The Development of a Small Molecule Transcriptional Activation Domain

by

Brian B. Brennan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2007

Doctoral Committee:

       Associate Professor Anna K. Mapp, Chair
       Professor Mark M. Banaszak Holl
       Professor Carol A. Fierke
       Assistant Professor Jorge A. Iniguez-Lluhi

© Brian B. Brennan

Acknowledgements

I would like to start by thanking all of those who made impacts on my professional development before I began at the University of Michigan. Fr. Sweeney S.J., Dr. Tribble, and Dr. Graham were truly inspirational as teachers and their passion for science will not be soon forgotten. Dr. Mota de Freitas was an excellent mentor as an undergraduate researcher and gave me tremendous guidance throughout my time at Loyola University.

My research at the University of Michigan has been extremely collaborative over the years and there are numerous people to thank. I worked closest with Aaron Minter, Steve Rowe, and Sara Buhrlage. You have all been tremendous colleagues and friends and I could not have gotten this far without you. Amelia, Marcelle, Alex, Zikiya, Bin, Jen, and Garrette: you were all here at the start and made my time in graduate school a great experience. I also want to thank the rest of the Mapp lab for all the support throughout the years.

I would like to thank the members of my committee: Dr. Fierke, Dr. Banaszak Holl, and Dr Iniguez. Your thought provoking questions and insight have been invaluable. Additionally I would like to thank Dr. Ansari at the University of Wisconsin Madison and Dr. Kodadek at the University of Texas Southwestern for allowing me to carry out research in their labs while I was at Michigan, it proved incredible useful as witnessed by the work in this thesis. Most of all, I would like to thank Anna. You have

been a tremendous role model throughout my time in graduate school. I can not thank you enough for all that you have done for me while I have been in your lab. You have been a tremendous mentor and shaped the way I think about science. I only hope that I have a fraction of the success you have had as I begin my own faculty career.

Miguel, Big Hört, Jeffrey, Dan, and Tom: I'll miss our trips to Ashley's and Dominick's (to discuss science of course). Jen Furchak: thanks for keeping me motivated as we applied for faculty positions. I could not have done it without you. Now if we can just find jobs at the same school… Most of all though Jen thanks for introducing me to my wife: Sarah. Sarah, you are the most supportive, loving, and caring person I have ever met. Meeting you was the best thing that happened to me in graduate school and I look forward to spending the rest of our lives together.

Lastly, I would like to thank my parents and siblings. Mom, Dad, John, Jenny, Gayle, and Liz: Your constant support for all my endeavors will never be forgotten.

Table of Contents

List of Figures

Chapter I

Chapter II

Chapter III

List of Abbreviations

A	alanine
AD	activation domain
AdML	adenovirus major-late promoter
AEEA	8-amino-3,6-dioxaoctanoic acid
AH	amphipathic helix
ATCC	American Type Culture Collection
ATF	artificial transcription factor
ATF14	artificial transcription factor-14
ATF29	artificial transcription factor-29
BL21(DE3)	protease deficient *Escherichia coli* cell line with T7 RNA polymerase
bp	base pair
BSA	bovine serum albumin
BGG	bovine gamma-globulin
C	cysteine
cAMP	3', 5'-cyclic adenosine monophosphate
CBP	CREB-binding protein
CDK8	cyclin-dependent kinase-8
CH_3CN	acetonitrile
CME	chromatin modifying enzymes
CMV	cytomegalovirus
c-Myb	c-myeloblastosis oncogene
CREB	cAMP response element-binding protein
D	aspartic acid
Da	dalton
DBD	DNA binding domain
DD	dimerization domain
DMEM	Dulbecco's modified Eagle's medium
DMF	N, N-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E	glutamic acid
eDHFR	E. coli dihydrofolate reductase
EDTA	ethylenediamine tetraacetic acid
ELISA	enzyme-linked immunosorbant assay
EMSA	electrophoretic mobility shift assay
ESI-MS	electrospray ionization-mass spectroscopy
F	phenylalanine
FBS	fetal bovine serum

FITC	fluorescein isothiocyanate
FKBP	FK506 binding protein
Fmoc	9-fluorenylmethoxy carbonyl
g	gram
G	glycine
Gal4$_{dd}$	Gal4 dimerization domain
GST	glutathione-S-transferase
h	hour
H	histidine
HBTU	2-(1 H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCl	hydrochloric acid
HeLa	Henrietta Lacks cervical cancer cell line
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
hGR	human glucocortocoid receptor
HIV	human immunodeficiency virus
HOBt	N-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HTH	helix-turn-helix
Hsp90	heat shock protein 90
I	isoleucine
IC50	inhibitory concentration 50%
IPTG	isopropyl-beta-D-thiogalactopyranoside
K	lysine
KBP 1.66	KIX-binding peptide 1.66
KBP 2.20	KIX-binding peptide 2.20
KBPo2	KIX-binding peptoid 2
K$_D$	equilibrium dissociation constant
k$_{off}$	dissociation rate constant
KCl	potassium chloride
L	leucine
L	liter
LB	Luria-Bertani media
LBD	ligand binding domain
M	methionine
M	molar
mg	milligram
mL	milliliter
mM	millimolar
μg	microgram
μL	microliter
MDM2	mouse double-minute 2
Med15P	Med15 potentiator
Med15DM	Med15 double mutant
Mtx	methotrexate
MW	molecular weight

N	asparagine
Na_2CO_3	sodium carbonate
NaOH	sodium hydroxide
NEAA	non-essential amino acids
ng	nanogam
nm	nanometer
nM	nanomolar
NMR	nuclear magnetic resonance
NTP	nucleotide triphosphate
OD_{600}	optical density at 600 nm
OD_{420}	optical density at 420 nm
ONPG	o-Nitrophenyl-beta-galactopyranoside
Opti-MEM	reduced-serum modified Eagle's medium
OxDex	oxidized dexamethasone
P	proline
P	total protein concentration in (mg/mL)
P201	peptide 201
p53	protein 53 kilodaltons
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PNA	peptide nucleic acid
pLysE	plasmid encoding high levels of T7 lysozyme
pLysS	plasmid encoding low levels of T7 lysozyme
PSG	penicillin-streptomycin-glutamine
Q	glutamine
R	arginine
REST/NRSF	repressor element-1 silencing transcription factor/neuron restrictive silencer factor
RLU	relative light units
RNA	ribonucleic acid
RPM	rotations per minute
S	serine
s	second
SAGA	Spt-Ada-Gcn5-Acetyltransferase complex
SDOM	standard deviation of the mean
SDS	sodium dodecyl sulfate
SRB4	suppressor of RNA polymerase B 4
SV40	simian virus 40
T	threonine
t	time
TBP	TATA binding protein
TIS	triisopropyl silane
TFIIB	transcription factor IIB
TFA	trifluoroacetic acid
TFO	triplex forming oligonucleotide
Tris	Tris(hydroxymethyl)aminomethane

U	uracil
UAS	upstream activating sequence
UV-Vis	ultraviolet-visible
V	valine
VP-16	viral protein-16
V_s	volume of lysate added
V_T	total volume
W	tryptophan
XL_Y	LTGLFVQDYLLPTCIP
XL_R	LTGLFVQDRLLPTCIP
Y	tyrosine

Abstract

The Development of a Small Molecule Transcriptional Activation Domain

by

Brian Brennan

Chair: Anna K. Mapp

The strict regulation of transcription is essential for the survival of an organism. Nearly every human disease is associated with mis-regulated transcription as either a cause or an effect. This realization has brought about significant research efforts to both better understand the mechanism of eukaryotic transcription and to create small molecules capable of modulating this complex process.

Transcriptional activator proteins play a critical role in the control of gene expression. Minimally composed of a DNA binding domain and an activation domain, they are responsible for the levels and time-course of expression of a particular gene. The creation of activator artificial transcription factors (ATFs), non-natural replacements for activators, has provided mechanistic details about this process and have potential as therapeutics to reprogram aberrant patterns of gene expression. Chapter I details our current understanding of activators and progress that has been made in the creation of activator ATFs.

Features of activators that contribute to their potency (the levels of transcription elicited) are poorly understood. Until recently, it was believed that the strength of an activator was directly correlated with its affinity for coactivator targets. Chapter II details our studies of activator ATFs that target the coactivator Med15 in order to reveal other functional contributors to activator potency. Specifically, these studies demonstrated the importance of binding site location and secondary interactions that prevent non-productive binding and proteolytic degradation to the potency of a transcriptional activator.

Endogenous activators bind to an overlapping set of coactivator targets yet exhibit very little sequence homology outside of a general amphipathic composition. The permissive nature of coactivator binding sites led us to design a series of small molecules which presented functionality commonly observed in amphipathic activators to be used as an activation domain. Chapters III and IV detail the cell-free and cell-based assays which demonstrate the utility of an isoxazolidine scaffold for this purpose. Remarkably, they represent the first small molecule activation domain with activity in living cells. Early mechanistic work suggests that they function by a mechanism analogous to their natural protein counterparts. Future studies of these molecules include in-depth mechanistic investigations as well as structural optimizations to increase potency, experiments that are detailed in Chapter V.

Chapter I

Current Picture of Transcriptional Activators

A. Significance

Even though every cell in the human body contains the same genetic information, each tissue is able to specialize to carry out specific functions. While the DNA code of a cell in the brain is the same as one in the heart for example, each is able to carry out its own unique task. To understand the origin of this cell type differentiation, one must look not only at the genetic code, but also at how this code is read, or transcribed. It is the differential expression of genes that accounts for cell type specificity and in turn, the vast specialization of each cell in the human body.

In 2001, the sequencing of the human genome was completed, putting forth the framework for an unprecedented understanding of the link between genetics and disease.[1] In fact, misregulation of gene expression is linked to almost every human disease, either as the cause for or effect of a particular disorder.[2,3] The discovery of molecules to reprogram patterns of gene expression would result in the development of powerful tools for the treatment of countless human diseases. While this long term goal certainly remains on the distant horizon, significant strides have been made towards its end.[4] The project herein describes efforts towards the development of a small molecule transcriptional activation domain, a molecule capable of upregulating gene expression when localized to DNA.

B. Modularity of Transcriptional Activators

In eukaryotic cells, transcriptional regulator proteins are responsible for the levels and time-course of gene expression.[5] Transcriptional regulators are capable of binding to a specific site on DNA and effecting gene transcription. Repressors typically inhibit transcription by recruiting corepressor proteins and, conversely, activators up-regulate transcription through binding and recruiting a host of coactivators. The proteins recruited by activators are responsible for modifying histones, allowing access to a particular gene, as well as forming the pre-initiation complex with RNA polymerase II, the first key step in gene transcription (Figure I-1).[4]
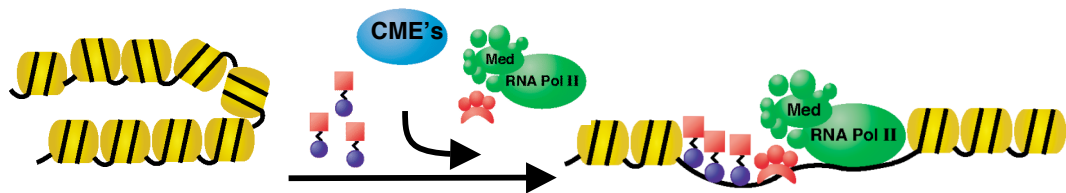
Figure I-1: Activators bind to DNA and recruit chromatin modifying enzymes (CME) and the RNA polymerase holoenzyme to a specific gene in order to up-regulate transcription.

Activator proteins are composed of at least two distinct domains. The DNA binding domain (DBD) is responsible for localizing the regulator to a specific site on DNA, imparting the majority of the gene targeting, and the activation domain (AD) determines the level of gene transcription through its contacts with other proteins.[6] The two domains are connected through covalent or non-covalent interactions in order to regulate a specific gene. There is often a linker region between the two domains that allows dimerization of the regulator, leading to increased potency. The separable nature of these domains was first demonstrated by the Ptashne lab in a series of "domain

2

swapping" experiments conducted in *Saccharomyces cerevisiae*.[7] In order to determine if

the two domains could function independently, a hybrid protein was constructed with the

DBD of the bacterial repressor LexA and the AD of the yeast activator protein Gal4, a

potent activator responsible for regulating the galactose metabolism genes in yeast. The

hybrid protein was able to activate a reporter gene bearing binding sites for the bacterial

repressor to the levels of the activator Gal4. Furthermore, activation was no longer

observed in genes bearing the Gal4 binding site. The separable nature of these two

domains suggested a straightforward mechanism for the construction of activator

artificial transcription factors (ATFs) incorporating the binding specificity of one protein

with the activity of another (Figure I-2).
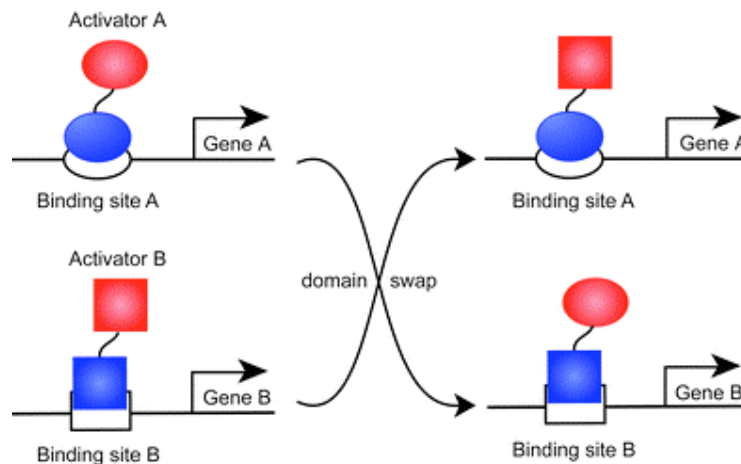
Figure I-2: The DNA binding domain (DBD) of an activator (blue) determines the gene to be modulated while the activation domain (AD) (red) determines potency. This allows for the creation activator ATFs with unique specificity and potency.

C. Transcription and Disease

As previously stated, aberrant patterns of gene expression can often lead to

disease. One disorder in which this link has been thoroughly investigated is

medulloblastoma, one of the most malignant pediatric brain tumors.[8] It is believed that this condition arises from undifferentiated cells in the cerebellum. These cells are often observed to overexpress repressor element-1 silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF), a global transcriptional repressor responsible for inhibiting transcription of most of the neuronal differentiation genes by binding to a 23bp consensus sequence. This protein, which is expressed at high levels in nearly all non-neuronal cells in the body, inhibits transcription through the recruitment of corepressor proteins such as CoREST, Sin3A, and histone deacetylaces (Figure I-3a).[8] In order to compete with this overactive repressor, Majumder and coworkers decided to infect medulloblastoma cells with an adenovirus harboring a chimeric activator ATF composed of the DBD of REST/NRSF and the AD of VP16, a potent viral coactivator protein.[9] It was thought that by taking advantage of the modular nature of transcriptional regulators and activating the genes that are being inhibited by overexpressed REST/NRSF, the neuronal cells could then differentiate and prevent tumorgenesis. While this was not the exact result that was observed, the REST-VP16 conjugate did lead to apoptosis in tumor cells both in vitro and in vivo (Figure I-3). It was speculated that the genetic instability of the medulloblastoma cell lines precluded differentiation, and in turn lead to cell death.[9]

a)   b)
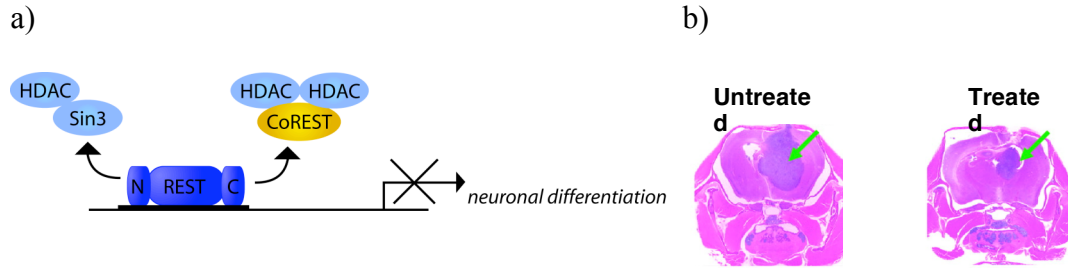
Figure I-3: (a) Over-expression of REST, a transcriptional repressor, is often observed in medulloblastoma. REST expression inhibits neuronal differentiation genes through the recruitment of a host of co-repressor proteins. (b) Treatment with REST-VP16 activates these inhibited genes leading to cell apoptosis and decreased tumor growth. Figure from (9).

The ability of this exogenous activator ATF to counteract an endogenous repressor is a remarkable discovery and truly shows the potential of transcription-targeted therapeutics. Further, the development of a small molecule which could be used in place of the chimeric protein is likely to have enormous advantages in terms of delivery and cellular stability.

D. Transcriptional Activators

The discovery of the modular nature of transcriptional activators initiated an intense effort to structurally and functionally characterize the DBD as well as the AD of natural activators in order to eventually create artificial replacements. Initial efforts towards characterizing the DBDs were quite successful leading to the discovery of numerous structural motifs, for example the helix-turn-helix motif observed in LexA and the $Zn(II)_2Cys$ binuclear cluster of Gal4.[10-12] Structural characterization of the ADs proved to be much more challenging. Activating regions generally remain unstructured in solution and become structured only upon binding their protein targets. Additionally, despite targeting similar coactivators, they exhibit very little sequence homology.[13-17]

Therefore, instead of structural categories, activators are typically characterized by the preponderance of certain amino acid residues such as proline-rich, glutamine-rich, or acid-rich.[5,18-20] The most well studied activators, including the yeast activator Gal4, typically fall into the class of amphipathic or acid-rich activators. These activation domains are found to contain a number of acidic and polar residues interspersed with hydrophobic amino acids (Figure I-4). It is thought that while the polar residues are important for interacting with the solvent, many of the most important protein contacts involve the hydrophobic groups. For example, mutations in any of three hydrophobic residues in the GCN4 central acidic activation domain (CAAD) (M107, L123, and F124) to alanine greatly diminish its ability to activate transcription *in vivo* and its ability to bind to coactivator targets *in vitro*.[21,22]

a)   b)

VP16
441 DFDLDMLG 448
p53
9   SVAPPLSQETFSDLW 23
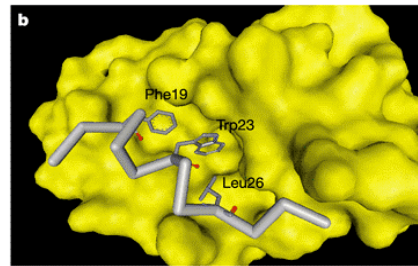
Figure I-4: a) Acidic activation domains are composed of interspersed polar (red) and hydrophobic (green) residues. b) Althoguh unstructured in solution, the p53 activation domain forms a helix upon binding to one of its protein targets (Mdm2). Figure Reprinted form (16) with permission.

Further obscuring mechanistic details of transcriptional activators is their promiscuous binding profiles that are often observed. While the DBD of an activator acts through targeting a specific binding site, the activation domain is capable of binding to numerous protein targets.[23] For instance, the Gal4 activation domain has been shown to interact *in vitro* with the transcriptional machinery proteins TBP, Med15, Cdk8, Srb4, SWI/SNF, and others as well as the chromatin modifying complex SAGA.[24] Attempts at

determining the functionally relevant targets are quite challenging due to the redundancy of the various binding surfaces. Disruption of one interaction can often be compensated with a non-physiologically relevant one. While the abundance of protein targets within and outside the transcriptional machinery makes constructing a detailed picture of activator function challenging, it is likely a necessary feature of these regulators. The initiation of transcription involves interactions between dozens of different proteins all of which must localize to a specific gene in a particular way.

E. Regulation of Activator Interactions

In addition to targeting coactivators, ADs also bind to proteins which control the time-course of gene activation. For example, Gal4 has also been shown to bind to Sug1 and Sug2, two components of the proteosome.[25,26] This interaction likely leads to degradation of the activator and a decrease in activated transcription. If the cell is to maintain strict control of gene expression, it must allow activation only when the gene product is needed and degradation of the activator serves as a useful mechanism.

Another method to prevent uncontrolled activation is through the binding of masking proteins. Masking proteins bind to an activator, preventing it from making key contact and/or trafficking to the nucleus. The usefulness of masking proteins as regulatory elements can be observed with the Gal4-Gal80 interaction. In the absence of galactose, when the galactose metabolism gene products are not needed, the Gal4 activation domain is tightly bound to Gal80 and incapable of activating transcription. However, the presence of galactose causes the Gal3 protein to interact with Gal80, leading to a conformational change and thereby allowing Gal4 to activate transcription.

This system serves to both prevent unnecessary transcription and to protect Gal4 from degradation.[27] Although inter-molecular masking is most common, there are also several examples of intra-molecular masking. For example, the Leu3 activator protein, responsible for the induction of leucine biosynthesis genes, contains an intra-molecular binding site for its activation domain.[28] This allows the protein to induce gene expression only in the presence of a necessary precursor. This intra-molecular masking has also since proven to be an extraordinary design feature for the creation of novel artificial transcriptional activators.[29]

F. Activation Domain Replacement

Transcription-based therapeutics capable of regulating specific genes to predetermined levels have the potential to treat a host of ailments. Towards this end, artificial replacements for both the DBD as well as the AD of transcriptional activators have been and continue to be investigated. Much of the early work towards the development of artificial replacements for transcriptional activators focused on the replacement of the DBD.[6] Since it is the DBD that imparts the majority of the specificity of transcriptional activators, discovering replacements capable of targeting specific DNA sequences is extremely important. Early work toward elucidating the structure and function of natural DBD was quite successful, leading to a detailed picture of DNA binding from a number of natural transcriptional activators as well as repressors (Figure I-5).[10] In order to reprogram aberrant patterns of gene expression, one could use a unique DBD capable of targeting a specific site on DNA coupled to appropriate activation or inhibition module.
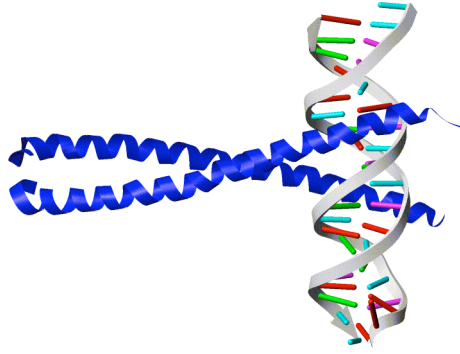
Figure I-5: The structure of the GCN4 DBD bound to its requisite DNA target. Figure reprinted from (30) with permission from Elsevier.

Genetic selection was initially used in order to create zinc finger protein DBDs capable of targeting unique sites within the genome.[31] Phage display was carried out in order to elucidate zinc fingers capable of binding to unique tri-nucleotide sequences. These zinc fingers could then be combined to target more specific sites. In theory, six zinc fingers would be capable of targeting an 18bp sequence, which would be specific within 68 billion base pairs, or 20 human genomes. While this method was found to be successful, the creation of non-peptidic DBDs carries more potential to be used as therapeutics. Because peptidic DBDs are poteolyzed in the cell, synthetic DBDs would have much longer lifetimes *in vivo* and therefore could much more readily be converted into drugs.

Efforts towards creating non-peptidic DBD were also quite successful. Triplex-forming oligonucleotides (TFOs) capable of recognizing unique DNA sequences through formation of Hoogsteen base pairs with double stranded DNA and peptide nucleic acids (PNAs) in which the phosphodiester bonds in the backbone have been replaced with amides were the first successful replacements discovered (Figure I-6).[32,33] Both of these

artificial replacements are able to bind to DNA in a highly specific manner. In fact, there

are numerous examples of activator ATFs in which these classes of DBDs are used.

a)

b)
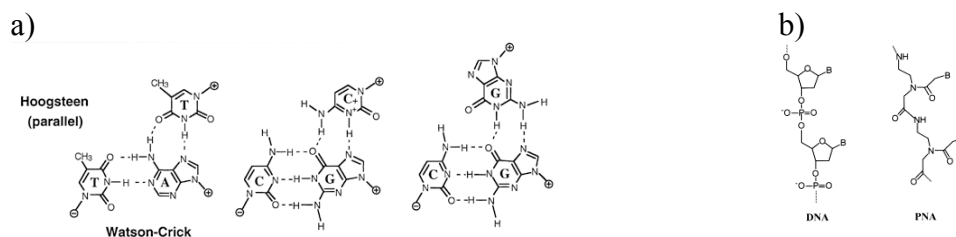
Figure I-6: (a) Triplex forming oligonucleotides take advantage of Hoogsteen base pairs in order to specifically interact with double stranded DNA. (b) Peptide nucleic acids contain an amide backbone in place of the sugar of DNA. This increases stability while allowing specific interactions with DNA and RNA. Figures from (31, 32).

Perhaps the most notable replacement for the DBD was discovered by the Dervan group

and was modeled after the natural product distamycin. These polyamide small molecules

are capable of recognizing and binding each of the four possible DNA base pair

combinations (Figure I-7).[34] They bind to the minor groove of DNA in a manner that

does not severely affect DNA's overall structure and have also been shown to activate

transcription when coupled to an appropriate AD.[35]

a)

b)
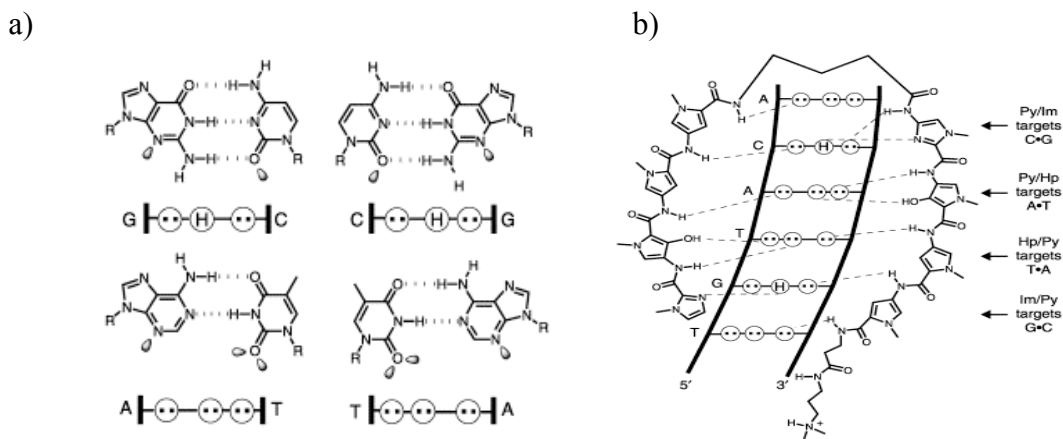
Figure I-7: (a) Pyrrole-imidazole polyamides are capable of binding to each of the four DNA base pair combination through the asymmetric hydrogen bond donors and acceptors

presented. (b) The polyamide binds as a dimer in the minor groove of DNA causing little distortion on the overall structure of DNA. Figure from (33).

The construction of artificial replacements for the DBD has been made possible by the vast structural data available on their natural counterparts. This has not, however, been the case for developing artificial replacements for the AD. As previously mentioned, AD's typically remain unstructured in solution and show significant secondary structure only when bound to a protein target. This has made it very challenging to determine the necessary structural motifs for activator binding and to incorporate them into small molecule replacements. Furthermore, the AD is responsible for far more than a simple binding interaction. It is the AD that controls the time course and levels of gene expression, so it must be able to respond to extra-cellular signals, undergo covalent modifications, traffic to the nucleus, and activate a gene to a predetermined level.[6] Integrating even some of these properties into small molecule replacements is certainly a formidable challenge and, in fact, at the onset of this research, there was not a single report of a small molecule capable of activating gene transcription.

G. Small Peptidic Activation Domains

Initial efforts towards creating artificial activation domains relied on the use of known activating regions from endogenous proteins. While the specific structures of activating regions has been very challenging to determine, the regions responsible for activating transcription in a particular activator protein could be determined from deletion experiments.[36] This allowed for the construction of the first artificial activators composed of protein DBD's coupled to peptidic AD's.[6] With further study, the peptidic

activation modules could be refined to very small sequences, such as the eight residues of the AD VP1 or the 14 residue AD ATF14, both derived from the potent viral co-activator VP16 (Figure I-8).[37] Acidic ADs are often composed of surreptitious repeats of eight to ten amino acids and the minimal repeat itself can function as an AD when fused to a DBD. A relatively small number of contacts are therefore necessary for an AD to interact with its coactivator target. This can also be observed with the interaction between Mdm2 and p53 (Figure I-4) where three key contacts are needed for the binding interaction.[16]

The p53●Mdm2 interaction also demonstrates another property of acidic activation domains: their propensity to form helices upon binding to their protein targets. This information in concert with the promiscuous nature of activator binding led the Ptashne lab to design a peptide predicted to form an amphipathic helix and to test its activity *in vivo*.[35,38] As many different activation domains are able to target an overlapping set of sites within the transcriptional machinery, it was thought that a peptide capable of loosely mimicking a motif observed in acidic activators would also be able to activate transcription. Remarkably, this 15 amino acid peptide named AH, for amphipathic helix, was able to activate transcription robustly in yeast.[38] This was the first example of a rationally designed AD and gave tremendous insight into the nature of natural activators; that presenting an amphipathic face was sufficient to form interactions with coactivator targets. If a small peptide designed to form a specific structural motif could activate transcription, then reproducing this affect with a small molecule could be envisioned (Figure I-8).

a)   b)

VP1:   DFDLDMLG

ATF14: CGSDALDDFDLDML

AH: ELQELQELQALLQQQ
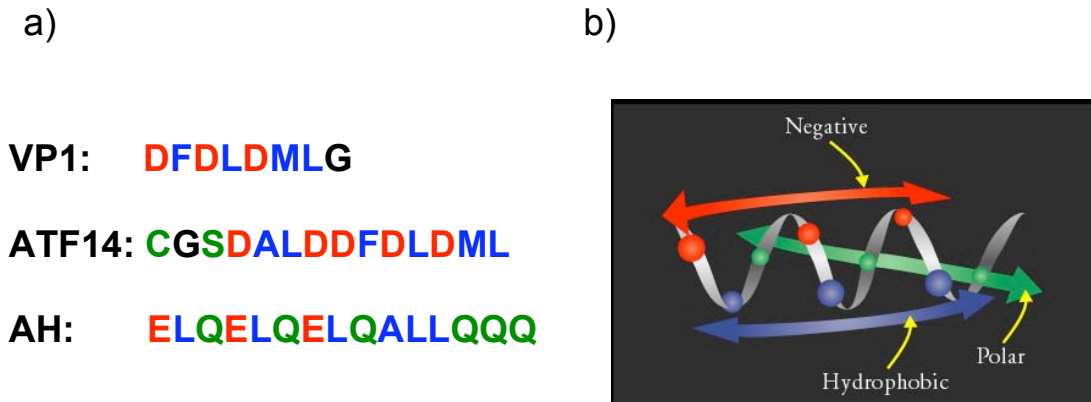
Figure I-8: (a) Amino acid sequence of two minimal activation modules (VP1 and ATF14) and a sequence designed to form an amphipathic helix (AH) with charged residues in red, polar in green, and hydrophobic in blue. All three of these sequences activate transcription when coupled to a DBD. (b) Model of AH, demonstrating its amphipathic character.

A significant amount of information about activation domains demonstrates their promiscuous binding nature, but it was a set of "activator bypass" experiments that showed targeting a single protein in the transcriptional machinery could suffice to activate transcription.[39,40] In these experiments, Med15 was expressed as a fusion protein with an endogenous DBD. Med15 is a component of the mediator, a complex which interacts directly with RNA Polymerase II and has since been shown to be a target of both natural and artificial AD's.[41] This hybrid protein was able to act as an activator ATF, demonstrating that localizing a single component of the transcriptional machinery to a gene was sufficient to activate transcription. This result was both an exciting discovery about activator function and also led the way to the development of numerous novel AD's which could uncover more important properties of activator proteins.

With this in mind, peptide selections have been used to create novel activation domains. In a peptide selection, the protein of interest is first immobilized on solid support and libraries of phage or bacteria expressing peptides on their coat proteins are

then added to the immobilized protein.[42,43] If the organism is capable of binding to the

protein of interest, it is trapped on solid phase and can be amplified. Through multiple

rounds of panning, followed by DNA sequencing, peptides capable of binding to the

protein with high affinity can be discovered. This approach was taken by Montminy and

coworkers to discover peptides capable of binding to the KIX domain of CBP, a

component of the eukaryotic transcriptional machinery.[44] The KIX domain has been

shown to be a target of numerous transcriptional activators including jun, tax, c-Myb,

HIV-tat, and others.[45,46] If small peptides were capable of making similar interactions

with KIX, then they should suffice as an AD. As expected, potent activation was

observed in mammalian cells when KIX binding peptides were localized to DNA. Upon

further analysis, these peptide ligands could present amphipathic nature when modeled to

a helical wheel (Figure I-9).[44] This result further demonstrated the ability to replace

protein ADs with small amphipathic molecules.

a)

KBP 1.66 SSLVDLIL

KBP 2.20 WAVYELLF

C-Myb   IKELELLL

b)

KBP 2.20   C-Myb
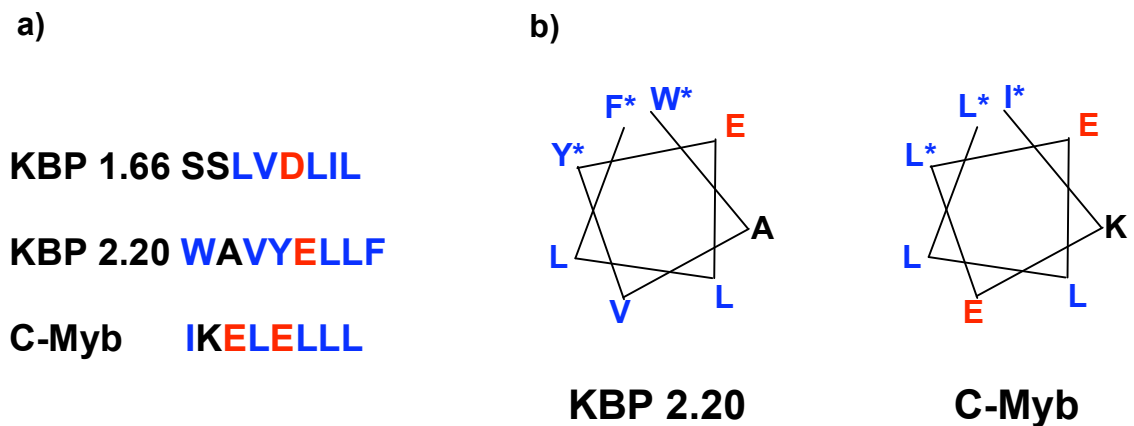
Figure I-9: (a) Sequences of KBP 1.66 and KBP 2.20, the peptides discovered through phage display against the KIX domain, as well as the sequence of c-Myb, an endogenous activator which targets the KIX domain. Acidic residues are in red and large hydrophobic residues in blue. (b) Helical wheel diagram demonstrating the hydrophobic face (starred residues) which interacts with the KIX domain.

Phage display also proved to be a useful technique for understanding the link between interactions with masking proteins and interactions important for transcriptional activation. As previously mentioned, in order to prevent nonspecific interactions and premature degradation and allow strict regulation of transcription, masking proteins which bind to activation regions are often employed. While it had been shown that ADs were capable of interacting with numerous components of the transcriptional machinery, chromatin modifying enzymes, and masking proteins, it was not known whether the same chemical interactions were used in all cases. To probe this question, the Kodadek lab used phage display to discover peptides which could interact with the Gal80 repressor.[47] If these peptides bound to a similar surface of Gal80 as the Gal4AD, one would expect them also to activate transcription. Alternatively, if Gal80 blocks Gal4 transcription by binding to a surface other than the AD, the ligands would be unlikely to activate. Kodadek found that these peptide ligands were capable of upregulating transcription and could prevent Gal4AD from binding to Gal80, demonstrating that the same surfaces of Gal4 interact with both coactivators and masking proteins. Further evidence to support this hypothesis was also revealed when it was shown that the Gal80 binding peptide interacted with the coactivator Med15, one of the targets of Gal4.

H. Non-Peptidic Transcriptional Activators:

Although peptidic activators have certainly provided a plethora of information about factors affecting potency, important binding interactions, and structural motifs of ADs, the next step towards transcription based therapeutics lies in the advancement of

non-peptidic replacements for the AD module. The first example of a non-peptidic

transcriptional activator was reported by the Ptashne lab in 2003.[48] As mentioned,

transcriptional activators must interact with DNA and components of the transcriptional

machinery. As numerous aptamers, DNA or RNA oligomers capable of binding to

protein surfaces, have been described, the Ptashne lab thought it might be possible to take

advantage of the binding properties of RNA and create an RNA AD. Indeed, a series of

hairpin oligonucleotides were capable of upregulating transcription in yeast when

localized to DNA. This remarkable finding further confirmed the permissive nature of

activator binding sites and demonstrated the utility of creating non-peptidic AD's.

Since this discovery, two other small molecule ADs and one peptoid AD have

been described. The first came from the Mapp research lab and is discussed in Chapter

III. This involves the use of an isoxazolidine scaffold bearing functionality often

observed in endogenous AD's.[49,50] This scaffold was chosen due to its ease of synthesis

and its ability to project functionality in a three dimensional array, similar to the helices

of natural AD's formed upon binding to their protein targets (Figure I-10).[51] Through the

small molecule's interactions with the shallow, permissive binding surfaces of

coactivators, it is able to activate transcription when localized to DNA. This activation is

observed both *in vitro* as well as in cell culture.

a)   b)
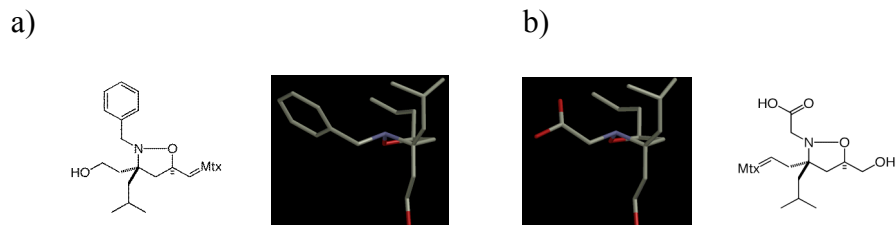
Figure I-10: The rigid isoxazolidine scaffold is capable of displaying varied
functionality which generally mimics common residues of endogenous activators.

Shortly after we reported this small molecule AD, the Uesugi lab described another molecule capable of modulating transcription. This molecule, named wrenchnolol, was initially discovered through a combination of rational design and screening for compounds capable of disrupting the interaction between the ESX activation domain and one of its coactivator targets Sur2 (Figure I-11a). Through binding of Sur2, wrenchnolol was capable of disrupting activated transcription by ESX and to a lesser extent, other activation domains. When wrenchnolol was localized to DNA through a polyamide DBD it was able to activate transcription, likely through these same binding interactions, yet its lack of cell permeability prevented its activity in cells.

The third example of a non-biopolymer activation domain was reported by the Kodadek lab in 2005. A peptoid screen was employed to discover molecules that could interact with the KIX domain of the coactivator CBP, a target of numerous natural activators (Figure I-11b). Peptoids, N-substituted glycine oligomers, are resistant to proteolysis and have proven useful as protein binding agents due to their unique secondary structure and the diverse functionality that can be incorporated within them. The peptoid screen identified two molecules capable of binding to the KIX domain, although one of them was extremely hydrophobic and appeared to interact solely through nonspecific hydrophobic interaction. The second peptoid, KBPo2, robustly activated transcription in cells when coupled to a protein DBD.[52]
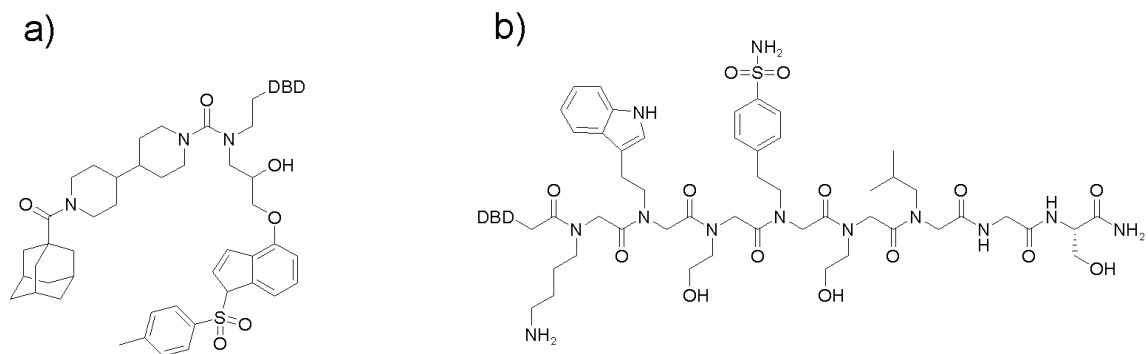
Figure I-11: a) Wrenchnolol and b) KBPo2 activate transcription when localized to DNA.

A more recent report by the Kodadek lab reveals a smaller peptoid screened against the same protein which upregulates transcription when coupled to a polyamide DBD, representing a fully non-biopolymer activator ATF.[53]

From the first reports of the unique, separable domains of transcriptional activators to the more recent replacements of both with small molecules, our view of transcriptional activators has changed dramatically over the last 20 years. We now have the ability to create artificial transcriptional activators composed of numerous different non-natural domains capable of controlling transcription of many different genes with varying potencies (Figure I-12).
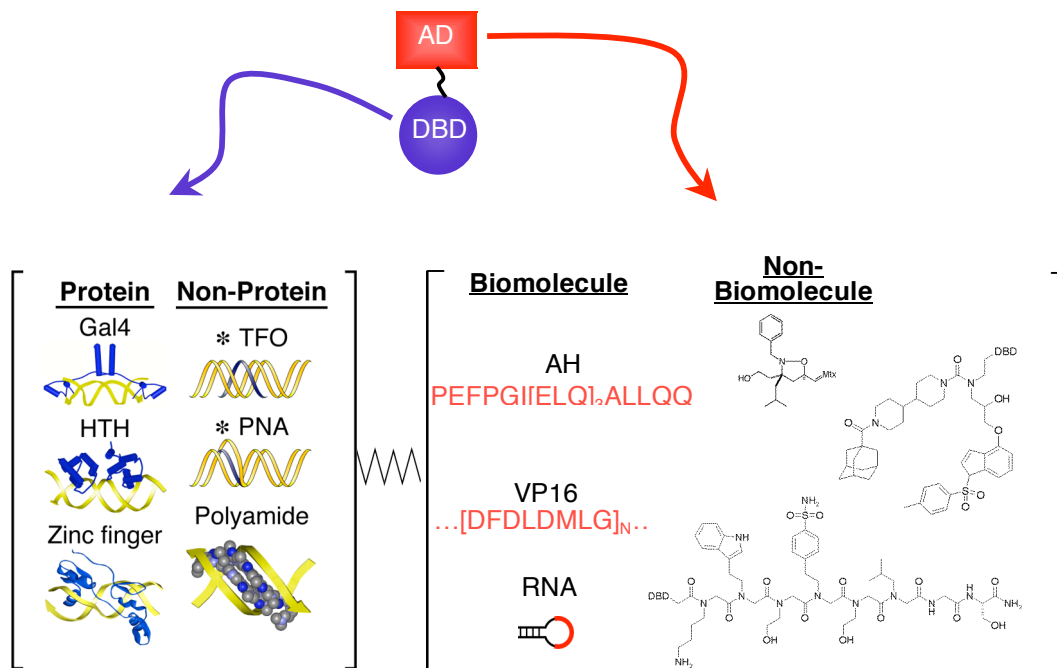
Figure I-12: Transcriptional activators are composed of a DBD and AD. Numerous artificial counterparts have been described for each domain resulting in the creation of activator ATFs.

In order to further expand available ADs, my research has focused on the development of a small molecule-based AD. Creating this activator required that a number of questions about the origin of activator potency be addressed. Initially we used small peptides to elucidate the effects of binding affinity and secondary interaction on activator potency (Chapter II). With this information in hand our lab designed and synthesized a small molecule we believed could modulate transcription (Chapter III). Indeed, this resulted in the first small molecule AD. After determining the activity of these small molecules *in vitro*, my work has centered on understanding how this small molecule could reconstitute the activity of a protein AD and furthermore if activation could be observed in a cellular context (Chapter IV). Our tremendous success with this

approach has not only expanded the repertoire of available AD's but has also paved the

way for other scaffolds designed to regulate transcription.

I. References

(1)   Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendl, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J. F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M. *Nature* **2001**, *409*, 860-921.

(2)   Chen, X.; Cheung, S. T.; So, S.; Fan, S. T.; Barry, C.; Higgins, J.; Lai, K. M.; Ji, J.; Dudoit, S.; Ng, I. O.; Van De Rijn, M.; Botstein, D.; Brown, P. O. *Mol Biol Cell* **2002**, *13*, 1929-39.

(3)   Perou, C. M.; Sorlie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Rees, C. A.; Pollack, J. R.; Ross, D. T.; Johnsen, H.; Aksien, L. A.; Fluge, O.; Pergamenschikov, A.; Williams, C.; Zhu, S. X.; Lonning, P. E.; Borresen-Dale, A. L.; Brown, P. O.; Botstein, D. *Nature* **2000**, *406*, 747-752.

(4)   Mapp, A. K.; Ansari, A. Z. *Acs Chemical Biology* **2007**, *2*, 62-75.

(5)   Ptashne, M.; Gann, A. *Genes & Signals*; Cold Spring Harbor Laboratory: New York, 2001.

(6)   Ansari, A. Z.; Mapp, A. K. *Curr Opin Chem Biol* **2002**, *6*, 765-772.

(7)   Brent, R.; Ptashne, M. *Cell* **1985**, *43*, 729-736.

(8)   Lawinger, P.; Venugopal, R.; Guo, Z. S.; Immaneni, A.; Sengupta, D.; Lu, W.; Rastelli, L.; Marin Dias Carneiro, A.; Levin, V.; Fuller, G. N.; Echelard, Y.; Majumder, S. *Nat Med* **2000**, *6*, 826-31.

(9) Fuller, G. N.; Su, X.; Price, R. E.; Cohen, Z. R.; Lang, F. F.; Sawaya, R.; Majumder, S. *Mol Cancer Ther* **2005**, *4*, 343-9.

(10) Garvie, C. W.; Wolberger, C. *Mol Cell* **2001**, *8*, 937-46.

(11) Luscombe, N. M.; Austin, S. E.; Berman, H. M.; Thornton, J. M. *Genome Biol* **2000**, *1*, REVIEWS001.

(12) Pan, T.; Coleman, J. E. *Proc Natl Acad Sci U S A* **1990**, *87*, 2077-81.

(13) Sigler, P. B. *Nature* **1988**, *333*, 210-2.

(14) Ferreira, M. E.; Hermann, S.; Prochasson, P.; Workman, J. L.; Berndt, K. D.; Wright, A. P. *J Biol Chem* **2005**, *280*, 21779-84.

(15) Giniger, E.; Ptashne, M. *Nature* **1987**, *330*, 670-672.

(16) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948-953.

(17) Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L. *Science* **1997**, *277*, 1310-1313.

(18) Courey, A. J.; Holtzman, D. A.; Jackson, S. P.; Tjian, R. *Cell* **1989**, *59*, 827-836.

(19) Gerber, H. P.; Seipel, K.; Georgiev, O.; Hofferer, M.; Hug, M.; Rusconi, S.; Schaffner, W. *Science* **1994**, *263*, 808-811.

(20) Mermod, N.; Oneill, E. A.; Kelly, T. J.; Tjian, R. *Cell* **1989**, *58*, 741-753.

(21) Drysdale, C. M.; Jackson, B. M.; McVeigh, R.; Klebanow, E. R.; Bai, Y.; Kokubo, T.; Swanson, M.; Nakatani, Y.; Weil, P. A.; Hinnebusch, A. G. *Molecular and Cellular Biology* **1998**, *18*, 1711-1724.

(22) Jackson, B. M.; Drysdale, C. M.; Natarajan, K.; Hinnebusch, A. G. *Molecular and Cellular Biology* **1996**, *16*, 5557-5571.

(23) Hall, D. B.; Struhl, K. *J Biol Chem* **2002**, *277*, 46043-50.

(24) Traven, A.; Jelicic, B.; Sopta, M. *EMBO Rep* **2006**, *7*, 496-9.

(25) Chang, C.; Gonzalez, F.; Rothermel, B.; Sun, L.; Johnston, S. A.; Kodadek, T. *J Biol Chem* **2001**, *276*, 30956-63.

(26) Archer, C. T.; Burdine, L.; Kodadek, T. *Mol Biosyst* **2005**, *1*, 366-72.

(27) Lohr, D.; Venkov, P.; Zlatanova, J. *FASEB J* **1995**, *9*, 777-87.

(28)   Wang, D.; Zheng, F.; Holmberg, S.; Kohlhaw, G. B. *J Biol Chem* **1999**, *274*, 19017-24.

(29)   Lum, J. K.; Majmudar, C. Y.; Ansari, A. Z.; Mapp, A. K. *Acs Chemical Biology* **2006**, *1*, 639-643.

(30)   Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223-37.

(31)   Beerli, R. R.; Barbas, C. F. *Nature Biotechnology* **2002**, *20*, 135-141.

(32)   Faria, M.; Giovannangeli, C. *Journal of Gene Medicine* **2001**, *3*, 299-310.

(33)   Nielsen, P. E. *Methods Enzymol* **2001**, *340*, 329-40.

(34)   Dervan, P. B. *Bioorg Med Chem* **2001**, *9*, 2215-35.

(35)   Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 3930-3935.

(36)   Ma, J.; Ptashne, M. *Cell* **1987**, *48*, 847-853.

(37)   Stanojevic, D.; Young, R. A. *Biochemistry* **2002**, *41*, 7209-16.

(38)   Ma, J.; Ptashne, M. *Cell* **1987**, *51*, 113-119.

(39)   Gaudreau, L.; Keaveney, M.; Nevado, J.; Zaman, Z.; Bryant, G. O.; Struhl, K.; Ptashne, M. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 2668-2673.

(40)   Cheng, J. X.; Gandolfi, M.; Ptashne, M. *Curr Biol* **2004**, *14*, 1675-9.

(41)   Myers, L. C.; Kornberg, R. D. *Annual Review of Biochemistry* **2000**, *69*, 729-749.

(42)   Smith, G. P.; Petrenko, V. A. *Chem Rev* **1997**, *97*, 391-410.

(43)   Lu, Z.; Murray, K. S.; Van Cleave, V.; LaVallie, E. R.; Stahl, M. L.; McCoy, J. M. *Biotechnology (N Y)* **1995**, *13*, 366-72.

(44)   Frangioni, J. V.; LaRiccia, L. M.; Cantley, L. C.; Montminy, M. R. *Nature Biotechnology* **2000**, *18*, 1080-1085.

(45)   Vendel, A. C.; Lumb, K. J. *Biochemistry* **2004**, *43*, 904-908.

(46)   Vendel, A. C.; McBryant, S. J.; Lumb, K. J. *Biochemistry* **2003**, *42*, 12481-12487.

(47)   Han, Y.; Kodadek, T. *The Journal of Biological Chemistry* **2000**, *275*, 14979-14984.

(48)   Saha, S.; Ansari, A. Z.; Jarrell, K. A.; Ptashne, M. *Nucleic Acids Research* **2003**, *31*, 1565-1570.

(49)   Buhrlage, S. J.; Brennan, B. B.; Minter, A. R.; Mapp, A. K. *Journal of the American Chemical Society* **2005**, *127*, 12456-12457.

(50)   Minter, A. R.; Brennan, B. B.; Mapp, A. K. *Journal of the American Chemical Society* **2004**, *126*, 10504-10505.

(51)   Minter, A. R.; Fuller, A. A.; Mapp, A. K. *Journal of the American Chemical Society* **2003**, *125*, 6846-6847.

(52)   Liu, B.; Alluri, P. G.; Yu, P.; Kodadek, T. *J Am Chem Soc* **2005**, *127*, 8254-5.

(53)   Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew Chem Int Ed Engl* **2007**, *46*, 2865-8.

Chapter II

Features Affecting Activator Potency[1]

A. Background

Transcriptional activators play a critical role in the regulation of gene expression.

The growing link between misregulation of transcription and disease has led to an effort

to create activator artificial transcriptional factors (ATFs) capable of repairing aberrant

expression patterns.[1-3] Minimally composed of a DBD and an AD, transcriptional

activators are responsible for regulating the time course and the levels of transcription of

specific genes.[4] The DBD imparts the majority of the specificity through localizing the

transcriptional activator to a specific site in the genome and the AD controls the potency

of the activator through contacts with chromatin modifying enzymes and components of

the RNA polymerase holoenzyme.[4] However, the factors that contribute to activator

potency – the extent to which an activator up-regulates transcription - remain poorly

understood. At the onset of this work, the prevailing model was that the affinity of an

activator for the transcriptional machinery directly correlated to the extent to which that

---

activator up-regulated transcription. The implication for the development of small molecule transcriptional activators was that they needed to interact with coactivator binding surfaces with high affinity (nanomolar or subnanomolar dissociation constants); this represented a potentially significant impediment to small molecule activation domains since the identification small molecules with this level of affinity for protein surfaces has been a longstanding challenge.

One aspect of the previous studies of activator-coactivator interactions that was problematic was that they were often carried out with transcriptional machinery proteins shown not to be physiologically relevant activator targets. TBP, for example, was often examined in these studies. In this Chapter, we describe the use of artificial transcriptional activators that specifically interact with the coactivator Med15 to demonstrate that affinity and potency are not always correlated. Further, these studies revealed that micromolar dissociation constants are sufficient for good activity. In addition, we found that secondary binding interactions, so-called 'masking' interactions, can also contribute significantly to the ability of an activator to up-regulate transcription

B. Beyond Binding Affinity

Although the promiscuous nature of AD binding has made it challenging to unambiguously identify which protein targets are most important for the initiation of transcription, the Mediator complex has emerged as a common target of numerous endogenous activators.[5-7] The Mediator is a multi-protein complex (20-25 proteins in yeast) believed to act as a bridge between DNA bound transcriptional activators and RNA polymerase II (Figure II-1). This complex is composed of a head and middle

region largely responsible for interacting with RNA polymerase II and a tail region which is proposed to be targeted by endogenous activators.[8]
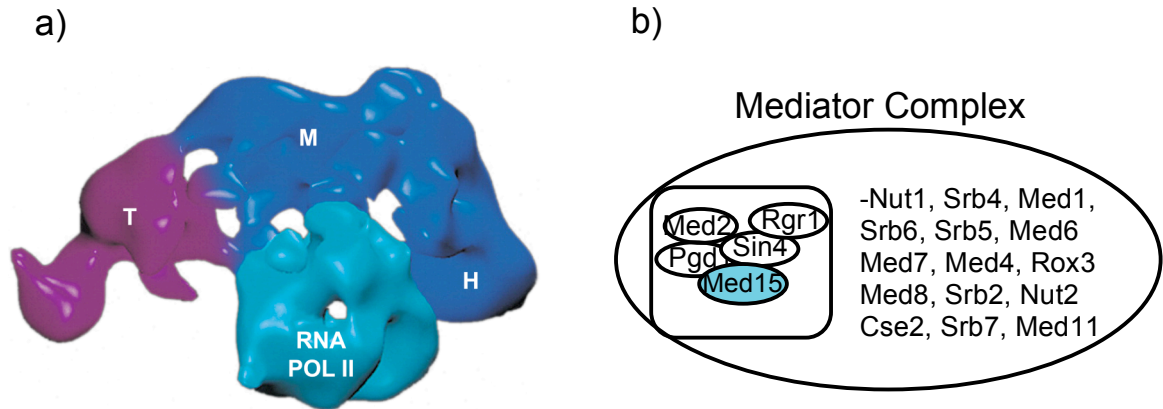
Figure II-1: a) Mediator, shown in this electron micrograph, is composed of a head (H) and middle (M) regions which contact RNA Polymerase II directly as well as a tail (T) region which is targeted by transcriptional activators. This figure was taken from (8) b) The yeast Mediator is composed of 20 proteins. Med15 (highlighted in blue) is located in the tail region and has been shown to interact with numerous different transcriptional activators.

Within the tail region of the yeast Mediator complex lies Med15, a protein that has been shown to interact with a number of artificial and natural ADs.[8,9] The most compelling evidence of interactions between Med15 and endogenous activators was recently presented by Hahn in a series of photo-cross-linking experiments. The incorporation of a cross-linking agent into the AD of Gal4 identified a Med15 interaction in the context of transcriptional initiation.[10] Further evidence of Med15 as an activator target came from a series of activator bypass experiments in which Med15 was shown to activate transcription when directly fused to a DBD.[11,12] The substantial amount of evidence for Med15 acting as a cellular target of activators prompted us to screen a peptide library against this protein in search of novel ADs. We inferred that these peptidic ligands would function as ADs through their interactions with Med15. A comparison of the affinity of these activation domains for Med15 versus their ability to up-regulate transcription would

26

then test the activity-affinity connection. The large diversity of a peptide library was also

likely to produce peptide sequences not commonly observed in known activating regions

with unique binding sites on the protein. In this case, we could also determine the

importance of binding site location on activator potency.[13]

We designed two libraries, each with four variable positions. Our first library

(AXXXXLSE) was designed to bias for peptides which would bind similar sites as

endogenous activators, as the LSE sequence is common among helix forming

amphipathic ADs. Our second library (AXXXXPSE) contained a proline which would

prevent helix formation and would probably produce ligands which would bind surfaces

of Med15 which are not targeted by endogenous activators. An ELISA screen was then

carried out by another graduate student (Aaron Minter) with GST-Med15 (186-619)

against the two peptide libraries. This region of Med15 was chosen for the screen since it

is the largest portion of Med15 that can easily be expressed. In addition, this region has

been shown to interact with endogenous and artificial ADs. The screen produced a total

of 37 peptide ligands (Figure II-2). Ligands were found to be either amphipathic in

sequence, hydrophobic, or contained an excess of positive charge.

Library 1

Library 2

Library 1: AXXXXLSE				Library 2: AXXXXPSE	
1. NITY	9. RARV	17. QRRV		22. SFHR	30. VLGW
2. ETVS	10. SHRT	18. SLNR		23. FWLF	31. IWLF
3. RRGV	11. GRRA	19. EDRR		24. NQQW	32. YFEV
4. TWRR	12. RVYR	20. RWTI		25. DLTM	33. VHPV
5. HTRH	13. AHRR	21. NLRT		26. FYRN	34. FNWE
6. RDRT	14. SRHR			27. MHFP	35. FTLW
7. FARR	15. RQRN			28. HYYY	36. LWFF
8. QRRT	16. YHRT			29. PVLG	37. IFFF
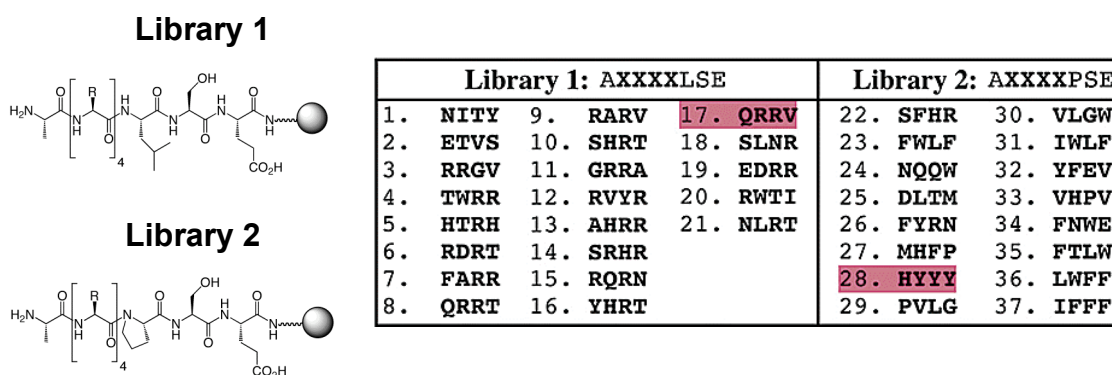
Figure II-2: Libraries used in the Med15 screen. The peptide screen produced 37 ligands which interact with Med15. The peptides capable of upregulating transcription when localized to DNA are highlighted.

Two members of my research lab (Garrette Belanger and Jenifer Lum) coupled each ligand to a protein DBD and tested its ability to activate transcription with a β-galactosidase assay. Two of these ligands were able to upregulate transcription in this context.
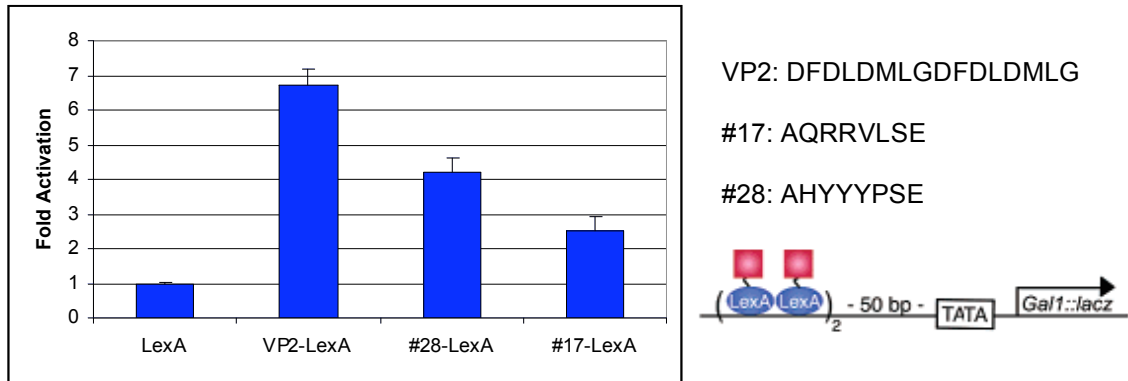
Figure II-3: Med15 Ligands #17 and #28 activate transcription when fused to the DBD of the bacterial protein LexA in a yeast strain bearing LexA binding sites upstream of a *lacZ* reporter gene.

Interestingly, despite their ability to activate transcription, neither activator resembles known activating regions. Peptide #17 bears a significant amount of positive charge and peptide #28 is extremely hydrophobic, two features not commonly observed with endogenous ADs. It was therefore likely that these peptides were targeting unique sites on Med15 which do not typically interact with endogenous ADs.

Along with Dr. Zhiqian Wu, a postdoctoral fellow in my research lab, I carried out fluorescence polarization binding experiments with the active peptidic ligands along with the controls VP2 and P201. We found that all of the ligands bound to Med15 with low micromolar dissociation constants; the most potent (XLY) and the least potent (#17) had nearly identical dissociation constants.

In order to determine if the peptides were targeting different surfaces on Med15, a set of competition binding experiments were performed by Dr. Zhiqian Wu. In these experiments, each peptide was fluorescently labeled and incubated with Med15 at a concentration that allowed saturated binding. Following this pre-incubation, an excess of unlabeled peptide was titrated in. As expected, unlabeled peptide of the same identity competed for binding and resulted in a decrease in fluorescence polarization. However, when any of the other peptides was titrated, no competition was observed, demonstrating the unique nature of the binding sites.

a)

AD	K$_D$ for Med15	Fold Activity
P201	2.2 ± 0.2 µM	>>100
VP2	7.3 ± 2.4 µM	6.9
#28	4.3 ± 1.4 µM	4.4
#17	2.4 ± 0.4 µM	2.3

b)
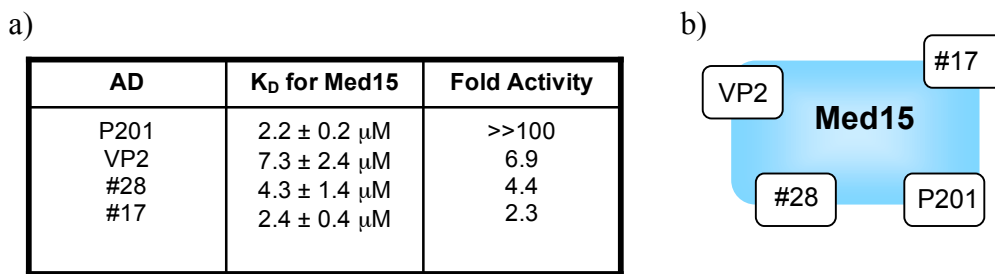
Figure II-4: (a) Dissociation constants for each of the ADs and their ability to upregulate transcription. (b) Each activation domain binds to unique sites on Med15.

Since each of the four activators bind to different sites on Med15, it is unlikely that binding affinity alone leads to the varied function observed. One striking result from this study was the potency of P201 when compared to the other peptidic ADs. Determining the factors that lead to this extraordinary activity could uncover design principles that could be incorporated into activator ATFs.

C. An Unusually Potent Peptidic Activation Domain

P201 was discovered by the Ptashne lab through a screen of random peptides that when fused to the first 100 residues of Gal4, comprising the DBD and dimerization domain (dd), were capable of upregulating transcription in a gene bearing the Gal4 upstream activating sequence (UAS). Remarkably, this small hydrophobic peptide with the sequence YLLPTCIP, lacking any acidic residues often observed in ADs, is capable of activating transcription as well as the full Gal4 activator protein (1-881) when fused to Gal4 (1-100).[9,14] Further studies also revealed that P201 was entirely dependent on the mediator component Med15. In fact, a single point mutation in Med15 (T322K) completely abrogates the ability of Gal4 (1-100)-P201 to activate transcription *in vivo* and the ability of P201 to bind to Med15 *in vitro*.[9,15] This is quite an unusual result since activators typically target numerous proteins in the transcriptional machinery and this can somewhat compensate for the loss of any single interaction.

Despite its unusual properties, P201 remains one of the two potent artificial activators with robust activity *in vivo*. The other is Gal4 (1-100) by itself in a yeast strain bearing a single point mutation in Med15 (N342V).[16,17] This mutant form of Med15 has been named Med15P for its ability to potentiate activation by Gal4 (1-100), a protein which does not act as a transcriptional activator in wild-type yeast. Through extensive NMR studies, the Ptashne lab determined the interaction between Med15P and the Gal4 dimerization domain (Gal4$_{dd}$) involved the largely hydrophobic amino acid residues of loop 1 and helix 3 binding to the hydrophobic surface created by the N342V mutation (Figure II-5).[17]
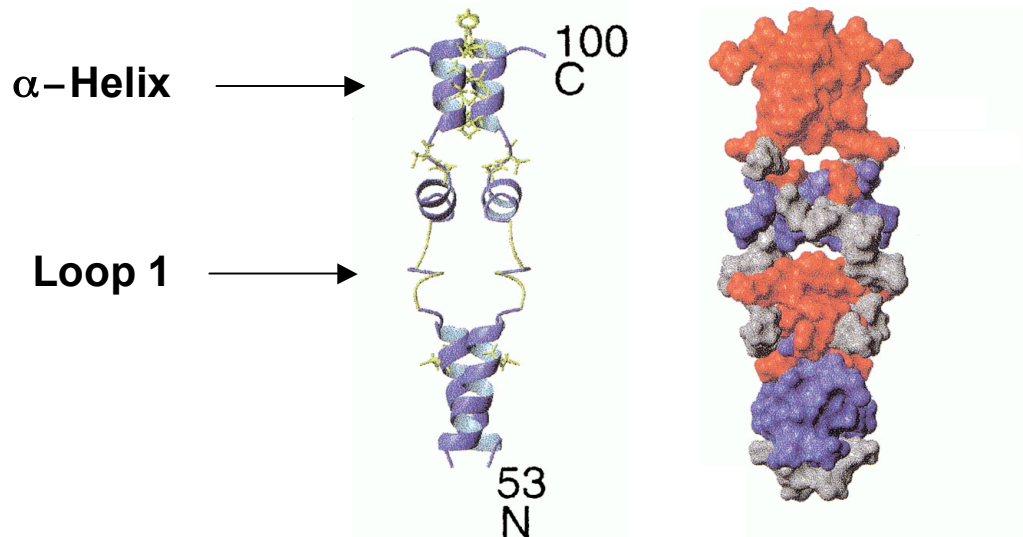
Figure II-5: (left) NMR structure of the Gal$_{dd}$ with helix 3 and Loop 1 indicated. (right) Residues found to be important for interacting with Med15P are colored in red. Figure from (14).

We initially hypothesized that the potent activity of Gal4 (1-100)-P201 might be the result of a combination of P201's affinity for Med15 as well as residual affinity of Gal4$_{dd}$ to wild type Med15. In order to address this possibility, a series of β-galactosidase assays were carried out in the lab of one of our collaborators, Dr. Aseem Ansari at the University of Wisconson, Madison. In these experiments, the Ansari lab examined the ability of Gal4 (1-100)-P201 and Gal4 (1-100) to activate transcription in yeast strains bearing wild type or mutant forms of Med15. The first mutant strain contained Med15 with the T322K mutation, the mutation which prevents P201 from interacting with Med15. The second mutant yeast strain contained the N342V mutation, the Med15P mutant which promotes an interaction with Gal4$_{dd}$. Thirdly, a yeast strain bearing Med15 double mutant (Med15DM) containing both T322K and N342V was tested.

As expected, our collaborators observed that Gal4 (1-100)-P201 activated transcription robustly in the wild-type Med15 strain and did not activate in the strain

bearing the T322K mutation (Figure II-6 A, B). Furthermore, Gal4 (1-100) was inactive in both of these strains, as they do not contain the mutation which promotes an interaction (Figure II-6 C, D). In addition, Gal4(1-100) was active in both the Med15P strain as well as the Med15DM strain, demonstrating that the second mutation does not prevent an interaction between Gal4$_{dd}$ and Med15P (Figure II-6 E, F). Although the Med15P strain contains binding sites for both Gal4 (1-100) as well as P201, activation elicited by Gal4 (1-100)-P201 was comparable to levels observed with the wild-type Med15 strain (Figure II-6 G). In the final experiment, the ability of Gal4 (1-100)-P201 to activate transcription in a yeast strain with Med15DM was assessed. Since Gal4 (1-100) is capable of activating transcription in this strain, it would be expected that despite the disruption of the P201-Med15 interaction, Gal4$_{dd}$ could interact and upregulate transcription. However, activation was severely diminished in this strain (Figure II-6 H). This suggests that P201 is able to block the interaction made by the Gal4$_{dd}$. Subsequent experiments were then designed to elucidate the nature of this interaction.
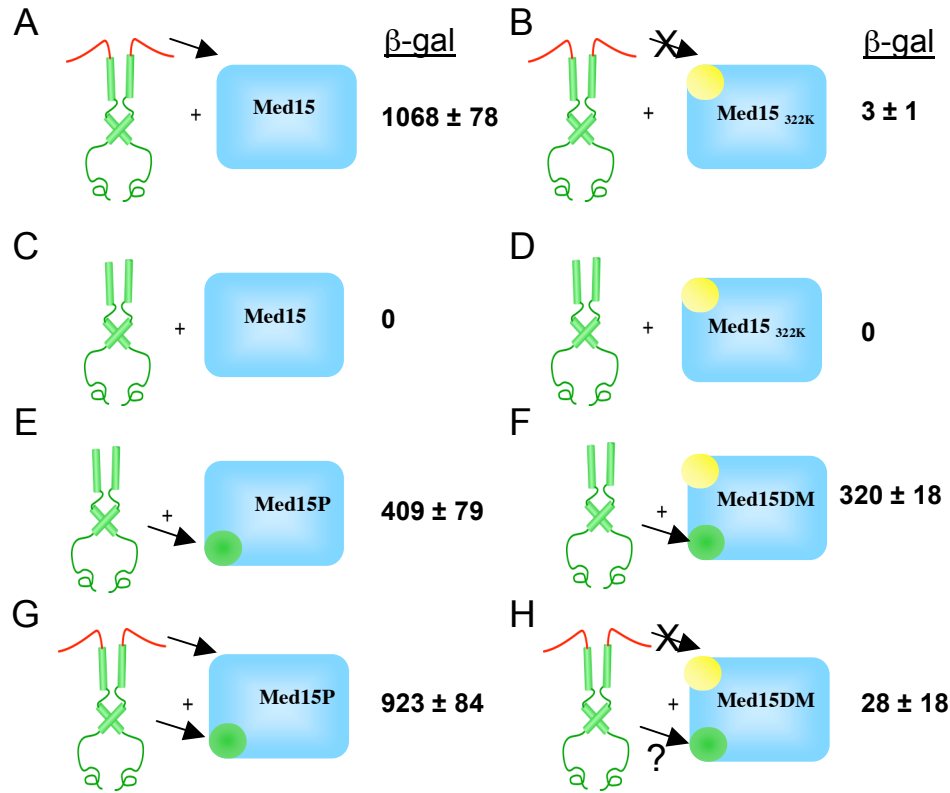
Figure II-6: A series of β-galactosidase assays carried out by our collaborators. Interactions are represented by arrows, and disrupted interactions are represented by crossed out arrows.

While the *in vivo* experiments described indicate the Gal4$_{dd}$ is incapable of making an interaction with wild type Med15, earlier mutagenesis experiments carried out by the Ptashne lab have shown that at least part of the Gal4$_{dd}$ is required for P201 function. In order to examine the role of helix 3 of the Gal4$_{dd}$ in P201-mediated activation, the Ansari lab performed a series of alanine scanning mutagenesis experiments. In order to prevent any potential impact of the scan on DNA binding, the DBD of Gal4 was replaced by the bacterial DBD LexA. This experiment indicated Gal4 (93-100) was essential for P201 function and therefore all *in vitro* binding experiments were performed with the minimal P201 peptide Gal4(93-100)-P201, LTGLFVQDYLLPTCIP), referred to as XL$_Y$ (Figure II-7).
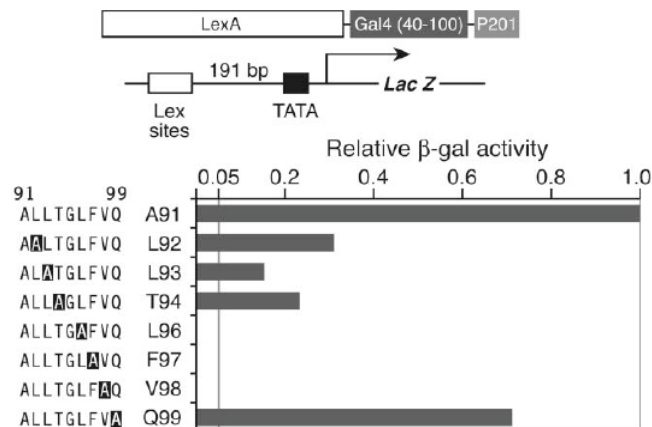
Figure II-7: Residues 93-100 of Gal4 are necessary for P201 function.

D. Binding Experiments

The activation data suggests an interaction between XL_Y and the $Gal4_{dd}$. In order to more fully investigate this interaction, Steven Rowe and I carried out a series of fluorescence polarization experiments. For this purpose, XL_Y was synthesized by standard methods and fluorescently labeled.[18] In addition, $Gal4_{dd}$ and the central portion of Med15 (residues 186-619) were expressed as GST fusion proteins and purified.

As demonstrated in figure II-8, XL_Y interacts with both Med15 and $Gal4_{dd}$ with micromolar binding affinity, with K_DS of 2.2 μM and 5 μM respectively (Figure II-8). Additionally, XL_Y binds to $Gal4_{dd}$ in the context of $Gal4_{dd}$-P201, albeit with a 3-4 fold reduction in binding affinity, indicating that the fused P201 does not prevent the XL_Y●$Gal4_{dd}$ interaction (Figure II-8). As a control, we also looked at the interaction of a mutant version of P201 which does not activate transcription *in vivo*. This peptide, referred to as XL_R, contains a tyrosine to arginine mutation, resulting in the sequence

34

LTGLFVQD<u>R</u>LLPTCIP. Binding assays with this version of P201 did not show appreciable binding to Med15, Gal4$_{dd}$, or Gal4$_{dd}$-P201.

Steve and I next tested to see if XL$_Y$ would interact with Med15 bearing either the N342V mutation (Med15P) singly or in combination with the T322K mutation (Med15DM). Consistent with the *in vivo* data, this mutation caused a 3-fold reduction in affinity for Med15P (Figure II-8). Also consistent with the *in vivo* results, XL$_Y$ did not bind to Med15DM within the limits of our assay.

XL$_Y$ + Med15

XL$_Y$ + Gal4$_{dd}$

K$_D$ 2.2 ± 0.4 μM

K$_D$ 5.0 ± 1.4 μM

[GST-Med15(186-619)] (μM)

[GST-Gal4(52-100)] (μM)

XL$_Y$ + Gal4$_{dd}$- P201

XL$_Y$ + Med15P

K$_D$ 16.9 ± 2.1 μM

K$_D$ 6.5 ± 2.1 μM

[GST-Gal4(52-100)-P201] (μM)

[GST-Med15P(186-619)] (μM)
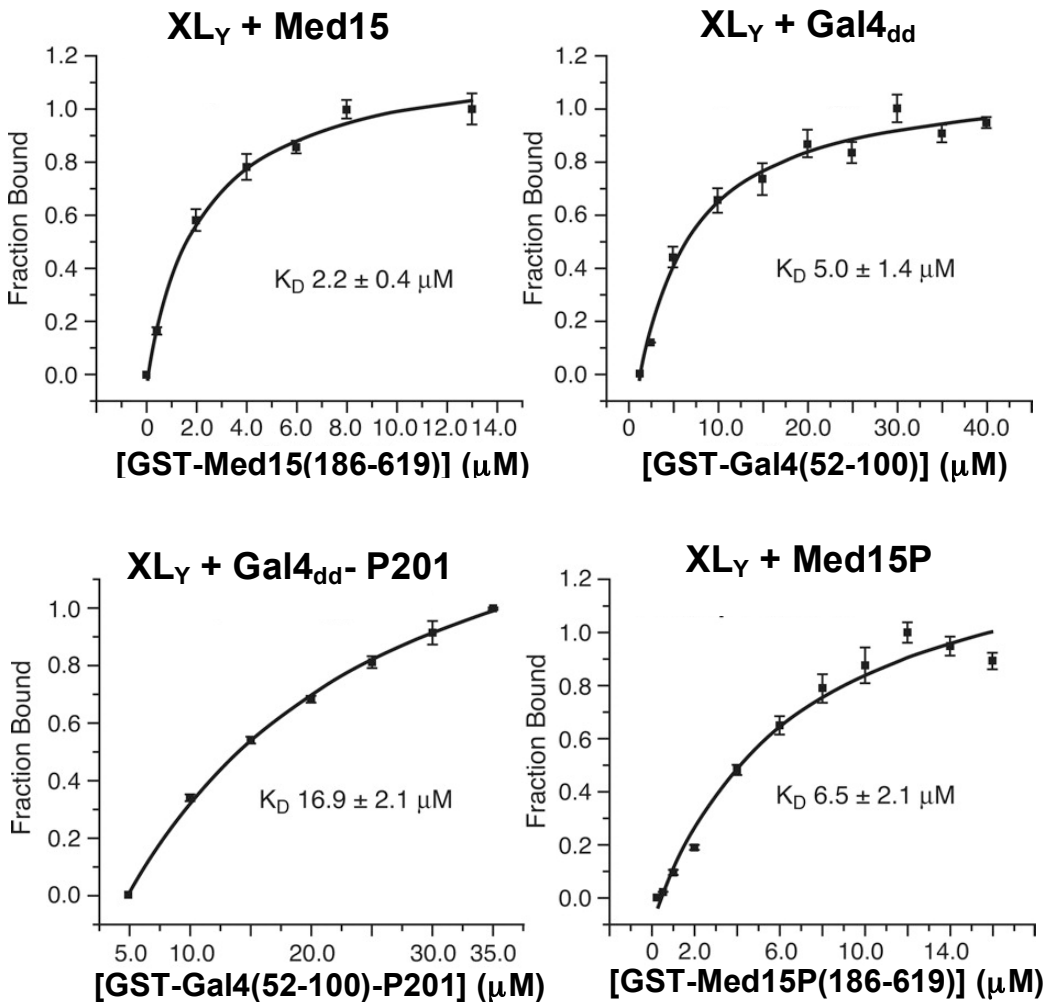
Figure II-8: Binding affinities of XL$_Y$ for Med15, Gal4$_{dd}$, Gal4$_{dd}$-P201, and Med15P.

NMR studies of the $Gal4_{dd}$ indicate the presence of several solvent exposed hydrophobic residues.[17] XL_Y is also a very hydrophobic peptide leading to the possibility that the interaction observed is simply a nonspecific hydrophobic interaction. In order to probe the specificity of the interaction between XL_Y and $Gal4_{dd}$, we decided to examine interactions between $Gal4_{dd}$ and the hydrophobic peptide discovered in the peptide screen previously discussed (AHYYYPSE). This peptide binds to Med15 with comparable affinity to XL_Y, yet does not activate transcription as robustly. If the origin of P201's activity lies in its ability to interact with $Gal4_{dd}$, then we would not expect the weaker activation domain AHYYYPSE to also interact. This was in fact the case, as illustrated in Figure II-9. Consistent with this result, it also does not inhibit activation by $Gal4_{dd}$ in yeast strains bearing Med15DM. In further experiments, we observed that XL_Y does not bind to bovine gamma globulin, a protein often observed to interact with hydrophobic surfaces (figure II-9). This data, taken together, demonstrates that the $Gal4_{dd} \bullet XL_Y$ binding event is not simply the result of nonspecific hydrophobic interactions.
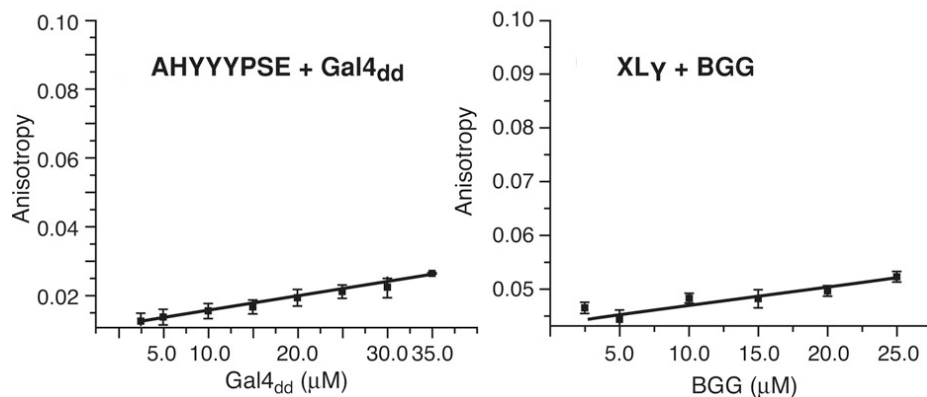
Figure II-9: The hydrophobic Med15 ligand does not interact with the Gal4DD. In addition, XLY does not bind nonspecifically to the hydrophobic protein bovine γ-globulin.

Another binding event of interest in this system is the interaction between $Gal4_{dd}$ and the various mutants of Med15. Due to the large size of these proteins, fluorescence polarization experiments could not be performed to determine the binding affinity. The affinities were thus determined by our collaborators at the University of Wisconsin, Madison, through the use of electrophoretic mobility shift assays (Figure II-10).   In this series of experiments, DNA bound Gal4 (1-100) (containing the Gal4 DBD as well as the $Gal4_{dd}$) was incubated with Med15 as well as varying concentrations of Med15P and Med15DM. Consistent with *in vivo* results, no interaction was observed between Gal4 (1-100) and Med15. However, introduction of the N342V mutation alone (Med15P) or in combination with the T322K (Med15DM) results in an interaction. Although the limited solubility of the Med15 proteins prevented the determination of the complete binding isotherm, 18 µM Med15P or Med15DM was sufficient for about half of the Gal4 (1-100)●DNA complex to form a ternary complex. If the half maximal binding can be assumed to be representative of the K_D for this interaction, it demonstrates the XL_Y has a higher affinity for $Gal4_{dd}$ than does Med15. This data is also consistent with *in vivo* observations that XL_Y can disrupt activation by $Gal4_{dd}$.
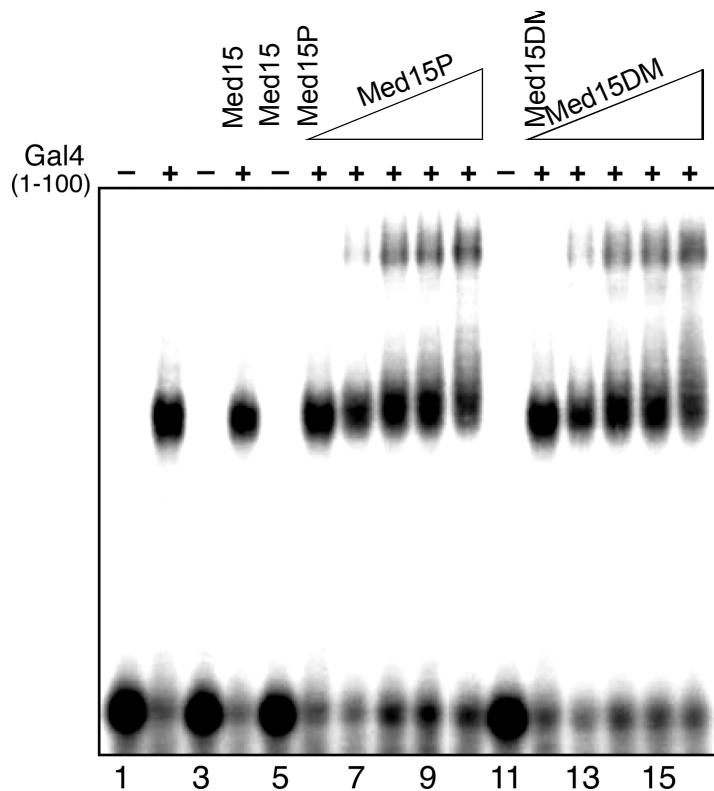
Figure II-10: Electrophoretic mobility shift assays demonstrating interactions between DNA bound Gal4 (1-100) and mutants of Med15. 30 nM Gal4 was used in all experiments, a concentration sufficient for saturation of the 30bp DNA oligonucleotide used. Concentrations in lanes 3,4,5, and 11 were 20 μM. Concentrations of Med15P and Med15DM in lanes 6-10 and 12-16 were 1, 3, 9, 18, and 20 μM respectively.

E. Identification of the XL$_Y$ Binding Site

Both the *in vivo* activation data as well as the *in vitro* binding data suggested an interaction between Gal4$_{dd}$ and XL$_Y$. Two questions that remained were whether this interaction contributed to XL$_Y$ function and where in the Gal4$_{dd}$ this interaction occurred. These questions were mainly addressed through a series of deletion and mutagenesis experiments carried out by our collaborators. In the first set of experiments P201 was attached to the LexA DBD along with fragments of the Gal4$_{dd}$. As observed in Figure II-11, no activation occurred until the entire dimerization domain of Gal4 (residues 50-100)

were present. The last experiment in the figure demonstrates the importance of helix 1

and loop 1, since constructs containing only helices 2 and 3 showed no activation.
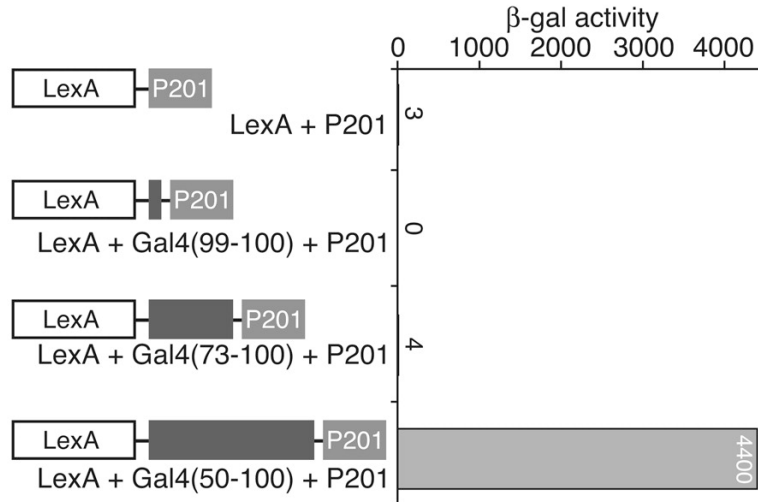
Figure II-11: A LexA-P201 chimera only activates transcription when the entire Gal4$_{dd}$
is also present.

In order to further refine the binding site, alanine scanning was performed. Again

the LexA DBD was used to prevent perturbation in Gal4 binding with the alanine scan in

the dimerization domain. This experiment demonstrates the importance of a number of

residues in loop 1 of the dimerization domain, namely residues 68, 69, 72, and 75 (Figure

II-12). Mutations in these residues greatly diminished the P201 mediated activated

transcription and is therefore the likely binding site of this peptide. To further probe the

importance of these residues, Steven Rowe expressed a mutant of Gal4$_{dd}$ containing the

L69A mutation. Indeed, this mutant did not interact with XL$_Y$, consistent with the *in vivo*

results.
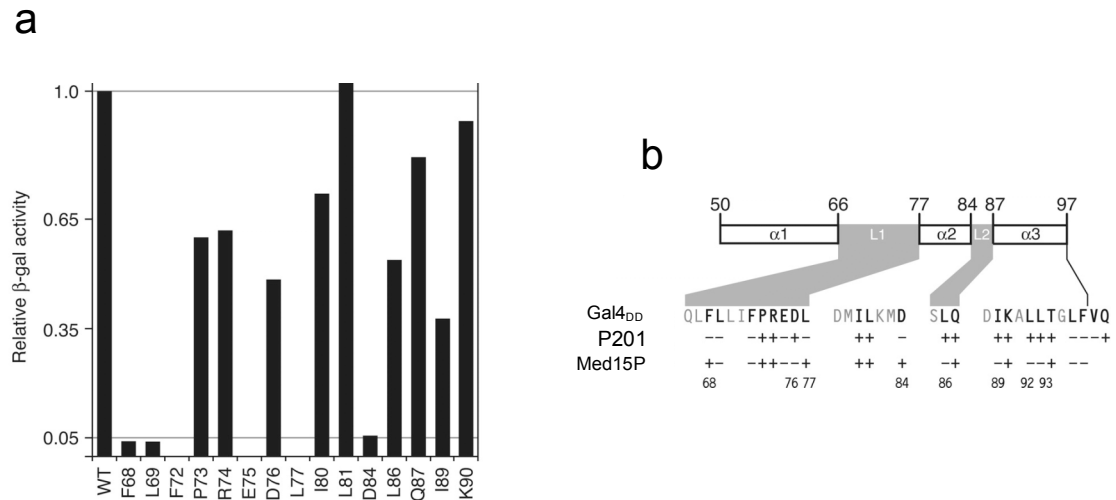
Figure II-12: a) Alanine scanning was used to uncover important residues for activation mediated by P201. b) Med15P and XL$_Y$ target unique yet overlapping sites on Gal4$_{dd}$.

We next compared the residues that were important for the Gal4$_{dd}$●XLY interaction to those previously found to be critical for the Gal4$_{dd}$●Med15P interaction. Some mutations, such as L86A and I89A, which completely disrupt the Gal4$_{dd}$●Med15P interaction *in vitro* and activation *in vivo,* did not severely affect P201 mediated activation. Others however, such as F68A, L77A, and D84A, completely abrogate P201 mediated transcription yet do not affect the Gal4$_{dd}$●Med15P interaction. Although there are a number of residues which, when mutated, disrupt both interactions, the differences indicate that P201 and Med15P bind to distinct, yet overlapping surfaces on Gal4$_{dd}$.

F. Dynamics of Binding

In order to more fully characterize the XL$_Y$●Gal4$_{dd}$ and the XL$_Y$●Med15 interactions, Steven Rowe and I measured the off-rates of XL$_Y$ with fluorescence polarization for both complexes. In these experiments, XL$_Y$ was fluorescently labeled and incubated with Gal4$_{dd}$ or Med15 at saturated binding conditions (18 μM Med15 and

50 μM Gal4$_{dd}$ with 25 nM XL$_Y$ and 10 nM XL$_Y$ respectively). Following incubation, a

large excess of unlabeled XL$_Y$ was titrated in. This resulted in a decrease in polarization

as the fluorescently labeled bound peptide was competed off with the unlabeled peptide.

As indicated in figure II-13, the dissociation of XL$_Y$ was significantly faster with Gal4$_{dd}$

(0.24 s^{-1}) than with Med15 (0.063 s^{-1}). In fact, the dissociation with Gal4$_{dd}$ was so fast,

that nearly all of the peptide had been competed off before our first measurement was

taken, thus the calculated k$_{off}$ for XL$_Y$●Gal4$_{dd}$ actually represents the lower limit of this

rate. In contrast, previous NMR measurements have shown the k$_{off}$ for Gal4$_{dd}$●Med15P

to be >5 s^{-1}. Although these measurements are not completely comparable since they

were taken with different experimental techniques, the off-rate of Med15P from Gal4$_{dd}$

appears to be much faster that of XL$_Y$ from Gal4$_{dd}$. This would help explain how XL$_Y$ is

capable of interfering with Gal4$_{dd}$ activation in yeast strains bearing Med15DM.
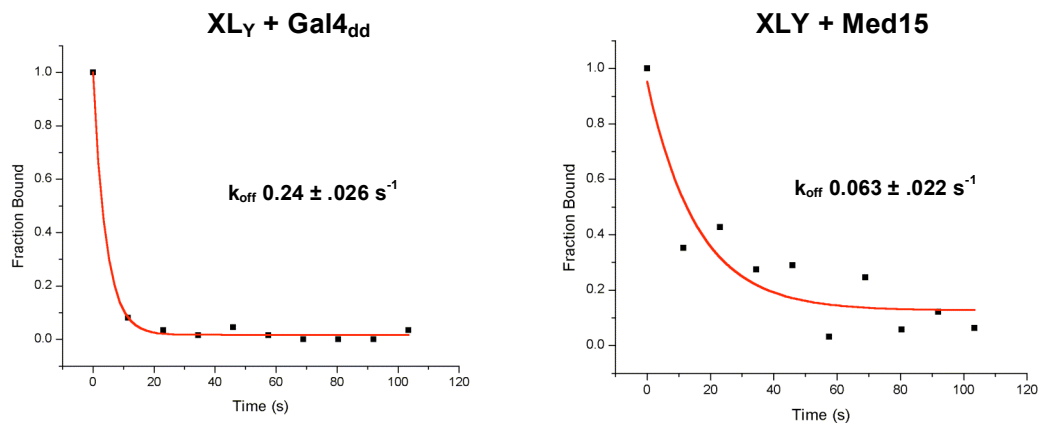
Figure II-13: Fluorescence polarization was used to determine the dissociation rates of XLY for Gal4$_{dd}$ and Med15.

G. Model of P201 Function

Both the *in vivo* and *in vitro* data suggest that XL_Y is capable of making transient interactions with both $Gal4_{dd}$ as well as Med15. This feature of transient concealment is commonly observed in endogenous activators, where hydrophobic residues critical for activation are often bound in intra- or intermolecular complexes preventing activation. When a gene product is needed, intracellular signals lead to the dissociation of this complex allowing transcriptional activation.[19-21] One such example can be observed with the Gal4 activator. In the absence of galactose, Gal4 is bound in a complex with the protein Gal80, preventing expression of the galactose metabolism genes. However, in the presence of galactose, Gal3 binds to Gal80, causing a conformational allowing Gal4 to upregulate transcription. Intramolecular examples of this type of 'masking' interaction also exist. This is seen with the yeast proteins Leu3 and Put3 which control amino acid metabolism.[21,22] In both cases, the full length protein is found to be inactive *in vitro* and *in vivo* due to masking. It is only the presence of a specific metabolite that the AD is revealed and stimulation of gene expression is observed. In both of these examples, another surface has co-evolved to bind and conceal the activating region.

Our data suggests a similar intramolecular binding interaction occurs with Gal4(1-100)-P201, albeit without a cellular signal, leading to its remarkable potency. Gal4(1-100)-P201 likely exists in two conformations (Figure II-14). In the "closed" conformation, XL_Y binds to the hydrophobic surface of loop one on $Gal4_{dd}$, preventing non-productive interactions and likely preventing premature degradation. In the "open" conformation, P201 binds to Med15 recruiting the RNA polymerase holoenzyme and

upregulating transcription. Through this coupled equilibria, Gal4(1-100)-P201 robustly activates transcription.
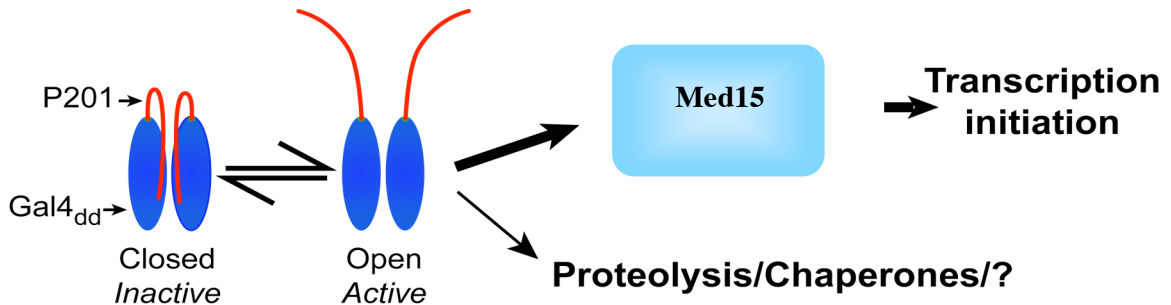
Figure II-14: Transient exposure of P201 leads to its remarkable activity *in vivo*.

H. Masking Interactions and Potency

Many activators interact with masking proteins to prevent activation except in the presence of a cellular signal. This allows the cell to maintain strict control over levels of gene expression. On the other hand, artificial activators typically operate outside of this regulatory network and lack inter- or intramolecular masking within the cell. While it was previously believed that these interactions were largely a mechanism of regulating transcription, Gal4(1-100)-P201demonstrates the important contributions these interactions can have on functional potency.

This property has also shown application in the construction of activator ATFs. Many of the inactive ligands from the Med15 screen contained hydrophobic amino acids in the randomized positions. It was therefore likely that the incorporation of an intramolecular binding surface would prevent nonproductive interactions and stabilize the activator in the cell. Two graduate students in my lab (Jenifer Lum and Chinmay Majmudar) tested these ligands in the context of Gal4(1-100). In an analogous manner to

Gal4(1-100)-P201, these inactive ligands were converted into potent transcriptional activators. Further studies also demonstrated binding between these ligands and Gal4$_{dd}$, indicating a similar interaction as observed with P201.[23]

I. Experimental Details

General peptide synthesis methods

Solid phase peptide synthesis was accomplished with FMOC-protected amino acids, HOBt, and HBTU purchased from Peptides International using standard methods either manually or using an Advanced ChemTech90 synthesizer using established protocols. Peptide-synthesis grade DMF, Triethylamine, and piperidine were purchased from Fisher and used without further purification. Coupling efficiencies were monitored using a ninhydrin test kit purchased from Fluka. All peptides were purified to homogeneity using reversed-phase HPLC on a C18 column with a gradient solvent system (buffer A: 0.1% TFA, buffer B: CH$_3$CN).

β-Galactosidase Assays

β-Galactosidase assays were performed in LS41 [JPY9::ZZ41, *Matα his3Δ200 leu2Δ1 trp1Δ63 ura3-52 lys2Δ385 gal4Δ11/ ZZ41*] which was generously provided by Dr. Aseem Ansari (University of Wisconsin). LS41 yeast was initially transformed with the expression vectors by Garrette Belanger using the Lithium Acetate method and plated on agar plates with the appropriate selection. Single colonies were then chosen to inoculate 5 mL starter cultures which were grown for 12-18 hours at 30ºC in SC media containing 2% Raffinose. The density of these cultures were then determined using

OD_{600} and they were then diluted to a concentration of 0.7×10^7 cells/mL in selection media supplemented with 2% Raffinose and 2% Galactose. These cultures were then grown for 12-18 hours at 30°C until the OD_{600} of the 10X dilution was between 0.4-0.7. Cells were harvested by centrifugation at 5000 RPM for 10 minutes at 4°C. Next, the cells were washed with 500 μL of sterile H20 and again harvested at 5000 RPM for 10 minutes at 4°C. Following the wash, the cells were resuspended in 150-250 μL of Breaking Buffer (100 mM Tris-Cl pH 8.0, 20% glycerol, Roche protease inhibitor cocktail) and placed in an eppendorf tube. In order to lyse the cells, glass beads were added and the cells were vortexed at maximum speed at 4°C for one minute followed by a one minute break. This process was repeated ten times to assure cell lysis. In order remove the lysed cells, a needle was heated in a flame and used to pierce the bottom of the eppendorf tube. This was then placed in a second tube and centrifuged for one minute at 4°C. The cellular debris was then pelleted by centrifuging at 14,000 RPM for ten minutes at 4°C. The clarified lysate was then transferred to a fresh tube and briefly vortexed to mix contents.

In order to determine the activity of the β-Galactosidase, 10-100 μL of lysate was diluted to 1 mL in Z-Buffer (40 mM Na_2HPO_4, 60 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM 2-Mercaptoethanol). These samples were then incubated at 30°C for 30 minutes. Following incubation, 200 μL of ONPG (4 mg/mL in Z buffrer) was added to each sample and timed. Once the samples began to turn yellow, the reactions were quenched with the addition of 500 μL of 1M Na_2CO_3 and the time was recorded. The OD_{420} was then determined for each sample. In order to normalize each assay, the total protein concentration was measured with a Bradford Assay. The activity of each sample

was then calculated with the equation: $(OD_{420} * V_T) / (P * V_S * t * 0.0045)$ where V_T is

the total volume in mL, P is the total protein concentration in mg/mL, V_S is the volume of

lysate added, and t is the time before quenching. The fold activities of each were

determined by dividing the activity of a sample with the activity observed with the DBD

alone. All experiments were performed in triplicate with the standard deviation of the

mean indicated.

Synthesis of Fluorescein Labeled Gal11 Ligands

The Gal11 ligands were synthesized on Rink amide resin and a cysteine residue

was incorporated at the carboxy terminus for subsequent functionalization. The peptides

were then labeled on resin with fluorescein-5-maleimide (Molecular Probes). At the

conclusion of the synthesis, the peptides were fully deprotected and cleaved from the

resin with treatment with 95% TFA/2.5% TIS/2.5% H_2O, precipitated with cold ether,

and purified to homogeneity using reversed-phase HPLC. Each peptide was characterized

using electrospray mass spectrometry.

Synthesis of Fluorescein Labeled XL_Y and XL_R

XL_Y and XL_R were also synthesized on Rink amide resin. Following the final Fmoc

deprotection, the N-terminal amine was labeled with the succinimidyl ester of

carboxyfluorescein (Molecular Probes). The peptides were then cleaved with 95%

TFA/2.5% TIS/2.5% H_2O, precipitated with cold ether and purified using reversed-phase

HPLC.

Expression and Purification of GST-Med15 (186-619)

The plasmid pGEXVsh1 was transformed into chemically competent BL21(DE3) pLysE

E. coli (Invitrogen) and cells were plated onto LB-agar plates supplemented with

ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL). Cultures (50 mL) from

single colonies were grown overnight at 37° C (275 rpm) in LB supplemented with

ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL) before addition to 1L of LB

supplemented with ampicillin (0.1 mg/mL). After 3 h, the cultures were cooled to 16 °C,

and expression was induced with IPTG (final concentration 0.5 mM) for 5 h. The cell

pellet was lysed using sonication, and the GST tagged protein was isolated from the cell

lysate using glutathione-sepharose beads (Amersham-Pharmacia Biotech). Elution from

the beads was accomplished with 50 mM Tris•HCl buffer (pH 8.0) containing 15 mM

glutathione and 0.1% NP-40. The protein solution thus obtained was concentrated to ~20-

25 μM and the buffer exchanged to storage buffer (PBS, pH 7.4, 1 mM DTT, 10%

glycerol (v/v), and 0.01% NP-40) using a Millipore Ultrafree centrifugal filter device.

The protein solution was stored in 50 μL aliquots at −80 °C until needed. The protein

concentration was measured using a Bradford assay (Bio-Rad) with BSA as the standard.

The identity and purity of the fusion protein was verified by reducing SDS-PAGE with

appropriate molecular weight standards.

Dissociation Constant Measurements

The dissociation constants of XL_Y and XL_R for GST-Med15(186-619), GST-

Med15P(186-619), GST-Med15DM(186-619), GST-Gal4(52-100), and GST-Gal4(52-

100) +P201 were determined by fluorescence polarization on a Spex Fluoromax-2 fluorimeter. All measurements on the Spex Fluoromax-2 were done at room temperature.

For each experiment, an aliquot of fluorescein-labelled peptide was dissolved in 1mL of storage buffer (PBS, pH 7.4, 1 mM DTT, 10% glycerol (v/v), and 0.01% NP-40) and the concentration determined on a Varian Cary 300 Bio UV-Vis spectrophotometer. 1.25 μM and 25 nM (for Med15 derivatives) or 10nM (for Gal4 derivatives) stocks of the labeled ligands in storage buffer were prepared. For each experiment on the Spex Fluoromax-2, the 1.25 μM stock solutions were then each diluted into three separate solutions in a Starna quartz sub-micro 40 mL 10 mm cell to yield a final concentration of 25nM (for Med15 derivatives) and 10nM (for the Gal4 derivatives) labeled peptide and between 14 μM and 40 μM protein (depending on the identity of the specific protein in a given experiment). All of the solutions were allowed to equilibrate for 10 minutes a room temperature prior to the first polarization measurement and all experiments were performed in triplicate.

Following determination of the polarization for each sample at their respective concentrations, each of the samples was diluted with 25 nM or 10 nM stock of labeled peptide such that the concentration of the labeled peptides remained constant throughout the course of the experiment. The diluted solutions were also allowed to reach equilibrium for 10 minutes prior to measurement of polarization. The process was repeated for all concentrations.

The data collected from these experiments were plotted in Origin 7.0. The data were fit to the equation $Y = (aX) / (b + X) + c$ (Where Y is the observed polarization at a given protein concentration X, a is the amplitude of the polarization change, b is the KD,

and c is the polarization of free ligand) using the Levenberg-Marquardt least squares method. Errors for individual data points were calculated as the standard deviation of the mean (SDOM) from three individual experiments.

k_{off} Measurments

All k_{off} measurements were performed on the PanVera Beacon 2000 fluorimeter with the instrument set at 25°C. In one set of triplicate experiments, fluorescein-labelled XL_Y at a concentration of 25 nM in storage buffer was incubated with 18 mM GST-Med15(186-619) for 10 minutes. Several successive polarization readings were taken to confirm the system had reached equilibrium. At that time, a 1700-fold excess of unlabeled XL_Y was added in a minimal amount of storage buffer. Polarization readings were taken every 11.5 seconds to monitor the release of the labeled peptide from the protein. An identical set of experiments was performed to determine the off-rate of XL_Y from $Gal4_{dd}$, except the concentration of the labeled peptide was adjusted to 10 nM and the protein concentration was changed to 50 μM. In addition, a set of control experiments was also carried out in which an identical volume of storage buffer was added to the equilibrated peptide-protein solutions to examine the effect of dilution on polarization. The changes in polarization due to dilution were then subtracted from the overall change. The data from these experiments were entered into Origin 7.0 and fit to a single phase exponential decay using the Levenberg-Marquardt least squares method.

J. References

(1)   Darnell, J. E. *Nature Reviews Cancer* **2002**, *2*, 740-749.

(2)   Duncan, S. A.; Navas, M. A.; Dufort, D.; Rossant, J.; Stoffel, M. *Science* **1998**, *281*, 692-695.

(3)   Pandolfi, P. P. *Oncogene* **2001**, *20*, 3116-3127.

(4)   Ptashne, M.; Gann, A. *Genes & Signals*; Cold Spring Harbor Laboratory: New York, 2001.

(5)   Jeong, C. J.; Yang, S. H.; Xie, Y. Q.; Zhang, L.; Johnston, S. A.; Kodadek, T. *Biochemistry* **2001**, *40*, 9421-9427.

(6)   Koh, S. S.; Ansari, A. Z.; Ptashne, M.; Young, R. A. *Molecular Cell* **1998**, *1*, 895-904.

(7)   Park, J. M.; Kim, H. S.; Han, S. J.; Hwang, M. S.; Lee, Y. C.; Kim, Y. J. *Molecular and Cellular Biology* **2000**, *20*, 8709-8719.

(8)   Myers, L. C.; Kornberg, R. D. *Annual Review of Biochemistry* **2000**, *69*, 729-749.

(9)   Lu, Z.; Ansari, A. Z.; Lu, X. Y.; Ogirala, A.; Ptashne, M. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 8591-8596.

(10)   Reeves, W. M.; Hahn, S. *Mol Cell Biol* **2005**, *25*, 9092-102.

(11)   Cheng, J. X.; Gandolfi, M.; Ptashne, M. *Curr Biol* **2004**, *14*, 1675-9.

(12)   Gaudreau, L.; Keaveney, M.; Nevado, J.; Zaman, Z.; Bryant, G. O.; Struhl, K.; Ptashne, M. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 2668-2673.

(13)   Wu, Z.; Belanger, G.; Brennan, B. B.; Lum, J. K.; Minter, A. R.; Rowe, S. P.; Plachetka, A.; Majmudar, C. Y.; Mapp, A. K. *Journal of American Chemical Society* **2003**, *125*, 12390-12391.

(14)   Lu, X. Y.; Ansari, A. Z.; Ptashne, M. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 1988-1992.

(15)   Lu, Z.; Rowe, S. P.; Brennan, B. B.; Davis, S. E.; Metzler, R. E.; Nau, J. J.; Majmudar, C. Y.; Mapp, A. K.; Ansari, A. Z. *Journal of Biological Chemistry* **2005**, *280*, 29689-29698.

(16)   Barberis, A.; Pearlberg, J.; Simkovich, N.; Farrell, S.; Reinagel, P.; Bamdad, C.; Sigal, G.; Ptashne, M. *Cell* **1995**, *81*, 359-368.

(17)   Hidalgo, P.; Ansari, A. Z.; Schmidt, P.; Hare, B.; Simkovich, N.; Farrell, S.; Shin, E. J.; Ptashne, M.; Wagner, G. *Genes & Development* **2001**, *15*, 1007-1020.

(18)   Chan, W. C. *Fmoc Solid-Phase Peptide Synthesis: A Practical Approach*; Oxford University Press: New York, 2000.

(19)   Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948-953.

(20)   Lue, N. F.; Chasman, D. I.; Buchman, A. R.; Kornberg, R. D. *Mol Cell Biol* **1987**, *7*, 3446-51.

(21)   Wang, D.; Zheng, F.; Holmberg, S.; Kohlhaw, G. B. *J Biol Chem* **1999**, *274*, 19017-24.

(22)   des Etages, S. A.; Falvey, D. A.; Reece, R. J.; Brandriss, M. C. *Genetics* **1996**, *142*, 1069-82.

(23)   Lum, J. K.; Majmudar, C. Y.; Ansari, A. Z.; Mapp, A. K. *Acs Chemical Biology* **2006**, *1*, 639-643.

Chapter III

A Small Molecule Transcriptional Activation Domain[2]

A. Background

The development of small molecules to perturb biological systems has typically depended on high affinity ligands for a specific binding site on a protein. While this approach is quite successful in a case where a single, well-defined binding site influences function, many biological processes are carried out by protein complexes whose assembly and function are initiated through a combination of weaker interactions with functionally redundant binding surfaces. In order to regulate these types of systems, small molecules must mimic the binding profile of their endogenous counterparts.

Transcriptional activation is an example of a system which is initiated by a network of low affinity interactions. For example the K_Ds of transcriptional activators for their protein targets have generally been found to be in the low micromolar range. This is likely an essential functional contributor as these proteins must mediate the assembly of the large transcriptional machinery complex onto DNA.[1] In addition, several lines of evidence suggest AD binding sites within the transcriptional machinery are somewhat functionally redundant. For example, ADs from VP16, Gcn4, and Gal4 target an

---

[2] Portions of this chapter were taken from Minter, A. R.; Brennan, B. B.; Mapp, A. K. *Journal of the American Chemical Society* **2004**, *126*, 10504-10505. and Buhrlage, S. J.; Brennan, B. B.; Minter, A. R.; Mapp, A. K. *Journal of the American Chemical Society* **2005**, *127*, 12456-12457. The isoxazolidines in Figure III-2 and III-11 were synthesized by Aaron Minter and Sara Buhrlage and I carried out all of the transcription assays.

overlapping group of transcriptional machinery protein targets yet exhibit very little

sequence homology outside of a general amphipathic composition.[2,3]

ADs are often composed of surreptitious repeats of 5-10 amino acid sequences

and the minimal repeat can itself act as an activation domain when fused to a DBD. This

demonstrates that relatively short sequences are capable of interacting with coactivator

targets, and through these interactions, upregulating transcription.[4-6] One such minimal

activation domain sequence which has been well studied both *in vitro* and *in vivo* is

ATF14 (CGSDALDDFDLDML), a sequence derived from the viral coactivator VP16.[7]

This minimal module thus serves as a useful guide for the design of small molecule ADs.

B. Isoxazolidine Scaffold

ADs typically bind to coactivator proteins in shallow binding surfaces. These

binding surfaces are promiscuous in nature and often allow interactions with numerous

different proteins which can present functionality in an amphipathic manner.[1,8,9] Our

strategy for the creation of a small molecule activation domain involves displaying

functionality observed in endogenous activators on a small molecule scaffold. We

therefore needed a scaffold that had good solubility properties, displayed functionality in

a three dimensional array, and could be obtained with relative synthetic ease. Five-

membered heterocycles have been shown in the past to exhibit all of these properties.

The heteroatom in the ring increases solubility, the constrained ring allows for three

dimensional projections, and numerous synthetic strategies are available for their

construction.

Five-membered heterocyclic rings have also been used by other groups for the disruption of protein•protein interactions. Perhaps the best example is a class of imidazoline compounds constructed by Liu and coworkers to disrupt the interaction between MDM2 and p53.[10] Overexpression of MDM2 has been associated with numerous cancers due to its ability to inhibit the tumor suppressor activator p53.[11] In disrupting this interaction, p53 is allowed to upregulate apoptotic genes and prevent uncontrolled growth. These imidazoline molecules, termed the Nutlins, bind tightly to MDM2 in the p53 binding site and therefore block MDM2's ability to mask the p53 activator (Figure III-1). Particularly relevant to our work, p53 forms a helix upon binding to MDM2, and it is this helix that is mimicked by the Nutlins and prevents the interaction. In an analogous manner, amphipathic activation domains form helices upon binding to coactivators. We therefore inferred that a five member heterocycle with appropriate functionality could generally mimic an amphipathic activation domain allowing it to interact with coactivator proteins an upregulate transcription when localized to DNA.
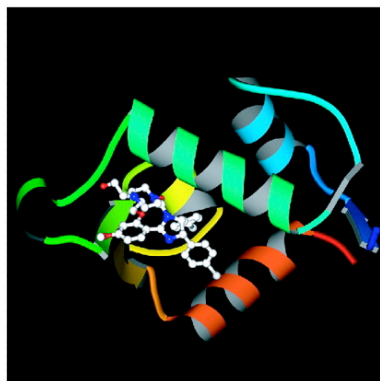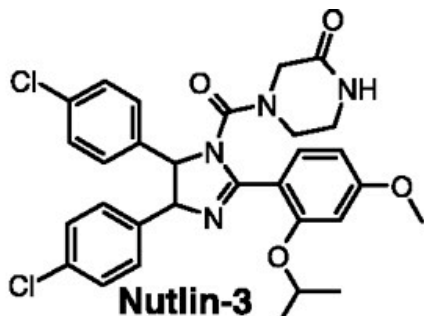

Figure III-1: Nutlin-3 was discovered through a screen of molecules which could disrupt the MDM2•p53 interaction. Nutlin-3 binds to MDM2 through mimicking the helix in p53 responsible for binding (right). Figure Reprinted from (10) with permission.

Key functional groups found in the minimal peptidic AD module ATF14 include

phenyl, hydroxyl, carboxylic acid, and isobutyl moieties.[12-15] In order to identify a

minimal functional unit for a small molecule-based activation domain, a series of

isoxazolidines were synthesized by another graduate student (Aaron Minter), in which

these important groups were appended to an isoxazolidine. In addition, each

isoxazolidine was also coupled to methotrexate (Mtx) in order to localize it to DNA for

testing as an AD. Our initial target compounds are shown in Figure III-2.
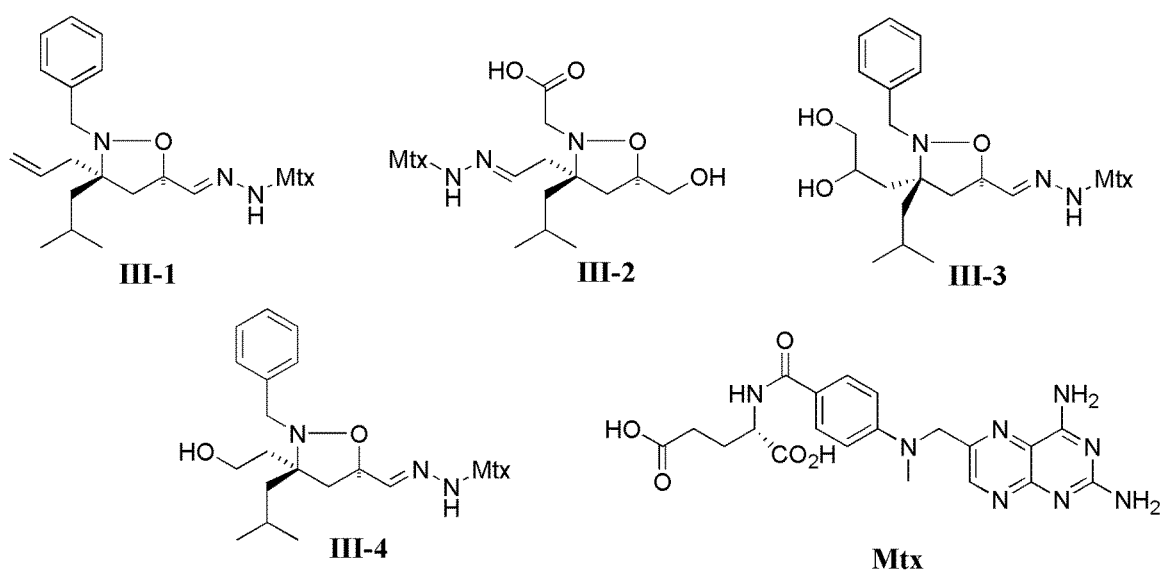
Figure III-2: Initial targeted isoxazolidines contained functionality often observed in
endogenous ADs.

Isoxazolidine **III-1** contains three hydrophobic groups (benzyl, allyl, and

isobutyl) appended at the N2 and C3 positions, while isoxazolidines **III-2**, **III-3**, and **III-4** contain a combination of polar (hydroxyl and carboxylic acid) and hydrophobic (benzyl

and isobutyl) functionalities at N2, C3, and C5. The core of the isoxazolidines was

obtained via a 1,3-dipolar cycloaddition, and the relative stereochemistry at C3 and C5

was set with a Grignard addition to the key intermediate (Figure III-3). Following the

Grignard addition, alkylation of N2 allows introduction of the benzyl or carboxylic acid

at that position (Figure III-3).[16-18] All the isoxazolidines were made as racemic mixtures. The final step of the synthesis of **III-1**-**III-4** was hydrazone formation with methotrexate hydrazide, used to localize the isoxazolidines to DNA in the functional assay. I then synthesized ATF14 through solid phase peptide synthesis and coupled it to Mtx for direct comparison with the small molecules.
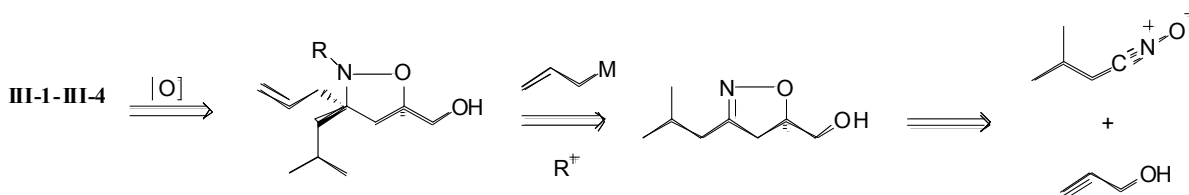
Figure III-3: Retrosynthetic strategy used to prepare isoxazolidines

C. Testing the Isoxazolidine Activation Domains

In order to determine if our small molecule ADs were capable of upregulating transcription, they had to be localized to DNA and transcription levels measured. As previously mentioned, naturally occurring transcriptional activators are composed of both a DBD as well as an AD. One possibility was to couple our small molecule directly to a protein DBD; however side reactions and purification steps make this a very difficult option. Alternatively, we decided to use a variation of a two-hybrid assay developed in the lab of Dr. Virginia Cornish at Columbia University.[19]

Traditionally, a two hybrid assay is used to determine if two proteins interact. Typically, Protein A is coupled to a DBD and Protein B is coupled to an AD. If Protein A and Protein B interact with each other, a complete activator is reconstituted and activation is observed in a reporter gene with binding sites for the DBD. In our case, the

first hybrid protein was composed of a fusion between the bacterial DBD of LexA and E.

coli dihydrofolate reductase (eDHFR).[19] The second hybrid in our case was not a protein,

but instead a fusion between the small molecule methotrexate and our isoxazolidines.

Methotrexate is an inhibitor of eDHFR with a K_D of approximately 1 nM.[20] Since the

isoxazolidine is covalently linked to methotrexate and methotrexate binds to eDHFR

which is expressed as a fusion protein with the LexA DBD, the isoxazolidine is localized

to DNA containing LexA binding sites (Figure III-4).
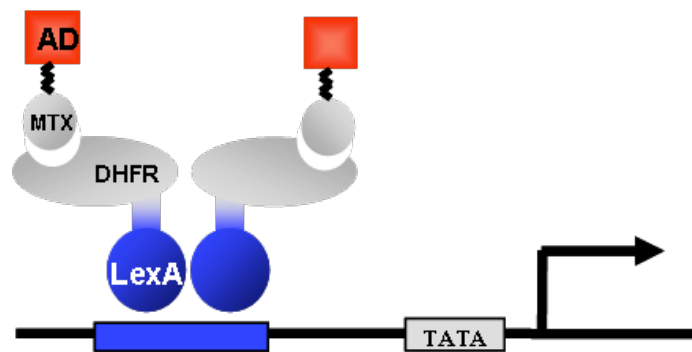
Figure III-4: A diagram demonstrating the method of localizing an activation domain to DNA in a "two-hybrid" type manner. The DBD of LexA (blue) is expressed as a fusion protein with bacterial DHFR (gray). This protein binds to DNA bearing LexA binding sites. Any molecule covalently linked to methotrexate (Mtx) is therefore localized to DNA through the interaction of Mtx and DHFR.

The next question to address was the method of monitoring the RNA output.

Typical methods of *in vitro* transcription involve the use of radioactive isotopes.[21-23]

Generally, radiolabeled nucleotides are incorporated into an RNA transcript. The

transcript can then be run on a gel to separate the incorporated from the free radioactive

nucleotides. This gel is exposed to X-ray film and the band intensity can be quantified

(Figure III-5).

DNA   RNA

Run reaction with → Stop reaction → Expose to X- → Quantify band
radioactive NTP's   & run PAGE   ray film   intensity

1. Activator
2. Nuclear extracts
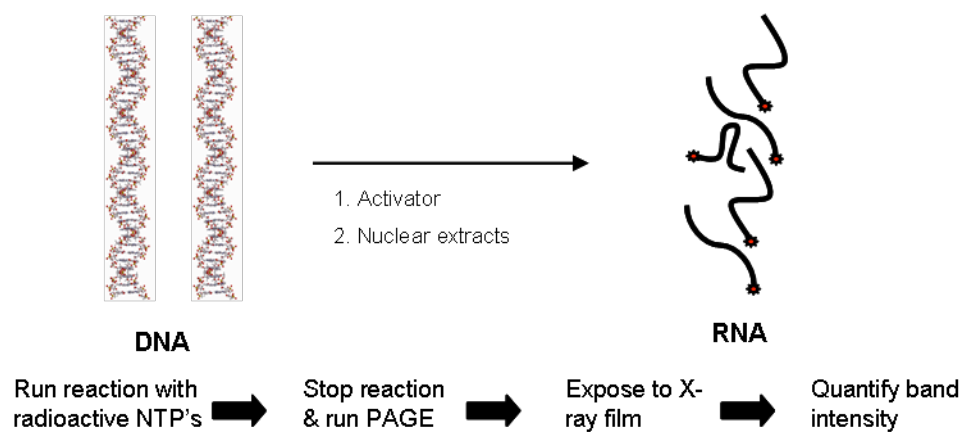
Figure III-5: Radioactive *in vitro* transcription assay.

Although the sensitivity of the radioactive assay makes it extremely valuable, there are some downfalls associated with it. This process can tend to be quite labor intensive and requires multiple RNA handling steps. Although we initially only had four compounds to test, the future plans involve testing libraries of small molecules as ADs. Therefore, we needed to develop a more high-throughput assay capable of testing numerous compounds at once.

One method of monitoring small amounts of RNA with high specificity and sensitivity involves the use of molecular beacons. Molecular beacons are hairpin oligonucleotides with a fluorophore on the 5' end and a quencher on the 3' end.[24] They are designed such that when a hairpin forms, the fluorophore and quencher are in close spatial proximity. Therefore, in this "closed" state, fluorescence intensity is low. The molecular beacon also contains a loop sequence capable of binding to an RNA or DNA. This target oligonucleotide base pairs with the molecular beacon, disrupting the internal hairpin and converting it into an "open" form.. It is more energetically favorable for the beacon to form the 10-20 bp interaction with the target than it is for it to form the internal

hairpin. In the open conformation, the distance between the fluorophore and quencher is increased leading to an increase in the amount of fluorescence observed. (Figure III-6).[25]

<p align="center">**Closed**   Target RNA or DNA →   **Open**</p>
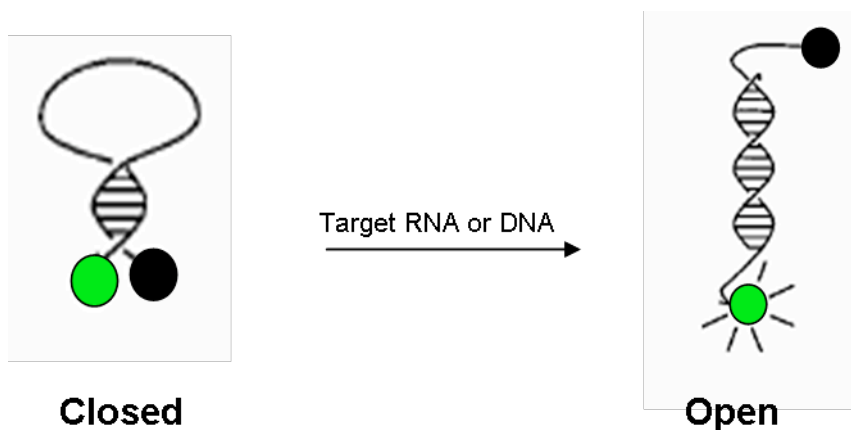
Figure III-6: Molecular beacons are an RNA or DNA oligonucleotide capable of forming a hairpin labeled with a fluorophore (green circle) and a quencher (black circle) at the 5' and 3' ends. Molecular beacons exist in two conformations. In the absence of an RNA or DNA target, they form a hairpin, termed the "closed" conformation, and little fluorescence emission is observed. However, in the presence of their oligonucleotide targets, the beacon converts into an "open" conformation and increased fluorescence is observed.

Molecular beacons have proven useful in numerous applications. The stability of the closed conformation allows them to have remarkable specificity for a target nucleic acid. Because of this, they can discriminate between a single mismatch and have been used to detect mutations. In addition, the presence of a strong quencher greatly decreases the background fluorescence of the probe and therefore allows for a large dynamic range.[26] This has made them useful for observing mRNA in a single living cell and for *in vitro* transcription assays.[23,27]

All of these properties make the molecular beacon an attractive method for monitoring activation of transcription of the isoxazolidines. In order to assure that fluorescence intensity could be observed at target levels expected in the assay (1-20 nM), a titration experiment was carried out. In this experiment, increasing concentrations of

target or a mismatch oligonucleotide were added to a constant concentration of molecular

beacon (100 nM) in transcription buffer (Figure III-7). Over this concentration range, an

increase in fluorescence intensity could be observed, confirming the usefulness of this

method for measuring transcription.
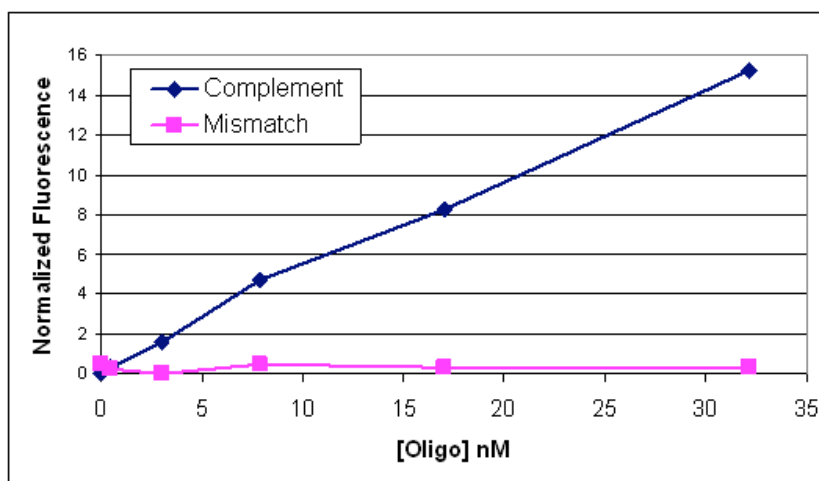
Figure III-7: Titration of complementary oligonucleotide with the molecular beacon
resulted in a linear dose dependent increase in fluorescent signal.

In addition to the use of this molecular beacon to monitor the levels of transcript, a

second molecular beacon was used to normalize the background fluorescence of each

sample. This sequence of the second beacon is not complementary with the sequence of

the transcript produced and thus should remain in the "closed" conformation. The

fluorophore on this beacon also fluoresces at a different wavelength and thus can by

monitored in the presence of the first beacon.

D. Results of Initial Compounds:

With the assay for measuring transcription in place, the individual isoxazolidines

could be tested. The plasmid used for measuring activation contained two LexA binding

sites upstream of an AdML promoter. Since LexA binds as a dimer, this resulted in four

isoxazolidines on each promoter. Following incubation of the DNA template with LexA-eDHFR and the small molecule, the transcription reaction was initiated by addition of HeLa nuclear extracts. mRNA production was directly measured with the molecular beacon and used to determine the activity of all compounds, displayed as a percent activation relative to the positive control ATF14 (Figure III-8).
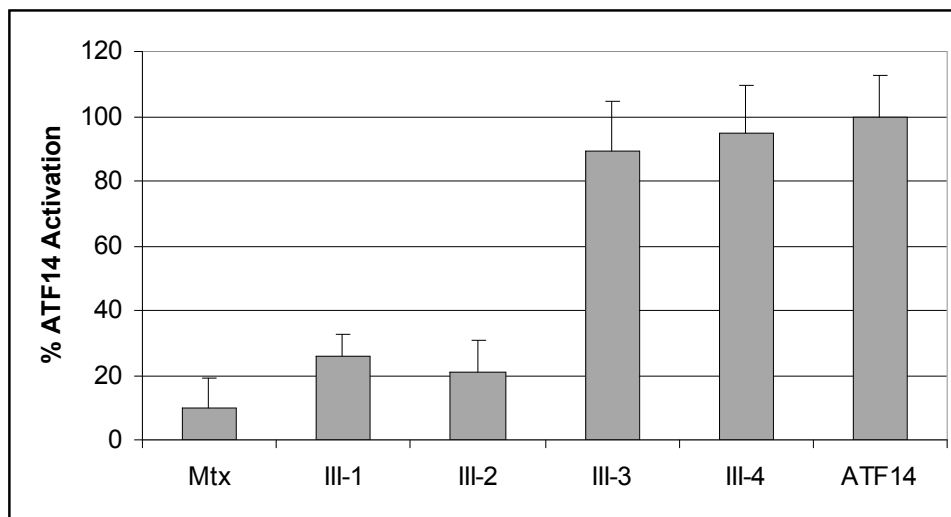
Figure III-8: Results from *in vitro* transcription assays. The activity of each compound represents the average of at least three individual experiments with the indicated error (SDOM).

Remarkably, isoxazolidine **III-4** was as active as the positive control ATF14 despite a considerable difference in size (MW 290 vs. 1674). Further it was found to be the most potent of all the isoxazolidines tested, with 5-7 fold activation over basal transcription. Similar to natural activation domains such as ATF14, a balance of hydrophobicity and polarity is important for overall potency. Substantially increasing the hydrophobicity (**III-1**) or the polarity (**III-2**) led to a dramatic decrease in function. In contrast, slightly increasing the polarity at C3 by the incorporation of an additional hydroxyl group (**III-3**) was well tolerated.

We next carried out experiments to further probe the mechanism of activation. The first of these examined the importance of localizing the small molecule to DNA. In order to address this, Aaron Minter synthesized a version of **III-4** without methotrexate attached, and thus would not bind to the protein DBD. As expected, the acetal version of the active compound (**III-5**) was unable to upregulate transcription, demonstrating the requirement for localization to DNA (Figure III-9a). Further, a competitive inhibition experiment was carried out in which an excess of (**III-5**) was added in the presence of the active compound. Competitive inhibition experiments with endogenous activators typically result in inhibition of transcription due to titration of activator binding sites within the transcriptional machinery.[28] In an analogous manner, the addition of 1 μM **III-5** inhibited the ability of **III-4** to activate transcription (Figure III-9b). In concert, these results suggest a mechanism of function in which the DNA bound isoxazolidine activate transcription through the recruitment of coactivator proteins to the promoter.

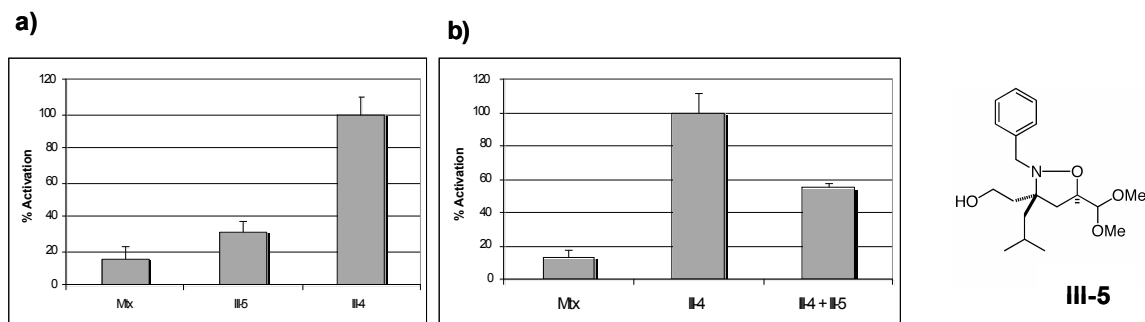**Figure III-9:** (a) Activation by the isoxazolidine is dependent on localization to DNA. (b) Addition of 1 μM free isoxazolidine (**III-5**) inhibits activation of **III-4**, likely through titration of coactivator binding sites.

E. Conclusions from Initial Studies

Two factors likely play a role in isoxazolidine **III-4** being nearly as active as ATF14 despite the size difference. As an organic molecule, **III-4** is resistant to

proteolytic degradation and thus probably has a longer lifetime in the transcription assay. In addition, structural studies suggest that while many natural activation domains are unstructured in solution, helix formation often accompanies activator●target interactions. In contrast, **III-4** likely populates conformations more closely related to the bound state due to structural constraints imposed by the ring. Both factors should also contribute to the function of **III-4** in cells and will be discussed in detail in Chapter IV. Furthermore, when sequences of minimal modules such as ATF14 are reiterated to make longer activation domains, the levels of transcription elicited increase synergistically. Thus considering **III-4** as a minimal functional module, it would be expected that oligomers of this isoxazolidine should exhibit increased potency. This represents the first example of a small molecule transcriptional activation domain and is a valuable tool for studying the mechanistic details of eukaryotic transcriptional activation.

F. Stereochemical Promiscuity in Artificial Transcriptional Activators

Endogenous activators target an overlapping set of coactivators despite little sequence homology, suggesting permissive binding sites. Further support of the permissive nature of coactivator binding sites was demonstrated by Verdine and Coworkers in a series of experiments carried out on a derivative of the activation domain of VP16.[29] In these experiments, Verdine used a modified two hybrid assay in order to determine if the enantiomer of the VP16 derived activation module ATF29 could upregulate transcription *in vitro* and in cell culture (Figure III-10a). If the binding sites for this AD were specific for the natural L version of the peptidic activator, one would not expect the D version to activate.
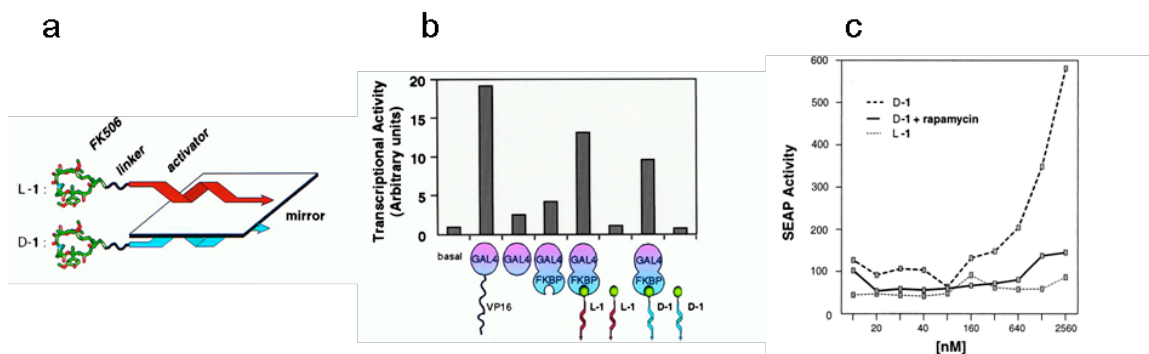
Figure III-10: a) The L and D versions of ATF29 were coupled to FK506 in order to localize them to DNA and test for activation. b) The natural L-version of the peptide, labeled L-1, and the unnatural D-version, labeled D-1, activated transcription *in vitro* only when coupled to FK506, demonstrating the need for DNA localization. c) Only D-1 activated transcription in cell culture, likely due to its proteolytic stability. Furthermore, activation is abolished in the presence of rapamycin, a compound which inhibits the interaction between FK506 and FKBP. Figure from (29).

As demonstrated in Figure III-10, both the L and D peptide activated transcription *in vitro* while only the D peptide activated transcription in cells. The latter result is likely due to proteolytic stability of the D peptide in cells as compared to the L version. Evidently, specific placement of functionality does not appear to be critical for the function of endogenous activators. In order to see if this was also the case for our small molecule activation domain, we decided to carry out a positional "mutagenesis" experiment.[30] Different analogues of **III-4**, our most active small molecule, were evaluated in which identical side chains were placed in various positions within the isoxazolidine scaffold. In our original experiments, all isoxazolidines were prepared as racemates and were tested as stereoisomeric mixtures. Sarah Buhrlage and Aaron Minter, fellow graduate students in the lab, therefore synthesized each enantiomer of **III-4** (**III-4a** and **III-4b**), as well as a diasteriomer (**III-6**), and two positional isomers (**III-7** and **III-8**) for this study (Figure III-11).
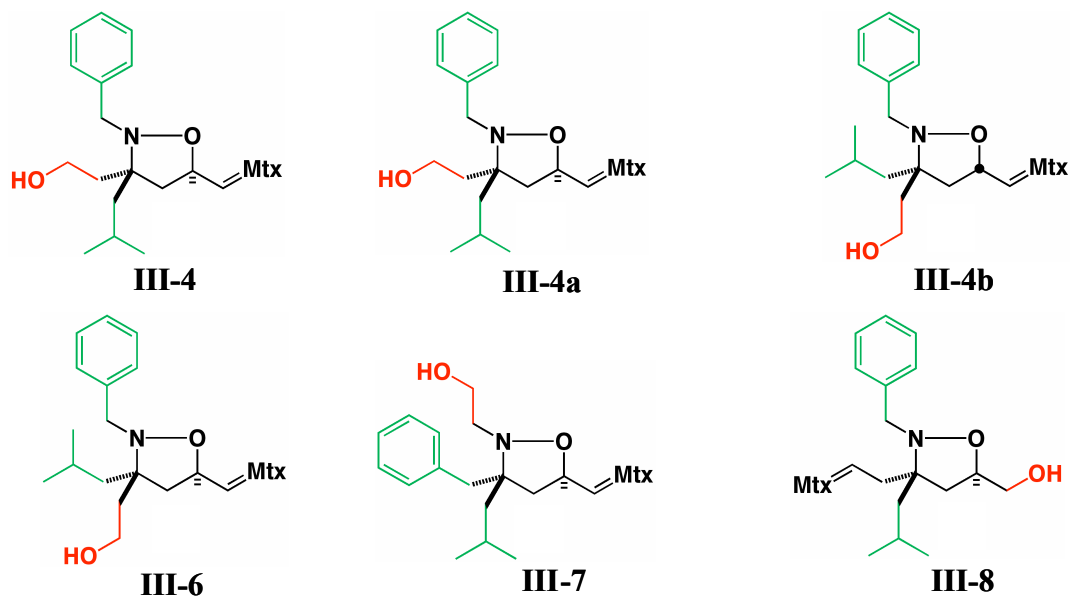
Figure III-11: Compounds used for positional "mutagenesis" study. Hydrophobic side chains of the isoxazolidine are colored green and polar side chains are red.

All compounds contain the same functional groups found in the original active compound (**III-4**) but in varying three dimensional orientations. Analogous to endogenous activators, we hypothesized that all molecules would function as ADs. Testing of the molecules was done in an identical manner as earlier described and the activity of each compound was compared to that of ATF14, our positive control.

We observed that the activities of the enantiomers **III-4a** and **III-4b** are indistinguishable from that of AD **III-4** containing a mixture of the two (Figure III-12). Isoxazolidines **III-6, III-7,** and **III-8** differed more significantly in the presentation of the amphipathic functional groups due to stereochemical changes (**III-6**) or positional changes (**III-7** and **III-8**). However, in line with our hypothesis, all function well as ADs. Only **III-8** showed noteworthy attenuation in activity with a 35% decrease in function.
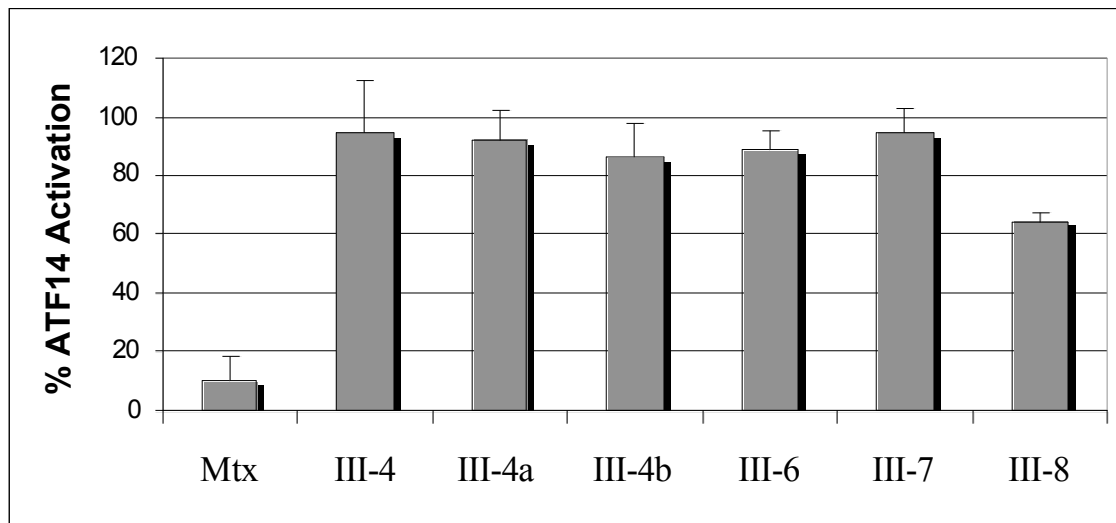
Figure III-12: Results from *in vitro* transcription assays of positional "mutants".

The conserved activity across isoxazolidines **III-4, III-4a, III-4b, III-6, III-7, and III-8** parallels the functional behavior of the endogenous amphipathic ADs that this molecular class was designed to mimic. For example **III-4a** and **III-4b** each activate transcription to similar levels, which was also observed with L and D ATF29 peptides previously discussed.[29] Among peptidic ADs, a variety of combinations of polar and hydrophobic amino acids function as ADs, but a hydrophobic/polar balance is conserved.[1,8,31] Like our small molecules, endogenous ADs share a common structural motif. Helix formation often accompanies binding of endogenous ADs to their protein targets.[32-34] Also similar to the isoxazolidines, mutations in natural ADs that disrupt the hydrophobic surface greatly impact potency.[6,35,36] Another remaining question is whether the isoxazolidines target similar binding surfaces as their endogenous counterparts. While this question remains to be fully answered, it will be addressed in more detail in Chapter IV. Overall, however, these data suggest that isoxazolidines are unlikely to be the only suitable scaffolds for the construction of small molecule ADs. Rather, a variety of appropriately functionalized, conformationally constrained small molecules should

66

function as well. While this hypothesis is not directly addressed in this work, it is a matter of ongoing investigation by other graduate students in the Mapp lab. This strategy obviates the need to identify high affinity ligands for a single protein target and takes advantage of the remarkable functional flexibility of the endogenous regulatory system.

G. Experimental Details

Synthesis of ATF14

The well-characterized activator ATF14 was used as the positive control in all in vitro transcription experiments. Solid phase peptide synthesis was accomplished with FMOC-protected amino acids, HOBt, and HBTU purchased from Peptides International using an Advanced ChemTech90 synthesizer using established protocols. Peptide-synthesis grade DMF, triethylamine, and morpholine were purchased from Fisher and used without further purification. Coupling efficiencies were monitored using a ninhydrin test kit (Fluka). Next, the protected peptide on solid phase was coupled to the γ-carboxylate of methotrexate using HOBt, HBTU, and triethylamine in DMF. Cleavage of the peptide from solid phase and deprotection was carried out in 95% TFA, 2.5% H_2O, and 2.5% triisopropylsilane. Methotrexate-coupled ATF14 was purified to homogeneity using reverse-phase HPLC on a C18 column with a gradient solvent system (buffer A: 0.1% TFA, buffer B: CH_3CN). Lastly, characterization was carried out using ESI-MS and UV-vis spectroscopy. UV (λ_{max} nm) 260, 306, 373; ESI-MS calculated for [Mtx-CGSDALDDFDLDML + 2 H]$^+$: 984.1, found 984.4.

General molecular biology methods

HeLa nuclear extracts, nucleotide triphosphates, and MgCl₂ were purchased from Promega and used in accordance with manufacturer instructions. Restriction endonucleases and T4 DNA Ligase were purchased from New England Biolabs and used as directed. Oligonucleotides, DH5α *E. coli*, and BL21(DE3) pLysE *E. coli* used for plasmid construction, amplification, and protein expression were obtained from Invitrogen. Molecular beacons were purchased from Biosearch Technologies. The plasmid encoding LexA-eDHFR (pWA02) was generously provided by Dr. Virginia Cornish (Columbia University). All other chemicals and supplies were purchased from Fisher unless otherwise noted.

Template construction

The DNA template for *in vitro* transcription was constructed from the pUC18 plasmid. First, the AdML promoter was inserted into the plasmid by annealing the oligonucleotides (5' - AGC TTT GAG GAC GAA CGC GCC CCC ACC CCC TTT TAT AGC CCC CCT TCA GGA ACA CCT GAG CCG ATT GCT GGC GAT CAA CGC GTA AAG CCG ATA GCC GAC - 3') and (5'- CTA GGT CGG CTA TCG GCT TTA CGC GTT GAT CGC CAG CAA TCG GCT CAG GTG TTC CTG AAG GGG GGC TAT AAA AGG GGT GGG GGC GCG TTC GTC CTC AA - 3') producing sticky ends that could be cloned into pUC18 following a double digestion with *Hind*III and *Xba*I. Next the LexA binding sites were inserted 60 base pairs from the TATA box by annealing (5' – GAT CCA CTG CTG TAT ATA AAA CCA GTG GTT ATA TGT ACA GTA GAC TGC TGT ATA TAA AAC CAG TGG TTA TAT GTA CAG TAG AGA

TCT T – 3') and (5' – AAT TAA GAT TCC TAC TGT ACA TAT AAC CAC TGG TTT
TAT ATA CAG CAG TCT ACT GTA CAT ATA ACC ACT GTT TTT ATA TAC AGC
AGT G – 3') producing sticky ends that could be cloned into the AdML containing
pUC18 plasmid following its digestion with *Bam*HI and *Eco*RI. Lastly, the molecular
beacon compliment was inserted by annealing (5'- TAT GAA AAA AAG TTA AGA
CCT ATG CTC GCT - 3') and (5'GCG CAG CGA GCA TAG GTC TTA ACT TTT
TTT CA – 3') producing sticky ends that could be inserted into the LexA and AdML
containing plasmid following its digestion with *Kas*I and *Nde*I. All intermediate and final
plasmids were amplified in DH5α *E. coli,* selected on LB-agar plates containing 0.1
mg/mL ampicillin, and isolated from cultures using a QIAprep Spin Miniprep Kit
(Qiagen). The sequences of the isolated plasmids were verified by sequencing at the
University of Michigan Core Facility.

LexA-eDHFR expression and purification

The LexA-eDHFR plasmid pWA02 was transformed into chemically competent
BL21(DE3) pLysE *E. coli* (Invitrogen) and cells were plated onto LB-agar plates
supplemented with ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL).
Cultures (50 mL) from single colonies were grown overnight at 37 °C (275 rpm) in LB
supplemented with ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL) before
addition to 1L of LB supplemented with ampicillin (0.1 mg/mL). After 3 h, the cultures
were cooled to 16 °C, and expression was induced with IPTG (final concentration 0.5
mM) for 5 h. Cells were harvested through centrifugation at 4 °C and 5000 RPM for 20
minutes. The cell pellet was resuspended in lysis buffer (50 mM NaH_2PO_4 (pH 8.0), 300

mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol, and 10% glycerol), lysed using sonication, and the His tagged protein was isolated from the cell lysate using Ni^{2+} charged agarose beads (Qiagen). The resin-bound protein was then washed with lysis buffer supplemented with 20 mM imidazole and eluted with lysis buffer supplemented with 250 mM imidazole. Storage buffer (40 mM Tris (pH 7.8), 100 mM KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT, and 25% glycerol) was then exchanged with the lysis buffer with the use of an Amicon® Ultra-4 centrifugal filter unit with a NMWL of 5,000 Da (Millipore). Lastly, the protein was snap frozen and stored at −78 °C in storage buffer. Protein purity was determined through SDS-PAGE and protein concentration was measured using Bradford reagent (BioRad).

In vitro transcription

The in vitro transcription reactions were monitored with molecular beacons and accomplished using standard protocols. For each reaction, 100 ng of template DNA was incubated with 50 nM LexA-eDHFR and 50 nM activator or negative control in a buffer containing 5 mM $MgCl_2$, 400 μM of each NTP, 10 μg of salmon sperm carrier DNA, 10 mM HEPES (pH 7.9), 50 mM KCl, 0.1 mM EDTA, 0.25 mM DTT, and 10% glycerol in 50 μL at 30 °C. All reactions were initiated by the addition of 8 units of HeLa nuclear extracts and were incubated at 30 °C for 60 min. Following the 60 min incubation, the reaction was terminated by the addition of NaCl at a final concentration of 500 mM. Molecular beacon was added in simultaneously at a concentration of 100 nM. Fluorescence signal was measured on a SpexFluoromax-2 with excitation and emission wavelengths at 492 nm and 520 nm respectively. An internal standard consisting of a mismatch molecular beacon with a Quasar 670 fluorophore (λ_{ex} = 649, λ_{em} = 670) was

also added to each reaction. Fluorescence intensity at 520 nm was divided by intensity at 670 nm to give a normalized intensity. To obtain fold activation, each individual experiment was divided by the average of two controls reactions containing methotrexate hydrazide.   For Figure III-8 and III-12, the fold activities were converted to percent activation relative to the activity of ATF14.

H. References

(1)   Ptashne, M.; Gann, A. *Genes & Signals*; Cold Spring Harbor Laboratory: New York, 2001.

(2)   Melcher, K. *Journal of Molecular Biology* **2000**, *301*, 1097-1112.

(3)   Brown, C. E.; Howe, L.; Sousa, K.; Alley, S. C.; Carrozza, M. J.; Tan, S.; Workman, J. L. *Science* **2001**, *292*, 2333-7.

(4)   Wu, Y. B.; Reece, R. J.; Ptashne, M. *Embo Journal* **1996**, *15*, 3951-3963.

(5)   Ohashi, Y.; Brickman, J. M.; Furman, E.; Middleton, B.; Carey, M. *Molecular and Cellular Biology* **1994**, *14*, 2731-2739.

(6)   Drysdale, C. M.; Duenas, E.; Jackson, B. M.; Reusser, U.; Braus, G. H.; Hinnebusch, A. G. *Molecular and Cellular Biology* **1995**, *15*, 1220-1233.

(7)   Stanojevic, D.; Young, R. A. *Biochemistry* **2002**, *41*, 7209-7216.

(8)   Giniger, E.; Ptashne, M. *Nature* **1987**, *330*, 670-672.

(9)   Ma, J.; Ptashne, M. *Cell* **1987**, *51*, 113-119.

(10)   Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. *Science* **2004**, *303*, 844-848.

(11)   Vogelstein, B.; Lane, D.; Levine, A. J. *Nature* **2000**, *408*, 307-310.

(12)   Gerber, H. P.; Seipel, K.; Georgiev, O.; Hofferer, M.; Hug, M.; Rusconi, S.; Schaffner, W. *Science* **1994**, *263*, 808-811.

(13)   Courey, A. J.; Holtzman, D. A.; Jackson, S. P.; Tjian, R. *Cell* **1989**, *59*, 827-836.

(14)   Courey, A. J.; Tjian, R. *Cell* **1988**, *55*, 887-898.

(15)   Mermod, N.; Oneill, E. A.; Kelly, T. J.; Tjian, R. *Cell* **1989**, *58*, 741-753.

(16)   Fuller, A. A.; Chen, B.; Minter, A. R.; Mapp, A. K. *Journal of the American Chemical Society* **2005**, *127*, 5376-5383.

(17)   Minter, A. R.; Brennan, B. B.; Mapp, A. K. *Journal of the American Chemical Society* **2004**, *126*, 10504-10505.

(18)   Minter, A. R.; Fuller, A. A.; Mapp, A. K. *Journal of the American Chemical Society* **2003**, *125*, 6846-6847.

(19)   Baker, K.; Bleczinski, C.; Lin, H.; Salazar-Jimenez, G.; Sengupta, D.; Krane, S.; Cornish, V. W. *Proc Natl Acad Sci U S A* **2002**, *99*, 16537-42.

(20)   Sasso, S. P.; Gilli, R. M.; Sari, J. C.; Rimet, O. S.; Briand, C. M. *Biochim Biophys Acta* **1994**, *1207*, 74-9.

(21)   Lue, N. F.; Flanagan, P. M.; Kelleher, R. J.; Edwards, A. M.; Kornberg, R. D. *Methods in Enzymology* **1991**, *194*, 545-550.

(22)   Lee, D. C.; Roeder, R. G. *Mol Cell Biol* **1981**, *1*, 635-51.

(23)   Liu, J.; Feldman, P.; Chung, T. D. *Anal Biochem* **2002**, *300*, 40-5.

(24)   Tyagi, S., Fred Russell Kramer *Nature Biotechnology* **1996**, *14*, 303308.

(25)   Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R. *Proc Natl Acad Sci U S A* **1999**, *96*, 6171-6.

(26)   Bratu, D. P. *Methods Mol Biol* **2006**, *319*, 1-14.

(27)   Peng, X. H.; Cao, Z. H.; Xia, J. T.; Carlson, G. W.; Lewis, M. M.; Wood, W. C.; Yang, L. *Cancer Res* **2005**, *65*, 1909-17.

(28)   Gill, G.; Ptashne, M. *Nature* **1988**, *334*, 721-724.

(29)   Nyanguile, O.; Uesugi, M.; Austin, D. J.; Verdine, G. L. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 13402-13406.

(30)   Buhrlage, S. J.; Brennan, B. B.; Minter, A. R.; Mapp, A. K. *Journal of the American Chemical Society* **2005**, *127*, 12456-12457.

(31)   Ma, J.; Ptashne, M. *Cell* **1987**, *48*, 847-853.

(32)   Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948-953.

(33)   Radhakrishnan, I.; Perez-Alvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. *J. Mol. Biol* **1999**, *287*, 859–865.

(34)   Parker, D.; Jhala, U. S.; Radhakrishnan, I.; Yaffe, M. B.; Reyes, C.; Shulman, A. I.; Cantley, L. C.; Wright, P. E.; Montminy, M. *Molecular Cell* **1998**, *2*, 353-359.

(35)   Regier, J. L.; Shen, F.; Triezenberg, S. J. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90*, 883-887.

(36)   Uesugi, M.; Verdine, G. L. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 14801-14806.

Chapter IV

Small Molecule Activation in Living Cells and Insights into their Cellular

Mechanism of Action[3]

A. Background

The ability to perturb biological processes with small molecules in living cells

presents many challenges not present in a cell-free assay. Factors such as cell

permeability, nuclear localization, and cellular stability can often have dramatic effects

on function.[1] Thus, the activity observed with the isoxazolidines in the cell-free assay

does not necessarily translate to activity in cells. For example, the small molecule

activation domain wrenchnolol (Figure IV-1) developed by Uesugi demonstrated modest

activity *in vitro* but failed to activate transcription in cells. Fluorescence microscopy

experiments suggested this was the result of poor cell permeability characteristics.[1]

---

[3] Portions of this chapter were taken from Rowe, S. P..; Casey, R. J.; Brennan, B. B.; Buhrlage, S. J.; Mapp, A.K. *Transcriptional Upregulation in Cells Mediated by a Small Molecule* **2007** *J. Am. Chem. Soc. accepted.* The experiments represented by Figure IV-5 and IV-6 were performed by Steven Rowe. I performed the experiments represented by Figure IV-2, IV-9, IV-11, IV-12, IV-13, and IV-14. Compounds **IV-1** and **IV-2** were synthesized by Ryan Casey and coupled to OxDex by Steve Rowe. The polyamide from compound **IV-5** was constructed by Dr. Anna Mapp and was conjugated to the isoxazolidine by Sara Buhrlage.
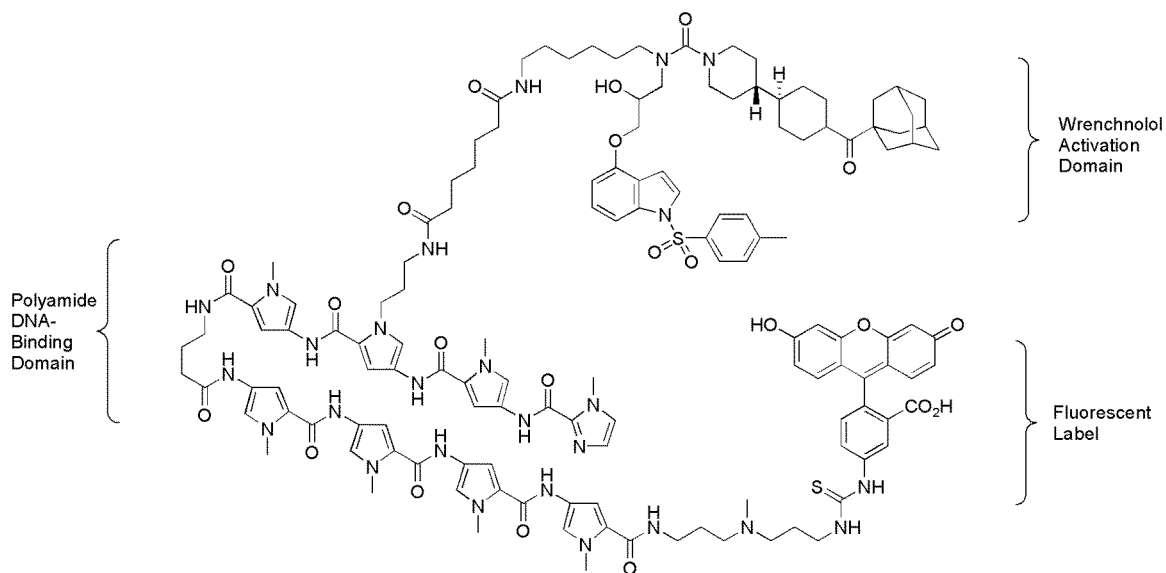
Figure IV-1: A wrenchnolol polyamide conjugate activates transcription in a cell-free system. The fluorescent label was attached in order to monitor cell permeability.[1]

A cell-free system also contains a much more limited range of binding partners. This is an important factor with ADs considering their promiscuous binding profile.[2] The transition from cell-free activity to cellular function can therefore be quite challenging, illustrated by the complete absence of cellularly active small molecule ADs identified to date. In this Chapter, we demonstrate that the isoxazolidine ADs function in both *S. cerevisiae* and human cells. Further, the use of cross interference experiments with isoxazolidines provides evidence of overlapping coactivator targets with an endogenous AD.

B. Cellular Activity in *Saccharomyces cerevisiae*

We chose to initially assess the cellular activity of the small molecules in yeast because the DNA binding strategy used in the cell-free system functions well in *S.*

cerevisiae; methotrexate conjugates cannot be used in mammalian cells due to toxicity.[3] There is significant conservation of the transcriptional machinery across eukaryotes and we therefore anticipated that the activity observed with mammalian nuclear extracts would also be seen in yeast as long as the small molecule was capable of transporting to the nucleus.[4-6] Since a number of methotrexate derivatives of comparable size have been shown to be cell permeable in yeast, we also anticipated permeability with our small molecule ADs.[7,8]

The positive control used in this assay was the peptide ATF14 coupled to methotrexate, the same as in our cell-free system. It was anticipated that due to the large size and potential proteolytic instability of this molecule that other positive controls would need to be used in the long term. In order to assess the cellular activity of **III-4** an assay was carried out in V784Y yeast, a strain harboring the LexA-eDHFR expression vector as well as an integrated β-galactosidase reporter gene with eight LexA binding sites 234 bp upstream of the transcription start site provided by Dr. Virginia Cornish. The methotrexate-isoxazolidine conjugates were dissolved in methanol and added to a 3 mL yeast culture in SC media to a final concentration of 1 μM. Following a 24 hour incubation at 30ºC, a quantitative β-galactosidase assay was performed to assay transcriptional activity (Figure IV-2).
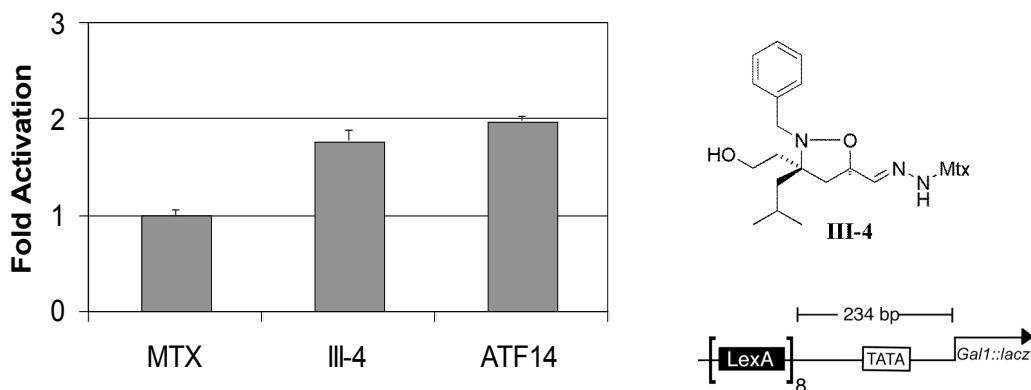
Figure IV-2: Activation of transcription by **III-4** in a yeast cell-based assay. All compounds were added to a final concentration of 1 μM. Each experiment was performed in triplicate with the indicated error (SDOM).

In this context, we observed a nearly two-fold activation of transcription from addition of III-4 to yeast cultures when compared to methotrexate alone, our negative control. Furthermore, the comparative activity between **III-4** and ATF14 was very similar to what was observed in the cell free system, albeit with attenuated overall levels of transcription.

One possibility for the levels of activation observed could be the stability of the compound in cells. As previously mentioned, a hydrazone linkage was used to couple methotrexate to the isoxazolidine. While this conjugate is likely stable over the short time course of the cell-free assay, hydrolysis in cells could partially or completely destroy this linkage. Although a much more stable amide bond was used for coupling ATF14 to methotrexate, proteolysis within the cellular environment likely leads to the degradation of this peptide resulting in the observed low levels of activation observed. In addition to these complications, it is possible that the compounds also have poor cell permeability.

C. Activity in HeLa Cells

Since the majority of the applications envisioned for the isoxazolidine ADs would be in human cells, we chose to investigate the activity in this context rather than optimizing activity in *S. cerevisiae*. As previously mentioned, the toxicity of methotrexate prevents its use in human cells, therefore a different DBD strategy is needed. We decided to utilize an analogous two-hybrid assay developed by the Kodadek group in which a fusion of the DBD Gal4(1-147) and the minimal ligand binding domain of the human glucocortocoid receptor (hGR) (499-777) is used to localize molecules tagged with dexamethasone (an hGR ligand) to Gal4-binding sites at a promoter on a reporter vector (Figure IV-3).[9]
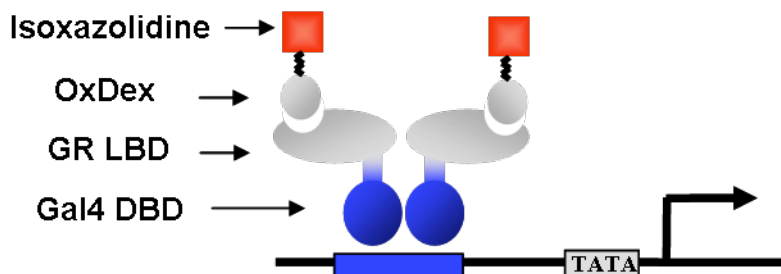

Figure IV-3: Method of localizing the isoxazolidines to DNA in the mammalian cell assay. OxDex represents an oxidized version of the steroid dexamethasone, GR LBD represents the minimal glucocortocoid ligand binding domain, and Gal4 DBD represents the Gal4 DNA binding domain.

For this assay, it was necessary to prepare isoxazolidine derivatives attached to a modified version of dexamethasone (OxDex) (Figure IV-4). Ryan Casey, Sara Buhrlage and Steven Rowe synthesized the new isoxazolidine constructs and Steven and I tested them for transcriptional activation.
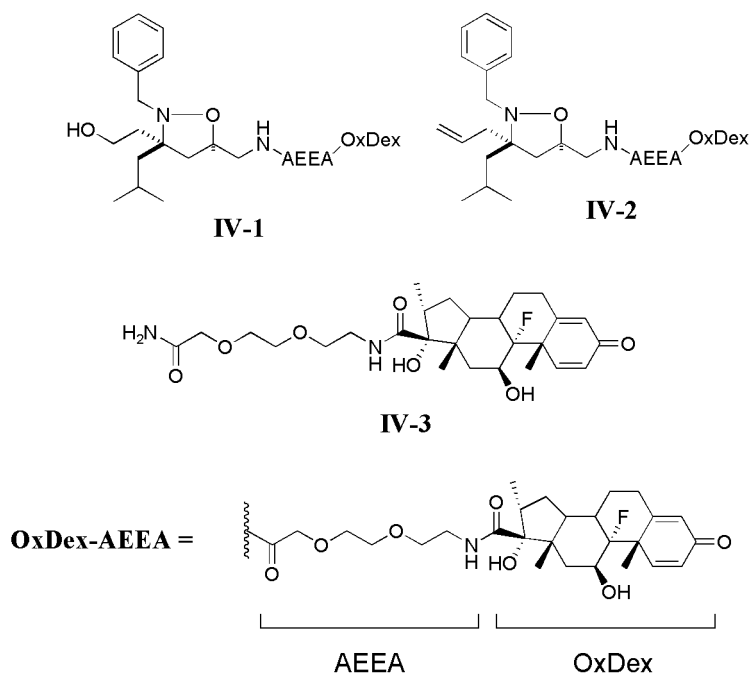
Figure IV-4: OxDex conjugated isoxazolidines tested in the mammalian cell assay.

A dual luciferase assay was used in order to determine the levels of transcription elicited by our small molecules. Prior to addition of the small molecule-OxDex conjugates, HeLa cells were plated on 96-well plates and cotransfected with three plasmids. The first plasmid encoded the gene for a constitutively expressed fusion protein, Gal4 DBD-hGR LBD, the second ecoded a gene expressing Firefly luciferase with five Gal4 binding sites upstream of the transcription start site, and the third encoded the gene for a constitutively expressed Renilla luciferase. Both Firefly and Renilla luciferase can be independently measured, which allows the use of Renilla luciferase as an internal transfection control to normalize against variations in transfection efficiency.

Following transfection, various concentrations of **IV-1**, **IV-2**, or **IV-3** were added to the HeLa cells as a DMSO solution. The final concentration of DMSO in all wells was 1% (vol/vol). In each case, the fold activation of transcription was determined by

dividing the luciferase activity of cells incubated with **IV-1** or **IV-2** by that of cells incubated with **IV-3** at the indicated concentrations (Figure IV-5).
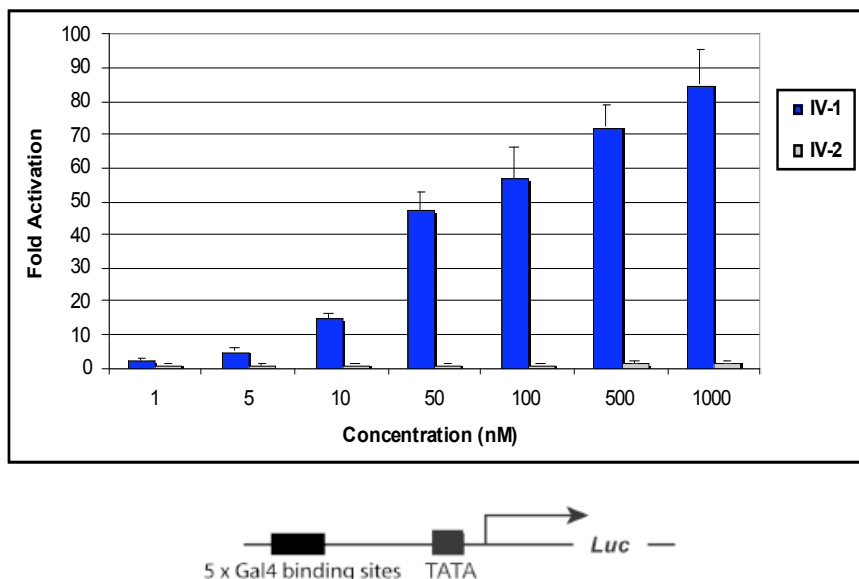

Figure IV-5: Results from the luciferase assay in HeLa cell culture.

As illustrated in figure IV-5, the amphipathic isoxazolidine (**IV-1**) exhibited measurable activity even at very low nanomolar concentrations (5-fold at 5 nM) and 80-fold activity at 1 μM. Additionally, over the concentration range tested, no phenotypic changes were observed with the cells and expression of the Renilla luciferase transfection control remained unaffected. Analogous to the cell-free data, **IV-2** does not function as a transcriptional activation domain at concentrations up to 1 μM. Similar to endogenous activators, the amphipathic characteristics of **IV-1** are important for the overall activity of this small molecule AD.

The DBD and the AD of natural activators typically function in an independent manner.[2] To test if the Gal4-hGR construct contributed to the activity observed with **IV-1** a competitive inhibition experiment was performed. A constant concentration (100

nM) of **IV-1** along with increasing concentrations of **IV-4,** the isoxazolidine that is not conjugated with OxDex, was added to a series of HeLa cells. As illustrated in Figure IV-6, at a concentration of 100 μM, transcriptional activation by **IV-1** was inhibited by approximately 70% (. These data are consistent with the DBD serving as an inert promoter localization scaffold.
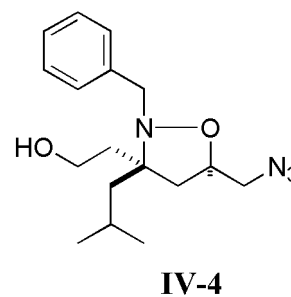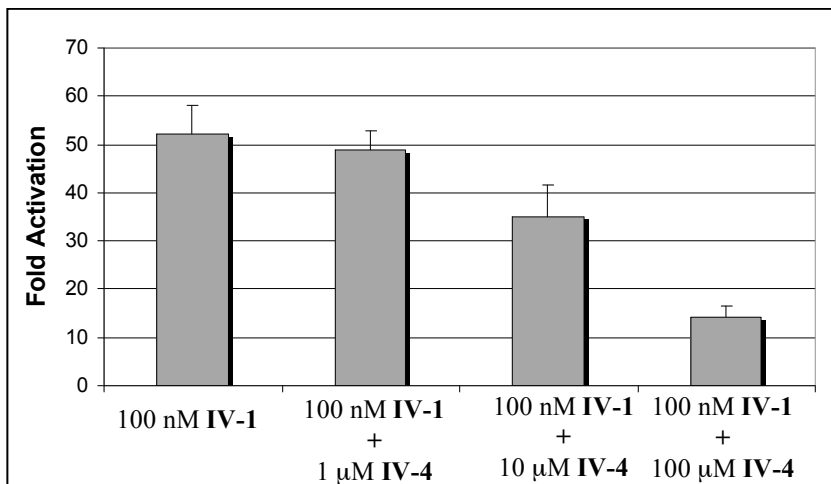

Figure IV-6: Activation by **IV-1** is inhibited by the addition of **IV-4**.

The isoxazolidine does require DNA localization in order to activate transcription as evidenced by the inability of **IV-4** to function as a transcriptional activation domain at concentrations up to 1 μM (Figure IV-7).
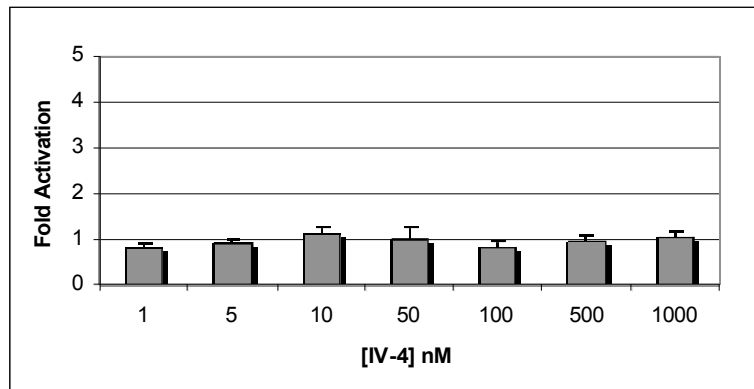
Figure IV-7: No activation is observed with the amphipathic isoxazolidine if it is not coupled to the OxDex DNA localization moiety.

This assay was initially developed by the Kodadek lab in order to discover peptoid ADs which could function in cell culture. Their initial screen yielded one peptoid with extremely potent activity (Figure IV-8).[9]

Figure IV-8: KIX binding peptoid KBPo2.[9]

Although this compound elicited nearly 900-fold activation at its maximal concentration, nearly 10 μM was needed to reach the EC_{50}. In order to more directly compare our amphipathic isoxazolidine activation domain (**IV-1**) with this peptoid, an assay was carried out over a larger concentration range and the EC_{50} calculated (Figure IV-9).

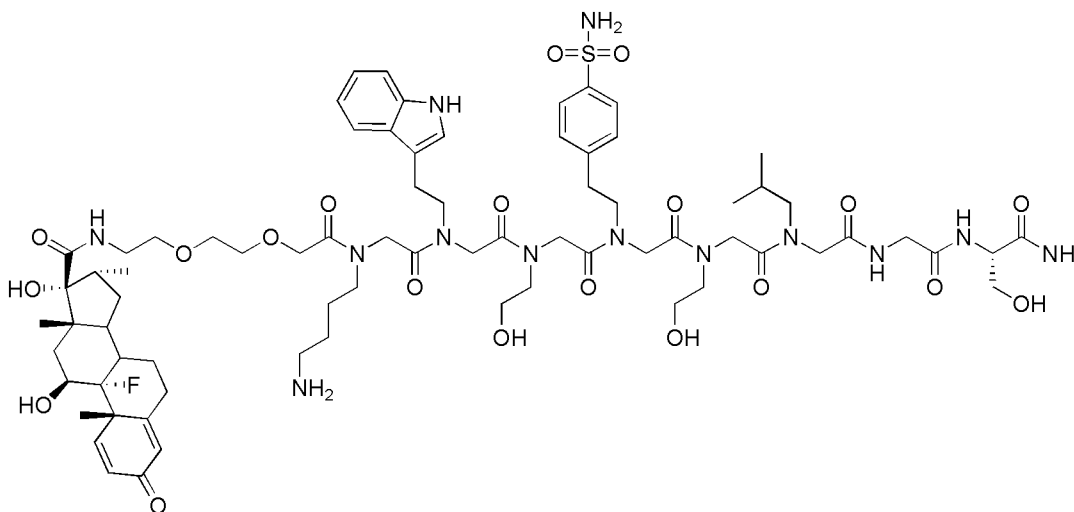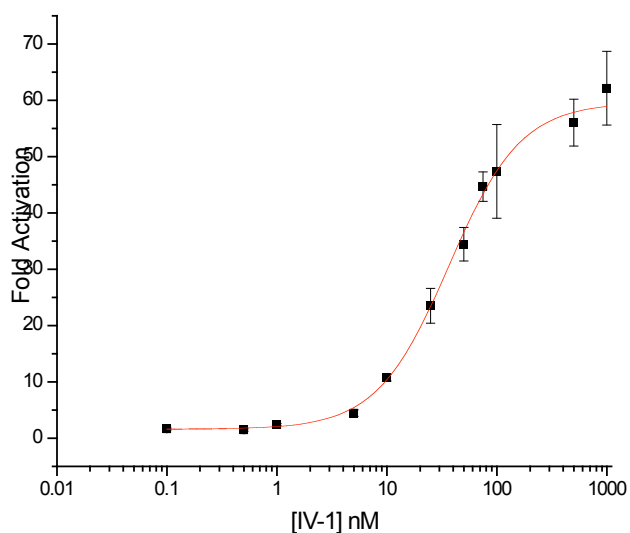**Figure IV-9:** The EC_{50} of **IV-1** is 36 nM. HeLa Cells are incubated with various concentrations of **IV-1** and then the fold activation is determined from the measured luciferase activity compared to that elicited by **IV-3**.

As demonstrated in Figure IV-9, the EC_{50} of the amphipathic isoxazolidine (**IV-1)** is approximately 36 nM. Remarkably, even at very low concentrations, **IV-1** is capable of mediating protein•protein interactions necessary for transcriptional activation. The decreased EC50 of **IV-**1 compared to KBPo2 is likely due to both enhanced cell permeability and the smaller size (1100 and 290 MW). The observed activity at such low concentrations will certainly be advantageous as we move to more complicated systems, where cellular delivery is often the limiting factor to potency.

D. A Fully Synthetic Activator ATF

One advantage to using the Gal4 DBD-hGR LBD construct in the cell-based assay is its ability to traffic to the nucleus. Like many other activator proteins, hGR binds to a masking protein in the absence of a cellular signal. In the presence of its ligand

dexamethasone, the hGR LBD is released from Hsp90 and trafficked to the nucleus.[10,11]

This assures nuclear localization of our isoxazolidine AD upon binding its target.

Although this is advantageous for the discovery of ADs, it circumvents an important

consideration in activator artificial transcription factor (ATF) design. The creation of

"drug-like" activator ATFs will require not only a small molecule AD, but also a small

molecule DBD.

One class of small molecule DBDs which have shown promise as activator ATFs

when conjugated to appropriate ADs are the polyamides developed by Dervan.[1,12-14]

These DBDs were originally inspired from the DNA binding natural product distamycin.

Through the use of pyrrole and imidazole groups on a polyamide backbone, these

molecules are capable of recognizing each of the four base pair combinations (A-T, T-A,

G-C, and C-G).[15] In order to determine if a polyamide conjugated isoxazolidine would

activate transcription in cell culture, Sara Buhrlage coupled the amphipathic

isoxazolidine to the polyamide ImPy7 and I constructed a reporter plasmid with

appropriate binding sites. This particular polyamide binds to double stranded DNA with

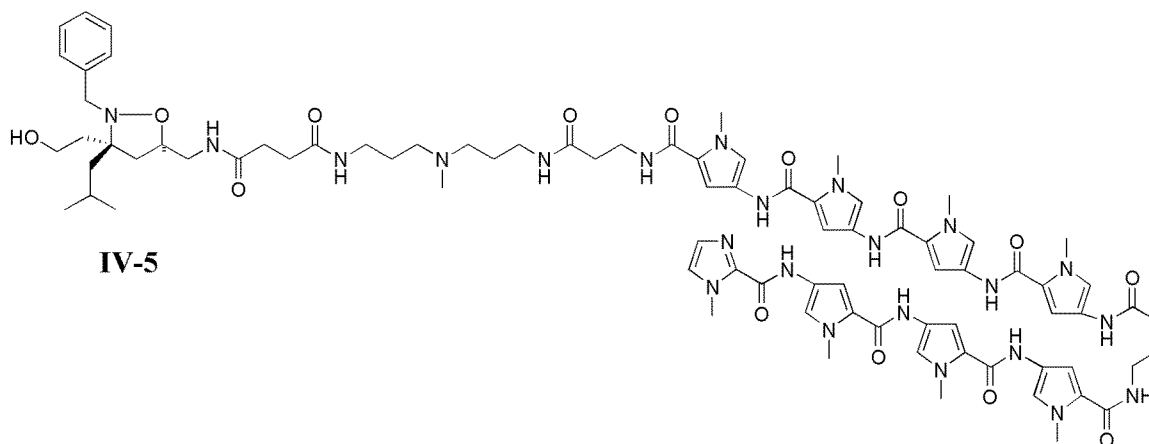the sequence WGWWWW where W represents A or T (Figure IV-10).

IV-5

Figure IV-10: Conjugate of polyamide (ImPy7) with the amphipathic isoxazolidine. This polyamide (ImPy7) binds to the DNA sequence WGWWWW where W represents A or T.

We expected to see a decrease in the activity of the isoxazolidine in this context for several reasons. First, the lack of DNA sequence specificity of this particular polyamide likely results in a substantial amount of binding to other sequences of DNA. Further, the polyamide binds its DNA sequence with much less affinity than the Gal4 protein DBD (about 100-fold weaker), resulting in the need for a higher concentration in the nucleus. The last major difference involves permeability of the molecule. The OxDex steroid is an extremely cell permeable compound and upon binding to its protein target is trafficked directly to the nucleus. The polyamide on the other hand has much lower permeability and requires transport across not only the cell membrane, but also the nuclear envelope. For example, the Kodadek peptoid which demonstrated nearly 900-fold activation in the two-hybrid assay showed only five-fold activity when coupled to a polyamide.[14]

For this experiment, HeLa cells were cotransfected with a plasmid containing either six or zero polyamide binding sites upstream of a firefly luciferase gene and the same Renilla transfection control as earlier described. The polyamide-isoxazolidine

conjugate was added as a DMSO solution at various concentrations resulting in a final DMSO concentration of 1% (vol/vol). The fold activities were determined by dividing the luciferase activity at each concentration by that measured for cells incubated with a free polyamide at the same concentration (Figure IV-11).
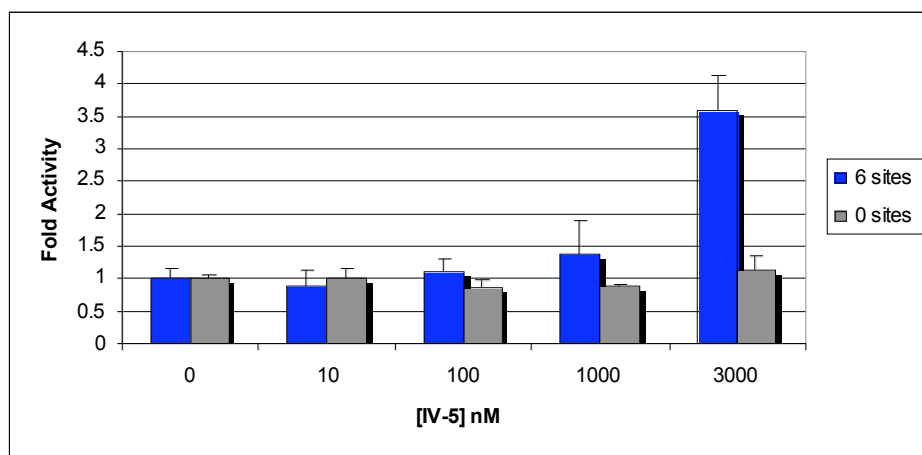
Figure IV-11: The completely synthetic activator ATF **IV-5** activates transcription in HeLa cells.

Consistent with our predictions, higher concentrations of **IV-5** were needed in order to observe activation as compared to the two-hybrid assay. Additionally, the fold-activation at the highest concentration is reduces (3.5-fold vs. 80-fold). Although this experiment needs to be repeated at higher concentrations to assess the EC_{50}, this is an extremely exciting initial result. This represented the first completely synthetic activator ATF composed of both a small molecule DBD and AD which demonstrated activity in living cells. It is likely that factors such as linker length and choice of polyamide DBD will result in more active constructs. Although much more work needs to be carried out on the polyamide-isoxazolidine, this represents the most "drug-like" activator ATF described and an enormous step towards the development of transcription-based therapeutics.

E. Insights into Cellular Mechanism of the Isoxazolidine ADs

All of the experiments thus far are consistent with the isoxazolidines functioning in a similar manner as endogenous ADs. We have demonstrated that the isoxazolidine AD and DBD act in an independent manner (Chapter III + IV), amphipathic characteristics are critical for activation (Chapter III + IV), and specific presentation of functionality is not required for function (Chapter III). This would suggest that the isoxazolidines also target similar surfaces as endogenous activators. In order to address this question, a series of inhibition experiments were performed in yeast and mammalian cells.

Since activator proteins target a limited number of coactivator binding surfaces, it is possible to completely titrate these sites and thus prevent recruitment of the holoenzyme by DNA-bound activators. This phenomenon, named squelching, was first recognized by Ptashne in an experiment with the potent activator Gal4.[16] They found that promoters bearing Gal4 binding sites were inhibited when Gal4 was overexpressed. This can be taken one step further to determine if different activators target the same coactivator binding sites. A cross interference experiment involves expressing or titrating free AD and observing the effect it has on a second DNA-bound activator. If the two ADs target the same set of protein surfaces, inhibition will likely be observed. Conversely, if they target different sites, transcription should remain unaffected.[17]

In order to determine if the small molecules are indeed targeting similar surfaces as endogenous ADs, I performed a series of cross interference experiments. The well studied amphipathic ADs of GCN4, Gal4, and VP16 were tested. Towards this end, Jenifer Lum cloned GCN4(107-144), Gal4(840-881), and VP2 (a peptide derived from

the VP16 AD) into plasmids as fusions with the DBD of Gal4 (residues 1-100). Each of these vectors was transformed into LS41 yeast, a strain bearing a β-galactosidase reporter with a promoter containing Gal4 binding sites, and plated on selective agar plates. Colonies were then chosen and grown in liquid media in preparation for a quantitative β-galactosidase assay. The next day, all yeast cultures were diluted to a density of 7×10^6 (cells / mL) as determined with the OD_{660}. **IV-6** was then added as a DMSO solution to a final concentration of 10 and 100 μM. Following a16 hr incubation at 30ºC the β-galactosidase activity was measured (Figure IV-12).
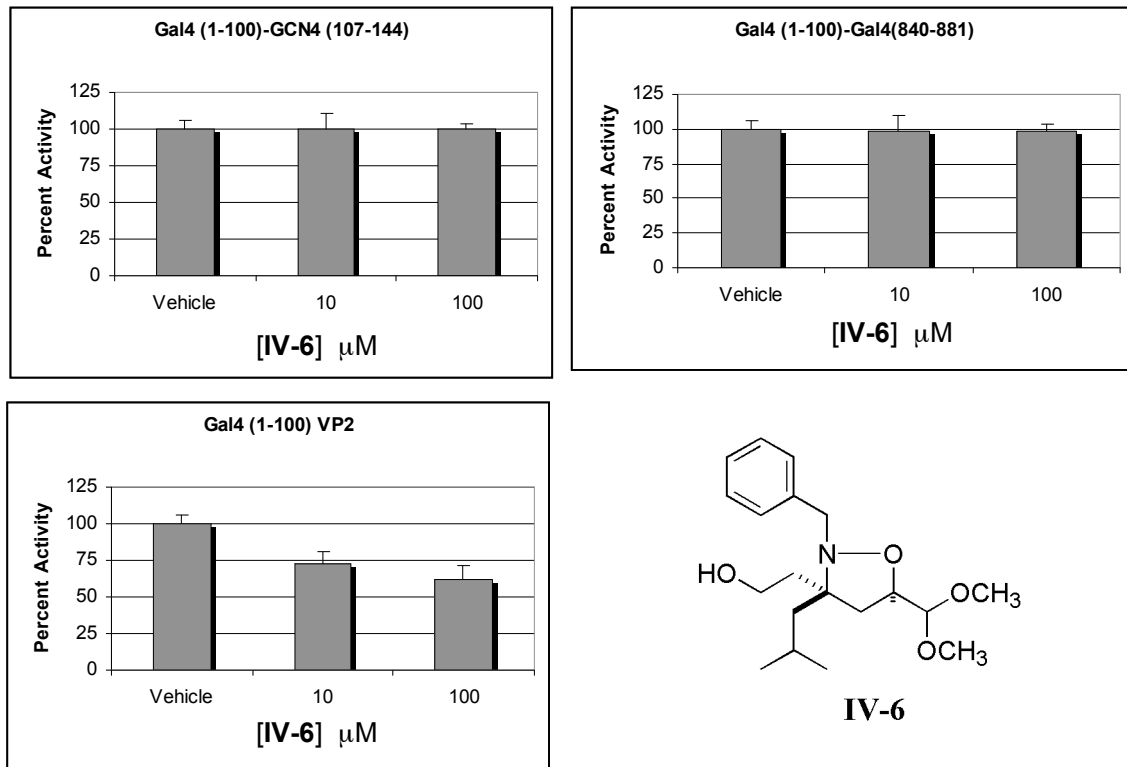

Figure IV-12: Cross interference experiments between **IV-6** and various amphipathic ADs.

In addition to these experiments, two more controls were performed. The first involved the use of the construct Gal4(1-100) without an attached activation domain (Figure IV-

13a). If inhibition were observed in this context, we could conclude **IV-6** disrupts basal transcription. The second control involved the use of an isoxazolidine we would not expect to function as an activator due to its polarity (**IV-7**) (Figure IV-13b).

a)

b)
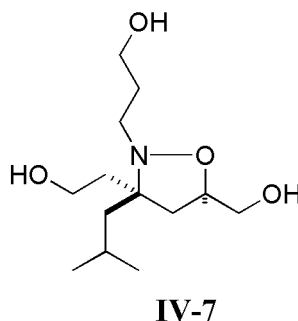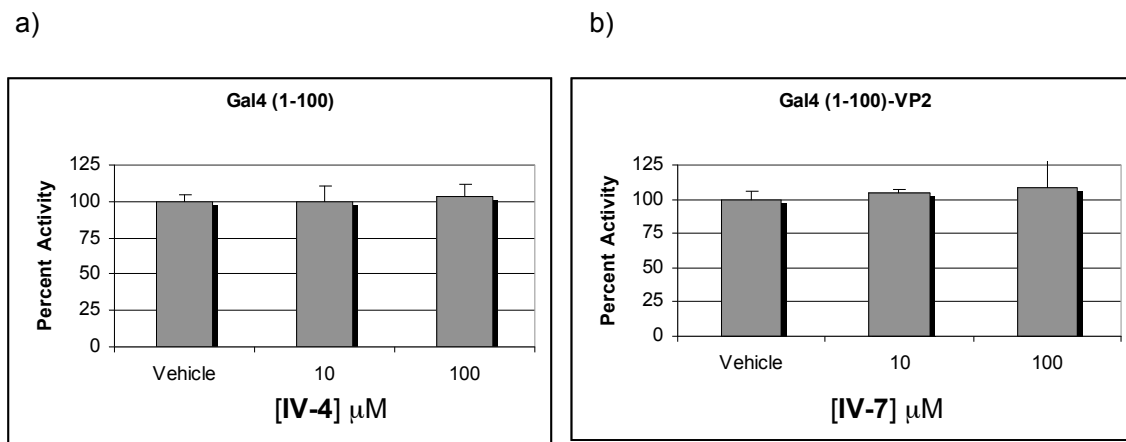
Figure IV-13: a) Inhibition of basal transcription was not observed with the addition of **IV-4**. b) The polar isoxazolidine (**IV-7**) does not interfere with VP2 dependent activation.

As demonstrated in Figure IV-12, the amphipathic isoxazolidine **IV-6** interferes with activation by VP2, decreasing transcriptional activation by about 50% at 100 μM, but not that of GCN4 or Gal4. The fact that the similar effects are observed for both GCN4 and Gal4 is not surprising since cross linking experiments have suggested they target four of the same protein complexes (Mediator, SAGA, NuA4, and TFIID) through three proteins (Tra1, Med15, and Taf12).[18,19] Studies on the VP16 AD have demonstrated these

interactions in addition to the complex TFIIB.[20] At least two possible models exist to explain the cross interference data. One possibility is that the isoxazolidine specifically target TFIIB, a complex not essential for activation by Gal4 or GCN4. A second possibility is that the affinity of the Gal4 and GCN4 ADs for their protein targets is much stronger than that of VP2 for identical targets. Distinguishing between these two models is challenging with the data I have described, although I feel it is unlikely the isoxazolidines target a single protein complex, considering the promiscuous binding profile observed with endogenous activators and is likely recapitulated with the isoxazolidine. For example, recent NMR experiments performed by Sara Buhrlage suggest a specific interaction between the isoxazolidines and the KIX domain of CBP, consistent with the small molecule binding several targets. In order to more fully characterize the interactions made by the isoxazolidines, cross-linking experiments could be performed with the small molecule in the context of transcriptional activation, a project currently under investigation in my lab.

Presumably, the amphipathic isoxazolidine targets similar protein surfaces in both yeast and mammalian cells. In order to investigate this, a similar cross interference experiment was performed in HeLa cells. This was accomplished with the cotransfection of HeLa cells with three plasmids. The first encoded Gal4(1-147)-VP2, the second contained Gal4 binding sites upstream of a luciferase reporter gene, and the third was the Renilla transfection control. Following the transfection, **IV-6** was added with a final concentration of 10 and 100 μM. The cells were incubated for 48 hours prior to measuring the luciferase activity (Figure IV-14).
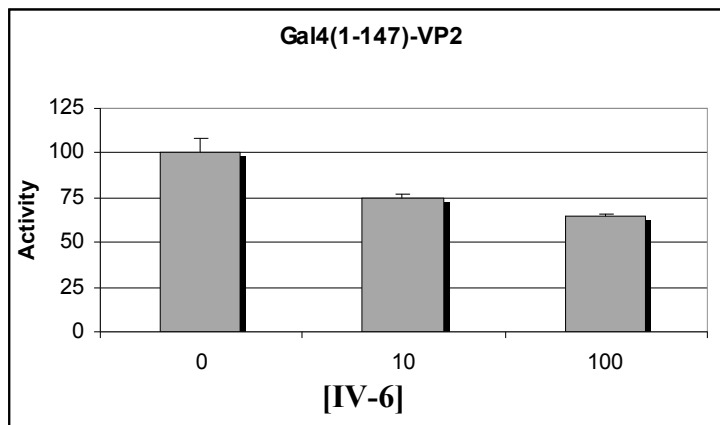
Figure IV-14: Activation of transcription by VP2 in HeLa cells is inhibited by the addition of **IV-6**.

As expected, similar inhibition was observed in mammalian cells as previously seen in yeast. This demonstrates that the amphipathic isoxazolidine likely targets similar proteins in yeast as mammalian cells.

In summary, the amphipathic isoxazolidine represents the first small molecule activation domain with activity in living cells. Furthermore, initial studies indicate function when conjugated to a polyamide DBD, representing an entirely small molecule activator ATF. All studies thus far are consistent with a model in which the amphipathic isoxazolidine targets similar surfaces as endogenous ADs. Through these interactions, the small molecule appears to recruit proteins and protein complexes to DNA and assist in the formation of the pre-initiation complex, the first mechanistic step in eukaryotic transcription. Although many questions about these small molecule ADs are currently being addressed by numerous graduate students in my lab (some of which are the focus of Chapter V), this work represents an enormous step forward in the field.

F. Experimental Details

Yeast Cell Based Assay

V784Y yeast, provided by Dr. Virginia Cornish, Columbia University, was plated on U⁻H⁻ SC agar plates. Individual colonies were chosen and grown overnight (~16 hours) in U⁻H⁻ SC media supplemented with 2% raffinose at 30°C. The following day the cell density was determined by OD_{660} measured on a Carey UV-Vis. The cells were then diluted to give a final concentration of 7 X 10^6 cells/mL in U⁻H⁻ SC media supplemented with 2% raffinose and 2% galactose in a volume of 3 mL. The compounds were then added to the cells as methanol solutions to a final concentration of 1 μM. All concentrations of methotrexate tagged compounds were determined with the characteristic UV-vis absorption of methotrexate at λ_{max} = 257, 302, and 370 nm with extinction coefficients of 23,000, 22,000, and 7,100 $M^{-1}cm^{-1}$. The cells were grown overnight (~16 hours) and the following day a b-galactosidase assay was performed as described in Chapter II.

Hela Cell Assay

HeLa cells (CCL-2) were purchased from the American Type Culture Center (ATCC) and upon arrival immediately plated onto a treated polystyrene Petri dish (Corning) with 10 mL of D-MEM supplemented with 10% FBS, NEAA, and PSG (Invitrogen). Following a 2 hour incubation to allow the cells to adhere, the media was replaced with fresh D-MEM. Cells were then grown at 37°C and 5% CO_2 to approximately 80% confluence.

Upon reaching the desired confluence, the D-MEM was removed from the cells and 2 mL of 0.25% trypisin added. The cells were incubated with the trypisin solution for 10 minutes at 37°C and 5% CO_2. Following the incubation D-MEM was added to the cells in order to remove them from the plate. The solution of media and cells were then centrifuged at 1000 RPM for 1 minute. The supernatant was removed and the cells were resuspended in fresh D-MEM.

The concentration of the cells was calculated using a Hausser Scientific improved Neubauer phase counting chamber hemocytometer. Based on the determined concentration, the cells were diluted to 150,000 cells/mL with the addition of more D-MEM. 100 mL of the diluted cell solution were then added to each well of a flat-bottom 96-well plate. The plated cells were incubated overnight (~16 hours) at 37°C and 5% CO_2.

The media was removed from the plated cells and they were washed once with Opti-MEM and then transfected. The transfection procedure consisted of mixing 50 ng of pCMV-GG (provided by Tom Kodadek), 50 ng pG5 (Promega), and 1 ng of pRLSV40 (Promega) in 50 μL of Opti-MEM for each well to be transfected. A second solution of 0.3 μL of Lipofectamine 2000 (Invitrogen) per well was diluted in 50 μL of Opti-MEM. After 15 minutes of incubation at room temperature, the two solutions were combined and allowed to incubate for another 15 minutes at room temperature. 100 μL of the combined plasmid Lipofectamine 2000 solution was then added to each well and the cells were incubated for 3 hours at 37°C and 5% CO_2.

The plasmids from the transfection served the following purposes. The pCMV-GG served as the expression vector, which expressed Gal4(1-147)-hGR(499-777) under

the control of the CMV promoter in order to localize the small molecules to DNA. The pG5 plasmid served as the reporter vector containing five Gal4 binding sites upstream of a Firefly luciferase reporter gene. The pRLSV40 served as a transfection control and expressed Renilla luciferase on a constitutively active SV40 promoter.

At the end of the three hour incubation, the transfection solution was removed and replaced with 100 mL D-MEM (+FBS +NEAA). To this media was added 1 μL of either DMSO or an appropriate concentration of one of the OxDex conjugated compounds. The cells were then incubated for 40 hours at 37°C and 5% CO_2.

Following the 40 hour incubation, the media was removed from each well of cells and replaced with 20 mL of passive lysis buffer (Promega). The cells were lysed by shaking at room temperature in the lysis buffer for 15 minutes. The dual luciferase assay was then performed using the Promega Dual-Luciferase Assay kit. The luminescence from each sample was measured on a Berthold B12 single cuvette luminometer.

Polyamide Activation Assay

This assay was carried out in the same manner as the HeLa cell assay with a few variations. The pTKGG plasmid was not used, since the DBD in this case was attached directly to the AD. Additionally, the pG5 plasmid was replaced with pE1b (a plasmid containing Firefly luciferase under the control of the E1b promoter) or pE1b-HPBX6 (a plasmid containing the Firefly luciferase reporter under the control of the E1b promoter with six polyamide binding sites upstream)

Plasmid Construction

Both plasmids used in the polyamide activation assay were constructed from pGL3-Basic vector containing the Firefly luciferase gene. Primers (5'-GA TCT TAG AGG GTA TAT AAT GGA TCA-3') and (5'-AG CTT GAT CCA TTA TAT ACC CTC TAA-3') were annealed and inserted into BglII/HindIII cleaved pGL-3 Basic vector to construct pE1b, which has an E1b promoter in front of Firefly luciferase. Primers (5'-CTAGC ATA ACA TTC CAT ATG TTA TAC ATA ACA TTC CAT ATG TTA TAC ATA ACA TTC CAT ATG TTA TAC A-3') and (5'-GATCT GTA TAA CAT ATG GAA TGT TAT GTA TAA CAT ATG GAA TGT TAT GTA TAA CAT ATG GAA TGT TAT G-3') were annealed and inserted into NheI/BglII cleaved pE1b vector to make pE1bHPBX6. The underlined residues in the oligo represent the polyamide binding sites. All plasmids were amplified in DH5α E. coli and sequenced at the University of Michigan sequencing core.

Inhibition Assays

The yeast inhibition assays were performed in LS41, a strain a β-galactosidase reporter gene downstream of five Gal4 binding sites (provided from Dr. Aseem Ansari, University of Wisconsin). The plasmids expressing the Gal4 fusion proteins were constructed by Jenifer Lum. LS41 was transformed with plasmids which expressed Gal4(1-100)-Gal4(840-881), Gal4(1-100)-GCN4(107-144), or Gal4(1-100)-VP2 and plated on selective plates. Individual colonies were chosen and grown in U⁻H⁻ SC media supplemented with 2% raffinose. The following day, the cells were diluted to 7×10^6

cells/mL and **IV-6** was added in DMSO resulting in a final concentration of 10 and 100

μM. The following day β-galactosidase assays were performed as previously described.

The HeLa cell inhibition assay was performed nearly identical as earlier described

except the commercial plasmid pM-VP2 was used in place of pTKGG. This plasmid

expressed Gal4(1-147)-VP2.

G. References

(1)   Kwon, Y.; Arndt, H. D.; Mao, Q.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M. *Journal of the American Chemical Society* **2004**, *126*, 15940-15941.

(2)   Mapp, A. K.; Ansari, A. Z. *Acs Chemical Biology* **2007**, *2*, 62-75.

(3)   Baker, K.; Bleczinski, C.; Lin, H.; Salazar-Jimenez, G.; Sengupta, D.; Krane, S.; Cornish, V. W. *Proc Natl Acad Sci U S A* **2002**, *99*, 16537-42.

(4)   Fischer, J. A.; Giniger, E.; Maniatis, T.; Ptashne, M. *Nature* **1988**, *332*, 853-856.

(5)   Kakidani, H.; Ptashne, M. *Cell* **1988**, *52*, 161-167.

(6)   Ma, J.; Przibilla, E.; Hu, J.; Bogorad, L.; Ptashne, M. *Nature* **1988**, *334*, 631-633.

(7)   Braun, P. D.; Barglow, K. T.; Lin, Y. M.; Akompong, T.; Briesewitz, R.; Ray, G. T.; Haldar, K.; Wandless, T. J. *J Am Chem Soc* **2003**, *125*, 7575-80.

(8)   Miller, L. W.; Sable, J.; Goelet, P.; Sheetz, M. P.; Cornish, V. W. *Angew Chem Int Ed Engl* **2004**, *43*, 1672-5.

(9)   Liu, B.; Alluri, P. G.; Yu, P.; Kodadek, T. *J Am Chem Soc* **2005**, *127*, 8254-5.

(10)   Bertorelli, G.; Bocchino, V.; Olivieri, D. *Pulm Pharmacol Ther* **1998**, *11*, 7-12.

(11)   Pratt, W. B.; Toft, D. O. *Endocr Rev* **1997**, *18*, 306-60.

(12)   Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 3930-3935.

(13)   Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202-205.

(14)   Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew Chem Int Ed Engl* **2007**, *46*, 2865-8.

(15)   Dervan, P. B. *Bioorg Med Chem* **2001**, *9*, 2215-35.

(16)   Gill, G.; Ptashne, M. *Nature* **1988**, *334*, 721-724.

(17)   Min, G.; Kim, H.; Bae, Y.; Petz, L.; Kemper, J. K. *J Biol Chem* **2002**, *277*, 34626-33.

(18)   Reeves, W. M.; Hahn, S. *Mol Cell Biol* **2005**, *25*, 9092-102.

(19)   Fishburn, J.; Mohibullah, N.; Hahn, S. *Mol Cell* **2005**, *18*, 369-78.

(20)   Hall, D. B.; Struhl, K. *J Biol Chem* **2002**, *277*, 46043-50.

Chapter V

Future Directions

Although a significant amount of progress has been made in the development of the isoxazolidine AD, several questions are currently under investigation. The amphipathic isoxazolidine (**IV-1**) is unlikely to be the sole small molecule AD with activity. Changes in hydrophobicity and electronics of the side chain functionality will likely produce small molecules with varied potency which can be used to regulate transcription to desired levels. Although potent activators can often regulate transcription 500-fold or more in cells, this level is not necessarily needed to be physiologically relevant. For example, a genome wide expression analysis of a various lung tumor cells demonstrated up- and down-regulation of most genes around five fold, levels likely achievable from our small molecule ADs.[1] Derivatives of **IV-1** will give greater insight into the mode of action of the isoxazolidines. Specifically, the incorporation of cross-linking agents onto the scaffold will elucidate direct cellular targets of this class of activators. Other areas to address involve the use of different DBD as well as testing the small molecule in a more complex biological system, such as a mouse model.

A. Other Isoxazolidine ADs

In order to determine features of the isoxazolidines that are important for activity, I have tested other derivatives recently synthesized by my co-workers, although a complete SAR analysis has not been completed. One region of the small molecule which can be derivatized with relative ease is the group attached to N2 of the isoxazolidine ring. I have found that changes in the N2 benzyl group have a dramatic effect on the overall potency of the AD in a cellular context (Figure V-1). For instance, conversion to a biphenyl or para-trifluoromethyl completely abrogates activity in cells. On the other hand, the napthyl derivative is still active, although at significantly attenuated levels. While these structural changes could have an effect on the cell permeability thereby decreasing activity, it seems more likely that the benzyl group at this position is important for target interactions. These are just a few of the derivative we would like to test to more fully investigate the isoxazolidine ADs.
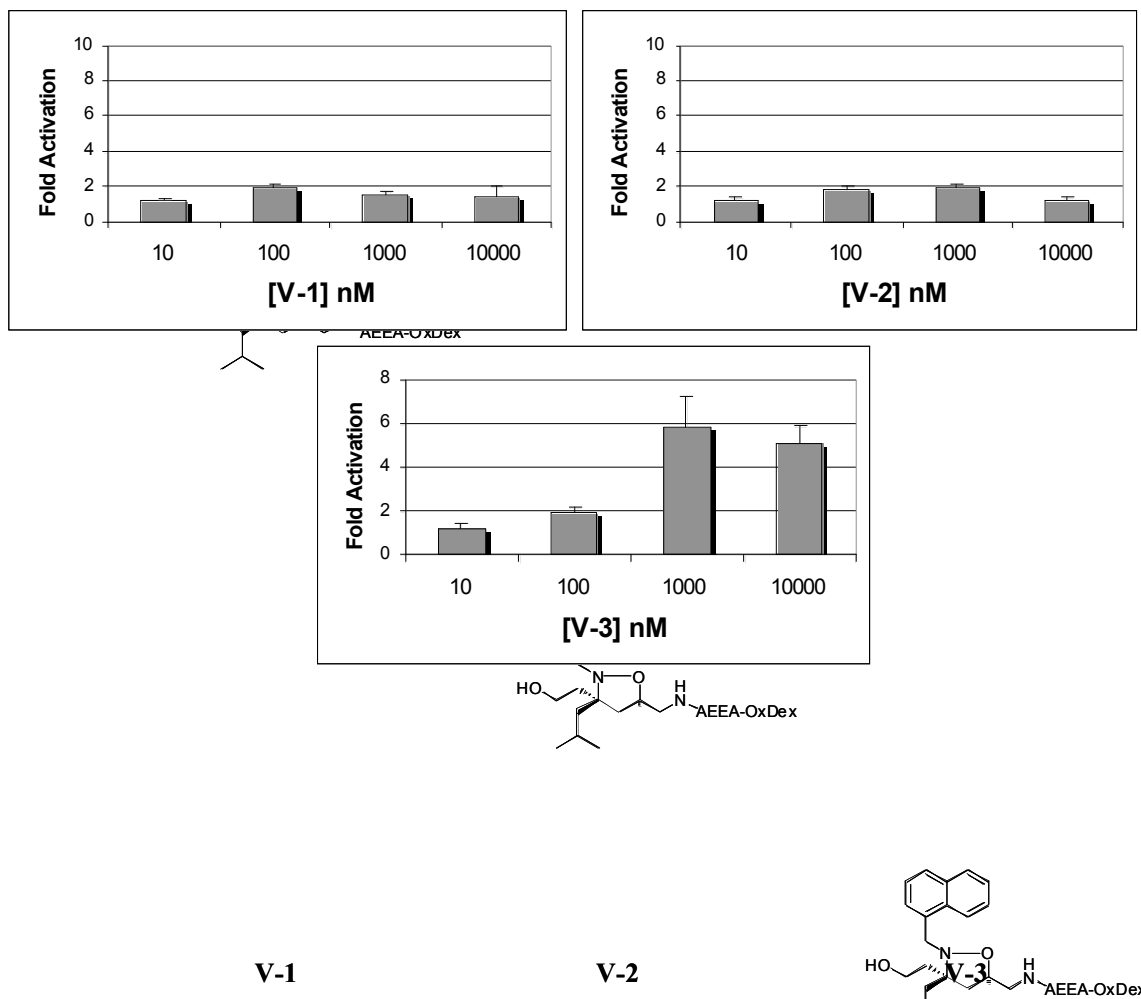
Figure V-1: Cellular activity of derivatives of the amphipathic isoxazolidines.

Supporting this model are NMR structural studies by my co-worker Sara Buhrlage of the

amphipathic isoxazolidine in complex with the coactivator CBP. Upon interaction with

CBP, the benzyl resonances shift significantly. However, additional data is required to

draw firm conclusions and future experiments examining a wider range of substitution at

N2 as well as alterations of the C3 substituents will be important to develop a better

model.

B. Cross-linking studies

The cross interference experiments descried in Chapter IV provide some information into possible cellular targets of the isoxazolidine, but cross-linking of the isoxazolidine in cell culture will result in more direct evidence. Towards this end, a graduate student in my lab (Caleb Bates) has been working on the synthesis of a benzophenone derivative of the isoxazolidine AD (Figure V-2).
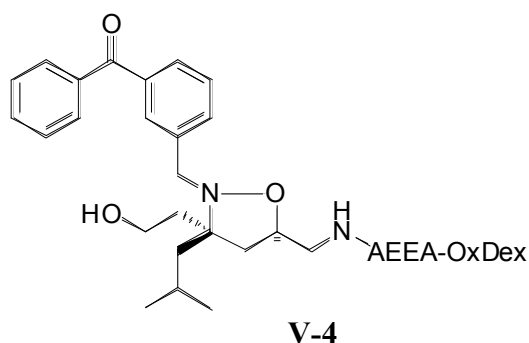

V-4

Figure V-2: A proposed isoxazolidine with a cross-linking agent attached to N3 in order to determine cellular targets.

Upon irradiation of this compound, the benzophenone moiety will cross-link proteins in close spatial proximity.[2] One possible complication involves the derivatization of the N3-benzyl group. As mentioned, this can have a negative impact on activity and could possibly prevent binding of **V-4** to its targets within the transcriptional machinery. If this ends up the case, the cross-linking moiety could be placed in other positions as to prevent interference of activation. A better understanding of the specific binding sites of the isoxazolidine ADs will assist in the design of more potent and specific ADs.

C. Variation in the DBD

The modular nature of transcriptional activators should allow for the incorporation of numerous different DBDs onto our isoxazolidine AD, allowing for a variety of mechanistic studies. This modular nature has been demonstrated in my recent work involving the polyamide DBD, however other choices might prove more successful. The low specificity and stability of the polyamide small molecule DBDs make them less attractive as a long term strategy. As described in Chapter I, another possibility is the use of triplex forming oligonucleotides (TFOs). These molecules are highly specific and stable both in cell culture and *in vivo*. A graduate student in my research lab (Gayle Gawlik) is currently investigating the utility of this DBD in the context of our isoxazolidine AD.

D. Activation of Transcription *in vivo*

Perhaps the most exciting future direction of this work involves the use of the isoxazolidine ADs in a mouse animal model. The first stage of this work involves creating a stably transfected cell line containing both the expression vector (Gal4-DBD hGR-LBD) as well as a luciferase reporter vector. Following the stable transfection, the HeLa cells can be xenografted into mice. If the small molecules can activate in this context, we would expect to see up-regulation of the luciferase reporter upon intra-peritoneal injection of the OxDex conjugated isoxazolidine. Remarkably, imaging technology allows for the visualization of the luciferase without the need for sacrificing the animals. Towards this end, we have begun a collaboration with another research group at the University of Michigan Cancer Center, the facility in which the testing will

take place. The observation of activity in a living biological system would be a

remarkable feat and move us one step closer to the creation of transcription based

therapeutics.

E. References

(1)   Garber, M. E.; Troyanskaya, O. G.; Schluens, K.; Petersen, S.; Thaesler, Z.; Pacyna-Gengelbach, M.; van de Rijn, M.; Rosen, G. D.; Perou, C. M.; Whyte, R. I.; Altman, R. B.; Brown, P. O.; Botstein, D.; Petersen, I. *Proc Natl Acad Sci U S A* **2001**, *98*, 13784-9.

(2)   Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. *Nat Methods* **2005**, *2*, 201-6.