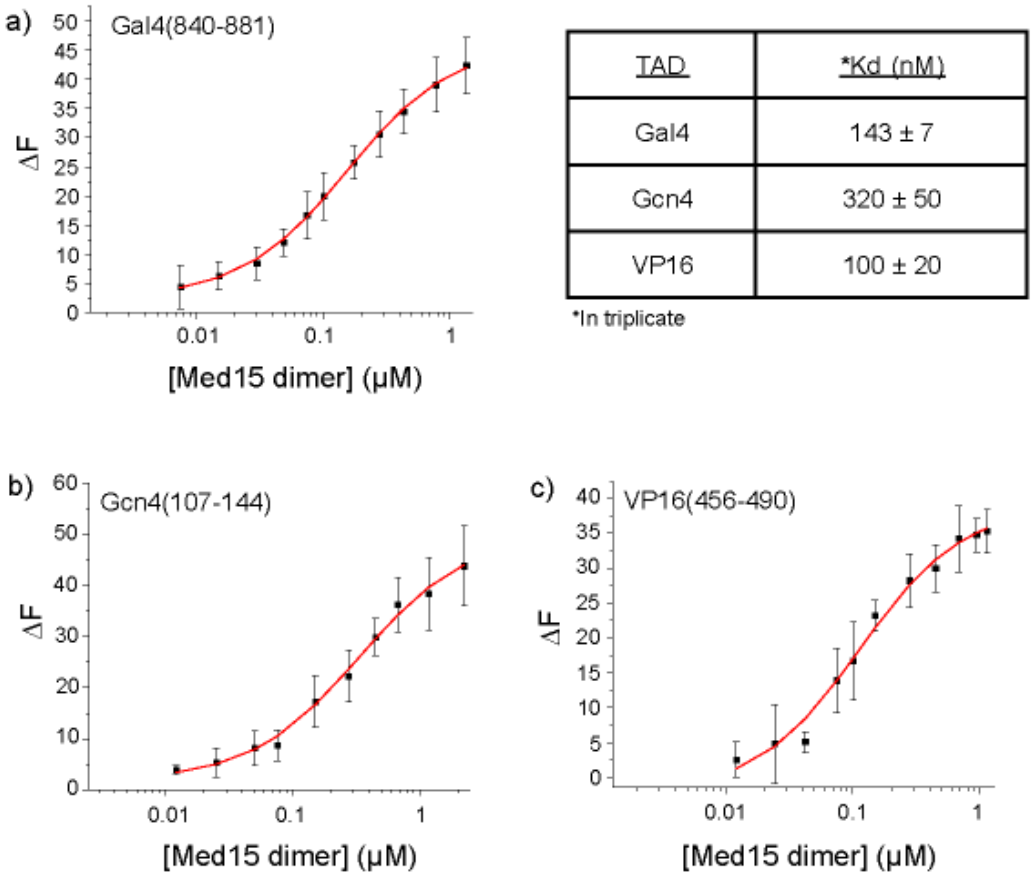
Figure 2.10: DNA-binding simulation studies. Using equation 11 (see Experimental Methods), the value of fraction bound = (y-c)/(b-c) was calculated at different values of x = $[activator]_{total}$ ranging from 5 nM to 250 nM (dimer) by setting the following parameters as constants: 1) a = $[DNA]_{total}$ = either 1nM or 25 nM, and 2) $K_{d,app}$ for activator binding to DNA = 15 nM.

pseudo-first order concentrations for the stopped-flow experiments) and the simulation studies were repeated, the amount of DNA present in the bound form was only reduced to ~84% (Figure 2.10). Therefore, a ratio of 4:1 activator dimer:duplex DNA was determined to be optimal for activator•DNA complex formation throughout the remainder of the binding experiments.

D.2. Overall affinity ($K_{d,apparent}$) of DNA•activator-Med15 interactions

The affinities of preformed activator•DNA complexes for the coactivator Med15 were then determined by fluorescence titration experiments. Titration of GST-tagged

Med15(1-345) dimer (see Figure 2.21 in the Experimental Methods) into a solution of pre-formed activator•DNA complex labeled with the BODIPY® FL fluorophore produced an increase in fluorescence signal (~20 %) until saturation is established at low micromolar concentrations. The three activators exhibit similar apparent dissociation constants ($K_{d,app}$) with Med15 in the submicromolar range, with the VP16- and Gal4-derived activators interacting more strongly than Gcn4 (Figure 2.11).



TAD	*Kd (nM)
Gal4	143 ± 7
Gcn4	320 ± 50
VP16	100 ± 20

*In triplicate

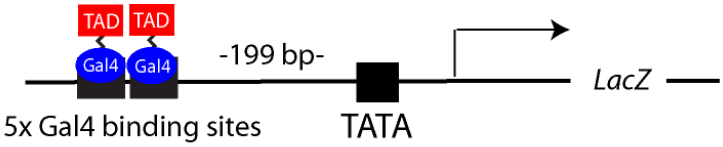
Figure 2.11: Dissociation constants of DNA-bound activators for Med15(Gal11). A solution of 25 nM 5'-BODIPY®FL-DNA•Gal4(1-100)-TAD complex was titrated with increasing amounts of GST-Med15(1-345), and the resultant fluorescence enhancement at each coactivator concentration was monitored on an Eclipse spectrofluorometer (Varian Corp). Each value is the average of three independent experiments with the indicated error (standard deviation). The solid line is a curve fit to these data according to a 1:1 binding interaction of DNA-bound activator to dimeric Med15.

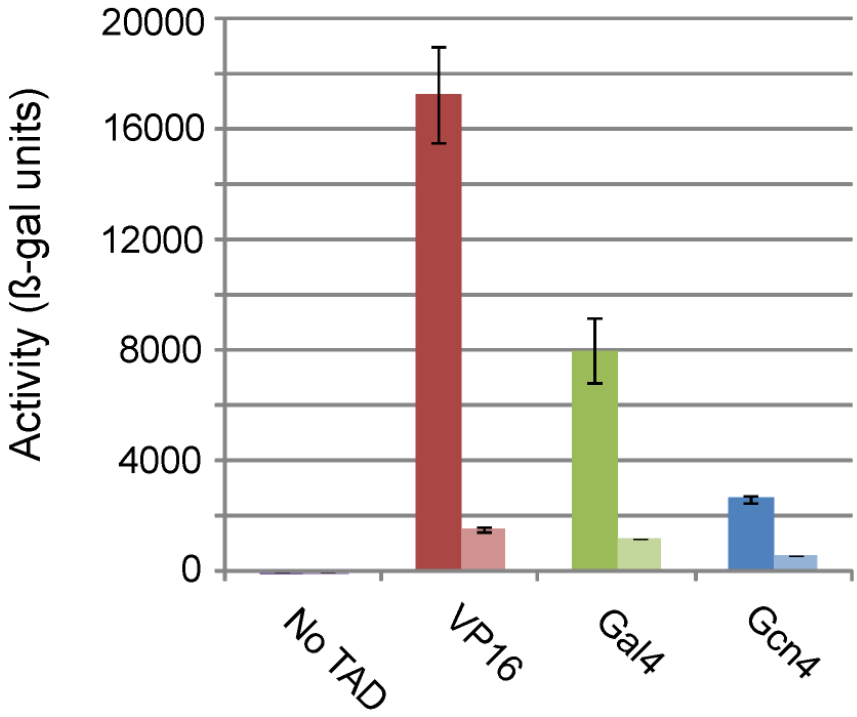
We then examined whether there was a positive correlation between the apparent affiinty ($K_{d,app}$) of the activators for Med15 and their activity. To this end, Figure 2.12 shows the results of a β-galactosidase assay in the presence of either full-length Med15, or Med15 with the N-terminal 345 amino acids deleted.[2] From these results, it is evident that all three activators are dependent on this N-terminal region for activity, and that the activity decreases from VP16 > Gal4 > Gcn4, which mirrors their binding affinities; a trend that is consistent with previous studies on TAD-target interactions (Chapter 1). However, the factors within the individual steps of the binding mechanism that are attributing to these differences in overall affinity still need to be determined.

D.3. Biphasic association of DNA•activator complexes to Med15

The transient association kinetics between the DNA-bound activators of Gal4, Gcn4, and VP16 and the N-terminus of Med15 were then monitored using stopped-flow fluorescence. Briefly, pre-formed DNA•activator complex and GST-Med15(1-345) were placed into separate syringes of a KinTek model SF-2001 stopped-flow and mixed under pseudo-first-order conditions in which Med15 is in excess. Once in the observation cell, the BODIPY® FL fluorophore attached to the duplex DNA was excited at 502 nm and an increase in fluorescence over time was monitored using a 510 nm long-pass filter. The resulting time courses obtained for all three activators best fit to a sum of two exponentials (equation 12) (Figures 2.13-2.15), which is consistent with a two-step binding mechanism. The residuals for both the single and double exponential fits are also provided to further illustrate the differences in the goodness of the fits.

[2] Experiment was performed by Ningkun Wang.





TAD	Activity (β -gal units)	
	Med15 wt	Med15Δ(1-345)
No TAD	3 ± 1	2 ± 1
VP16	17302 ± 1767	1555 ± 62
Gal4	8001 ± 1187	1182 ± 33
Gcn4	2651 ± 131	580 ± 12

Figure 2.12: Specificity of natural TADs for Med15(Gal11). Activities of the Gal4(1-100)-TAD activators (VP16(456-490) > Gal4(840-881) > Gcn4(107-144)) in the presence of full-length Med15 (dark colored bars). Deletion of the N-terminal 345 amino acids of Med15 results in a loss of activity of all three activators (light colored bars). Each β-galactosidase assay was performed at least in triplicate with the indicated error (standard deviation of the mean (SDOM).

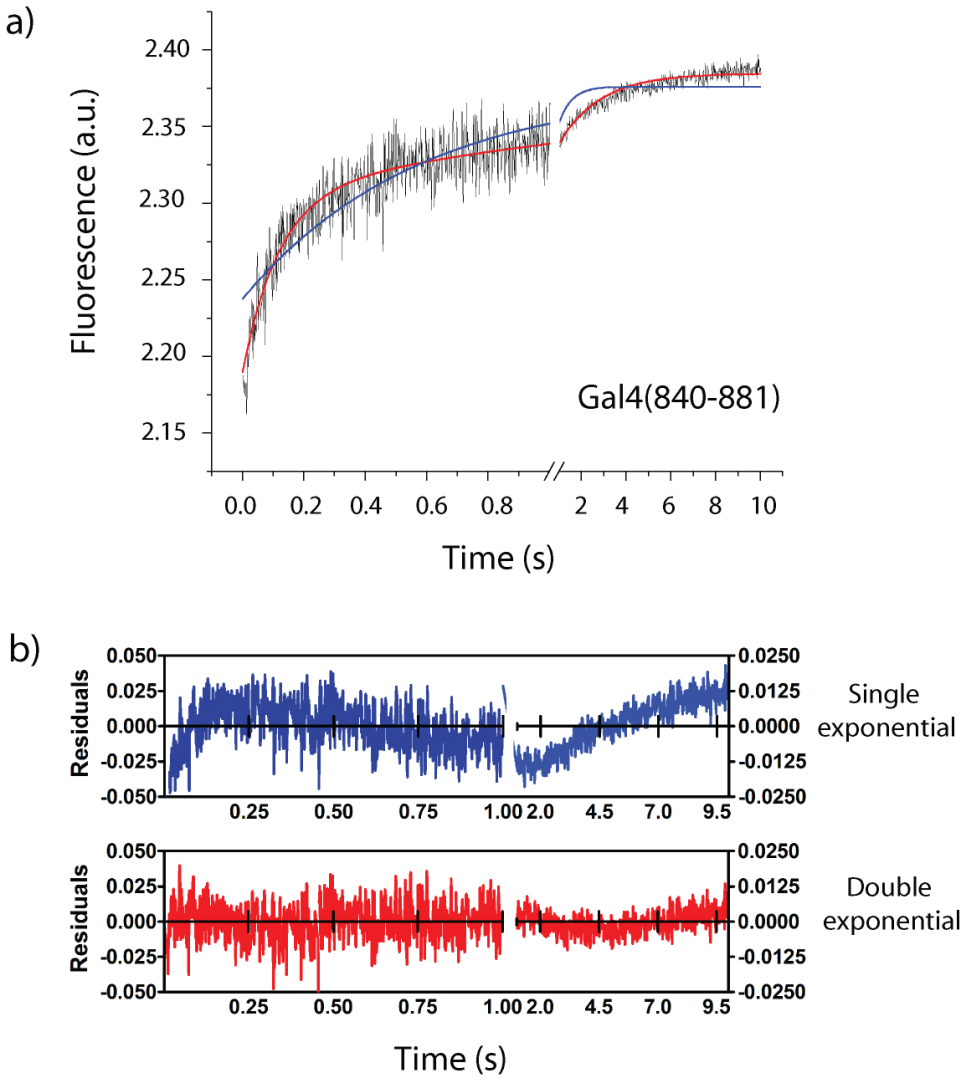


Figure 2.13: Stopped-flow kinetic studies of the DNA-bound Gal4(1-100)-Gal4(840-881) activator complex binding to Med15(1-345). (a) Binding time course for 5'-BODIPY®FL-DNA•Gal4(1-100)-Gal4(840-881) activator complex (25 nM) mixed with Med15 (0.625 μM) in a stopped-flow apparatus in DNA-binding buffer at 25 °C and monitoring the fluorescence (λ_{ex} = 502 nm; λ_{em} > 510 nm). The blue line superimposed on the time course is the best fit of the sum of one exponential to the data. The red line is the best fit of the sum of two exponentials to the data, with observed rates $k_{obs,1}$ = 8.53 ± 0.44 s^{-1} and $k_{obs,2}$ = 0.54 ± 0.04 s^{-1}, and amplitude terms A_1 = 0.120 ± 0.003 and A_2 = 0.079 ± 0.002. (b) Residuals for the single exponential fit (blue line) or double exponential fit (red line) to the time course in (a) reveals that a double exponential is the best fit to the data. Values for the first and second time frames are plotted with respect to the left and right axes, respectively.

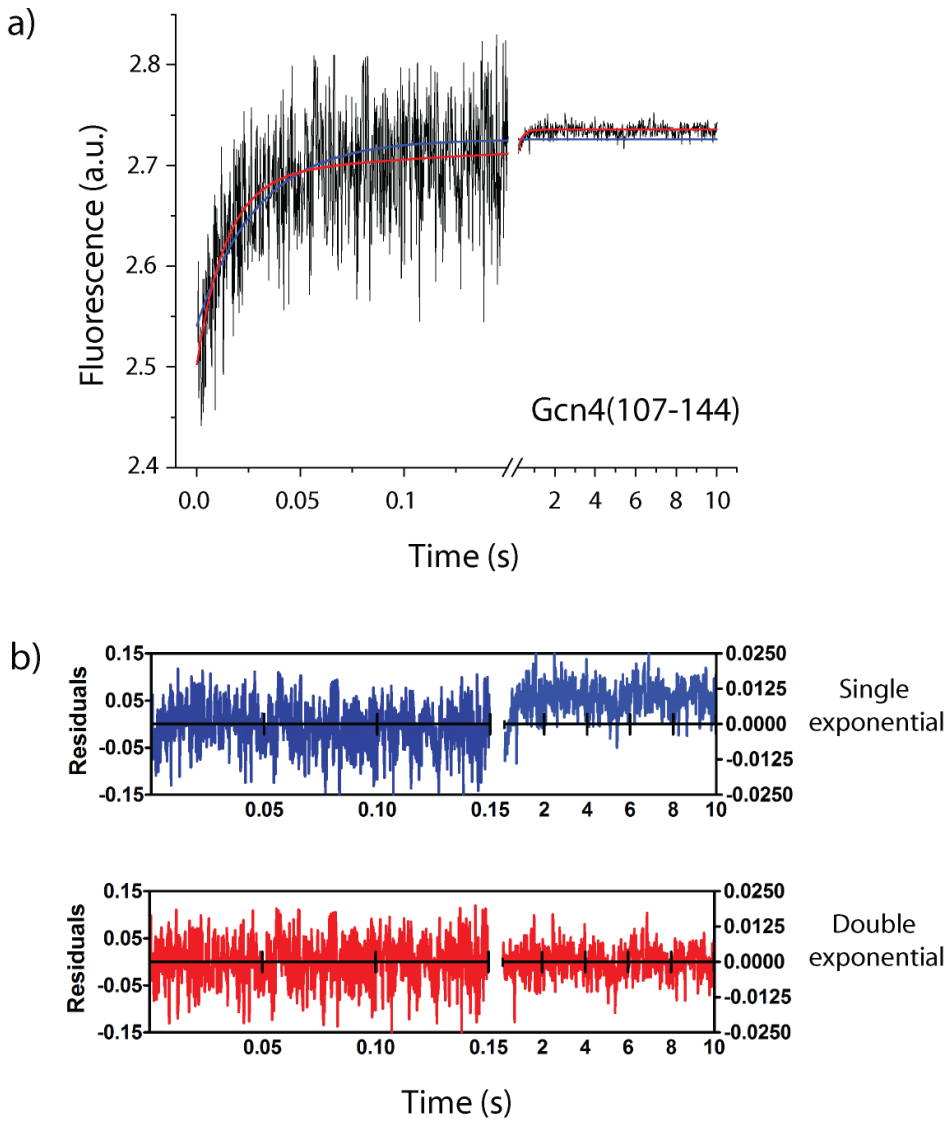


Figure 2.14: Stopped-flow kinetic studies of the DNA-bound Gal4(1-100)-Gcn4(107-144) activator complex binding to Med15(1-345). (a) Binding time course of 5'-BODIPY®FL-DNA•Gal4(1-100)-Gcn4(107-144) activator complex (25 nM) mixed with Med15 (0.75 μM) in a stopped-flow apparatus in DNA-binding buffer (at 25 °C and monitoring the fluorescence (λ_{ex} = 502 nm; λ_{em} > 510 nm). The blue line superimposed on the time course is the best fit of the sum of one exponential to the data. The red line is the best fit of the sum of two exponentials to the data, with observed rates $k_{obs,1}$ = 69.77 ± 8.25 s^{-1} and $k_{obs,2}$ = 3.14 ± 1.49 s^{-1}, and amplitude terms A_1 = 0.20 ± 0.01 and A_2 = 0.050 ± 0.008. (b) Residuals for the single exponential fit (blue line) or double exponential fit (red line) to the time course in (a) reveals that a double exponential is the best fit to the data. The single exponential fit does not describe the fluorescence changes that occur in the second time frame. Values for the first and second time frames are plotted with respect to the left and right axes, respectively.

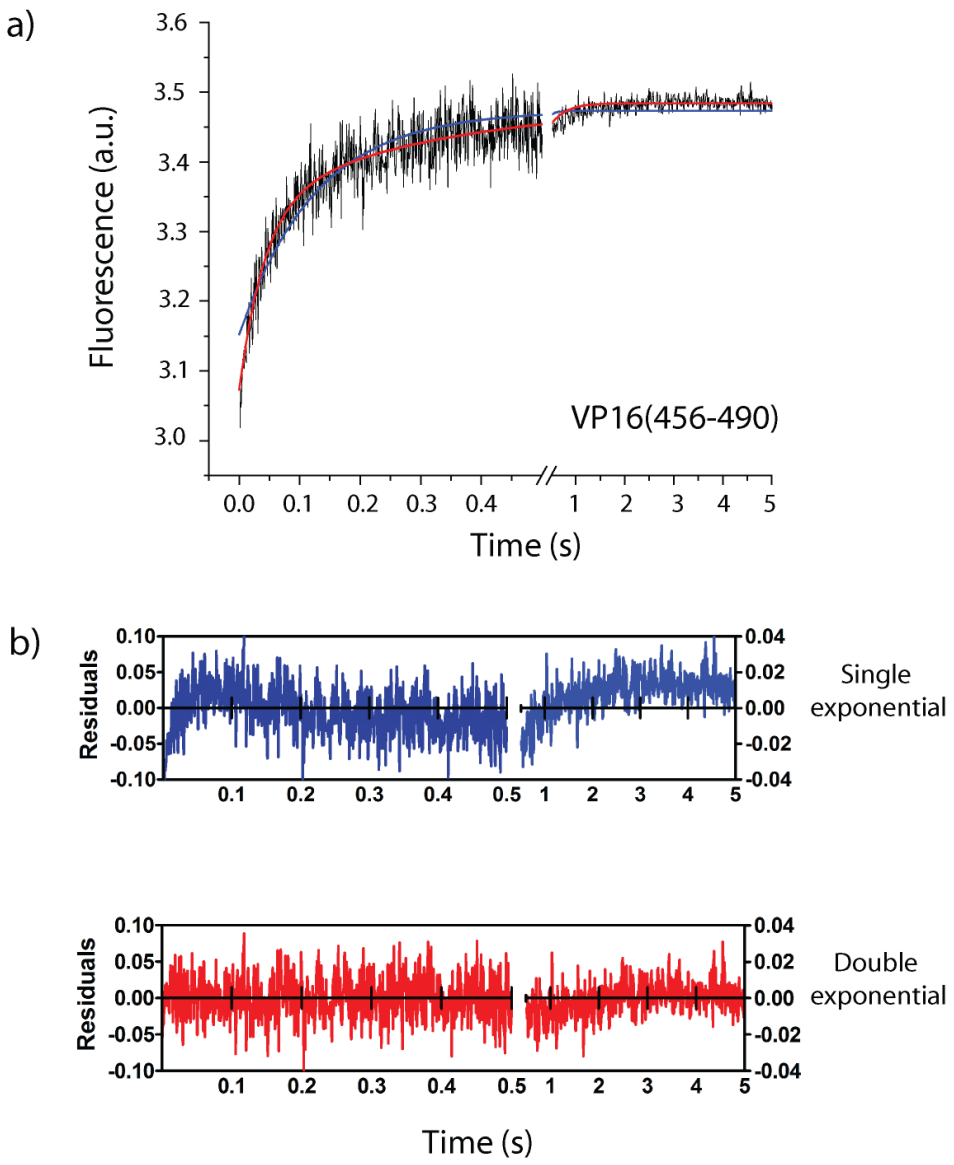


Figure 2.15: Stopped-flow kinetic studies of the DNA-bound Gal4(1-100)-VP16(456-490) activator complex binding to Med15(1-345). (a) Binding time course of 5'-BODIPY®FL-DNA•Gal4(1-100)-VP16(456-490) activator complex (25 nM) mixed with Med15 (0.50 μM) in a stopped-flow apparatus in DNA-binding buffer (at 25 °C and monitoring the fluorescence (λ_{ex} = 502 nm; λ_{em} > 510 nm). The blue line superimposed on the time course is the best fit of the sum of one exponential to the data. The red line is the best fit of the sum of two exponentials to the data, with observed rates $k_{obs,1}$ = 23.49 ± 1.76 s^{-1} and $k_{obs,2}$ = 3.10 ± 0.26 s^{-1}, and amplitude terms A_1 = 0.27 ± 0.01 and A_2 = 0.15 ± 0.01. (b) Residuals for the single exponential fit (blue line) or double exponential fit (red line) to the time course in (a) reveals that a double exponential is the best fit to the data. Values for the first and second time frames are plotted with respect to the left and right axes, respectively.

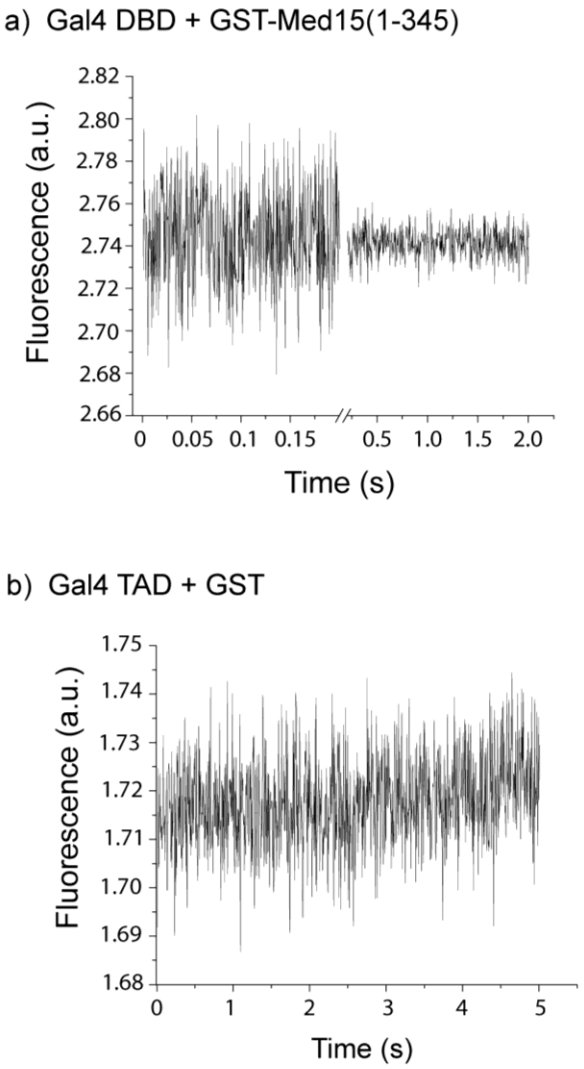


Figure 2.16: Negative control stopped-flow experiments. (a) No fluorescence enhancement is observed in the time-course when 5'-BODIPY®FL-DNA•Gal4(1-100) complex (25 nM) is mixed with Med15 (0.375 μM) at 25 °C. (b) No fluorescence enhancement is observed in the time-course when 5'-BODIPY®FL-DNA•Gal4(1-100)-Gal4(840-881) activator complex (5 nM) is mixed with GST (0.25 μM) at 25 °C.

From these fits, values for the observed rates for the faster phase ($k_{obs,1}$) and slower phase ($k_{obs,2}$) were also obtained, which are unique to the identity of the TAD. Because the observed rates are functions of all the microscopic rate constants that are reversibly linked, this indicates that the values of these different rate constants are unique to each TAD as well. Analogous experiments performed on Gal4(1-100) lacking a TAD or with GST alone produced no increase in fluorescence over background (Figure 2.16), further illustrating that the observed fluorescence change is dependent on a TAD-Med15 interaction.

D.4. Concentration dependence of $k_{obs,1}$ and $k_{obs,2}$

The key to obtaining mechanistic information from transient kinetic methods is to examine the concentration dependence of the observed rates (k_{obs})[28]. Therefore, time-courses were collected for each of the three activators over a range of Med15 concentrations (in duplicate), all of which exhibited biphasic behavior (data not shown). The observed rate for the faster phase ($k_{obs,1}$) was then plotted as a function of [Med15], and the results for a single concentration-dependence study are shown in Figure 2.17. As seen in Figure 2.17, $k_{obs,1}$ for all three activators is linearly dependent on the concentration of Med15, which is consistent with a bimolecular collision step. In addition, the positive slope reflects the apparent bimolecular association rate constant (9 - 66 x 10^6 $M^{-1}s^{-1}$) (Table 2.2) which is in the range of a diffusion-controlled process (10^6 – 10^7 $M^{-1}s^{-1}$) [52]. The observed rate for the slower phase ($k_{obs,2}$) was then plotted as a function of [Med15] as well, and the results for a single concentration-dependence study are shown in Figure 2.17. As seen in Figure 2.17, $k_{obs,2}$ for all three activators does not

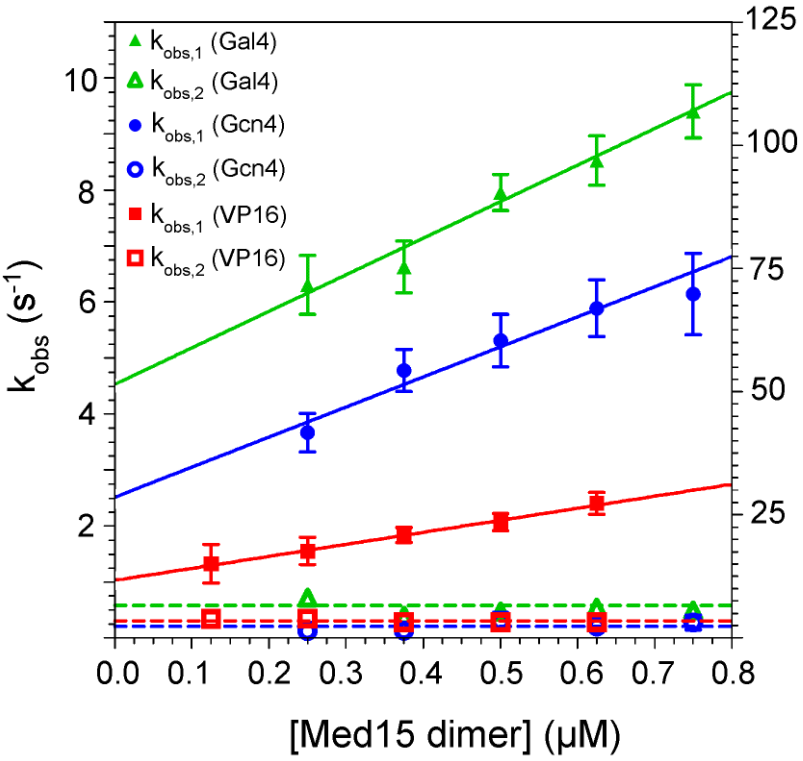


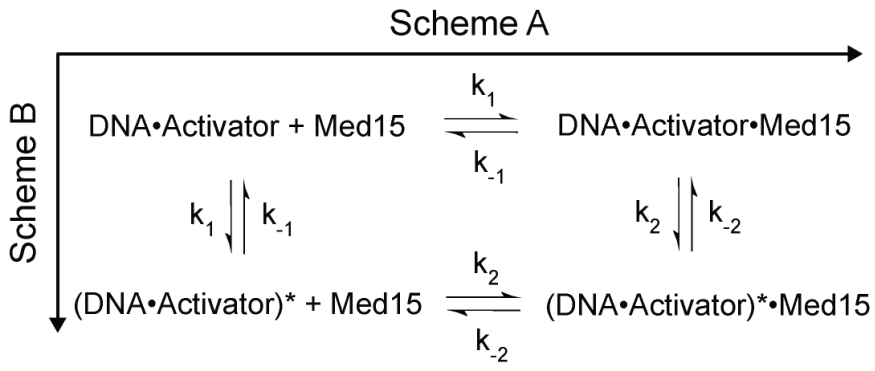
Figure 2.17: Dependence of k_{obs} on the concentration of Med15. Plot of $k_{obs,1}$ and $k_{obs,2}$ of the activators Gal4 (left axis), Gcn4 (right axis) and VP16 (right axis) versus dimeric Med15 (0.125-0.75 μM). The standard errors are indicated. The solid line is a curve-fit to these data, and the dotted line is the average $k_{obs,2}$.

show a dependence on the concentration of Med15, which is consistent with the occurrence of a rate-limiting conformational change step.

Table 2.2: Experimentally determined kinetic constants

TAD	k_{on} ($\mu M^{-1}s^{-1}$)	$k_{obs,1}^{(y\text{-intercept})}$ (s^{-1})	$k_{obs,2}^{(avg)}$ (s^{-1})
VP16	23 ± 2	12.1 ± 0.7	3.3 ± 0.2
Gal4	9.1 ± 0.8	3.6 ± 0.4	0.55 ± 0.04
Gcn4	66 ± 8	26 ± 4	2.6 ± 0.4

a)



b) Equations for the binding model in Scheme A

(1) $k_{obs,1} \approx k_1[Med15] + k_{-1} + k_2 + k_{-2}$

(2) $k_{obs,2}^{max} \approx k_2 + k_{-2}$

(3) $K_1 = k_1/k_{-1}$

(4) $K_2 = k_2/k_{-2}$

(5) $K_{d,app} = 1/[K_1(1 + K_2)]$

c) Equations for the binding model in Scheme B

(6) $k_{obs,1} \approx k_1 + k_{-1} + k_2[Med15] + k_{-2}$

(7) $k_{obs,2}^{max} \approx k_1$

(8) $K_1 = k_1/k_{-1}$

(9) $K_2 = k_2/k_{-2}$

(10) $K_{d,app} = (1 + K_1)/(K_1K_2)$

Figure 2.18: Limiting two-step binding models for DNA-bound activators interacting with Med15 (Schemes A and B). (a) In Scheme A, the conformational change occurs after an initial binding event, whereas in Scheme B, the DNA-bound activator undergoes a conformational change prior to associating with Med15. (b) Equations used to calculate the microscopic rate and equilibrium constants, according to the binding model presented in Scheme A. (c) Equations used to calculate the microscopic rate and equilibrium constants, according to the binding model presented in Scheme

D.5. Two-step binding models: Conformational change occurs after (Scheme A) or before (Scheme B) the bimolecular association step

Although the data presented thus far supports a two-step DNA•activator-Med15 binding mechanism involving a fast, bimolecular association step and a slow, conformational change step, the sequence of events has yet to be determined; i.e., the data is consistent with two limiting models in which the conformational change within the complex occurs either solely after (Scheme A) or before (Scheme B) the binding step (Figure 2.18a) [28]. Using the experimentally obtained values for $K_{d,app}$, k_{on}, $k_{obs,1}^{(y\text{-intercept})}$ and $k_{obs,2}^{(max)}$ (Table 2.2), the microscopic rate constants k_1, k_{-1}, k_2, k_{-2} were calculated according to the binding models in Schemes A (equations (1)-(5), Figure 2.18b) and B (equations (6)-(10), Figure 2.18c) [28]; the results are summarized in Tables 2.3 and 2.4, respectively, and the details of the calculations performed are available within the Experimental Methods.

Table 2.3: Microscopic rate constants for the binding reaction in Scheme A*

TAD	k_1 ($\mu M^{-1}s^{-1}$)	k_{-1} (s^{-1})	k_2 (s^{-1})	k_{-2} (s^{-1})
VP16	23 ± 2	8.9 ± 0.7	2.4 ± 0.6	0.9 ± 0.3
Gal4	9.1 ± 0.8	3.0 ± 0.4	0.32 ± 0.06	0.23 ± 0.06
Gcn4	66 ± 8	24 ± 4	0.28 ± 0.08	2.3 ± 0.9

*A global-fit of these rate constants as fixed parameters to experimental time-courses are presented in Figure 2.22 within the Experimental Methods.

Table 2.4: Microscopic rate constants for the binding reaction in Scheme B

TAD	k_1 (s^{-1})	k_{-1} (s^{-1})	k_2 ($\mu M^{-1}s^{-1}$)	k_{-2} (s^{-1})
VP16	3.3 ± 0.2	8.2 ± 0.9	23 ± 2	0.7 ± 0.2
Gal4	0.55 ± 0.04	2.8 ± 0.5	9 ± 1	0.2 ± 0.1
Gcn4	2.6 ± 0.4	21 ± 5	66 ± 8	2.3 ± 0.9

Simulation studies were performed with the calculated rate constants for Schemes A and B as fixed parameters in order to generate time courses that could be compared to the experimental data (See Experimental Methods).[3] While the simulated curves for both Schemes A and B possess values for $k_{obs,1}$ and $k_{obs,2}$ that correspond well with the experimental data at various Med15 concentrations (Figure 2.23), the amplitude changes (A_1 and A_2) for Scheme A better agree with what is observed experimentally for all three activators (Figures 2.24 – 2.26). Although these differences are not significant enough to confidently distinguish between the two mechanistic schemes, taken together with the bulk of the structural evidence presented previously that supports the implications of a conformational change occurring *after* binding (Scheme A), we will further discuss the data in that context.

D.5.A. Additional two-step binding models

Additional binding models that could explain the biphasic association between that of DNA-bound activators and dimeric GST-Med15(1-345), represented as Schemes C-E in Figure 2.19, are not consistent with the experimental data. For example, Scheme C proposes that the unimolecular step can be attributed to an intramolecular binding event within a DNA•activator•Med15 intermediate. However, this phenomenon is unlikely due to the fact that the observed rate for the unimolecular step is slower than that of the bimolecular step. Next, because the observed rates for both the fast and slow steps along the reaction pathway are not linearly dependent on Med15 concentration, this argues against the existence of two different populations of DNA•activator complexes that can bind intermolecularly to Med15 at different rates (Scheme D). Finally, Scheme E

[3] Simulation studies were performed by Ningkun Wang.

proposes that binding of the 21 nM of preformed DNA•activator complex to Med15 shifts the equilibria and results in the complexation of the remaining 4 nM of free DNA with excess unbound activator. Thus, because a fluorescence change is only observed with DNA-bound activator interacting with Med15, the slow step would be newly formed DNA•activator complex binding to Med15. However, this would lead to the same percent fluorescence change as would be observed in the fast step, and therefore the amplitude of the slow step should be only ~20% of the amplitude of the fast step; this model does not recapitulate the behavior of the three activators (Figures 2.24 – 2.26).

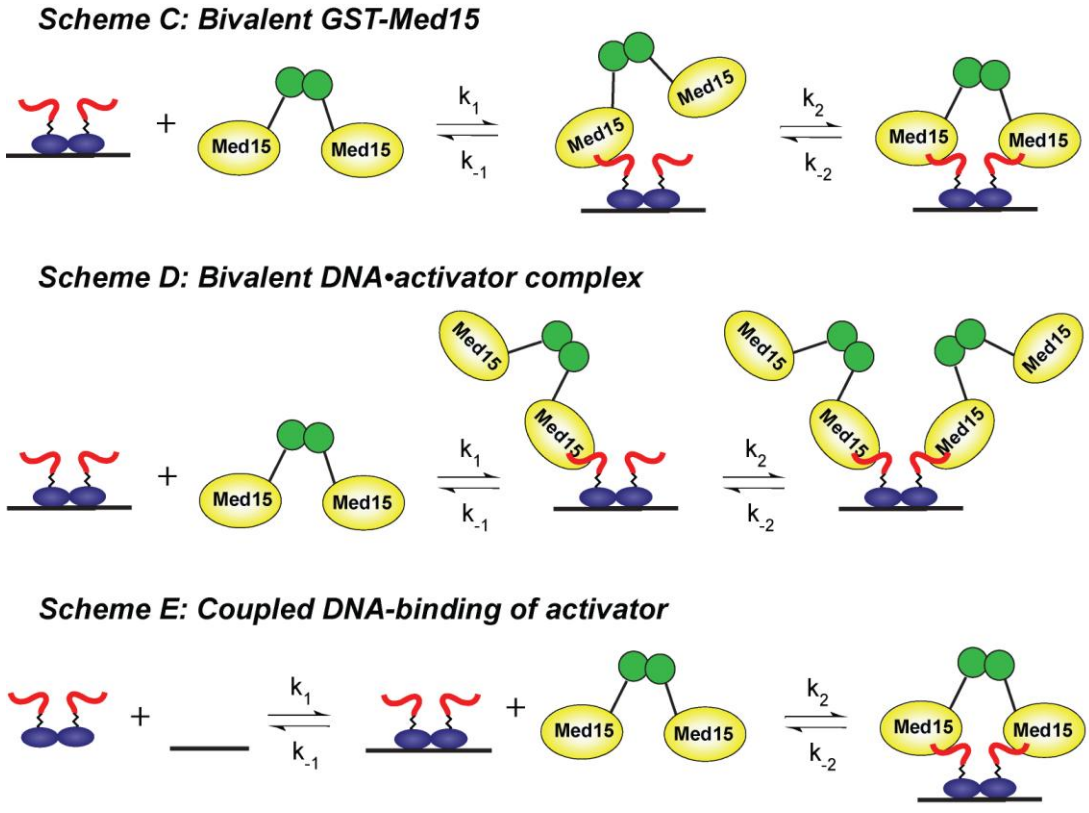


Figure 2.19: Additional two-step binding models for DNA-bound activators interacting with Med15 (Schemes C-E).

E. Discussion: Conformational change step is a defining feature of activators

A comparison of these rate (k_1, k_{-1}, k_2, k_{-2}) and equilibrium (K_1, K_2) constants for Scheme A revealed key trends that provide insight into the differential activity of the three activators Gal4, Gcn4, and VP16. As previously stated, the TAD with a more favorable $K_{d,app}$ for Med15 correlates with a TAD that is more active (VP16 > Gal4 > Gcn4) (Figure 2.12). However, a closer look at the thermodynamics of the individual steps within the kinetic mechanism reveals that this correlation is governed *not* by the equilibrium constant for the initial bimolecular association step (K_1), which is invariant among all three TADs, but by that of the subsequent conformational change step (K_2) (Figure 2.20). In particular, a greater value for K_2 (more favorable conformational change step) correlates with a more active TAD.[4] Furthermore, a more active TAD also possesses a larger ratio of k_2/k_{-1}, indicating that its intermediate DNA•activator•Med15 complex preferentially forms the final conformer relative to dissociation (Figure 2.20); just as this partition ratio correlates with activator potency, it is also utilized by enzymes such as T7 DNA polymerase to enhance binding specificity [53-54]. The remainder of this section will focus on hypotheses for how these differences translate to the ability of these activators to recruit the transcriptional machinery to a gene promoter.

E.1. Factors influencing the partition ratio (k_2/k_{-1})

According to Scheme A, the observed intermediate DNA•activator•Med15 complexes are not committed along the reaction coordinate to form the final complex (i.e., $k_{-1} > k_2$) [55], and thus we propose that the preference by which the intermediate undergoes

[4] The same trend is observed for data analyzed in accordance with Scheme B (data not shown). In other words, the major conclusion that it is the differences in the rate of the conformational change step that correlate with transcriptional activity is model independent.

a)

TAD	Scheme A		
	K_1 (μM^{-1})	K_2	k_2/k_{-1}
VP16	2.6 ± 0.3	2.8 ± 0.6	0.27 ± 0.07
Gal4	3.0 ± 0.5	1.4 ± 0.2	0.10 ± 0.02
Gcn4	2.8 ± 0.5	0.12 ± 0.03	0.012 ± 0.004

b)

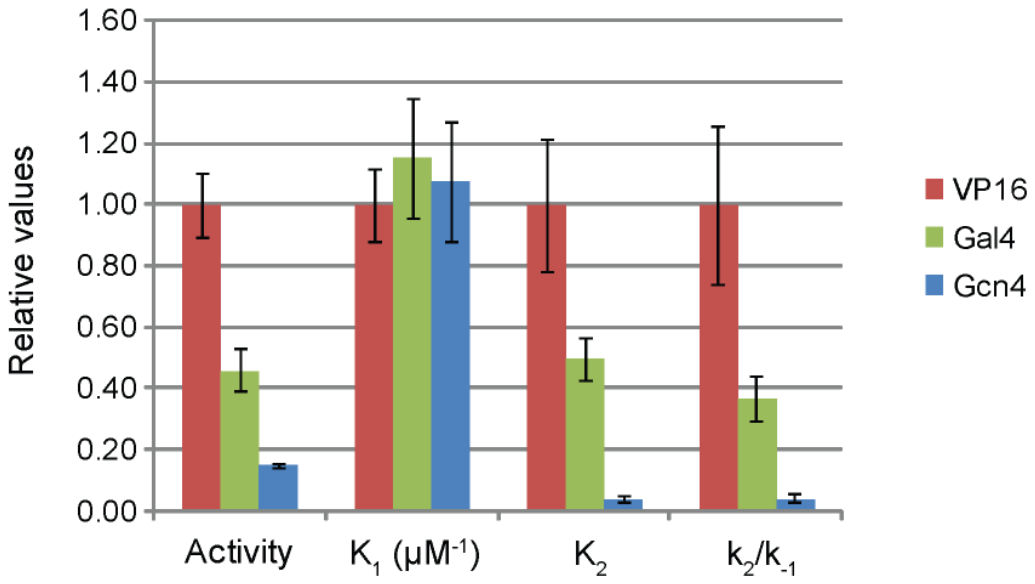


Figure 2.20: Conformational change step correlates with activating potential. (a) Values for the microscopic equilibrium constants for Schemes A. (b) Comparison of the equilibrium constants (K_1 and K_2) and partition coefficients (k_2/k_{-1}) for the binding mechanism presented in Scheme A to that of activator potency. Values are normalized to the most potent activator, VP16.

reorganization (k_2) relative to dissociation (k_{-1}) plays an important role in its ability to recruit the transcriptional machinery to a gene promoter. Thus, we hypothesize that one contributing factor to the differential partition ratios (k_2/k_{-1}) observed for the TADs of Gal4, Gcn4, and VP16 is the differences in the stabilities of the intermediates that are formed, with a greater structural propensity within the TAD resulting in a more stable

intermediate. For instance the most potent activator VP16, whose residues 472-479 within the isolated VP16 TAD have a propensity to form a relatively well-defined helix [25], has the largest partition ratio, although additional structural studies will be required to refine this model. Even further, the differences in the values of k_1 among the activators, which suggest that a conformational change step is occurring during the time-domain of the fast bimolecular collision step (either before or after the binding event), also supports this hypothesis.

E.2. Factors influencing K_2

The favorability of the conformational change within the complex (K_2) is unique to each TAD sequence and correlates with their activating potential. This data suggests that the primary amino acid sequence of each amphipathic TAD (Figure 2.2a) doesn't simply encode the linear arrangement of its acidic, polar, and hydrophobic residues; rather, it is encoding the identity and quantity of potential attractive and repulsive interactions it has to utilize during the formation of the final bound conformation of the complex [20]. A reorganization of the TAD within the low-affinity intermediate complex is hypothesized to occur in a number of consecutive elementary steps of non-covalent bond-making and bond-breaking events fueled by its translational and rotational diffusion motions, the driving force of which is that each sequential step is distinguished from the previous one by stronger binding energy and longer duration (and the entropic gain from solvent release). In this view, the process of "sequential selection" is critical to achieve specificity within the complex, and it is a prime opportunity for TADs to deviate not only from one another, but between different coactivator targets as well.

F. Conclusions/Future Directions

Overall, we have observed that the TADs of Gal4, Gcn4, and VP16 bind to Med15 through the same two-step binding mechanism comprised of a diffusion-controlled bimolecular collision step and a slower conformational change step. We propose to take this model one step further, by hypothesizing that the favorability by which amphipathic TADs are capable of forming their intermolecular interactions within the binding site of a shared coactivator target during the conformational change step is unique to each TAD sequence and dictates their activating potential. Future studies (discussed further in Chapter 4 of this dissertation) will entail examining if this is a common binding mechanism that holds for other coactivator targets (such as Tra1 of the SAGA complex), as well as comparing the kinetics of complex formation for these other coactivators with what we have obtained for Med15. Furthermore, current work is underway to develop a FRET approach that can be used to detect what is occurring at the TAD•target interface, with the goal of defining the order in which the events occur within the binding mechanism. Finally, the dynamics of the TAD portions of the activator constructs will be compared via NMR methods through analysis of their residual dipolar couplings (RDCs), in order to determine the different structural propensities of these polypeptide sequences.[5]

[5] This project, which will be carried out by Ningkun Wang and Amanda Dugan, is a collaborative effort with Dr. Hashim Al-Hashimi.

G. Experimental Methods

Table of plasmids used in this study[*]

Plasmid name	Function
pGal4(1-100)-Gal4(840-881) [56], pGal4(1-100)-Gcn4(107-144) [56], pGal4(1-100)-VP16(456-490) [57]	ARS/CEN yeast expression plasmid under the control of a β-actin promoter with a HIS^+ selection marker (NYC317)
ycplac111-Med15 [51], ycplac111-Med15Δ(1-345) [51]	ARS/CEN yeast expression plasmid under the control of a Med15 native promoter with a LEU^+ selection marker
pMCSG9-Gal4(1-100)-Gal4(840-881) [58], pMCSG9-Gal4(1-100)-Gcn4(107-144) [58], pMCSG9-Gal4(1-100)-VP16(456-490) [58]	Expresses activators fused to the His_6-MBP solubility tag in *E. coli*
pGEX-Med15(1-345) [59]	Expresses Med15 fused to the GST solubility tag in *E. coli*

*The sequences of all plasmids were verified by sequencing at the University of Michigan Core Facility (Ann Arbor, MI).

β-galactosidase assays

LS41ΔMed15 [JPY9::ZZ41, *Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4* URA::pZZ41 Med15::TRP] yeast was co-transformed with plasmids encoding each Gal4(1-100)-TAD fusion and a fragment of Med15 using the LiOAc method or by electroporation, and transformed colonies were selected by growth on synthetic complete (SC) media containnig 2% gucose lacking uracil, tryptophan, histidine, and leucine. The activity of each activator construct was monitored using quantitative liquid β-galactosidase assays in accordance with established methods [60]. Freshly transformed colonies were used to inoculate 5 mL cultures of SC media containing 2% raffinose and lacking 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 0.15-0.20 (10-fold dilution). 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 total protein concentration of the extract, measured using a Bradford assay kit (Bio-Rad) with BSA as the standard.

Protein expression and purification

<u>His_6-MBP-tagged activators:</u> Expression of the activators fused to a His_6-MBP tag was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen) as previously described [61]. Maltose-binding protein (MBP) has been demonstrated to effectively enhance the solubility of aggregation-prone proteins [58]; in addition, fusion of solubility tags to the N-terminus of the Gal4 DBD has been reported previously not to impact activator function [62-63]. Briefly, cultures (50 mL) from single colonies were grown overnight at 37 °C (250 rpm) in Lennox L Broth (Research Products International) supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL) before dilution (50-fold) into 8 x 50 mL cultures of Lennox L Broth supplemented with ampicillin (100 μg/mL). After an OD_{600} of 0.4 was reached, expression was induced with IPTG (final concentration 1 mM) in the presence of 20 μM $ZnSO_4$ for 5 hours. Each cell pellet was resuspended in 10 mL lysis buffer A (10 mM Tris, pH 8.0 at 4 °C, 500 mM NaCl, 10% glycerol (v/v), 10 mM β-

ME, 0.1% Tween* 20 (v/v), and Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the His-tagged protein was isolated using Ni-NTA Agarose (Qiagen). The cell lysate was incubated with 200 μL of Ni-NTA beads for 1 hour at 4 °C. The beads were washed 8 times with 1 mL wash buffer A (20 mM Tris, pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 1 mM β-ME, 0.1% Tween* 20 (v/v), 30 mM imidazole). The protein was eluted from the beads by incubation at 4 ºC overnight with 1 mL elution buffer A (20 mM Tris pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 250 mM imidazole). The protein solution was buffer exchanged into storage buffer A (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM β-ME, 1 mM EDTA, 20 μM $ZnSO_4$) using a PD-10 column (GE Healthcare), and the protein concentration was measured using absorbance at 280 nm. The identity and purity (>90%) of the protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

GST-tagged Med15(1-345): Expression of GST-Med15(1-345) was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen). Glutathione S-transferase (GST) has been demonstrated to effectively enhance the solubility and stability of aggregation-prone proteins [64]; in addition, fusion of GST to the N-terminus of Med15(Gal11) has been reported previously not to impact activator binding to this target protein [49, 65]. Briefly, cultures (50 mL) from single colonies were grown overnight at 37 °C (250 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL) before dilution (100-fold) into 4 x 1 L of Select APS Super Broth supplemented with ampicillin (100 μg/mL). After an OD_{600} of 0.3 was reached, the cultures were cooled for 45 min at 16 °C (150 rpm), and expression was induced with

IPTG (final concentration 0.1 mM) for 5-6 hours at 250 rpm. Each cell pellet was resuspended in 25 mL lysis buffer B (100 mM PBS pH 7.4 (Pierce), 0.2% NP-40 Substitute (Fluka), 10% glycerol (v/v), 1 mM DTT and Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the GST-tagged protein was isolated using Glutathione Sepharose 4B (GE Healthcare). The cell lysate was incubated with 2 x 1 mL of glutathione beads for 1 hour at 4 °C. The beads were washed 6 times with 10 mL wash buffer B (100 mM PBS pH 7.4 (Pierce), 0.2% NP-40 Substitute (Fluka), 10% glycerol (v/v), 1 mM DTT), and the protein was eluted from the beads by incubation at 4 ºC overnight with 1 mL elution buffer B (50 mM Tris pH 8.0 at 4 °C, 0.015 M reduced glutathione, 0.1% NP-40 Substitute). Additional protein was eluted from the column by twice incubating the beads with elution buffer for 1 hour at 4 °C. The protein samples were combined and concentrated using a Centriprep 10K centrifugal filter device before buffer exchange into storage buffer B (10 mM PBS pH 7.4 (Pierce), 10% glycerol (v/v), 0.01% NP-40 Substitute, 1 mM DTT) using a PD-10 column (GE Healthcare). The protein was then concentrated using a Vivaspin 30K centrifugal filter device, and the protein concentration was measured using absorbance at 280 nm. The identity and purity (>85%) of the protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

Analytical gel filtration: Analytical gel filtration was performed to determine the oligomeric state of GST-Med15(1-345) over the concentration range used in the fluorescence stopped-flow kinetic experiments described below. A final concentration of 1.5 μM and 0.5 μM (based on monomer concentrations) of GST-Med15 (1-345) was run

through a Superose 6 gel filtration column equilibrated with DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β-mercaptoethanol, 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) [49]. At the highest concentration of GST-Med15(1-345) tested (1.5 μM), only a single peak eluted as detected by UV (280 nm), with a projected molecular weight of 114 kDa as determined from molecular weight standards (Figure 2.21). These data are consistent with a dimeric state of GST-Med15(1-345) (monomer = 62.5kD, dimer = 125 kDa). This same species was the only species observed at the lowest concentration tested (0.5 μM) as determined by a western blot probing for GST on eluted fractions (data not shown).

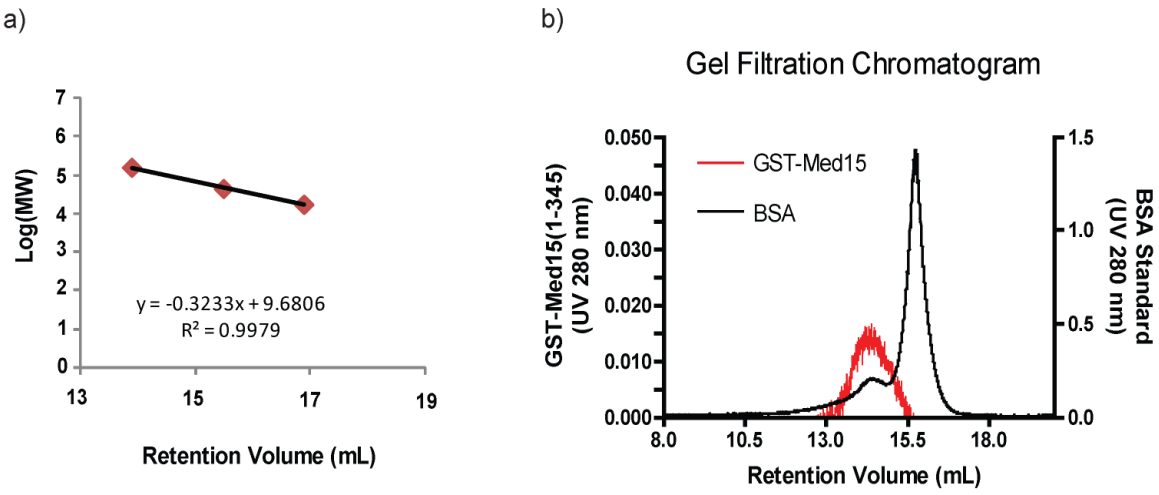


Figure 2.21: Analytical gel filtration data for GST-Med15(1-345). (a) Calibration curve for log (molecular weight) plotted against retention volume obtained with standards γ-globulin, ovalbumin and myoglobin. (b) Chromatogram of 1.5 μM (based upon a monomer) GST-Med15 (red) compared to chromatogram of BSA (black) (monomer (66.4 kD) and dimer (132.8 kD)).

Fluorescently-labeled Gal4 oligonucleotides

The 20 bp oligonucleotide 5'-TCC GGA GGA CTG TCC TCC GG-3' was purchased from Invitrogen with a 5'-modification of either a Fluorescein or BODIPY®FL fluorophore. The fluorescently labeled oligonucleotide was then annealed with an unlabeled complementary oligonucleotide 5'-GCC GGA GGA CAG TCC TCC GG-3' in annealing buffer (10 mM HEPES pH 7.5, 150 mM NaCl) by heat denaturation for 7 min at 95 ºC, followed by cooling at room temperature for 30 min and 4 ºC for 30 min [1].

Fluorescence polarization assays

Annealed 5'-Fluorescein-labeled, double-stranded oligonucleotides (45 μM) were diluted in DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β-mercaptoethanol, 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) [49] to a concentration of 1.25 nM. Then 200 μL of the DNA solution was added to a series of 50 μL solutions of varying activator concentrations in storage buffer A to obtain the final concentrations indicated in Figure 2.8 in a total volume of 250 μL. The samples were incubated for 30 min at room temperature before the degree of fluorescence polarization was measured on a Beacon 2000 instrument (PanVera Corp). The observed mP values were plotted as a function of activator concentration, and a binding isotherm that accounts for ligand depletion [66] (assuming a 1:1 binding model of dimeric activator to duplex DNA) was fit to the data using Origin 7.0 software to obtain the apparent equilibrium dissociation constant, K_d:

$$y = c + ((b - c)/(2a)) \times [(Kd + a + x) - \sqrt{(Kd + a + x)^2 - 4ax}] \quad (11)$$

where “a” and “x” are the total concentrations of duplex DNA and dimeric activator, respectively, “y” is the observed polarization at any activator concentration, “b” is the maximum observed polarization value, and “c” is the minimum observed polarization value. Each data point in Figure 2.8 is an average of three independent experiments with the indicated error (standard deviation).

Fluorescence titration assays

GST-Med15(1-345) was titrated (without exceeding a 5% volume increase) into the following solution to obtain the final Med15 concentrations indicated in Figure 2.11: 25 nM 5’-BODIPY®FL-labeled, double-stranded oligonucleotides pre-complexed to 100 nM dimeric activator in DNA-binding buffer, from which the DNA is estimated to be 84% bound according to simulations studies performed with Equation 11 (Figure 2.10). The fluorescence intensity of BODIPY®FL was monitored on an Eclipse spectrofluorometer (Varian Corp) (λ_{ex} = 500 nm, λ_{em} = 512 nm; 5 nm bandpass). The fluorescence intensity (F_i) was corrected for dilution effects, then normalized to that of the DNA•activator complex (F_0), such that $\Delta F = F_i - F_0$ [67]. The observed relative fluorescence enhancement, ΔF, was plotted as a function of Med15 concentration, and a binding isotherm that accounts for ligand depletion (equation 11) (assuming a 1:1 binding model of DNA•activator complex to dimeric GST-Med15) was fit to the data using Origin 7.0 software to obtain the apparent equilibrium dissociation constant, K_d. The parameters “a” and “x” represent the total concentrations of DNA•activator complex and dimeric Med15, respectively, “y” is the observed relative fluorescence enhancement at any Med15 concentration, “b” is the maximum observed relative fluorescence

enhancement value, and "c" is the minimum observed relative fluorescence enhancement value. Each data point in Figure 2.11 is an average of three independent experiments with the indicated error (standard deviation).

Fluorescence stopped-flow kinetic experiments

Stopped-flow experiments were performed on a KinTek model SF-2001 stopped-flow equipped with 75W Xe arc lamp in two-syringe mode. BODIPY®FL was excited at 502 nm and its emission was monitored at wavelengths > 510 nm using a long-pass filter. All kinetic traces reported are an average of four to six independent determinations. Sum of exponentials was fit to the transient kinetic time courses, F(t), to obtain the fluorescence amplitude (A) and the observed rate, k_{obs}, for each exponential phase where F(0) is the initial fluorescence intensity, and t, time:

$$F(t) = \sum A_n[1 - \exp(-k_{obs,n}t)] + F(0) \qquad (12)$$

Data fitting was performed using Origin 7.0 software, and the reported errors are the asymptotic standard errors. The dependence of the observed rates on Med15 concentration was plotted using GraphPad Prism 4.0 software. Each kinetic constant reported in Table 2.2 is an average of two independent experiments with propagation of the experimental error.

Calculation of the microscopic kinetic rate and equilibrium constants for the two-step binding models

Scheme A:

The full solutions to the rate equations describing the binding mechanism presented in Scheme A (Figure 2.18a), in which a conformational change occurs *after* the bimolecular collision step, have been published elsewhere [44]. However, approximate rates for $k_{obs,1}$ and $k_{obs,2}$ can be obtained from equations 1 (Figure 2.18b) and 13, respectively [44].

$$k_{obs,2} \approx \frac{k_1[Med15](k_2 + k_{-2}) + k_{-1}k_{-2}}{k_1[Med15] + k_{-1} + k_2 + k_{-2}} \quad (13)$$

The experimentally obtained values for $K_{d,app}$, k_{on}, $k_{obs,1}^{(y\text{-intercept})}$ and $k_{obs,2}$ summarized in Table 2.2 were then used to calculate the microscopic rate (k_1, k_{-1}, k_2, k_{-2}) and equilibrium (K_1 and K_2) constants as follows, according to equations (1)-(5) presented in Figure 2.18b:

Equation 1 describes the linear dependence of $k_{obs,1}$ on the concentration of Med15 (Figure 2.17), in which the slope of the line defines the microscopic rate constant k_1, and the y-intercept of the line reflects a composite value for k_{-1}, k_2, and k_{-2}. Furthermore, equation 13 defines the maximum rate of $k_{obs,2}$ (Figure 2.17) (when the concentration of Med15 is infinitely high) as the sum of the microscopic rate constants for the conformational change step, k_2 and k_{-2} (Equation 2). Therefore, subtraction of the average value of $k_{obs,2}$ from the $k_{obs,1}^{y\text{-intercept}}$ allows for an estimation of the microscopic rate constant k_{-1}. Using these values for k_1 and k_{-1}, the microscopic equilibrium constant for the bimolecular binding step, K_1, can be calculated according to Equation 3. The microscopic equilibrium constant for the conformational change step, K_2, can then be calculated using the values for K_1 and $K_{d,app}$ according to Equation 5. Finally, the solution

to the system of simultaneous equations composed of Equations 2 and 4 provides values for the microscopic rate constants k_2 and k_{-2}. The calculated values for the microscopic rate constants k_1, k_{-1}, k_2, and k_{-2} for Scheme A are summarized in Table 2.3, with propagation of the experimental error.

Scheme B:

For the binding mechanism presented in Scheme B (Figure 2.18a), in which a conformational change occurs *before* the bimolecular collision step, approximate rates for $k_{obs,1}$ and $k_{obs,2}$ can be obtained from equations 6 (Figure 2.18c) and 14, respectively [28].

$$k_{obs,2} \approx \frac{k_1(k_{-2} + k_2[Med15]) + k_{-1}k_{-2}}{k_1 + k_{-1} + k_2[Med15] + k_{-2}} \quad (14)$$

The experimentally obtained values for $K_{d,app}$, k_{on}, $k_{obs,1}^{(y\text{-intercept})}$ and $k_{obs,2}$ summarized in Table 2.2 were then used to calculate the microscopic rate (k_1, k_{-1}, k_2, k_{-2}) and equilibrium (K_1 and K_2) constants as follows, according to equations (6)-(10) presented in Figure 2.18c:

Equation 6 describes the linear dependence of $k_{obs,1}$ on the concentration of Med15 (Figure 2.17), in which the slope of the line defines the microscopic rate constant k_2, and the y-intercept of the line reflects a composite value for k_1, k_{-1}, and k_{-2}. Furthermore, equation 14 defines the maximum rate of $k_{obs,2}$ (Figure 2.17) (when the concentration of Med15 is infinitely high) as the microscopic rate constant k_1 (Equation 7). Therefore, subtraction of k_1 from the $k_{obs,1}^{y\text{-intercept}}$ allows for an estimation of $k_{-1} + k_{-2}$, whose values

can be calculated using Equation 10 to construct a system of simultaneous equations. The calculated values for the microscopic rate constants k_1, k_{-1}, k_2, and k_{-2} for Scheme B are summarized in Table 2.4, with propagation of the experimental error.

Berkeley Madonna

Global-fit analysis:

Experimental time courses obtained by stopped-flow spectroscopy for each of the three activators binding to Med15 were imported into the Berkeley Madonna software and fit to the mechanism shown in Scheme A (Figure 2.22). Only the change of fluorescent intensity ("In") between different species (a, b and c) were varied during the fit, while all other parameters were held as fixed parameters (Table 2.3). Below are shown the equations used to fit the Gal4 data, where:

DG = DNA•Gal4
M = Med15
DGM = DNA•Gal4•Med15
DGM1 = DNA•Gal4•Med15'
In = total fluorescent intensity; The resulting values of a, b and c were similar among varying Med15 concentrations for each activator.

```
METHOD RK4
STARTTIME = 0
STOPTIME = 5
DT = 0.002

In=a*DG+b*DGM+c*DGM1
a=1.2e8
b=1.3e8
c=1.4e8
```

```
RXN1 = K1f*DG*M - K1r*DGM
K1f =0.91e+07
K1r = 3.0
INIT DG = 2.5e-08
INIT DGM = 0
INIT M = 5e-07
d/dt(DG) = -RXN1
d/dt(DGM) = +RXN1-RXN2
d/dt(M) = -RXN1
```

```
RXN2 = K2f*DGM - K2r*DGM1
K2f = 0.32
K2r = 0.23
INIT DGM1 = 0
d/dt(DGM1) = +RXN2
```

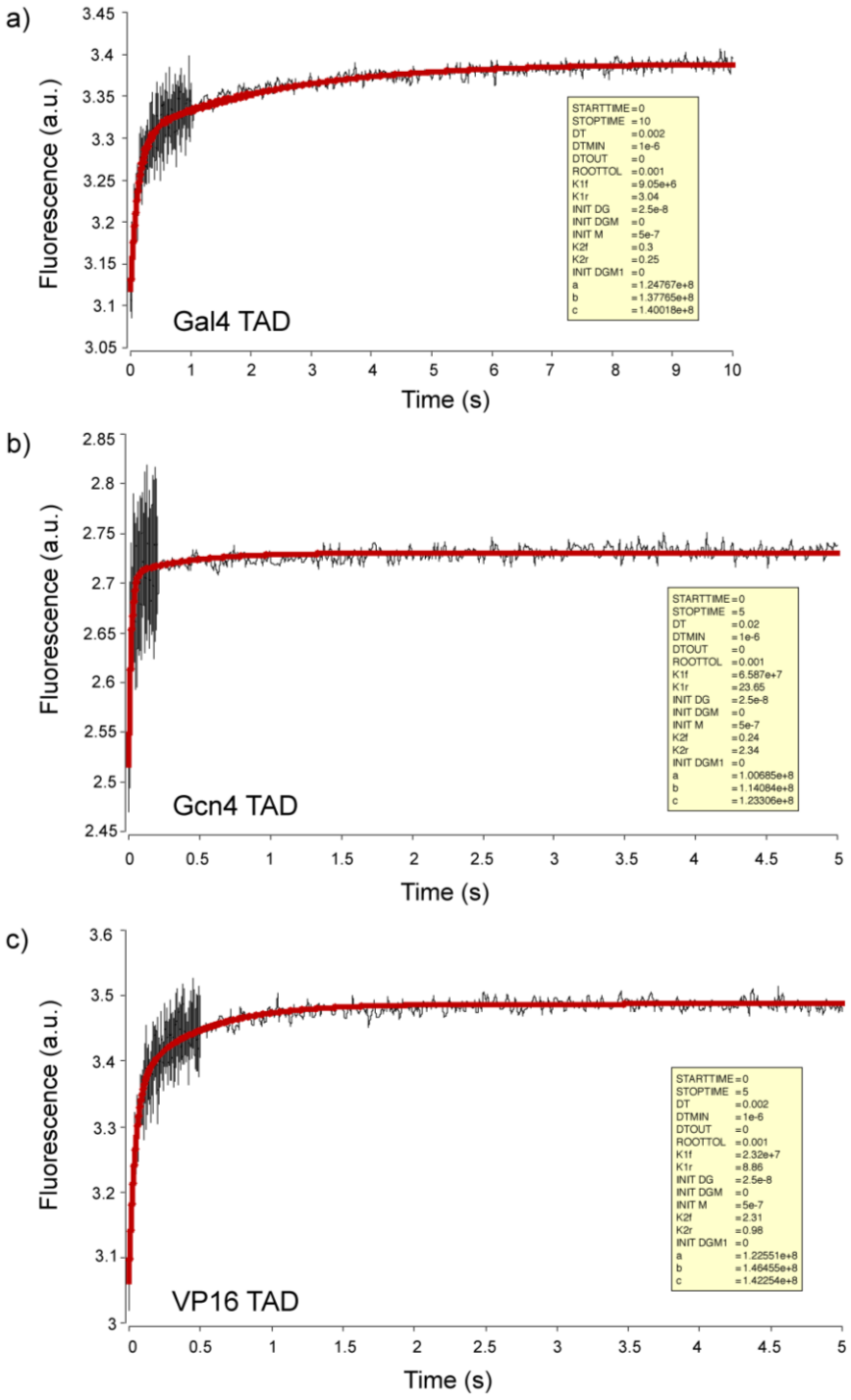


Figure 2.22: Global fit of the microscopic rate constants for Scheme A to experimental time courses. Global fit of time courses (red line) for binding of each of the three DNA-bound TADs to 0.50 μM Med15 according to Scheme A. Experimental data were fit using Berkeley Madonna software.

Simulations:

Simulated time courses were then generated by numerical integration without approximations using Berkeley Madonna software. To achieve this, chemical reactions were entered for either Scheme A or B shown in Figure 2.18a, while the following parameters were set as fixed values: 1) the microscopic rate constants k_1, k_{-1}, k_2, k_{-2} (shown in Tables 2.3 and 2.4 for Schemes A and B, respectively), 2) initial concentrations for both the DNA•activator complex (25 nM) and the varying Med15 concentrations, and 3) an initial fluorescence intensity value for the DNA•activator complex obtained previously by global-fit analysis. These simulated curves were then fit by non-linear regression to a double exponential (equation 12) to obtain values for the observed rates $k_{obs,1}$ and $k_{obs,2}$ (Figure 2.23) and amplitude terms (Figures 2.24 – 2.26) for comparison with the experimental data.

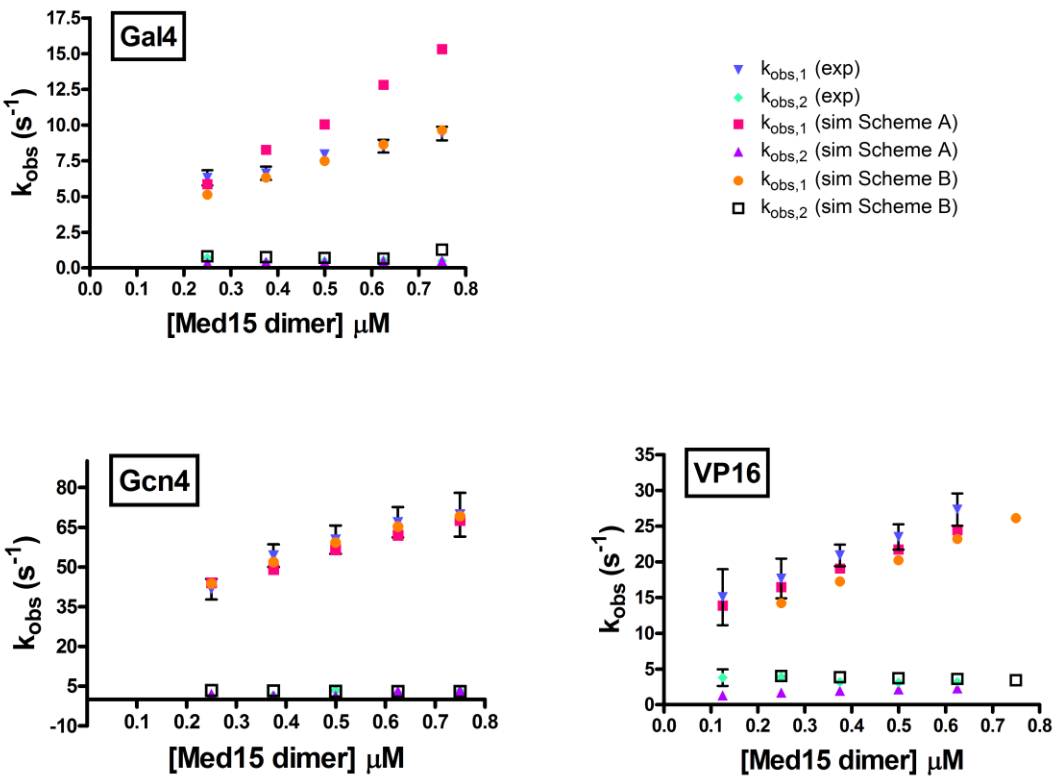


Figure 2.23: Simulated observed rates (k_{obs}) according to Schemes A and B. Simulated time courses for the three DNA•activator complexes binding to various Med15 concentrations according to Schemes A and B both possess observed rates ($k_{obs,1}$ and $k_{obs,2}$) that are consistent with the experimental data. The values for $k_{obs,1}$ show a linear dependence on Med15 concentration, while the values for $k_{obs,2}$ show no dependence on Med15 concentration.

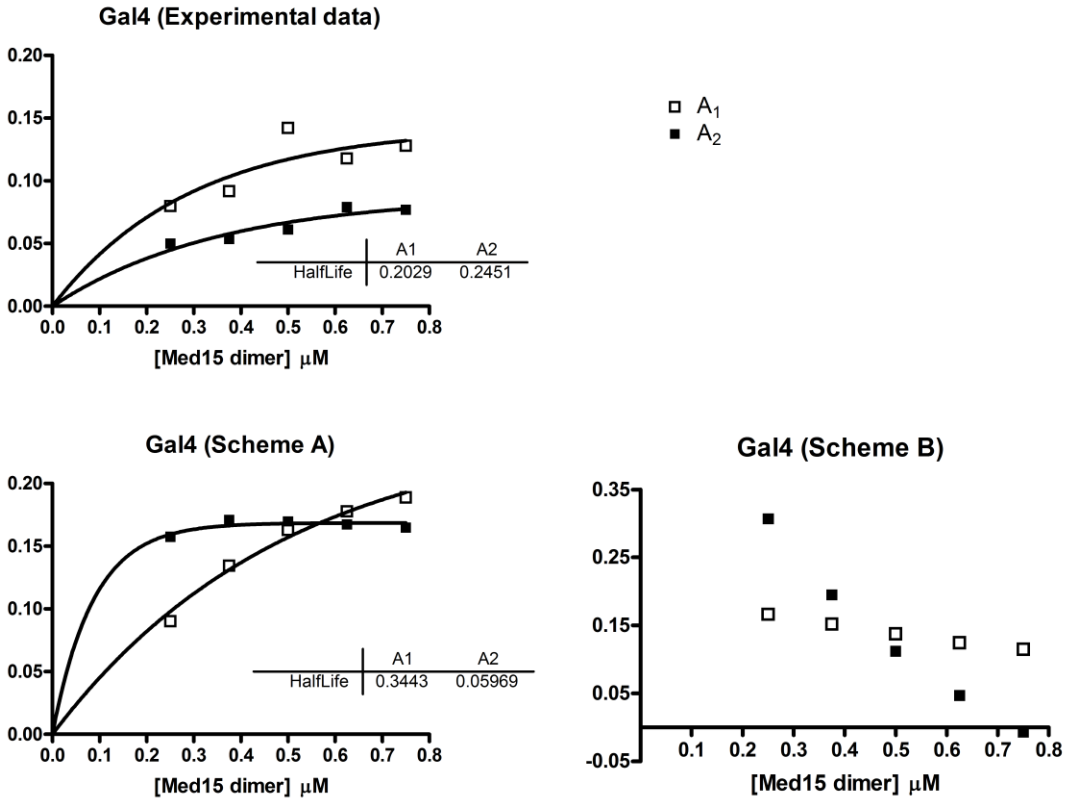


Figure 2.24: Simulated amplitude terms according to Schemes A and B for the Gal4-derived activator data. Simulated time courses for DNA-bound Gal4(1-100)-Gal4(840-881) binding to various Med15 concentrations according to Scheme A possess amplitude terms (A_1 and A_2) that better agree with what is observed experimentally than that of Scheme B.

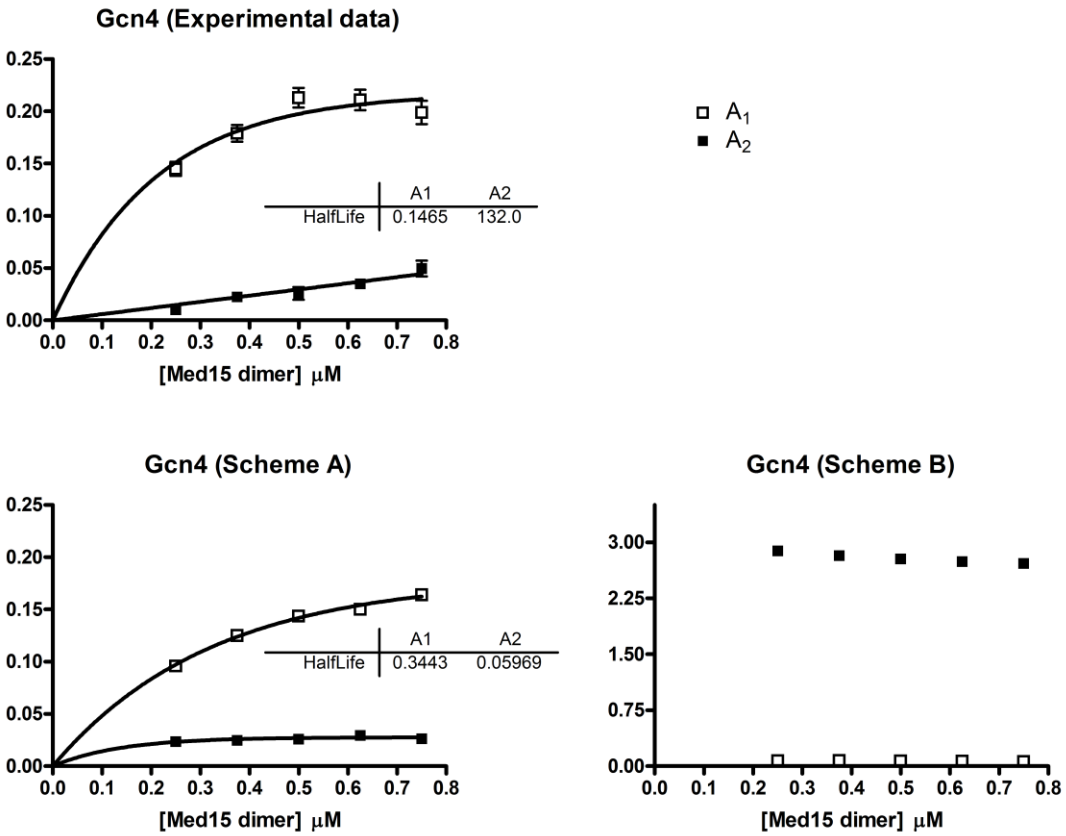


Figure 2.25: Simulated amplitude terms according to Schemes A and B for the Gcn4-derived activator data. Simulated time courses for DNA-bound Gal4(1-100)-Gcn4(107-144) binding to various Med15 concentrations according to Scheme A possess amplitude terms (A_1 and A_2) that better agree with what is observed experimentally than that of Scheme B.

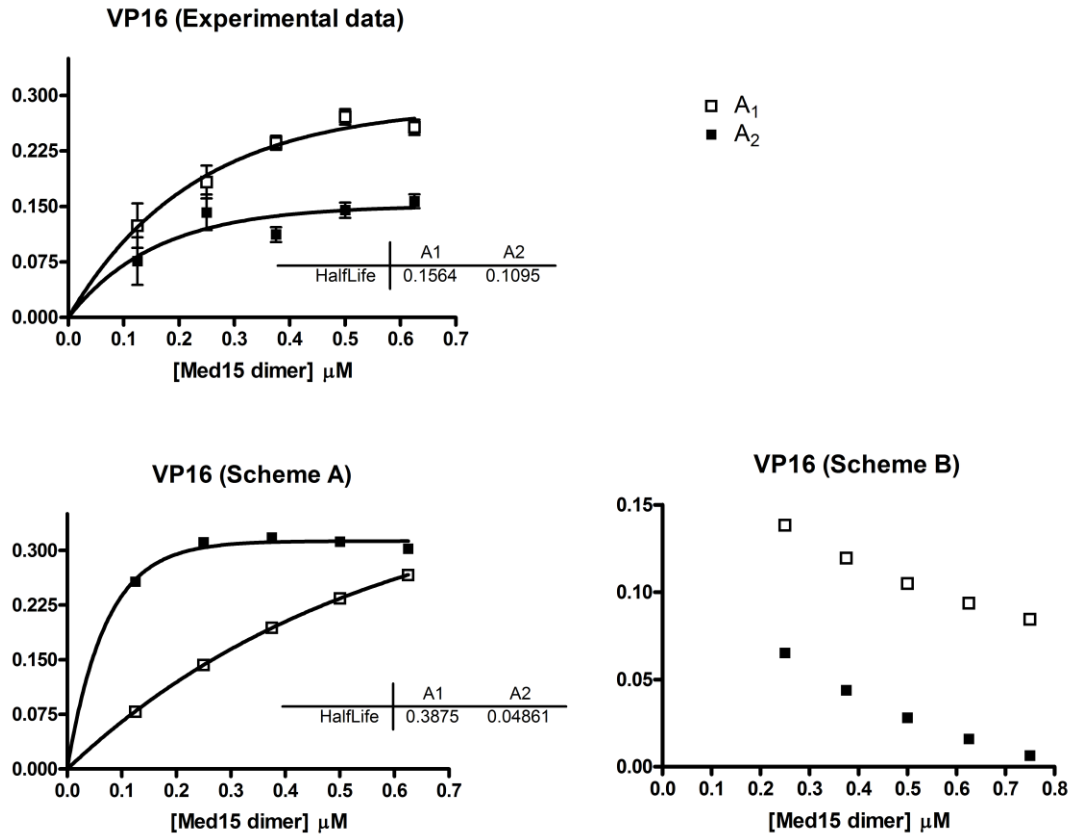


Figure 2.26: Simulated amplitude terms according to Schemes A and B for the VP16-derived activator data. Simulated time courses for DNA-bound Gal4(1-100)-VP16(456-490) binding to various Med15 concentrations according to Scheme A possess amplitude terms (A_1 and A_2) that better agree with what is observed experimentally than that of Scheme B.

H. References

1. Liang, S.D., et al., *DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA.* Mol Cell Biol, 1996. **16**(7): p. 3773-80.
2. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
3. Schmitz, M.L., et al., *Structural and functional analysis of the NF-kappa B p65 C terminus. An acidic and modular transactivation domain with the potential to adopt an alpha-helical conformation.* J Biol Chem, 1994. **269**(41): p. 25613-20.
4. Ferreira, M.E., et al., *Mechanism of transcription factor recruitment by acidic activators.* J Biol Chem, 2005. **280**(23): p. 21779-84.
5. Sugase, K., H.J. Dyson, and P.E. Wright, *Mechanism of coupled folding and binding of an intrinsically disordered protein.* Nature, 2007. **447**(7147): p. 1021-5.
6. Donaldson, L. and J.P. Capone, *Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1.* J Biol Chem, 1992. **267**(3): p. 1411-4.
7. Ansari, A.Z., R.J. Reece, and M. Ptashne, *A transcriptional activating region with two contrasting modes of protein interaction.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13543-8.
8. Wells, M., et al., *Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5762-7.
9. Ptashne, M. and A. Gann, *Genes & Signals*. 2001, New York: Cold Spring Harbor Laboratory.
10. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
11. Fuxreiter, M., et al., *Malleable machines take shape in eukaryotic transcriptional regulation.* Nat Chem Biol, 2008. **4**(12): p. 728-37.
12. Uesugi, M., et al., *Induced alpha helix in the VP16 activation domain upon binding to a human TAF.* Science, 1997. **277**(5330): p. 1310-3.
13. Thoden, J.B., et al., *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. **283**(44): p. 30266-72.
14. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. **274**(5289): p. 948-53.
15. Chang, J., et al., *Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein.* J Biol Chem, 1995. **270**(42): p. 25014-9.
16. Langlois, C., et al., *NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53.* J Am Chem Soc, 2008. **130**(32): p. 10596-604.
17. Kumar, S., et al., *Folding and binding cascades: dynamic landscapes and population shifts.* Protein Sci, 2000. **9**(1): p. 10-9.

18. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Curr Opin Struct Biol, 2009. **19**(1): p. 31-8.
19. Boehr, D.D., R. Nussinov, and P.E. Wright, *The role of dynamic conformational ensembles in biomolecular recognition.* Nat Chem Biol, 2009. **5**(11): p. 789-96.
20. Demchenko, A.P., *Recognition between flexible protein molecules: induced and assisted folding.* J Mol Recognit, 2001. **14**(1): p. 42-61.
21. Kanelis, V., J.D. Forman-Kay, and L.E. Kay, *Multidimensional NMR methods for protein structure determination.* IUBMB Life, 2001. **52**(6): p. 291-302.
22. Wuthrich, K., *Protein structure determination in solution by NMR spectroscopy.* J Biol Chem, 1990. **265**(36): p. 22059-62.
23. Baldwin, A.J. and L.E. Kay, *NMR spectroscopy brings invisible protein states into focus.* Nat Chem Biol, 2009. **5**(11): p. 808-14.
24. Korzhnev, D.M. and L.E. Kay, *Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding.* Acc Chem Res, 2008. **41**(3): p. 442-51.
25. Kim, D.H., et al., *Multiple hTAF(II)31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16.* BMB Rep, 2009. **42**(7): p. 411-7.
26. Fuxreiter, M., et al., *Preformed structural elements feature in partner recognition by intrinsically unstructured proteins.* J Mol Biol, 2004. **338**(5): p. 1015-26.
27. Radhakrishnan, I., et al., *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions.* Cell, 1997. **91**(6): p. 741-52.
28. Johnson, K.A., *Transient-state kinetic analysis of enzyme reaction pathways.*, in *The enzymes*, D.S. Sigman, Editor. 1992, Academic: New York. p. 1-61.
29. Malmqvist, M. and R. Karlsson, *Biomolecular interaction analysis: affinity biosensor technologies for functional analysis of proteins.* Curr Opin Chem Biol, 1997. **1**(3): p. 378-83.
30. van der Merwe, P.A., *Surface Plasmon Resonance*, in *Protein-Ligand interactions: hydrodynamics and calorimetry*, S. Harding and B.Z. Chowdhry, Editors. 2001, Oxford University Press: New York. p. 137-170.
31. *BIAevaluation Software Handbook*. 1997.
32. Paal, K., P.A. Baeuerle, and M.L. Schmitz, *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): p. 1050-5.
33. Wu, Y., R.J. Reece, and M. Ptashne, *Quantitation of putative activator-target affinities predicts transcriptional activating potentials.* EMBO J, 1996. **15**(15): p. 3951-63.
34. Morton, T.A., D.G. Myszka, and I.M. Chaiken, *Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration.* Anal Biochem, 1995. **227**(1): p. 176-85.
35. O'Shannessy, D.J. and D.J. Winzor, *Interpretation of deviations from pseudo-first-order kinetic behavior in the characterization of ligand binding by biosensor technology.* Anal Biochem, 1996. **236**(2): p. 275-83.

36. Hermann, S., K.D. Berndt, and A.P. Wright, *How transcriptional activators bind target proteins.* J Biol Chem, 2001. **276**(43): p. 40127-32.
37. Lavery, D.N. and I.J. McEwan, *Functional characterization of the native NH2-terminal transactivation domain of the human androgen receptor: binding kinetics for interactions with TFIIF and SRC-1a.* Biochemistry, 2008. **47**(11): p. 3352-9.
38. Lavery, D.N. and I.J. McEwan, *Structural characterization of the native NH2-terminal transactivation domain of the human androgen receptor: a collapsed disordered conformation underlies structural plasticity and protein-induced folding.* Biochemistry, 2008. **47**(11): p. 3360-9.
39. Gestwicki, J.E., H.V. Hsieh, and J.B. Pitner, *Using receptor conformational change to detect low molecular weight analytes by surface plasmon resonance.* Anal Chem, 2001. **73**(23): p. 5732-7.
40. Phillips, K.S. and Q. Cheng, *Recent advances in surface plasmon resonance based techniques for bioanalysis.* Anal Bioanal Chem, 2007. **387**(5): p. 1831-40.
41. Rich, R.L. and D.G. Myszka, *Survey of the year 2006 commercial optical biosensor literature.* J Mol Recognit, 2007. **20**(5): p. 300-66.
42. Fierke, C.A. and G.G. Hammes, *Transient kinetic approaches to enzyme mechanisms.* Methods Enzymol, 1995. **249**: p. 3-37.
43. Johnson, K.A., *Advances in transient-state kinetics.* Curr Opin Biotechnol, 1998. **9**(1): p. 87-9.
44. Johnson, K.A., *Rapid kinetic analysis of mechanochemical adenosinetriphosphatases.* Methods Enzymol, 1986. **134**: p. 677-705.
45. Hong, M., et al., *Structural basis for dimerization in DNA recognition by Gal4.* Structure, 2008. **16**(7): p. 1019-26.
46. Zhang, F., et al., *A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p.* Mol Cell Biol, 2004. **24**(15): p. 6871-86.
47. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. **25**(20): p. 9092-102.
48. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. **18**(3): p. 369-78.
49. Jeong, C.J., et al., *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. **40**(31): p. 9421-7.
50. Park, J.M., et al., *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. **20**(23): p. 8709-19.
51. Majmudar, C.Y., et al., *A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry.* Angew Chem Int Ed Engl, 2009. **48**(38): p. 7021-4.
52. Northrup, S.H. and H.P. Erickson, *Kinetics of protein-protein association explained by Brownian dynamics computer simulation.* Proc Natl Acad Sci U S A, 1992. **89**(8): p. 3338-42.
53. Tsai, Y.C. and K.A. Johnson, *A new paradigm for DNA polymerase specificity.* Biochemistry, 2006. **45**(32): p. 9675-87.

54. Donlin, M.J., S.S. Patel, and K.A. Johnson, *Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction.* Biochemistry, 1991. **30**(2): p. 538-46.
55. Schreiber, G., *Kinetic studies of protein-protein interactions.* Curr Opin Struct Biol, 2002. **12**(1): p. 41-7.
56. Lum, J.K., et al., *Converting inactive peptides into potent transcriptional activators.* ACS Chem Biol, 2006. **1**(10): p. 639-43.
57. Plasmids described were constructed using standard molecular biology techniques.
58. Donnelly, M.I., et al., *An expression vector tailored for large-scale, high-throughput purification of recombinant proteins.* Protein Expr Purif, 2006. **47**(2): p. 446-54.
59. Lum, J.K., *Peptides as tools to probe transcriptional activator function.*, in *Medicinal Chemistry*. 2007, University of Michigan: Ann Arbor.
60. Burk, D., D. Dawson, and T. Stearns, *Methods in Yeast Genetics*. 1 ed. 2000, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
61. Reece, R.J., R.J. Rickles, and M. Ptashne, *Overproduction and single-step purification of GAL4 fusion proteins from Escherichia coli.* Gene, 1993. **126**(1): p. 105-7.
62. Ferdous, A., et al., *The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II.* Mol Cell, 2001. **7**(5): p. 981-91.
63. Archer, C.T., et al., *Physical and functional interactions of monoubiquitylated transactivators with the proteasome.* J Biol Chem, 2008. **283**(31): p. 21789-98.
64. Smith, D.B. and K.S. Johnson, *Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.* Gene, 1988. **67**(1): p. 31-40.
65. Hidalgo, P., et al., *Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain.* Genes Dev, 2001. **15**(8): p. 1007-20.
66. Volkman, H.M., S.E. Rutledge, and A. Schepartz, *Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain.* J Am Chem Soc, 2005. **127**(13): p. 4649-58.
67. Lohman, T.M. and D.P. Mascotti, *Nonspecific ligand-DNA equilibrium binding parameters determined by fluorescence methods.* Methods Enzymol, 1992. **212**: p. 424-58.

CHAPTER 3

ISOLATION OF PEPTIDIC LIGANDS THAT TARGET SURFACES ON DNA-BINDING PROTEINS

A. Introduction

Diseased cells possess different transcription profiles relative to their normal counterparts, spurring considerable interest in the discovery of molecules that can correct these errant transcription patterns for use as mechanistic tools and as therapeutic agents [1-5]. One significant challenge in the development of such artificial transcription factors (ATFs) is a lack of small molecules that can be used to localize them to a gene promoter in a cellular context [5]. We propose a novel approach to accomplish this task which utilizes ligands that can interact with endogenous DNA-bound proteins. To this end, we have used an *in vitro* phage display selection with a random 12 amino acid peptide library to isolate ligands that are capable of interacting with surfaces on DNA-binding proteins. Future studies will entail the implementation of a selection using a conformation-constrained peptide library to obtain ligands that may possess increased stability and specificity within the cellular milieu. In particular, protein scaffolds that promote helix stabilization within the ligand would aid in the future identification of their peptidomimetic or small molecule replacements.

B. Artificial DNA-binding domains (DBDs)

The function of a natural transcription factor can be reconstituted minimally by an ATF that contains a DNA-binding domain (DBD) and a regulatory domain (i.e., an activation or repression domain), which are tethered through either covalent or non-covalent interactions [6-7]. In practice, ATFs that utilize protein DBDs have been the most successfully utilized in a cellular context. For instance, zinc finger (ZF) proteins (Chapter 1) composed of 3–6 fingers with apparent dissociation constants in the picomolar to nanomolar range for DNA sequences 9-18 bp in size, respectively, have been implemented in activator and repressor ATFs that function in mammalian cell culture [8-11]. One such activator ATF which contains a six-finger DBD that binds within the γ-globin promoter can increase fetal hemoglobin levels 7–16 fold in human erythroleukemia cells [8]. Additionally, a repressor ATF engineered with a six-finger DBD that binds within the HIV-1 (5'LTR) promoter was able to inhibit viral replication by 75% in infected cells [11]. Furthermore, ZF-based activator ATFs have even been shown to function in a mouse model, in which they were able to upregulate expression of the endogenous VEGF-A gene (vascular endothelial growth factor) and induce angiogenesis. Despite these achievements, delivery still remains a challenge, as these exogenous proteins had to be administered by viral vectors [12].

The incorporation of small molecule DBDs into ATFs has the potential to bypass the need for delivery strategies altogether. To date, the most successful compounds that can be programmed to target a specific DNA sequence are the hairpin polyamides developed by Dervan and coworkers (Chapter 1) [13]. However, the size of the polyamide is the major determining factor in its cellular permeability, with large hairpins (10-ring) exhibiting poor cellular uptake and small hairpins (6-ring) exhibiting excellent cellular

uptake [14]. To this end, Kodadek and coworkers were able to successfully construct a cell-permeable activator ATF composed of an eight-ring hairpin polyamide (ImPy7) that is covalently linked to a hexameric peptoid (TBHK6) capable of binding to the core KIX domain of the coactivator CBP [15]. However, this polyamide-peptoid construct was able to upregulate expression of 45 endogenous genes at least 3-fold in mammalian cell culture due to the low specificity of its 6 bp recognition sequence (i.e., 5'-WGWWWW-3' (W = A or T)). Thus, an alternative approach may be needed in order to obtain cell-permeable molecules that can achieve DNA-binding specificities which rival that of the proteins they are trying to replace.

B.1. Hijacking ATFs

Rather than developing ligands that can directly target DNA, a novel approach for localizing an artificial transcriptional regulatory domain to the promoter of a gene is to utilize ligands that can target an endogenous DNA-bound protein, in order to "hijack" its exquisite DNA-binding specificity (Figure 3.1). In the remainder of this section, potential applications for this technique will be provided.

B.1.A. Activator ATFs

One application of this method is to localize a transcriptional activation domain to an endogenous transcriptional repressor. For instance, REST/NRSF is a global transcriptional repressor that binds to a 21-bp recognition sequence located within the regulatory regions of neuronal differentiation genes, and has been found to be overexpressed in the majority of human medulloblastoma tumors [16]. Majumder and

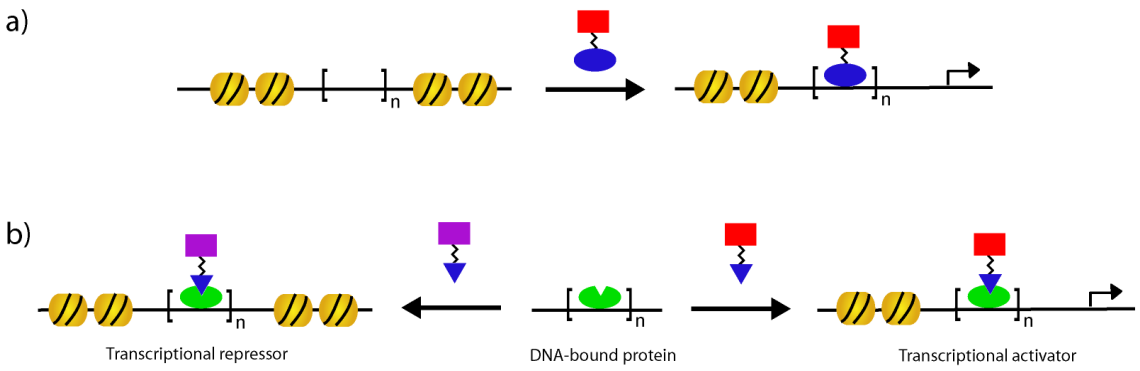


Figure 3.1: Schematic of hijacking a DNA-bound protein. a) A traditional ATF with a heterologous TAD (red square) and DBD (blue circle) can target available binding sites across all cell types. b) A peptidic ligand (blue triangle) that is capable of interacting with a DNA-bound protein (green circle), which is attached to either a transcriptional activation domain (red square) or repression domain (purple square), creates a new class of ATFs that can function only in those cell lines that possess the targeted protein.

coworkers have demonstrated that adenovirus-mediated expression of an activator ATF comprised of the DNA-binding domain (DBD) of REST/NRSF and the transcriptional activation domain (TAD) of VP16 (i.e., REST-VP16) within these medulloblastoma cell lines stimulated the expression of endogenous neuronal genes and triggered apoptosis through the activation of caspase cascades [17]; this, in turn, results in the inhibition of the growth of established tumors in nude mice, as well as an initial block of their intracranial tumorigenic potential upon implantation [16]. Furthermore, because adjacent normal cerebellum tissue sections were negative for REST/NRSF expression, targeting of the VP16 activation domain to endogenously bound REST using our proposed system would have the potential to constrict its function specifically to cancerous tissue within the brain.

B.1.B. Repressor ATFs

Another application would be to localize a transcriptional repression domain to an endogenous transcriptional activator. For instance, the cAMP-response element-binding protein (CREB) is a ubiquitously expressed transcriptional activator that binds to CREs found in the promoter regions of many cellular and viral genes. In particular, expression of human T-cell leukemia virus type 1 (HTLV-1) is controlled by the assembly of a Tax/CREB multiprotein complex; the interaction between Tax and CREB is highly specific, involving amino acid residues Ala-Ala-Arg at positions 282-284 (AAR) on CREB located immediately upstream of its DNA-binding domain [18]. The formation of this protein complex thus expands the DNA recognition properties of CREB to bind selectively on the 21-bp viral CRE repeat, with Tax binding to the minor groove of GC-rich sequences flanking both the 5' and 3' of the 8-bp CRE core. Therefore, targeting a transcriptional repression domain to this Tax/CREB complex with our proposed system would have potential anti-viral therapeutic applications that possess an increased specificity for infected cells over that of alternative inhibitive approaches. For example, Nyborg and Dervan designed hairpin polyamides to disrupt the interaction of Tax at each of the three non-identical viral CREs present within the HTLV-1 promoter [19]. However, those that effectively inhibited Tax-mediated transcription *in vitro* had no effect on virion production in cell culture (despite their ability to enter the nucleus of HTLV-1 infected T-cells); one possible explanation may be that because the targeted DNA sequences were only ~6 bp in length, the polyamides were sequestered instead by additional binding sites located within the cellular genome. Furthermore, the Tax/CREB complex mediates high-level transcription by recruitment of full-length CBP/p300 [20]. However, due to the fact that a global coactivator such as CBP/p300 is capable of

interacting with several different TADs (e.g., c-Jun, c-Myb, and p53), this is a significant obstacle for the identification of small-molecule inhibitor that is specific for this interaction [21].

C. Isolation of peptidic ligands that target surfaces on DNA-binding proteins

The screening of random peptide libraries is a powerful method for the discovery of protein-protein interaction surfaces, and molecular biology techniques have enabled the facile generation of libraries with high sequence diversity. Therefore, in this section the implementation of the yeast two-hybrid (Y2H) and phage display systems for the isolation of 12-16 amino acid peptides that bind to surfaces on DNA-binding proteins will be described. Overall, we found that an *in vivo* Y2H screen did not yield ligands that are capable of binding to these types of targets (possibly due to technical limitations of the method itself), but rather an *in vitro* phage display affinity selection was a much more effective technique.

C.1. Two-hybrid

Two-hybrid screening is a technique that is primarily used for protein interaction mapping under cellular conditions. To achieve this, a known bait gene is expressed as a fusion to a DNA-binding domain (DBD) and a library of prey genes are expressed as fusions to a transcriptional activation domain (TAD), which upon formation of a positive protein-protein interaction reconstitutes a transcriptional activator protein [22]. However, the two-hybrid system has also been amenable to the screening of peptidic libraries as well [23]. For instance, in the Matchmaker Random Peptide Library (Clontech) for two-

hybrid screening in yeast, random 16-residue peptides are expressed as C-terminal fusions to the Gal4 TAD and assayed for their ability to interact with a bait gene fused to the C-terminus of the Gal4 DBD by the upregulation of such integrated reporter genes as *HIS3* (encoding a protein required for histidine synthesis), *ADE2* (conferring a pink or red colony color to colonies growing on media low in adenine), and *LacZ* (encoding the enzyme β-galactosidase) (Figure 3.2) [24].

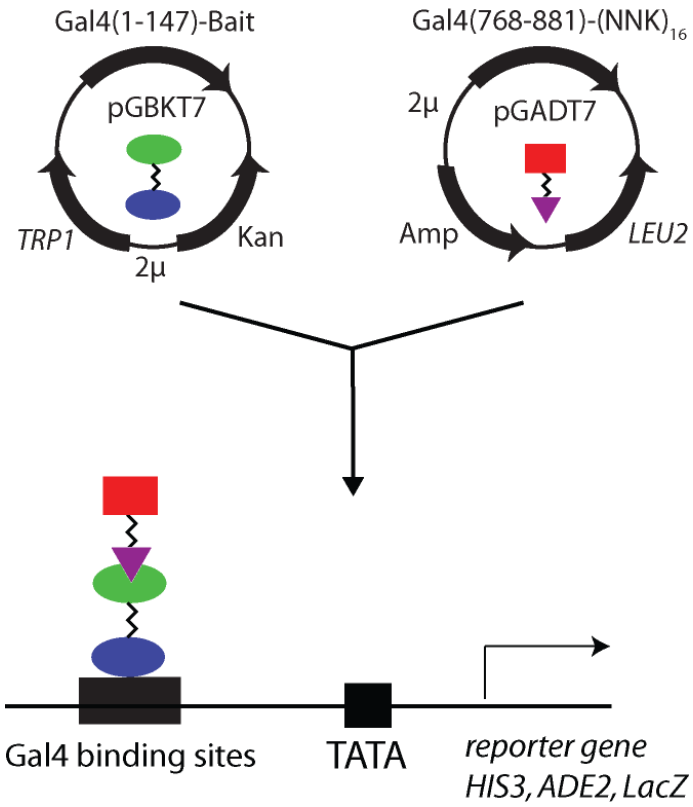


Figure 3.2: Schematic of the Matchmaker Random Peptide Library (Clontech) for two-hybrid screening in yeast. Plasmids encoding the Gal4(DBD)-bait fusion and Gal4(TAD)-peptide fusions are transformed into yeast. If a positive interaction occurs between the bait and prey, they will reconstitute a transcriptional activator and upregulate the *HIS3*, *ADE2*, and *LacZ* reporter genes.

The first report of utilization of this yeast two-hybrid (Y2H) random peptide library was by Fields and coworkers to screen for ligands that bind to the retinoblastoma tumor suppressor protein (Rb) [23]. From 3 million transformant clones that were screened (out of a library complexity of 10^7 sequences), seven $His^+/LacZ^+$ colonies were identified, each encoding peptides containing the Leu-X-Cys-X-Glu motif utilized by the various known protein binding partners of Rb. In particular, the crystal structure of the N-terminal region of the SV40 large T antigen binding to the pocket domain of Rb (Figure 3.3) illustrates that its Leu-X-Cys-X-Glu motif is encompassed within an extended loop (residues 103-117) that interacts with a shallow surface groove on the B box of Rb created by five helices [25]. This interface accounts for approximately two-thirds of the

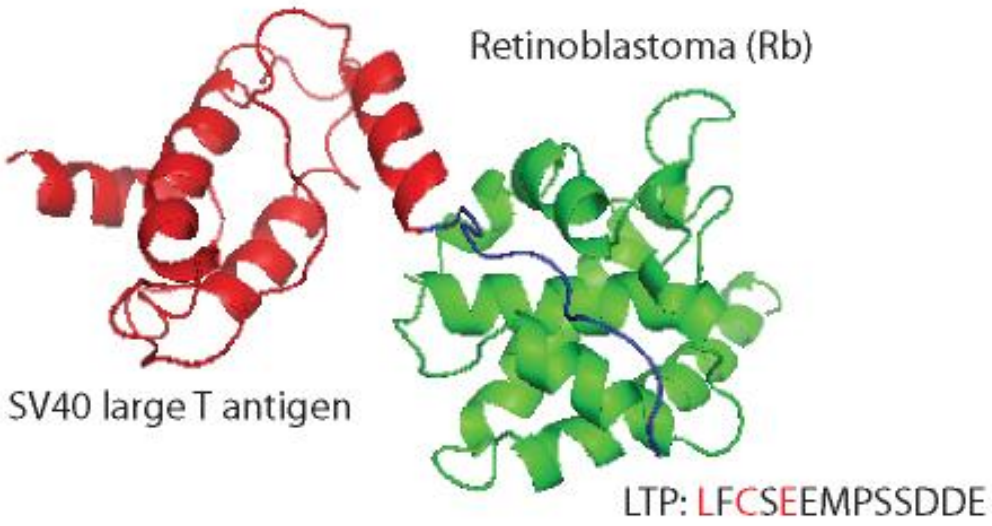


Figure 3.3: Retinoblastoma-SV40 large T antigen interaction. Ribbon diagram of the crystal structure of the Rb•SV40 large T antigen complex, in which the N-terminal region of the SV40 large T antigen (red) binds through an extended loop (blue, residues 103-117) to the B box of the pocket domain of Rb (green) (PDB ID: 1GH6). The Large T antigen Peptide (LTP) with its LxCxE motif was able to interact with a Gal4(DBD)-Rb fusion in a Y2H setting when attached to the C-terminus of the Gal4(TAD).

total buried surface area of the overall binding interaction between these two proteins, and is what the peptides are most likely mimicking; furthermore, one can easily envision replacement of the four-helix core domain (residues 7-102) with the Gal4 TAD being amenable to reconstitution of activator function. However, regardless of the fact that all seven peptides (P1-P7) possessed this conserved Leu-X-Cys-X-Glu motif, only P1 was shown to be as active in a β-galactosidase assay as that of the Large T antigen Peptide (LTP) derived from the SV40 large T antigen, in addition to being ~40-fold more active than the weakest peptide P7 [23]. Comparison of this activity data with affinity data shows that P1 binds to Rb via surface plasmon resonance (SPR) methods with a $K_{d,app}$ of 13 – 23 μM, a mutant peptide P1-a1 that is ~4-fold less active than P1 binds with a $K_{d,app}$ of 61 – 76 μM, and binding of P7 to Rb is undetectable. Overall, these results illustrate that using this approach, peptides sequences can be obtained that appear to target the same binding surface, but possess a range of activities and/or affinities.

In addition, Tanamoto and coworkers used this system to screen for peptides that are capable of interacting with the extracellular domain of the toll-like receptor 4 (TLR4), in order to discover novel inhibitors of the innate immune response induced by the endotoxin lipopolysaccharide (LPS) [26]. Six His^+/ADE^+ colonies were identified, whose peptide sequences possessed no significant homology to any known TLR4-associated proteins. Notably, one of these novel peptides, STM28, was capable of specifically inhibiting LPS-induced processes such as TNF-α production and NF-κB activation in macrophages, as well as protected mice from septic shock. This example illustrates the applicability of this method to the identification of novel ligand sequences.

Based on these precedents, we applied this method to directly screen for peptidic ligands that can localize a transcriptional activation domain to a DNA-binding protein *in vivo*. The bait genes chosen for screening with this method were: 1) The DNA-binding domain of the mammalian transcriptional repressor REST/NRSF, due to its potential therapeutic applications within medulloblastoma cell lines [16-17], and 2) the bacterial transcriptional activator SoxS, in order to obtain peptides that can reconstitute its interaction with the nonlinear recognition module from the αCTD of the bacterial RNA polymerase [27].

C.1.A. REST/NRSF

The transcriptional repressor REST/NRSF binds as a monomer to a 21-bp RE1/NRSE site through a central cluster of eight zinc fingers, whereby its N- and C- terminal repression domains can recruit corepressor complexes (Figure 3.4) [28]. For example, an N-terminal fragment of REST (residues 1-600) is able to repress Gal4-mediated expression of reporter genes from a $(RE1)_3$UAS-regulated promoter by recruitment of the Sin3 complex in yeast. However, removal of the N-terminal repression domain (residues 1-141) and the lysine-rich domain (residues 446-600) resulted in a complete loss of repression. Therefore, in our initial screen, rather than being forced to compete with this repression pathway, we decided to focus our efforts on the isolation of peptide ligands that are capable of interacting with surfaces on the central DNA-binding domain of REST. Furthermore, while no obvious structural motifs are present within the amino-terminal repressor domain of REST, the minimal domain (residues 38-57) consists mostly of hydrophobic amino acids that are present as an alpha-helix when bound to a

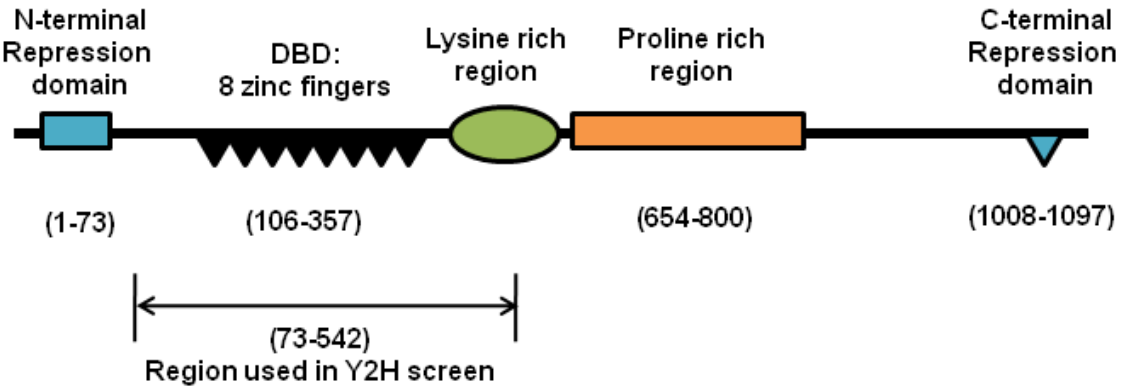


Figure 3.4: Map of the transcriptional repressor REST/NRSF. REST/NRSF possesses a central DNA-binding domain (DBD) composed of 8 zinc fingers, as well as N- and C-terminal repression domains.

hydrophobic cleft within the mSin3B PAH1 domain [29]; these characteristics mimic those observed within transcriptional activation domains. Thus, this domain is an unlikely target for short peptide ligands.

The Y2H screen of REST DBD with a random 16-residue peptide library was carried out according to the procedure outlined in Figure 3.5. The REST DBD (residues 73-542) [17, 30-31] was fused to the C-terminus of the Gal4 DBD. Its expression in the yeast strain AH109 was verified by a western blot, and it failed to upregulate expression of the *HIS3* reporter after 7 days (data not shown). The Matchmaker Random Peptide Library (Clontech) was sequentially transformed into yeast harboring the Gal4(DBD)-REST(73-542) expression plasmid with an efficiency of ~4,500 cfu/µg DNA (i.e., ~450,000 clones) and incubated at 30°C for 2 weeks. During this time, 23 yeast colonies that were able to upregulate the *HIS3* reporter were twice restreaked onto fresh media, 10 of which were also able to upregulate the *ADE2* reporter (white to light pink in color). The plasmids encoding the Gal4 TAD-peptide fusions were isolated from these 10 colonies and sequenced, the results of which are presented in Table 3.1.

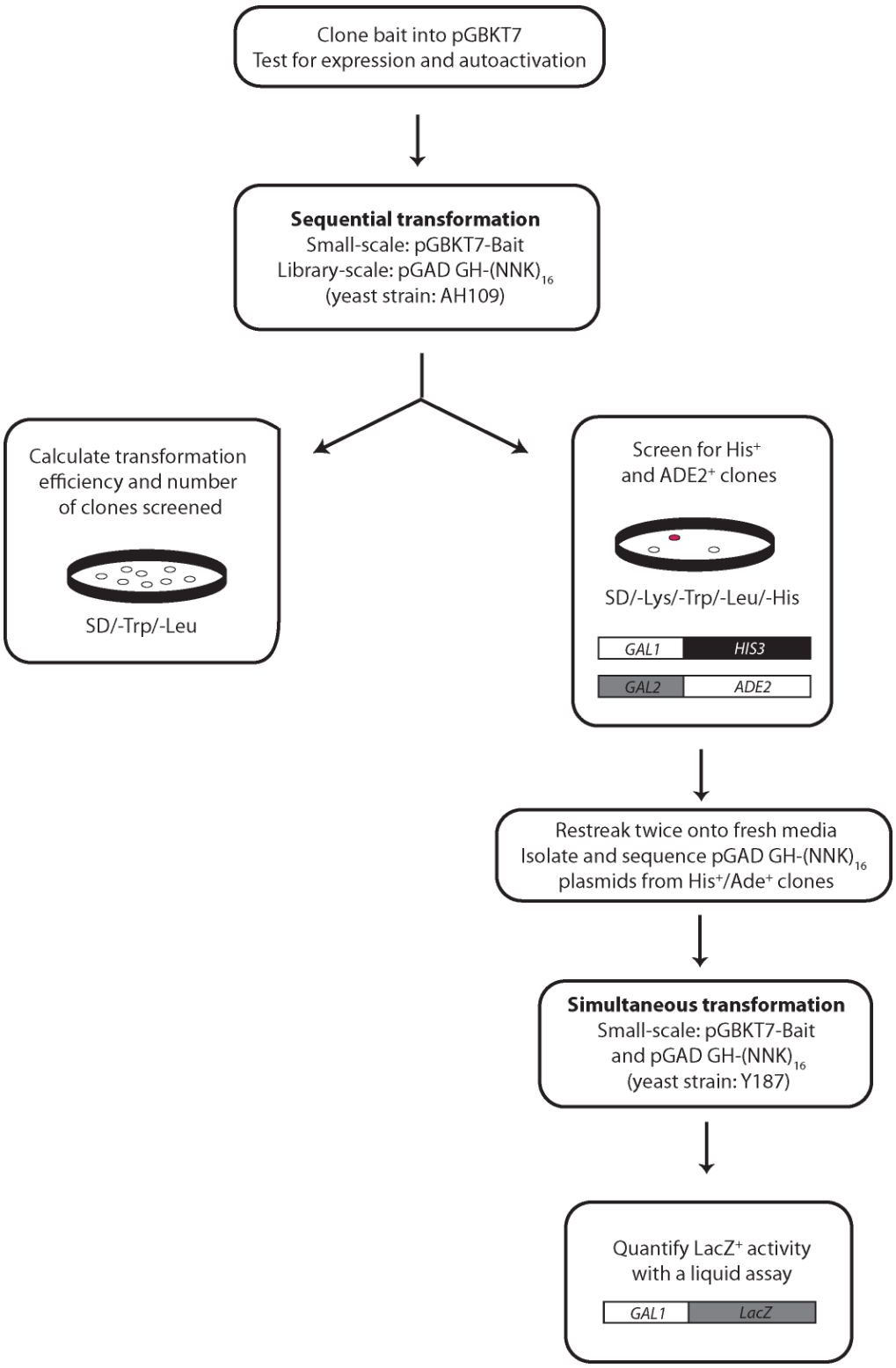


Figure 3.5: Flow-chart for the Y2H screening of a bait protein with the Matchmaker Random Peptide Library (Clontech).

Table 3.1: Peptide sequences identified in a Y2H screen of REST DBD

Batch.colony[a]	**Day**[b]	**Color**	**Sequence**[c]
15.1	5	White	GGGGGRVGRLCVRVCVGIPNQPHE
11.1	7	White	KRNMEKGQWVKGYSQD
15.2	7	Light pink	LVNDPEPMSAKRRFAA
15.3	9	White	GTSVWTSMRLLEYCFP
10.1	11	White	DEGGDKRGAGRLWGLR
6.2	13	Light pink	PSAVDFATTVCRCCGV or TRASGQRICCRVRGLSRNSYNTANTQVRSL
3.1	14	White	Not sequenced
16.1	14	White	Not sequenced
16.2	14	White	Not sequenced
19.2	14	White	Not sequenced

a: Transformants were screened in 20 batches of 5 plates. The colony number from each batch of plates was recorded.
b: The number of days of growth at 30°C until the colony was restreaked onto fresh media.
c: Four colonies for each TAD-peptide construct were sequenced; 50% of the colonies for construct 6.2 that were sequenced contained these peptides.

In order to verify the validity of these peptide-REST interactions, the Gal4(DBD)-REST(73-542) and Gal4(TAD)-peptide plasmids were transformed into the yeast strain Y187 to quantify their ability to upregulate the *LacZ* reporter gene. However, the peptides were unable to upregulate expression of β-galactosidase to detectable levels using either the substrates *o*-nitrophenyl-β-D-galactopyranoside (ONPG) or chlorophenol red-β-D-galactopyranoside (CPRG) (similar results with the peptides were also observed in the absence of bait protein, i.e. Gal4(DBD)), even though the interaction of Rb and P1 (which served as the positive control) was able to upregulate *LacZ* expression 12.7 ± 0.4 and 9.0 ± 0.2 β-gal units, respectively. Thus, a peptide sequence was not identified that is able to interact with the DNA-binding domain of REST/NRSF and upregulate expression of the *LacZ* reporter gene to levels rivaling that of our positive control. One possible explanation for these results may be that due to the high concentrations of the bait and prey proteins present during the initial screening process (200 copies per cell when using 2μ based plasmids), this may have allowed for the detection of weak bait-peptide

interactions that produce β-galactosidase levels that are below the sensitivity limits of the aforementioned substrates [32]. Rather, such interactions might be better suited for quantification via real-time PCR of the *LacZ* mRNA levels produced, or through the utilization of a more sensitive chemiluminescent substrate. In addition, due to limitations imposed by the transformation efficiency of the plasmids encoding the prey library into the yeast host, it is possible that not enough sequences were screened in order to identify one which possesses the chemical properties needed to mediate a high-affinity interaction with the bait protein. Thus, future efforts to repeat a Y2H screen on the REST DBD with the random 16-residue peptide library will focus initially on increasing the transformation efficiency through electroporation (rather than chemical) methods.

An alternative explanation for the inability of this Y2H screening method to identify peptide sequences that are capable of interacting with the REST DBD is that it may not possess a readily targetable binding site on its surface. In order to test this theory, the bacterial transcriptional activator SoxS was chosen as the next bait protein, which participates in a known binding interaction with the αCTD of the bacterial RNA polymerase [27].

C.1.B. SoxS

The bacterial transcriptional activator SoxS is a relatively small protein (only 107 amino acids) that is synthesized in *Escherichia coli* as a defense against oxidative stress [33]. It binds as a monomer to a degenerate 20 bp sequence, and interacts with the DNA binding determinant (a.k.a., the 265 determinant (residues 264-269 and 296-302)) of the αCTD of the RNA polymerase (Figure 3.6a-b) [27]. Furthermore, this protein-

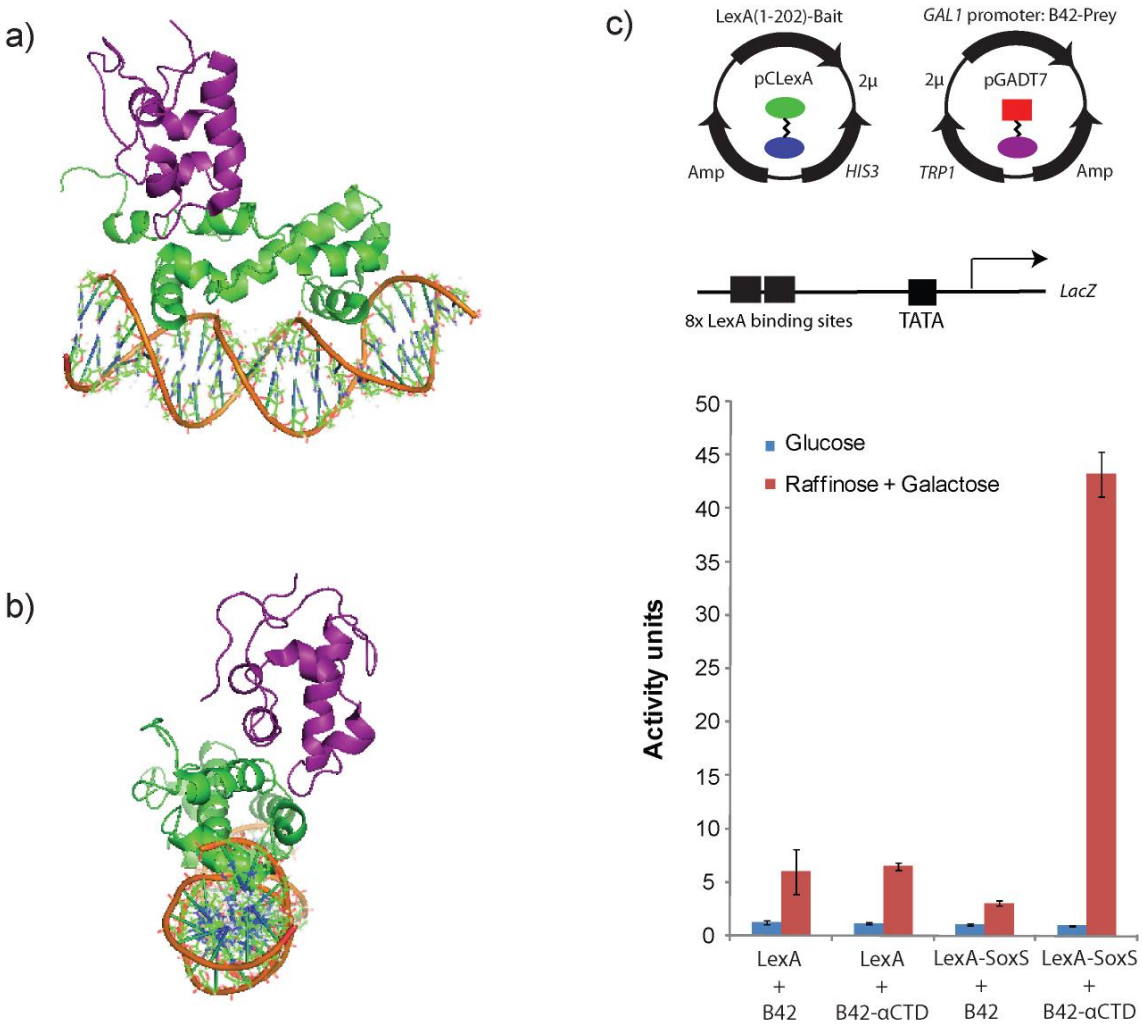


Figure 3.6: SoxS-αCTD interaction. a) Model of the DNA•MarA•αCTD ternary complex (PDB ID:1XS9). MarA (green), whose amino acid sequence is 42% identical to SoxS, interacts with the 265 DNA binding determinant of the αCTD of RNA polymerase (purple). b) Side view of a. c) Results from a liquid β-galactosidase assay for the Y2H interaction of SoxS and αCTD. The galactose-induced expression of the prey fusion B42-αCTD in yeast harboring the bait fusion LexA-SoxS results in the upregulation of a *LacZ* reporter gene that is under the control of LexA binding sites. Error is indicated as standard deviation of the mean (SDOM).

protein binding interaction has been observed in a Y2H setting; in particular, constitutively expressed SoxS fused to the C-terminus of the DNA-binding protein LexA (bait) is able to upregulate expression of a *LacZ* reporter gene upon the induced expression of the αCTD fused to the C-terminus of the B42 TAD (prey), as determined through a liquid β-galactosidase assay (Figure 3.6c) [33]. Therefore, this was an excellent model system in which to test the two-hybrid methodology for our purposes.

The Y2H screen of SoxS with a random 16-residue peptide library was carried out according to the procedure outlined in Figure 3.5. The Gal4(DBD)-SoxS fusion protein (bait) failed to upregulate expression of the *HIS3* reporter in the yeast strain AH109 after 2 weeks at the lowest concentration of 3-aminotriazole (3-AT) tested (2.5 mM). Additionally, since 3-AT is a competitive inhibitor of the *HIS3* gene product, its presence can also help increase the stringency against potential weak binding interactions that can result in false-positives. The Matchmaker Random Peptide Library (Clontech) was sequentially transformed into yeast harboring the Gal4(DBD)-SoxS expression plasmid with an efficiency of ~4,000 cfu/µg DNA (i.e., ~400,000 clones) and incubated at 30°C for 1 week. During this time, 9 yeast colonies were able to upregulate the *HIS3* reporter, but possessed an abnormal morphology and were unable to grow when restreaked onto fresh media. These results illustrate a further technical limitation of the system, in which a deleterious phenotypic response can prevent further manipulation of the microbial host.

C.1.C. Conclusions

Overall, screening DNA-binding proteins with a random peptide library via a yeast two-hybrid system may not be the optimal method for the isolation of ligands that can

bind to surfaces located on these targets. To a large extent, these results can be attributed to technical limitations of performing such a screen *in vivo*. As will be described in the next section, many of these issues can be addressed through the utilization of an *in vitro* phage display selection technique.

C.2. Phage display

Phage display is a technique in which a library of random proteins or peptides can be encoded within the genome of bacteriophage such that they are expressed on its surface as a fusion to a coat protein [34]. For instance, the Ph.D.-12 library (New England Biolabs) possesses a complexity of ~2.7 x 10^9 dodecapeptide sequences which are each expressed on the N-terminus of the pIII coat protein of an M13 phage (i.e., present as 5 copies per phage) (Figure 3.7a) [35]. This phage library can then be enriched for those sequences that are capable of binding to a target protein of interest through an *in vitro* selection process called panning (Figure 3.7b) [34-35]. Briefly, panning is carried out by first incubating the phage library with a target that is immobilized on a solid support, whereupon the unbound phage are washed away and the specifically-bound phage are eluted and amplified for another round. After 3-4 rounds of panning, the pool should be enriched in phage that bind to the target. Individual clones can then be sequenced and the identity of the proteins or peptides determined.

With regards to its applicability for identifying peptide sequences that can interact with DNA-binding proteins, phage display techniques using a random peptide library have been used to isolate conformational probes that can specifically recognize the different liganded states of the estrogen receptor (ER) [36]. For instance, many of the

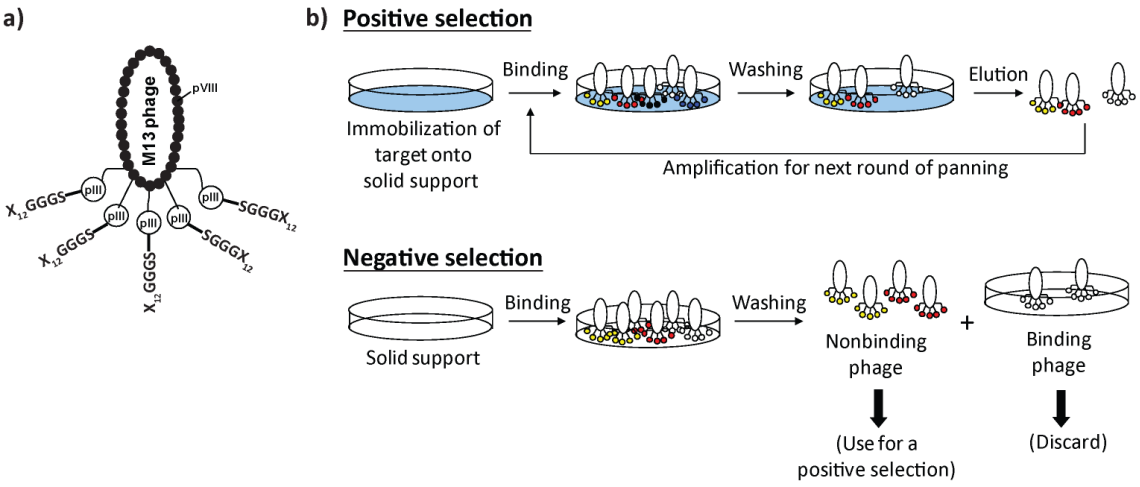


Figure 3.7: Random 12-mer phage display library. a) Schematic of an M13 bacteriophage engineered to express a random 12-mer peptide on the N-terminus of the pIII coat protein (5 copies per phage). b) Outline of panning. In a positive selection, a target that is immobilized on a solid support is incubated with the phage library, whereupon the unbound phage are washed away and the specifically-bound phage are eluted and amplified for another round. Peptide sequences that can interact with components other than the desired target can be removed by the incorporation of a negative selection after the first panning. In a negative selection, a selection is performed in the absence of the target, in which nonbinding phage are used for another round of panning.

sequences identified in the presence of estradiol were that of the LxxLL motif found within nuclear receptor coactivators; these sequences bind to a compact hydrophobic groove within the AF2 region of the ligand-binding domain (LBD) that is exposed upon agonist binding [37]. Furthermore, one peptide identified in the presence of both estradiol and tamoxifen called αII (LTSRDFGSWYA) binds in an AF2-independent manner to a unique surface on the LBD (with an apparent dissociation constant of 34 μM) [36, 38]. The αII peptide adopts an extended conformation that binds across a shallow hydrophobic depression on the β-hairpin face, whereby no single interaction appears to dominate (Figure 3.8); this example illustrates the type of binding site one would expect to find on the surfaces of DNA-binding proteins.

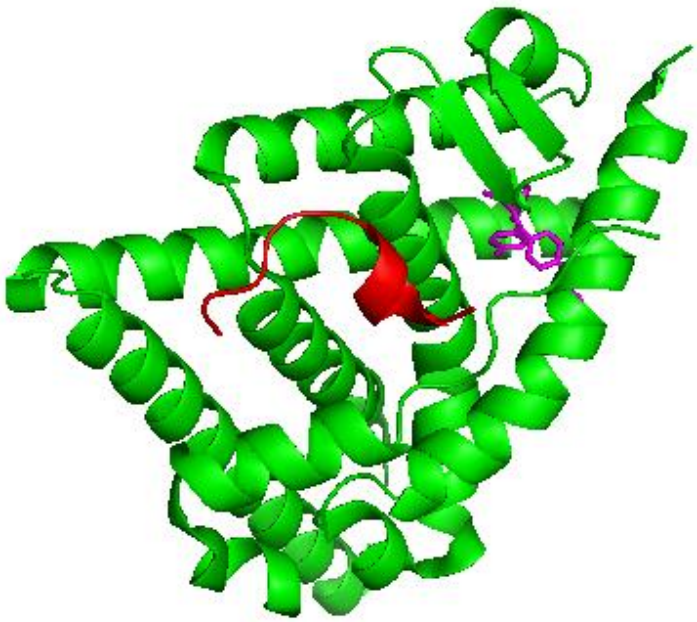


Figure 3.8: αII-ERα interaction. Ribbon diagram of the crystal structure of the αII peptide (red) bound to the LBD of ERα (green) in the presence of tamoxifen (cyan) (PDB ID: 2BJ4).

Thus, our goal was to determine the applicability of an *in vitro* phage display selection for the isolation of peptidic ligands that can bind to surfaces on other DNA-binding proteins beyond nuclear receptors, as well as the transferability of these interactions to an *in vivo* setting. The first target chosen was the DNA-binding domain of the yeast transcriptional activator Gal4, which has a known binding surface that can be utilized during transcriptional activation. Once this technique was established, it was then used to isolate *novel* peptide sequences that are capable of interacting with the bacterial repressor LexA.

C.2.A. Gal4(1-100)

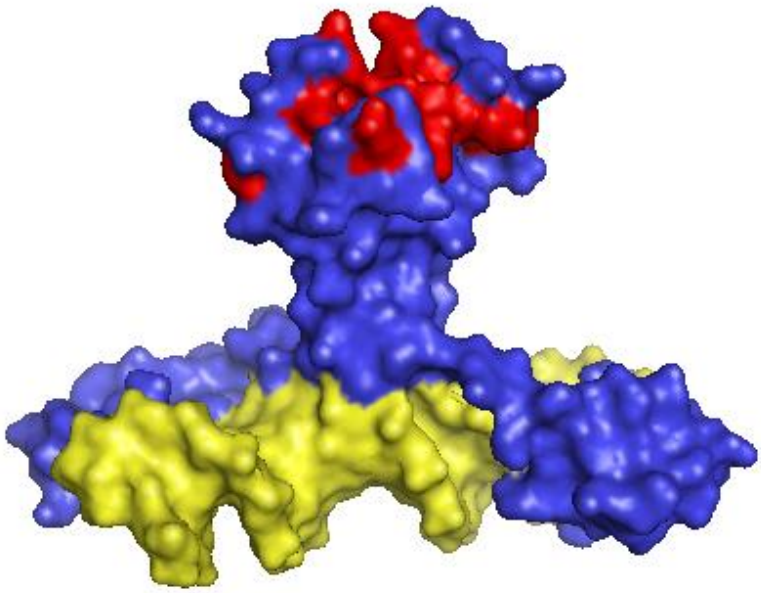


Figure 3.9: Gal11P binding surface on Gal4(1-100). Surface rendering of the crystal structure of a Gal4(1-100) dimer (blue) bound to DNA (yellow) (PDB ID: 3COQ). In red are the 13 residues on the Gal4-dd that are essential for Gal11P binding *in vitro* and activation *in vivo*.

The first 100 residues of the yeast transcriptional activator Gal4, which encompass a DNA binding domain (residues 1-50) and a dimerization domain (residues 50-100) [39], is an important component within two of the most potent artificial transcriptional activators to date. In the first example, Gal4(1-100) gains a novel transcriptional activating capability in yeast cells carrying a single point mutation in the Med15(Gal11) protein of the Mediator complex [40]. This is a result of the fact that Gal11P (P standing for transcriptional potentiator), in which the N342 residue is replaced with one of any number of hydrophobic amino acids, is able to bind to the Gal4 dimerization domain (Gal4-dd), whereas the wild-type coactivator is not (Figure 3.9).

In the second example, attachment of the hydrophobic peptide P201 (YLLPTCIP) to the C-terminus of Gal4(1-100) creates a transcriptional activator as potent as that of full-length Gal4 [41]. This is not only due to the ability of Gal4(93–100)+P201, referred to as XL_Y, to interact with the transcriptional machinery target Med15(Gal11) (K_d = 2.2 ± 0.4 μM) [42], but also due to its ability to interact with the Gal4-dd at the solvent-exposed hydrophobic loop 1 (K_d = 5.0 ± 1.4 μM) [43]. This XL_Y•Gal4-dd complex formation is hypothesized to serve as a masking interaction which protects the TAD from proteolysis and non-productive binding within the cellular milieu. In support of this model, in the absence of this intramolecular masking interaction the 16-amino acid peptide XL_Y is unable to interact with Gal4(1-100) in a yeast two-hybrid setting (Figure 3.10).

In order to use phage display to isolate ligands that are capable of interacting with this hydrophobic binding surface, Gal4(1-100) with an N-terminal His_6-tag was immobilized onto Ni-NTA resin and subjected to 4 rounds of biopanning with $2x10^{11}$ input phage from the Ph.D.-12 library (New England Biolabs), the second of which was a negative selection to remove those phage capable of interacting with the Ni-NTA resin. Twenty-four phage that were eluted from the 4th round of biopanning were then tested for their ability to bind to the target in an enzyme-linked immunosorbent assay (ELISA) (Figure 3.11). In general, at dilution ranges of ~10^9-10^{10} the phage exhibit a binding preference for His_6-Gal4(1-100) immobilized onto a Ni-NTA coated surface over that of nonspecifically bound bovine serum albumin (BSA). Although, sequencing of the peptides failed to reveal a consensus sequence, phage #1 (GHPQEYLLQTVH) possesses homology to XL_Y (Figure 3.12a), suggesting that a subset of peptides could be capable of

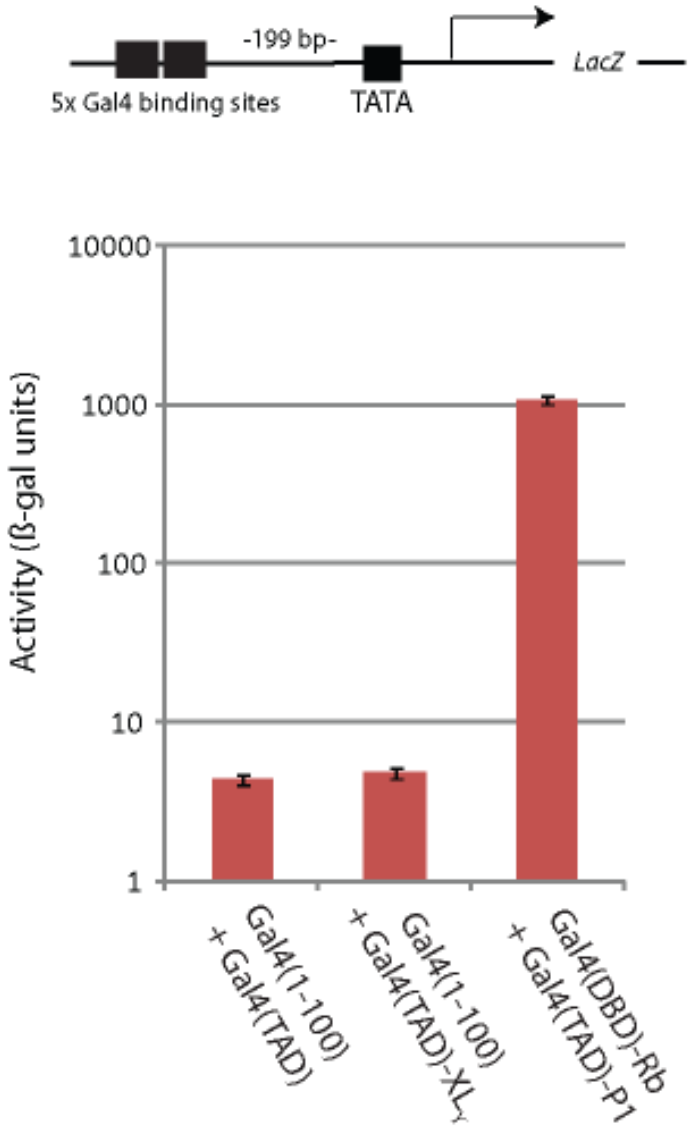


Figure 3.10: Results from a liquid β-galactosidase assay for the Y2H interaction of XL_Y and Gal4(1-100). The constitutive expression of the prey fusion Gal4(TAD)-XL_Y in yeast harboring Gal4(1-100) is unable to upregulate a *LacZ* reporter gene that is under the control of Gal4 binding sites (compared to that of the Gal4(DBD)-Rb/Gal4(TAD)-P1 positive control). Error is indicated as standard deviation of the mean (SDOM).

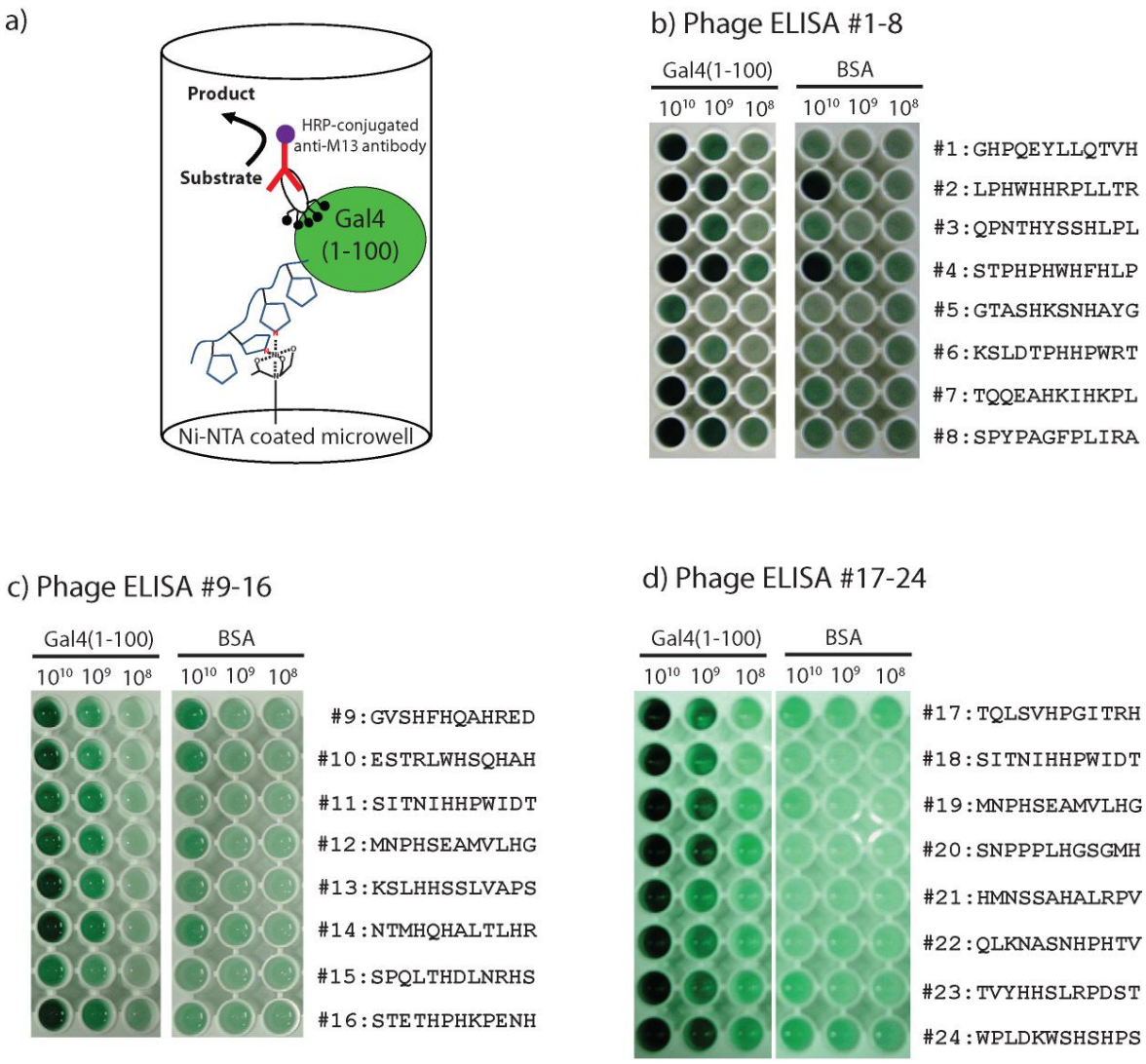


Figure 3.11: ELISA of individual phage that bind Gal4(1-100). a) Schematic of the phage ELISA. Gal4(1-100) with an N-terminal His_6-tag was immobilized onto a Ni-NTA microwell. Phage capable of binding to Gal4(1-100) were identified using an HRP conjugated anti-M13 antibody and an ABTS/H_2O_2 substrate. b) ELISA of phage #1-8. The first three columns contain immobilized His_6-Gal4(1-100) and the second three columns are non-specifically coated with BSA. The phage were serially diluted from 10^{10}-10^8 phage/well as indicated above the figure. Sequences of the peptides attached to the pIII coat protein of each phage are indicated to the right of the figure. c) ELISA of phage #9-16, as described in b. d) ELISA of phage #17-24, as described in b.

interacting via the same surface on Gal4-dd. Furthermore, these results illustrate that a phage display random peptide library can be used to isolate peptides that can bind to surfaces on DNA-binding proteins, and this method can be applied to targets with unknown binding sites (such as the bacterial repressor LexA).

C.2.A.1. Artificial TADs

Phage display has previously been used to discover novel TAD sequences by the selection against masking proteins, which are proteins that bind to TADs and inhibit their exposure to the transcriptional machinery until signaling cues trigger its release. In particular, Kodadek and coworkers isolated two 20-mer peptides (G80BP-A and –B) from a selection against Gal4's masking protein Gal80, which were shown through competition studies to bind to the same site targeted by the Gal4 TAD [44]. Furthermore, these peptides were able to upregulate expression of a reporter gene in yeast when attached to a proteinacious DBD, thus suggesting that those chemical properties used to interact with Gal80 are similar to those which mediate protein-protein interactions within the transcriptional machinery; to this end, the authors demonstrate that G80BP-A is capable of binding to the N-terminus of the coactivator Med15(Gal11), a target of the Gal4 TAD, with a K_d = ~1.2 μM.

Therefore, we hypothesized that peptides that are capable of binding to the masking interaction surface present on Gal4(1-100) would possess chemical properties similar to that of XL_Y, and thus be able to function as transcriptional activation domains. For instance, the hydrophobic residues of XL_Y that are required for its interaction with Med15(Gal11) (tyrosine, for example) are also required for XL_Y to bind to Gal4-dd [43].

Therefore, peptides #1 (GHPQE**Y**LLQTVH) and #8 (SP**Y**PAGFPLIRA), which possess such a conserved tyrosine residue in their sequence (Figure 3.12a), were attached to the C-terminus of Gal4(1-100) and assayed for their ability to activate transcription in yeast. As illustrated in Figure 3.12b, peptides #1 and #8 were able to upregulate expression of the *LacZ* reporter gene approximately 3- and 6-fold (respectively); in addition, the portion of the peptide sequence that is the P201 mimic possessed a 10-fold enhancement in potency (approximately 30- and 60-fold, respectively), most likely due to its ability to better interact with the Gal4-dd.

a)

```
XL_Y: LTGLFVQDYLLPTCIP
 #1:     GHPQEYLLQTVH
 #8:         SPYPAGFPLIRA
```

b)

Construct	Fold Activation
Gal4(1-100)	1.0 ± 0.1
Gal4(1-100)-P201	4330.7 ± 447.2
Gal4(1-100)-GHPQEYLLQTVH	2.9 ± 0.3
Gal4(1-100)-YLLQTVH	28.6 ± 2.3
Gal4(1-100)-SPYPAGFPLIRA	5.9 ± 0.7
Gal4(1-100)-YPAGFPLI	61.1 ± 5.8

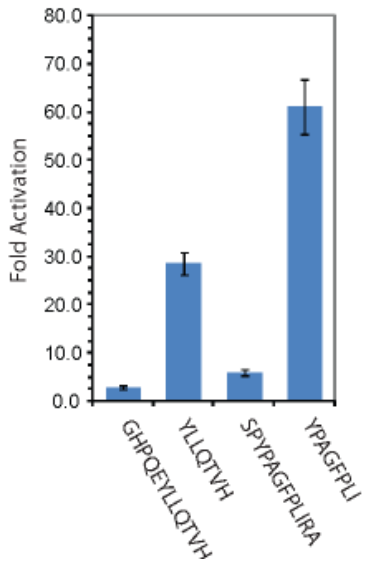


Figure 3.12: Phage #1 and #8 that bind to Gal4(1-100) can activate transcription. a) Sequence alignment of XL_Y, phage #1, and phage #8. Polar residues that are identical are shown in red and hydrophobic residues that are identical are shown in blue. b) Results from a liquid β-galactosidase assay for peptides #1 and #8 attached to the C-terminus of Gal4(1-100). Peptides #1 and #8 were able to upregulate expression of the *LacZ* reporter gene approximately 3- and 6-fold (respectively), while the portion of the peptide sequence that is the P201 mimic possessed a 10-fold enhancement in potency. Error is indicated as standard deviation of the mean (SDOM).

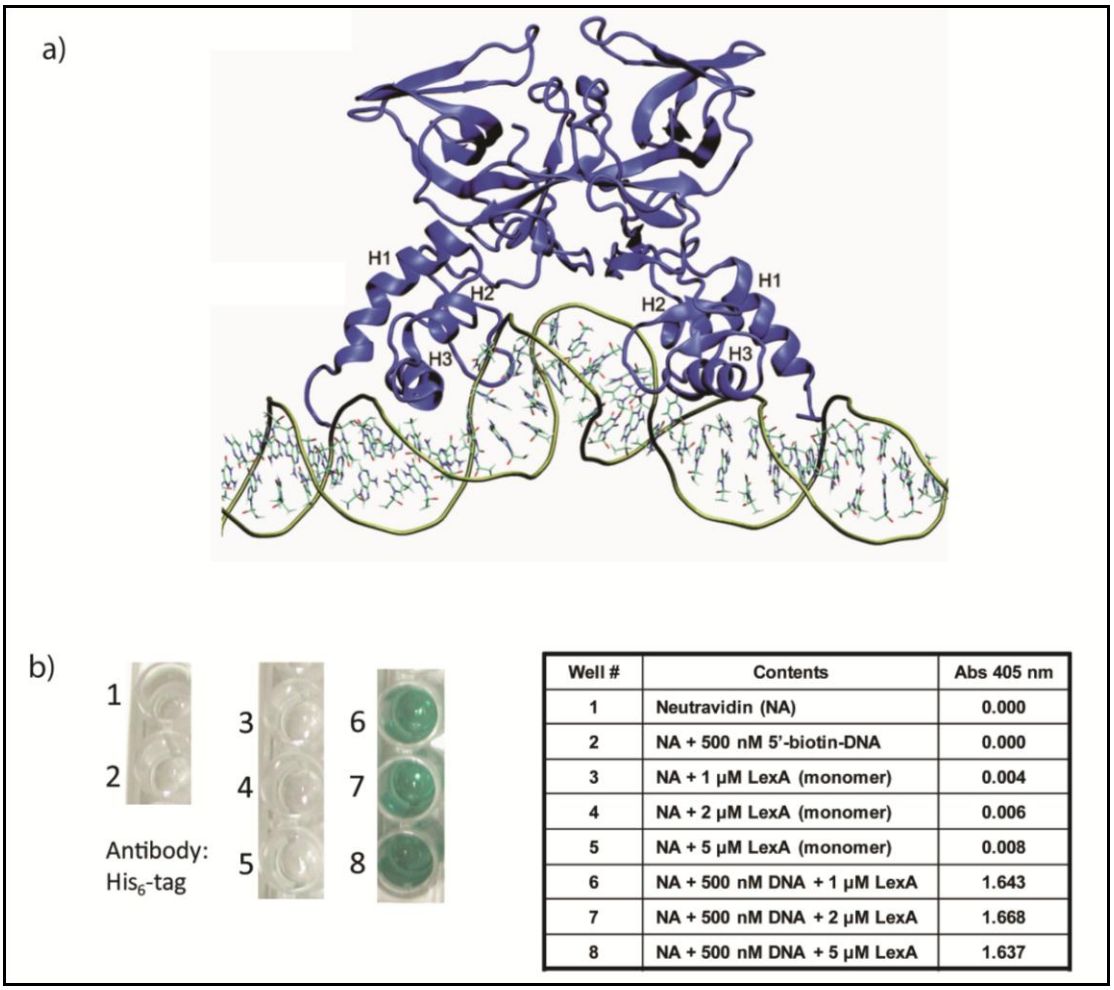
C.2.B. LexA

The bacterial repressor LexA is a 202 amino acid protein that binds as a dimer to its cognate DNA site through an N-terminal helix-turn-helix motif, which is connected via a flexible linker to a C-terminal dimerization domain (Figure 3.13a) [45]. In order to repress transcription of its targeted set of genes in bacteria, LexA binds to the promoters and blocks access of the RNA polymerase, until cellular cues instigate LexA's self-cleavage and degradation; therefore, it should function as a DNA-binding protein devoid of transcriptional activity when introduced into a eukaryotic system.

In order to select for ligands in the context in which the target will be presented from a gene promoter, recombinant LexA was immobilized to a NeutrAvidin-coated microwell via a biotin-labeled double-stranded oligonucleotide containing a single LexA dimer binding site (Figure 3.13b). This immobilized DNA-bound LexA was exposed to 3 rounds of panning (including a negative selection in round 2), and was compared to the results of an analogous biopanning against LexA nonspecifically bound to the solid support. Figures 3.14 and 3.15 show that enrichment of LexA binding phage only occurred when it was tethered to the solid support via a linker,[1] possibly due to its denaturation upon nonspecific binding to the solid support.

Twenty-four phage from the 3rd round of biopanning were characterized in an ELISA (Figure 3.16-3.18). The peptides were able to bind equally well to LexA in the presence and absence of its DNA binding site, but did not bind to BSA (with the exception of phage #12). Taking into consideration recent modeling studies which demonstrate that the DNA-binding domain of LexA rotates with respect to the

[1] The ELISAs in Figures 3.14 and 3.15 were performed by Tonia Buchholz.



Well #	Contents	Abs 405 nm
1	Neutravidin (NA)	0.000
2	NA + 500 nM 5'-biotin-DNA	0.000
3	NA + 1 µM LexA (monomer)	0.004
4	NA + 2 µM LexA (monomer)	0.006
5	NA + 5 µM LexA (monomer)	0.008
6	NA + 500 nM DNA + 1 µM LexA	1.643
7	NA + 500 nM DNA + 2 µM LexA	1.668
8	NA + 500 nM DNA + 5 µM LexA	1.637

Figure 3.13: DNA-bound LexA. a) Proposed model of the LexA repressor bound to DNA from molecular dynamics simulations. With kind permission from Springer Science+Business Media: Cellular and Molecular Life Sciences, The bacterial LexA transcriptional repressor, v66, 2009, 82-93, Butala, M.; Žgur-Bertok, D.; Busby, S.J.W., Figure 2B. b) ELISA with an anti-His6 antibody demonstrating that His6-LexA can bind specifically to 5'-biotin-DNA (containing a single LexA binding site) which has been immobilized onto a NeutrAvidin (NA) coated microwell.

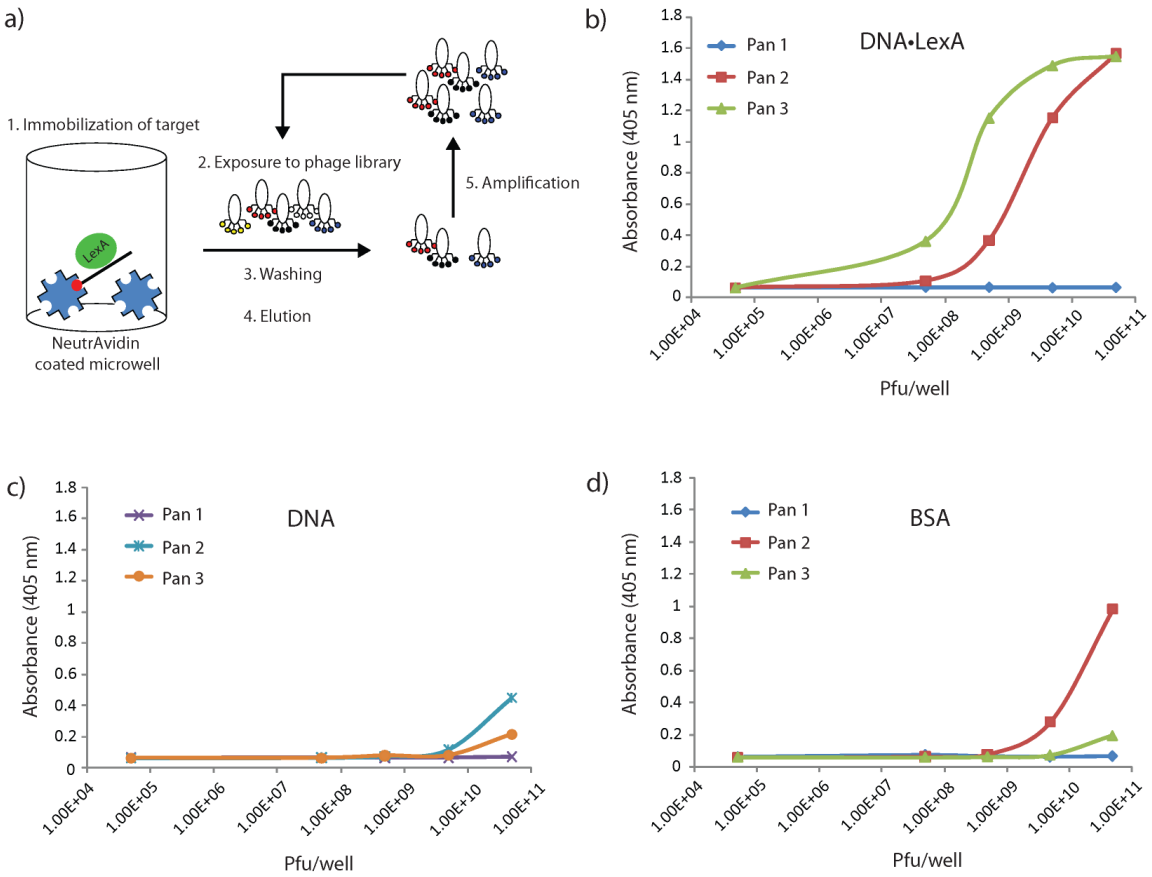


Figure 3.14: Phage display selection against DNA-bound LexA. a) Schematic of the panning procedure against DNA-bound LexA. 5'-biotin-DNA•LexA complex was immobilized onto a NeutrAvidin-coated microwell. The phage library was added, the unbound phage were washed away, and the specifically-bound phage were eluted and amplified for another round. 3 rounds of panning were performed (including a negative selection in round 2). b-d) Phage ELISA to test for enrichment of LexA-binding peptide sequences. Serial dilutions of 5 x 10^7 – 5 x 10^{10} phage/well of the amplified phage eluted from rounds 1-3 were tested for their ability to interact with the DNA•LexA complex (b), DNA (c), or BSA (d). Round 3 phage showed a greater ability to interact with the DNA•LexA complex over that of DNA alone or BSA. Therefore, the selection process was successful.

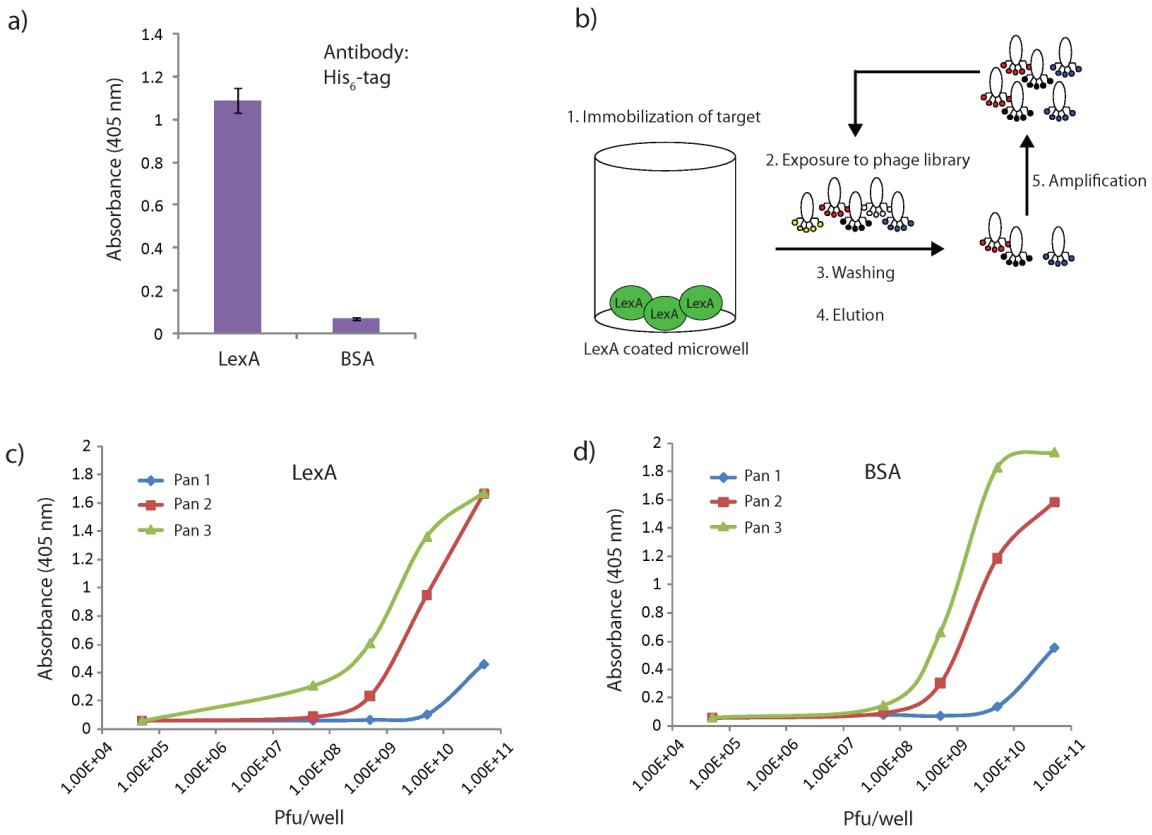


Figure 3.15: Phage display selection against LexA. a) ELISA with an anti-His_6 antibody demonstrating that His_6-LexA can bind non-specifically to the surface of a microwell (compared to BSA as a negative control). b) Schematic of the panning procedure against LexA. LexA(1-202) was immobilized onto a microwell, the phage library was added, the unbound phage were washed away, and the specifically-bound phage were eluted and amplified for another round. 3 rounds of panning were performed (including a negative selection in round 2). c-d) Phage ELISA to test for enrichment of LexA-binding peptide sequences. Serial dilutions of $5 \times 10^7 - 5 \times 10^{10}$ phage/well of the amplified phage eluted from rounds 1-3 were tested for their ability to interact with the LexA (c) or BSA (d). Round 3 phage showed no greater ability to interact with LexA than that of BSA. Therefore, the selection process was not successful.

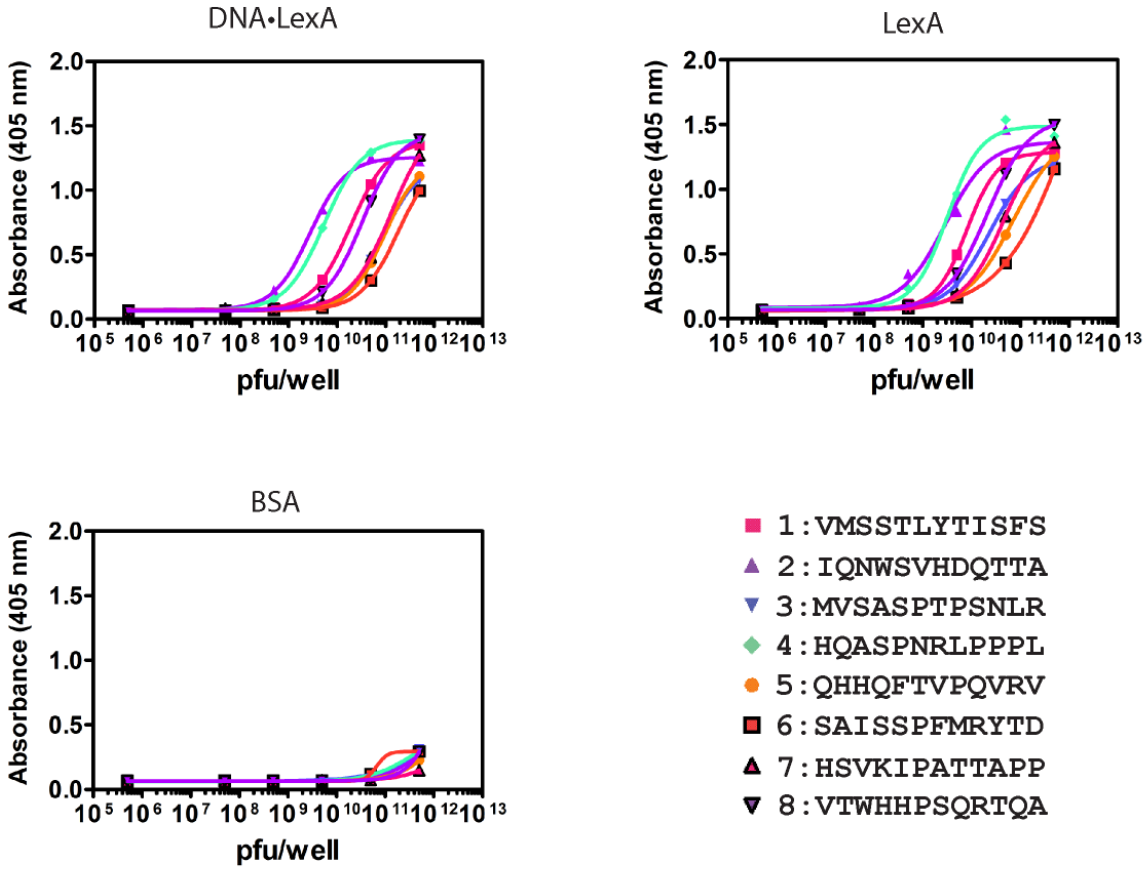


Figure 3.16: ELISA of individual phage #1-8 that bind LexA. Serial dilutions of 5 x 10^7 – 5 x 10^{11} phage/well of amplified phage #1-8 (sequences shown) were tested for their ability to interact with (a) DNA-bound LexA, (b) LexA, or (c) BSA that was immobilized onto a microwell. Detection of the bound-phage was achieved with an HRP conjugated anti-M13 antibody and an ABTS/H_2O_2 substrate. The absorbance at 405 nm was measured using an absorbance plate reader (Tecan Genios Pro). The data was curve-fit (solid line) to a sigmoidal dose-response (variable slope) in order to obtain the LogEC50 values presented in Table 3.2.

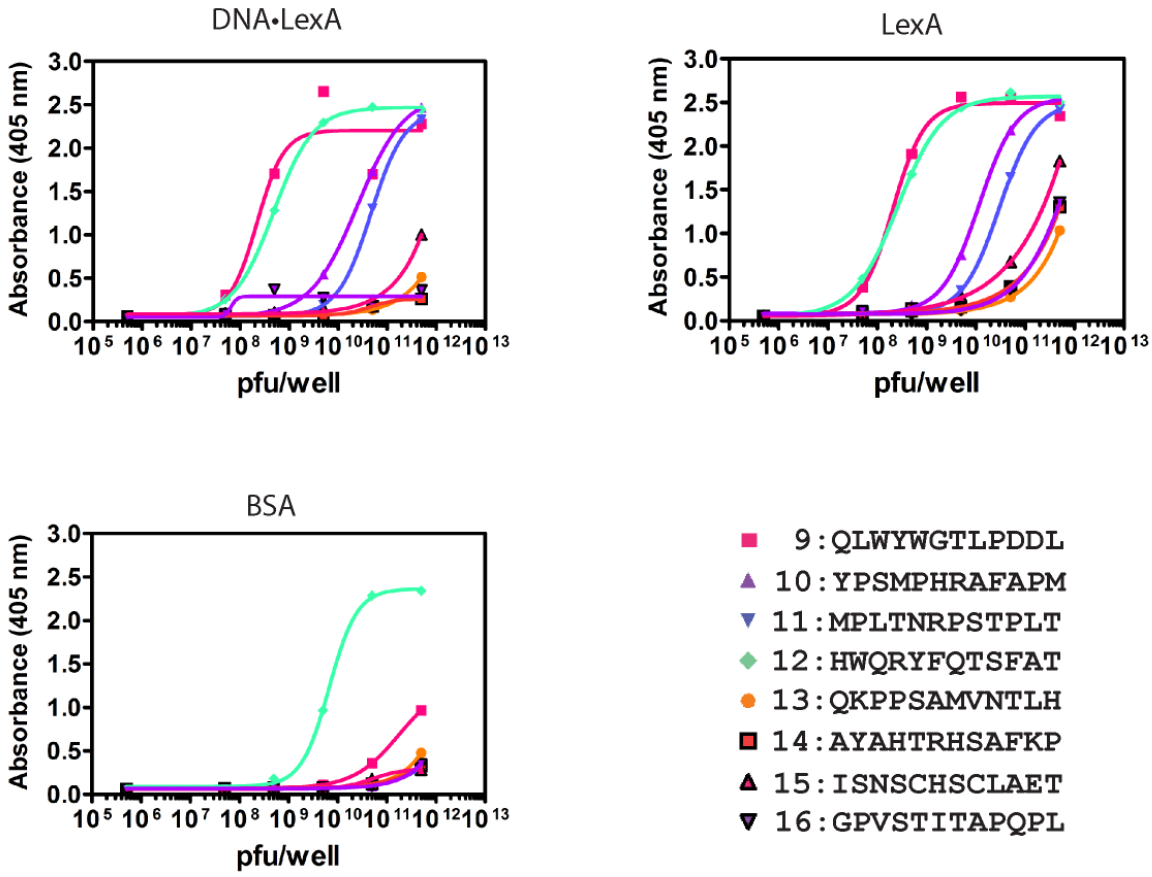


Figure 3.17: ELISA of individual phage #9-16 that bind LexA. Serial dilutions of 5 x 10^7 – 5 x 10^{11} phage/well of amplified phage #9-16 (sequences shown) were tested for their ability to interact with (a) DNA-bound LexA, (b) LexA, or (c) BSA that was immobilized onto a microwell. Detection of the bound-phage was achieved with an HRP conjugated anti-M13 antibody and an ABTS/H_2O_2 substrate. The absorbance at 405 nm was measured using an absorbance plate reader (Tecan Genios Pro). The data was curve-fit (solid line) to a sigmoidal dose-response (variable slope) in order to obtain the LogEC50 values presented in Table 3.2.

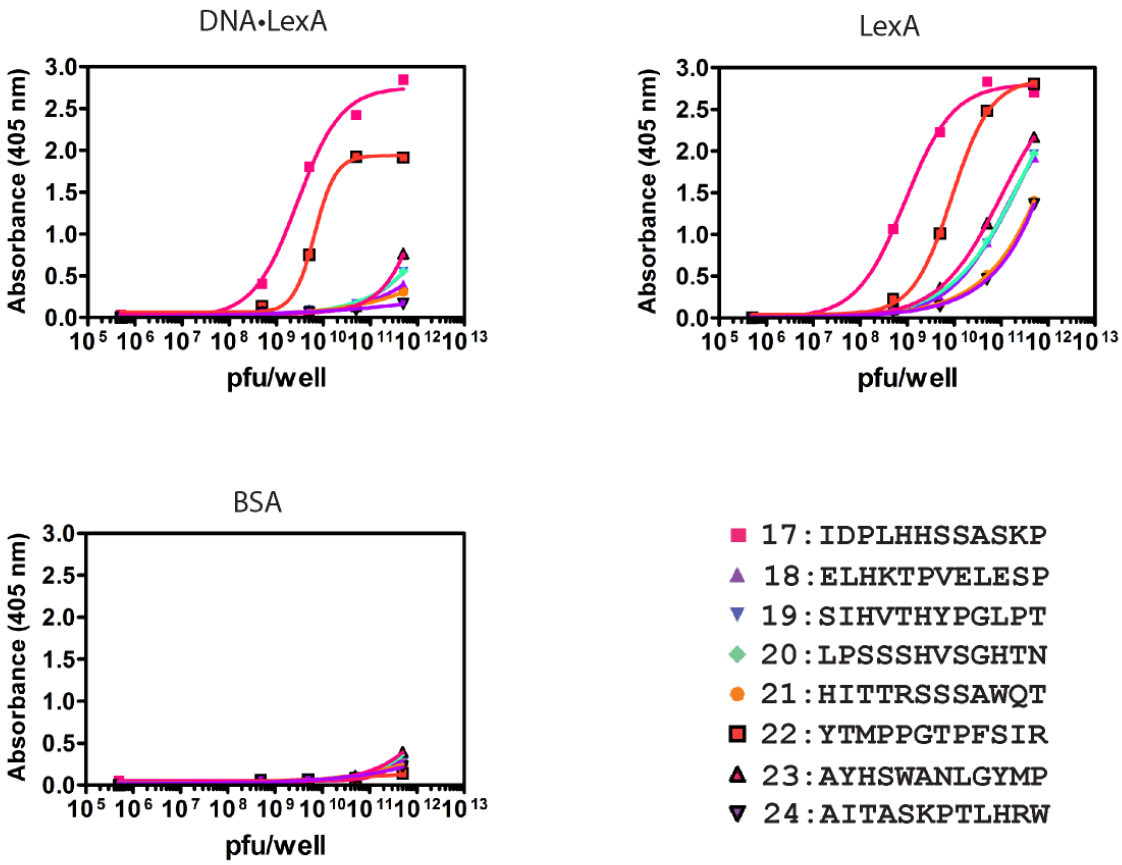


Figure 3.18: ELISA of individual phage #17-24 that bind LexA. Serial dilutions of 5 x 10^7 – 5 x 10^{11} phage/well of amplified phage #17-24 (sequences shown) were tested for their ability to interact with (a) DNA-bound LexA, (b) LexA, or (c) BSA that was immobilized onto a microwell. Detection of the bound-phage was achieved with an HRP conjugated anti-M13 antibody and an ABTS/H_2O_2 substrate. The absorbance at 405 nm was measured using an absorbance plate reader (Tecan Genios Pro). The data was curve-fit (solid line) to a sigmoidal dose-response (variable slope) in order to obtain the LogEC50 values presented in Table 3.2.

dimerization domain upon sequence-specific binding to DNA [45], these results may be an indication that the peptides are targeting the invariable dimerization domain. The five sequences showing significant specific binding and with the lowest EC50 values (< $1x10^{10}$; Table 3.2, highlighted in yellow) were then tested for their ability to interact with DNA-bound LexA in a Y2H setting, but were unable to upregulate above basal levels (data not shown). Possible factors which may be attributed to the inability of the peptides to transition from an *in vitro* to an *in vivo* interaction will be addressed in the next section.

Table 3.2: Peptide sequences identified in an *in vitro* phage display selection against DNA-bound LexA.

Phage	Sequence	DNA•LexA	LexA	BSA
		Log EC50 ± SEM	Log EC50 ± SEM	Log EC50 ± SEM
1	VMSSTLYTISFS	10.28 ± 0.01	9.90 ± 0.01	> 11.70
2	IQNWSVHDQTTA	9.44 ± 0.05	9.46 ± 0.28	> 11.70
3	MVSASPTPSNLR	10.95 ± 0.02	10.35 ± 0.01	> 11.70
4	HQASPNRLPPPL	9.73 ± 0.01	9.50 ± 0.11	11.56 ± 0.20
5	QHHQFTVPQVRV	10.98 ± 0.01	10.86 ± 0.01	> 11.70
6	SAISSPFMRYTD	11.27 ± 0.05	> 11.70	10.83 ± 4.57
7	HSVKIPATTAPP	11.12 ± 0.09	10.67 ± 0.01	> 11.70
8	VTWHHPSQRTQA	10.53 ± 0.01	10.32 ± 0.01	> 11.70
9	QLWYWGTLPDDL	8.35 ± 0.49	8.32 ± 0.12	11.26 ± 0.24
10	YPSMPHRAFAPM	10.40 ± 0.01	10.07 ± 0.03	> 11.70
11	MPLTNRPSTPLT	10.67 ± 0.01	10.48 ± 0.02	> 11.70
12	HWQRYFQTSFAT	8.68 ± 0.02	8.42 ± 0.06	9.83 ± 0.04
13	QKPPSAMVNTLH	> 11.70	> 11.70	> 11.70
14	AYAHTRHSAFKP	10.64 ± 0.07	> 11.70	> 11.70
15	ISNSCHSCLAET	> 11.70	> 11.70	10.71 ± 0.06
16	GPVSTITAPQPL	7.85 ± 183.32	> 11.70	> 11.70
17	IDPLHHSSASKP	9.45 ± 0.19	8.96 ± 0.14	11.36 ± 0.23
18	ELHKTPVELESP	> 11.70	11.15 ± 0.25	> 11.70
19	SIHVTHYPGLPT	> 11.70	11.49 ± 0.15	> 11.70
20	LPSSSHVSGHTN	> 11.70	11.49 ± 0.29	> 11.70
21	HITTRSSSAWQT	> 11.70	> 11.70	> 11.70
22	YTMPPGTPFSIR	9.82 ± 0.11	9.95 ± 0.06	> 11.70
23	AYHSWANLGYMP	> 11.70	11.01 ± 0.06	> 11.70
24	AITASKPTLHRW	> 11.70	> 11.70	> 11.70

C.2.C. Conclusions

Overall, a random peptide phage display selection was able to identify peptides that can bind to surfaces on DNA-binding proteins. However, these ligands that were isolated *in vitro* were unable to reconstitute the interaction in a cellular environment. Possible explanations for this result could be attributed to their proteolytic degradation or an inability to locate their functional binding partner due to non-productive binding with other proteins. In addition, because the peptides were displayed from the pIII coat protein as 5 copies throughout their identification/characterization, there may be an avidity effect needed to mediate the binding interaction, which is not present when expressed as the Gal4(TAD)-peptide fusion.

D. Conclusions/Future directions

In conclusion, DNA-binding proteins have proven to be difficult targets for peptidic ligand binding in comparison to enzymes or protein receptors with defined binding pockets. We have found that an affinity selection, rather than a screening approach, of a random peptide library was able to isolate ligands that are capable of binding to the surfaces present on these types of targets. However, peptides that are 12 amino acids in length that are capable of binding to their targets *in vitro* do not transition well to an *in vivo* interaction, most likely due to problems associated with their proteolytic stability and binding specificity within the cellular milieu.

One approach for increasing the stability of the peptide sequences identified thus far is to incorporate them into larger protein scaffolds; these larger constructs can then be expressed in cell culture as fusions to a transcriptional regulatory domain and assayed for

their ability to interact with the promoter-bound target. For instance, when the previously described ERα-binding peptide αII was used to replace NR boxes 1 and 2 within the coactivator TIF2 to create TIF2αII, this engineered coactivator was able to increase ERα activity in the presence of the antagonist tamoxifen (even with a truncated ERα lacking its AF1 domain) [38]. Presumably, these results are due to the ability of the αII peptide and ERα to reconstitute their binding interaction. In addition, when Mekalanos and co-workers screened a library of random 16 amino acid peptides in bacterial cells for their ability to interact with and inhibit the activity of thymidylate synthase (ThyA), it was as fusions to the protein scaffold thioredoxin [46]. Thioredoxin was chosen because of its stability and tolerance of insertions within the active site of the enzyme; thus, tethering both ends of the peptide provides a higher degree of conformational constraint that would not be present as a fusion to its N- or C-termini, which may increase the specificity/affinity of the peptide for its target. Finally, Nolan and coworkers expressed a random peptide library in mammalian cells that included the flanking sequence EFLIVIKS (which forms stable dimers in solution) before and after the random peptide sequence to serve this purpose as well [47].

Future directions for this project, which will be addressed in more detail in Chapter 4 of this dissertation, will focus on performing the phage display selections using a random peptide library that is conformationally-constrained. In particular, helix stabilized ligands would be optimal candidates for the future production of peptidomimietic and small molecule replacements.

E. Experimental Methods

Table of plasmids used in this study[*]

Plasmid name	Origin	Restriction Sites	Function
pGBT9-Rb	Clontech	NA	2μ yeast expression plasmid under the control of a constitutive *ADH1* (~400 bp) promoter with a TRP^+ selection marker
pGAD GH-P1	Clontech	NA	2μ yeast expression plasmid under the control of an ethanol-repressed *ADH1* (full-length) promoter with a LEU^+ selection marker
pGBKT7-REST pGBKT7-SoxS	This study	EcoRI/ BamHI	2μ yeast expression plasmid under the control of a constitutive *ADH1* (~700 bp) promoter with a TRP^+ selection marker
pCLexA-SoxS	This study	EcoRI/ BamHI	2μ yeast expression plasmid under the control of a constitutive *ADH1* promoter with a HIS^+ selection marker
pYESTrp2-αCTD	This study	HindIII/ XhoI	2μ yeast expression plasmid under the control of an inducible *GAL1* promoter with a TRP^+ selection marker
pGADT7-XL_Y	This study	NdeI/ BamHI	2μ yeast expression plasmid under the control of an ethanol-repressed *ADH1* (full-length) promoter with a LEU^+ selection marker
NYC317-stop codon NYC317-P201 NYC317-phage#1 NYC317-phage#8 NYC317-Y1 NYC317-Y8	This study	XbaI/ SalI	ARS/CEN yeast expression plasmid under the control of a β-actin promoter with a HIS^+ selection marker; the designated peptides are expressed directly on the C-terminus of Gal4(1-100) (without any amino acids in between for a linker)
pRSETA-Gal4 pRSETA-LexA	This study	BamHI/ EcoRI	Expresses proteins in *E. coli* with an N-terminal His_6 affinity tag

*Plasmids described were constructed using standard molecular biology techniques. The sequences of all plasmids were verified by sequencing at the University of Michigan Core Facility (Ann Arbor, MI).

Table of yeast strains used in this study

Yeast strain	Genotype	Application
AH109	Mata, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::*GAL1*$_{UAS}$-*GAL1*$_{TATA}$-*HIS3*, MEL1 *GAL2*$_{UAS}$-*GAL2*$_{TATA}$-*ADE2*, URA3::*MEL1*$_{UAS}$-*MEL1*$_{TATA}$-*lac*Z	Y2H random peptide library screen
Y187	Matα, ura3-52, his3-200, ade 2-101, trp 1-901, leu 2-3, 112, gal4Δ, met–, gal80Δ, URA3::*GAL1*$_{UAS}$-*GAL1*$_{TATA}$-*lac*Z, MEL1	Y2H random peptide library screen
EGY48 /pSH18-34	MATα, ura3, his3, trp1, LexA$_{op\ (x6)}$-LEU2 (pSH18-34: lacZ under control of lexA$_{op(x8)}$, URA3, amp^{r})	Detection of the SoxS-αCTD Y2H interaction
LS41	JPY9::ZZ41, Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4 URA::pZZ41 (pZZ41 contains two LexA binding sites 50 bp upstream of the *GAL1*$_{TATA}$, and five Gal4 binding sites 191 bp upstream of the *GAL1*$_{TATA}$)	Test the activity of the phage display peptides that bind to Gal4(1-100)

Liquid β-Galactosidase assays

Plasmids were transformed into yeast using the LiOAc method or by electroporation; the transformed colonies were selected by growth on synthetic complete (SC) media containing 2% glucose and lacking the appropriate amino acid(s). Freshly transformed colonies were then used to inoculate 5 mL cultures of SC media containing either 2% raffinose (or 2% glucose for the Y2H screen) and lacking the appropriate amino acids. The cultures were incubated overnight at 30°C with agitation, and subsequently used to inoculate 5 mL cultures of SC media containing 2% raffinose + 2% galactose (or 2% glucose for the Y2H screen) and lacking the appropriate amino acids. Again, these cultures were incubated overnight at 30°C with agitation until an OD_{660} of 0.3-0.5 (10-fold dilution) was reached. 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 (ONPG) (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 total protein concentration of the extract, measured using a Bradford assay kit (Bio-Rad) with BSA as the standard. If further sensitivity was required, a portion of the cell extract could then be used to measure β-galactosidase activity via incubation with chlorophenol red-β-D-galactopyranoside (CPRG) according to the manufacturer recommended protocol (Stratagene High Sensitivity β-Galactosidase Assay Kit).

Matchmaker Random Peptide Library Y2H Screen

A yeast two-hybrid screen was performed with the Matchmaker Random Peptide Library following the manufacturer recommended protocols (Clontech). Briefly, yeast strain AH109 (whose glycerol stock had been streaked onto synthetic complete (SC) media containing 2% glucose and lacking lysine) was transformed by LiOAc methods with the pGBKT7 plasmid encoding the bait gene fusion. Transformed colonies were selected by growth on synthetic complete (SC) media containing 2% glucose lacking lysine and tryptophan. The random 16 amino acid peptide library that is ligated into the pGAD GH vector was then transformed, using the Yeastmaker Yeast Trasnformation System 2 kit (Clontech), into the AH109 strain harboring the pGBKT7-bait plasmid. The 15 mL of resuspended cells were plated using glass beads onto 5x20 batches of 150-mm plates, whereby a positive bait-peptide interaction was selected for by growth at 30 °C on

synthetic complete (SC) media containing 2% glucose lacking lysine, tryptophan, leucine, and histidine. In addition, in order to calculate the transformation efficieny, 150 uL of resuspended cells were also plated onto a 100-mm plate of synthetic complete (SC) media containing 2% glucose lacking lysine, tryptophan, and leucine. Colonies from the screen that grew in the absence of histidine were restreaked onto fresh media. Cells from this restreaked plate were then resuspended in 50 uL of sterile water, serially diluted to 1/1000, and 50 uL of this solution was plated onto fresh media in order to yield single colonies. Four colonies from each plate were then grown in synthetic complete (SC) media containing 2% glucose lacking leucine in order to select for the pGAD GH plasmid encoding the peptide, which was subsequently isolated using a Zymoprep II yeast plasmid miniprep kit. The pGAD GH peptide plasmids were then transformed into SmartCells™ chemically competent *E. coli*. (Genlantis) and further selected by the ampicillin resistence marker, whereby they were isolated using a Qiagen Miniprep Kit. The purified plasmids were sequenced by the the University of Michigan Sequencing Core to obtain the identity of the peptide. Finally, in order to quantitate the bait-peptide interaction using a liquid β-galactosidase assay, the pGAD GH peptide plasmids were cotransformed with the pGBKT7-bait plasmid into the yeast strain Y187 by electroporation methods and selected by growth on synthetic complete (SC) media containing 2% glucose lacking tryptophan, leucine, and uracil.

Protein expression and purification

$\underline{His_6\text{-Gal4(1-100):}}$ Expression of His_6-Gal4(1-100) was carried out in BL21(DE3)pLysS *E. coli* cells as previously described [48]. Briefly, cultures (50 mL) from single colonies

were grown overnight at 37 °C (250 rpm) in Lennox L Broth (Research Products International) supplemented with ampicillin (100μg/mL) and chloramphenicol (34 μg/mL) before dilution (50-fold) into a 1L culture of Lennox L Broth supplemented with fresh antibiotic. After an OD_{600} of 0.4 was reached, expression was induced with IPTG (final concentration 1 mM) in the presence of 20 μM $ZnSO_4$ for 2-3 hours. The cell pellet was resuspended in 25 mL lysis buffer A (10 mM Tris, pH 8.0 at 4 °C, 500 mM NaCl, 10% glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), 10 mM imidazole, and a Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the His_6-tagged protein was isolated using Ni-NTA Agarose (Qiagen). The protein was bound to the resin by batch absorption (15 mL of a 50% slurry that was pre-washed with lysis buffer), and then subjected to FPLC purification with a BioRad system. The column was run at 2 mL/min, in which wash buffer A (20 mM Tris, pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 30 mM imidazole) was run for 50 min, followed by elution buffer A (20 mM Tris, pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 250 mM imidazole) for another 50 min. Collected protein fractions were combined, concentrated to 1 mL in an Amicon 5K centrifugal filter device (Millipore), and buffer exchanged three times into 15 mL of storage buffer A (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM β-ME, 1 mM EDTA, 20 μM $ZnSO_4$). The protein concentration was measured using a Bradford assay kit (Bio-Rad) with BSA as the standard.

His_6-LexA(1-202): Expression of His_6-LexA(1-202) was carried out in C41(DE3)pLysS *E. coli* cells. Briefly, cultures (50 mL) from single colonies were grown overnight at 37 °C (250 rpm) in Lennox L Broth (Research Products International) supplemented with

ampicillin (100μg/mL) and chloramphenicol (34 μg/mL) before dilution (40-fold) into a 1L culture of Lennox L Broth supplemented with fresh antibiotic. After an OD_{600} of 0.3 was reached, the culture was cooled for 15 min at 16 °C (100 rpm), and expression was induced with IPTG (final concentration 0.5 mM) for 3-4 hours at 250 rpm. The cell pellet was resuspended in 20 mL lysis buffer B (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 % glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), and a Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the His_6-tagged protein was isolated using Ni-NTA Agarose (Qiagen) [49]. The cell lysate was incubated with 2 x 1 mL of Ni-NTA beads for 2 hours at 4 °C. The beads were washed 3 times with 3 mL wash buffer B (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 % glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), 20 mM imidazole). The protein was eluted from the beads by incubation at 4 ºC overnight with 1 mL elution buffer B (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 % glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), 250 mM imidazole). Additional protein was eluted by twice incubating the beads with elution buffer for 1 hour at 4 °C. The protein samples were combined, concentrated to 1 mL using an Amicon 10K centrifugal filter device, and buffer exchanged three times into 15 mL of storage buffer B (100 mM Tris-Cl (pH 7.5), 500 mM NaCl, 1 mM EDTA, 1 mM DTT) [50]. The protein concentration was measured using a Bradford assay kit (Bio-Rad) with BSA as the standard.

Biotin-labeled LexA oligonucleotides

The oligonucleotide 5'-CTA TAC TGT ATA TAA AAC CAG TGG TTA TATGTA CAG TAA TCC-3’ was purchased from Invitrogen with a 5’-modification of biotin and

annealed with an unlabeled complementary oligonucleotide in annealing buffer (10 mM HEPES pH 7.5, 150 mM NaCl) by the following method: heat denaturation for 7 min at 95 ºC, cooling at room temperature for 30 min, and cooling at 4 ºC for 30 min.

Phage Display

Biopanning against Gal4(1-100): Phage display was performed using the Ph.D.-12 Phage Display Peptide Library Kit using the manufacturer recommended protocols (New England Biolabs (NEB)). Briefly, 50 uL of a 50% aqueous suspension of Ni-NTA Agarose (Qiagen) was washed once with 1 mL wash buffer (10 mM PBS, 0.1% Tween-20) and pelleted in a centrifuge at 3000 rpm for 30 sec. The supernatant was removed and the resin was blocked in 1 mL of blocking buffer (10 mM PBS, 0.2% non-fat milk) at 4°C for 1 hour. The resin was then washed 4 x 1mL, and 200 uL of 10 nM Gal4 (monomer) was added and incubated for 15 min at room temperature (mixing occasionally). After washing 4 x 1 mL, the phage library was diluted 20-fold, and 200 uL of this solution was incubated with the resin for 20 min at room temperature (mixing occasionally). The resin was washed 10 x 1 mL, and the binding phage were eluted in 1 mL of 0.2 M Glycine-HCl (pH 2.2) + 1 mg/mL BSA for 10 min at room temperature. The eluant was transferred to a microfuge tube and neutralized with 150 μL of 1M Tris-HCl (pH 9.1). 10 uL of the eluant was saved for titering. The rest of the eluant was amplified in ER2738 *E. coli*, precipitated using PEG/NaCl, and ~1-2 x 10^{11} pfu were used for a 2nd round of panning, consisting of a negative selection was performed against the blocked Ni-NTA agarose. The nonbinding phage were amplified and used for a 3rd round of panning consisting of a positive selection, but this time using 0.5% Tween-20 in

the wash buffer. Finally, a 4th round of panning was performed analogous to the 3rd round, from which individual plaques were selected and amplified, and their DNA extracted for sequencing at the University of Michigan Sequencing Core.

Biopanning against LexA(1-202): Phage display was performed using the Ph.D.-12 Phage Display Peptide Library Kit using the manufacturer recommended protocols (New England Biolabs). Briefly, to pan against LexA that has been immobilized in its DNA-bound form, a NUNC Maxisorp well was first incubated overnight at 4°C with 150 µL of 100 µg/mL of NeutrAvidin (Pierce) in 0.1 M $NaHCO_3$ (pH 8.6). The well was then washed with wash buffer (4 x 200 µL) (10 mM PBS, 0.1% Tween-20), and 100 µL of 4.5 µM 5'-biotin duplex DNA in 10 mM PBS was added and incubated for 1 hour at room temperature. The well was washed 4 x 200 µL, and 300 µL of sterile filtered blocking buffer (10 mg/mL BSA in PBS) was added and incubated for 1 hour at room temperature. The well was washed 6 x 200 µL, and 100 uL of 9 µM LexA (monomer) (diluted in wash buffer) was added and incubated for 1 hour at room temperature. The well was again washed 6 x 200 µL, and 100 µL of a 10-fold dilution of the original phage library was added and incubated for 1 hour at room temperature. The well was washed 10 x 200 µL, and the binding phage were eluted with 100 µL of 0.2 M Glycine-HCl (pH 2.2) for 10 min with agitation. The supernatant was transferred to a microfuge tube and neutralized with 15 uL of 1M Tris-HCl (pH 9.1). 5 uL of the eluant was saved for titering. The rest was amplified in ER2738 *E. coli*, precipitated using PEG/NaCl, and ~1-2 x 10^{11} pfu were used for a 2nd round of panning. In the 2nd round of panning, a negative selection was performed against the blocked NeutrAvidin•5'-biotin-DNA complex, and the supernatant

was immediately used for a positive selection against DNA-bound LexA before amplification. The amplified phage from the 2nd round were then used in a 3rd round of panning consisting of a positive selection. Enrichment of the libraries was determined with an ELISA, whereupon individual plaques could then be selected, amplified, and their DNA extracted for sequencing at the University of Michigan Sequencing Core. A similar protocol was used to pan against 100 uL of 9 μM LexA (monomer) that had been non-specifically bound to the well overnight at 4°C.

ELISA

Immobilized Gal4(1-100): 200 uL of 10 μg/μL (or 0.6 μM) His_6-Gal4(1-100) was added to a Ni-NTA coated 96-well plate and incubated overnight at 4 °C. The plate was blocked with blocking buffer (10 mM PBS, 0.2% nonfat milk) for 1 hour. Subsequently, 200 μL of serial dilutions of 10^{10} - 10^{18} phage were added to the appropriate wells in wash buffer (10 mM PBS, 0.2% Tween-20) and incubated for 1 hour. After 6 washes, 200 μL of anti-M13-HRP antibody (1:5000 dilution) was added in blocking buffer to each well for 1 hour. After 6 washes, the phage were detected using 200 μL of 2,2'-Azino-bis(3-Ethylbenzthiazoline-6-Sulfonic Acid) (ABTS) solution in sodium citrate with hydrogen peroxide as per the NEB phage display manual. The absorbance at 405 nm was measured using an absorbance plate reader (Tecan Genios Pro).

Immobilized LexA(1-202): To probe LexA in its DNA-bound form, a NUNC Maxisorp 96-well plate was incubated overnight at 4°C with 150 μL of 10 μg/mL of NeutrAvidin (Pierce) in 0.1 M $NaHCO_3$ (pH 8.6). The wells were then washed with wash buffer (3 x

200 μL) (10 mM PBS, 0.5% Tween-20), and 100 μL of 0.5 μM 5'-biotin duplex DNA in 10 mM PBS was added and incubated for 1 hour at room temperature. The wells were washed 3 x 200 μL, and 200 μL of sterile filtered blocking buffer (10 mg/mL BSA in PBS) was added and incubated for 1 hour at room temperature. The wells were washed 3 x 200 μL, and 100 uL of 1 μM LexA (monomer) diluted in wash buffer was added and incubated for 1 hour at room temperature. The wells were washed 3 x 200 μL, and 100 μL of serial dilutions of phage were added to the appropriate wells in wash buffer and incubated for 1 hour at room temperature. After 3 washes, 100 μL of anti-M13-HRP antibody (1:5000 dilution) was added in wash buffer to each well and incubated for 1 hour at room temperature. After 3 more washes, the phage were detected using 200 μL of 2,2'-Azino-bis(3-Ethylbenzthiazoline-6-Sulfonic Acid) (ABTS) solution in sodium citrate with hydrogen peroxide as per the NEB phage display manual. The absorbance at 405 nm was measured using an absorbance plate reader (Tecan Genios Pro). These results were compared to those in which a well was either non-specifically coated with 150 μL of 1 μM LexA or 10 mg/mL BSA in PBS.

F. References

1. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
2. Chen, X., et al., *Gene expression patterns in human liver cancers.* Mol Biol Cell, 2002. **13**(6): p. 1929-39.
3. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. **2**(10): p. 740-9.
4. Pandolfi, P.P., *Transcription therapy for cancer.* Oncogene, 2001. **20**(24): p. 3116-27.
5. Wands, A.M. and A.K. Mapp, *Transcription-based Therapeutics*, in *Wiley Encyclopedia of Chemical Biology*, T.P. Begley, Editor. 2008, John Wiley & Sons, Inc.: Hoboken, NJ.
6. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
7. Ptashne, M. and A. Gann, *Genes & Signals*. 2001, New York: Cold Spring Harbor Laboratory.
8. Graslund, T., et al., *Exploring strategies for the design of artificial transcription factors: targeting sites proximal to known regulatory regions for the induction of gamma-globin expression and the treatment of sickle cell disease.* J Biol Chem, 2005. **280**(5): p. 3707-14.
9. Falke, D., et al., *Design of artificial transcription factors to selectively regulate the pro-apoptotic bax gene.* Nucleic Acids Res, 2003. **31**(3): p. e10.
10. Corbi, N., et al., *The artificial zinc finger coding gene 'Jazz' binds the utrophin promoter and activates transcription.* Gene Ther, 2000. **7**(12): p. 1076-83.
11. Reynolds, L., et al., *Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1615-20.
12. Rebar, E.J., et al., *Induction of angiogenesis in a mouse model using engineered transcription factors.* Nat Med, 2002. **8**(12): p. 1427-32.
13. Dervan, P.B., R.M. Doss, and M.A. Marques, *Programmable DNA binding oligomers for control of transcription.* Curr Med Chem Anticancer Agents, 2005. **5**(4): p. 373-87.
14. Liu, B. and T. Kodadek, *Investigation of the relative cellular permeability of DNA-binding pyrrole-imidazole polyamides.* J Med Chem, 2009. **52**(15): p. 4604-12.
15. Xiao, X., et al., *Design and synthesis of a cell-permeable synthetic transcription factor mimic.* J Comb Chem, 2007. **9**(4): p. 592-600.
16. Fuller, G.N., et al., *Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST)/neuron-restrictive silencer factor, which can be functionally countered by REST-VP16.* Mol Cancer Ther, 2005. **4**(3): p. 343-9.
17. Immaneni, A., et al., *REST-VP16 activates multiple neuronal differentiation genes in human NT2 cells.* Nucleic Acids Research, 2000. **28**(17): p. 3403-3410.

18. Adya, N., et al., *Expansion of CREB's DNA recognition specificity by Tax results from interaction with Ala-Ala-Arg at positions 282-284 near the conserved DNA-binding domain of CREB.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5642-6.
19. Livengood, J.A., et al., *Paradoxical effects of DNA binding polyamides on HTLV-1 transcription.* Front Biosci, 2004. **9**: p. 3058-67.
20. Geiger, T.R., et al., *The human T-cell leukemia virus type 1 tax protein confers CBP/p300 recruitment and transcriptional activation properties to phosphorylated CREB.* Mol Cell Biol, 2008. **28**(4): p. 1383-92.
21. Majmudar, C.Y. and A.K. Mapp, *Chemical approaches to transcriptional regulation.* Curr Opin Chem Biol, 2005. **9**(5): p. 467-74.
22. Young, K.H., *Yeast two-hybrid: so many interactions, (in) so little time.* Biol Reprod, 1998. **58**(2): p. 302-11.
23. Yang, M., Z. Wu, and S. Fields, *Protein-peptide interactions analyzed with the yeast two-hybrid system.* Nucleic Acids Res, 1995. **23**(7): p. 1152-6.
24. *MATCHMAKER Random Peptide Library*, in *User Manual*. 2001.
25. Kim, H.Y., B.Y. Ahn, and Y. Cho, *Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen.* EMBO J, 2001. **20**(1-2): p. 295-304.
26. Sugiyama, K., M. Muroi, and K. Tanamoto, *A novel TLR4-binding peptide that inhibits LPS-induced activation of NF-kappaB and in vivo toxicity.* Eur J Pharmacol, 2008. **594**(1-3): p. 152-6.
27. Dangi, B., et al., *Versatility of the carboxy-terminal domain of the alpha subunit of RNA polymerase in transcriptional activation: use of the DNA contact site as a protein contact site for MarA.* Mol Microbiol, 2004. **54**(1): p. 45-59.
28. Roopra, A., et al., *Transcriptional repression by neuron-restrictive silencer factor is mediated via the Sin3-histone deacetylase complex.* Mol Cell Biol, 2000. **20**(6): p. 2147-57.
29. Nomura, M., et al., *The neural repressor NRSF/REST binds the PAH1 domain of the Sin3 corepressor by using its distinct short hydrophobic helix.* J Mol Biol, 2005. **354**(4): p. 903-15.
30. Chong, J.A., et al., *REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons.* Cell, 1995. **80**(6): p. 949-57.
31. Tapia-Ramirez, J., et al., *A single zinc finger motif in the silencing factor REST represses the neural-specific type II sodium channel promoter.* Proc Natl Acad Sci U S A, 1997. **94**(4): p. 1177-82.
32. Vidalain, P.O., et al., *Increasing specificity in high-throughput yeast two-hybrid experiments.* Methods, 2004. **32**(4): p. 363-70.
33. Shah, I.M. and R.E. Wolf, Jr., *Novel protein--protein interaction between Escherichia coli SoxS and the DNA binding determinant of the RNA polymerase alpha subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-element-containing promoters to SoxS-dependent promoters during oxidative stress.* J Mol Biol, 2004. **343**(3): p. 513-32.
34. Willats, W.G., *Phage display: practicalities and prospects.* Plant Mol Biol, 2002. **50**(6): p. 837-54.
35. *Ph.D.-12 Phage Display Peptide Library Kit*, in *Instruction Manual*. 2006.

36. Norris, J.D., et al., *Peptide antagonists of the human estrogen receptor.* Science, 1999. **285**(5428): p. 744-6.
37. Paige, L.A., et al., *Estrogen receptor (ER) modulators each induce distinct conformational changes in ER alpha and ER beta.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3999-4004.
38. Kong, E.H., et al., *Delineation of a unique protein-protein interaction site on the surface of the estrogen receptor.* Proc Natl Acad Sci U S A, 2005. **102**(10): p. 3593-8.
39. Hong, M., et al., *Structural basis for dimerization in DNA recognition by Gal4.* Structure, 2008. **16**(7): p. 1019-26.
40. Hidalgo, P., et al., *Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain.* Genes & Development, 2001. **15**(8): p. 1007-1020.
41. Lu, X.Y., A.Z. Ansari, and M. Ptashne, *An artificial transcriptional activating region with unusual properties.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(5): p. 1988-1992.
42. Lu, Z., et al., *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): p. 8591-8596.
43. Lu, Z., et al., *Unraveling the mechanism of a potent transcriptional activator.* J Biol Chem, 2005. **280**(33): p. 29689-98.
44. Han, Y. and T. Kodadek, *Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast.* J Biol Chem, 2000. **275**(20): p. 14979-84.
45. Knegtel, R.M., et al., *A model for the LexA repressor DNA complex.* Proteins, 1995. **21**(3): p. 226-36.
46. Blum, J.H., et al., *Isolation of peptide aptamers that inhibit intracellular processes.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2241-6.
47. Xu, X., et al., *Dominant effector genetics in mammalian cells.* Nat Genet, 2001. **27**(1): p. 23-9.
48. Reece, R.J., R.J. Rickles, and M. Ptashne, *Overproduction and single-step purification of GAL4 fusion proteins from Escherichia coli.* Gene, 1993. **126**(1): p. 105-7.
49. *The QIAexpressionist,* in *Handbook.* 2003.
50. Brooks, P.C., F. Movahedzadeh, and E.O. Davis, *Identification of some DNA damage-inducible genes of Mycobacterium tuberculosis: apparent lack of correlation with LexA binding.* J Bacteriol, 2001. **183**(15): p. 4459-67.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

A. Introduction

Transcriptional activators control the gene expression patterns of cells by regulating the timing and extent to which mRNA levels are upregulated in response to signaling cues [1]. They achieve this with a modular architecture, in which one domain of the protein (the DBD) is responsible for localizing to a specific DNA sequence within the genome, while another domain (the TAD) is responsible for recruiting multiprotein complexes that aid in the assembly of the pre-initiation complex (PIC) at the gene promoter [1-2]. Experiments described in Chapter 2 revealed that the favorability of a conformational change step during activator•coactivator complex formation correlates with activator potency (i.e., levels of transcriptional output). Thus, future efforts will focus on further characterizing the roles that the TADs and coactivators play throughout this process, both of which are important to consider when designing exogenous molecules that can effectively orient themselves within this pathway.

Furthermore, much work has been done on the development of small molecules for use as DBDs in activator ATFs that can be programmed to display the proper recognition surfaces to interact with a desired DNA sequence [3]; although size limitations imposed to maintain cell permeability has thus limited their targeting specificity in relation to the proteins they are trying to replace. Instead, in Chapter 3 we have proposed an alternative

method which relies on the interaction of a ligand with an endogenous DNA-bound protein in order to localize a TAD to the promoter of a gene. To this end, while we have shown that a phage display selection technique is able to isolate peptidic ligands that can target surfaces on DNA-binding proteins, we hypothesize that a conformation-constrained peptide library instead will result in the identification of ligands with higher specificity and stability for use in a cellular context.

B. Activator•coactivator complex formation

Transcription is initiated through a series of coupled binding equilibria between transcriptional activators and their array of protein targets within the transcriptional machinery. Thus, we hypothesized that the life-times of these interactions should be more revealing of differences in activator potency (i.e., transcriptional output) than that of equilibrium binding measurements (apparent affinities). We therefore performed a transient kinetic analysis on the interactions of the TADs of Gal4, Gcn4, and VP16 with their shared coactivator target Med15 [4-8]. In our experimental set-up, the TADs were attached to the Gal4(1-100) DBD and equilibrated with fluorescently labeled DNA, which was then mixed in a fluorescence stopped flow spectrometer with the N-terminal 345 amino acids of Med15. The Gal4, Gcn4, and VP16 DNA-bound TADs all exhibited an increase in fluorescence over time at each Med15 concentration tested, producing association time-courses that are biphasic and best fit by a double exponential. Our data is consistent with a two-step binding mechanism, in which a fast bimolecular collision step is either preceded or followed by a slow a conformational change step. However, regardless of which of these limiting binding models is chosen for analysis, a more

favorable conformational change occurring within the complex correlates with a TAD sequence that is more potent. However, as will be described in the remainder of this section, the contributions made by the TAD and Med15 during this process need to be further resolved by additional experiments.

B.1. Future directions

B.1.A. Order of events within the two-step binding mechanism

It has been established through NMR studies that a small percent of isolated TAD conformers may possess secondary structural elements within the regions utilized for coactivator binding [9-10]; although these pre-formed motifs do not encompass the whole polypeptide segment found structured within the final bound complexes [11-12]. Thus, this implicates the involvement of a conformational change step occurring in the TAD structure *after* binding, in order to better maximize intermolecular interactions available within the coactivator binding site. In order to definitively observe such a step within the binding mechanism, efforts in our lab are currently focused on developing a FRET readout between the TAD-Med15 protein interface for use in the fluorescence stopped-flow experiments.

FRET has been utilized previously to observe the interaction network of transcriptional machinery proteins, such as the KID•KIX interaction [13], the Gal4•Tra1 and Gal4•Gal80 interactions [14], and the Gal3•Gal80 interaction [15]. In particular, we propose to use the tetracysteine-biarseincal system developed by Tsien and coworkers [16] to fluorescently label the coactivator Med15 with ReAsH to serve as a FRET acceptor for EGFP-tagged activators. Previously, through a combination of *in vitro* cross-

linking, multiplexed mass spectrometry, and genetic approaches, Mapp and coworkers have identified the functionally important binding sites of Gal4, Gcn4 and VP16 as residing in the amino terminus of Med15(Gal11) (first 400 residues) [8]. Thus, we believe placement of the CCPGCC tetracysteine motif within loop regions that are proximal to these identified activator binding sites will be tolerated for efficient labeling with the ReAsH dye [17-18]. To achieve this, we can use the crystal structure of residues 1-100 of Med15 (i.e., the KIX domain) [19], as well as the recently described Phyre structure prediction program [20], to design optimal points for tag placement within residues 1-345 of Med15. Current work is also underway to visualize these TAD-Med15 interactions *in vivo* in the context of full-length proteins that are located within the transcriptional machinery (See Appendix).

B.1.B. Significance of TAD structure

Analysis of TAD•Med15 complex formation according to a model in which a conformational change occurs *after* binding implicates that the stability of intermediate species is important in determining the partition ratio of the TAD (i.e., with a more stable intermediate preferring to undergo a conformational change over that of dissociation). One hypothesis to explain how the TADs of Gal4, Gcn4, and VP16 can form intermediates of differing stabilities with the shared coactivator Med15, as implied from their differing partition ratios, is that they possess different structural propensities in their unbound forms that are inherent to their primary amino acid sequences. Therefore, the characterization of the excited state conformers of the TADs used in our kinetic analysis would provide further insight into the mechanism of TAD•target complex formation. For

instance, Han and coworkers have determined that the region within the isolated TAD of VP16 (residues 412-490) that is found to interact with targets (i.e., 472-479 or DFEFEQMF) forms a well-defined helix in ~5% of the conformers at equilibrium [9], however, similar studies should be performed on the fragment used in our studies, residues 456-490, in order to determine if the context of the sequence has a drastic effect on this propensity, as well as on the TADs of Gal4 (residues 840-881) and Gcn4 (residues 107-144). In particular, because the nine residues 862-870 of the Gal4 TAD, or DDVYNYLFD, are expected to form an α-helix upon binding to a target protein as suggested by the crystal structure of the *K. lactis* Gal4•Gal80 proteins [21], it will be interesting to see if this region exists as a pre-structured motif. Even further, it will be interesting to see if there is a correlation between the % of excited state conformers present at equilibrium and the stability of the intermediate involved during coactivator binding. Also, the effects on this structural propensity upon restricting the movement of the TAD by tethering it to a DBD (and thus limiting the number of conformers it can sample) could be determined through a measurement of its residual dipolar couplings (RDCs), as has been achieved previously with the TAD of p53 [10].

Finally, the TADs used in activator ATFs containing non-natural DBDs are generally small peptide sequences taken from a larger, naturally occurring sequence [22-24]. Therefore, similar structural and kinetic characterizations should be performed on these types of TADs as well, such as H2 of VP16 (residues 469-485) [9, 11] and its peptidomimetic replacements (β-peptide and peptoid versions of this sequence) [25], in order to determine if they retain the properties of the larger sequences they were obtained from.

B.1.C. Significance of the conformational change step

The conformational change step within the TAD•Med15 complex is a defining step within their binding mechanism, presumably for efficiently recruitment of the Mediator complex to the promoter through a stable interaction. However, it has yet to be determined if this is true for the interaction of the TADs of Gal4, Gcn4, and VP16 with other shared coactivators, such as that of Tra1 which resides in the SAGA histone modifying complex. If a two-step binding mechanism is in fact a commonality, it would be interesting to observe how the binding site identity (i.e., the composition of amino acids available to make TAD contacts) affects the favorability of the conformational change step.

Furthermore, a conformational change in Med15 may also be occurring upon TAD binding, in order to mediate a signal transduction from the TAD to the rest of the protein complex through an interaction with this target. Indeed, the binding of activators to different subunits of the human CRSP-Med complex induce distinct conformational states (Figure 4.1) [26-27]. To this end, we propose to use our tetracysteine-biarsenical labeling system within the N-terminus of Med15 in combination with an EGFP tag fusion to either the N- or C-terminus of the protein to detect a possible conformational change upon TAD binding via intramolecular FRET; similar studies have been performed to monitor the conformational change within TFIIB from a closed to an open state upon TBP•DNA binding [28-29]. In addition, a gel ‘super shift’ assay can be performed with DNA-bound activators and Med15 to ensure tag placement does not disrupt this interaction [21, 28, 30].

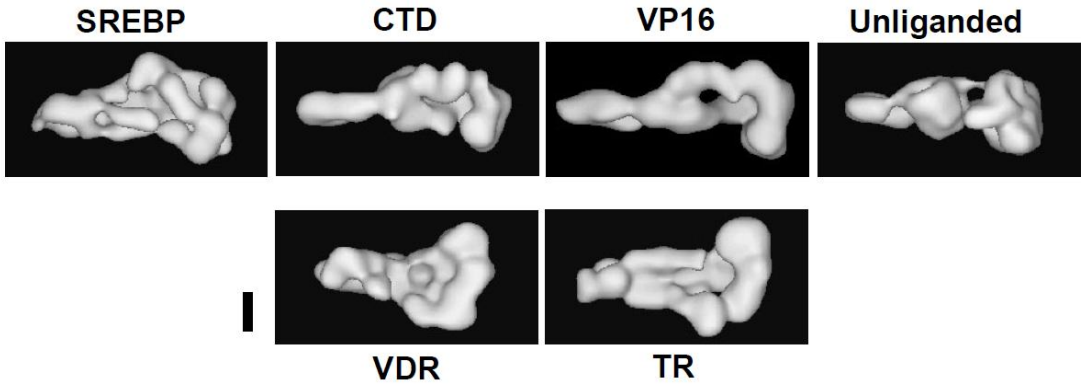


Figure 4.1: Conformational states of the CRSP-Med complex induced by activators. VP16 interacts with the Med78 or Med97 subunit, SREBP-1a interacts with Med105, and TR and VDR both interact with Med220. Activators that interact with different subunits within CRSP-Med seem to induce different conformational shifts within the complex. Scale bar, 75 Å. Reprinted by permission from Macmillan Publishers Ltd: Nature Structural and Molecular Biology, Taatjes, D.J.; Schneider-Poetsch, T.; Tjian, R., Distinct conformational states of nuclear receptor-bound CRSP-Med complexes, copyright (2004). From Science, Taatjes, D.J.; Näär, A.M.; Andel III, F.; Nogales, E.; Tjian, R., Structure, function, and activator-induced conformations of the CRSP coactivator, 1058-1062, copyright (2002). Reprinted with permission from AAAS.

C. Ligands that target DNA-binding proteins

One significant challenge in the development of artificial transcription factors (ATFs) is a lack of small molecules that can be used to localize them to a gene promoter in a cellular context. We propose to circumvent this problem by utilizing ligands that can target endogenous DNA-bound proteins. To this end, we found that an *in vitro* phage display selection technique was able to isolate peptides 12 amino acids in length that can interact with the surfaces of Gal4(1-100) and LexA(1-202). However, these peptides did not translate well to a cellular context, possibly due to non-productive interactions and instability due to proteolysis. One approach to address these issues is to instead isolate ligands from a conformation-constrained random peptide library. Even further, selecting for helix stabilized ligands would pave the way for the construction of peptidomimetic

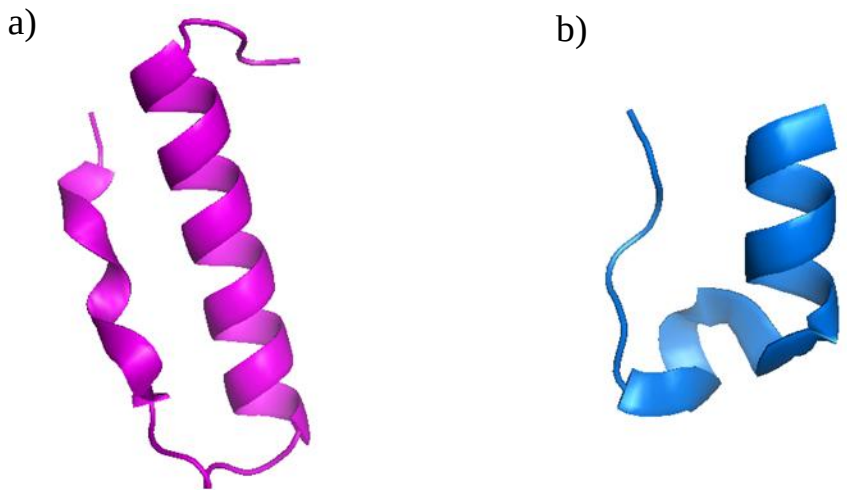


Figure 4.2: Helix-stabilizing using mini-proteins. a) Ribbon diagram of the structure of the avian pancreatic protein (PDB ID: 1PPT). b) Ribbon diagram of the structure of the Trp-cage protein (PDB ID: 1LTY).

replacements, such as β-peptides, peptoids, and stapled peptides, as well as small molecules**.**

C.1. Future directions

C.1.A. Conformation-constrained peptides

Disulfide-bridges and loop regions of protein scaffolds are approaches that have been integrated into phage display selections to obtain high-affinity ligands [31-32]. Even further, because α-helices have been found to play a major role in mediating protein-protein interactions (with the average length spanning 8-12 residues), miniature proteins that display stable, solvent-exposed helical scaffolds that are readily amenable to diversification have been used for the selection of such recognition elements (Figure 4.2) [25, 32]. For instance, Schepartz and coworkers have successfully utilized the aPP (avian pancreatic protein) scaffold, consisting of a single α-helix stabilized by hydrophobic

interaction with a type II polyproline helix (Figure 4.2a) [33-34], while Johnson and coworkers have successfully utilized the Trp-cage scaffold derived from a 20-residue C-terminal sequence of extendin-4 (Figure 4.2b) [35].

The isolation of α-helical peptidic ligands that can interact with surfaces on DNA-binding proteins would aid in the discovery of peptidomimetic and small molecule replacements. For instance, β-peptides and peptoid analogs are helical foldamers with enhanced stability in cell culture [25, 36]. In addition, a covalent bond formed between positions i and i+4 of the helix via an olefin-metathesis reaction has also been proven to stabilize its formation [37]. Bradner and coworkers have successfully constructed such a hydrocarbon-stapled peptide, SAHM1, that permeates T-cell acute lymphoblastic leukaemia (T-ALL) cultured cells and induces anti-proliferative effects by interacting with the NOTCH transcription factor complex (Kd = 0.12 ± 0.02 μM) [38]. Finally, the peptides obtained from a phage display selection can also be used discover small molecules that can compete for binding in a fluorescence polarization assay [39].

D. Conclusions

Overall, this thesis addresses topics that will facilitate the development of more efficient activator ATFs. In particular, a better understanding of the structural features required for a conformational change during TAD•target complex formation, as well as the involvement of the conformational change within the coactivator itself, will be needed in order to design molecules that can mimic these functions. In addition, new approaches for localizing the TAD to the promoter of a gene of interest in a cell with a cell-permeable molecule will be needed for future applications as therapeutic agents.

E. References

1. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. **2**(1): p. 62-75.
2. Ptashne, M. and A. Gann, *Genes & Signals*. 2001, New York: Cold Spring Harbor Laboratory.
3. Ansari, A.Z. and A.K. Mapp, *Modular design of artificial transcription factors.* Curr Opin Chem Biol, 2002. **6**(6): p. 765-72.
4. Jeong, C.J., et al., *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. **40**(31): p. 9421-7.
5. Park, J.M., et al., *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. **20**(23): p. 8709-19.
6. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. **25**(20): p. 9092-102.
7. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. **18**(3): p. 369-78.
8. Majmudar, C.Y., et al., *A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry.* Angew Chem Int Ed Engl, 2009. **48**(38): p. 7021-4.
9. Kim, D.H., et al., *Multiple hTAF(II)31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16.* BMB Rep, 2009. **42**(7): p. 411-7.
10. Wells, M., et al., *Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5762-7.
11. Uesugi, M., et al., *Induced alpha helix in the VP16 activation domain upon binding to a human TAF.* Science, 1997. **277**(5330): p. 1310-3.
12. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. **274**(5289): p. 948-53.
13. Mayr, B.M., G. Canettieri, and M.R. Montminy, *Distinct effects of cAMP and mitogenic signals on CREB-binding protein recruitment impart specificity to target gene activation via CREB.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10936-41.
14. Bhaumik, S.R., et al., *In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer.* Genes Dev, 2004. **18**(3): p. 333-43.
15. Wightman, R., R. Bell, and R.J. Reece, *Localization and interaction of the proteins constituting the GAL genetic switch in Saccharomyces cerevisiae.* Eukaryot Cell, 2008. **7**(12): p. 2061-8.
16. Adams, S.R., et al., *New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications.* J Am Chem Soc, 2002. **124**(21): p. 6063-76.
17. Nakanishi, J., et al., *FRET-based monitoring of conformational change of the beta2 adrenergic receptor in living cells.* Biochem Biophys Res Commun, 2006. **343**(4): p. 1191-6.

18. Goodman, J.L., D.B. Fried, and A. Schepartz, *Bipartite tetracysteine display requires site flexibility for ReAsH coordination.* Chembiochem, 2009. **10**(10): p. 1644-7.
19. Thakur, J.K., et al., *A nuclear receptor-like pathway regulating multidrug resistance in fungi.* Nature, 2008. **452**(7187): p. 604-9.
20. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server.* Nat Protoc, 2009. **4**(3): p. 363-71.
21. Thoden, J.B., et al., *The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p.* J Biol Chem, 2008. **283**(44): p. 30266-72.
22. Stanojevic, D. and R.A. Young, *A highly potent artificial transcription factor.* Biochemistry, 2002. **41**(23): p. 7209-7216.
23. Mapp, A.K., et al., *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): p. 3930-3935.
24. Ansari, A.Z., et al., *Towards a minimal motif for artificial transcriptional activators.* Chemistry & Biology, 2001. **8**(6): p. 583-592.
25. Henchey, L.K., A.L. Jochim, and P.S. Arora, *Contemporary strategies for the stabilization of peptides in the alpha-helical conformation.* Curr Opin Chem Biol, 2008. **12**(6): p. 692-7.
26. Taatjes, D.J., et al., *Structure, function, and activator-induced conformations of the CRSP coactivator.* Science, 2002. **295**(5557): p. 1058-62.
27. Taatjes, D.J., T. Schneider-Poetsch, and R. Tjian, *Distinct conformational states of nuclear receptor-bound CRSP-Med complexes.* Nat Struct Mol Biol, 2004. **11**(7): p. 664-71.
28. Zheng, L., et al., *FRET evidence for a conformational change in TFIIB upon TBP-DNA binding.* Eur J Biochem, 2004. **271**(4): p. 792-800.
29. Heyduk, T., *Measuring protein conformational changes by FRET/LRET.* Curr Opin Biotechnol, 2002. **13**(4): p. 292-6.
30. Hidalgo, P., et al., *Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain.* Genes Dev, 2001. **15**(8): p. 1007-20.
31. Uchiyama, F., et al., *Designing scaffolds of peptides for phage display libraries.* J Biosci Bioeng, 2005. **99**(5): p. 448-56.
32. Binz, H.K., P. Amstutz, and A. Pluckthun, *Engineering novel binding proteins from nonimmunoglobulin domains.* Nat Biotechnol, 2005. **23**(10): p. 1257-68.
33. Chin, J.W. and A. Schepartz, *Concerted evolution of structure and function in a miniature protein.* J Am Chem Soc, 2001. **123**(12): p. 2929-30.
34. Rutledge, S.E., H.M. Volkman, and A. Schepartz, *Molecular recognition of protein surfaces: high affinity ligands for the CBP KIX domain.* J Am Chem Soc, 2003. **125**(47): p. 14336-47.
35. Herman, R.E., et al., *The Trp cage motif as a scaffold for the display of a randomized peptide library on bacteriophage T7.* J Biol Chem, 2007. **282**(13): p. 9813-24.
36. Rowe, S.P. and A.K. Mapp, *Assessing the permissiveness of transcriptional activator binding sites.* Biopolymers, 2008. **89**(7): p. 578-81.

37. Blackwell, H.E., et al., *Ring-closing metathesis of olefinic peptides: design, synthesis, and structural characterization of macrocyclic helical peptides.* J Org Chem, 2001. **66**(16): p. 5291-302.
38. Moellering, R.E., et al., *Direct inhibition of the NOTCH transcription factor complex.* Nature, 2009. **462**(7270): p. 182-8.
39. Reed, D., et al., *Identification and characterization of the first small molecule inhibitor of MDMX.* J Biol Chem, 2010. **285**(14): p. 10786-96.

APPENDIX[1]

[1] In collaboration with Dr. Malathy Krishnamurthy and Dr. Ibrahim Bori. Dr. Bori was responsible for synthesizing the ReAsH dye, while Dr. Krishnamurthy assisted with the remainder of the experiments.

A. Introduction

Transcriptional activators recruit a variety of coactivators to the gene promoter in order to initiate gene expression [1, 2]. Several studies have attempted to identify the binding partners of transcriptional activators, but these approaches are limited by the use of *in vitro* techniques and non-physiological conditions that result in the identification of false positives [3]. Moreover, *in vivo* approaches such as ChIP do not distinguish the between direct and indirect binding partners thus restricting their usefulness for bonafide target identification [4]. One approach to identify the direct binding partners of proteins in a cell is to use FRET (Förster resonance energy transfer) [5]. In this approach, the candidate proteins are labeled with donor and acceptor fluorophores respectively or expressed as fusions to fluorescent proteins in cells and an interaction between the two proteins is detected by the occurrence of a fluorescent signal [6].

Few reports exist of successfully using FRET to observe TAD-target interactions in live cells [7-10]. This is most likely due to complications that arise upon fusing a large autofluorescent protein (AFP) to a coactivator target that resides within a large multiprotein complex. To address this issue, we propose attaching the coactivators instead to a small tetracysteine-containing peptide sequence that can interact with a small molecule fluorophore. This technique would thus enable us to observe those dynamic interactions that we are kinetically characterizing *in vitro* (Chapter 4) in a more transcriptionally relevant setting.

B. FRET: tetracysteine/biarsenical method

FRET has proven to be an enormously powerful tool for the characterization of macromolecular interactions *in vitro* and in cell culture [5, 6, 11]. Briefly, this technique relies upon the nonradiative transfer of energy from a donor to acceptor fluorophore when positioned within a critical distance (~10-100 Å) and orientation from each other. The transfer efficiency is dependent on the inverse sixth power of intermolecular separation, thus making it a sensitive technique for the investigation of biological phenomena that produce changes in molecular proximity (e.g., protein-protein interactions).

Traditionally, the use of this technique in cell culture has relied upon labeling each of the binding partners with an autofluorescent protein tag (AFP). These tags are large in size (~26 kDa), the fusion of which to a desired protein has the potential of disturbing its function. For instance, attachment of an AFP probe can lead to protein misfolding [12] or can perturb protein localization [13], thereby altering cellular processing events. Furthermore, the probability of complications arising is particularly true for proteins that normally exist in complexes. For example, in the budding yeast *Saccharomyces cerevisiae*, expression of a GFP fusion to tubulin fails to complement tubulin null mutants [14]. This phenomenon was also observed in the work of Green et al who carried out FRET experiments in *S. cerevisiae* in which various coactivators in the transcriptional machinery were labeled with EYFP; the other FRET partner, the transcriptional activator Gal4 was labeled with ECFP [8]. From this experiment only a single coactivator (Tra1) produced a FRET signal, despite considerable evidence of a Gal4 targeting at least three distinct coactivators during the process of transcriptional activation.

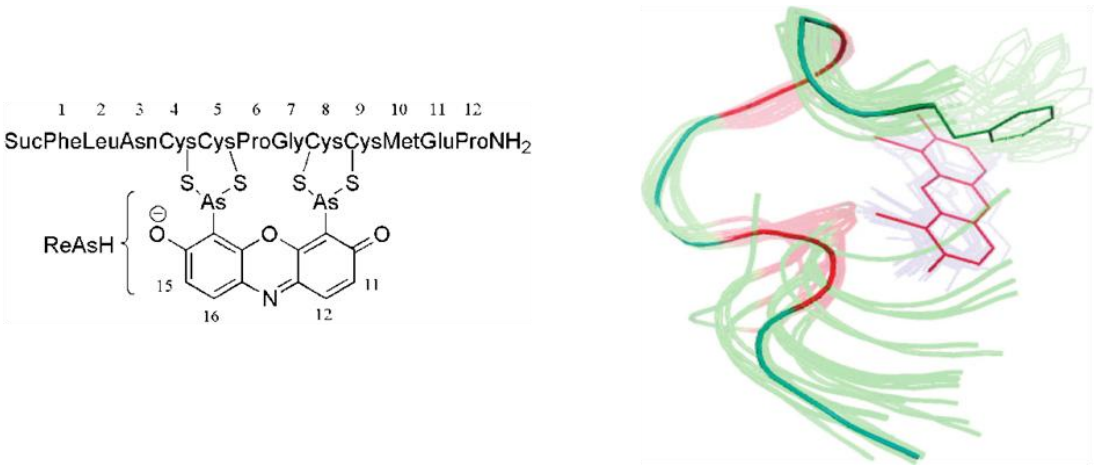


Figure A.1: NMR structure of ReAsH bound to the peptide FLNCCPGCCMEP reveals a hairpin turn. An ensemble of the 30 best structures appear as semi-transparent, while the average peptide backbone structure is shown as a solid line. Cysteine atoms are shown in red. Adapted with permission from Journal of the American Chemical Society, Madani, F.; Lind, J.; Damberg, P.; Adams, S.R.; Tsien, R.Y.; Gräslund, A.O., Hairpin Structure of a Biarsenical−Tetracysteine Motif Determined by NMR Spectroscopy, 4613-4615. Copyright 2009 American Chemical Society.

We propose to circumvent this issue through the use of a far smaller peptide tag that can be post-translationally modified by the addition of membrane-permeable biarsenical molecules, most notably the green and red dyes 'FlAsH' and 'ReAsH', that become fluorescent upon chelation to their targeted sequence. More specifically, the tag is a 12-residue peptide sequence that includes four cysteines (**CC**PG**CC**) and can easily be incorporated into the coactivators of interest (Figure A.1) [15, 16]. This tetracysteine/biarsenical system has been used for labeling proteins in microorganisms [14, 17] and mammalian cell culture [13, 18]. In addition, FlAsH is a good FRET acceptor from CFP [19, 20], while ReAsH is a FRET acceptor from longer wavelength donors such as GFP and YFP [21, 22], thus making amenable to the study of protein-protein interactions.

C. ReAsH labeling of a nuclear protein in *Saccharomyces cerevisiae*

Our initial efforts have focused on labeling a transcriptional activator with a C-terminal EGFP-tetracysteine tag with ReAsH within the model organism of *Saccharomyces cerevisiae*; this experiment would allow us not only to determine if this method is capable of selectively labeling proteins that are localized within the nucleus of the cell (which has not been demonstrated before in yeast), but would also demonstrate the maximum degree of transfer efficiency that can be achieved between the fluorophores (i.e., upon intramolecular FRET). Once this is achieved, we could then move the tetracysteine/biarsenical tag to a coactivator of interest in order to observe intermolecular FRET upon its interaction with an EGFP-tagged activator. We chose ReAsH as our small molecule probe due to previous reports that it possesses a higher affinity and selectivity for the CCPGCC motif over that of FlAsH, thus making it more applicable for the labeling of low abundance proteins (e.g., such as those found within the transcriptional machinery) [23]. In addition, yeast strains in which endogenous transcriptional activators have been tagged on their C-termini with GFP are commercially available (Invitrogen), and could therefore be utilized in subsequent experiments with this set-up. However, as will be described in the remainder of this section, initial labeling attempts call into question the general cell permeability of these types of molecules within *Saccharomyces cerevisiae*.

C.1. EGFP as a FRET donor

The FRET donor EGFP was fused to the C-terminus of transcriptional activators consisting of the TADs from Gal4 (residues 840-881) and Gcn4 (residues 107-144)

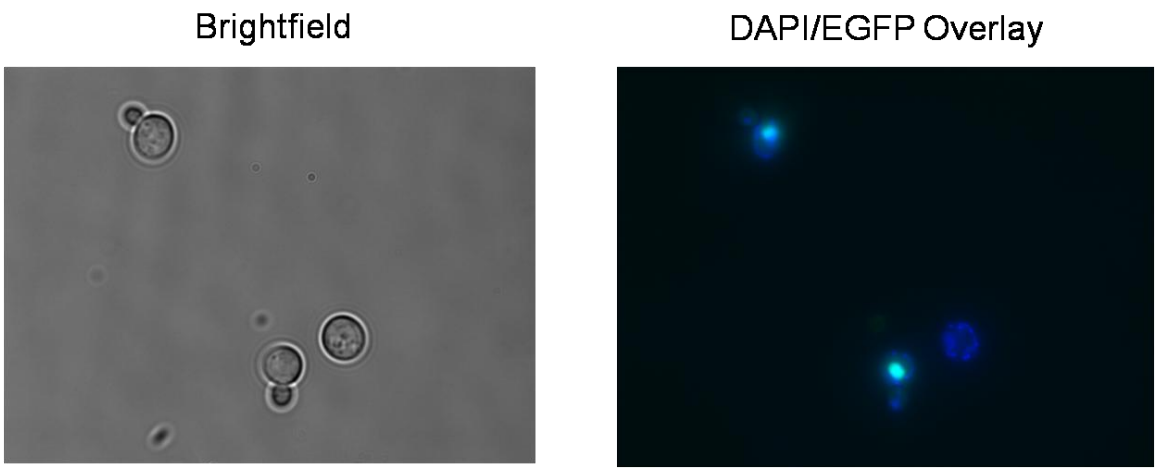


Figure A.2: Live cell imaging of yeast expressing an EGFP-tagged transcriptional activator. Imaging was performed on a Nikon Eclipse TE2000-U fluorescence microscope with an inverted oil immersion 100x objective and DAPI and GFP filter cube sets. Yeast cultures cotransformed with the activator Gal4(1-100)-Gcn4(107-144)-EGFP and FLN**CC**PG**CC**MEP-Med15 were grown to mid-late log phase in SC drop-out media, DAPI stained for 30 min, then 1 µL was placed on a microscope slide, covered with a coverslip, and mounted for imaging.

attached to the Gal4 DBD (residues 1-100). As illustrated in Figure A.2, the EGFP-tagged Gcn4-derived activator is able to localize to the nucleus of *S. cerevisiae* when visualized using fluorescence microscopy; similar results are obtained for the Gal4-derived activator (data not shown). In order to determine what effect this tag has on the ability of the activators to interact with their protein targets within the transcriptional machinery (i.e., their transcriptional potency), the constructs were tested for their ability to upregulate expression of a *LacZ* reporter gene under the control of Gal4 binding sites (Figure A.3). Furthermore, in order to determine the effect that placement of the tetracysteine motif would have on activator-coactivator interactions, the EGFP-tagged activators were expressed in a yeast strain in which the coactivator Med15 [24-28] was expressed from its native promoter with an N-terminal tetracysteine tag sequence

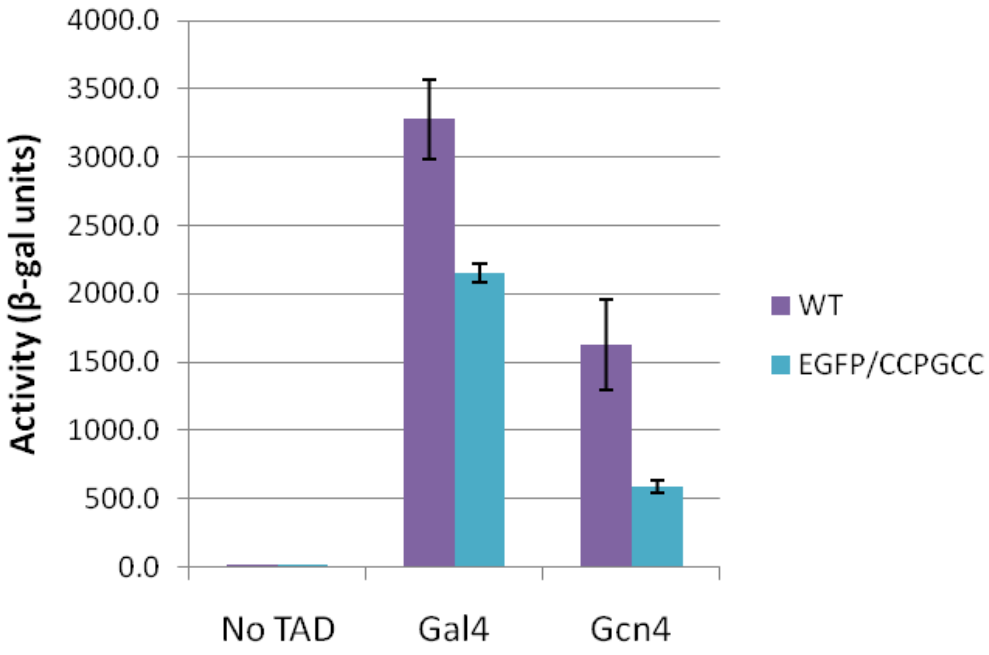


Figure A.3: Results from a liquid β-galactosidase assay for EGFP-tagged activators. The presence of protein tags on an activator-coactivator pair (i.e., an EGFP tag on the C-terminus of the TADs of Gal4 and Gcn4 and a FLN**CC**PG**CC**MEP tag on the N-terminus of Med15) results in a reduction in activity compared to that of WT levels (untagged proteins).

FLN**CC**PG**CC**MEP [21]. From Figure A.3, the Gal4 and Gcn4 activators are still able to upregulate transcription in the presence of the EGFP and CCPGCC tags (compared to Gal4(1-100)), although retaining 66% and 36% of their wild-type levels, respectively. However, we are currently determining the effect of linkers between the proteins and their tags on these results [21, 29].

C.2. ReAsH as a FRET acceptor

ReAsH is more completely 4,5-bis(1,3,2-dithiarsolan-2-yl)-resorufin; this molecule is not fluorescent when bound to 1,2-ethanediol, the form in which it is synthesized and stored, but becomes fluorescent when it complexes with the optimized tetracysteine motif

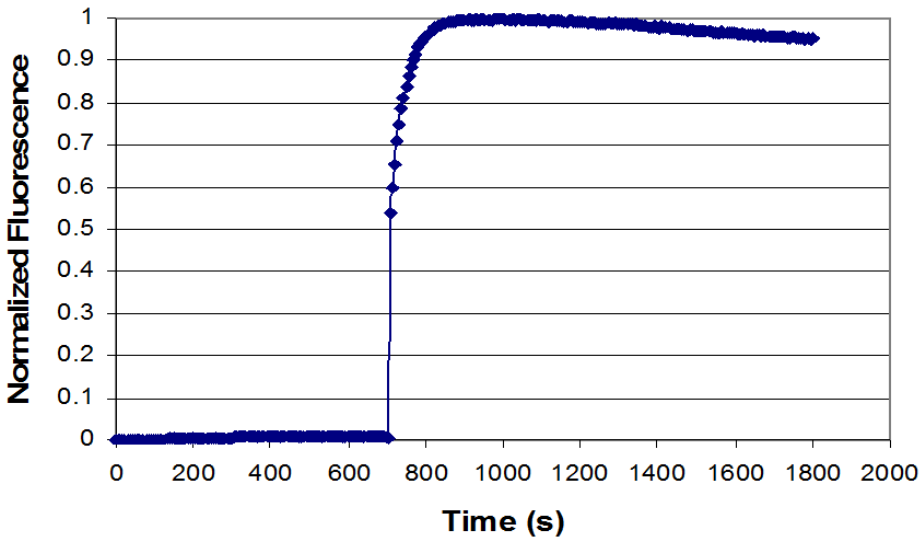


Figure A.4: Time course for the reaction of ReAsH-EDT2 with the tetracysteine peptide, FLNCCPGCCMEP. Fluorescence enhancement of ReAsH is observed upon binding to a tetracysteine peptide. The emission intensity of the following solution at 608 nm was monitoring continuously at 6 second intervals with an excitation at 597 nm: t = 0 sec, Buffer (100 mM MOPS (pH 7.2), 10 mM MES, 10 μM EDT); t = 120 sec, ReAsH compound was added; t = 700 sec, tetracysteine peptide was added.

FLN**CC**PG**CC**MEP [15, 21, 30] (Figure A.4). Therefore, in order to determine the ability of this molecule to label the same motif when attached to a protein localized within the nucleus of budding yeast, the FLN**CC**PG**CC**MEP tag was attached to the C-terminus of the EGFP-tagged Gcn4-derived activator. ReAsH labeling of this activator construct was pursued following previously reported protocols for implementation of the biarsenical-tetracysteine system in *Saccharomyces cerevisiae* [14, 31], in addition to varying such parameters as concentration of ReAsH (ranging from 1 – 10 μM) and labeling time (ranging from 15 minutes to 12 hours) using yeast in their logarithmic growth phase; however, the only fluorescence observed with the mCherry Ex/Em filter set was that of sporadically brightly stained dead or dying cells, probably by exposure of hydrophobic

sites that bind the dye (Figure A.5) [32]. Most likely, these results indicate that the ReAsH molecule is unable to permeate the cell wall of the live cells to reach its target [32].

D. Future Directions

Thus, in order to determine if the ReAsH molecule is able to permeate the yeast strain, labeling of a more abundant protein may be required, such as that of β-tubulin which can be visualized within microtubular structures upon labeling with FlAsH [14]. Furthermore, future studies will also focus on using this technology to label transcriptional proteins within mammalian cell culture. In particular, the abundance of structural data available on the interaction of that various domains of the mammalian coactivator CBP/p300 with transcriptional activators makes it the perfect model system to exploit in order to validate this approach [33-39].

E. References

1. Ptashne, M. and A. Gann, *Transcriptional activation by recruitment.* Nature, 1997. 386(6625): p. 569-577.
2. Ptashne, M. and A. Gann, *Genes & Signals*. 2001, New York: Cold Spring Harbor Laboratory.
3. Mapp, A.K. and A.Z. Ansari, *A TAD further: exogenous control of gene activation.* ACS Chem Biol, 2007. 2(1): p. 62-75.
4. Hall, D.B. and K. Struhl, *The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo.* J Biol Chem, 2002. 277(48): p. 46043-50.
5. Sekar, R.B. and A. Periasamy, *Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations.* J Cell Biol, 2003. 160(5): p. 629-33.
6. Berney, C. and G. Danuser, *FRET or no FRET: a quantitative comparison.* Biophys J, 2003. 84(6): p. 3992-4010.
7. Mayr, B.M., G. Canettieri, and M.R. Montminy, *Distinct effects of cAMP and mitogenic signals on CREB-binding protein recruitment impart specificity to target gene activation via CREB.* Proc Natl Acad Sci U S A, 2001. 98(19): p. 10936-41.
8. Bhaumik, S.R., et al., *In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer.* Genes Dev, 2004. 18(3): p. 333-43.
9. Jiang, F., et al., *Gene activation by dissociation of an inhibitor from a transcriptional activation domain.* Mol Cell Biol, 2009. 29(20): p. 5604-10.
10. Wightman, R., R. Bell, and R.J. Reece, *Localization and interaction of the proteins constituting the GAL genetic switch in Saccharomyces cerevisiae.* Eukaryot Cell, 2008. 7(12): p. 2061-8.
11. Giepmans, B.N., et al., *The fluorescent toolbox for assessing protein location and function.* Science, 2006. 312(5771): p. 217-24.
12. Waldo, G.S., et al., *Rapid protein-folding assay using green fluorescent protein.* Nat Biotechnol, 1999. 17(7): p. 691-5.
13. Dyachok, O., et al., *Oscillations of cyclic AMP in hormone-stimulated insulin-secreting beta-cells.* Nature, 2006. 439(7074): p. 349-52.
14. Andresen, M., R. Schmitz-Salue, and S. Jakobs, *Short tetracysteine tags to beta-tubulin demonstrate the significance of small labels for live cell imaging.* Mol Biol Cell, 2004. 15(12): p. 5616-22.
15. Adams, S.R., et al., *New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications.* J Am Chem Soc, 2002. 124(21): p. 6063-76.
16. Madani, F., et al., *Hairpin structure of a biarsenical-tetracysteine motif determined by NMR spectroscopy.* J Am Chem Soc, 2009. 131(13): p. 4613-5.
17. Hwang, R.D., C.C. Chen, and D.A. Knecht, *ReAsH: another viable option for in vivo protein labelling in Dictyostelium.* J Microsc, 2009. 234(1): p. 9-15.
18. Ju, W., et al., *Activity-dependent regulation of dendritic synthesis and trafficking of AMPA receptors.* Nat Neurosci, 2004. 7(3): p. 244-53.

19. Hoffmann, C., et al., *A FlAsH-based FRET approach to determine G protein-coupled receptor activation in living cells.* Nat Methods, 2005. 2(3): p. 171-6.
20. Nakanishi, J., et al., *FRET-based monitoring of conformational change of the beta2 adrenergic receptor in living cells.* Biochem Biophys Res Commun, 2006. 343(4): p. 1191-6.
21. Martin, B.R., et al., *Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity.* Nat Biotechnol, 2005. 23(10): p. 1308-14.
22. Gaietta, G.M., et al., *Golgi twins in late mitosis revealed by genetically encoded tags for live cell imaging and correlated electron microscopy.* Proc Natl Acad Sci U S A, 2006. 103(47): p. 17777-82.
23. Chen, B., et al., *Identification of an orthogonal peptide binding motif for biarsenical multiuse affinity probes.* Bioconjug Chem, 2007. 18(4): p. 1259-65.
24. Reeves, W.M. and S. Hahn, *Targets of the Gal4 transcription activator in functional transcription complexes.* Mol Cell Biol, 2005. 25(20): p. 9092-102.
25. Jeong, C.J., et al., *Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator.* Biochemistry, 2001. 40(31): p. 9421-7.
26. Park, J.M., et al., *In vivo requirement of activator-specific binding targets of mediator.* Mol Cell Biol, 2000. 20(23): p. 8709-19.
27. Fishburn, J., N. Mohibullah, and S. Hahn, *Function of a eukaryotic transcription activator during the transcription cycle.* Mol Cell, 2005. 18(3): p. 369-78.
28. Majmudar, C.Y., et al., *A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry.* Angew Chem Int Ed Engl, 2009. 48(38): p. 7021-4.
29. Sheff, M.A. and K.S. Thorn, *Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae.* Yeast, 2004. 21(8): p. 661-70.
30. Adams, S.R. and R.Y. Tsien, *Preparation of the membrane-permeant biarsenicals FlAsH-EDT2 and ReAsH-EDT2 for fluorescent labeling of tetracysteine-tagged proteins.* Nat Protoc, 2008. 3(9): p. 1527-34.
31. Rice, M.C., et al., *In vitro and in vivo nucleotide exchange directed by chimeric RNA/DNA oligonucleotides in Saccharomyces cerevisae.* Mol Microbiol, 2001. 40(4): p. 857-68.
32. Griffin, B.A., et al., *Fluorescent labeling of recombinant proteins in living cells with FlAsH.* Methods Enzymol, 2000. 327: p. 565-78.
33. Radhakrishnan, I., *Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions.* Cell, 1997. 91: p. 741-752.
34. Ema, M., et al., *Molecular mechanisms of transcription activation by HLF and HIF1alpha in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300.* Embo J, 1999. 18(7): p. 1905-14.
35. Van Orden, K., et al., *Binding of p53 to the KIX domain of CREB binding protein. A potential link to human T-cell leukemia virus, type I-associated leukemogenesis.* J Biol Chem, 1999. 274(37): p. 26321-8.
36. Goodman, R.H. and S. Smolik, *CBP/p300 in cell growth, transformation, and development.* Genes Dev, 2000. 14(13): p. 1553-77.

37. Chan, H.M. and N.B. La Thangue, *p300/CBP proteins: HATs for transcriptional bridges and scaffolds.* J Cell Sci, 2001. 114(Pt 13): p. 2363-73.
38. Ruas, J.L., L. Poellinger, and T. Pereira, *Role of CBP in regulating HIF-1-mediated activation of transcription.* J Cell Sci, 2005. 118(Pt 2): p. 301-11.
39. Ramirez, J., *Molecular Characterization of HTLV-1 Tax Interaction with the KIX Domain of CBP/p300.* Journal of Molecular Biology, 2007. 372: p. 958-969.