Targeting the Activator Interaction Domain of Mediator Subunit Med25

by

Steven Michael Sturlis

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2016

Doctoral Committee:

Professor Anna K. Mapp, Chair
Professor Carol A. Fierke
Associate Professor Jorge A. Iñiguez-Lluhí
Assistant Professor Brent R. Martin

Acknowledgements

Completing my doctorate is perhaps the most challenging thing that I have accomplished in my brief twenty-seven years and I will be the first to admit that there were plenty of days that I didn't really think I would see it through to the end. Despite the number of times that I swore "today's the day" while getting ready to leave my apartment, the remarkable people around me and the wonderful support system that they represented over the last five and a half years were pivotal in getting me to the finish line. Not to mention that the majority of them probably would have tried to chain me to my bench if I had actually tried to leave.

First and foremost, I need to thank my research advisor, Professor Anna K. Mapp, without whom I literally could not have completed my PhD. The summer that she let me spend in her lab as an REU student was pivotal in my decision to not only attend graduate school, but also helped me to identify Michigan as the ideal institution for my graduate work. Her guidance over the last five and a half years have taught me how to truly think like a scientist, researcher, and at times, educator. Perhaps the single greatest lesson she taught me in my time working in her lab was how to persevere even when projects and experiments deviate significantly from the plan. I will always value her mentorship and the support that she has given me during my tenure at Michigan. I am particularly thankful for her guidance with issues that extended beyond the laboratory from career development to the two-body opportunity (as she would say). Thank you.

I also want to thank the other members of my thesis committee Dean Carol A. Fierke, Professor Jorge A. Iñiguez-Lluhí, and Professor Brent R. Martin. Your helpful suggestions and constructive criticism throughout the years made me a better scientist and pushed me to think critically about my work.

My research advisor while I was a student at Illinois Wesleyan University was Professor Brian B. Brennan and I also owe him a colossal debt of gratitude for helping to

shape me into the scientist that I am today. I was Brian's first student and was a biology major hell-bent on going to medical school when I started working in his lab as a sophomore. Through the course of several illuminating conversations he first convinced me to change my major to chemistry with the promise that it would teach me more about problem solving than a degree in biology could, before ultimately guiding me down a path that led to a summer REU at Michigan working for Anna and ultimately my decision to pursue a PhD rather than an MD. His mentorship throughout my academic tenure as both an undergraduate and graduate student were invaluable and I am profoundly grateful for his friendship.

I was incredibly fortunate to be surrounded by an amazing group of colleagues throughout my time in the Mapp lab who made coming into the lab a joy every day. I'd like to thank Professor Aaron Van Dyke and Dr. Jonas Hojfeldt, who took me under their wings and mentored me when I was but a lowly roton. I also need to thank Jonas for being such an amazing mentor during the summer I worked in the Mapp lab as an REU student as he immediately treated me as a peer and left me clamoring to return to Michigan as a graduate student. I will cherish the time I spent and the remarkably diverse conversations that I had with Paul and JP who started in the Mapp lab at the same time as I did and with my degree completed the last of the "Young Guns" have truly and finally become the the "Old Guard". I've known JP for going on a decade now and his continued friendship through the years is something I will always value. Paul and I collaborated extensively on significant portions of the work published in this thesis, an experience which I thoroughly enjoyed. I also need to thank Andy and Matt for their contributions to this project and for being excellent cubby-mates and their awesome podcast recommendations. To Rachel, Meg, Cassie, Laura, and Omari thank you for your humor and making the Mapp lab such a great environment to work in. I also need to collectively thank the past members of the Lab for their guidance and support throughout the years.

I am truly blessed to have been born into such an amazingly strong and supportive family and there is no way that I could have made it this far without them. Mom and Dad, you kept me going when things got difficult and I cannot possibly express in words how much your love and support over the last twenty-seven years mean to me. I want you both to know that you raised me right and I am the man I am today only because of you.

Molly, Nolan, and Becky, you are the best siblings a guy could ask for (even those of you who made the misguided decision to attend *The* Ohio State University). I love you all.

Finally, I have to thank my better half. Megan, you have been my rock throughout this entire process and I want to first apologize for the amount of crazy that you have had to deal with, particularly over the last month and a half while I was writing this dissertation. I have loved every minute of our relationship and without your constant love and support there is no way I could have made it this far. I am in constant awe of your strength, dedication, and compassion. You make me a better man and I cannot wait to see what the future holds for us. I love you. A lot.

To everyone who has helped me get to this point, from teachers in my earlier years to family, friends, and colleagues that I have not explicitly mentioned I am eternally grateful for your support and guidance. Forever go blue!

Table of Contents

List of Figures

List of Tables

List of Abbreviations

$^{1}H,^{15}N$-HSQC	$^{1}H,^{15}N$-Heteronuclear Single Quantum Coherence
AcID	Activator Interaction Domain (Med25)
ATF6α	Activating transcription factor 6 α
CBP	CREB Binding Protein
CH1	Cystein-histidine rich 1
co-IP	Coimmunoprecipitation
DBD	DNA-Binding Domain
DMEM	Dubelco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
E1A	Adenovirus-encoded E1A 13S protein
ESX	Epithelial specific transcription factor
FBS	Fetal bovine serum
FP	Fluorescence polarization
GACKIX	Gal11, Arc105, CBP/p300, kinase-inducible domain interacting (KIX)
GST	Glutathione transferase
GTF	General transcription factor
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HIF1-α	Hypoxia inducible factor 1 α
HPLC	High-performance liquid chromatography
HSV	Herpes simplex virus
HTS	High-throughput screen
IC_{50}	Half-maximal inhibitory concentration
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
KIX	Kinase inducible domain interacting domain of CBP

MD	Molecular dynamics
MDM2	Murine double minute 2
Med25	Mediator subunit 25
MLL	Multi-lineage leukemia
MMP	Matrix metalloproteinase
mRNA	Messenger RNA
MS/MS	Tandem mass spectrometry
Ni-NTA	Nickel-Nitrilotriacetic Acid
NMR	Nuclear Magnetic Resonance
NR	Nuclear receptor
NR	Nuclear Receptor
PC4	Positive Cofactor 4
PDB	Protein DataBank
pKID	Phosphorylated kinase inducible domain (CREB)
PPI	Protein-protein interaction
PrOF	Protein-observed fluorine NMR
RARα	Retinoic acid receptor α
RNA	Ribonucleic acid
RNAPII	RNA Polymerase II
RPM	Revolutions per minute
RT-qPCR	Real time quantitative polymerase chain reaction
siRNA	Small interfering RNA
TAD	Transcriptional Activation Domain
TAF	TBP associated factor
TBP	TATA binding protein
TFIIA	Transcription factor II A
UV/Vis	Ultraviolet/visible spectroscopy
VP16	Herpes simplex virus protein 16
VWA	Von Willebrand Factor Type A

List of Appendices

Abstract

Transcription, the process by which the expression of every gene encoded in an organism's DNA is regulated, is critical to all cellular processes. As a result of this fundamental importance, transcription is tightly regulated at the molecular level by the assembly of dynamic protein complexes at gene promoters. That the dysregulation of transcription is a cause or consequence of nearly all human diseases underscores the importance of this regulation. Given the dire consequences of the loss of transcriptional regulation in diseases like cancer, it is apparent that small molecules designed to attenuate aberrant transcriptional processes would be particularly useful as mechanistic probes or potential therapeutics.

A mechanism by which this inhibition could be achieved is to target the interactions between DNA-bound transcriptional activators and critical coactivator proteins required for the assembly of the transcriptional machinery at target genes. These interactions typically occur over broad surfaces areas with only moderate affinity, making them challenging to target, though recent advances in PPI inhibitor discovery such as exploiting conformational plasticity within target proteins has led to some success. In particular, the work described in this thesis is focused on disrupting the interactions between a number of transcriptional activators and the Activator Interaction Domain (AcID) of Mediator subunit Med25. This domain contains a unique protein fold not found in any other known transcriptional coactivators, leading to the hypothesis that AcID can be selectively targeted. The activators that bind to the AcID motif have been implicated in a number of cellular processes including the hijacking of transcriptional machinery by viral activators, endoplasmic reticulum stress response, and most interestingly the progression of malignant cells to a metastatic phenotype. Thus, selective inhibitors against this domain will be useful as mechanistic probes or even potential therapeutics against these processes.

Towards this goal, studies designed to elucidate the underlying molecular features that define activator•AcID interactions were performed. These experiments identified a minimal interaction surface within the transcriptional activation domains of an activator binding partner and that electrostatic contacts are critical for the interaction of activators with the domain. Furthermore, hotspot analysis of the VP16 TAD suggests that the interaction occurs over broad surfaces within AcID. These molecular features, combined with data supporting the hypothesis that the AcID motif is relatively plastic and contains distinct activator binding surfaces that may be in allosteric communication provide several potential mechanisms by which AcID-dependent interactions might be disrupted.

In an effort to identify small molecule inhibitors of the AcID motif, screening campaigns against the VP16•AcID and ERM•AcID interactions were completed. In the case of the former, screening identified a number of relatively selective inhibitors belonging to the depside and depsidone classes of natural products, which have previously been identified as inhibitors of activator•coactivator interactions. Biophysical and biochemical studies suggest that these inhibitors perturb AcID-dependent interactions by binding covalently to lysine residues within the putative activator binding sites, thereby potentially disrupting critical electrostatic contacts while simultaneously occluding surfaces necessary for activator binding. Additional evidence suggests that these inhibitors may allosterically induce structural shifts within AcID that contribute to their inhibitory activity. Building on these results, the interaction between ERM, a member of the PEA3 subfamily of ETS transcription factors linked to metastasis, and AcID was screened against a library of natural product extracts in an effort to identify novel non-covalent inhibitors of the AcID motif. Through a series of counter screens and stringent hit filtering steps, a promising extract with potent activity *in vitro* against the ERM•AcID interaction has been identified and is currently undergoing deconvolution to identify the structures of the active natural products within the extracts. The activity and mechanism of action for these molecules will be elucidated and they will subsequently be used as molecular probes to explore the importance of PEA3 subfamily-regulated transcriptional programs in biological processes such as progression to a metastatic phenotype.

CHAPTER 1

Targeting Transcriptional Protein-Protein Interaction Networks[1]

1.1 Abstract

Transcription is a critical process required for maintaining homeostasis and ensuring proper function within all organisms. Given its central importance, transcription is tightly regulated at the level of individual promoters through the formation of dynamic protein complexes, and the dysregulation of the assembly process contributes to a host of diseases in humans. Thus, modulating aberrant transcription through small molecules that target the assembly of these regulatory complexes is an attractive therapeutic goal. It is also a difficult one given the large surface areas and modest affinities of the protein-protein interactions that govern the formation of these complexes. Nonetheless, recent advances in targeting protein interaction networks, such as exploiting compositional and conformational dynamics, have led to some success in targeting these difficult interactions.

In this chapter, the various proteins required for regulated transcription at individual genes are first outlined. The mechanisms by which transcription becomes dysregulated and the functional consequences of this loss of transcriptional control are then discussed. Finally, we describe different approaches for the development of small molecule inhibitors of transcriptional regulators, with a particular focus on inhibitors of activator•coactivator interactions. The lessons learned through the successes and pitfalls of these approaches are ultimately applied to the overall goal driving the work contained in this thesis, namely the development of small molecule inhibitors of a unique transcriptional coactivator domain.

[1] Portions of this chapter were adapted from: Mapp, Anna K.; Pricer, Rachel; Sturlis, Steven. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **2015**, *11*, 891-894.

1.2 Introduction

The existence of endogenous factors that functionally connect DNA-encoded information and protein levels was first postulated by Monod and Jacob in 1961.[1] In the ensuing decades, evidence has mounted that these factors—transcription factors—are also critical drivers of human disease. The overexpression, underexpression, and formation of fusion proteins of transcription factors underpin a range of human diseases, and thus these proteins have high intrinsic value as therapeutic targets. Given the fundamental importance of these proteins, the logical questions are why there are no drugs that directly target these transcription factors and why there are so few quality probe molecules to further dissect the function of these factors. These questions are particularly urgent given the avalanche of new data regarding transcription factor localization through the transformative technological advances in sequencing and genetic manipulation that have occurred in the past decade.

1.3 Overview of Transcriptional Activators

Every cellular organism contains all of the requisite instructions to maintain homeostasis and proper function encoded in the form of DNA. Proteins known as transcriptional activators recognize and bind to specific sequences in the promoters or enhancers of cognate genes and recruit a suite of coregulatory proteins that modulate the transcriptional output of these genes by modifying chromatin structure and/or assembling the RNA Polymerase II (RNAPII) pre-initiation complex.[2]

Transcription is the process by which genetic information encoded within DNA is converted to a transferrable signal through which gene products, namely proteins, are produced. This critical process is completed using a specific class of proteins termed transcriptional activators, which are modular proteins minimally composed of two domains: a DNA-binding domain (DBD) and a transcriptional activation domain (TAD).[3]

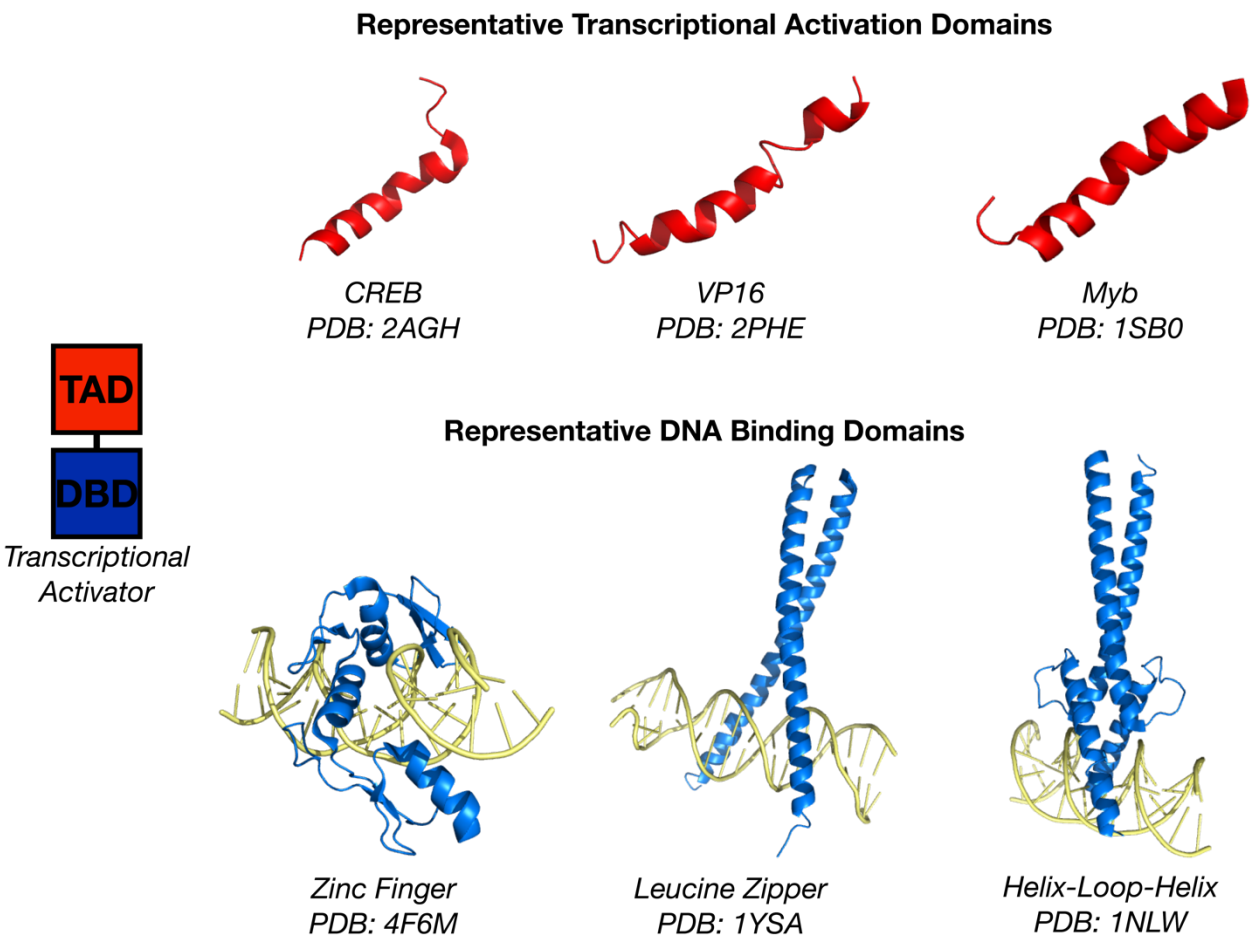


***Figure 1.1*- Transcriptional Activators are Minimally Composed of a Transcriptional Activation Domain and a DNA Binding Domain** A schematic representation of the modular nature of transcriptional activators is shown on the left. Several structures of representative TADs (red) and DBDs (blue) are shown on the right. The TAD structures are derived from solution NMR experiments of the indicated activators in complex with binding partners. The DBD structures are derived from x-ray crystal structures of the transcriptional regulators Kaiso (zinc finger), GCN4 (leucine zipper), and Mad/Max in complex with DNA.

DNA-Binding Domains

DNA binding domains recognize specific sequences on DNA in target genes upstream of the transcription start site. Several folds capable of binding DNA have been identified within DBDs such as helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix motifs.[4] In all of these motifs, the secondary structure of the domain positions the side chains of critical amino acids within the major groove of DNA, allowing for intermolecular contacts between the DBD and specific bases of DNA and the phosphate backbone. Generally, this proper spatial orientation is achieved through the use of alpha-helical secondary structure, though several examples of beta-turns being used for protein-

DNA interactions have been reported.[5] Furthermore, the majority of the specificity for a particular promoter sequence arises from hydrogen bonding with most protein-DNA interactions utilizing ten to twelve hydrogen bonds.[4] Thus, various transcription factors can attain exquisite selectivity for DNA recognition through slight differences in the secondary structure of interaction surfaces or the identity of specific amino acids in the proper spatial orientation for interaction with the nucleobases or phosphate backbone. Representative examples of the structures of zinc finger[6], leucine zipper[7], and helix-loop-helix[8] DBDs are shown in Figure 1.1 above.

Transcriptional Activation Domains

Following localization to the proper genomic sequence through the DBD, the TAD recruits coregulatory protein complexes or the transcriptional machinery through a network of protein-protein interactions. These domains have been historically classified as acidic, glutamine-rich, or proline-rich dependent upon the high prevalence of these amino acids within the TAD sequences.[9] The acidic activation domains, also called amphipathic activation domains, represent the largest and best-studied class and includes transcription factors such as the viral activator VP16[10,11], the tumor suppressor p53[12], and the yeast activator Gal4[13]. TADs tend to be intrinsically disordered, often existing as random coils in isolation and adopting α-helical secondary structures following contact with their binding partners as in the cases of the pKID domain of CREB binding to the KIX domain of CBP/p300[14], HIF-1α binding to the CH1 domain of CBP/p300[15], and Gcn4 binding to Med15[16]. Structures of several representative TADs, including the pKID domain of CREB[17], VP16[10], and Myb[18], in complex with coactivator binding partners are shown in Figure 1.1 above. In fact, kinetic and thermodynamic studies of three distinct TADs binding to Med15 suggest that activator•coactivator binding is biphasic, with an initial diffusion limited binding event followed by conformational changes in the complex.[16] Intrinsic disorder within TADs is hypothesized to be functionally advantageous in that the domains can bind to multiple coregulatory partners by adopting unique and complementary conformations, allowing for the possibility of transcriptional regulation through competition for necessary cofactors. Additionally, disorder allows the TAD to interact over a significantly broader surface area than is possible for preordered proteins,

providing selectivity for binding partners without the use of prototypical 'hotspot' regions.[19] Finally, many transcriptional activators, such as VP16[20], the glucocorticoid receptor[21], and the androgen receptor[22], contain more than one activation domain that is capable of independently functioning, allowing for recruitment of multiple coregulatory proteins or stronger interactions with their binding partners. Furthermore, evidence suggests that these distinct TADs may have unique functions in a gene dependent context. For example, DNA microarrays and RT-qPCR analysis demonstrated the TADs activation function 1 (AF1) and activation function 2 (AF2) of the transcriptional activator Glucocorticoid Receptor (GR) were independently required for different subsets of GR regulated genes based upon the sequence and structure of the glucocorticoid response element within those genes.[23]

1.4 Coregulatory Proteins and the Pre-Initiation Complex

Beyond sequence-specific activators that recognize and bind to DNA, transcriptional regulation requires additional proteins that modulate the transcription of cognate genes. Coactivators may physically bridge the DNA bound activator to the transcriptional machinery or modify chromatin such that transcription is more facile, whereas corepressors may block the TAD from contacting necessary cofactors or modify chromatin such that transcription is repressed. In the following section we discuss three classes of coregulatory proteins: (1) masking proteins, (2) the Mediator complex and other adaptor coactivators, and (3) chromatin modifying enzyme complexes, as well as the components of the RNAPII pre-initiation complex.

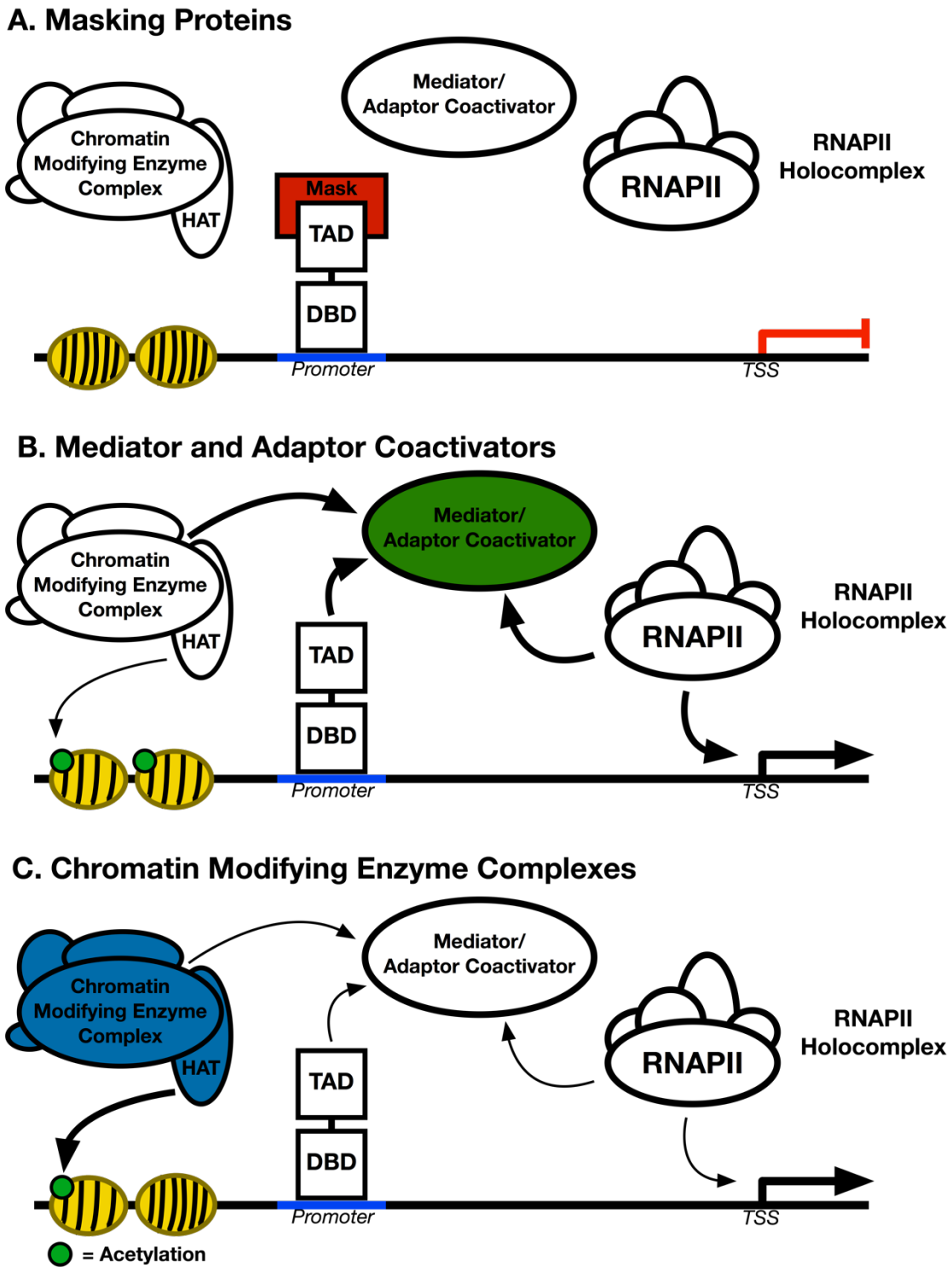


***Figure 1.2*- Examples of Coregulatory Protein Classes** (A) Masking proteins bind to the TAD of a transcriptional activator, preventing it from binding to other required elements of the transcriptional machinery. (B) Adaptor coactivators, including the Mediator Complex, act as scaffolding platforms that allow for the assembly of multiprotein complexes required for transcription. (C) Chromatin modifying enzyme complexes alter histone protein tails or nucleosome organization to activate or suppress transcription. Histone Acetyl Transferase activity is used as a representative example.

Masking Proteins

Perhaps the most intuitive class of coregulatory proteins are the corepressor masking proteins that recognize and bind to the TADs of transcriptional activators, thereby physically blocking the potential interactions between the TAD and the rest of the transcriptional machinery. An excellent example of this type of coregulator is the yeast protein Gal80, which interferes with the transcription of yeast *GAL* genes by masking the TAD of the activator Gal4. Upon exposure to galactose, Gal80 undergoes a conformational change that unmasks the Gal4 TAD resulting in the upregulation of the expression of *GAL* genes by the unmasked activator.[24] In humans, MDM2 forms a complex with the TAD of the transcriptional activator p53, which functions as a tumor suppressor and regulates cellular responses to cell damage. Similar to Gal4•Gal80, MDM2 blocks transcriptional activity of p53 by masking the TAD, but within normal cells dissociates following a phosphorylation event on MDM2 in response to DNA damage signals.[12] Additionally, MDM2 modulates the activity of p53 by altering the lifetime of the activator through its E3 ubiquitin ligase activity, targeting p53 for proteasomal degradation.[25]

Mediator and Adaptor Coactivators

Contrary to masking proteins, the primary function of which are to bind and sequester the TAD of activators from contacting other elements of the transcriptional machinery, a second class of coregulatory proteins function to link components of the transcriptional machinery to one another by directing the assembly of large multiprotein complexes.[26]

The Mediator complex is a megadalton assembly of twenty-six protein subunits in yeast with an additional five subunits found in higher order eukaryotes, the recruitment of which is required for the expression of virtually all protein coding genes.[27] This complex functions by making protein-protein interactions with DNA-bound transcriptional activators and components of the general transcriptional machinery, bridging the activators to the RNAPII holocomplex and enabling activated transcription. [28,29] Structural characterization of the full intact Mediator complex has proven to be quite challenging, given the flexibility, low abundance, and heterogeneity of the complex. However, most of

the subunits have been localized to approximate locations within the complex and comprise three modules termed the head, middle, and tail, as represented below.[30]

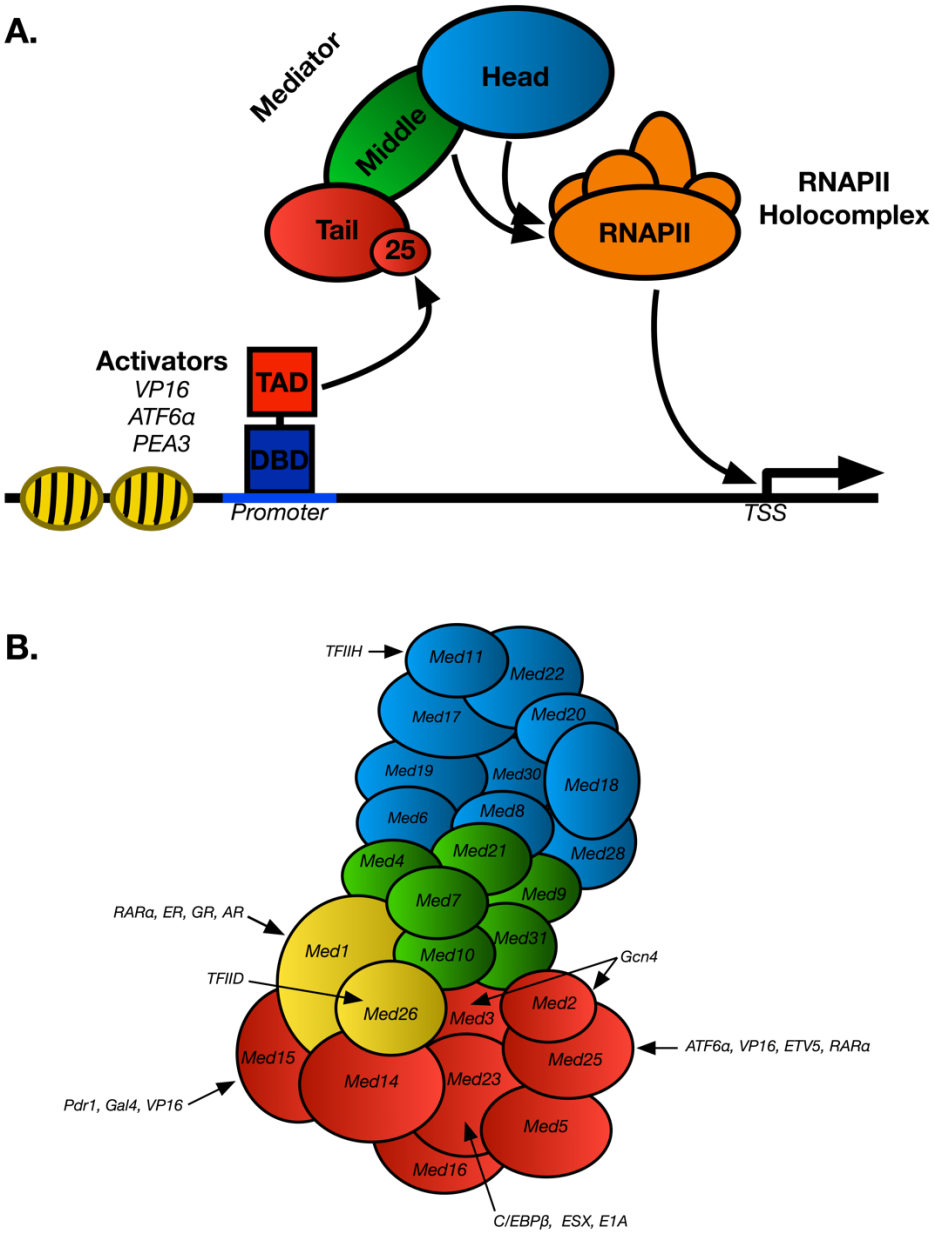


***Figure 1.3*- The Mediator Complex Bridges DNA-Bound Activators to the RNAPII PIC** (A) DNA-bound activators such as VP16, ATF6α, and the PEA3 subfamily contact the Mediator tail module at subunit 25, while the head and middle modules contact components of the RNAPII holoenzyme, aiding in the assembly of the PIC at the transcription start site. (B) General organization of the Mediator complex as reported by Malik et. al.[30] Head module components are shown in blue, middle module components in green, and tail module components in red. Med1 and Med26 (yellow) are not uniformly found in the complex are believed to lie at the junction of the tail and head modules. Protein binding partners for specific subunits are indicated using arrows.

The head and middle sections of the Mediator complex are responsible for contacting the general transcription factors of the PIC, while the highly modular tail contacts various DNA-bound transcriptional activators.[31,32] For example, a direct interaction has recently been demonstrated between Med26 and the RNAPII pre-initiation complex component TFIID.[33] Additionally, TFIIH has been reported to interact with Med11, further aiding in the assembly of the PIC at target gene promoters.[34,35] The specific interactions by which elements of the RNAPII holoenzyme interact with the head and middle modules of Mediator are still being elucidated, but the role of these modules in linking Mediator to RNAPII has been further validated by structural techniques such as electron microscopy that examine whole Mediator modules. Additionally, biochemical experiments such as activated transcription assays with Mediator subunit or general transcription factor depleted nuclear extracts have been used to identify key interactions between subunits required for function.[36-38] For example, a minimal functional core of the Mediator head module comprised of Med17, Med11, and Med22 has been reported to bind TBP with a near 1:1 stoichiometry, with transcriptional activity in yeast nuclear extracts directly correlated with the amount of TBP present in the extracts.[37] Similar experiments demonstrated that a recombinantly expressed Mediator head module-bound RNAPII subunits Rbp4 and Rbp7, further supporting the role of Mediator in stabilizing the assembly of the PIC at target gene promoters. [37,39] Additionally, specific protein-protein interactions have been identified between various activators and Mediator subunits such as nuclear receptors RARα[40,41], ERα[42], AR[43], and GR[44] contacting Med1; Gcn4 contacting Med2 and Med3[45]; Pdr1[46], VP16[47], and Gal4 interacting with Med15; C/EBPβ[48], ESX[49], E1a[50] interacting with Med23; and ATF6α[51], VP16[52], and ETV5[53] contacting Med25 physically linking the aforementioned activators to the rest of the Mediator complex and the general transcriptional machinery. Beyond physically linking activators to the RNAPII PIC, an additional function of the Mediator complex is that it can utilize chromatin looping in order to connect far upstream regulatory elements to the transcription start site.[32,54]

The master coactivator CBP/p300 is another excellent example of an adaptor coactivator. This large protein contains at least six domains (NRID, CH1, KIX, CH2, CH3, and IBiD) that have been validated as interaction partners for a variety of activators

including numerous nuclear receptors[55], c-Jun[56], c-Myb[18], CREB[14], p53[57], and HIF-1α[15]. For the KIX domain alone more than fifteen unique binding partners have been identified, underscoring the role of CBP/p300 as a 'master coactivator' capable of binding a plethora of transcriptional activators.[18,58-64] Beyond interacting with DNA-bound activators, CBP/p300 has been shown to bind other elements of the transcriptional machinery, allowing for the assembly of diverse multiprotein complexes. For example, elements of the RNAPII pre-initiation complex such as TBP, TFIIB, TFIIE, and TFIIF have been reported to bind to CBP/p300, aiding in the assembly of the PIC at gene promoters. [55,65] Beyond acting as a bridge between DNA-bound activators and the RNAPII PIC, CBP/p300 also functions as a scaffold that allows for the recruitment of other necessary coactivator proteins for DNA-bound activators, such as the interaction with p160 coactivators like SRC1 in androgen receptor (AR) dependent transcription.[66] Furthermore, evidence suggests that the N-terminal domain of CBP contacts the Mediator complex at subunit Med25[41], which may indirectly aid in the recruitment of Mediator to gene promoters without a direct contact to a DNA-bound activator. Perhaps the most critical function of CBP/p300, beyond its involvement in the assembly of multimeric protein complexes, is its enzymatic ability to acetylate chromatin at gene promoters through the activity of its histone acetyltransferase domain.[67] The importance of this function is discussed in greater detail in the following section.

Chromatin Modifying Enzyme Complexes

Beyond recruiting coregulators that block or recruit the PIC, many activators also recruit proteins or complexes that harbor enzymatic function and modify the local structure of chromatin. In the nucleus of eukaryotic organisms, DNA is wrapped around an octet of histone proteins, creating a protein-DNA complex referred to as a nucleosome that can be visualized as a 'bead on a string'.[68] Transcriptional activity is effectively modulated based upon how tightly DNA is wrapped around histones and how tightly packaged nucelosomes are in reference to one another. Post-translational modification of histone proteins, including lysine acetylation, phosphorylation, ubiquitylation, and sumoylation determine how tightly DNA is wrapped around the histone octet.[69] For example, histone acetyltransferases (HAT) such as CBP or the nuclear receptor

coactivator NCOA1 acetylate N-terminal tails of histone H3 and H4, which relieves electrostatic interactions between the negatively charged phosphate backbone of DNA and the basic lysine residues within the histone tail. [70,71] This results in a looser association between DNA and the histones and allows for greater accessibility of cis-regulatory elements and upregulation of gene expression. In converse, histone deacetylases (HDAC) such as HDAC1, HDAC2, and HDAC3 can remove acetyl modifications from the histone tails, thereby promoting a tighter interaction between DNA and the histones resulting in downregulation of gene expression.[72] Thus, enzymatic modification of epigenetic marks on histone proteins confers an additional mechanism by which transcription is tightly regulated. As an aside, evidence also suggests that HATs and HDACs can also act upon non-histone substrates, such as the androgen receptor and p53, in order to modify protein-protein interaction networks or protein half-life. [73-75] Methylation of lysine residues within the histone tail also affects transcriptional throughput, though the functional consequence is not as intuitive as differences in acetylation state and depends upon the number of methyl modifications and the location of these modifications within the histone tail. [76,77] For example, H3K27 mono-methylation activates transcription while tri-methylation of the same residue results in repression of transcription.[78] In addition to recruiting epigenetic modifiers that alter histone marks, many activators also recruit nucleosome remodeling complexes that alter the physical packaging of nucleosomes into chromatin, such as the SWI/SNF and Mi2/NuRD families of protein complexes to positively or negatively affect the transcription of target genes.[79] A table containing examples of several histone modifications, the enzyme responsible, and their functional consequences on transcription is presented below.

Table 1.1- Examples of Histone Modifications, Their Effectors, and Functional Consequences

Modification	Site	Enzyme	Functional Consequence
Acetylation	H3K4	Esa1	Transcriptional Activation
	H3K9	Gcn5, SRC-1	Transcriptional Activation
	H3K14	Gcn5, Esa1, TIP60, SRC-1, p300, CBP	Transcriptional Activation, DNA Repair
	H3K23	Gcn5, Sas3, p300, CBP	Transcriptional Activation, DNA Repair
	H3K27	Gcn5	Transcriptional Activation
	H4K5	Esa1, TIP60, p300	Transcriptional Activation, DNA Repair
	H4K8	Gcn5, PCAF, Esa1, TIP60, p300	Transcriptional Activation, DNA Repair
	H4K12	Esa1, TIP60, p300	Transcriptional Activation, DNA Repair
Methylation	H3K4	Set9, MLL	Transcriptional Activation, Transcriptional Elongation
	H3K9	Suv39h, Clr4	Transcriptional Repression (3-Me)
	H3K27	Ezh2	Transcriptional Activation (1-Me), Transcriptional Repression (3-Me)
	H4K20	Suv4-20h	Transcriptional Repression (3-Me)
Phosphorylation	H2AS1	MSK1	Transcriptional Repression
	H2AS139	DNA-PK	DNA Repair
	H2BS14	Mst1	Apoptosis
	H2BS33	TAF1	Transcriptional Activation
	H3S10	Snf1	Transcriptional Activation

The RNA Polymerase II Pre-Initiation Complex

Following the recruitment of the necessary coactivator proteins, the final step in transcription is the recruitment of the enzymatic complex responsible for the actual reading of the gene and synthesis of the mRNA. This has been termed the pre-initiation complex and it is composed of RNAPII, itself a multimeric complex of twelve subunits, and a variety of accessory proteins and protein complexes termed the general transcription factors (GTF). Specifically, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH are the essential GTFs and carry out a variety of functions including stabilization of RNAPII at the transcription start site, ATP-dependent helicase activity responsible for unwinding the DNA double helix to form the transcription bubble, enzymes responsible for nucleotide excision repair to replace mismatched or mutated nucleotides, and enabling promoter clearance by RNAPII.[80] The table below defines the function of each the required GTFs.[81]

Table 1.2- General Transcription Factors of the RNAPII Pre-Initiation Complex

	Protein Composition	Function
TFIIA	p39, p19, p12	Stabilizes TATA-TBP interaction, *Not essential for PIC assembly at all promoters*
TFIIB	p33	Start site selection, Stabilizes TATA-TBP interaction, Recruits RNAPII and TFIIF
TFIID	TBP, TAFs* 1-14	Core promoter (TATA) binding factor, Kinase, Ubiquitin activation and conjugation, Histone acetyltransferase activity
TFIIE	p56, p34	Recruits TFIIH, Promoter clearance by RNAPII
TFIIF	RAP30, RAP74	Binds and recruits RNAPII to promoter, Recruits TFIIE and TFIIH, Assists TFIIB and RNAPII in start site selection, Assists RNAPII promoter clearance, Promotes RNAPII-mediated transcript elongation
TFIIH	p89, p80, p62, p52, p44, p40/CDK7, p38/Cyclin H, p34, p32/MAT1, p85/TFB5	ATPase activity for transcription initiation and promoter clearance, Helicase activity to open promoter assisting PIC assembly, Nucleotide excision repair capability, Kinase for RNAPII C-terminal domain phosphorylation, E3 ubiquitin ligase activity

* = TBP Associated Factor

Thus, in conclusion, transcriptional activation requires a veritable host of proteins in order to ensure the proper regulation of gene expression. These proteins interact directly to form large multimeric complexes at the promoters of target genes in order to stabilize the RNAPII pre-initiation complex at the transcriptional start site. Notably, these

complexes are highly dynamic, owing to the large surface areas and weak affinities with which many of these interactions occur, and modular, with different cellular or promoter contexts requiring unique components. [19,82] The dynamics and modular nature of these complexes are discussed in more depth in Section 1.6 and Section 1.7. As a result, these interactions each represent potential points of intervention for small molecule inhibitors in the modulation of transcriptional activity. For example, targeting activator•masking protein interactions could be used to reactivate an aberrantly suppressed transcriptional activator, as in the case of the p53•MDM2 interaction.[83] Conversely, targeting specific activator•coactivator interactions for inhibition could be used to suppress the activity of an overexpressed transcriptional activator.[84] Targeting chromatin modifying complexes is an area of very active research at present and has seen promising results as inhibitors advance to the clinic and enter the marketplace, such as Vorinostat which has been approved for treatment of cutaneous T cell lymphoma. [85-87] In the next section the functional consequences and mechanisms of transcriptional dysregulation are outlined.

1.5 Dysregulation of Transcription Contributes to Disease

Given the clear importance of transcription in maintaining normal cellular function, it is unsurprising that the dysregulation of this process can have dire consequences on human health. Perhaps the most prominent example of this phenomenon is the variety of pathogenic mechanisms by which dysregulated transcription leads to cancer, including abnormal expression of transcription factors, the formation of fusion genes through chromosomal translocations, and mutations within transcription factors that alter their function.

Abnormal Expression of Transcription Factors

Abnormal changes in the expression profiles of proteins responsible for regulating transcription is a common mechanism by which human malignancies occur. For example, the c-Myc oncogene is amplified in lung[88], breast[89], and colon cancers[90] resulting in overexpression of c-Myc. In turn, an overabundance of c-Myc results in overexpression of genes implicated in cell cycle regulation, such as the cyclins and cyclin dependent kinases, leading to loss of cell cycle control. Other genomic targets for c-Myc activation

include regulators of metabolic processes such as glycolysis, and genes that have been linked to immortalization in malignant tissues.[91] Importantly, loss of c-Myc function has been recently demonstrated to result in tumor regression, underscoring the importance of this activator as a therapeutically relevant target. [92,93] As additional examples, c-Myb overexpression has been linked to lymphocyte transformation to a leukemic state and overexpression of the androgen receptor is required for transformation in prostatic malignancies and sensitizes cells to low levels of circulating androgens thereby circumventing the most common therapeutic strategy in treatment of PCa. [94,95] Overexpression of the PEA3 subfamily of Ets transcriptional activators as a result of aberrant PI3K/RAS signaling cascades has been linked to tumorigenesis and metastatic processes in breast and prostatic carcinomas.[96]

Conversely, underexpression of transcriptional activators responsible for regulating tumor suppressor transcriptional programs can also lead to malignancies. For example, the retinoblastoma protein (RB), a tumor suppressor whose function is to regulate cell cycle progression, is deleted or mutated in 90% of lung cancers, resulting in loss of critical cell cycle progression checkpoints ultimately leading to unrestricted cell growth. [97,98] Similar loss of normal RB expression has also been demonstrated to result in retinoblastoma and osteosarcomas in children.[99] Similarly, loss of normal KLF6 expression in prostatic tissues has been linked to diminished p21 expression and a subsequent loss of cell cycle control.[100] As a final example, overexpression of the p53 corepressor MDM2 results in a decrease in cellular p53 concentrations as a consequence of proteasomal degradation triggered by MDM2 dependent ubiquitylation, resulting in loss of p53 dependent tumor suppressor gene programs.[101]

Chromosomal Translocations

One mechanism by which transcription becomes dysregulated is through chromosomal abnormalities that arise following translocation events in which a coding region of a particular gene becomes associated with regulatory elements or coding regions of another gene. In the case of the former proteins under tight transcriptional regulation may become fused to regulatory elements from a far more actively transcribed gene while the latter may result in fusion proteins with aberrant function. For example, all

cases of Burkitt's Lymphoma require a chromosomal translocation of c-Myc to the regulatory regions that control expression of the immunoglobulin heavy or light chains in lymphocytes. Thus, c-Myc is upregulated far above normal levels resulting in transformation of healthy lymphocytes to the lymphoma phenotype as a result of excessive activity in c-Myc transcriptional programs.[102] A similar example of this phenomenon is the TMPRSS2:ERG gene fusion found in a majority of aggressive prostate cancer phenotypes, which results in the expression of a truncated form of ERG that drives the activation of invasion and metastasis related gene programs.[103] In the case of alveolar rhabdomyosarcoma, a pediatric cancer of soft tissues that can occur in any anatomical location, chromosomal translocations result in the fusion of the PAX3 or PAX7 DBD to the TAD of FKHR, a member of the forkhead transcription factor family. The PAX3/PAX7 TAD is negatively regulated through interactions with the N-terminus of the activator, but the chimeric protein containing the FKHR TAD lacks this regulatory intramolecular interaction, resulting in abnormally high expression of PAX3/PAX7 gene programs ultimately leading to a malignant phenotype.[104]

Mutations in transcription factors

An additional mechanism by which transcriptional activation may become dysregulated is through mutations within transcriptional activators that alter their protein interaction networks or function. As a prototypical example, p53 is the most commonly mutated gene in human cancers, with more than fifty percent of all malignancies harboring mutations in this activator.[105] As a result of the clinical implications associated with these mutations, p53 is particularly well studied and provides excellent examples of several mechanisms by which activator mutations generally lead to loss of transcriptional regulation. Mutations within p53 have been linked to four distinct mechanisms by which malignant phenotypes are promoted, as shown in Figure 1.4 below. In the first, mutations within the p53 DBD impairs DNA recognition of normal p53-target gene regulatory elements while simultaneously enabling recognition of non-native promoters. As a result, wild type p53 target genes responsible for tumor suppressor activity are downregulated, while oncogenic proteins may be upregulated by the intact p53 TAD as a consequence of non-native DNA recognition. [106,107] Attempts at identifying specific DNA recognition

sequences for p53 mutations in the DNA-binding domain are ongoing, but have yet to define a consensus sequence. In fact, these types of mutations initially led to the classification of p53 as an oncogene as opposed to a tumor suppressor.

A second mechanism by which p53 mutation results in malignant transformation is the development of new protein-protein interactions between p53 and other transcriptional activators, resulting in enhanced transcriptional throughput, as in the case of mutant p53•nuclear factor Y (NF-Y) interactions, which may cause loss of cell cycle control following DNA damage events.[108] More specifically, these interactions recruit the potent p53 TAD to NFY regulated genes through p53•NFY•DNA interactions with the functional consequence being overexpression of NFY regulated genes as the p53 TAD upregulates transcription more efficiently than the NFY TAD.

A closely related, but functionally distinct mechanism involves p53 mutations that allow for non-native protein-protein interactions with other transcription factors that mitigate their normal function by interfering with DNA binding or the recruitment of other transcriptional coregulators, as in the case of mutant p53 interfering with the normal function of p63 and p73.[109]

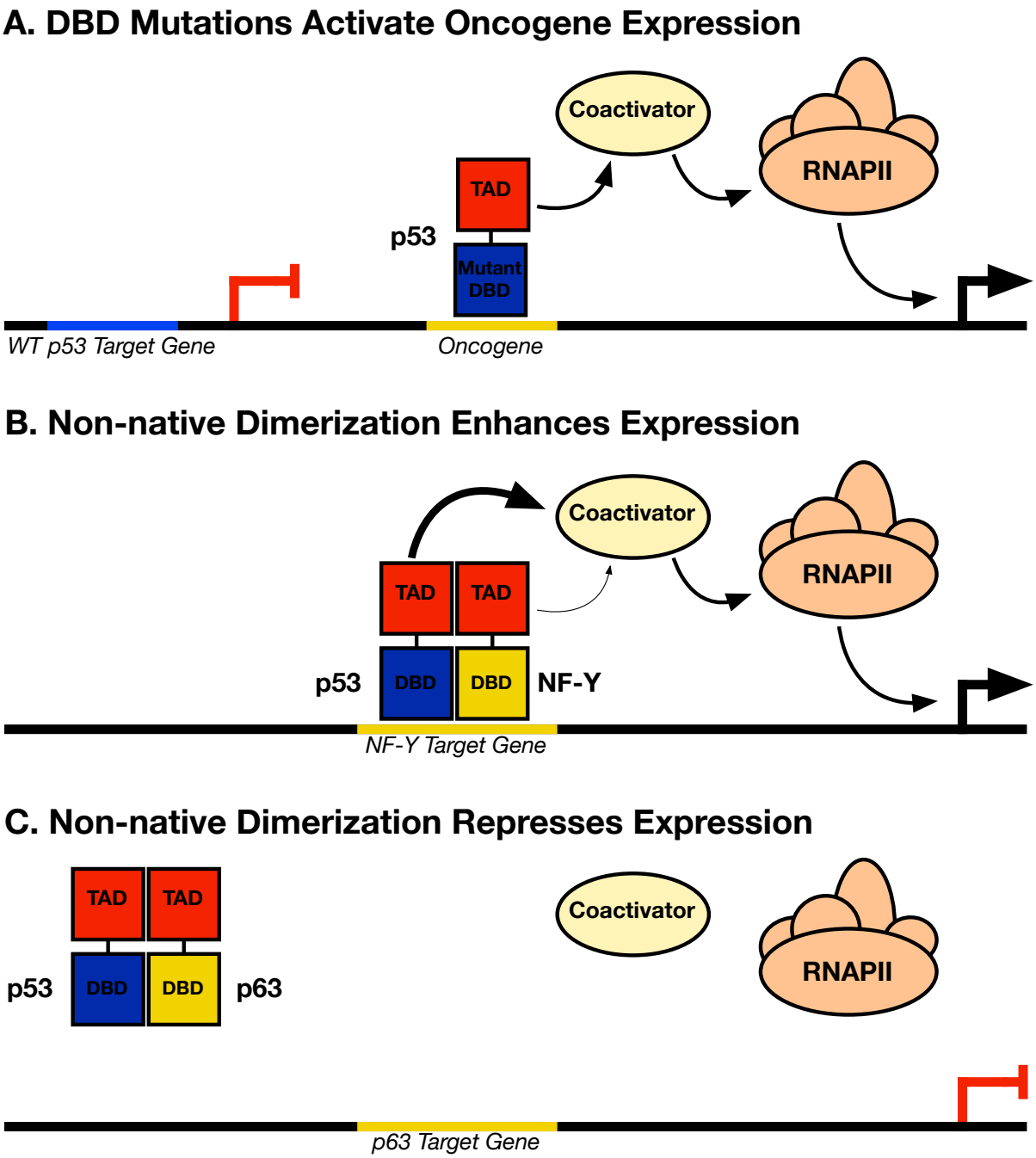


***Figure 1.4*- Functional Consequences of p53 Activator Mutations** (A) Mutations within the DBD of p53 prevent it from activating expression of native tumor suppressor target genes and instead activates expression of oncogenes through non-native DBD•DNA interactions. (B) Mutations may cause p53 to dimerize with non-native partners such as NF-Y, localinzing the dimer to NF-Y target genes. The p53 TAD is more efficient at activating transcription than the TAD of NF-Y, leading to upregulated expression of NF-Y target genes. (C) Mutations may cause p53 to dimerize with non-native partners such as p63, preventing the localization of p63 to target genes.

In addition to interfering with normal transcriptional activity, mutant p53 can also bind to components of other important, non-transcriptionally related, cellular pathways such as

interfering with the MRE11-RAD50-NSB1 complex, resulting in diminished DNA-repair through homolgous recombination.[110] These specific mechanisms are not all-inclusive, but instead illustrate the variety of pathways by which mutations in transcriptional regulators may result in malignancies.

Notably, in all three of the above common pathogenic mechanisms (overexpression of activators, chromosomal translocations, and activator mutations) small molecules that modulate aberrant transcriptional activator function would be useful mechanistic probes and promising therapeutic leads. Despite their potential, transcription factor–targeting molecules have remained elusive; in more conventional terms, transcription factors have largely been classified as undruggable. The origin of this description becomes clear as one examines the possible avenues for altering transcription factor activity. Except for nuclear receptors, transcription factors do not have native small-molecule ligands; thus, the primary options for their alteration involve manipulation of the complex network of protein-protein and protein–nucleic acid interactions by which transcription factors function.

1.6 Features of Protein-Protein Interaction Networks in Transcription

As p53 is an exceptionally well studied transcriptional activator, the PPI network required for its regulation has been thoroughly characterized and provides insight to the wide range of affinities and surface areas of the complexes required for transcriptional regulation *in general*, as shown in Figure 1.5-A below. The complexes formed between p53 and its regulatory partners such as MDM2 are typically high affinity with a relatively small surface area, and their interaction energy largely resides in a small number of residues. These are structurally well-organized interfaces and are highly amenable to structural characterization.[12] As with many activators, p53 typically functions as a multimer, and the homo-oligomerization interface is another high-affinity, well-organized interface, although it takes place over a considerably larger binding surface.[111] The final group of PPIs that activators such as p53 utilize are those with coactivator proteins, such as CBP/p300, typically involved as part of the assembly of the transcriptional machinery in the early stages of transcription as discussed above.[60] These interactions are the least

well characterized, likely because the binding partners are conformationally dynamic, inhibiting high-resolution structural studies, and many are of only moderate affinity.[112]

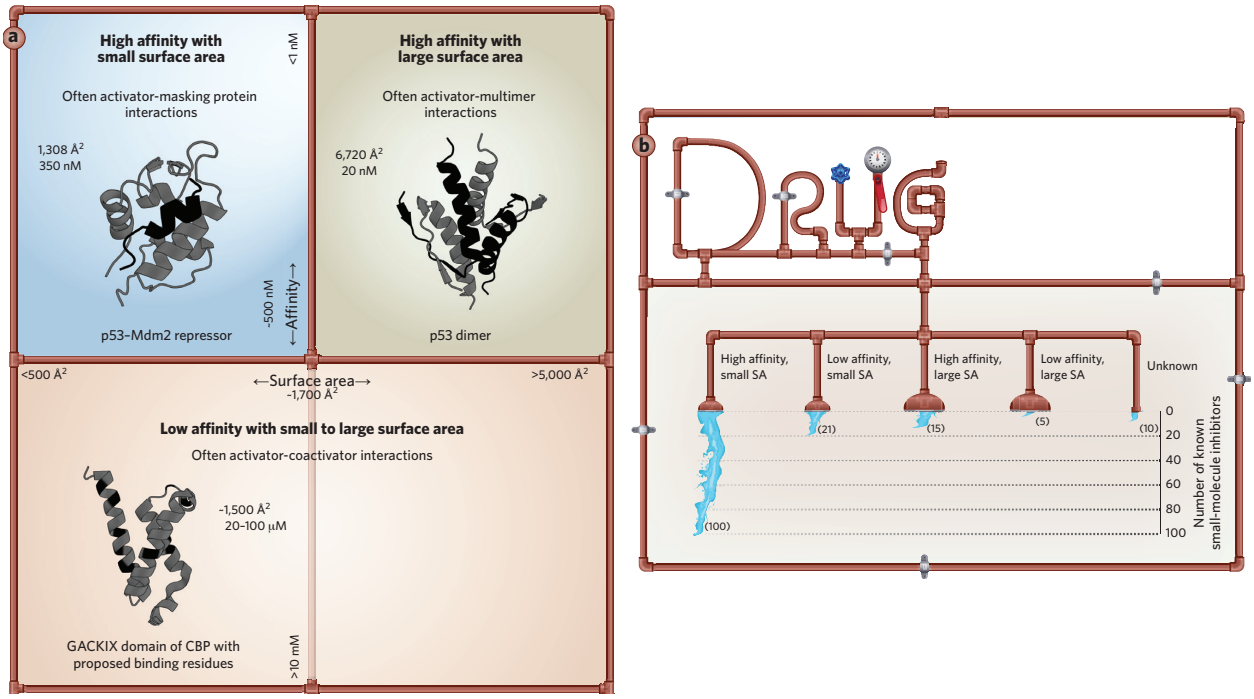


***Figure 1.5*- The Chemical Space of Protein-Protein Interactions** (A) Protein-protein interactions can be classified based upon the surface area over which they occur (x-axis) and the affinity of the complex (y-axis). PDB Codes: p53•MDM2- 1YCR, p53 Dimer- 1PET, KIX- 2LQH. (B) The development of small molecule PPI inhibitors have been spurred by advances in inhibitor design and screening methodologies. Success in targeting high affinity, low surface area interactions has far outpaced discovery of inhibitors for broader or weaker interactions. *82,113*

Building on this example, the interaction between p53 and MDM2 serves as a useful case study for the successful targeting of transcriptionally relevant PPIs. This is a high-affinity complex that masks the p53 activation domain, preventing p53 from functioning as a transcriptional activator and regulating its lifetime through its ubiquitylation state.[101] It is also highly ordered, with a focused interaction surface area of <1,800 A^2 that has been amenable to structural characterization and has characteristics similar to those of receptor-ligand interactions. Specifically, a crystal structure of the N-terminal TAD of p53 in complex with MDM2 was solved and suggested that the interaction is mediated primarily by only three residues (F19, W23, and L26). [12,114] The importance of these residues was further supported by functional data wherein transcriptional activation by p53 requires that these residues be intact.[115] As a result, p53•MDM2 and the closely related p53•MDMx complex have been readily targeted with several unique

small molecules.[105] These scaffolds include the benzodiazepinediones[116], terphenyls[117], and chalcones[118], and were derived as a result of information gleaned from the development of peptidic inhibitors and the first synthetic inhibitor nutlin.[83] Perhaps the best example of p53-MDM2 inhibitors to date are the spirooxindole derived molecules that have single-digit nanomolar affinity for MDM2, maintain bioavailability, and have a good pharmacodynamic profile leading to Phase I clinical trials.[119] Thus, transcription factor PPIs characterized by high affinity and a small interaction surface area are very targetable, largely owing to the advances in PPI inhibitor discovery strategies over the past decade. This reflects the overall success in targeting high-affinity, small-surface-area PPIs in many functional contexts, as these types of interactions most closely mirror enzyme-ligand interactions (Figure 1.5-B).

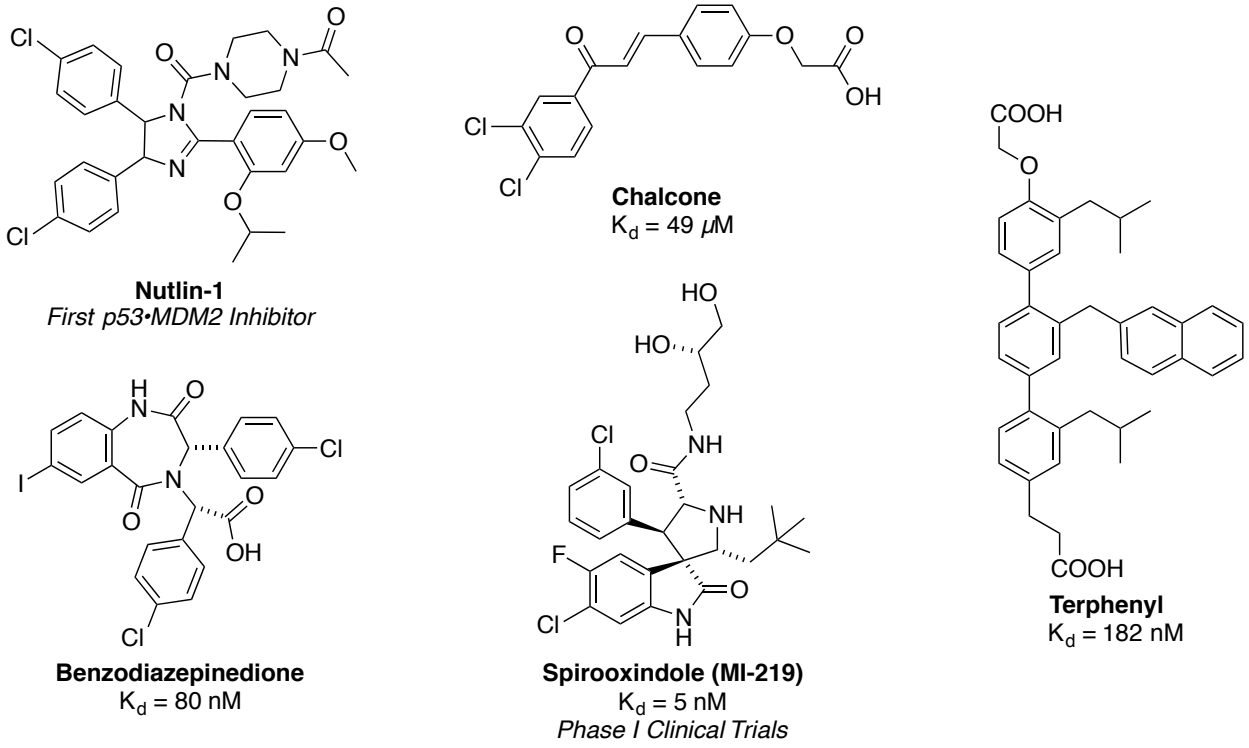


***Figure 1.6*- Representative Examples of p53•MDM2 Inhibitor Scaffolds** Given the low surface area and high affinity of the p53•MDM2 interaction a number of effective small molecule scaffolds have been developed to inhibit this interaction. Representative examples of each class are shown with their affinities for MDM2.

Perhaps not surprisingly, PPI complexes that utilize broader interaction surfaces and occur with weaker affinity have been targeted far less successfully, as shown in Figure 1.5-B. [82,113] As mentioned above, these broad and weak interactions lead to distinct

networks of PPIs being used to assemble the transcriptional machinery. The transient and conformationally dynamic complexes are formed with various coactivators and coactivator complexes. Traditional probe discovery or design methods are ill equipped to target these complexes because one or both binding partners are classified as intrinsically disordered proteins, and the complexes often form transiently in the cell, with affinities one to two orders of magnitude weaker than p53-MDM2–type interactions.[120] Further, the interaction surface is often considerably larger, with interaction energies shared over a greater number of amino acids, defying hotspot analysis.

1.7 Targeting Activator•Coregulator Interactions

Despite the inherent challenges in targeting the interactions between transcriptional activators and the coregulatory machinery, doing so offers an opportunity to achieve context specificity in perturbing transcriptional processes.

Exploiting compositional dynamics in transcriptional complexes

Many of the complexes involved in transcription are dynamic in composition and comprise a core enzymatic function flanked by scaffolding proteins and exchangeable modules. The BAF-type chromatin remodeling complexes that play a key role in transcription initiation are an excellent example. Most share a core enzymatic subunit (the ATPase Brg), but there are exchangeable subunits that vary according to tissue.[121] As an example, subunits within the BAF complex are exchanged as embryonic stem cells differentiate into neural progenitors and again as neural progenitors differentiate into neurons. Specifically, the neural progenitor BAF (npBAF) complex contains the BAF45a and BAF53a subunits and is responsible for maintaining self-renewal potential through enhanced NOTCH signaling by binding to the NICD•RBPj complex thereby stabilizing the interaction with DNA.[122] Notably, loss of BAF45a or BAF53a in neural progenitors resulted in a loss of proliferative capacity, while overexpression of BAF45a enhanced proliferative capacity. Following differentiation into neurons, BAF45a and BAF53a are exchanged for BAF45b and BAF53b, respectively, to form the neuronal BAF (nBAF) complex, which assists in neuron-specific functions such as dendritic outgrowth.[123]

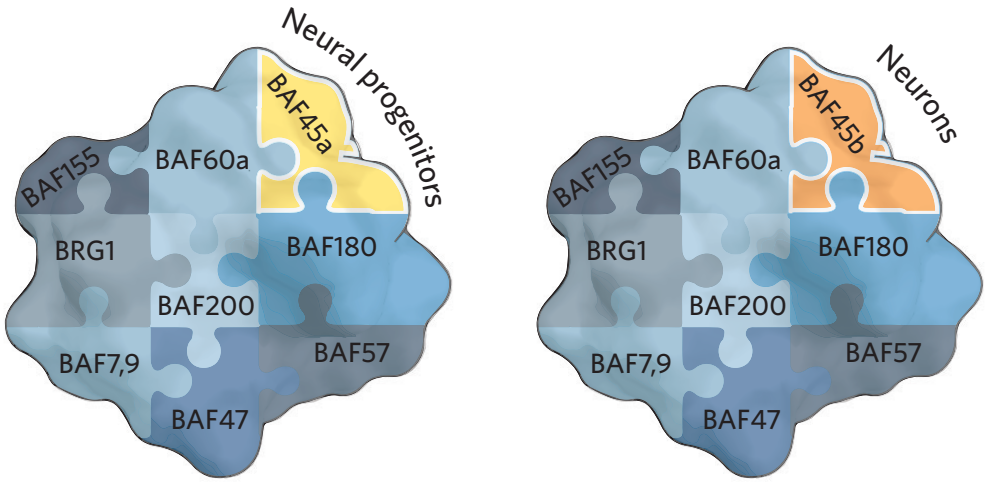


***Figure 1.7*- Transcriptional Complexes are Dynamic in Composition** The BAF chromatin remodeling complex is a representative example of the dynamic composition of transcriptional complexes. The core enzymatic subunit BRG1 is found in all BAF complexes, but exchangeable subunits such as BAF45a or BAF45b are found in the complexes in a tissue-dependent manner.

Evidence has emerged that transcription factors target both the enzymatic component and exchangeable modules in such complexes as part of the assembly of the transcriptional machinery. In yeast, the application of covalent chemical cross-linking to map contacts with the BAF-type complex Swi/Snf revealed that the transcription factors Gal4 and VP16 each contact the core enzymatic functionality (Snf1) in addition to auxiliary factor Snf5.[124] Thus, one could imagine blocking a PPI that would affect the localization of a complex at particular gene promoters by targeting tissue or gene specific components of the complex, but would leave the core enzyme unaltered. As a result, the core complex would remain active at other promoters or within other tissues that utilize a complex with a distinct selection of components.

A second group of context-dependent transcription factor–coactivator PPIs are those used by viruses to hijack the transcriptional apparatus in infected tissues. Arora and Pan recently demonstrated that an inhibitor of a complex formed between the human papillomavirus transcription factor E6 and p300 restores the ability of p53 to function in human papillomavirus–positive head and neck cancers and, in doing so, blocks tumorigenicity. [125]

Thus, an ongoing challenge is to provide a more comprehensive map of the network of transcription factor PPIs at gene promoters as there remain more questions than answers for most transcription factors. In doing so, unique complex compositions

may be identified in a promoter or tissue specific manner, widening the avenues for target identification. For example, even in the case of the well-studied transcription factor p53, it is not clear which of its coactivator complexes are the most critical to block, either alone or in combination; considering its interaction with the master coactivator CBP/p300 alone, p53 binds *in vitro* with four of the activator interaction motifs within CBP/p300, but the functional relevance of each of those interactions is not yet defined.[126] Identifying critical contact points within CPB/p300 for the N-terminal TAD and developing small molecule inhibitors of those domains could provide a useful mechanism for blocking the oncogenic activity of mutant p53, as discussed above. Thus, this is an area in which chemical biology tools such as covalent chemical capture or high-quality small-molecule probes will be invaluable.

Exploiting conformational plasticity of coactivators

In addition to compositional dynamics within transcription factor complexes, many of the individual subunits exhibit significant conformational plasticity as a means to interact specifically with a variety of transcriptional activators. The low energy barriers between individual conformations mean that each participant can use the same group of amino acids to recognize a variety of binding partners, with each complex requiring a distinct conformation. As such, synthetic regulation of this conformational plasticity to direct complex assembly represents a second exciting opportunity to specifically target transcription pathways.

A foundational example of this phenomenon is the GACKIX (Gal11, Arc105, CBP/p300, kinase-inducible domain interacting) domain of the master coactivator and histone acetyltransferase CBP/p300. GACKIX is highly plastic, and its two transcription factor–binding surfaces can accommodate more than 15 distinct transcription factor binding partners, in the context of binary or ternary complexes, that are involved in a variety of physiological processes and implicated in diseases from cancer to neuropathic pain, as shown in Figure 1.8-A.[64] Furthermore, the two binding surfaces are allosterically connected, allowing for cooperative binding of particular transcription factor–GACKIX pairs with a wide range of enhancement. For example, a pre-formed complex of the KIX domain and the lysine methyltransferase MLL, a coactivating histone modifying enzyme,

resulted in the enhanced affinity of the transcriptional activators c-Myb and the pKID domain of CREB by 2.5-fold and 2-fold, respectively.[127] Experimental and computational studies indicate that the mechanistic origin of the variable cooperativity is differential stabilization of the ternary complex. [128,129] Thus by intercepting a particular conformer, a small molecule could have either a positive or a negative influence on the binary and ternary transcription factor–coactivator assembly process. Certainly the field can be guided by the success seen in targeting particular conformational states of kinases or of the protein folding machinery for enhanced selectivity and context-specific effects on downstream processes. [130,131]

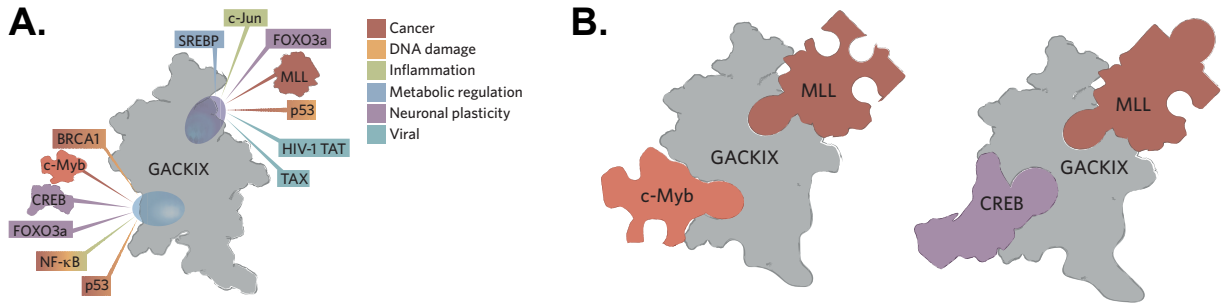


***Figure 1.8*- Transcriptional Complexes are Dynamic in Conformation** (A) The KIX domain of CBP/p300 utilizes two binding sites to interact with greater than fifteen distinct transcriptional activators. A representative sample of these activators and their biological implications is shown. (B) Ternary complexes formed between KIX and two transcription factors requires unique conformational signatures as a result of allosteric communication between the two binding sites.

Screening techniques that directly address the conformational plasticity of coactivators have been effective for identifying small-molecule modulators that capture distinct conformers. The site-directed fragment screening strategy of tethering first developed at Sunesis is one such strategy. When applied to the conformationally dynamic GACKIX motif, for example, researchers identified chemical cochaperones that stabilize a range of GACKIX conformations and dictate the formation of particular GACKIX–transcription factor assemblies either positively or negatively. [129,132] Specifically, the disulfide containing fragment **1-10** has been demonstrated to elicit negative cooperativity of pKID binding when it is tethered to a mutated cysteine (L664C) within the MLL interaction surface. Interestingly, when they same fragment is tethered to a different mutated cysteine within the same interface (N627C), pKID affinity is actually enhanced, suggesting that 1-10 is stabilizing unique conformations of KIX depending upon it's

location. Furthermore, covalent modification of KIX L664C with the 1-10 disulfide resulted in considerably enhanced stability, allowing for crystallization of this highly plastic domain and the first ever reported crystal structure, further underscoring the utility of Tethering as a strategy for identifying small molecule modulators of activator•coactivator interactions.

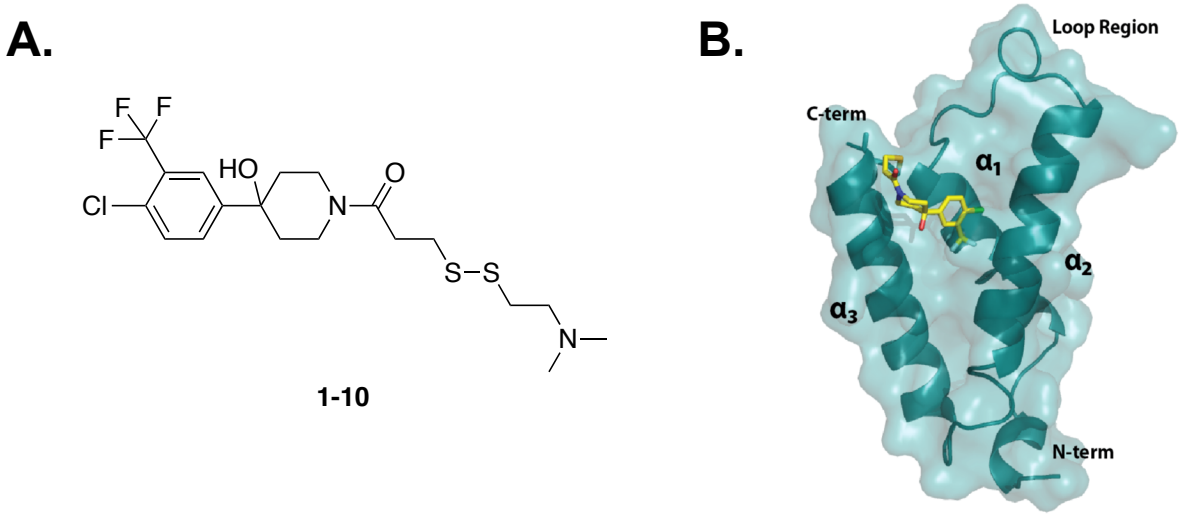


***Figure 1.9*- The Chemical Cochaperone 1-10 Binds to and Stabilizes KIX** (A) Structure of the disulfide containing 1-10 chemical cochaperone identified through a tethering screen of the KIX domain containing engineered cysteine residues. (B) Crystal structure of the 1-10 bound stabilized KIX conformation. PDB ID: 4I90

Standard binding screens can also be adapted to discover modulators that capture unique coactivator conformations through the triaging of primary screening hits that mimic native binding partners. Using this process, the natural products sekikaic acid and the related lobaric acid capture a conformation of the GACKIX motif that showed greatly attenuated ternary complex formation.[133] More explicitly, these natural products exhibited a mixed mode of action in which inhibition was achieved both orthosterically at the MLL interaction surface of KIX, against which natural products were screened, as well as allosterically at the pKID interaction surface.

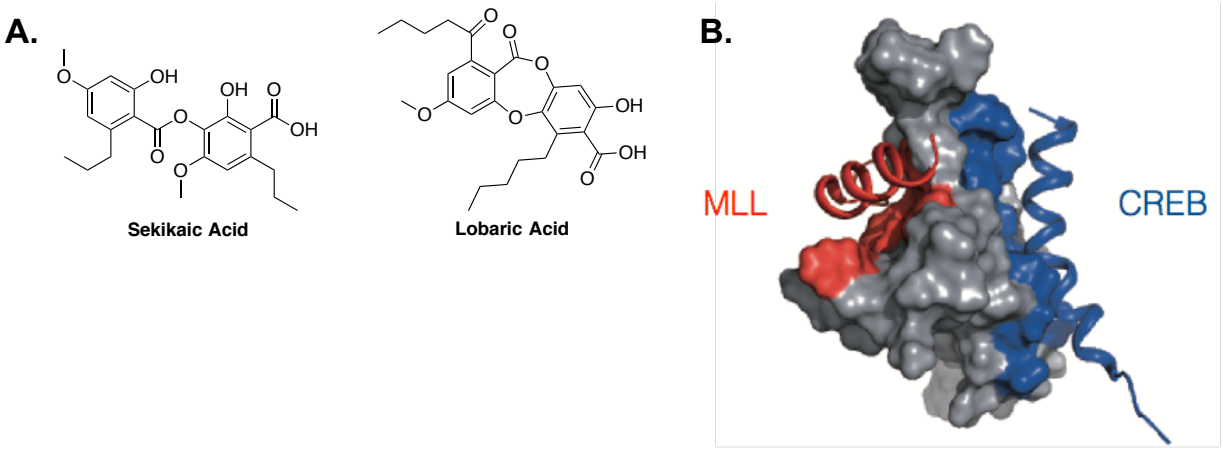


***Figure 1.10*- Sekikaic Acid and Lobaric Acid are Mixed Orthosteric/Allosteric Inhibitors of the KIX Domain** (A) Structures of the depside sekikaic acid and depsidone lobaric acid. (B) These natural product small molecule inhibitors of the KIX domain bind at the MLL site, inhibiting it orthosterically, while also inducing conformational changes within the protein that perturb binding events at the distal CREB site. PDB ID: 2AGH

One advantage from this type of strategy is that small-molecule modulators are likely to exhibit enhanced specificity because they are targeting unique coactivator conformations; this is true of sekikaic acid and lobaric acid. A second advantage is that it presents a generalizable mechanism by which allosteric modulators of highly challenging binding interfaces can be discovered. Again using GACKIX as an example, the binding site used by the oncogene c-Myb is shallow and lacks significant topology. However, the second binding site within GACKIX, which is targeted by sekikaic acid and lobaric acid, is smaller and better defined. By targeting a more druggable binding surface, we hypothesize that allosteric networks within the coactivator can be exploited to affect less amenable distal binding sites.

1.8 Emergent Technologies and Alternative Approaches

Recent advances in chemical biology, has led to enormous progress in targeting protein-protein interactions, and this has fueled successes in developing small-molecule modulators of a key, readily targetable subset of transcriptional activator PPIs. The recognition that the dynamic composition and structure of transcription factor–coactivator complexes offers advantages in terms of specificity and function in recent years now opens the door for similar success in targeting significantly more challenging interactions. As molecules move toward clinical applications, a question that has yet to be answered

is one of potency. At least *in vitro*, transcription factor–coactivator complexes are an order of magnitude or more weaker than complexes such as p53–MDM2. Whether it is possible to obtain synthetic modulators that exceed the affinity of the native ligands is still an open question. This is an area in which allosteric modulators may offer significant advantages. NMR screening techniques that rely on conformational dynamics, such as protein-observed fluorine (PrOF) NMR spectroscopy, have proven powerful for focusing on molecules that capture coactivators in a particular conformational space and will be a key discovery tool.[134] In addition, screening formats such as small-molecule microarrays that facilitate the interrogation of transcription factor binding in the presence of endogenous competing factors enable specificity to be assessed much earlier in the discovery process. [75,135]

While we have chosen to focus primarily upon protein-protein interactions between transcriptional activators and other elements of the transcriptional machinery, there is still significant untapped potential in the broader network of transcription factor interactions as the field continues to define these critical connections. Approaches such as targeting transcription factor dimerization motifs or preventing promoter localization by blocking DNA binding of transcription factors are additional strategies being explored that may also allow for the successful modulation of transcriptional processes for mechanistic insight or therapeutic gain.

1.9 Thesis Summary

The overall goal of the research presented in this thesis is to identify small molecule inhibitors of a transcriptional coactivator, Med25, by targeting the Activator Interaction Domain (AcID) within this protein. AcID is a critical point of contact for a number of transcriptional activators with diverse biological functions, but is poorly understood at present. Thus, we hypothesize that small molecule inhibitors of the domain will be particularly useful as mechanistic probes in the elucidation of the role of Med25 in the regulation of transcriptional programs that affect cellular processes such as endoplasmic reticulum stress response, viral infection, and the development of metastatic phenotypes in a number of cancers.

Towards this goal, we first explore in Chapter 2 the molecular underpinnings that define interactions between transcriptional activators and the activator interaction domain of Med25, with the goal of identifying mechanisms by which we might inhibit activator•AcID interactions. In Chapter 3, we apply the insights gained regarding interactions between activators and AcID to develop an assay adapted to high-throughput screening and describe the discovery of first-generation AcID inhibitors. A hypothetical mechanism of action by which these molecules inhibit activator•AcID interactions is also discussed. Finally, in Chapter 4 we describe a screen of a library composed of natural products extracted from diverse organisms against an interaction between a family of activators linked to the progression of cancer to a metastatic phenotype and AcID with the goal of identifying second generation AcID inhibitors.

1.10 References

1. JACOB, F. & MONOD, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3,** 318–356 (1961).

2. Gann, M. P. A. A. Genes and Signals. 1–209 (2005).

3. Lee, L. W. & Mapp, A. K. Transcriptional switches: chemical approaches to gene regulation. *Journal of Biological Chemistry* **285,** 11033–11038 (2010).

4. Pabo, C. O. & Sauer, R. T. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61,** 1053–1095 (1992).

5. Steitz, T. A. Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. *Q. Rev. Biophys.* **23,** 205–280 (1990).

6. Buck-Koehntop, B. A. *et al.* Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. *Proceedings of the National Academy of Sciences* **109,** 15229–15234 (2012).

7. Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* **71,** 1223–1237 (1992).

8. Nair, S. K. & Burley, S. K. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **112,** 193–205 (2003).

9. Johnson, P. F., Sterneck, E. & Williams, S. C. Activation domains of transcriptional regulatory proteins. *The Journal of Nutritional Biochemistry* **4,** 386–398 (1993).

10. Jonker, H. R. A., Wechselberger, R. W., Boelens, R., Folkers, G. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44,** 827–839 (2005).

11. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277,** 1310–1313 (1997).

12. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274,** 948–953 (1996).

13. Giniger, E. & Ptashne, M. Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. *Nature* **330,** 670–672 (1987).

14. Parker, D. *et al.* Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular cell* **2,** 353–359 (1998).

15. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci USA* **99,** 5271–5276 (2002).

16. Wands, A. M. *et al.* Transient-state kinetic analysis of transcriptional activator·DNA complexes interacting with a key coactivator. *Journal of Biological Chemistry* **286,** 16238–16245 (2011).

17. De Guzman, R. N., Goto, N. K., Dyson, H. J. & Wright, P. E. Structural basis for cooperative transcription factor binding to the CBP coactivator. *J. Mol. Biol.* **355,** 1005–1013 (2006).

18. Zor, T., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337,** 521–534 (2004).

19. Berlow, R. B., Dyson, H. J. & Wright, P. E. Functional advantages of dynamic protein disorder. *FEBS Letters* **589,** 2433–2440 (2015).

20. Cress, W. D. & Triezenberg, S. J. Critical structural elements of the VP16 transcriptional activation domain. *Science* **251,** 87–90 (1991).

21. Veleiro, A. S., Alvarez, L. D., Eduardo, S. L. & Burton, G. Structure of the Glucocorticoid Receptor, a Flexible Protein That Can Adapt to Different Ligands. *ChemMedChem* **5,** 649–659 (2010).

22. Bevan, C. L., Hoare, S., Claessens, F., Heery, D. M. & Parker, M. G. The AF1 and AF2 domains of the androgen receptor interact with distinct regions of SRC1. *Molecular and Cellular Biology* **19,** 8383–8392 (1999).

23. Rogatsky, I. *et al.* Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc Natl Acad Sci USA* **100,** 13845–13850 (2003).

24. Pilauri, V., Bewley, M., Diep, C. & Hopper, J. Gal80 dimerization and the yeast GAL gene switch. *Genetics* **169,** 1903–1914 (2005).

25. Hock, A. & Vousden, K. H. Regulation of the p53 pathway by ubiquitin and related proteins. *Int. J. Biochem. Cell Biol.* **42,** 1618–1621 (2010).

26. Näär, A. M., Lemon, B. D. & Tjian, R. Transcriptional coactivator complexes. *Annu. Rev. Biochem.* **70,** 475–501 (2001).

27. Lariviere, L., Seizl, M. & Cramer, P. A structural perspective on Mediator function. *Curr. Opin. Cell Biol.* **24,** 305–313 (2012).

28. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30,** 235–239 (2005).

29. Ebmeier, C. C. & Taatjes, D. J. Activator-Mediator binding regulates Mediator-cofactor interactions. *Proceedings of the National Academy of Sciences* **107,** 11283–11288 (2010).

30. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Reviews Genetics* **11,** 761–772 (2010).

31. Malik, S. & Roeder, R. G. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30,** 256–263 (2005).

32. Poss, Z. C., Ebmeier, C. C. & Taatjes, D. J. The Mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **48,** 575–608 (2013).

33. Takahashi, H. *et al.* Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146,** 92–104 (2011).

34. Esnault, C. *et al.* Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Molecular cell* **31,** 337–346 (2008).

35. Seizl, M., Larivière, L., Pfaffeneder, T., Wenzeck, L. & Cramer, P. Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. *Nucleic Acids Res.* **39,** 6291–6304 (2011).

36. Elmlund, H. *et al.* The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. *Proc Natl Acad Sci USA* **103,** 15788–15793 (2006).

37. Cai, G. *et al.* Mediator head module structure and functional interactions. *Nat. Struct. Mol. Biol.* **17,** 273–279 (2010).

38. Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nature Reviews Molecular Cell Biology* **16,** 155–166 (2015).

39. Cai, G., Imasaki, T., Takagi, Y. & Asturias, F. J. Mediator structural conservation and implications for the regulation mechanism. *Structure* **17,** 559–567 (2009).

40. Yuan, C. X., Ito, M., Fondell, J. D., Fu, Z. Y. & Roeder, R. G. The TRAP220 component of a thyroid hormone receptor- associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-dependent fashion. *Proc Natl Acad Sci USA* **95,** 7939–7944 (1998).

41. Lee, H.-K. H., Park, U.-H. U., Kim, E.-J. E. & Um, S.-J. S. MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* **26,** 3545–3557 (2007).

42. Kang, Y. K., Guermah, M., Yuan, C.-X. & Roeder, R. G. The TRAP/Mediator coactivator complex interacts directly with estrogen receptors alpha and beta through the TRAP220 subunit and directly enhances estrogen receptor function in vitro. *Proc Natl Acad Sci USA* **99,** 2642–2647 (2002).

43. Wang, Q., Sharma, D., Ren, Y. & Fondell, J. D. A coregulatory role for the TRAP-mediator complex in androgen receptor-mediated gene expression. *Journal of Biological Chemistry* **277,** 42852–42858 (2002).

44. Hittelman, A. B., Burakov, D., Iñiguez-Lluhí, J. A., Freedman, L. P. & Garabedian, M. J. Differential regulation of glucocorticoid receptor transcriptional activation via AF-1-associated proteins. *EMBO J.* **18,** 5380–5388 (1999).

45. Zhang, F., Sumibcay, L., Hinnebusch, A. G. & Swanson, M. J. A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Molecular and Cellular Biology* **24,** 6871–6886 (2004).

46. Thakur, J. K. *et al.* A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **452,** 604–609 (2008).

47. Park, J. M. *et al.* In vivo requirement of activator-specific binding targets of mediator. *Molecular and Cellular Biology* **20,** 8709–8719 (2000).

48. Mo, X., Kowenz-Leutz, E., Xu, H. & Leutz, A. Ras induces mediator complex exchange on C/EBP beta. *Molecular cell* **13,** 241–250 (2004).

49. Asada, S. *et al.* External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc Natl Acad Sci USA* **99,** 12747–12752 (2002).

50. Stevens, J. L. *et al.* Transcription control by E1A and MAP kinase pathway via Sur2 mediator subunit. *Science* **296,** 755–758 (2002).

51. Sela, D. *et al.* Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6α. *J. Biol. Chem.* **288,** 26179–26187 (2013).

52. Yang, F., DeBeaumont, R., Zhou, S. & Näär, A. M. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc Natl Acad Sci USA* **101,** 2339–2344 (2004).

53. Verger, A. *et al.* The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41,** 4847–4859 (2013).

54. Carlsten, J. O. P., Zhu, X. & Gustafsson, C. M. The multitalented Mediator complex. *Trends Biochem. Sci.* **38,** 531–537 (2013).

55. Chan, H. M. & La Thangue, N. B. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell. Sci.* **114,** 2363–2373 (2001).

56. Bannister, A. J., Oehler, T., Wilhelm, D., Angel, P. & Kouzarides, T. Stimulation of c-Jun activity by CBP: c-Jun residues Ser63/73 are required for CBP induced stimulation in vivo and CBP binding in vitro. *Oncogene* **11,** 2509–2514 (1995).

57. Gu, W., Shi, X. L. & Roeder, R. G. Synergistic activation of transcription by CBP and p53. *Nature* **387,** 819–823 (1997).

58. Zhong, H., Voll, R. E. & Ghosh, S. Phosphorylation of NF-kappa B p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300. *Molecular cell* **1,** 661–671 (1998).

59. Radhakrishnan, I. *et al.* Structural analyses of CREB-CBP transcriptional activator-coactivator complexes by NMR spectroscopy: implications for mapping the boundaries of structural domains. *J. Mol. Biol.* **287,** 859–865 (1999).

60. Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry* **48,** 2115–2124 (2009).

61. Wang, F. *et al.* Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proceedings of the National Academy of Sciences* **109,** 6078–6083 (2012).

62. Ramírez, J. A. & Nyborg, J. K. Molecular characterization of HTLV-1 Tax interaction with the KIX domain of CBP/p300. *J. Mol. Biol.* **372,** 958–969 (2007).

63. Campbell, K. M. & Lumb, K. J. Structurally distinct modes of recognition of the KIX domain of CBP by Jun and CREB. *Biochemistry* **41,** 13956–13964 (2002).

64. Thakur, J. K., Yadav, A. & Yadav, G. Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res.* **42,** 2112–2125 (2014).

65. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14,** 1553–1577 (2000).

66. Waters, L. *et al.* Structural diversity in p160/CREB-binding protein coactivator complexes. *J. Biol. Chem.* **281,** 14787–14795 (2006).

67. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384,** 641–643 (1996).

68. Thoma, F. & Koller, T. Influence of histone H1 on chromatin structure. *Cell* **12,** 101–107 (1977).

69. Jenuwein, T. Translating the Histone Code. *Science* **293,** 1074–1080 (2001).

70. Bedford, D. C. & Brindle, P. K. Is histone acetylation the most important physiological function for CBP and p300? *Aging (Albany NY)* **4,** 247–255 (2012).

71. Gronemeyer, H., Gustafsson, J.-Å. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3,** 950–964 (2004).

72. de Ruijter, A. J. M., van Gennip, A. H., Caron, H. N., Kemp, S. & van Kuilenburg, A. B. P. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem. J.* **370,** 737–749 (2003).

73. Lavery, D. N. & Bevan, C. L. Androgen receptor signalling in prostate cancer: the functional consequences of acetylation. *Journal of Biomedicine and Biotechnology* **2011,** 862125 (2011).

74. Tang, Y., Zhao, W., Chen, Y., Zhao, Y. & Gu, W. Acetylation is indispensable for p53 activation. *Cell* **133,** 612–626 (2008).

75. Pop, M. S. *et al.* A small molecule that binds and inhibits the ETV1 transcription factor oncoprotein. *Mol. Cancer Ther.* **13,** 1492–1502 (2014).

76. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21,** 381–395 (2011).

77. Lan, F. & Shi, Y. Epigenetic regulation: methylation of histone and non-histone proteins. *Sci. China, C, Life Sci.* **52,** 311–322 (2009).

78. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129,** 823–837 (2007).

79. Vignali, M., Hassan, A. H., Neely, K. E. & Workman, J. L. ATP-dependent chromatin-remodeling complexes. *Molecular and Cellular Biology* **20,** 1899–1910 (2000).

80. Conaway, R. C. & Conaway, J. W. General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* **62,** 161–190 (1993).

81. Thomas, M. C. & Chiang, C.-M. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41,** 105–178 (2006).

82. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chemical Biology* **7,** 1311–1320 (2012).

83. Vassilev, L. T. *et al.* In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303,** 844–848 (2004).

84. Zhang, M. *et al.* Genetic and chemical targeting of epithelial-restricted with serine box reduces EGF receptor and potentiates the efficacy of afatinib. *Mol. Cancer Ther.* **12,** 1515–1525 (2013).

85. Kouraklis, G., Misiakos, E. P. & Theocharis, S. Histone deacetylase inhibitors as a potential therapeutic agent for human cancer treatment. *Targeted Oncology* (2006).

86. Manzo, F., Tambaro, F. P., Mai, A. & Altucci, L. Histone acetyltransferase inhibitors and preclinical studies. *Expert Opin Ther Pat* **19,** 761–774 (2009).

87. Duvic, M. & Vu, J. Update on the treatment of cutaneous T-cell lymphoma (CTCL): Focus on vorinostat. *Biologics* **1,** 377–392 (2007).

88. Little, C. D., Nau, M. M., Carney, D. N., Gazdar, A. F. & Minna, J. D. Amplification and expression of the c-myc oncogene in human lung cancer cell lines. *Nature* **306,** 194–196 (1983).

89. Münzel, P., Marx, D., Köchel, H., Schauer, A. & Bock, K. W. Genomic alterations of the c-myc protooncogene in relation to the overexpression of c-erbB2 and Ki-67 in human breast and cervix carcinomas. *J. Cancer Res. Clin. Oncol.* **117,** 603–607 (1991).

90. Augenlicht, L. H. *et al.* Low-level c-myc amplification in human colonic carcinoma cell lines and tumors: a frequent, p53-independent mutation associated with improved outcome in a randomized multi-institutional trial. *Cancer Res.* **57,** 1769–1775 (1997).

91. Dang, C. V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Molecular and Cellular Biology* **19,** 1–11 (1999).

92. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).

93. Lin, C. Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151,** 56–67 (2012).

94. Pattabiraman, D. R. *et al.* Interaction of c-Myb with p300 is required for the induction of acute myeloid leukemia (AML) by human AML oncogenes. *Blood* **123,** 2682–2690 (2014).

95. Visakorpi, T. *et al.* In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat. Genet.* **9,** 401–406 (1995).

96. Kinoshita, J. *et al.* Clinical significance of PEA3 in human breast cancer. *Surgery* **131,** S222–5 (2002).

97. Salgia, R. & Skarin, A. T. Molecular abnormalities in lung cancer. *J. Clin. Oncol.* **16,** 1207–1217 (1998).

98. Gouyer, V. *et al.* Mechanism of retinoblastoma gene inactivation in the spectrum of neuroendocrine lung tumors. *Am. J. Respir. Cell Mol. Biol.* **18,** 188–196 (1998).

99. Hansen, M. F. *et al.* Osteosarcoma and retinoblastoma: a shared chromosomal mechanism revealing recessive predisposition. *Proc Natl Acad Sci USA* **82,** 6216–6220 (1985).

100. Narla, G. *et al.* KLF6, a candidate tumor suppressor gene mutated in prostate cancer. *Science* **294,** 2563–2566 (2001).

101. Kubbutat, M. H., Jones, S. N. & Vousden, K. H. Regulation of p53 stability by Mdm2. *Nature* **387,** 299–303 (1997).

102. Boxer, L. M. & Dang, C. V. Translocations involving c-myc and c-myc function. *Oncogene* **20,** 5595–5610 (2001).

103. Tomlins, S. A. *et al.* Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia* **10,** 177–188 (2008).

104. Barr, F. G. Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. *Oncogene* **20,** 5736–5746 (2001).

105. Khoo, K. H., Hoe, K. K., Verma, C. S. & Lane, D. P. Drugging the p53 pathway: understanding the route to clinical efficacy. *Nat Rev Drug Discov* **13,** 217–236 (2014).

106. Ludwig, R. L., Bates, S. & Vousden, K. H. Differential activation of target cellular promoters by p53 mutants with impaired apoptotic function. *Molecular and Cellular Biology* **16,** 4952–4960 (1996).

107. Weisz, L., Oren, M. & Rotter, V. Transcription regulation by mutant p53. *Oncogene* **26,** 2202–2211 (2007).

108. Di Agostino, S. *et al.* Gain of function of mutant p53: the mutant p53/NF-Y protein complex reveals an aberrant transcriptional mechanism of cell cycle regulation. *Cancer Cell* **10,** 191–202 (2006).

109. Gaiddon, C., Lokshin, M., Ahn, J., Zhang, T. & Prives, C. A subset of tumor-derived mutant forms of p53 down-regulate p63 and p73 through a direct interaction with the p53 core domain. *Molecular and Cellular Biology* **21,** 1874–1887 (2001).

110. Song, H., Hollstein, M. & Xu, Y. p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. *Nature Cell Biology* **9,** 573–580 (2007).

111. Lee, W. *et al.* Solution structure of the tetrameric minimum transforming domain of p53. *Nat. Struct. Biol.* **1,** 877–890 (1994).

112. Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* **45,** 6873–6888 (2006).

113. Cesa, L. C., Mapp, A. K. & Gestwicki, J. E. Direct and Propagated Effects of Small Molecules on Protein-Protein Interaction Networks. *Front Bioeng Biotechnol* **3,** 119 (2015).

114. Böttger, A. *et al.* Design of a synthetic Mdm2-binding mini protein that activates the p53 response in vivo. *Curr. Biol.* **7,** 860–869 (1997).

115. Lin, J., Wu, X., Chen, J., Chang, A. & Levine, A. J. Functions of the p53 protein in growth regulation and tumor suppression. *Cold Spring Harb. Symp. Quant. Biol.* **59,** 215–223 (1994).

116. Grasberger, B. L. *et al.* Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *J. Med. Chem* **48,** 909–912 (2005).

117. Yin, H. *et al.* Terphenyl-based helical mimetics that disrupt the p53/HDM2 interaction. *Angew. Chem. Int. Ed. Engl.* **44,** 2704–2707 (2005).

118. Stoll, R. *et al.* Chalcone derivatives antagonize interactions between the human oncoprotein MDM2 and p53. *Biochemistry* **40,** 336–344 (2001).

119. Shangary, S. *et al.* Temporal activation of p53 by a specific MDM2 inhibitor is selectively toxic to tumors and leads to complete tumor growth inhibition. *Proceedings of the National Academy of Sciences* **105,** 3933–3938 (2008).

120. Fuxreiter, M. *et al.* Malleable machines take shape in eukaryotic transcriptional regulation. *Nat. Chem. Biol.* **4,** 728–737 (2008).

121. Hargreaves, D. C. & Crabtree, G. R. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res.* **21,** 396–420 (2011).

122. Lessard, J. *et al.* An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55,** 201–215 (2007).

123. Wu, J. I. *et al.* Regulation of dendritic development by neuron-specific chromatin remodeling complexes. *Neuron* **56,** 94–108 (2007).

124. Krishnamurthy, M. *et al.* Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo. *ACS Chemical Biology* **6,** 1321–1326 (2011).

125. Xie, X. *et al.* Targeting HPV16 E6-p300 interaction reactivates p53 and inhibits the tumorigenicity of HPV-positive head and neck squamous cell carcinoma. *Oncogene* **33,** 1037–1046 (2014).

126. Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M. & Wright, P. E. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proceedings of the National Academy of Sciences* **107,** 19290–19295 (2010).

127. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *Journal of Biological Chemistry* **277,** 43168–43174 (2002).

128. Law, S. M., Gagnon, J. K., Mapp, A. K. & Brooks, C. L. Prepaying the entropic cost for allosteric regulation in KIX. *Proceedings of the National Academy of Sciences* **111,** 12067–12072 (2014).

129. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. *Proceedings of the National Academy of Sciences* **111,** 12061–12066 (2014).

130. Leonard, S. E., Register, A. C., Krishnamurty, R., Brighty, G. J. & Maly, D. J. Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors. *ACS Chemical Biology* **9,** 1894–1905 (2014).

131. Cesa, L. C. *et al.* Inhibitors of difficult protein-protein interactions identified by high-throughput screening of multiprotein complexes. *ACS Chemical Biology* **8,** 1988–1997 (2013).

132. Wang, N. *et al.* Ordering a dynamic protein via a small-molecule stabilizer. *J. Am. Chem. Soc.* **135,** 3363–3366 (2013).

133. Majmudar, C. Y. *et al.* Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300. **51,** 11258–11262 (2012).

134. Gee, C. T., Koleski, E. J. & Pomerantz, W. C. K. Fragment screening and druggability assessment for the CBP/p300 KIX domain through protein-observed 19F NMR spectroscopy. *Angew. Chem. Int. Ed.* **54,** 3735–3739 (2015).

135. Bradner, J. E., McPherson, O. M. & Koehler, A. N. A method for the covalent capture and screening of diverse small molecules in a microarray format. *Nat Protoc* **1,** 2344–2352 (2006).

CHAPTER 2

Molecular Underpinnings of Activator•Med25 Interactions[2]

2.1 Abstract

The Activator Interaction Domain (AcID) of Mediator subunit Med25 is a critical coactivator binding partner for a variety of transcriptional activators including the herpes simplex viral activator VP16, the PEA3 subfamily of Ets transcription factors, and the endoplasmic reticulum stress response transcription factor ATF6α. AcID utilizes a unique fold found in no other characterized coactivator and as such, very little is understood about specific molecular requirements for activator•AcID interactions. In this chapter, specific characteristics of these interactions, such as the minimal sequence of the VP16 TAD capable of interacting with the domain and the role of electrostatics in promoting these interactions are explored. Additionally, alanine scanning mutagenesis studies demonstrate that these interactions occur over broad surface areas with no defined hot-spot contacts. Collectively, these data also suggest at least two distinct binding surfaces that are allosterically connected. Taken together, the data indicate that this motif will be most effectively targeted by allosteric modulators, a strategy that is employed in the work of Chapters 3 and 4.

2.2 Introduction

Transcription is the critical process by which information encoded within an organism's genome is converted to mRNA before being ultimately translated into protein and plays a fundamental role in virtually all cellular processes. This process is tightly regulated and requires the assembly of large and dynamic protein complexes at gene

[2] The research described in this chapter was a collaborative effort. Steven M. Sturlis and Paul A. Bruno expressed and purified AcID protein and completed the VP16 truncation analysis. Matthew S. Beyersdorf expressed and purified the AcID R466E mutant and completed direct binding assays with the mutant protein.

promoters through the use of specific protein-protein interactions between complex components.[1] In particular, interactions between transcriptional activators and coactivators are a fundamental requirement for activated transcription at a majority of gene promoters and represent a particularly attractive target for inhibitor development given the ability to achieve excellent context specificity, as discussed in Chapter 1. However, the large surface areas over which these interactions occur and the relatively low affinities of activators for coactivator binding partners complicates small molecule inhibitor discovery.[2] Therefore, in order to identify attractive features for potential exploitation in a small molecule inhibition strategy, it is first necessary to determine the molecular underpinnings for a particular activator•coactivator interaction. In this case our goal is to define the interactions between transcriptional activators and the Activator Interaction Domain of Med25, with the ultimate goal of discovering small molecule modulators of these complexes (Chapter 3, 4).

Mediator Subunit 25 (Med25)

Med25, previously known as ARC92 or DRIP97, is a component of the megadalton Mediator complex, a critical coactivator for the expression of virtually protein coding genes (described in Chapter 1).[3] This subunit is a 92 kDa protein found only in higher eukaryotes and is reliant upon three key domains to link transcriptional activators to the rest of the Mediator complex, as shown in Figure 2.1 (A).[4] At the N-terminus (residues 1-229) is the von Willebrand factor type A domain (VWA), which is responsible for the association of Med25 with the larger Mediator complex, though the interaction partners within the rest of the Mediator complex are currently unknown.[5] A LXXLL motif within the nuclear receptor box (residues 646-650) is capable of interactions with retinoic acid receptor α (RARα) and estrogen receptor α (ERα).[6,7] Finally, the Activator Interaction Domain (AcID, a.k.a VP16 binding domain, VBD; residues 395-545) is responsible for the association of Med25 with transcription factors such as VP16[8], the PEA3 subfamily of ETS transcriptional activators[9], the endoplasmic reticulum stress response transcriptional activator ATF6α[10], and the N-terminal domain of CBP/p300[6]. Though the specific contact points within the Mediator complex have yet to be elucidated, proteomic experiments to determine the relative binding locations of the various Mediator subunits have indicated

that Med25 is localized within the tail module of the complex, consistent with other subunits responsible for linking Mediator to DNA-bound transcriptional activators such as Med15 and Med23.[11]

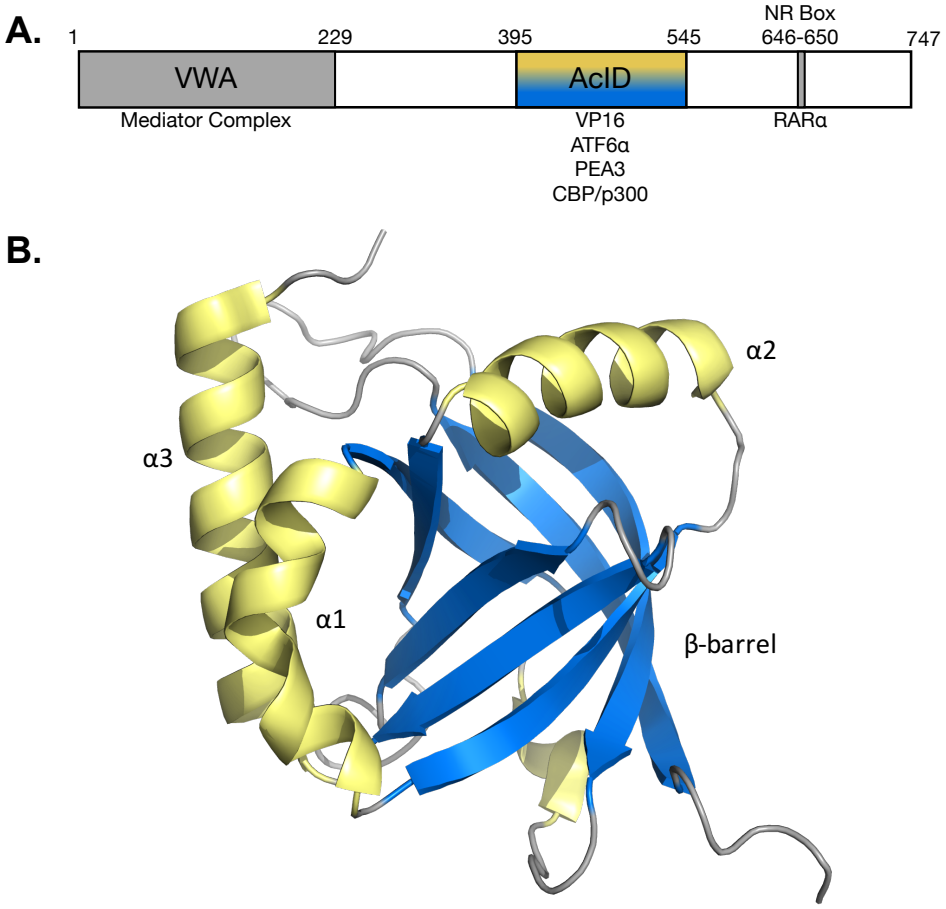


***Figure 2.1*- Med25 Organization and AcID Structure.** (A) Schematic representation of Med25 domain organization, with known interaction partners listed below the domains. (B) Cartoon representation of AcID. α-helices are shown in yellow, while the β-barrel is shown in blue.

Recently, solution NMR structures of AcID have been reported by several research groups.[8,12,13] The studies identify the domain as being composed of a seven-stranded β-barrel flanked by three α-helices, as shown in Figure 2.1 (B). Interestingly, the architecture of ACID represents a novel activator target fold not previously reported. In particular, the presence of a β-barrel as the defining structural feature of the domain is unusual, as most activator-targeted coactivator domains are highly α-helical in nature, as shown in Figure 2.2. [14-16]

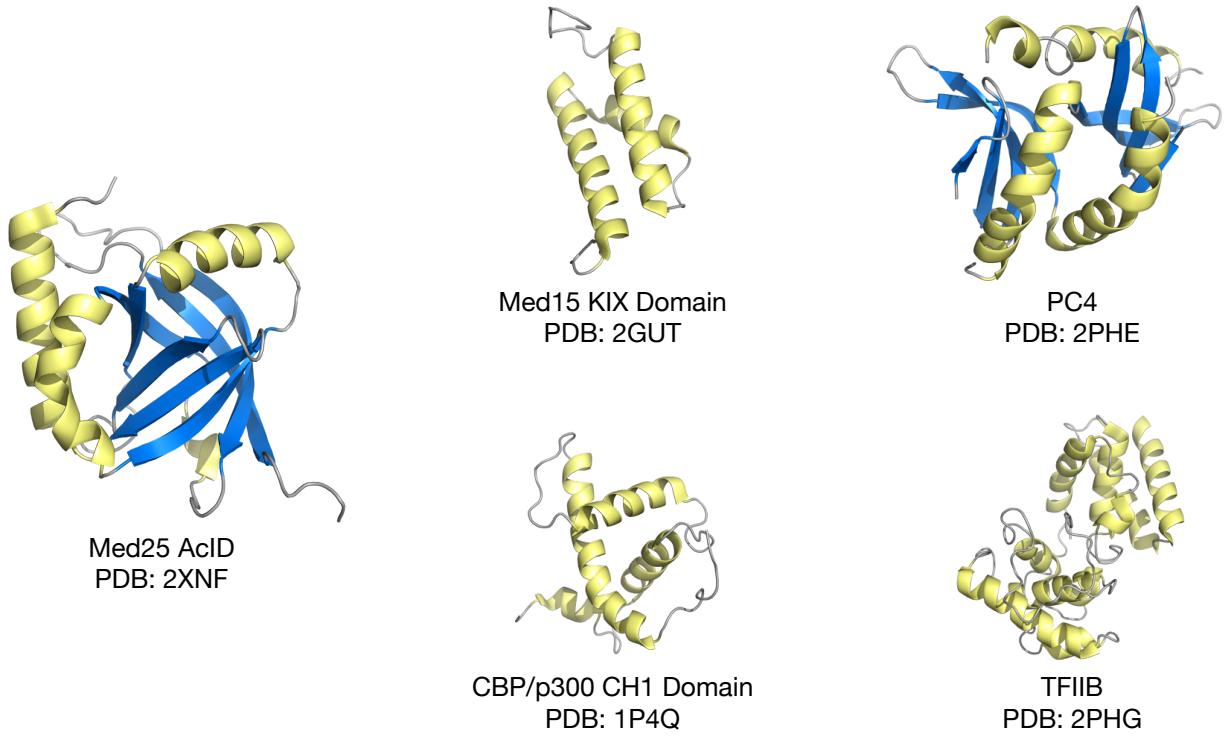


***Figure 2.2*- Comparison of AcID to other Activator-Targeted Coactivator Folds** The presence of a β-barrel as a defining structural feature of AcID is highly unusual for coactivator proteins. As the four examples on the right show, most coactivator folds are highly α-helical, though PC4 demonstrates that β-sheets are not necessarily uncommon.

In fact, a search of known protein structures reveals that there are only two other known domains with similar architecture to the ACID, both of which are found in a protein known as PTOV1, which is an overexpressed protein of unknown function found in cancerous prostate tissue. [8,12,17] Thus, AcID is an attractive potential target for small molecule inhibitors given that: (1) the unique nature of the domain's fold should allow for the discovery of highly selective small molecules with few off target effects, and (2) development of small molecule inhibitors of activator•AcID interactions would be useful in answering mechanistic questions about this poorly understood domain and may serve as potential leads in novel therapeutic strategies for diseases linked to AcID dependent transcriptional activators. However, the dearth of information surrounding activator•AcID interactions requires that their molecular underpinnings be defined in order to identify potential molecular features for exploitation and to develop an effective strategy for small molecule inhibitor discovery. This is particularly important given the unusual nature of the

AcID fold, suggesting that the domain may interact with activator binding partners in unique ways not typical of most coactivator proteins.

The VP16•AcID Interaction and Structural Insights

Of the four AcID-dependent interactions identified above, the interaction between the viral activator VP16 and AcID is perhaps the best studied. VP16 is a critical component of the herpes simplex virus (HSV) responsible for the expression of immediate early viral genes, which are important for the functional switch from latent to lytic infection within host cells.[18] This domain is rich in acidic amino acids, making it a prototypical member of the acidic class of transcriptional activators, though its function is also dependent upon specific hydrophobic residues.[19] Furthermore, it has been demonstrated that the VP16 TAD contains two subdomains termed H1 (residues 410-452) and H2 (residues 453-490) that are capable of independently initiating transcription using distinct molecular mechanisms.[20,21] Interestingly, there is significant evidence that free VP16 TAD is flexible and relatively disordered, but adopts an α-helical conformation upon binding to target proteins.[15] This observation has been made for interactions between VP16 and coactivator proteins with α-helical secondary structure, thus the importance of the putative VP16 α-helices for interactions with the β-barrel of AcID will be an important point of inquiry in this chapter.

The discovery that VP16 interacts with Med25 was initially made based on the observation that an affinity column composed of an immobilized VP16-GST fusion was capable of purifying the Mediator complex from HeLa cell nuclear extracts[22], which was shown to by mediated by Med25 following a more stringent analysis.[23] More recently, $^{1}H,^{15}N$-HSQC NMR chemical shift perturbation studies of ^{15}N-labeled AcID in complex with unlabeled VP16 TAD have been reported, revealing two putative binding sites for the VP16 H1 andH2 subdomains, as shown below in Figure 2.3. [8,12,13]

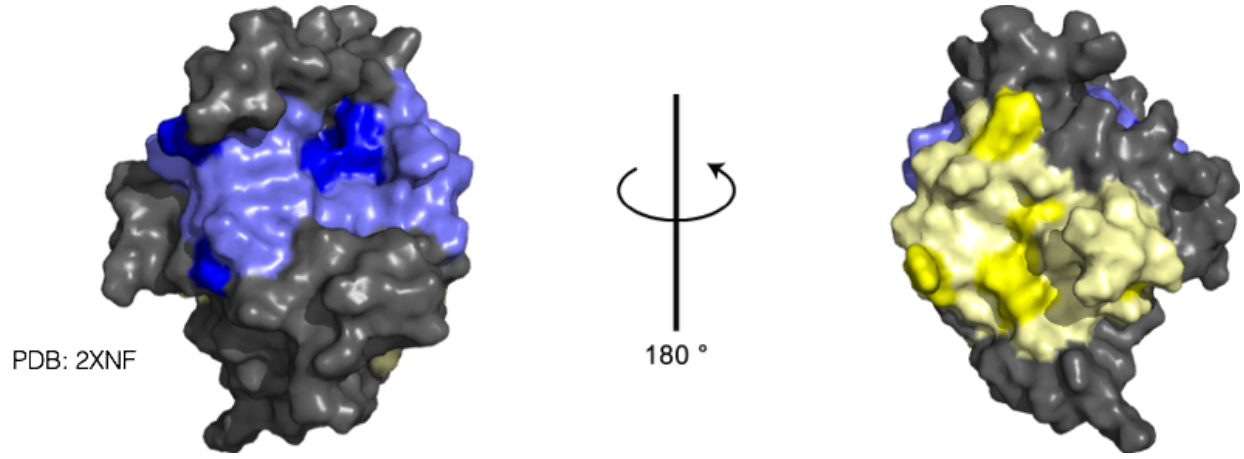


***Figure 2.3*- Med25 AcID Contains Two Binding Sites that Interact with the VP16 TAD** The VP16 TAD contacts two distinct surfaces of the AcID domain located on opposite faces of the protein. The darker shaded sites in each surface represent the residues most strongly perturbed in HSQC experiments in which the VP16 TAD was titrated.

More specifically, an examination of the chemical shift perturbations in the NMR spectrum of AcID following the titration of full length VP16 TAD suggested that the binding interface occurs over a broad surface of the protein. Subsequent titration of isolated VP16 H1 or H2 subdomains demonstrated that the subdomains perturbed individual subsets of the chemical shifts induced by the binding of the full length TAD. Together, these data indicated that the H1 and H2 subdomains of the VP16 TAD contact distinct binding sites on opposite faces of AcID, as shown in Figure 2.3. Residues that were most significantly perturbed in the HSQC experiments are highlighted in bright blue or yellow[8,12], while the lighter shades represent the full estimated interaction surfaces. Measurements using PyMol indicate that both surfaces cover approximately 1800 $Å^2$, suggesting that activator•AcID interactions occur over very broad surface areas. Though there is information available with regards to the effects of VP16 binding on the structure of AcID, very little is known with regards to the molecular determinants within the TAD that are required for binding to AcID.

The relative importance of these binding surfaces for interaction with VP16 is a matter of debate. In one report, mutation of the H1 subdomain had a minimal impact on transcriptional activity of a Gal4-VP16 fusion in yeast nuclear extracts, while mutation of the H2 subdomain resulted in a significant decrease in activity. Similarly, the H1

subdomain alone was insufficient for transcriptional activation while the H2 subdomain was capable of moderate activation, suggesting that the H2 binding surface may be a more functionally significant binding surface.[12] Alternative evidence suggests that the H1 subdomain binds with significantly higher affinity than the H2 subdomain and also undergoes more significant structural shifts following binding to AcID based on HSQC studies using ^{15}N-labeled VP16. [8,23]

In this Chapter we aim to elucidate underlying molecular features of activator interactions with the Activator Interaction Domain (AcID) of Med25 in order to identify potential mechanisms by which small molecule inhibitors could inhibit these interactions. For example, we will identify the minimal sequences of the VP16 TAD that interact with the AcID motif. We will also determine whether activator•AcID interactions occur primarily through the use of focused hotspot residues or over broad surfaces. Finally, we will explore the importance of electrostatic interactions in activator•AcID complexes, given the highly basic character of AcID and acidic character of known activator binding partners.

2.3 Results and Discussion

Determining the Minimal VP16 Interaction Surface

Following expression of the AcID protein, we next sought to determine the critical VP16 sequences that are capable of interacting with AcID in order to determine important molecular features for recognition and affinity. Previous studies have demonstrated that two regions within the VP16 TAD can adopt an α-helical structure.[15,24] ^{1}H,^{15}N-HSQC perturbation experiments of VP16(413-451), the H1 subdomain, demonstrated particularly strong structural changes in residues 435-442. This region encompasses one of the purported α-helices and the lack of perturbation in the N-terminal region of the TAD suggests that it may be dispensable for binding. Furthermore, the observation that the chemical shifts in residues 435-442 were only observable following titration of AcID supports previous observations that the VP16 TAD is intrinsically disordered and only adopts secondary structure following interaction with a coactivator binding partner.[8] Analogous studies completed with VP16(452-490), the H2 subdomain, failed to identify specific structural changes within the TAD due to significant line broadening that precluded prediction of VP16 H2 secondary structure following binding to AcID. However,

the approximately 1:1 stoichiometry required to saturate AcID, indicated that the VP16 H2•AcID interaction is relatively high affinity.[12]

For our purposes, the VP16 TAD was truncated into three peptides spanning the full length of the TAD, VP16(413-437), VP16(438-464), and VP16(465-490). Fluorescein was appended to the N-terminus of the peptides following a β-alanine linker and the affinity of the peptides was assessed for AcID using fluorescence polarization.

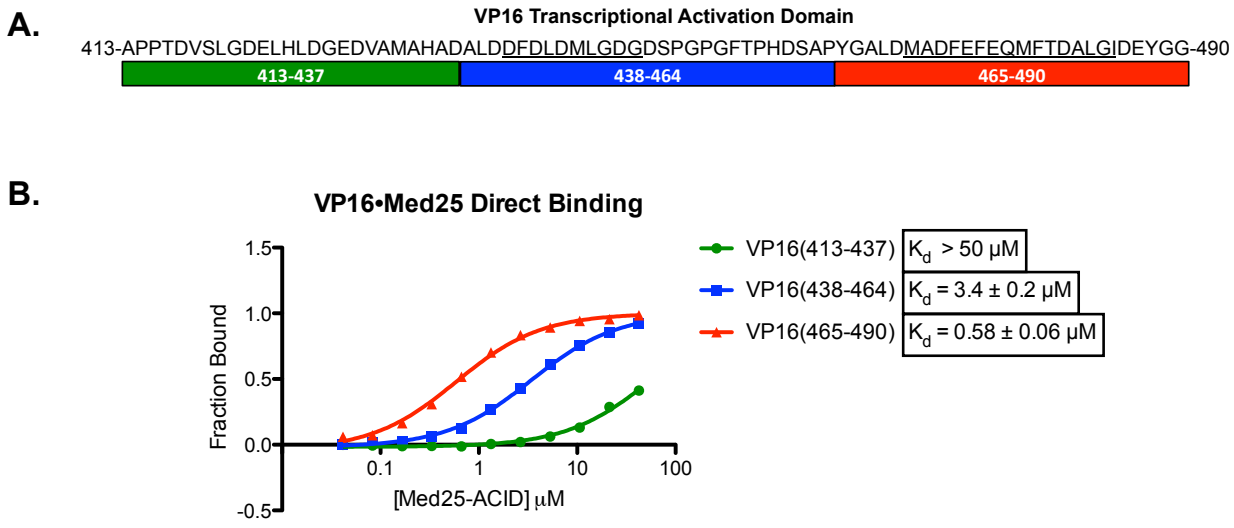


***Figure 2.4*- Initial Truncation of the VP16 TAD and Determination of Affinity for AcID** (A) Sequence of the VP16 TAD with the indicated truncations represented by the colored bars beneath the sequence. Purported α-helical regions within the TAD are indicated by underlining within the TAD. (B) Direct binding curves for the initial VP16 truncation color coded to match the sequences in (A). Curves represent the mean values of three independent experiments with error bars indicating the standard deviation of the fraction bound at each concentration of the purified AcID protein. *(With Paul A. Bruno)*

The N-terminus of the VP16 TAD contributes very little to the interaction between the VP16 TAD and AcID, as evidenced by the low affinity of VP16(413-437) for the purified domain. In fact, an accurate K_d for VP16(413-437) could not be determined due to limitations in concentrating the expressed AcID protein. VP16(438-464) and VP16(465-490) bound to AcID with K_d values of 3.4 μM and 0.58 μM, indicating that these are relatively high affinity interactions compared to several other activator•coactivator interactions.[25-27] Interestingly, these two peptides contain sequences that have been previously reported to be functionally independent and capable of adopting α-helical secondary structure upon binding to target coactivators.[15,24] Thus, a potential hypothesis is that VP16•AcID interactions may also be dependent upon α-helical secondary structure

within the TAD contacting the β-barrel of AcID. Such a binding mode would be similar to other activator•coactivator interactions in that the TAD becomes α-helical following contact with its binding partner, but unique in that the TAD contacts a β-barrel as opposed to helices within the coactivator.

In order to begin to test this hypothesis, the VP16 TAD was further truncated to produce peptides that contained only the purported α-helices and three residues on either side in order to allow for maximal potential helicity within the peptides. While these experiments do not directly examine the required secondary structure of AcID binding partners, they are useful in further minimizing the VP16 surfaces that interact efficiently with AcID. These shorter peptides, VP16(438-454) and VP16(467-488) were conjugated to fluorescein and their affinities for AcID were again measured by fluorescence polarization.

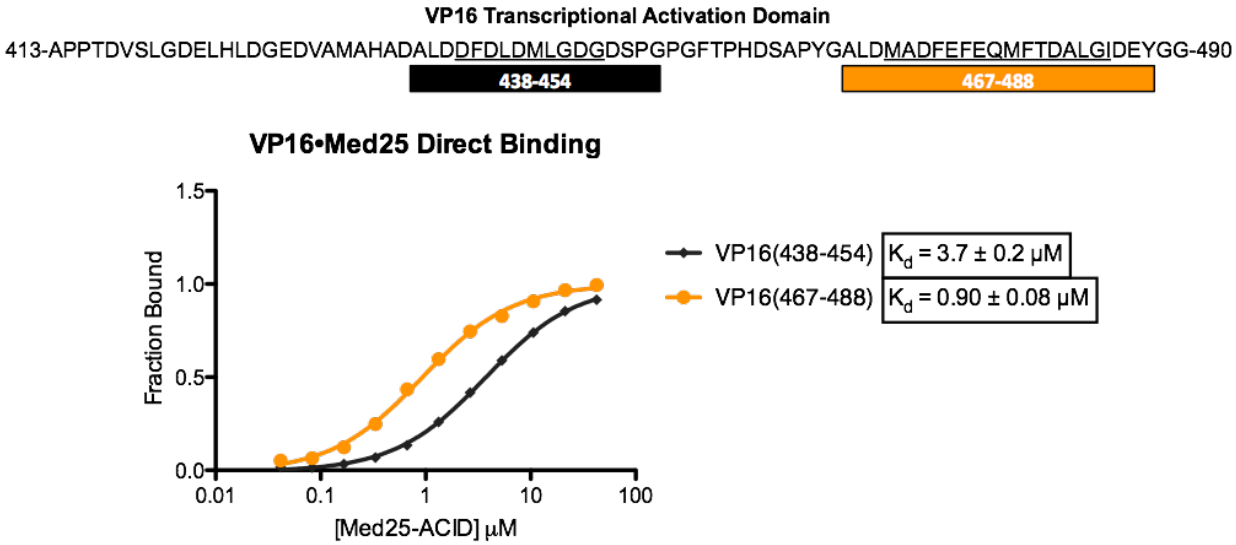


***Figure 2.5*- Direct Binding Assays of VP16(438-454) and VP16(467-488) to AcID** The affinity of VP16(438-454) and VP16(467-488) were determined using fluorescence polarization. Curves represent the mean values of three independent experiments with error bars indicating the standard deviation of the fraction bound at each concentration of the purified AcID protein. *(With Paul A. Bruno)*

As shown in Figure 2.5, the truncated tracers maintained comparable binding affinity to their larger counterparts. Specifically, VP16(438-454) is within error of VP16(438-464), and VP16(467-488) has an only slightly lower affinity than VP16(465-490) with K_d values of 0.90 μM compared to 0.58 μM. A possible explanation for the slight difference between VP16(467-488) and VP16(465-490) is that the theoretical binding site on AcID for the H2 derived peptides is comprised almost entirely of the β-barrel and thus

lacks significant topology, unlike the binding site for H1 derived peptides which bind to a surface that contains a significant cleft formed by the β-barrel and helix α3. Thus, the VP16 H2 derived peptides may be more dependent upon weak hydrophobic interactions, though the difference between the two affinities is insignificant enough that the data support the hypothesis that the purported α-helical region of the TAD is a key contributor to the VP16(465-490)•AcID interaction.

The affinities of these shorter peptides for AcID appears to further support the hypothesis that the domain contains two independent binding sites. Other VP16 coactivator binding partners, such as PC4, require the simultaneous binding of both of the α-helical sequences to maintain a functionally relevant affinity. More specifically, the full length VP16 TAD binds to PC4 with an affinity of 0.7 ± 0.2 μM, while the H1 subdomain does not bind with a detectable affinity and the H2 subdomain binds with an affinity of 15 ± 6 μM, more than twenty-fold weaker than the full length TAD.[15] Thus, the relatively high affinity of VP16(438-454) and VP16(467-488) for AcID supports that both of these sequences are capable of making necessary intermolecular contacts.

As a final experiment to determine the minimal interaction surface of VP16 for AcID, an eight amino acid peptide was synthesized based upon the purported α-helical sequence in the H2 subdomain. VP16(472-479) has been examined before in the literature and was found to be capable of stimulating transcription, indicating that this sequence represents a minimal functional unit of the VP16 TAD.[24]

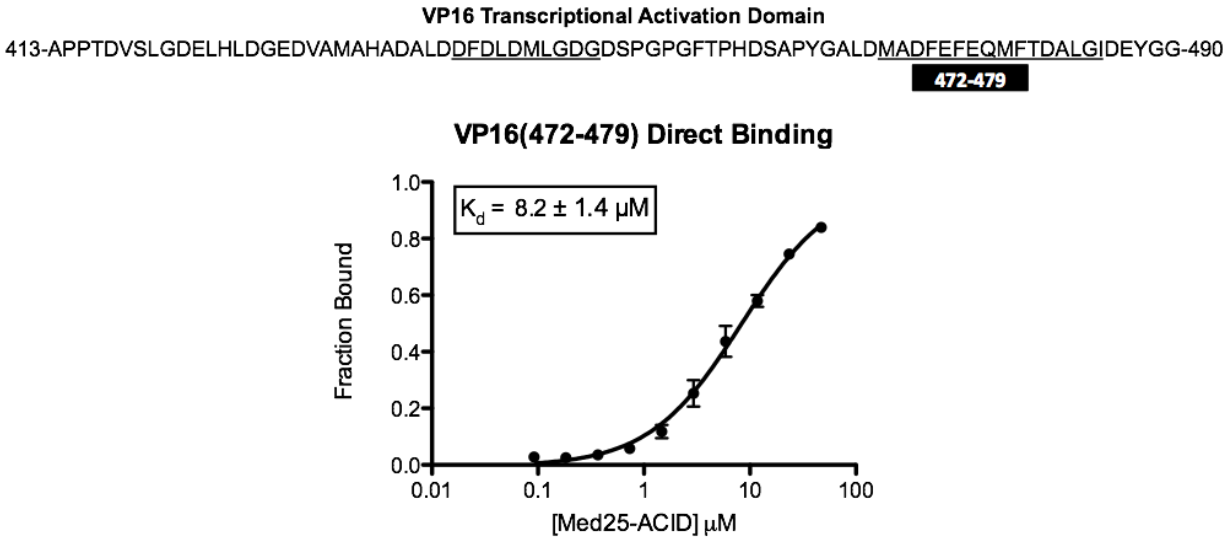


Figure 2.6- **Direct Binding of VP16(472-479) to AcID** The affinity of VP16(472-479) was determined using fluorescence polarization. Curves represent the mean values of three independent experiments with error bars indicating the standard deviation of the fraction bound at each concentration of the purified AcID protein.

Even this significantly truncated peptide is capable of interacting with AcID with modest affinity, with a K_d value of 8.2 μM representing only a ten-fold loss in affinity from VP16(467-488). Thus, the collective truncation data demonstrates that the VP16•AcID interaction does not require the full length TAD or even the intact H1 or H2 subdomains and can be minimized to relatively short sequences containing previously identified functional units within the TAD that may adopt an α-helical conformation.

The transcriptional activation domains of activators frequently adopt an α-helical conformation in order to efficiently bind to coactivator binding partners. For example, a phosphorylation event at serine 133 within the CREB kinase inducible domain (KID) stabilizes α-helical structure allowing the TAD to contact the coactivator CBP/p300.[28] Thus, given the prominent role that this secondary structure appears to play in activator•coactivator interactions in general, the finding that the minimal VP16 interaction surfaces for the AcID motif may also adopt α-helical secondary structure is unsurprising and suggests that activator binding partners of AcID may be α-helical despite the interaction occurring at a surface primarily composed of a β-barrel. Though most coactivator domains are highly α-helical in terms of secondary structure, as shown in Figure 2.2, reported molecular docking experiments combined with HSQC chemical shift

perturbation experiments suggest that the VP16 TAD adopts an α-helix and binds to a positive 'channel' formed by a series of highly basic β-sheets within a homodimer of the coactivator PC4, further supporting this hypothesis.[15] Additional analysis including mutations to disrupt helical propensity or NMR solution structures of ^{15}N labeled peptides in complex with the domain will be completed to further explore this hypothesis.

Furthermore, this trend appears to apply to activators that bind to AcID beyond VP16, with ERM/ETV5 as an excellent example.

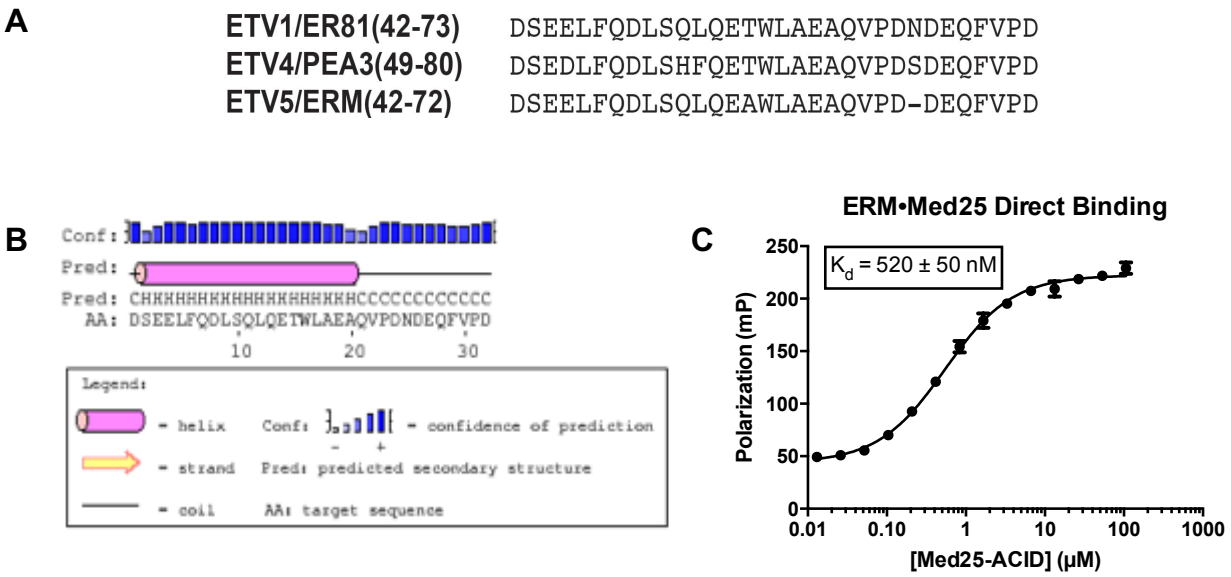


***Figure 2.7*- The TAD of PEA3 Family Members Bind to AcID.** (A) The three transcriptional activators that comprise the PEA3 subfamily contain TADs that are highly conserved. (B) Secondary structure predictions demonstrate with a high degree of confidence that the ERM TAD adopts an α-helical conformation at its N-terminus. (C) The ERM TAD, residues 38-68, bind to AcID with relatively high affinity.

As shown in Figure 2.7 (A), the TADs of the three members that comprise the PEA3 subfamily are highly conserved and highly acidic, similar to the VP16 TAD. In (B) secondary structure prediction suggests with high confidence that the N-terminal sequence of the ERM TAD, ERM(38-68), adopts an α-helical secondary structure, and in (C) the binding curve of the fluorescein labeled peptide shows that the ERM TAD binds to AcID with an affinity of approximately 520 nM, consistent with literature reports.[9] Recently published NMR studies of the ERM TAD in complex with AcID further supports that the TAD is disordered in solution, but adopts an α-helical conformation following contact and sampling of AcID binding surfaces.[29] Though additional analysis is necessary, such as mutations to disrupt helical propensity or truncations similar to those

described above for VP16, the emerging model is that ERM binds to AcID using predominantly α-helical secondary structure.

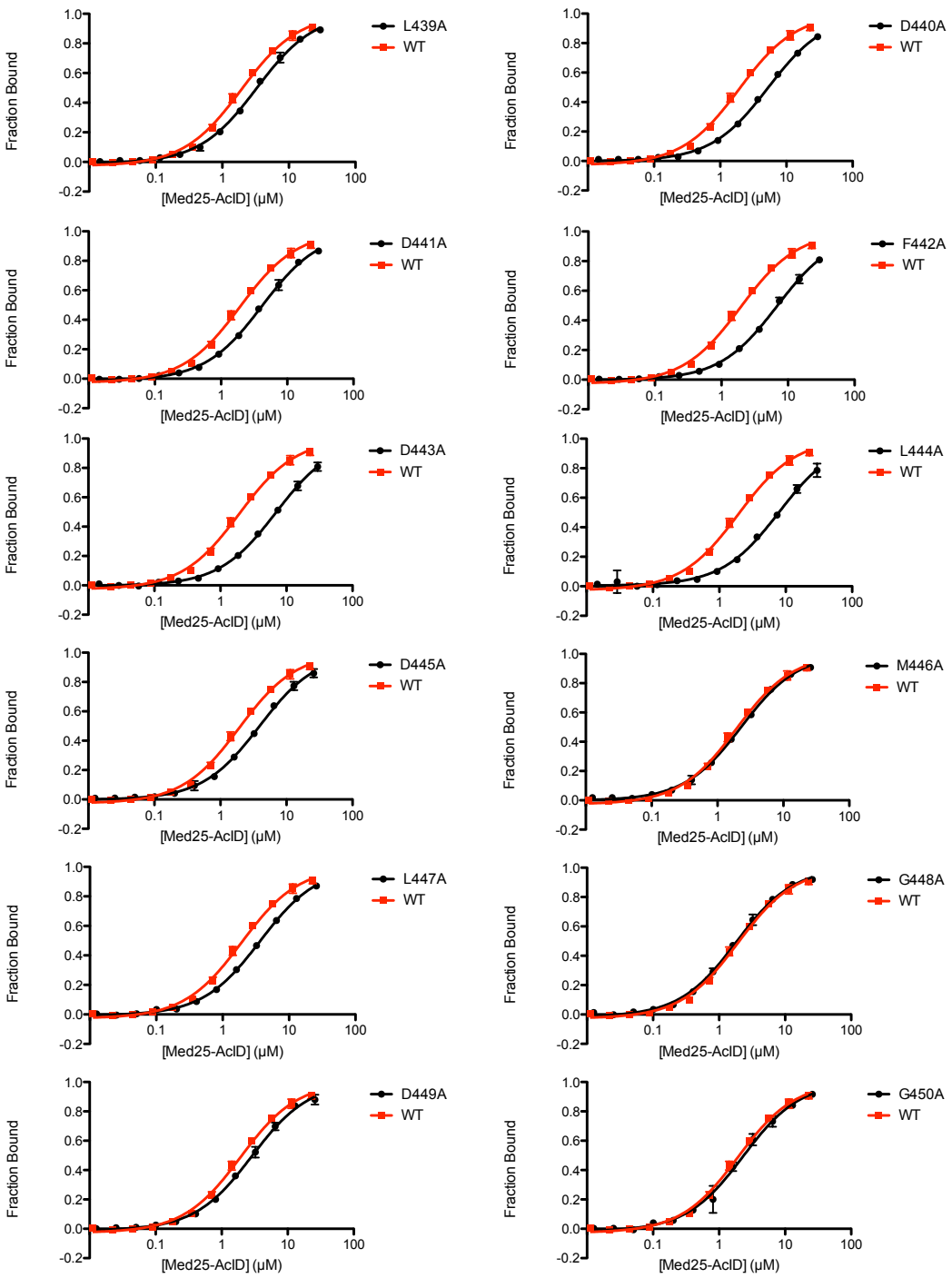
VP16•AcID Hotspot Analysis Using Alanine Scanning Mutagenesis

Based on the observation that the majority of the interaction could be recapitulated with sequences containing the putative α-helical portions of the VP16 TADs, we next sought to determine if specific residues within those sequences are particularly critical for the majority of the binding energy in order to identify potential hotspot residues within the interaction. In order to accomplish this, an alanine scan of VP16(438-454) and VP16(467-488) was completed. Specifically, this hotspot analysis was completed by synthesizing fluorescein labeled peptides in which each residue within the VP16(438-454) and VP16(467-488) sequence was sequentially mutated to an alanine, determining the affinity of the mutant for AcID by FP, and comparing the mutant affinity to the wild type affinity in order to determine the individual contribution for each residue to the binding energy. Alanine scanning hotspot analysis is a common method to determine the energetic contributions of individual amino acids in protein-protein interactions.[30,31]

A

438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454
A	L	D	D	F	D	L	D	M	L	G	D	G	D	S	P	G

B



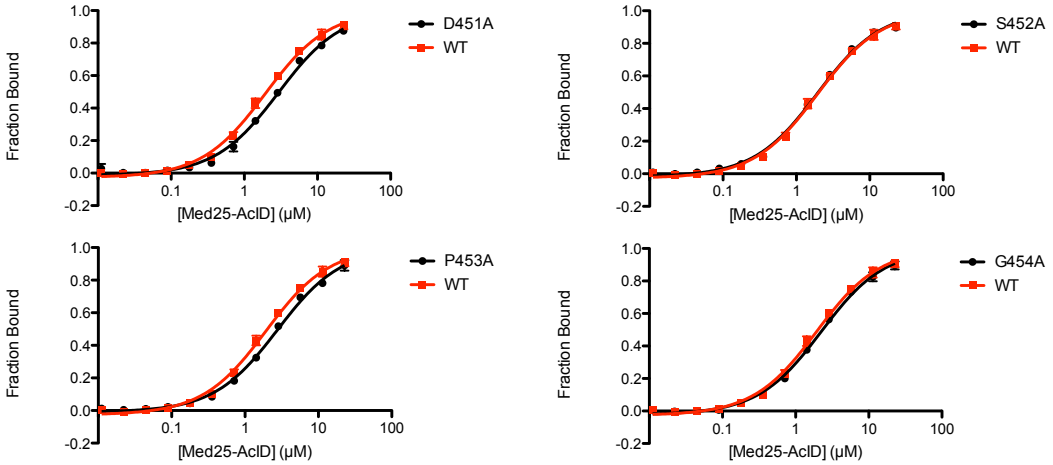


***Figure 2.8*- Alanine Scan of VP16(438-454).** (A) The sequence of VP16(438-454) is shown. Each residue was sequentially mutated to an alanine, with the exception of 438, which would generate the parent sequence. (B) Direct binding curves of each mutant compared to the wild type peptide.

***Table 2.1*- Summary of VP16(438-454) Alanine Scan Results**

Residue	Kd (μM)	Kd / Kd(WT)	ΔΔG (kcal/mol)	ΔΔG/ΔG(WT)*
A	1.92 ± 0.20	1.00	0.000	0.000
L	3.29 ± 0.28	1.71	0.319	0.041
D	5.30 ± 0.42	2.76	0.602	0.077
D	4.16 ± 0.35	2.17	0.458	0.059
F	6.92 ± 0.63	3.60	0.760	0.097
D	6.98 ± 0.63	3.64	0.765	0.098
L	7.95 ± 1.42	4.14	0.842	0.108
D	3.90 ± 0.34	2.03	0.420	0.054
M	2.28 ± 0.19	1.19	0.102	0.013
L	3.66 ± 0.24	1.91	0.382	0.049
G	1.81 ± 0.15	0.94	-0.035	-0.004
D	2.86 ± 0.28	1.49	0.236	0.030
G	2.30 ± 0.37	1.20	0.107	0.014
D	2.92 ± 0.39	1.52	0.248	0.032
S	1.93 ± 0.19	1.01	0.003	0.000
P	2.76 ± 0.26	1.44	0.215	0.028
G	2.23 ± 0.20	1.16	0.089	0.011

* ΔG(WT) = 7.8 kcal/mol

In order to be considered a hotspot, mutation of the residue in question to alanine should result in a loss of binding energy such that the ΔΔG between the wild type and alanine mutant should be greater than 1.5 kcal/mol.[30,32] None of the mutated residues within the VP16(438-454) sequence exhibits a ΔΔG greater than 0.9 kcal/mol. Thus, the VP16(438-454) interaction does not contain any specific hotspot residues, per accepted definitions, but instead relies upon a broad interaction network where many residues contribute to binding. In fact, the affinity of all of the mutated peptides were lower for AcID with the exception of G448A and S452A, with the remaining mutants demonstrating

anywhere from a 1.2-fold to 4.1-fold decrease in affinity for AcID. This finding is consistent with most activator•coactivator interactions in that it occurs with weak to moderate affinity over a broad surface area.[2] Though there may not be hot spots, mutation of F442, D443, and L444 to alanine result in the most significant decrease in affinity for AcID. In fact, comparing the $\Delta\Delta G$ values for these mutations to the ΔG for the binding of the wild-type peptide to AcID, suggests that these three residues account for greater than 30% of the total interaction energy. Thus, these residues appear to disproportionately contribute to the binding of the peptide to AcID and may constitute a critical surface within the interaction. Experiments in which all three residues are mutated to alanine may demonstrate the collective importance of these residues. Furthermore, though many of the acidic residues within the sequence appear to contribute to the affinity for AcID, D443 has nearly twice the binding energy of any of the other acidic residues, suggesting that D443 is involved in a particularly important electrostatic or hydrogen-bonding interaction.

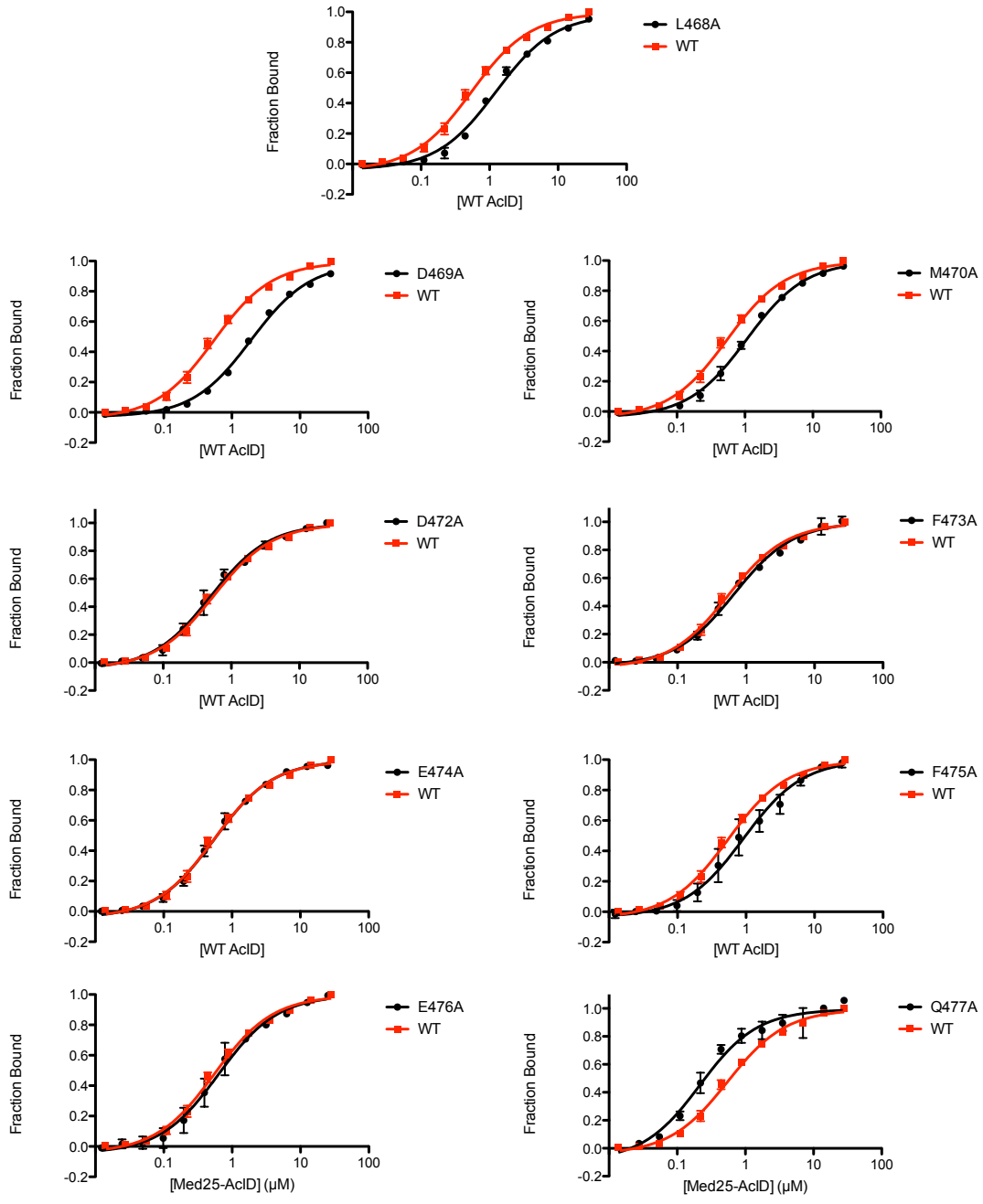
Previous mutational analysis of the VP16 H1 subdomain indicate that L439, F442, and L444 are critical for the transcriptional activity of the VP16 TAD in yeast nuclear extracts, with F442 being the most critical of the three.[19,33] Alanine is a well tolerated mutation in terms of helical propensity, indicating that observed effects on transcriptional activity or binding affinity do not result from loss of α-helical secondary structure, but are instead likely a result of a loss of critical contacts with the bulky hydrophobic side chains of these residues. While L439 does not appear to contribute significantly to affinity for AcID, the observation that F442 and L444 are among the most important contributors to the interaction with AcID is consistent with these previous findings and suggests that VP16(438-454) may generally bind to coactivator partners with a conserved structure.

In addition to VP16(438-454), an alanine scan was also completed for the sequence VP16(467-488) to determine if perhaps the H2 binding site had more specific molecular requirements than the H1 binding site, the results of which are reported in Figure 2.9 and Table 2.2.

A

467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488
A	L	D	M	A	D	F	E	F	E	Q	M	F	T	D	A	L	G	I	D	E	Y

B



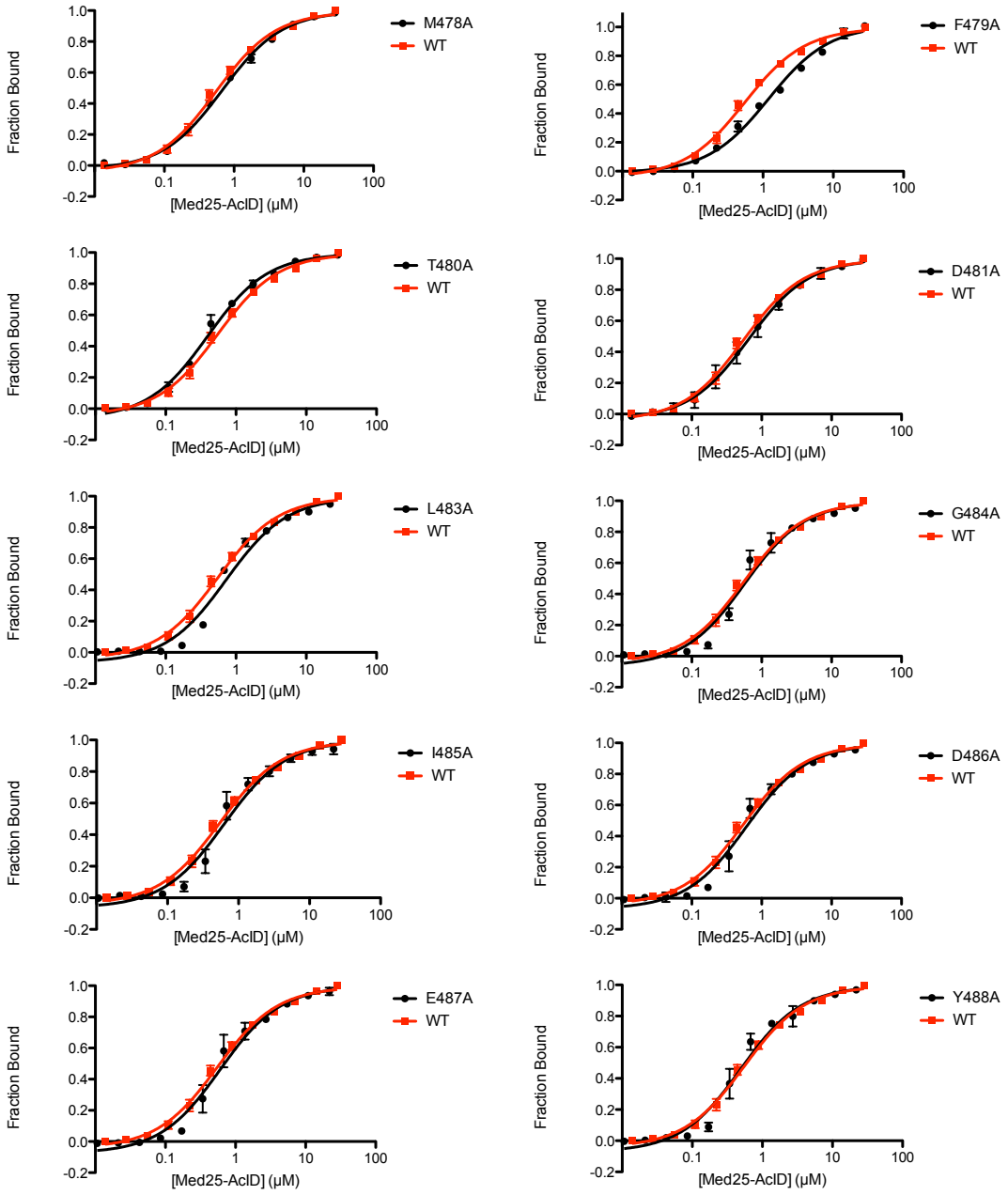


***Figure 2.9*- Alanine Scan of VP16(467-488).** (A) The sequence of VP16(467-488) is shown. Each residue was sequentially mutated to an alanine, with the exception of 467, 471, and 482 which would generate the parent sequence. (B) Direct binding curves of each mutant compared to the wild type peptide.

***Table 2.2*- Summary of VP16(467-488) Alanine Scan Results**

Residue	Kd (nM)	Kd / Kd(WT)	ΔΔG (kcal/mol)	ΔΔG/ΔG(WT)*
A	540 ± 60	1.00	0.000	0.000
L	1220 ± 170	2.26	0.483	0.056
D	1900 ± 180	3.52	0.745	0.087
M	1100 ± 110	2.04	0.422	0.049
A	-----	-----	-----	-----
D	490 ± 80	0.91	-0.058	-0.007
F	660 ± 100	1.22	0.119	0.014
E	540 ± 60	1.00	0.000	0.000
F	910 ± 220	1.69	0.309	0.036
E	620 ± 130	1.15	0.082	0.010
Q	200 ± 50	0.37	-0.588	-0.069
M	660 ± 60	1.22	0.119	0.014
F	1200 ± 180	2.22	0.473	0.055
T	380 ± 50	0.70	-0.208	-0.024
D	630 ± 100	1.17	0.091	0.011
A	-----	-----	-----	-----
L	720 ± 170	1.33	0.170	0.020
G	560 ± 140	1.04	0.022	0.003
I	620 ± 160	1.15	0.082	0.010
D	600 ± 140	1.11	0.062	0.007
E	590 ± 140	1.09	0.052	0.006
Y	480 ± 120	0.89	-0.070	-0.008

* ΔG(WT) = 8.55 kcal/mol

Similar to VP16(438-454), none of the residues in VP16(467-488) met the accepted criteria to be considered a hot-spot as the largest ΔΔG for any of the mutants was 0.7 kcal/mol. Thus, much like VP16(438-454), many of the residues within VP16(467-488) contribute weakly to binding, indicating that the interaction of this sequence also occurs over a broad surface area. Generally, the individual amino acids within this peptide appear to contribute less to the interaction energy than the residues within the VP16(438-454) peptide. This is not entirely unexpected, as the H2 interaction surface is comprised predominantly of the β-barrel, whereas the H1 interaction surface contains a more topographically defined cleft created by the interface between the β-barrel and helix three. Thus, the shallower and less defined topography of the β-barrel likely requires that most residues contribute approximately equally to binding. Similar to the VP16(438-454)•AcID interaction, a small cluster of three amino acids appear to contribute disproportionately to the interaction. Specifically, L468, D469, and M470 account for 20% of the total interaction energy after comparing the ΔΔG values for these residues to the ΔG for the binding of the wild type peptide to the domain. Of these three, the most important residue for the interaction appears to be D469, which exhibits a 3.5-fold decrease in affinity for

AcID following mutation to alanine. It is therefore possible that this residue in particular makes a critical electrostatic or hydrogen-bonding contact with AcID, as the other six acidic residues within the peptide do not demonstrate anywhere near as strong of an importance to the binding energy of the AcID interaction.

Previous mutational analysis of the VP16 TAD revealed that F473, F475, F479, and E476 were amongst the most important residues for transcriptional activation in reporter assays that utilized the VP16 TAD.[34] Of these functionally important residues, only F479 exhibits a greater than two fold decrease in affinity for AcID following mutation to alanine. Thus, these data would appear to suggest that VP16(467-488) may interact with the H2 binding site of AcID in a unique manner not previously observed for other VP16•coactivator interactions. Taken together, the hot-spot analyses of VP16(438-454) and VP16(467-488) suggest that the VP16•AcID interactions occur over broad surface areas with many residues contributing weakly to the interaction with small clusters of sequential amino acids representing 'hot-patches' that disproportionately contribute to the binding energy.

Electrostatic contacts are critical for the binding of VP16 to AcID

Based upon the highly basic character of AcID, due to the eleven lysine residues and six arginine residues within the protein, and the fact that many of the activator binding partners are highly acidic in nature, the importance of electrostatic interactions in the binding of activators to AcID was next examined. Furthermore, electrostatics have previously been demonstrated as an important factor in VP16•coactivator interactions, including interactions with PC4 and TFIIB.[15] In order to assess the importance of electrostatic interactions in VP16•AcID interactions, the affinity of VP16(465-490) was measured in the presence of various concentrations of NaCl, NaSCN, and $MgCl_2$, as shown in Figure 2.10. In electrostatically driven interactions, increasing salt concentrations leads to the shielding of charged amino acids, resulting in decreased affinity.[35,36]

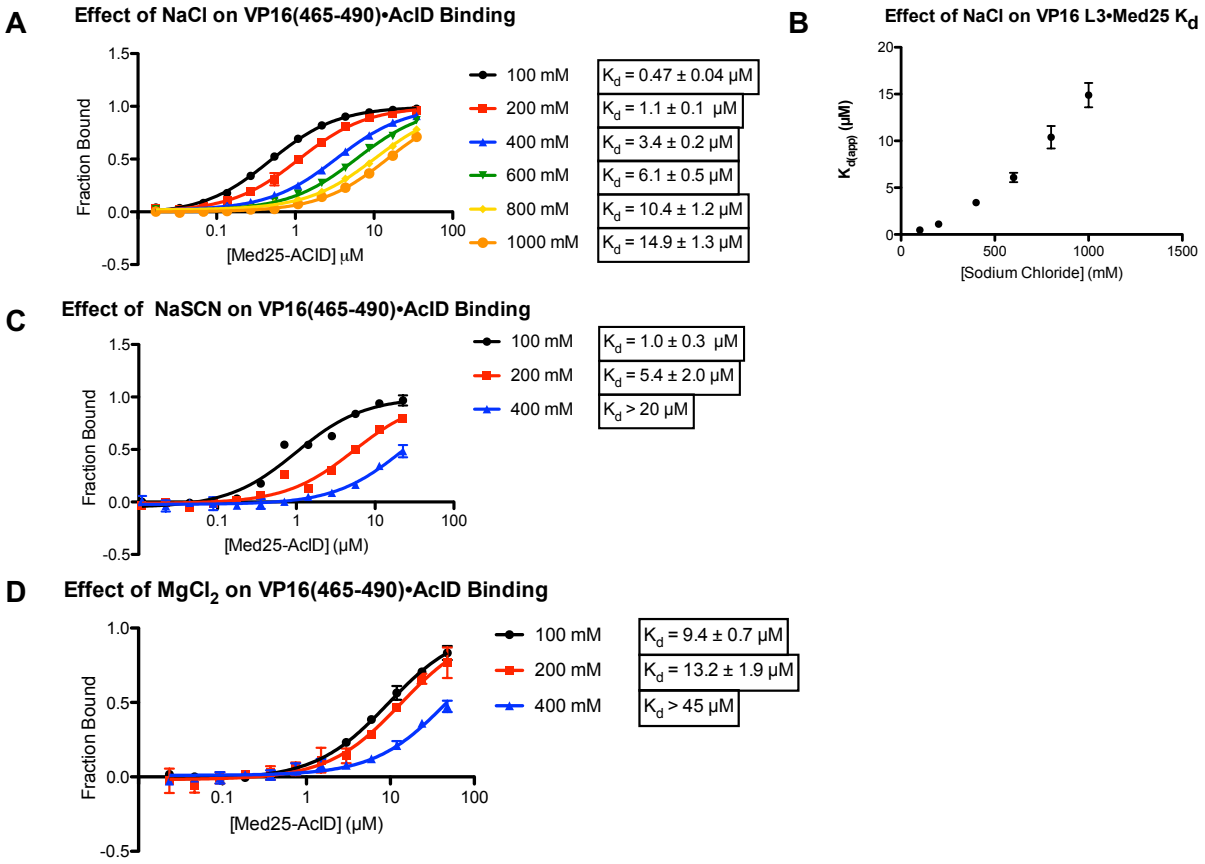


***Figure 2.10*- Effect of Salt Concentration on Direct Binding of VP16(465-490) to AcID.** (A) Affinity of VP16(465-490) was measured in the presence of various concentrations of sodium chloride. (B) Plot of apparent K_d value vs. sodium chloride concentration. (C) Affinity of VP16(465-490) was measured in the presence of various concentrations of sodium thiocyanate. (D) Affinity of VP16(465-490) was measured in the presence of various concentrations of magnesium chloride.

Increasing the concentration of NaCl in the assay buffer in which the fluorescence polarization assays were run shows a clear effect on the affinity of VP16(465-490) for AcID, with the interaction becoming weaker in the presence of higher salt concentrations. A ten-fold increase in the concentration of salt from 100 mM NaCl to 1 M salt results in a more than thirty-fold decrease in affinity of the peptide for AcID. Thus, the strong dependence of the interaction on salt concentration indicates that the VP16(465-490)•AcID interaction is strongly dependent on electrostatic interactions.

Additionally, the effects of NaSCN and $MgCl_2$ concentration on the affinity of VP16(465-490) were determined. These salts were selected because they differentially shield cationic amino acid side chains and anionic side chains, respectively. The interaction strength between salt ions and proteins is directly related to the size and

polarizability of the ions. Specifically, the positively-charged side chains of arginine and lysine residues are shielded more effectively by large, polarizable anions such as thiocyanate ($^{-}$SCN). Conversely, the negatively-charged side chains of acidic residues like aspartate and glutamate are shielded more effectively by small, less polarizable cations such as magnesium (Mg^{+2}).[37] Therefore, testing the effect of NaSCN and $MgCl_2$ on the affinity of VP16(465-490) for AcID may provide insight as to whether the anionic side chains of the TAD or the cationic side chains of AcID contribute more strongly to the interaction. In contrast to sodium chloride, both of these salts had significantly pronounced effects with 200 mM NaSCN and $MgCl_2$ resulting in a ten-fold to twenty-five-fold decrease in affinity compared to the affinity in 100 mM NaCl, further supporting the hypothesis that VP16•AcID interactions rely strongly upon electrostatic contacts. The apparent stronger effect of magnesium chloride relative to sodium thiocyanate suggests that anionic side chains may contribute more significantly to binding, though the exact molecular mechanism by which this occurs is unclear.

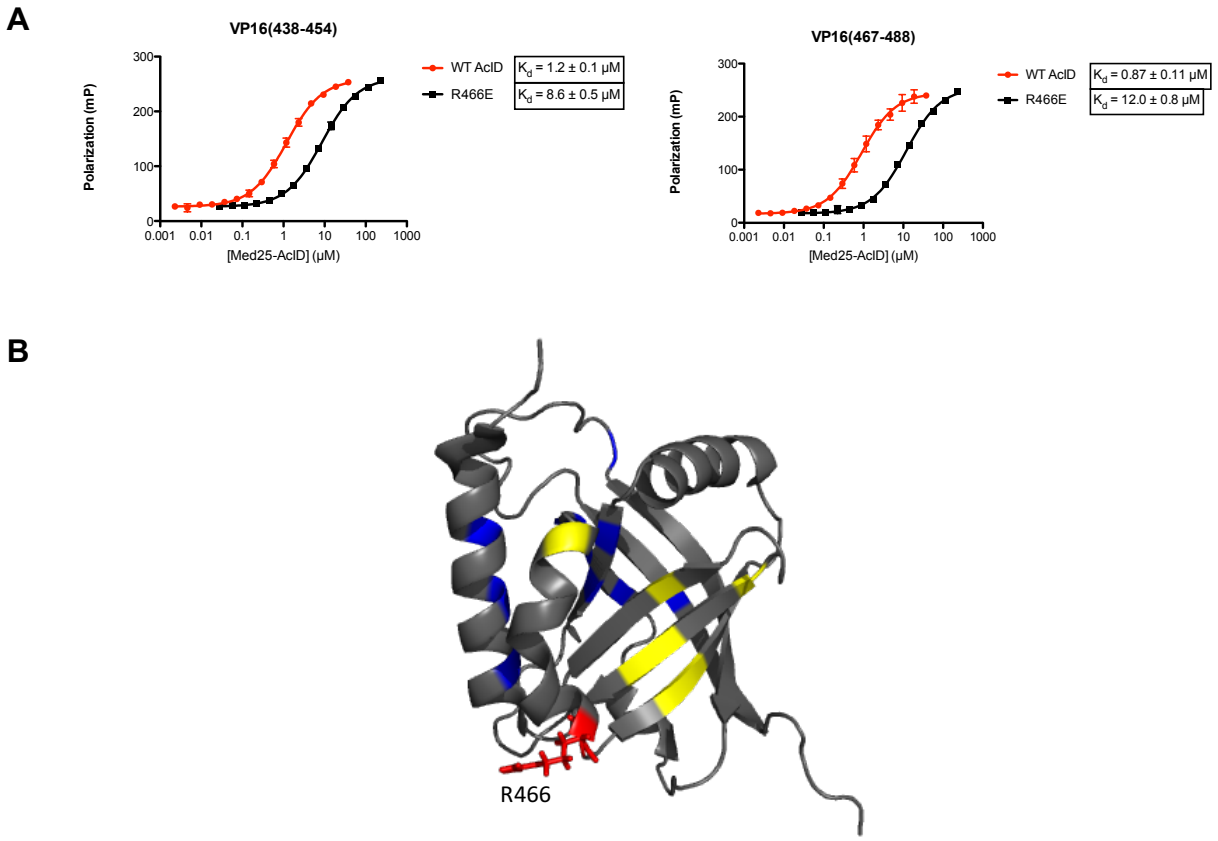


***Figure 2.11*- Effect of R466E Mutation on Binding of VP16(438-454) and (467-488)** Completed by Matthew S. Beyersdorf. (A) Direct binding of VP16(438-454) and VP16(467-488) to the AcID mutant R466E compared to the wild type protein. Curves represent the mean values of three independent experiments with error bars indicating the standard deviation of the fraction bound at each concentration of the purified AcID protein. (B) Structure of AcID with the position of R466 shown in red. The blue residues are those that are heavily perturbed following binding of VP16 H1 and the yellow residues are those that are heavily perturbed following the binding of VP16 H2.

The mutation R466E was initially reported following structural work completed by Cramer and colleagues.[12] This mutation was reported to abrogate the ability of the VP16 TAD to bind to AcID in gel supershift assays and was also incapable of squelching transcription in yeast nuclear extracts by a Gal4(DBD)-VP16(TAD) fusion protein, in contrast to the wild type protein. However, a more specific analysis of the effect on VP16 TAD affinity for this mutant was not reported. In Figure 2.11, the direct binding of VP16(438-454) and VP16(467-488) to the R466E mutant is compared to the affinity for the wild type AcID. This mutation results in an approximately 7.2-fold decrease in affinity for VP16(438-454) (8.6 μM vs. 1.2 μM) and an approximately 13.8-fold decrease in affinity for VP16(467-488) (12.0 μM vs. 0.87 μM), further underscoring the importance of electrostatic contacts between the VP16 TAD and AcID. The more pronounced effect on

the binding of VP16(467-488) is unsurprising given the proximity of R466 to the VP16 H2 interaction surface as shown in panel (B) of Figure 2.11. The charge inversion introduced by this mutation likely results in a repulsive effect that diminishes the ability of VP16(467-488) to bind to the H2 interaction surface.

The molecular cause for the impact on the binding of VP16(438-454) is less immediately apparent, but one possibility is that the charge inversion affects the tertiary structure of the domain. Though R466 is located nearer to the H2 binding site, it is near the interface between the two binding surfaces. Thus, minor changes in the orientation of helix one and helix three may affect the affinity of both VP16(438-454) and (467-488). For example, D529 is located at the bottom of helix three in close proximity to R466. Though the distance between these residues it too large to allow for formation of a salt bridge based on measurements of the various NMR structures, the charge inversion present in the mutant may distort the position of helix three, which is a significant component of the VP16 H1 interaction surface based upon HSQC NMR perturbation studies of VP16 H1 binding to AcID.[8] Further, the mutation R469E, which was selected based on its close spatial orientation to R466, had no significant effect on the affinity of VP16(438-454) or (467-488). The absence of any perturbation of the affinity for either of the VP16 derived peptides thus further supports the notion that the R466E mutation results in a perturbation of the tertiary structure of the domain. Solution of the NMR structure of this mutant or a comparative analysis of HSQC perturbations elicited following incubation with VP16 TAD would be useful in further corroborating this hypothesis.

2.4 Conclusions

The Activator Interaction Domain of Med25 contains an uncommon fold that has recently emerged as an important coactivator binding partner for a number of transcriptional activators. Previous reports suggest that the domain contains two distinct binding sites for VP16 located on opposite faces of the motif and that the TAD binds over a very large surface area estimated to be approximately 3600 $Å^2$. Serial truncation of the VP16 TAD, as reported in this chapter, revealed that short segments of the TAD could recapitulate the majority of the binding affinity, suggesting that the the critical contacts within the interaction surface occur over a significantly more concise surface than is

required for binding of the full TAD. Thus, the fact that activator•AcID interactions can be localized to relatively concise surfaces suggested that it would be possible to identify small molecule inhibitors of these interactions.

Consistent with other activator•coactivator interactions, alanine scanning mutagenesis analyses of the minimal VP16 interaction sequences revealed that these binding partners do not contain specific hotspot residues, suggesting that the interactions are dependent upon weak contacts from each of the amino acids within these peptides. This was particularly apparent for the H2 derived minimal peptide as only one of the alanine mutations resulted in a greater than two-fold decrease in affinity. Though no single residue was of critical importance for the interaction, small clusters of sequential amino acids in each of the H1 and H2 derived peptides contribute disproportionately to the affinity of the peptide for the domain, suggesting that they constitute surfaces of particular importance to the interactions. The further minimization of the binding surfaces further underscores the potential for identification of small molecule inhibitors of these interactions.

The highly acidic nature of most activator binding partners for the AcID motif and the highly basic character of the domain itself, along with previous observations of VP16•coactivator interactions, suggested that electrostatic interactions may play an important role in activator•AcID interactions. Increasing the concentration of salt within the binding assays resulted in a significant decrease in the affinity of the activator-derived peptides for the domain, with 1M NaCl resulting in a greater than thirty-fold decrease in the affinity of VP16(465-490) for the AcID motif. Thus, this suggests that electrostatic interactions play a particularly important role in the binding of activators to AcID and that small molecule inhibitors that interfere with these electrostatic contacts may be effective. Furthermore, the AcID charge inversion point mutation R466E results in a significant decrease in affinity for the VP16 TAD, suggesting that this residue may be involved in a particularly important electrostatic interaction with the VP16 TAD. The observation that this mutation affects affinity of peptide ligands for both of the AcID binding sites suggests that this residue may also be important to the secondary or tertiary structure of the domain and that the two binding sites may be allosterically connected.

Taken together, these data provide insight into the mechanisms by which we might be able to use small molecules to disrupt activator•AcID interactions. For example, molecules that disrupt critical electrostatic contacts between activators and the domain by targeting positive surfaces on AcID may be particularly effective. Additionally, given that activators do not appear to bind through the use of critical hotspot residues but instead require broad interaction surfaces, molecules that can induce allosteric structural shifts within the domain may offer an attractive mechanism for inhibiting these interactions.

Assessing the Presence of Distinct Interaction Surfaces on AcID

An alternative area of research within our laboratory is focused on further validating the hypothesis that AcID contains two distinct interaction surfaces that can independently bind activators. This study is being completed using a peptide, VP16(438-454) G450C, that covalently binds to AcID at the H1 interaction site through the formation of a disulfide between an engineered cysteine within the TAD and a native cysteine residue within the binding site, as has been previously reported for studies involving 3-phosphoinositide-dependent kinase 1.[38] This work and the findings relevant to the molecular underpinnings of the VP16•AcID interaction are summarized in this section.

The affinity of VP16(438-454) and VP16(467-488) was determined for unlabeled AcID and AcID labeled with the G450C mutant peptide using fluorescence polarization, as shown in Figure 2.12.

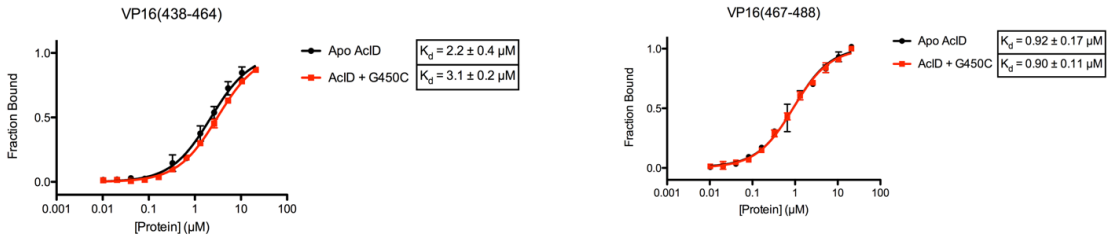


***Figure 2.12*- Direct Binding of VP16(438-454) and VP16(467-488) to the AcID•VP16 G450C Complex.** AcID was covalently labeled using disulfide exchange with VP16 mutant G450C and the affinity of VP16(438-454) and VP16(467-488) was assessed for the covalent complex and compared to unlabeled AcID. Curves represent the mean values of three independent experiments with error bars indicating the standard deviation of the polarization at each concentration of the purified AcID protein. *(Completed by Andrew R. Henderson)*

The VP16 H2 derived peptide, VP16(467-488), binds to the G450C•AcID complex with essentially no loss in affinity relative to unlabeled AcID, supporting the hypothesis that the domain contains two functionally distinct binding sites. Interestingly, VP16(438-454) binds to the G450C•AcID complex with a mere 1.4-fold decrease in affinity. One possible explanation for the lack of perturbation in the affinity of VP16(438-454) for the covalently labeled complex is that it may be capable of binding to the VP16 H2 binding site in the event that the H1 binding site is already occupied, given the similarity between the two sequences.

^{1}H,^{15}N-HSQC NMR perturbation studies of the G450C•AcID complex were also completed to identify regions perturbed by the covalent binding of the peptide, as shown in Figure 2.13.

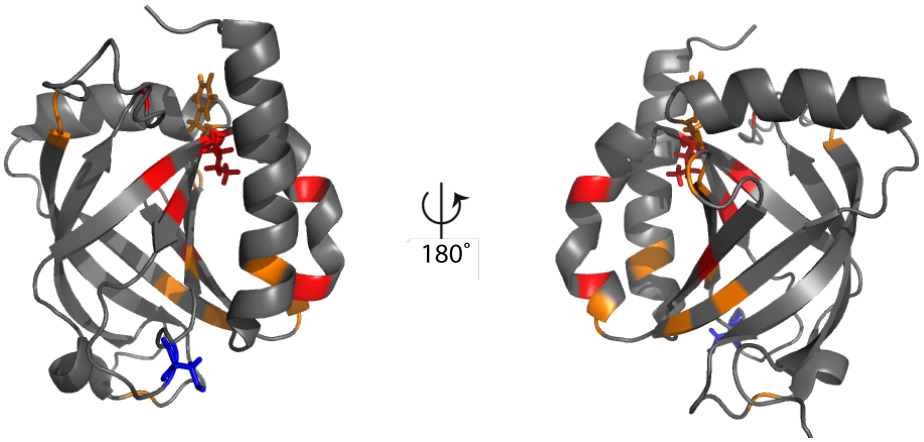


***Figure 2.13*- ^{1}H,^{15}N-HSQC NMR Perturbation of VP16 G450C•AcID Complex.** Structure of AcID with perturbations induced by labeling with the VP16(438-454) G450C peptide. Heavily perturbed (>2 standard deviations) residues are shown in red, while significantly perturbed (>1 standard deviation) residues are shown in orange. The H1 binding interface is shown on the left and the H2 binding interface on the right. *(Completed by Andrew R. Henderson and Felicia Gray)*

A number of the perturbed residues in the G450C•AcID complex are consistent with the perturbations previously reported for binding of VP16 H1 to AcID and thus represent a subset of the observed shifts for binding of the full H1 subdomain. Interestingly, a number of perturbed residues, such as I453 (red sticks in Figure 2.13) and F494 (orange sticks in Figure 2.13) are located at the interface of the H1 and H2 binding sites. Furthermore, perturbations in residues within the H2 binding site indicate that the binding of VP16(438-454) to the H1 interaction surface is capable of inducing chemical shifts in the H2 binding site. Thus, these data suggest that not only does AcID contain

two functionally distinct binding surfaces, but that these surfaces may also be within allosteric communication.

Taken together, the data presented in this section are consistent with the hypothesis that AcID contains two distinct binding sites capable of interacting with transcriptional activators. Furthermore, these data suggest that AcID may be a relatively plastic protein fold with the two binding sites potentially allosterically connected. Thus, the plasticity present within AcID following the binding of the covalent peptide to the H1 binding site suggests that it may be possible to identify small molecule inhibitors that can induce allosteric changes within the domain, providing a potential approach to disrupting the broad interaction surfaces required for activator binding.

2.5 Materials and Methods

Plasmids

Plasmid pET21b-Med25(394-543)-His_6, henceforth referred to as pAcID-His_6, was a generous gift from Patrick Cramer.[12] pAcID(R466E) was generated by Matthew S. Beyersdorf using site directed mutagenesis as previously described.[39]

Protein Expression and Purification

AcID ($Med25_{394-543}$) protein was expressed and purified together with Paul A. Bruno. Plasmid pAcID-His_6 was transformed into heat-shock competent Rosetta pLysS cells (Novagen), streaked onto LB Agar plates containing ampicillin and chloramphenicol, and incubated at 37 °C overnight. The next evening, a 25 mL Terrific Broth (TB) starter culture with 0.1 mg/mL ampicillin and 0.034 mg/mL chloramphenicol was then inoculated with a colony selected from the LB Agar plate and incubated at 37 °C overnight. The next morning, 5 mL from the starter culture was added to 1L TB containing ampicillin and bacteria were grown at 37 °C to an OD_{600} of 0.8. Temperature was reduced to 18 °C and protein expression was induced upon addition of IPTG to a final concentration of 0.5 mM. Cells were incubated overnight at 18 °C. The 1 L cultures were then collected and centrifuged at 6000xg for 20 mins at 4 °C. Cell pellets were stored at -80 °C prior to purification. The harvested pellet was thawed on ice and resuspended in 20 mL of lysis buffer (50 mM phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 6.8). Cells were then lysed by sonication on ice and cellular lysates were cleared by centrifugation at 9500 rpm for 20 min at 4 °C. The supernatant lysate was then added to Ni-NTA beads (Qiagen) and incubated for 1 hour at 4 °C. The resin was pelleted by centrifugation at 2500 rpm for 2 min at 4 °C and washed with wash buffer (50 mM phosphate, 300 mM sodium chloride, 30 mM imidazole, pH 6.8) a total of five times. Protein was then eluted with 2 mL of elution buffer (50 mM phosphate, 300 mM sodium chloride, 400 mM imidazole, pH 6.8) a total of three times. Eluent was then pooled and purified by cation exchange FPLC (Source 15S, GE Healthcare) using a gradient of Buffer B (50 mM phosphate, 100 mM NaCl, 1 mM DTT, pH 6.8) in Buffer A (50 mM phosphate, 1 mM DTT). The FPLC purified protein was then dialyzed into storage buffer (10 mM phosphate, 50 mM NaCl, 10% v/v glycerol, 0.001% v/v NP-40, pH 6.8) overnight, concentrated, aliquoted, and stored at -20 °C. Final

protein was greater than 90% pure as determined by coomassie stained polyacrylamide gel. Protein concentration was determined by UV-Vis spectroscopy using an extinction coefficient, ε = 22,460 $M^{-1}cm^{-1}$.

AcID(R466E) was expressed and purified as described above by Matthew S. Beyersdorf.

Peptide Synthesis

Peptides were synthesized on CLEAR amide resin (Peptides International) using standard HBTU/HOBT/DIEA coupling conditions as previously described.[40] TFA cleaved peptides were purified using reverse-phase HPLC (Agilent 1260) on a C_{18} Poroshell column (Agilent) using 20 mM ammonium acetate and acetonitrile as the eluent with the gradient indicated for each peptide. Following HPLC purification, fractions were pooled, lyophilized, and reconstituted in DMSO. The concentration of fluorescein labeled peptides was then determined by UV-Vis spectroscopy following 1:1000 dilution of the DMSO stock into 10 mM PBS, pH 7.4 using ε = 72,000 $M^{-1}cm^{-1}$, per the manufacturer (Pierce). Peptide identity was confirmed by electrospray mass spectrometry. Analytical HPLC spectra of all synthesized peptides can be found in **Appendix A**.

Flo-VP16(413-437) was synthesized and purified by Paul A. Bruno. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 413-437 to produce FITC-βA-APPTDVSLGDELHLDGEDVAMAHAD. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(438-464) was synthesized and purified by Paul A. Bruno. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 438-464 to produce FITC-βA-ALDDFDLDMLGDGDSPGPGFTPHDSAP. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(465-490) was synthesized and purified by Paul A. Bruno. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 465-

490 to produce FITC-βA-YGALDMADFEFEQMFTDALGIDEYGG. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(438-454) was synthesized and purified as described above. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 438-454 to produce FITC-βA-ALDDFDLDMLGDGDSPG. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(467-488) was synthesized and purified as described above. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 467-488 to produce FITC-βA-ALDMADFEFEQMFTDALGIDEY. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(472-479) was synthesized and purified as described above. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing VP16 residues 472-479 to produce FITC-βA-DFEFEQMF. A gradient of 10-30% acetonitrile over 20 min was used for HPLC purification.

Flo-ERM(38-68) was synthesized and purified as described above. Fluorescein isothiocyanate and β-alanine were coupled to a sequence containing ERM residues 38-68 to produce FITC-βA-ALDMADFEFEQMFTDALG. A gradient of 10-40% acetonitrile over 30 min was used for HPLC purification.

Flo-VP16(438-454) alanine scanning mutants were synthesized and purified as described for Flo-VP16(438-454). Fluorescein isothiocyanate and β-alanine were coupled to the sequences indicated in the table below. A gradient of 10-30% acetonitrile over thirty minutes was used for HPLC purification.

VP16(438-454) L439A:	AADDFDLDMLGDGDSPG
VP16(438-454) D440A:	ALADFDLDMLGDGDSPG
VP16(438-454) D441A:	ALDAFDLDMLGDGDSPG

VP16(438-454) F442A:	ALDDADLDMLGDGDSPG
VP16(438-454) D443A:	ALDDFALDMLGDGDSPG
VP16(438-454) L444A:	ALDDFDADMLGDGDSPG
VP16(438-454) D445A:	ALDDFDLAMLGDGDSPG
VP16(438-454) M446A:	ALDDFDLDALGDGDSPG
VP16(438-454) L447A:	ALDDFDLDMAGDGDSPG
VP16(438-454) G448A:	ALDDFDLDMLADGDSPG
VP16(438-454) D449A:	ALDDFDLDMLGAGDSPG
VP16(438-454) G450A:	ALDDFDLDMLGDADSPG
VP16(438-454) D451A:	ALDDFDLDMLGDGASPG
VP16(438-454) S452A:	ALDDFDLDMLGDGDAPG
VP16(438-454) P453A:	ALDDFDLDMLGDGDSAG
VP16(438-454) G454A:	ALDDFDLDMLGDGDSPA

Flo-VP16(467-488) alanine scanning mutants were synthesized and purified as described for Flo-VP16(467-488). Fluorescein isothiocyanate and β-alanine were coupled to the sequences indicated in the table below. A gradient of 10-40% acetonitrile over thirty minutes was used for HPLC purification.

VP16(467-488) L468A:	AADMADFEFEQMFTDALGIDEY
VP16(467-488) D469A:	ALAMADFEFEQMFTDALGIDEY
VP16(467-488) M470A:	ALDAADFEFEQMFTDALGIDEY
VP16(467-488) D472A:	ALDMAAFEFEQMFTDALGIDEY
VP16(467-488) F473A:	ALDMADAEFEQMFTDALGIDEY
VP16(467-488) E474A:	ALDMADFAFEQMFTDALGIDEY
VP16(467-488) F475A:	ALDMADFEAEQMFTDALGIDEY
VP16(467-488) E476A:	ALDMADFEFAQMFTDALGIDEY
VP16(467-488) Q477A:	ALDMADFEFEAMFTDALGIDEY
VP16(467-488) M478A:	ALDMADFEFEQAFTDALGIDEY
VP16(467-488) F479A:	ALDMADFEFEQMATDALGIDEY
VP16(467-488) T480A:	ALDMADFEFEQMFADALGIDEY
VP16(467-488) D481A:	ALDMADFEFEQMFTAALGIDEY

VP16(467-488) L483A:	ALDMADFEFEQMFTDAAGIDEY
VP16(467-488) G484A:	ALDMADFEFEQMFTDALAIDEY
VP16(467-488) I485A:	ALDMADFEFEQMFTDALGADEY
VP16(467-488) D486A:	ALDMADFEFEQMFTDALGIAEY
VP16(467-488) E487A:	ALDMADFEFEQMFTDALGIDAY
VP16(467-488) Y488A:	ALDMADFEFEQMFTDALGIDEY

Fluorescence polarization direct binding assays

Direct binding assays were performed in triplicate with a final sample volume of 20 µL in a low volume, non-binding, 384-well black plate (Corning). FITC-labeled peptides were diluted in assay buffer (10 mM PBS, 100 mM NaCl, 10 % glycerol, 0.001% NP-40 pH 6.8) to a centration of 40 nM. 10 µL of AcID protein was serially diluted two-fold on the 384-well plate for the number of data points indicated for each experiment using assay buffer. 10 µL of the diluted fluorescent peptide stock was then added to each well of diluted protein for a final tracer concentration of 20 nM. An additional well containing 10 µL buffer and 10 µL fluorescent peptide was prepared for use as a 'tracer only control' to determine optimal gain settings on the plate reader. Samples were then incubated for 30 minutes at room temperature before fluorescence polarization was measured on a Pherastar plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm. A binding isotherm that accounts for ligand depletion (assuming a 1:1 binding model of peptide to ACID) was fit to the observed polarization values as a function of ACID to obtain the apparent equilibrium dissociation, K_d:

$$y = c + (b - c) \times \frac{(K_d + a + x) - \sqrt{(K_d + a + x)^2 - 4ax}}{2a}$$

Where "a" and "x" are the total concentrations of fluorescent peptide and AcID, respectively, "y" is the observed anisotropy at a given AcID concentration, "b" is the maximum observed anisotropy value, and "c" is the minimum observed anisotropy value. Each data point is an average of three independent experiments with the indicated error

representing the standard deviation of the three replicates. All curves and calculations were generated using GraphPad Prism 5.

Hotspot Analysis

Direct binding experiments for each of the alanine mutants of VP16(438-454) and VP16(467-488) were completed as described above. The calculated K_d value for each mutant was then compared to the calculated K_d value for the wild-type peptide in order to calculate the $\Delta\Delta G$ of each residue using the following equation:

$$\Delta\Delta G = RT \ln \frac{K_{d_{mut}}}{K_{d_{WT}}}$$

Where, "R" is the ideal gas constant in units of $kcalK^{-1}mol^{-1}$, "T" is the temperature in kelvin, "Kd_{WT}" is the calculated K_d value of the WT peptide binding to AcID, and "Kd_{mut}" is the calculated K_d value of the indicated alanine mutant.

Salt Titration Direct Binding Experiments

The direct binding experiments were performed as noted above with several modifications. The salt concentration of assay buffer was adjusted to the indicated concentrations by adding an appropriate volume of assay buffer containing 5 M of the indicated salt prior to adding components to the 384-well plate. For NaSCN and $MgCl_2$ titrations, AcID protein was dialyzed into storage buffer containing 100 mM of the appropriate salt in place of NaCl. All curves and calculations were generated using GraphPad Prism 5.

2.6 References

1. Gann, M. P. A. A. Genes and Signals. 1–209 (2005).

2. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chemical Biology* **7,** 1311–1320 (2012).

3. Myers, L. C. & Kornberg, R. D. Mediator of transcriptional regulation. *Annu. Rev. Biochem.* **69,** 729–749 (2000).

4. Bourbon, H.-M. Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.* **36,** 3993–4008 (2008).

5. Mittler, G. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* **22,** 6494–6504 (2003).

6. Lee, H.-K. H., Park, U.-H. U., Kim, E.-J. E. & Um, S.-J. S. MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* **26,** 3545–3557 (2007).

7. Shi, Z., Yang, W., Goldstein, J. A. & Zhang, S.-Y. Med25 is required for estrogen receptor alpha (ERα)-mediated regulation of human CYP2C9 expression. *Biochem. Pharmacol.* **90,** 425–431 (2014).

8. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **18,** 410–415 (2011).

9. Verger, A. *et al.* The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41,** 4847–4859 (2013).

10. Sela, D. *et al.* Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6α. *J. Biol. Chem.* **288,** 26179–26187 (2013).

11. Malik, S. & Roeder, R. G. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30,** 256–263 (2005).

12. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18,** 404–409 (2011).

13. Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *Journal of Structural Biology* **174,** 245–251 (2011).

14. Freedman, S. J. *et al.* Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2. *Nat. Struct. Biol.* **10,** 504–512 (2003).

15. Jonker, H. R. A., Wechselberger, R. W., Boelens, R., Folkers, G. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44,** 827–839 (2005).

16. Yang, F. *et al.* An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442,** 700–704 (2006).

17. Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* **20,** 1455–1464 (2001).

18. Wysocka, J. & Herr, W. The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. *Trends Biochem. Sci.* **28,** 294–304 (2003).

19. Regier, J. L., Shen, F. & Triezenberg, S. J. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc Natl Acad Sci USA* **90,** 883–887 (1993).

20. Walker, S., Greaves, R. & O'Hare, P. Transcriptional activation by the acidic domain of Vmw65 requires the integrity of the domain and involves additional determinants distinct from those necessary for TFIIB binding. *Molecular and Cellular Biology* **13,** 5233–5244 (1993).

21. Ikeda, K., Stuehler, T. & Meisterernst, M. The H1 and H2 regions of the activation domain of herpes simplex virion protein 16 stimulate transcription through distinct molecular mechanisms. *Genes Cells* **7,** 49–58 (2002).

22. Näär, A. M. *et al.* Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature* **398,** 828–832 (1999).

23. Yang, F., DeBeaumont, R., Zhou, S. & Näär, A. M. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc Natl Acad Sci USA* **101,** 2339–2344 (2004).

24. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277,** 1310–1313 (1997).

25. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *Journal of Biological Chemistry* **277,** 43168–43174 (2002).

26. Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry* **48,** 2115–2124 (2009).

27. Zor, T., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337,** 521–534 (2004).

28. Parker, D. *et al.* Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular cell* **2,** 353–359 (1998).

29. Landrieu, I. *et al.* Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **43,** 7110–7121 (2015).

30. Golden, M. S. *et al.* Comprehensive experimental and computational analysis of binding energy hot spots at the NF-κB essential modulator/IKKβ protein-protein interface. *J. Am. Chem. Soc.* **135,** 6242–6256 (2013).

31. Zerbe, B. S., Hall, D. R., Vajda, S., Whitty, A. & Kozakov, D. Relationship between hot spot residues and ligand binding hot spots in protein-protein interfaces. *J Chem Inf Model* **52,** 2236–2244 (2012).

32. Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267,** 383–386 (1995).

33. Cress, W. D. & Triezenberg, S. J. Critical structural elements of the VP16 transcriptional activation domain. *Science* **251,** 87–90 (1991).

34. Sullivan, S. M. *et al.* Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Res.* **26,** 4487–4496 (1998).

35. Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M.-F. & Ahmadian, M. R. An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Molecular cell* **20,** 313–324 (2005).

36. Perez-Jimenez, R., Godoy-Ruiz, R., Ibarra-Molero, B. & Sanchez-Ruiz, J. M. The efficiency of different salts to screen charge interactions in proteins: a Hofmeister effect? *Biophys. J.* **86,** 2414–2429 (2004).

37. Zhang, J. *Protein-Protein Interactions in Salt Solutions*. (2012).

38. Sadowsky, J. D. *et al.* Turning a protein kinase on or off from a single allosteric site via disulfide trapping. *Proceedings of the National Academy of Sciences* **108,** 6056–6061 (2011).

39. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chemical Biology* **7,** 1345–1350 (2012).

40. Majmudar, C. Y., Wang, B., Lum, J. K., Håkansson, K. & Mapp, A. K. A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew. Chem. Int. Ed.* **48,** 7021–7024 (2009).

CHAPTER 3

Depside and Depsidone Inhibitors of Activator•Med25 Interactions[3]

3.1 Abstract

The central hypothesis guiding the work described in this chapter is that the AcID motif of Med25 is amenable to the development of selective small molecule inhibitors that perturb the binding of activators to the domain. The results discussed previously in Chapter 2 suggest that small molecules that allosterically alter the conformation of the protein or disrupt critical electrostatic contacts between activators and the AcID motif will be particularly effective. In order to identify these inhibitors, a screen of biologically active small molecules was screened using a fluorescence polarization based binding assay and hits were subsequently filtered to eliminate non-selective inhibitors of AcID, ultimately providing lead molecules that belong to the depside and depsidone classes of natural products. The most potent lead molecule, norstictic acid, is a covalent inhibitor of AcID and reacts with reactive lysine side chains near surfaces important for the binding of transcriptional activators. Interestingly, the binding of the inhibitor to one activator binding site is capable of simultaneously perturbing binding at the second activator binding site, consistent with a model in which norstictic acid functions as a mixed orthosteric/allosteric inhibitor. Finally, preliminary evidence supports a mechanistic model in which norstictic

[3] The research described in this chapter was a collaborative effort. Steven M. Sturlis and Paul A. Bruno synthesized and purified fluorescent tracers, expressed and purified protein, optimized the FP HTS assay, performed the primary screen, and analyzed covalent adducts by mass spectrometry. Paul A. Bruno completed the cellular assays, prepared samples for HSQC analysis, and expressed Med15 protein. Jessica Gagnon (Brooks Lab, University of Michigan) completed MD simulations of norstictic acid in complex with AcID. Felicia Gray (Cierpicki Lab, University of Michigan) performed HSQC analysis of AcID•norstictic acid adducts. Giselle Tamayo-Castillo (INBio) provided additional samples of norstictic acid. A manuscript is in preparation based on portions of this chapter: Sturlis, Steven M.*; Bruno, Paul A.*; Gray, Felicia; Gagnon, Jessica K.; Tamayo-Castillo, Giselle; Sherman, David H.; and Mapp, Anna K. Identifying Depside and Depsidone Inhibitors of Med25 AcID-mediated Transcription. *Manuscript in preparation.*

acid and the closely related molecule psoromic acid are capable of perturbing AcID-dependent transcriptional processes in a cellular context.

3.2 Introduction

Given the fundamental role that transcription plays in all cellular processes and the fact that transcriptional dysregulation has been linked to virtually all human diseases as either a cause or consequence, small molecule inhibitors of transcriptional processes would be remarkably useful as mechanistic probes or novel therapeutic agents.[1-4] Transcription is an inherently difficult process to target; this is largely due to the fact that, with the notable exception of nuclear receptors, transcriptional activators do not require the binding of small molecules for their function. The primary avenue by which they can be targeted involves disrupting the dynamic network of protein-protein interactions that govern the assembly of the large multimeric protein complexes required for the expression of target genes.[5] Targeting transcription is further complicated by the characteristics of the majority of the protein-protein interactions involved in transcriptional regulation. Typically they occur over broad surface areas with weak to moderate affinity and defy hotspot analysis, leading to their frequent characterization as 'undruggable'.[6] Despite these inherent challenges, transcriptional inhibitors are attractive as mechanistic probes and/or as therapeutic agents as the disruption of transcriptionally relevant protein-protein interactions offers an unparalleled opportunity to achieve context specificity, as discussed in Chapter 1. Furthermore, recent advances in discovery strategies for protein-protein interaction inhibitors have made the task of developing transcriptional modulators more readily achievable.[7] In particular, protein-protein interactions between activators and coactivators are attractive targets as their disruption will allow for the inhibition of specific activator-driven gene expression programs.

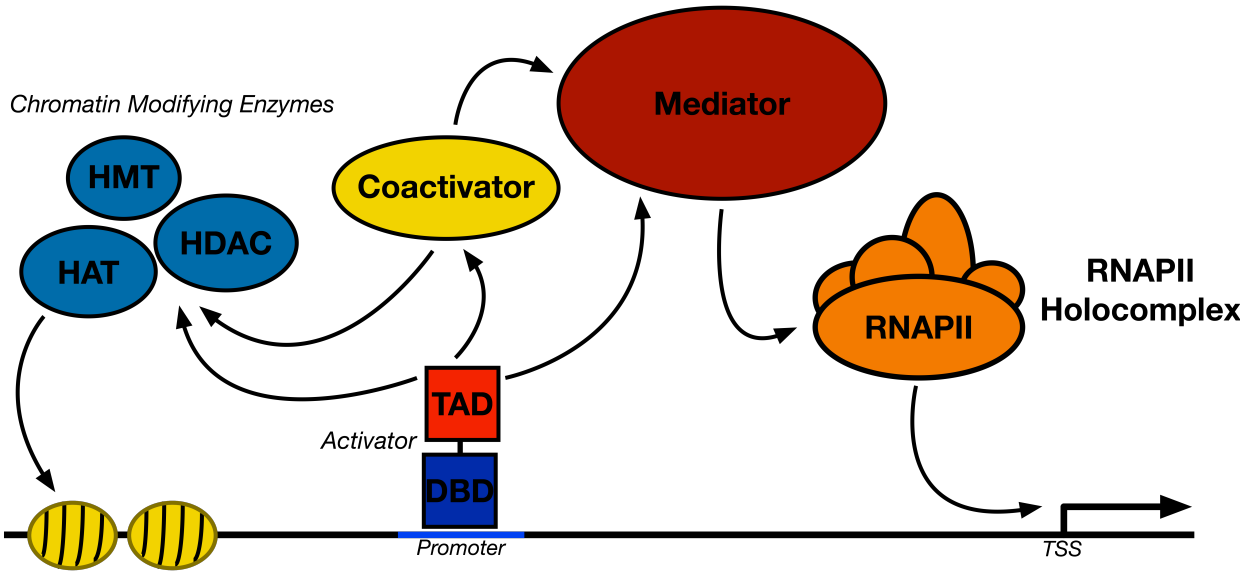


Figure 3.1- Transcription Requires the Assembly of Large Multimeric Protein Complexes Activators bind to DNA through sequence-specific interactions between their DNA binding domains (DBD) and elements within the promoter regions of target genes. The transcriptional activation domains (TAD) of these activators in turn make critical contacts with other elements of the transcriptional machinery including, but not limited to: chromatin modifying enzymes or complexes that alter the local chromatin structure, enabling or suppressing transcription of the target gene[8,9]; the Mediator complex, which bridges the DNA-bound activator to RNA Polymerase II and the general transcription factors[10,11]; or alternative coactivator proteins that may function as 'bridging' proteins that facilitate interactions between the activator and other elements of the transcriptional machinery.[12]

Success in Targeting Transcriptional Coactivators

Despite the significant challenges, a number of small molecule transcriptional modulators have been identified in the past decade based upon design strategies that mimic the structure of transcriptional activation domains.[7,13-16] These small molecules inhibit activator•coactivator interactions by binding to target coactivators at surfaces required for interaction with native activator binding partners. The first reported example of this strategy was a series of compounds built upon an isoxazolidine core to which various functional groups such as hydroxyl, phenyl, or isobutyl could be appended. These molecules mimic critical residues of transcriptional activation domains and have been shown to be capable of activating transcription to comparable levels as a minimal sequence of the potent VP16 TAD following their localization to DNA.[17] Subsequent experiments, including pull-down experiments from cellular lysates with biotinylated molecule and fluorescence polarization assays with purified domains, revealed that one of these isoxazolidines, **iTAD 1**, was capable of binding to the KIX domain of the master

coactivator CBP. Furthermore, $^{1}H,^{15}N$-HSQC chemical shift perturbation studies revealed that the binding of iTAD 1 to the KIX domain induced similar shifts as the binding of native activator binding partners, suggesting that the molecules bind to similar surfaces within the domain.[18]

The ability of isoxazolidine TAD mimetics to inhibit activator•coactivator interactions by competing for necessary coactivator binding surfaces has also been investigated. ErbB2 is a surface receptor tyrosine kinase, the overexpression of which leads to an enhanced metastatic phenotype and resistance to chemotherapies in breast cancer. [19-21] The transcriptional activator ESX has been demonstrated as a central regulator in the expression of ErbB2 and is dependent upon interactions with Mediator subunit Med23 for its transcriptional activity. [22,23] Furthermore, inhibition of the ESX•Med23 complex has been shown to reduce the proliferative capacity of ErbB2 overexpressing breast cancer cell lines, underscoring the role of ESX and ErbB2 in breast cancer.[22,24] Consistent with these prior reports, a biphenyl containing isoxazolidine similar in structure to iTAD 1 was capable of inhibiting the formation of the ESX•Med23 and led to a decrease in the expression of ErbB2 and a decrease in the cellular proliferation of ErbB2 dependent breast cancer cell lines such as SKBR3.[14] Interestingly, the unmodified iTAD 1 molecule exhibited no effect against ErbB2 expression or the proliferation of SKBR3 cells, presumably as a result of its ten fold lower affinity for Med23 relative to the affinity of the biphenyl containing molecule, demonstrating that this class of transcriptional inhibitor is tunable and dependent upon specific functional groups for their activity. Recently, the biphenyl isoxazolidine ESX inhibitor was shown to act synergistically with afatinib, an irreversible EGFR/ErbB2 tyrosine kinase inhibitor, in an *in vivo* mouse model of head and neck squamous cell carcinoma, underscoring the potential utility of these transcriptional inhibitors.[25]

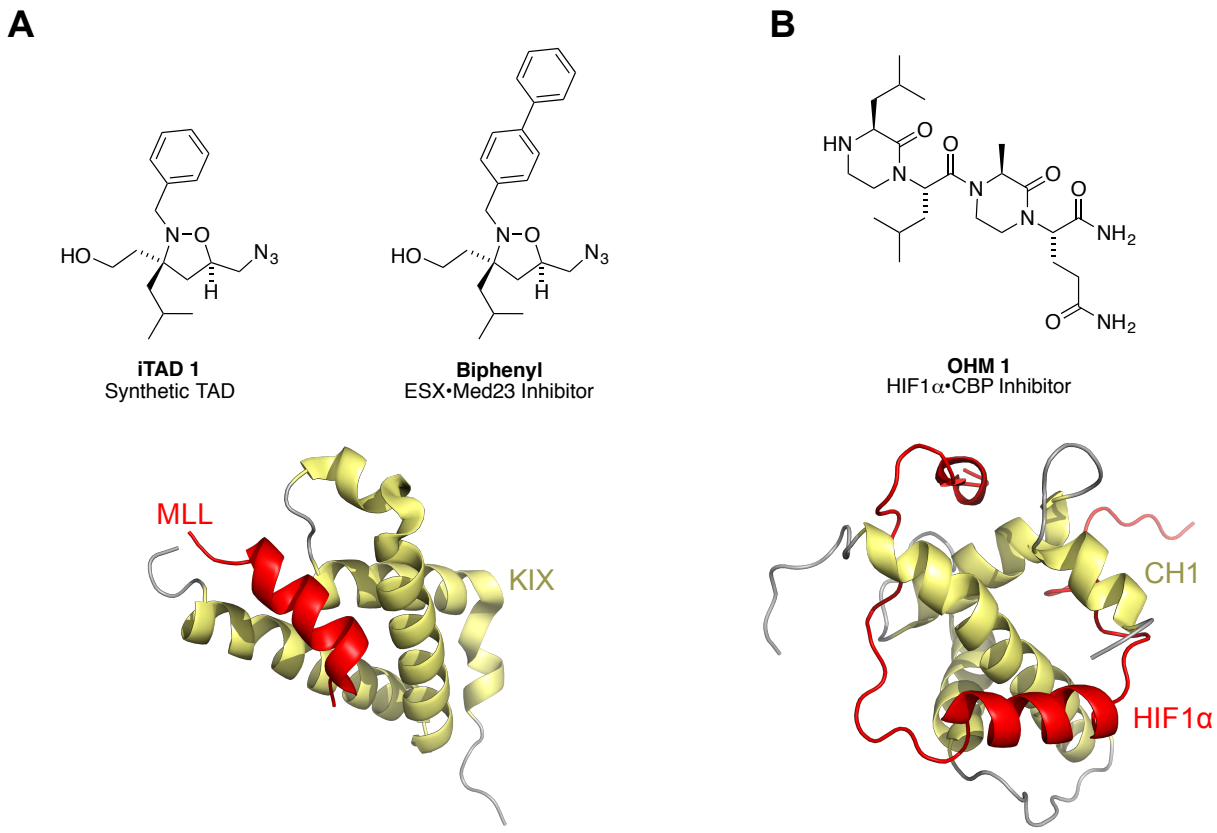


Figure 3.2- Examples of TAD Mimetic Transcriptional Inhibitors (A) Examples of molecules belonging to the isoxazolidine class of transcriptional inhibitors. iTAD 1 functions as a synthetic activation domain upon localization to DNA by binding to the KIX domain of CBP/p300. HSQC data indicates that the molecule binds to the surface of KIX (yellow) utilized by the transcription factor MLL (red) (PDB ID: 2AGH). Biphenyl is an inhibitor of the ESX•Med23 interaction and has been demonstrated to downregulate the expression of ESX target genes, such as ErbB2. (B) Example of an oligooxopiperazine helix mimetic that has been shown to inhibit the interaction between HIF1α (red) and the CH1 domain of CBP (yellow), thereby downregulating the expression of HIF1α dependent genes (PDB ID: 1P4Q).

Another synthetic scaffold that has been effectively exploited in the development of activator TAD mimetics is derived from the functionalization and oligomerization of oxopiperazines. These molecules, termed oxopiperazine helix mimetics (OHM), offer significant advantages over similar molecules that utilize aromatic scaffolds as the oxopiperazine backbone is chiral, positioning side chains more effectively in three dimensional space allowing for more efficient inhibition of three dimensional binding surfaces with greater specificity. Furthermore, an oxopiperazine dimer approximately spans the length of an eight-mer helix, positioning side-chain-like functionalities in a similar spatial orientation as the i, i+4, and i+7 positions within α-helices. Biophysical

methods including circular dichroism, COSY NMR spectroscopy, and NOESY NMR spectroscopy confirm that the OHMs adopt α-helical like geometries.[13]

Recently, the oxopiperazine helix mimetic **OHM 1** has been shown to inhibit hypoxia inducible factor 1 (HIF1α) regulated gene expression and has been show to reduce tumor growth rates in mouse MDA-MB-231 xenograft models.[26] Under hypoxic conditions that frequently accompany the growth of solid tumors the expression of HIF1α is induced, resulting in the upregulation of transcriptional programs that drive angiogenesis and contribute to invasion and altered energy metabolism in cancer.[27] Transcriptional activation by HIF1α requires an interaction with the cysteine-histidine rich 1 (CH1) domain of the master coactivator CBP/p300. [28,29] OHM 1 effectively inhibits the HIF1α•CH1 interaction by mimicking critical hot-spot residues located within an eight-mer α-helix of the C-terminal TAD of HIF1α required for interaction with CH1.[26] Thus, isoxazolidine and oxopiperazine derived TAD mimetics demonstrate that synthetic small molecules are capable of inhibiting activator•coactivator interactions and that these small molecules are useful as mechanistic probes, even in *in vivo* contexts.

More recently developed strategies have sought to inhibit interactions between activators and coactivators through small molecules that do not merely mimic structural features of transcriptional activation domains, but instead induce structural shifts that interfere with protein-protein interactions in target coactivators through allosteric mechanisms. An example is the disulfide-containing small molecule fragment **1-10**, which was originally identified in a Tethering screen of the KIX domain of CBP/p300.[30] The KIX domain contains two distinct binding sites for transcriptional activators that are within allosteric communication. More specifically, the binding of an activator, such as MLL, to the domain induces allosteric shifts within the protein that enhance the binding of a second ligand at an alternative site, such as the pKID domain of the activator CREB. [31,32] The binding of 1-10 to an engineered cysteine mutation within the MLL binding site has been shown to significantly inhibit the binding of MLL to KIX through orthosteric inhibition, while also inhibiting the binding of pKID to the secondary site by inducing allosteric changes within the domain. [30,33] Though this inhibition was a modest two-fold decrease in affinity for the pKID domain, 1-10 serves as a useful example of a small molecule that

can induce allosteric changes within a coactivator that modulates its ability to bind a native activator partner.

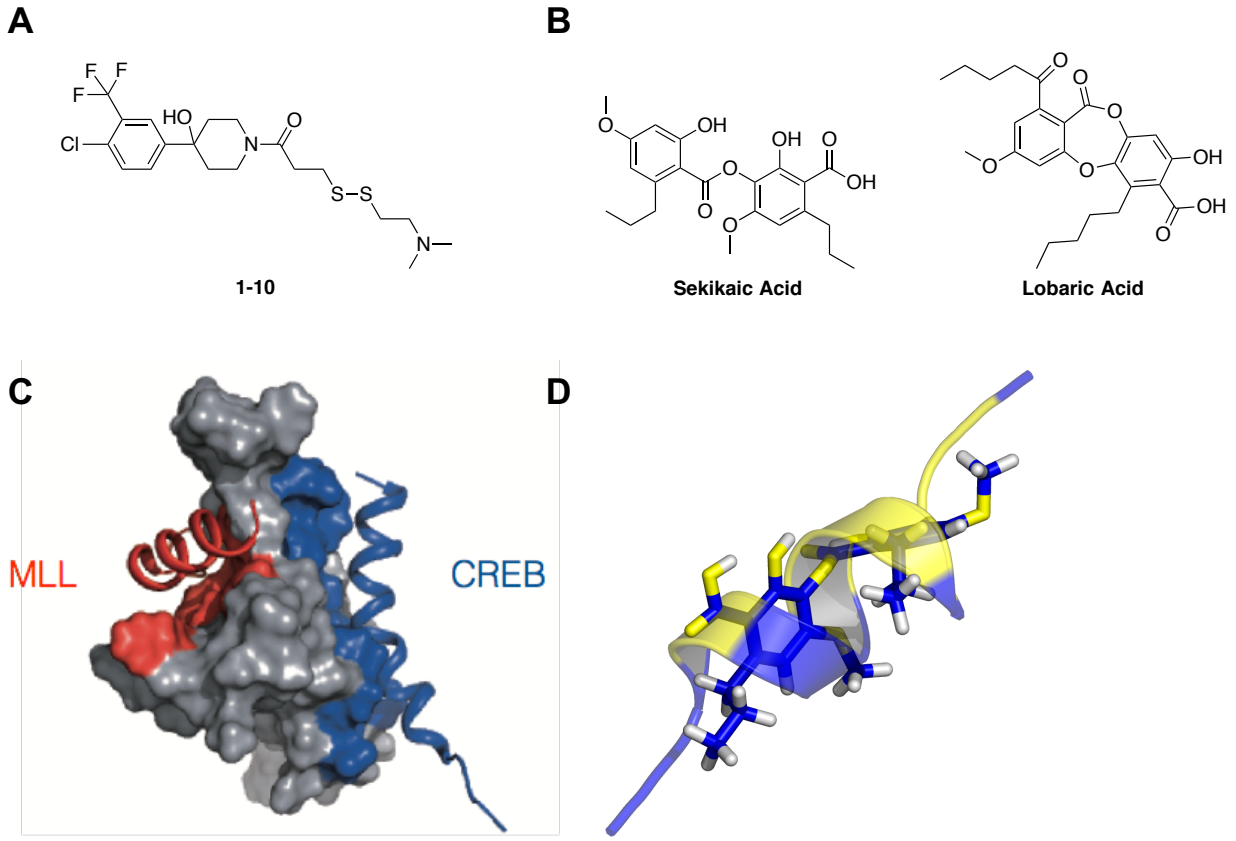


Figure 3.3- Examples of Allosteric Transcriptional Inhibitors (A) 1-10 is a small molecule fragment that binds to an engineered cysteine mutant of the KIX domain through disulfide exchange. (B) Sekikaic acid and lobaric acid are small molecules belonging to the depside and depsidone class of lichen-derived natural products. (C) The KIX domain of CBP/p300 with the MLL binding site indicated in red and the binding site for the pKID domain of the activator CREB indicated in blue. (D) Molecular dynamics simulations suggest that depsides such as sekikaic acid can adopt α-helical like conformations. Sekikaic acid is shown overlaid with an amphipathic helix of the p53 TAD.

The lichen-derived natural products sekikaic acid and lobaric acid were recently discovered as inhibitors of the CBP/p300 KIX domain.[34] These molecules were identified by a fluorescence polarization-based high-throughput screen of the MLL•KIX interaction, which occurs at a flexible interface of the domain (**Figure 3.3 C**). Screening of a 50,000 compound small molecule library failed to identify any potential lead molecules, but subsequent screening of a 15,000-sample natural product extract library identified sixty-four extracts with inhibitory activity. This observation underscores the potential of natural products as inhibitors of protein-protein interactions as they tend to be more complex and three-dimensional than 'drug-like' molecules typical of commercially available screening

libraries. An iterative screening strategy in which these extracts were then tested against two other protein-protein interactions and one protein-DNA interaction identified two extracts with specificity for the KIX domain. Compound isolation and characterization identified sekikaic acid as the most abundant component of these extracts with lobaric acid subsequently identified as a result of its structural similarities. These compounds were then tested against the pKID•KIX interaction, which occurs at a second broader and shallower binding site on the KIX domain, and were found to inhibit this interaction in addition to the MLL•KIX interaction. Subsequent biophysical characterization demonstrated that the inhibition of the pKID•KIX interaction was achieved through allosteric shifts within the protein following the binding of the molecule to the MLL interaction site. Molecular dynamics simulations of sekikaic acid suggests that the structure adopts a three-dimensional orientation similar to an α-helix with the aryl substituent arrangement resulting in amphipathic character, as shown in Figure 3.3 D. Taken together, these observations demonstrate for the first time that small molecules can inhibit activator•KIX interactions at both activator binding sites through a mixed orthosteric/allosteric mode of action.

The above examples of small molecule transcriptional inhibitors demonstrate that the inhibition of activator•coactivator interactions can be achieved by developing small molecules that: (1) compete with native activator binding partners for critical binding surfaces within coactivators or (2) by inducing allosteric changes within the activator binding domains of coactivators that alter critical activator binding surfaces.

The Activator Interaction Domain of Med25 as a Target for Small Molecule Inhibitors

The Activator Interaction Domain (AcID) of the Mediator subunit, Med25, contains a unique protein fold consisting of a seven-stranded β-barrel flanked by three α-helices, as shown in Figure 3.4 A. [35-37] This particular fold has not been previously observed in any other known coactivator protein and has been identified in only one other protein of unknown function that is overexpressed in cancerous prostate tissue. [35,36,38] Known activator binding partners of AcID include the PEA3 subfamily of ETS transcription factors that have been implicated in tumor progression and metastatic phenotypes in breast and prostate cancer[39-41], the endoplasmic reticulum stress response transcription factor

ATF6α[42], and the potent viral activator VP16 that is responsible for the functional switch from latent to lytic infection in herpes simplex virus infections. [35,36,43,44]

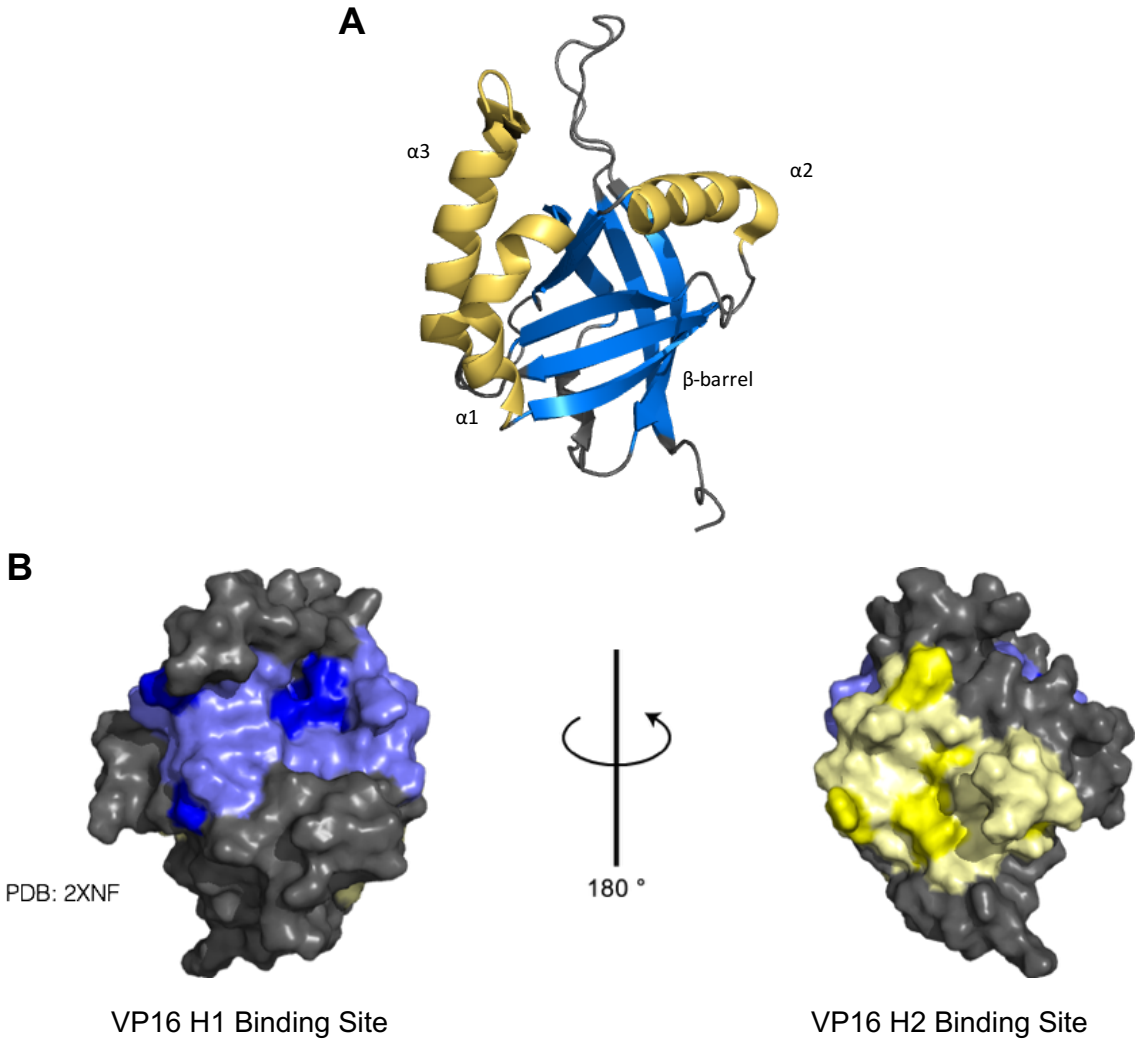


Figure 3.4- The Activator Interaction Domain (AcID) of Mediator Subunit 25 (A) A cartoon schematic of the AcID motif reveals that the unique fold consists of a seven-stranded β-barrel flanked by three α-helices. (B) $^{1}H,^{15}N$-HSQC NMR chemical shift perturbation studies of the binding of the VP16 TAD to AcID reveals two distinct binding sites for the H1 and H2 subdomains of the TAD.

Recent NMR studies demonstrated that the VP16 TAD binds to AcID over a broad surface with individual titration of the H1 or H2 subdomains binding to distinct interaction surfaces on opposite faces of the protein, as shown in panel B of Figure 3.4. [35,36] Results discussed in Chapter 2 suggest that the VP16•AcID interaction is predominantly dependent upon reported α-helices within the VP16 TAD and that the individual VP16 H1 and H2 binding sites within AcID may be capable of allosteric communication. Thus, the AcID motif of Med25 is an attractive target for the development of small molecule

inhibitors given that: (1) the unique nature of the domain's fold should allow for the discovery of highly selective small molecules with few off target effects, (2) previous success in targeting α-helix mediated activator•coactivator interactions has been achieved, and (3) development of small molecule inhibitors of activator•AcID interactions would be useful in answering mechanistic questions about this poorly understood domain, could be used to validate future discoveries of activator binding partners for AcID, and could serve as potential leads in novel therapeutic strategies for diseases linked to AcID dependent transcriptional activators.

Given the broad surface area over which activators interact with the AcID motif and the conformational plasticity within the domain, particularly the allosteric communication between the two putative binding sites, inhibitors that function through an allosteric mechanism of action will likely be particularly effective at disrupting activator•AcID interactions. In order to select lead compounds with this activity, hits identified in the primary screen are tested for inhibition against functionally related activator•coactivator interactions. This filter eliminates compounds that merely mimic TADs or amphipathic helices, resulting in lead compounds with excellent selectivity profiles for the target of interest and likely function through allosteric mechanisms. This screening strategy is discussed in the context of a larger screening campaign in Chapter 4.

3.3 Results and Discussion

VP16 is perhaps the best studied example of the amphipathic acidic class of activators and given its demonstrated interaction with the Activator Interaction Domain (AcID) of Med25, serves as a useful binding partner for the identification of inhibitors of activator•AcID interactions.[35-37,44-46] One of the primary concerns with small molecule inhibitors is that non-specific inhibition of alternative targets may result in significant off-target effects that preclude their use as mechanistic probes or potential lead compounds in novel therapeutic strategies. In order to combat this potential issue in the context of activator•coactivator interaction inhibitors, a fluorescence polarization based assay adapted to a high-throughput format and a series of counter-screens that will test the ability of identified hits to inhibit unrelated and related activator•coactivator interactions will be employed. In this manner, hits will be filtered to remove those compounds that

merely mimic the VP16 transcriptional activation domain or those that non-specifically block transcriptionally relevant protein-protein interactions. Our lab has previously utilized this strategy to identify inhibitors of other activator•coactivator interactions.[34]

High-Throughput Screen Assay Development

As was previously discussed at length in Chapter 2, a truncation study of the VP16 TAD was completed to identify specific elements of the domain that contribute strongly to interaction with the AcID motif of Med25. This data is presented again in Figure 3.5 for the sake of clarity.

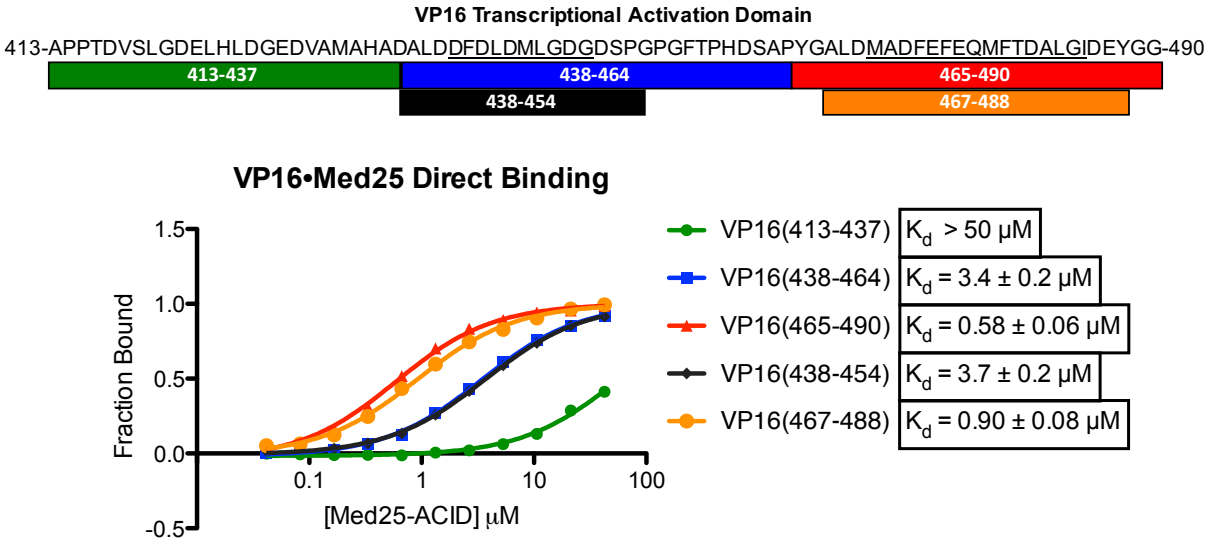


***Figure 3.5*- Fluorescent Tracers Derived from the Truncation of the VP16 TAD** The VP16 TAD was truncated into three peptides of approximately equally length and an additional two peptides were synthesized based on purported α-helices within the TAD (underlined sequences within the TAD). The peptides were conjugated to fluorescein and the K_d of each peptide was determined for AcID using fluorescence polarization. Curves represent the means of three independent experiments with error bars representing the standard deviation of the fraction bound at the indicated concentration of AcID protein. Curves were fit using GraphPad Prism 5. *(With Paul A. Bruno)*

VP16(465-490) was selected for high-throughput screening because it is the sequence that binds to the domain with the highest affinity, suggesting that this particular TAD sequence is critical for interaction with the domain, and the dynamic range for binding of the tracer to AcID spans approximately 250 mA units.

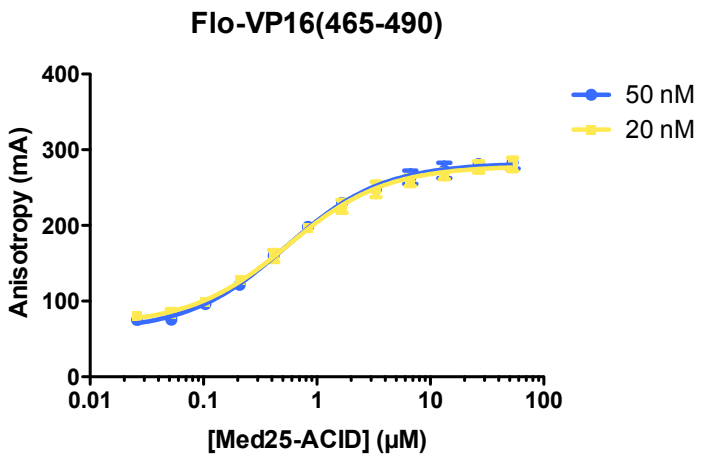
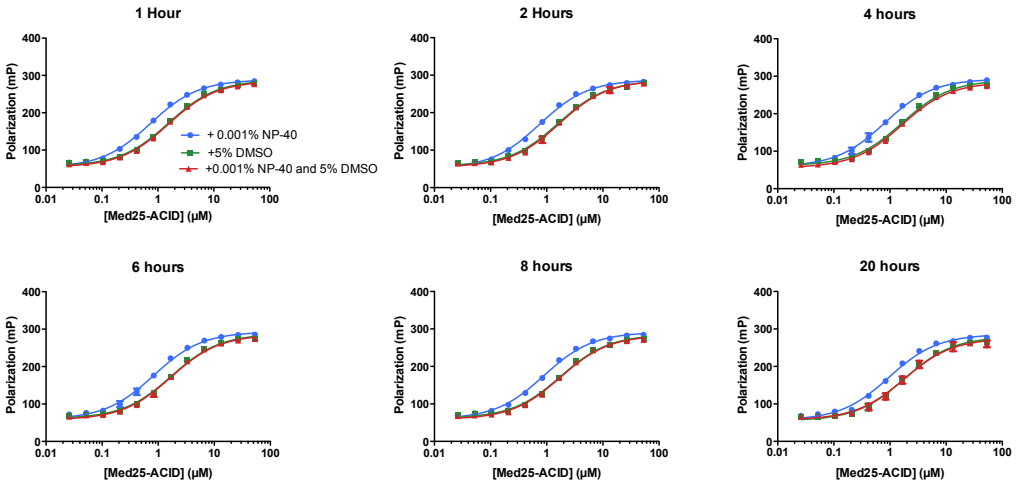


Figure 3.6- Effect of Tracer Concentration on Dynamic Range The binding curves of VP16(465-490) at tracer concentrations of 50 nM and nM are compared. Curves represent the means of three independent experiments with error bars representing the standard deviation of the anisotropy at the indicated concentration of AcID protein. Curves were fit using GraphPad Prism 5.

Following the selection of VP16(465-490) as the fluorescent tracer for use in the screen, we next sought to determine if the concentration of tracer present in the assay had a significant effect on the dynamic range of the binding curve. Minimizing the concentration of tracer in the FP assay would reduce the amount of material required for the screen, saving time and resources. In Figure 3.6 the binding curves of VP16(465-490) to AcID with 20 nM tracer and 50 nM tracer present are compared. Excess tracer did not further enhance the dynamic range relative to that observed with 20 nM tracer, and the lower concentration was thus used in the screen.

Subsequently, the assay was further optimized for HTS by testing the effects of NP-40 and DMSO on the interaction of VP16(465-490) and AcID and the stability of the assay over time. Very low concentrations (0.001% v/v) of the detergent NP-40 is included in the assay buffer in order to minimize aggregation of the protein or small molecules and to minimize non-specific interactions between the protein and equipment used to dispense assay components.[47] Since small molecules are added from concentrated DMSO stocks, demonstrating that DMSO does not have deleterious effects on the assay provides further confidence in identified hit molecules. Finally, assay plates may sit at room temperature for a period of several hours while waiting to be read by plate reader, so it is necessary that the assay be stable over a significant period of time. The effects of DMSO, NP-40, and the stability of the assay over time is shown below in Figure 3.7.



	K_d of Flo-VP16(465-490)-ACID interaction					
	1 hour	2 hour	4 hour	6 hour	8 hour	20 hour
+ NP-40	0.72 μM	0.75 μM	0.80 μM	0.81 μM	0.88 μM	0.95 μM
+DMSO	1.5 μM	1.6 μM	1.6 μM	1.7 μM	1.7 μM	1.8 μM
+ NP-40 and DMSO	1.5 μM	1.7 μM	1.6 μM	1.7 μM	1.7 μM	1.8 μM

Figure 3.7- Assay Stability Time Course, DMSO Effects, and NP-40 Effects on Tracer Affinity. The effects of 0.001% NP-40 and 5% DMSO are shown in isolation (blue and green curves, respectively) or in combination (red curve). Additionally, the stability of the assay is shown after 1, 2, 4, 6, 8, and 20 hours. Curves represent the means of three independent experiments with error bars representing the standard deviation of the anisotropy at the indicated concentration of AcID protein. Curves were fit using GraphPad Prism 5. *(With Paul A. Bruno)*

The 0.001% NP-40 had a negligible impact on the affinity of VP16(465-490) for AcID as the K_d value was 0.72 μM in the presence of the detergent, compared to 0.58 μM without NP-40 present. DMSO at 5% v/v had a slightly stronger, though still negligible, effect as the K_d value shifted to 1.5 μM at this concentration of DMSO. The standard protocol for compound addition to the assay results in a 1% final DMSO concentration, but demonstrating the minimal perturbation caused by a 5% final concentration would allow for testing compounds at a higher concentration in the event that the standard protocol failed to identify any lead molecules. Furthermore, the simultaneous presence of NP-40 and DMSO did not result in effects on the assay different than either component in isolation. Finally, the assay is highly stable over time as the K_d value shifted less than thirty percent even after being incubated at room temperature for twenty hours.

Additionally, NP-40, DMSO, and extended incubations had no effect on the dynamic range of the experiment, indicating that the assay is well adapted to HTS.

The final parameter to optimize prior to screening was the concentration of AcID protein to use in the assay. In order to accomplish this, the calculated Z' scores for 500 nM, 850 nM, and 2.5 μM AcID in the presence of 20 nM VP16(465-490) fluorescent tracer were compared. The Z' score is a metric, measured from zero to one with one representing a theoretically perfect assay, that defines the quality of an assay based upon the dynamic range and the reproducibility of the positive and negative controls as defined by their standard deviations. Increasing protein concentrations results in larger dynamic ranges, and thus larger Z' scores. Standard convention for high-throughput screening dictates that Z' scores greater than 0.6 indicate an excellent assay.[48] These scores are calculated using **Equation 3.1**:

$$Z' = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

$$\sigma = standard\ deviation\ of\ positive\ or\ negative\ control$$

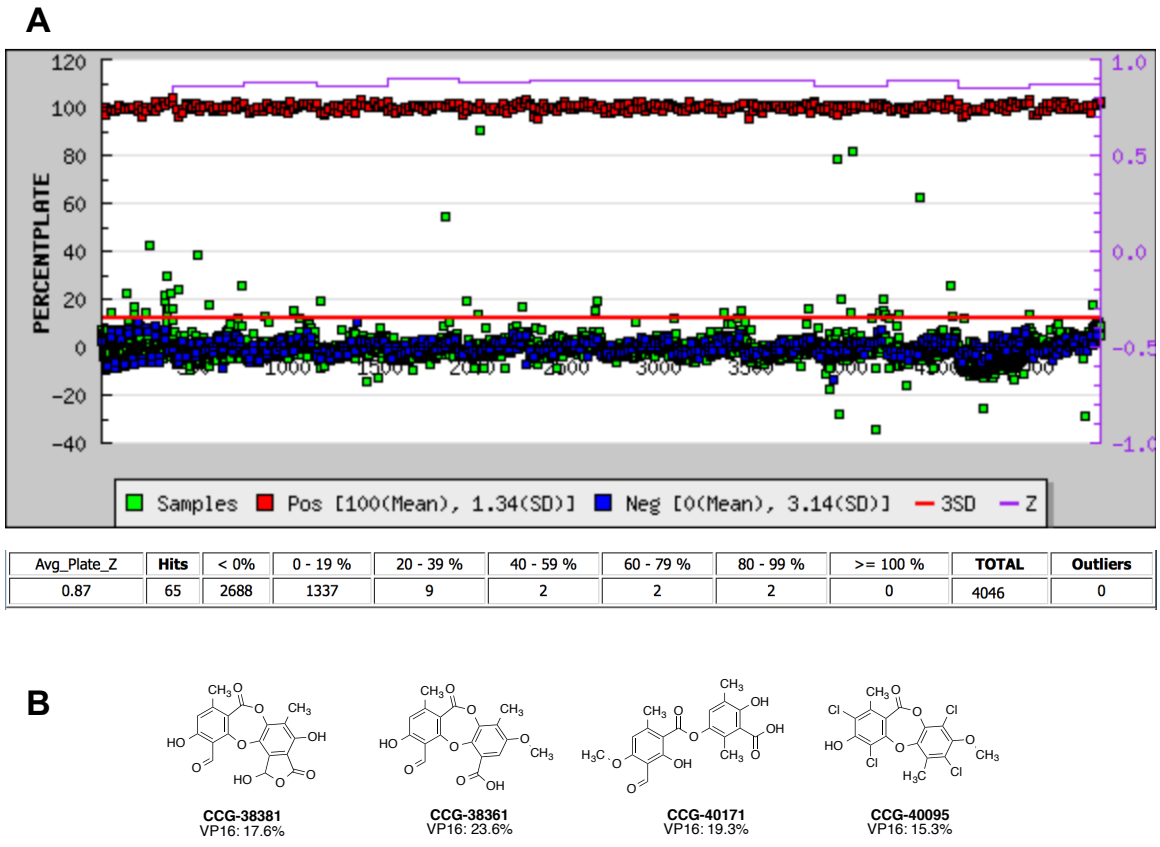
$$\mu = mean\ value\ of\ positive\ or\ negative\ control$$

Within the context of the fluorescence polarization assay, the positive control was defined as the polarization of the tracer alone and the negative control was defined as the polarization of the tracer bound to AcID protein in the presence of one percent DMSO. AcID concentrations of 500 nM, 850 nM, and 2.5 μM produced Z' scores of 0.76, 0.81, and 0.88, respectively. Furthermore, the dynamic range of the assay with 500 nM, 850 nM, and 2.5 μM spanned 70, 95, and 136 mP, respectively. Thus, the screen was run with an AcID concentration of 850 nM as these conditions produced an excellent Z' score and demonstrated a broad dynamic range without requiring large amounts of purified protein. Taken together, the above data is consistent with a robust and high-fidelity assay that is well suited for high-throughput screening.

Primary Screen for Inhibitors of VP16(465-490)•AcID

Following the successful optimization of an FP assay well suited for HTS applications, a primary screen at the University of Michigan Center for Chemical Genomics (CCG) was completed. In this screen, 4,046 compounds were tested at a

concentration of 20 μM from the MS Spectrum 2000, Focused Collections, and BioFocus NCC libraries, which include known bioactive molecules, secondary metabolites, natural products, and FDA approved drugs. The primary screening campaign had an average Z' score of 0.87, indicating an excellent assay, and a 1.6% hit rate. For the purposes of this screen, a hit was defined as any compound that resulted in inhibition greater than three standard deviations above the negative control, which corresponded to approximately ten percent inhibition. Following the primary screen, hits were filtered and compounds with known chemically reactive properties as well as those compounds that demonstrated native fluorescence greater than ten percent of the fluorescence produced by the tracer were removed. Following the filtering step, we identified the compounds **CCG-38381** (norstictic acid), **CCG-38361** (psoromic acid),**CCG-40171** (baeomycesic acid), and **CCG-40095**, as shown in Figure 3.8.



Avg_Plate_Z	Hits	< 0%	0 - 19 %	20 - 39 %	40 - 59 %	60 - 79 %	80 - 99 %	>= 100 %	TOTAL	Outliers
0.87	65	2688	1337	9	2	2	2	0	4046	0

Figure 3.8- Primary Screen of Bioactive Compounds. (A) Campaign view of the primary screen against VP16(465-490)•AcID. Red points represent positive controls, blue points represent negative controls, green points are the inhibition induced by test compounds, and the red line represents the threshold for hits. (B) Structures of the most promising hits following filtering. The percent inhibition of the VP16(465-490)•AcID interaction is denoted below the CCG identifier. *(With Paul A. Bruno)*

All four of these lead compounds belong to a class of molecules known as depsides and depsidones, which are natural products derived from lichens and were well studied by Emil Fischer.[49] This result was particularly exciting given that previously identified molecules belonging to this class have been reported as effective inhibitors of α-helix dependent activator•coactivator interactions by our group.[34] Molecular dynamics simulations of the depside sekikaic acid revealed that the central ester can rotate, displaying the aromatic side chains in an orientation similar to an α-helix. Thus, an intriguing possibility is that this class of molecules represent a privileged scaffold for the inhibition of α-helix dependent protein-protein interactions, though additional experimentation will be necessary to more definitively demonstrate this.

Hit Confirmation and Selectivity Studies of Norstictic Acid and Psoromic Acid

In order to confirm that norstictic acid and psoromic acid are inhibitors of the VP16•AcID interaction, samples of both molecules were acquired from commercial sources and their activity against the VP16(465-490)•AcID FP assay was retested. Unfortunately, we were unable to secure a commercial source for baeomycesic acid and so we were unable to further confirm it as an AcID inhibitor. Freshly purchased norstictic acid and psoromic acid were tested for inhibition of VP16(465-490)•AcID as well as VP16(438-464)•AcID in order to determine if the molecules are site selective inhibitors of the AcID motif, as shown in Figure 3.9.

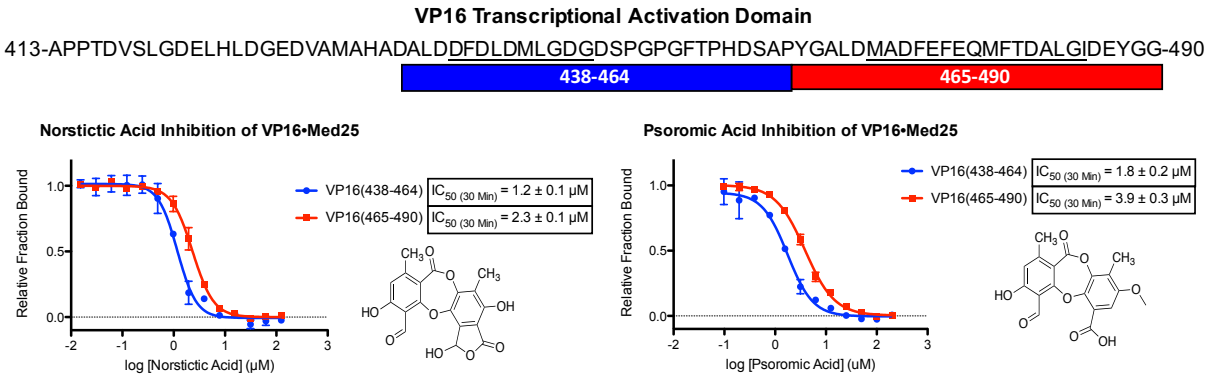


***Figure 3.9*- Validation of Norstictic and Psoromic Acid with Commercial Compounds.** The ability of Norstictic and psoromic acid to inhibit VP16(438-464)•AcID and VP16(465-490)•AcID was tested with freshly prepared compound obtained from commercial sources. Curves represent the mean values of three independent experiments and vertical error bars represent the standard deviations of

the fraction bound at the indicated concentrations of small molecule. Curves were fit with GraphPad Prism 5.

The freshly prepared stocks of norstictic acid and psoromic acid were significantly more potent than the compounds held within the CCG libraries. Norstictic acid inhibits VP16(438-464)•AcID with an $IC_{50\ (30\ Min)}$ value of 1.2 μM and VP16(465-490)•AcID with an $IC_{50\ (30\ Min)}$ value of 1.8 μM compared to 13 μM for the library stocks (Figure 3.13). Thus, the fresh compound is more potent by approximately ten-fold. Psoromic acid inhibits VP16(438-464)•AcID with an $IC_{50\ (30\ Min)}$ value of 2.3 μM and VP16(465-490)•AcID with an $IC_{50\ (30\ Min)}$ value of 3.9 μM compared to 107 μM for the library stock (Figure 3.13), representing a greater than twenty-five fold enhancement in potency. The largely enhanced potency observed with freshly purchased compounds is consistent with norstictic acid and psoromic acid as inhibitors of VP16•AcID interactions.

In order to determine whether these molecules are selective for inhibition of AcID-dependent interactions, the ability of the molecules to inhibit alternative activator•coactivator interactions was determined. Two CBP/p300 KIX domain-dependent interactions were tested to demonstrate that the molecules do not function as generic TAD or amphipathic helix mimetics and the VP16•Med15 interaction was tested to demonstrate that the molecules do not specifically mimic the VP16 TAD structure. These experiments support the hypothesis that norstictic acid and psoromic acid are selective inhibitors of the AcID motif, suggesting that they may function by allosterically inhibiting activator interactions with the domain which is further explored in subsequent sections of this Chapter.

Specifically, the ability of norstictic acid and psoromic acid to inhibit the activator•coactivator interactions between MLL and the KIX domain of CBP/p300, as well as the interaction between the phosphorylated activation domain of CREB (pKID) and the KIX domain was determined. Our lab and others have extensively studied these interactions and demonstrated that they rely upon α-helical secondary structure within the TADs of the activator binding partner.[50,51] Thus, given that these interactions occur with similar structural features to the VP16•AcID interaction and the demonstrated utility of these binding assays, the MLL•KIX and pKID•KIX interactions are useful in determining the specificity of norstictic acid and psoromic acid for the AcID motif of Med25. As shown

in panels B and C of Figure 3.10, neither norstictic acid nor psoromic acid demonstrated any significant perturbation of the interaction between MLL and KIX or the pKID domain of CREB and KIX. Thus, these data are consistent with the hypothesis that norstictic acid and psoromic acid are selective inhibitors of AcID and do not function as generic TAD mimics.

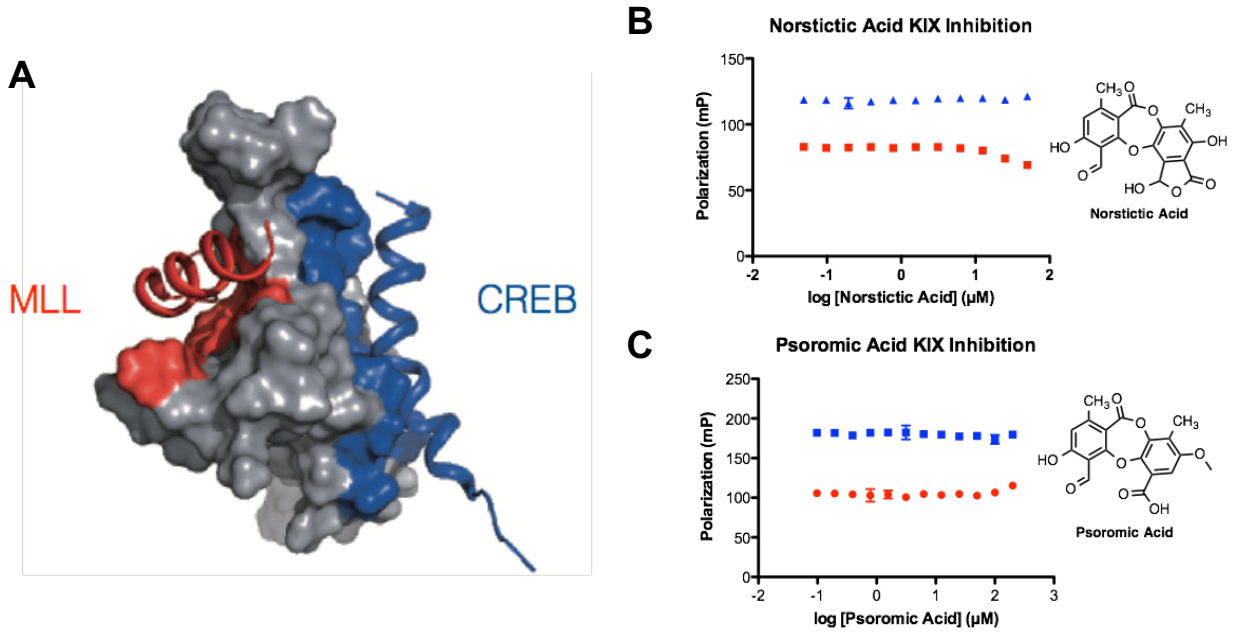


***Figure 3.10*- Effects of Norstictic and Psoromic Acid on Activator•KIX Interactions.** (A) Binding of MLL and the pKID domain of CREB occurs at distinct surfaces on the KIX domain and requires that the TADs adopt an α-helical secondary structure. (B) Inhibition of norstictic acid against the MLL•KIX interaction (Red curve) and the pKID•KIX interaction (Blue curve) (C) Inhibition of psoromic acid against the MLL•KIX interaction (Red curve) and the pKID•KIX interaction (Blue curve). Curves represent the mean values of three independent experiments and were plotted using GraphPad Prism 5.

In order to further confirm that norstictic acid and psoromic acid are selective for AcID, their ability to inhibit the interaction between VP16 and Mediator Subunit 15 was also determined. Testing this particular interaction will allow us to determine if the identified lead molecules are merely acting as VP16 mimetics, as comparable inhibition of VP16•Med15 interactions would indicate. In order to test the inhibition of VP16•Med15 interactions, Med15(1-345), which contains a GACKIX domain and the A and B boxes which serve as VP16 interaction sites, was expressed.[52] The expressed protein was then confirmed to be binding-competent by testing the direct binding of the VP16 derived

peptide VP2 (DFDLDMLGDFDLDMLG), which is a single repeat of VP16(441-448), as well as VP16(438-454) and VP16(467-488).

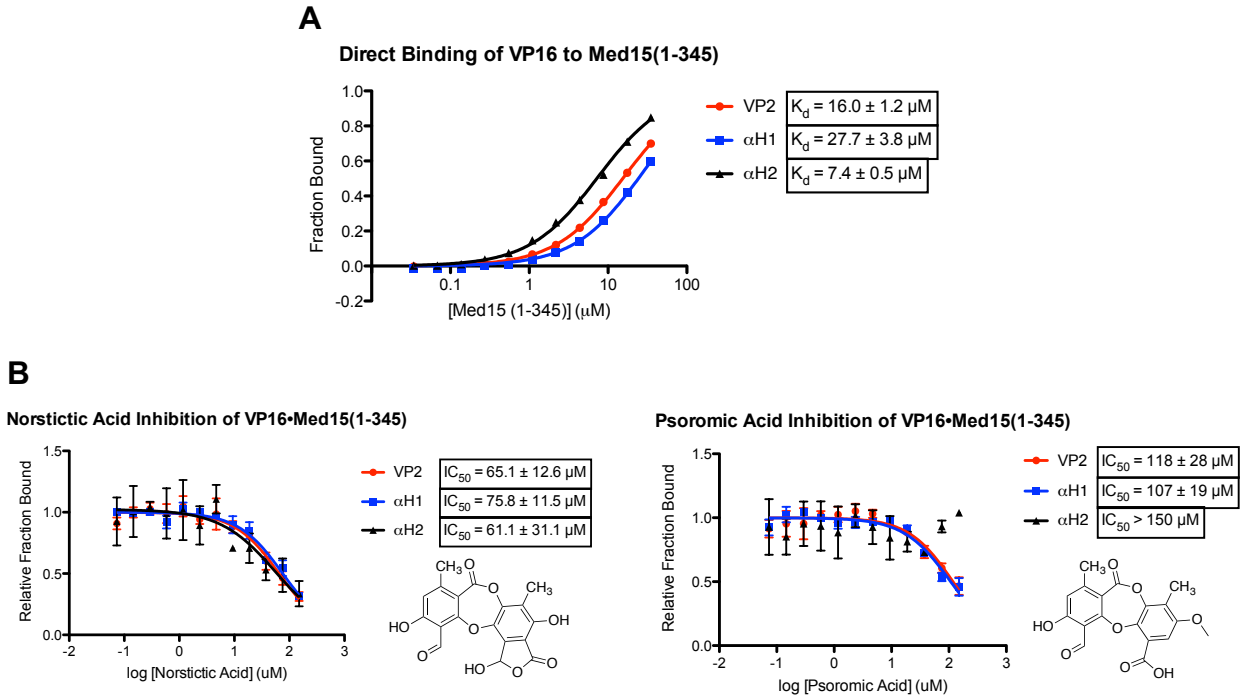


***Figure 3.11*- Effects of Norstictic and Psoromic Acid on VP16•Med15(1-345) Interactions.** (A) Direct binding of VP2, VP16(438-454) (αH1), and VP16 (467-488) (αH2) to Med15(1-345). (B) Inhibition of VP16•Med15 interactions by norstictic acid and psoromic acid. Curves represent the mean values of three independent experiments, and error bars represent the standard deviation of the fraction of tracer bound at the indicated concentration of protein or small molecule. Curves were fit with GraphPad Prism 5. *(With Paul A. Bruno)*

As shown in panel A of Figure 3.11, VP2 bound to Med15 with a K_d value of 16.0 μM, consistent with reported literature values.[52] VP16(438-454) and VP16(467-488), denoted as αH1 and αH2 on the curves, bound to Med15 with K_d values of 27.7 and 7.4 μM, respectively, indicating that the protein was properly folded and capable of binding VP16 derived peptides. Inhibition curves of the three VP16•Med15 interactions demonstrate that norstictic acid and psoromic acid are relatively weak inhibitors of VP16•Med15 interactions with IC_{50} values greater than 60 μM and 105 μM respectively; values that are significantly larger than those observed for the inhibition of VP16•AcID interactions. Thus, it is possible that the small molecules weakly mimic some aspects of the VP16 TAD, though the significantly greater potency against VP16•AcID interactions suggests that they are more selective for AcID than one of its native activator binding

partners. Taken together, the above data demonstrate that norstictic acid and psoromic acid are effective and selective inhibitors of VP16•AcID interactions and do not merely mimic the structure of an amphipathic helix generally or the VP16 TAD specifically. The possibility that this selectivity is achieved through the induction of allosteric changes within the domain will be explored in subsequent sections of this Chapter.

Substructure Search of CCG Libraries for Depside and Depsidone-like Molecules

Based upon these observations, a substructure search of the small molecule libraries held by the CCG was completed using the depside and depsidone core scaffold defined in panel A of Figure 3.12 in order to identify closely related compounds.

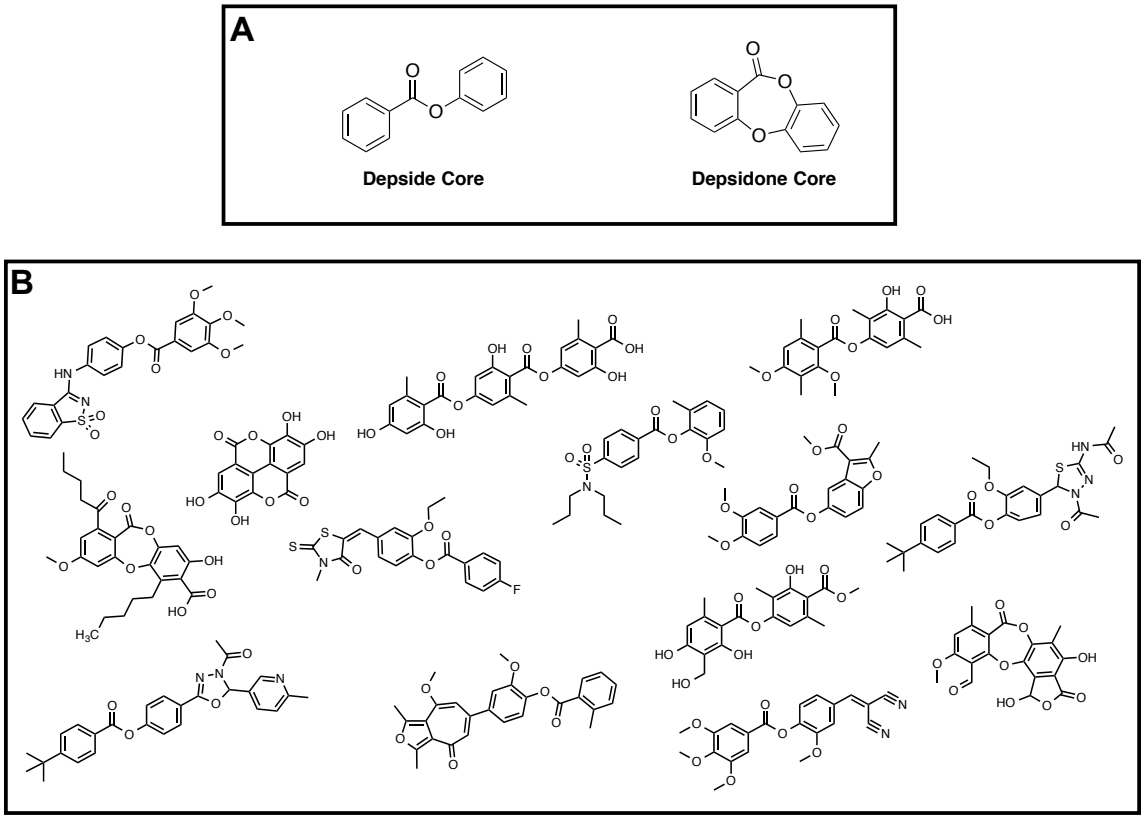


***Figure 3.12*- Substructure Search of CCG Libraries for Depside/Depsidone Core.** (A) The core depside and depsidone scaffolds that were utilized in the substructure search. (B) Compounds identified within the CCG libraries that contain the core depside and depsidone scaffold. *(With Paul A. Bruno)*

This substructure search revealed sixteen additional molecules that contain the depside or depsidone core with unique patterns of aryl substituents. These compounds, in addition to the four identified in the primary screen, were then subjected to a dose-

response analysis. Additionally, the order of addition of assay components was reversed for the dose-response assays, with the addition of compound to buffer to first check for native fluorescence of the compounds, followed by the addition of the tracer to check for fluorescence quenching caused by the molecules. None of the compounds demonstrated native fluorescence outside of acceptable limits nor significant fluorescence quenching of the tracer. Inhibition curves of active compounds were subsequently fit to determine approximate IC_{50} values, as shown in Figure 3.13.

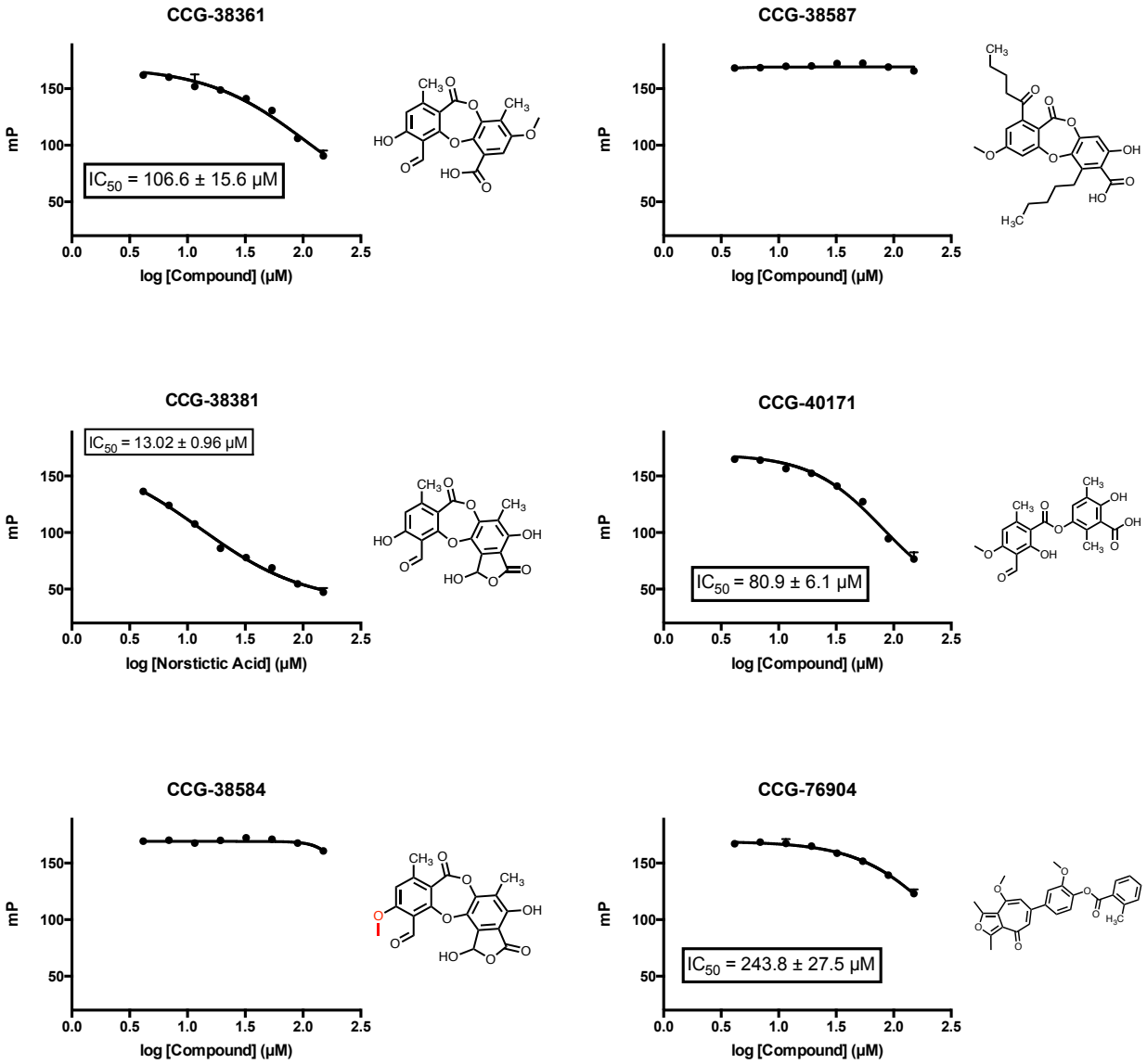


***Figure 3.13*- Dose Response Curves of Selected Compounds from Substructure Search.** Inhibition assays were completed in duplicate at eight concentrations of inhibitor ranging from 150 μM

to 4.1 μM. Curves were fit using GraphPad Prism 5 to determine approximate IC_{50} values. CCG-38587, lobaric acid, is a previously identified activator•coactivator interaction inhibitor, the compounds in the left column are closely related depsidones, and CCG-40171, baeomycesic acid, is the lone active depside. *(With Paul A. Bruno)*

The majority of the twenty compounds tested did not demonstrate significant inhibition of VP16(465-490)•AcID, with the exception of the compounds identified in the primary screen. A number of the compounds lacked particularly complex or complete substitution of the aryl rings that comprised the core scaffold, suggesting that this substitution is critical for function. Furthermore, compound **CCG-38587** (lobaric acid), which was previously identified as an inhibitor of activator•coactivator interactions, failed to demonstrate any inhibition of the VP16•AcID interaction over the tested concentration range. This suggests that the identity of the substituents on the aryl rings are capable of conferring specificity. Ultimately, the substructure search failed to produce additional compounds from those identified in the primary screen, but did provide some preliminary structure-activity-relationship data, as shown in in Figure 3.14.

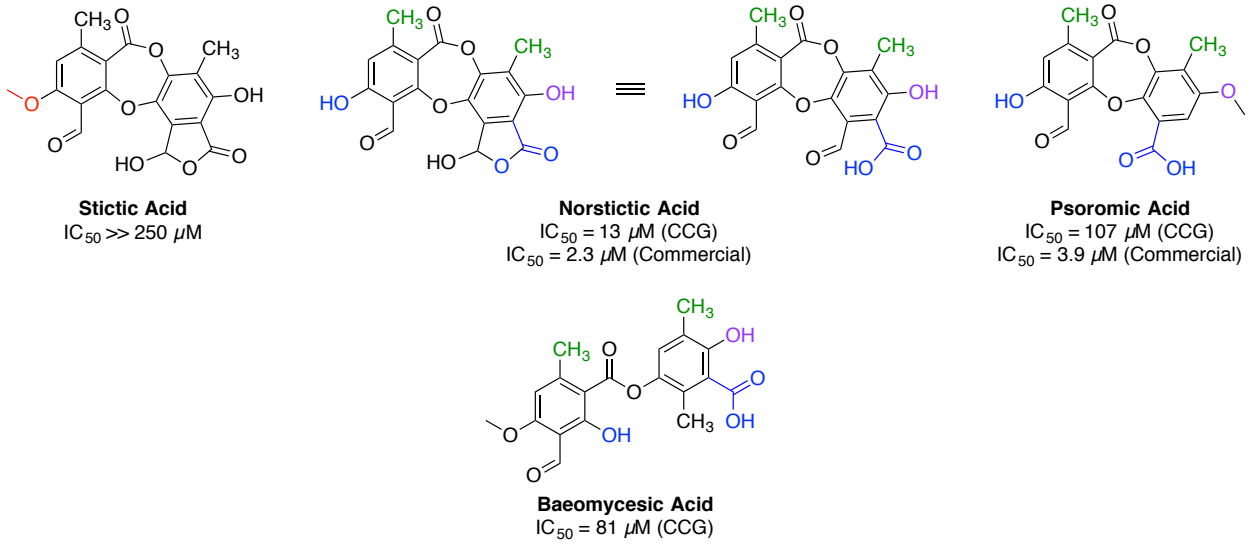


***Figure 3.14*- Limited Structure Activity Relationship for Potential Lead Molecules.** Comparisons of the activity of closely related molecules in the dose response screen provides some insight into important structural elements of the potential lead molecules. Potential hydrophobic contacts are shown in green and important polar functionalities are shown in blue and purple.

Norstictic acid, psoromic acid, and baeomycesic acid are similarly substituted, suggesting that these functionalities may be important for the specific inhibition of AcID dependent interactions. All of the hits contain an *ortho*-phenol substituted aldehyde.

Notably, stictic acid, an analogue of norstictic acid is methylated at the *ortho*-phenol and lacks all inhibitory activity, suggesting that a free phenol at this position is indispensable for activity. The three lead compounds also contain hydrophobic methyl functionalities (green functional groups on the structures in Figure 3.14) on the same face of the molecules, while the opposite face of the molecules contain an acidic functionality in addition to the polar phenol group (blue functional groups on the structures in Figure 3.14). Thus, the presence of a hydrophobic face and a distinct hydrophilic face suggests that these molecules may be amphipathic in nature, similar to the VP16 TAD, perhaps explaining their relatively potent activity. The acidic functionality on norstictic acid exists as a lactonol that results from an intramolecular reaction with a second aldehyde. Finally, the three lead molecules also contain a polar functional group, either a phenol or methylated phenol, adjacent to one of the methyl groups (purple functional groups on the structures in Figure 3.14). The fact that this functional group may be a free phenol or a methylated phenol suggests that it may function as a hydrogen bond acceptor and does not necessarily need to act as a hydrogen bond donor.

Mechanism of Action Elucidation for Norstictic Acid

After confirming that norstictic acid and psoromic acid are inhibitors of the VP16•AcID interaction *in vitro* and that the inhibition was selective, we next sought to determine the mode of action by which inhibition is achieved. The majority of these experiments were completed with norstictic acid, given its greater potency in inhibiting VP16•AcID interactions. Molecular dynamics (MD) simulations were completed in collaboration with the Brooks lab at the University of Michigan to attempt to identify potential binding surfaces for norstictic acid. The MD simulations revealed that the highest density of molecule clusters around surfaces rich in lysine residues, which is unsurprising given the acidic, negatively charged, functionalities present on norstictic acid and the positively charged side chains of the lysine residues. The clustering of norstictic acid near regions rich in lysine residues and the lowest energy conformations of the molecule at those surfaces on AcID are shown below in Figure 3.15.

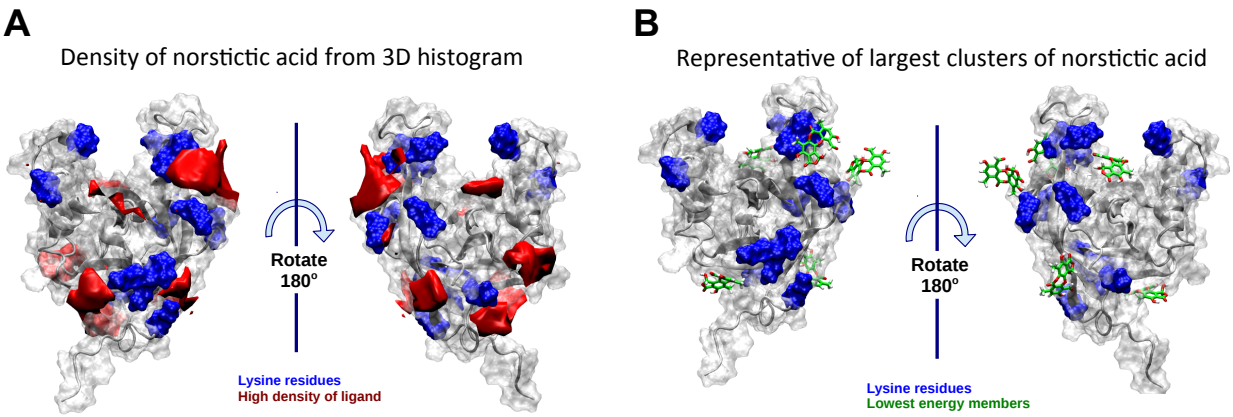


Figure 3.15- ***Molecular Dynamics Simulations of Norstictic Acid Localization.*** (A) Molecular dynamics simulation that reveals the predicted locations with the highest norstictic acid density (Red) on the AcID surface. Blue surfaces indicate lysine residues (B) Molecular dynamics simulation that shows the lowest energy conformations (Green) of norstictic acid within the regions of highest density on the AcID surface. Blue surfaces indicate lysine residues. *(Completed by Jessica K. Gagnon)*

In order to support the hypothesis that electrostatic interactions between the small molecule and protein play a critical role in the inhibition of VP16•AcID interactions by norstictic acid, a sodium chloride titration experiment was completed for the inhibition of VP16(465-490)•AcID by norstictic acid. The ability of norstictic acid to inhibit VP16(465-490)•AcID in the presence of sodium chloride concentrations ranging from 100 mM to 1000 mM is reported below in Figure 3.16.

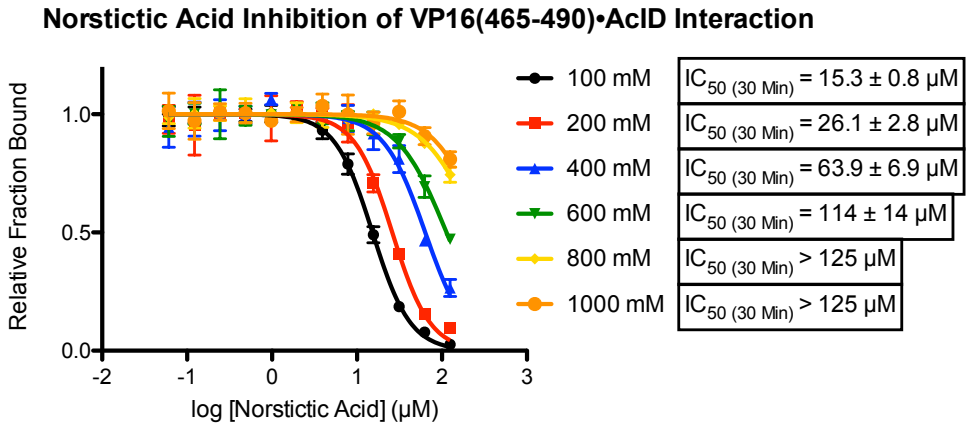


***Figure 3.16*- Effect of NaCl on Norstictic Acid Inhibition of VP16(465-490)•AcID.** Inhibition of VP16(465-490)•AcID at various concentrations of sodium chloride ranging from 100 mM to 1000 mM. Curves represent the mean values of three independent experiments with error bars representing the standard deviation of the fraction of tracer bound at the indicated concentration of small molecule. Curves were fit with GraphPad Prism 5.

There is a clear salt dependence on the inhibition of VP16(465-490)•AcID by norstictic acid as the $IC_{50\ (30\ Min)}$ values increase significantly with increasing concentrations of sodium chloride. The $IC_{50\ (30\ Min)}$ value has shifted to larger than 110 μM in the presence of 600 mM sodium chloride and cannot be accurately determined in the presence of greater than 800 mM salt due to solubility constraints of the small molecule in the assay buffer. Thus, these data support the hypothesis that the interaction of norstictic acid with the AcID motif is dependent upon electrostatic interactions between the small molecule and protein.

One of the defining molecular features of the three compounds identified as lead molecules in the primary screen is the presence of an aldehyde functionality on one of the aryl rings, though in the case of norstictic acid there are two aldehyde functionalities, one of which is involved in an intramolecular reaction to generate the lactonol ring. Aldehydes are highly electrophilic and their presence in other inhibitors has been previously demonstrated to result in the covalent binding of the inhibitor to the target protein through the formation of a Schiff base, a fully substituted imine, through the dehydration reaction of the aldehyde with the nucleophilic amine of lysine residues. [53,54] Thus, we next sought to determine if norstictic acid was capable of covalently binding to AcID. Notably, such a modification would not only immobilize the inhibitor on the surface of the protein, but would also result in the loss of a positive charge on the surface of the domain.

Given the demonstrated importance of electrostatic contacts in the interaction between the VP16 TAD and AcID, the loss of critical electrostatic contacts through covalent binding of norstictic acid with nucleophilic lysine residues may be a significant contributor to the inhibitory activity of the molecule.

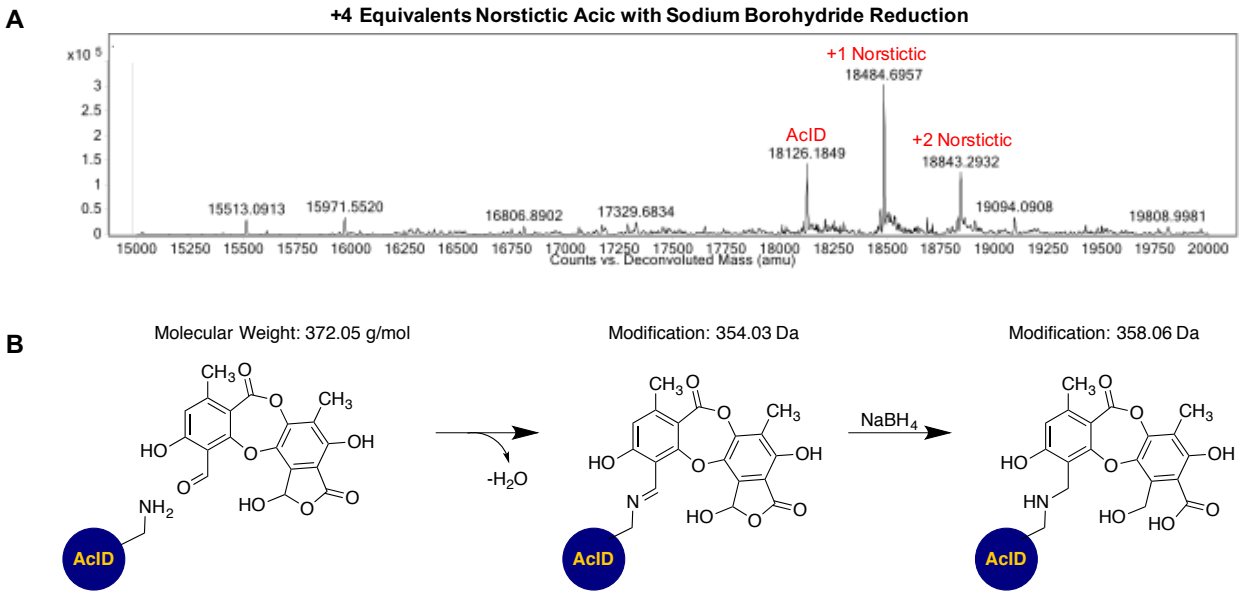


***Figure 3.17*- Covalent Labeling of AcID by Norstictic Acid.** (A) Mass spectrometry results that demonstrate norstictic acid covalently binds to the AcID motif. Protein was incubated with four equivalents of molecule and reduced with sodium borohydride. (B) Reduction of the Schiff base formed following the reaction between the protein and norstictic acid results in reduction of the second aldehyde in addition to the imine. *(With Paul A. Bruno)*

The results shown in Figure 3.17, confirm that norstictic acid is capable of covalently labeling AcID. Purified AcID protein was incubated with four equivalents of norstictic acid and treated with sodium borohydride to reduce the resultant imine. The labeled complex was then subjected to mass spectrometric analysis, revealing that the domain (18,126 Da) could be labeled with multiple equivalents of small molecule as demonstrated by the presence of peaks in the spectrogram at molecular weights of 18,484 Da and 18,843 Da, corresponding to the addition of one and two norstictic acid molecules, respectively. The protein was labeled with four equivalents of molecule as this was the stoichiometry of molecule present relative to protein at the IC_{50} value of norstictic acid. Of note, the reduction following reaction of norstictic acid with the domain also reduces the second aldehyde within the lactonol ring, explaining why reduction results in a mass increase of four Daltons, as opposed to the expected two Dalton increase. The observation that the resultant covalent adduct can be reduced is consistent with the hypothesis that the covalent interaction occurs through the formation of an imine. Additionally, the covalent adduct formed between norstictic acid and the AcID motif is particularly stable as it can be observed by mass spectrometric analysis even in the

absence of sodium borohydride reduction, as is shown later in this chapter in Figure 3.21 and Figure 3.27.

In order to demonstrate that the observed inhibition of AcID-dependent interactions by norstictic acid is not a result of the nonspecific labeling of lysine residues within the domain, the ability of alternative aldehyde containing small molecules to inhibit VP16•AcID interactions was next determined. Benzaldehyde, which has been previously demonstrated to modify lysine residues, was tested as in order to determine the general ability of aromatic aldehydes to modify the AcID motif, as shown below in Figure 3.18.[54]

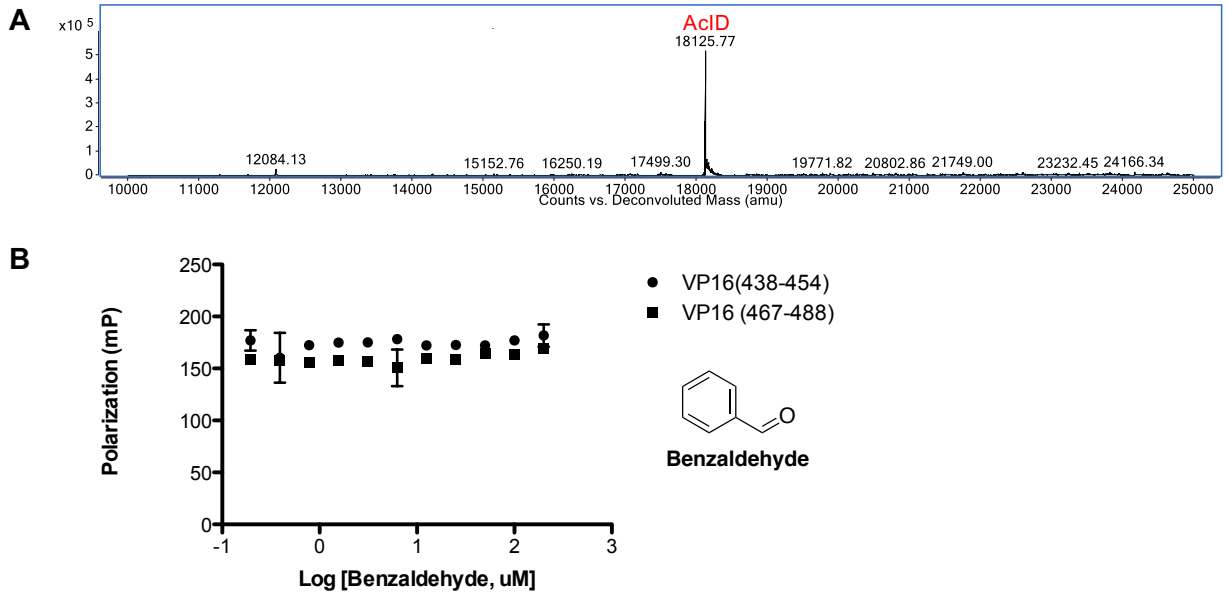


Figure 3.18- Benzaldehyde Does Not Inhibit or Covalently Label the AcID Motif (A) Mass spectrometric analysis of AcID in the presence of four equivalents of benzaldehyde. The expected mass of the adduct was 18,232 Da. (B) Inhibition experiments of benzaldehyde against the VP16(438-454) and VP16 (467-488) interactions with AcID.

Mass spectrometric analysis of AcID protein treated with four equivalents of benzaldehyde revealed that the small molecule is incapable of covalently modifying the domain, consistent with the hypothesis that norstictic acid specifically modifies the AcID motif. Furthermore, benzaldehyde fails to inhibit the interaction of both VP16(438-454) and VP16(467-488) with AcID, providing further evidence that the activity of norstictic acid against the domain is not solely dependent upon the reactivity of its aldehydes.

In addition to benzaldehyde, which is significantly less structurally complex than norstictic acid, the ability of atranorin, a closely related depside to norstictic acid, to

covalently modify the AcID motif and inhibit the VP16•AcID interaction was next determined, as shown below in Figure 3.19.

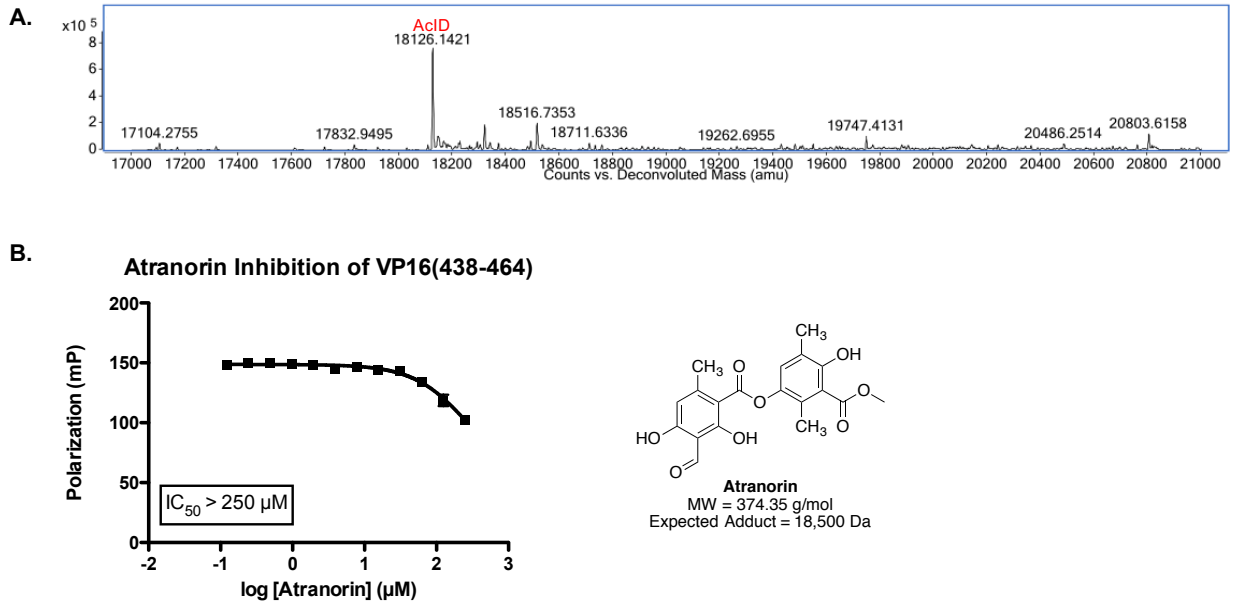


Figure 3.19- Atranorin Does Not Inhibit or Covalently Label the AcID Motif (A) Mass spectrometric analysis of AcID in the presence of four equivalents of Atranorin. The expected mass of the covalent adduct was 18,500 Da. (B) Inhibition of the VP16(438-464)•AcID interaction by atranorin.

Atranorin contains an ortho-phenolic aldehyde, similar to norstictic acid, but is incapable of covalently modifying the AcID motif, further supporting the hypothesis that norstictic acid specifically modifies AcID. Furthermore, atranorin is not an effective inhibitor of the VP16•AcID interaction as the IC_{50} of the small molecule was determined to be in excess of 250 µM. Atranorin contains a methyl ester as opposed to a free carboxylic acid and thus may lack a critical electrostatic interaction with the domain that is important for inhibitory activity, as has been previously demonstrated for norstictic acid. Taken together, the inability of other aldehyde containing small molecules to inhibit the VP16•AcID interaction or covalently modify the AcID motif is consistent with the hypothesis that norstictic acid is a selective inhibitor of the domain.

Following the observation that norstictic acid is a covalent inhibitor of AcID, we next sought to determine how quickly maximal inhibition was achieved. Optimizing incubation time to allow for complete formation of the imine between the molecule and protein would allow for a more accurate determination of the full inhibitory potential of the molecule. In order to accomplish this, the IC_{50} of norstictic acid against VP16(465-

490)•AcID was measured at various time points to determine at which point maximal inhibition was achieved. The results of this time course study are shown below in Figure 3.20.

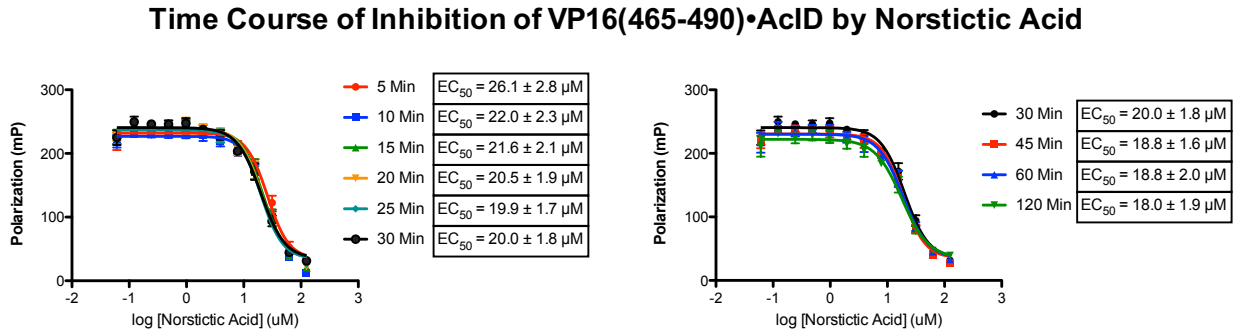


***Figure 3.20*- Norstictic Acid Inhibition Time Course of VP16(465-490)•AcID.** The IC_{50} of norstictic acid was determined for inhibition of the VP16(465-490)•AcID interaction at various time points to determine how quickly inhibition was achieved. Curves represent the mean values of three independent experiments, with vertical error bars representing the standard deviation of the mean polarization at the indicated concentration of norstictic acid. Curves were fit using GraphPad Prism 5.

The time course reveals that norstictic acid inhibits VP16(465-490)•AcID over a fairly rapid time frame, with maximal inhibition achieved after approximately ten minutes as the IC_{50} values determined after this point are within error. At the later time points the dynamic range of the binding curve decreases slightly, most likely as a result of photobleaching of the fluorescein tag on the peptide due to repeated measurements of the same samples, which may provide an explanation for the slight decrease in IC_{50} values at late time points. Thus, these data suggest that the aldehyde is particularly reactive as maximal inhibition is achieved rapidly. Importantly, previously completed experiments testing the inhibition of norstictic acid against the VP16•AcID interaction were all incubated for thirty minutes prior to analysis which was well beyond the time required to achieve maximal inhibition.

$^{1}H,^{15}N$-Heteronuclear Single Quantum Coherence (HSQC) NMR Studies of AcID Perturbations Induced by Norstictic Acid

Norstictic acid has been identified as a potent and selective inhibitor of VP16•AcID interactions. Molecular dynamics simulations have demonstrated that norstictic acid clusters near positively charged surfaces on the AcID protein likely through critical electrostatic contacts, a hypothesis that has been further supported by demonstrating that

increasing salt concentration perturbs the activity of norstictic acid. Furthermore, it has been demonstrated that norstictic acid is a covalent inhibitor that labels AcID likely through a reaction with the nucleophilic amine of lysine residues on the domain, further underscoring the importance of the localization of the small molecule to lysine containing regions of the protein. In order to provide further evidence that norstictic acid clusters near lysine rich regions of the domain and to determine if norstictic acid induces allosteric changes to the structure of the AcID motif, $^{1}H,^{15}N$-HSQC NMR perturbation experiments of the AcID protein in complex with norstictic acid were completed. In order to accomplish this, ^{15}N labeled AcID protein was expressed and purified and incubated with norstictic acid prior to NMR analysis. Mass spectrometric characterization of the complexes is shown in Figure 3.21.

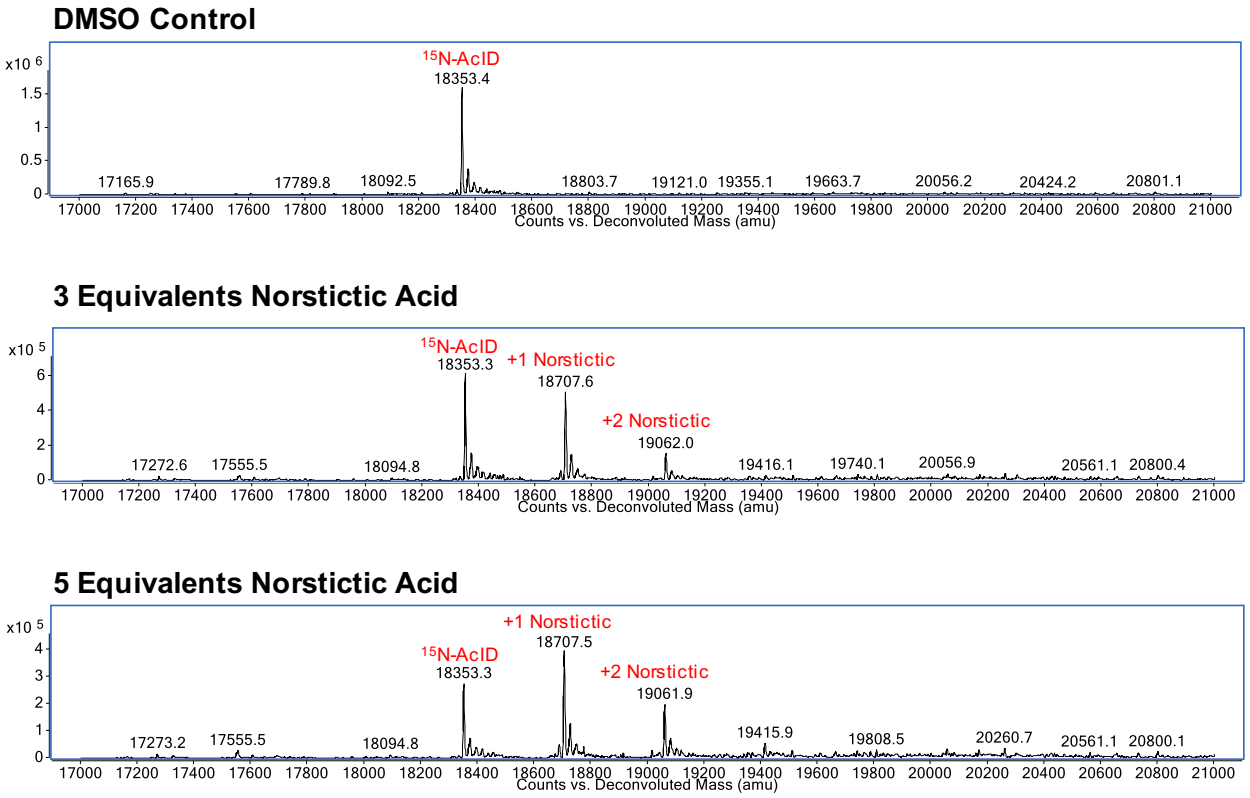


***Figure 3.21*- Confirmation of Norstictic Acid Labeling for HSQC Analysis.** ^{15}N labeled AcID protein was treated with three or five equivalents of norstictic acid or DMSO as a control and incubated at room temperature for two hours. The samples were then analyzed by mass spectrometry prior to confirm sufficient labeling prior to analysis by NMR. *(Completed by Paul A. Bruno)*

^{15}N-labeled AcID was incubated with DMSO as a vehicle only control, three equivalents of norstictic acid, or five equivalents of norstictic acid for two hours at room

temperature in order to analyze the range of chemical shifts induced by the formation of a covalent complex of AcID and norstictic acid. These three samples were then subjected to mass spectrometric analysis to confirm that a covalent complex had been formed. The DMSO control sample showed only unlabeled ^{15}N-AcID (18,353 Da), while samples treated with norstictic acid contained singly labeled complexes (18,707 Da) or doubly labeled complexes (19,062 Da). As expected, incubation with five equivalents resulted in a higher degree of labeling for both the singly bound and doubly bound complexes. The three samples were then immediately subject to NMR analysis following the confirmation of complex formation as the buffer these samples were prepared in lacked glycerol and NP-40, which are critical for maintaining solubility of the AcID protein over time. Thus, by minimizing the time from sample preparation to NMR analysis, the potential loss of signal that results from the precipitation of the protein•norstictic acid complex over time was also minimized. ^{1}H,^{15}N-HSQC peaks were then assigned using previously published chemical shifts for the AcID motif and the spectra were overlaid for the sake of comparison, as shown in **Figure 3.22**.[35]

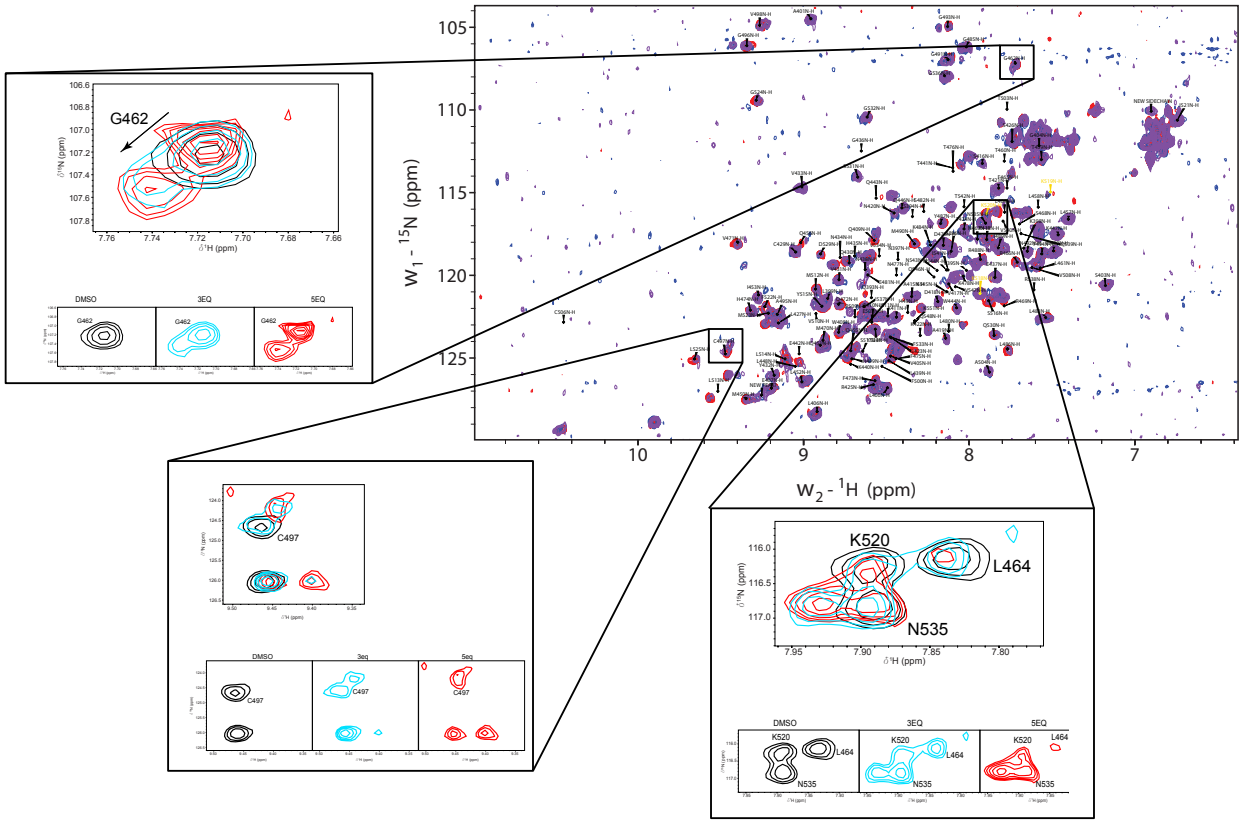


Figure 3.22- **^{1}H,^{15}N-HSQC NMR Spectrum of Norstictic Acid Labeled AcID Protein** The full ^{1}H,^{15}N-HSQC spectra of ^{15}N-labeled AcID treated with DMSO (red spectrum), three equivalents of norstictic acid (purple spectrum), or five equivalents of norstictic acid (blue spectrum) is shown in the upper right hand corner. Additionally, representative chemical shifts (G462, C497, and L464/K520/N535) are shown as inserts. *(Completed by Felicia Gray)*

The most immediate observation is that the addition of norstictic acid to the AcID protein produces spectra with relatively small changes in the backbone amide signals and that these signals remain well resolved. This observation suggests that covalently labeled AcID protein is not destabilized, as such an event would result in the time-dependent clustering of residues in the ^{1}H,^{15}N-HSQC spectrum. Additionally, the observation that induced chemical shifts were minor in terms of intensity suggests that the binding of Norstictic AcID induces subtle structural changes, as opposed to large scale shifts in the global protein structure which would have resulted in significantly larger chemical shifts. Examples of the intensity of observed chemical shifts are shown as inserts in Figure 3.22. A significant majority of the amide backbone resonances were assigned using previously published results, but not every peak was effectively assigned as several of them were significantly broadened by the addition of norstictic acid. Particularly notable shifts were

observed for AcID residues W408, Q409, G462, L483, G485, L486, G491, G493, C496, C497, V498, L513, S516, K520, and N535. These residues were mapped onto the AcID structure and compared to observed chemical shifts induced by the binding of the VP16 TAD to the H1 and H2 interaction surfaces of AcID in Figure 3.23.

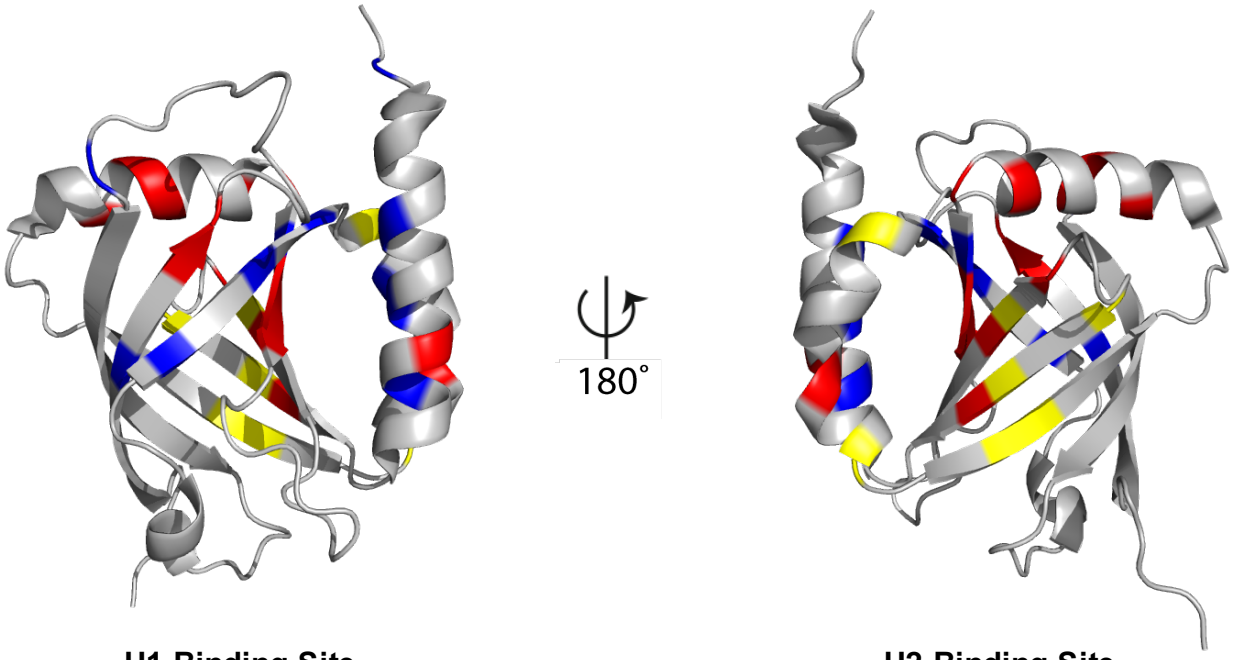


***Figure 3.23*- $^{1}H,^{15}N$-HSQC NMR Perturbation of AcID Induced by Norstictic Acid Binding.** Significantly perturbed residues are shown in red on the reported structure of the AcID motif. Additionally, the induced chemical shifts following the binding of the VP16 TAD to the H1 and H2 binding sites are shown in blue and yellow, respectively. Shifts in helix α-2 (perpendicular to the β-barrel), are unique to norstictic acid binding events.

Interestingly, the majority of the strongest chemical shifts induced by binding of norstictic acid to AcID occurred at surfaces within the H1 and H2 binding sites, suggesting that the inhibitor is binding to the protein in locations required for interaction with VP16. Furthermore, norstictic acid does not appear to be inducing structural changes at positions far from the VP16 TAD binding sites, further supporting the notion that the preferred binding locations of the small molecule on AcID is similar to the binding sites of the native ligands. The observation of induced chemical shifts at both the VP16 H1 and H2 binding sites is consistent with previous observations that norstictic acid is capable of perturbing binding events at both surfaces. Whether these perturbations are a result of orthosteric inhibition at both sites following the binding of norstictic acid to lysine residues

within both surfaces or as a result of allosteric changes that result from the binding of the inhibitor to alternative surfaces cannot be determined based on the HSQC results alone.

Interestingly, the binding of norstictic acid to the AcID motif induces a number of chemical shifts within helix α2, which is located at the top of the β-barrel at the interface of the H1 and H2 binding sites. These chemical shifts are not observed following the binding of either VP16 H1 or H2, suggesting that the binding of norstictic acid may result in unique structural changes not caused by the binding of native ligands. Thus, this data provides evidence that norstictic acid may be exerting its effects through allosteric shifts in the protein in addition to the direct competition for critical binding surfaces required by native ligands, such as VP16. Such a mechanism of action may provide some explanation for the observed selectivity of the molecules against the domain. Two flexible loop regions are located near the top of the H1 and H2 binding sites in close proximity to helix α2, the orientations of which may be affected by shifts within α2. Significant additional experimentation will be required to confirm the role of helix α2 in a potential allosteric mechanism of inhibition, including mutational analyses and further structural elucidation using alternative NMR techniques.

Given that norstictic acid covalently binds to AcID by forming an imine with reactive lysine residues, the chemical shifts induced by the binding of the inhibitor were compared to the locations of the eleven lysine residues within the domain, as shown in Figure 3.24, in order to attempt to determine which lysine residues might be involved in imine formation with norstictic acid.

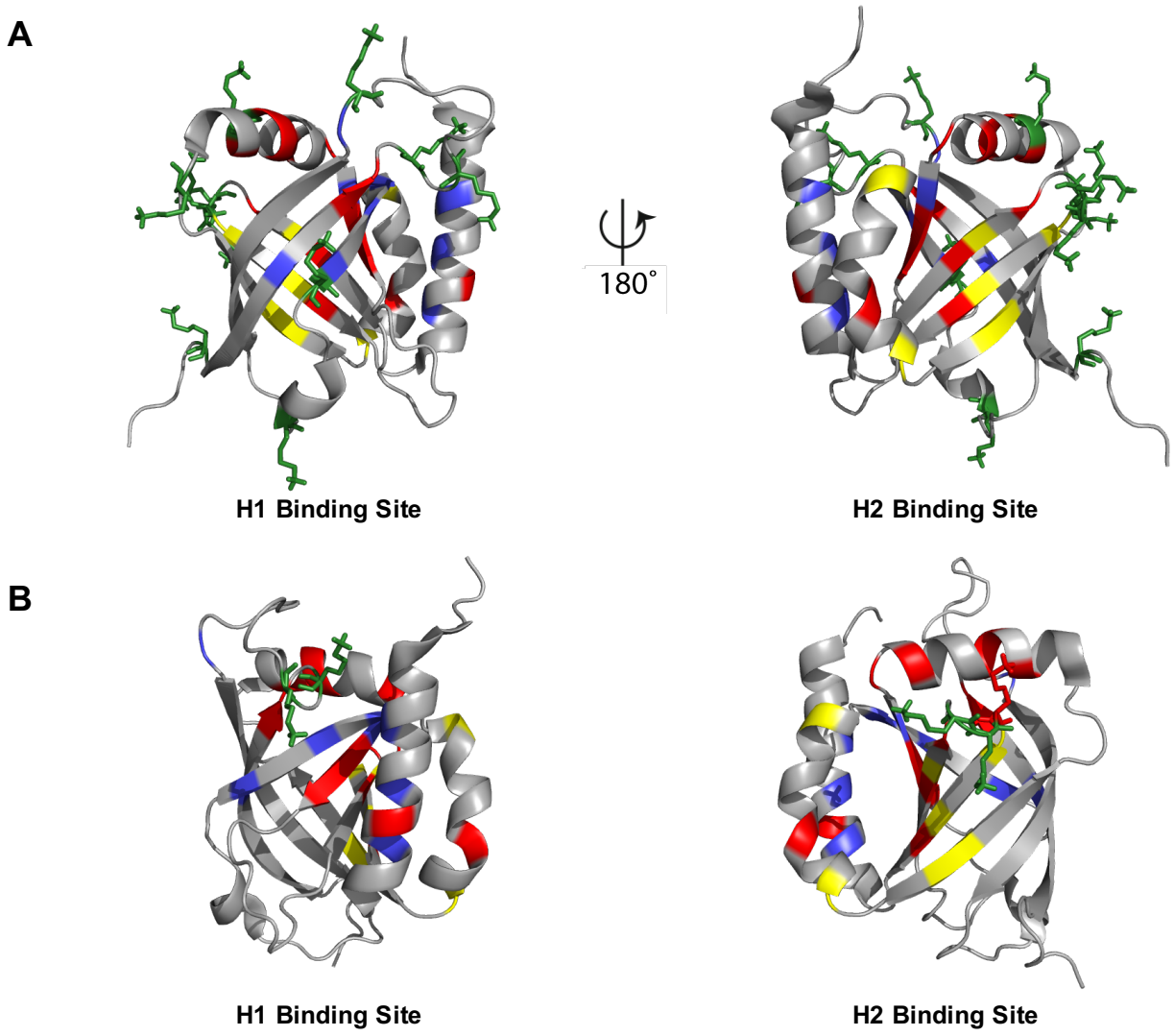


***Figure 3.24*- Proximity of Lysine Residues to Norstictic Acid Induced Chemical Shifts** (A) The eleven lysine residues within the AcID motif are shown as green sticks, while residues perturbed by the binding of norstictic acid are shown in red, and perturbations induced by binding of the VP16 TAD to the H1 and H2 binding sites are shown in blue and yellow, respectively. (B) K411 and K413 (green sticks) are located on a flexible loop within the H1 binding site of AcID. K518 (green sticks), K519 (green sticks), and K520 (red sticks) are located on a small flexible loop within the H2 binding site of the domain.

Highlighting the location of the lysine residues within the AcID motif in conjunction with mapping the chemical shifts induced by the binding of norstictic acid reveals a number of lysine residues in close proximity to the inhibitor induced perturbations. In particular, K411 and K413 are located on a large, highly flexible loop within the H1 binding site that may account for the observed chemical shifts induced by norstictic acid following imine formation. Furthermore, K518, K519, and K520 are located on a short flexible loop within the H2 binding site that may similarly account for chemical shifts following imine

formation. The close spatial arrangement of these residues, K411/K413 and K518/K519/K520, also result in significant regions of dense positive charge on the AcID surface, which may assist in the localization of norstictic acid to preferred binding locations through critical electrostatic contacts. Additionally, K484 is a potential residue of interest, as it is located on helix α2 and lies at the interface of the H1 and H2 binding sites on the AcID protein. Thus, based on the combined evidence provided by MD simulations, NMR perturbations, and mass spectrometric analysis of the covalent norstictic acid•AcID complex, it was hypothesized that K411, K413, K18, K519, and K520 may be potential sites for imine formation with norstictic acid.

Mutational Analysis to Identify Sites of Covalent Modification by Norstictic Acid

In order to attempt to support the hypothesis that these lysine residues may function as inhibitor binding sites, a mutational analysis was next completed in which the nucleophilic lysine residues were mutated to non-nucleophilic arginine residues. This mutation conserves the presence of a positive charge at the aforementioned locations on the AcID motif, and thus should have a minimal impact on the binding of the VP16 TAD as electrostatic contacts should be maintained, but removes the ability of norstictic acid to covalently bind through the formation of an imine as arginine lacks a nucleophilic amine. The ability of norstictic acid to inhibit the binding of VP16(438-454) and VP16(467-488) to the mutant proteins was then assessed. A decrease in the potency of the molecule against the mutant proteins would suggest that the inhibitor has lost a site of covalent binding. In order to accomplish this, site-directed mutagenesis was performed on the plasmid encoding the wild type domain in order to generate a series of single and multiple mutants, which were then expressed and purified following standard procedures.

Initially, lysine residues K518, K519, and K520 within the H2 binding site were mutated as the presence of three sequential lysine residues results in a region of densely packed positive charge, representing a highly attractive site for norstictic acid. Three single point mutants, K518R, K519R, and K520R; a double point mutant K518R/K519R; and a triple point mutant K18R/K519R/K520R were generated. As a note on the naming conventions we have elected to use in referring to the lysine to arginine mutants, KK518RR refers to the double mutant K518R/K519R and KKK518RRR refers to the triple

mutant K518R/K519R/K520R. After expressing and purifying the mutants, their ability to bind VP16(438-454) and VP17(467-488) was tested in order to confirm that the mutations did not significantly alter the structure or binding competence of the domain, as shown below in Figure 3.25 and summarized in Table 3.1.

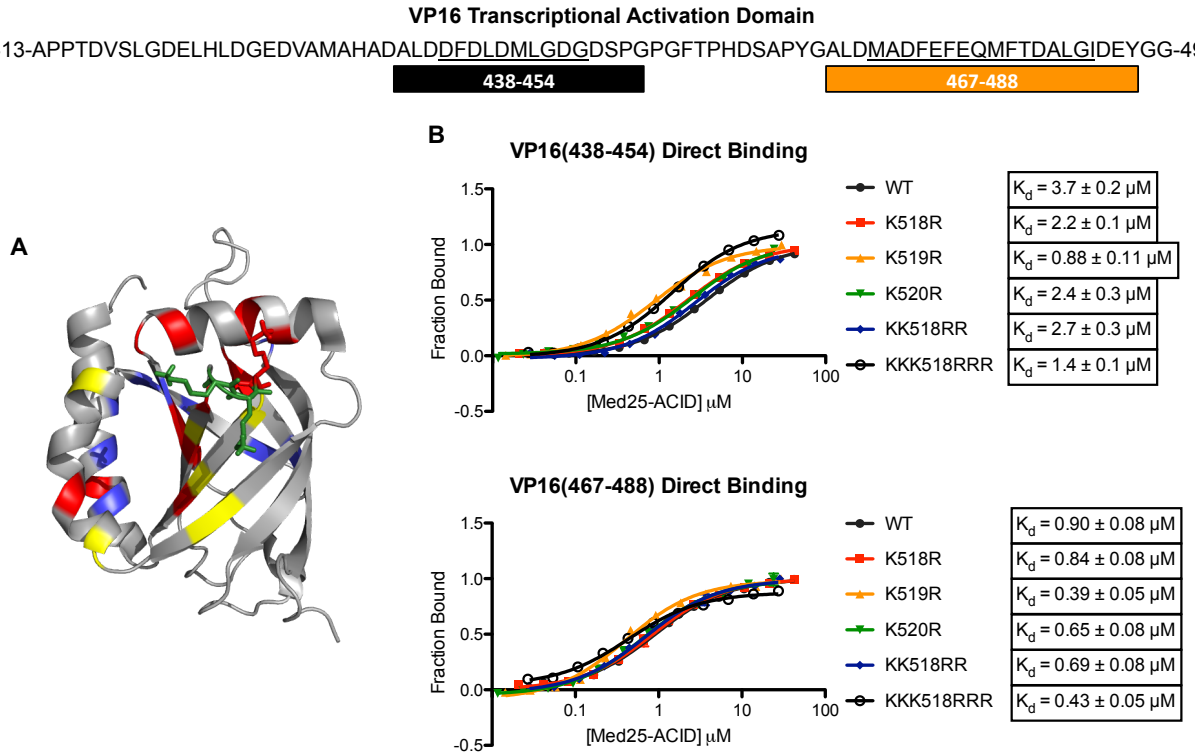


***Figure 3.25*- Direct Binding of VP16(438-454) and VP16(467-488) to H2 Lysine Mutants** (A) Structure of the AcID motif with perturbations induced by the binding of norstictic acid shown in red, perturbations induced by the binding of VP16 H1 and H2 shown in blue and yellow, respectively, and K518/K519/K520 shown as sticks within the H2 binding site. (B) Direct binding curves of VP16(438-454) and VP16(467-488) to various AcID constructs. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the fraction of tracer bound at the indicated AcID concentration. Curves were fit with GraphPad Prism 5.

Table 3.1- Effect of H2 Binding Site Lysine to Arginine Mutations on VP16 Affinity

	VP16(438-454)		*VP16(467-488)*	
	Kd	**Fold Change**	**Kd**	**Fold Change**
WT	3.7 ± 0.2	---	0.90 ± 0.08	---
K518R	2.2 ± 0.1	-1.68	0.84 ± 0.08	-1.07
K519R	0.88 ± 0.11	-4.20	0.39 ± 0.05	-2.31
K520R	2.4 ± 0.3	-1.54	0.65 ± 0.08	-1.38
KK518RR	2.7 ± 0.3	-1.37	0.69 ± 0.08	-1.30
KKK518RRR	1.4 ± 0.1	-2.64	0.43 ± 0.05	-2.09

As expected, the conservative lysine to arginine mutations did not have a deleterious effect on the affinity of VP16(438-454) or VP16(467-488) for AcID. In fact, the affinity of the VP16 derived peptides for the domain was slightly enhanced following mutation of the indicated lysine residues to arginine, though this enhancement is not significant and the mechanism by which this occurs is unclear.

The effect that these mutations had on the activity of norstictic acid against the interaction was then assessed by determining the $IC_{50\ (30\ Min)}$ values for the inhibition of the interaction between VP16(438-454) or VP16(467-488) and AcID. In all cases the molecule was tested against a VP16•AcID complex in which the tracer was 50% bound and all samples were incubated for thirty minutes to account for potential time-dependent inhibition. The results of these experiments are shown below in Figure 3.26 and summarized in Table 3.2.

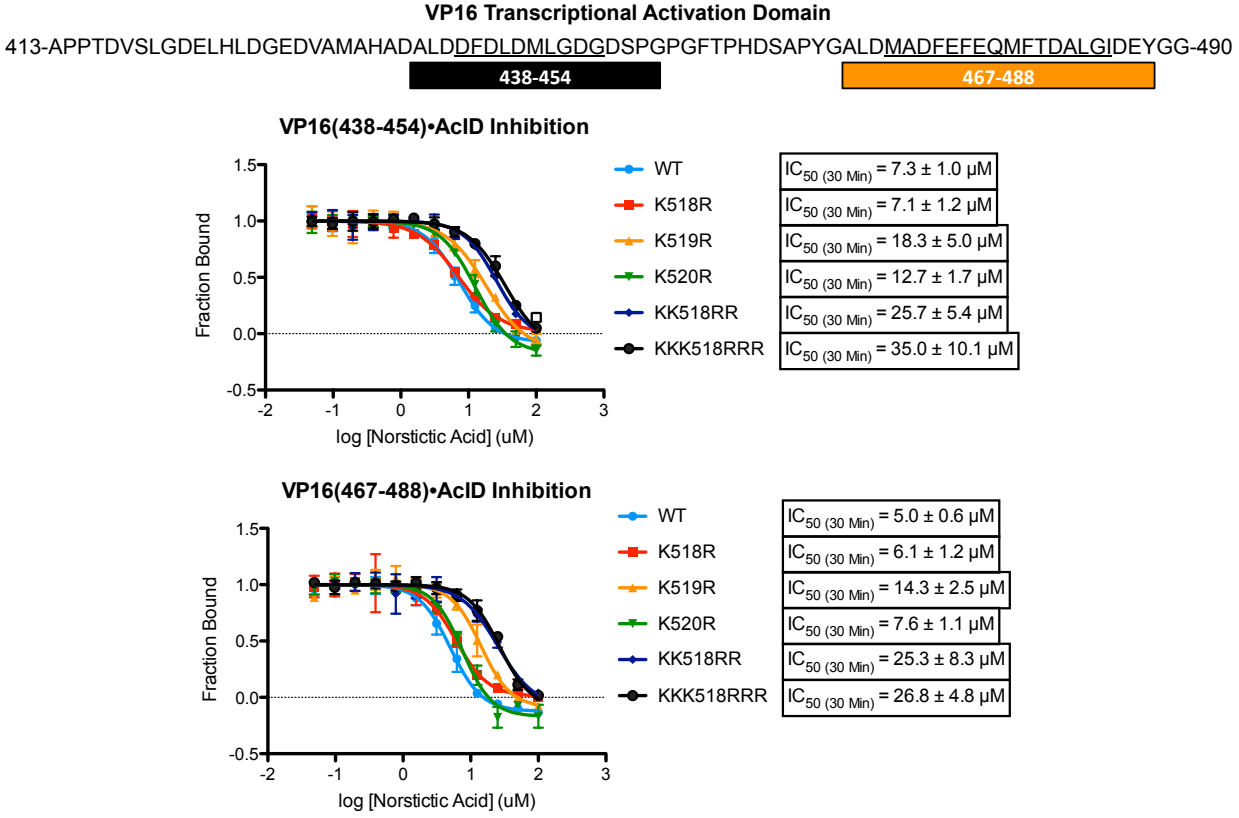


***Figure 3.26*- Effect of H2 Lysine Mutations on Norstictic Acid Inhibition of VP16•AcID.** Competition experiments of norstictic acid for the VP16(438-454)•AcID and VP16(467-488)•AcID

interactions with the various AcID constructs. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the relative fraction of tracer bound in the presence of the indicated concentration of norstictic acid. Curves were fit using GraphPad Prism 5.

Table 3.2- Effect of H2 Binding Site Lysine to Arginine Mutations on Norstictic Acid Inhibition

	VP16(438-454)		*VP16(467-488)*	
	IC50 (30 Min)	**Fold Change**	**IC50 (30 Min)**	**Fold Change**
WT	7.3 ± 1.0	---	5.0 ± 0.6	---
K518R	7.1 ± 1.2	-1.03	6.1 ± 1.2	1.22
K519R	18.3 ± 5.0	2.51	14.3 ± 2.5	2.86
K520R	12.7 ± 1.7	1.74	7.6 ± 1.1	1.52
KK518RR	25.7 ± 5.4	3.52	25.3 ± 8.3	5.06
KKK518RRR	35.0 ± 10.1	4.79	26.8 ± 4.8	5.36

These data demonstrate that the sequential mutation of lysine residues within the K518/K519/K520 loop region results in the significant loss of inhibition by norstictic acid. The triple mutant KKK518RRR results in an approximately five-fold decrease in the potency of the small molecule for AcID interactions with both VP16(438-454) and VP16(467-488). The individual point mutations of K518, K519, and K520 fail to recapitulate this loss in activity, suggesting that the residues may be interchangeable as sites for imine formation with norstictic acid, which is reasonable given their immediate proximity. Of the three single point mutants, only K519R demonstrated significant perturbation in the activity of norstictic acid, suggesting that this residue may be of particular importance. The side chain of K519 points into the H2 binding surface and thus in prime position for disruption of H2 binding site dependent interactions following imine formation with norstictic acid, whereas K518 and K520 point out of the binding site and are in close proximity to helix α2. The double mutant KK518RR failed to fully recapitulate the loss of activity observed with the KKK518RRR mutant, but had significantly stronger effects than any of the single mutants, further supporting the hypothesis that the three lysine residues within this loop may be interchangeable as sites for the covalent attachment of norstictic acid.

The observation that mutations within the H2 binding site also affect the ability of norstictic acid to perturb the binding of VP16(438-454) to the H1 binding site of AcID was perhaps unexpected, but is consistent with an allosteric mechanism of action for inhibition

by norstictic acid. As was discussed previously, the binding of norstictic acid induces a number of significant chemical shifts within helix α2, which is located at the interface between the H1 and H2 binding sites, that may result in structural changes within both binding sites. Thus, the loss of norstictic acid binding at K518, K519, or K520 may mitigate the induced α2 structural changes, thereby disrupting structural changes within the H1 binding site resulting in a concomitant decrease in potency against interactions that occur at the H1 site.

Following the observation that mutation of lysine residues to arginine within the H2 binding site led to a decrease in the potency of norstictic acid, it was next sought to confirm that this effect was a result of the decreased covalent binding of norstictic acid to the mutants, relative to the wild type domain. In order to accomplish this, each of the mutants was incubated with four equivalents of norstictic acid at room temperature for two hours and resulting complexes were analyzed by mass spectrometry in order to determine the degree of labeling of the AcID protein by norstictic acid. The results of these experiments are shown below in Figure 3.27.

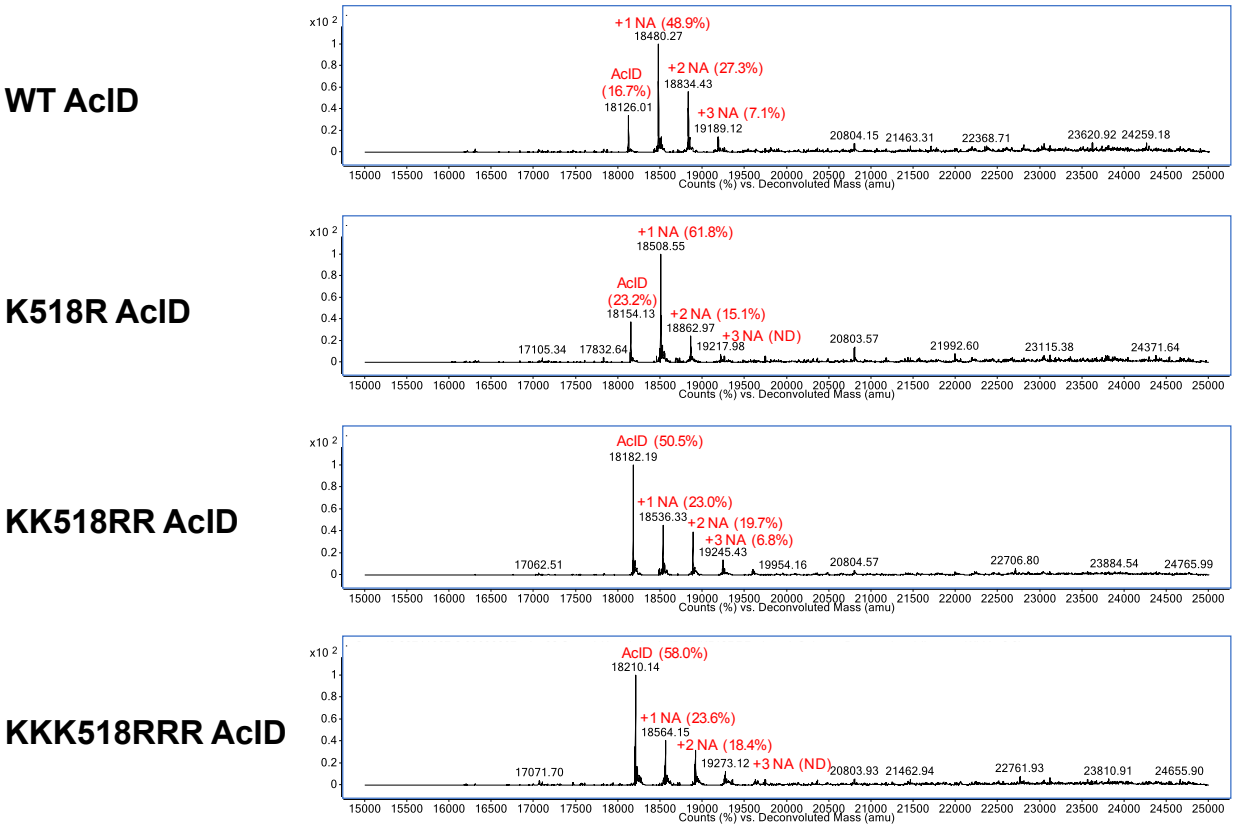


***Figure 3.27*- Effect of Lysine Mutation on the Covalent Labeling of AcID by Norstictic Acid.** The indicated AcID construct was incubated with four equivalents of norstictic acid at room temperature for two hours prior to being subjected to analysis by mass spectrometry. Peaks of interest are labeled in the spectra with 'AcID' referring to the mass of the unlabeled construct and '+X NA' referring to the number of norstictic acid molecules bound to the protein. Percentages refer to the approximate composition of the indicated species in the sample as determined by individual peak intensity over the total intensity of peaks of interest.

Consistent with the observations from the competition experiments with the various AcID mutants, mutation of lysine residues within the H2 binding site resulted in a decrease in the degree of covalent labeling of AcID by norstictic acid. The single point mutant K518R did not significantly alter the labeling of the domain, consistent with the hypothesis that the three lysine residues may be interchangeable sites for imine formation. The double mutant KK518RR and triple mutant KKK518RRR resulted in significantly reduced labeling of the domain by norstictic acid, consistent with the observations that these mutations also reduced the potency of the molecule in inhibiting VP16•AcID interactions. Taken together, these data suggest that the three lysine residues K518, K519, and K520 are likely sites for the covalent binding of norstictic acid to the AcID motif. However, the triple mutant KKK518RRR is still labeled to a moderate degree by norstictic acid,

suggesting that these residues are not the only binding site for norstictic acid on the domain.

The $^{1}H,^{15}N$-HSQC perturbation experiments also indicated that K411 or K413 may also be sites for the covalent binding of norstictic acid, given their proximity to a number of observed chemical shifts induced by the binding of the inhibitor. In order to confirm the role of these residues in the activity of norstictic acid, the same battery of experiments was completed for K411 and K413 as had been completed for K518, K519, and K520. After expressing and purifying the AcID mutant K411R/K413R the ability of VP16(438-454) and VP16(467-488) to bind to the mutant domain was assessed. The results of those experiments are shown below in Figure 3.28 and summarized in Table 3.3.

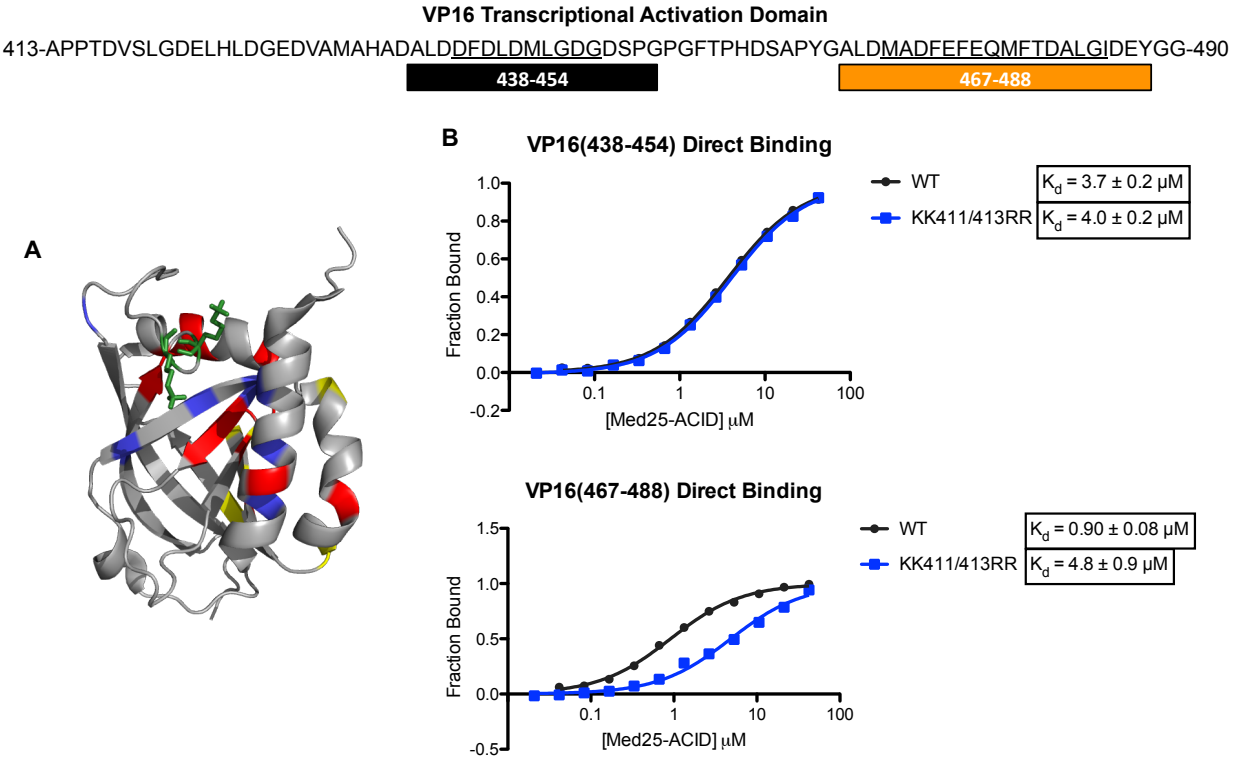


***Figure 3.28*- Direct Binding of VP16(438-454) and VP16(467-488) to H1 Lysine Mutants.** (A) Structure of the AcID motif with perturbations induced by the binding of norstictic acid shown in red, perturbations induced by the binding of VP16 H1 and H2 shown in blue and yellow, respectively, and K411/K413 shown as sticks within the H1 binding site. (B) Direct binding curves of VP16(438-454) and VP16(467-488) to wild type and K411R/K413R AcID. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the fraction of tracer bound at the indicated AcID concentration. Curves were fit with GraphPad Prism 5.

Table 3.3- Effect of H1 Binding Site Lysine to Arginine Mutations on VP16 Affinity

	VP16(438-454)		*VP16(467-488)*	
	Kd	**Fold Change**	**Kd**	**Fold Change**
WT	3.7 ± 0.2	---	0.90 ± 0.08	---
K411R/K413R	4.0 ± 0.2	1.08	4.8 ± 0.9	5.33

The binding of VP16(438-454) is unperturbed by the mutations K411R and K413R as the K_d value of the peptide for the mutant protein is within error of the K_d value for the wild type protein. This was the expected result as the mutation to arginine maintains the positive charge, and thus any critical electrostatic contacts with the VP16 TAD. The binding of VP16(467-488) is moderately perturbed by the mutations, which is unexpected given that the peptide likely binds to the H2 binding site and not near the mutated residues. This finding suggests that the mutation of K411 and K413 may have induced minor structural changes within the H2 binding site, which is plausible given that K411 and K413 are located on a large and highly flexible loop near the interface of the H1 and H2 binding sites. Additional mechanistic and structural studies will be necessary to determine the source of this perturbation, though it appears to support the hypothesis that AcID is relatively plastic and the distinct binding surfaces may be in allosteric communication.

The ability of norstictic acid to perturb interactions between the K411R/K413R double mutant and VP16(438-454) and VP16(467-488) was then compared to the activity against the interactions with the wild type protein, as shown below in Figure 3.29 and summarized in Table 3.4. In all cases the molecule was tested against a VP16•AcID complex in which the tracer was 50% bound and all samples were incubated for thirty minutes to account for potential time-dependent inhibition.

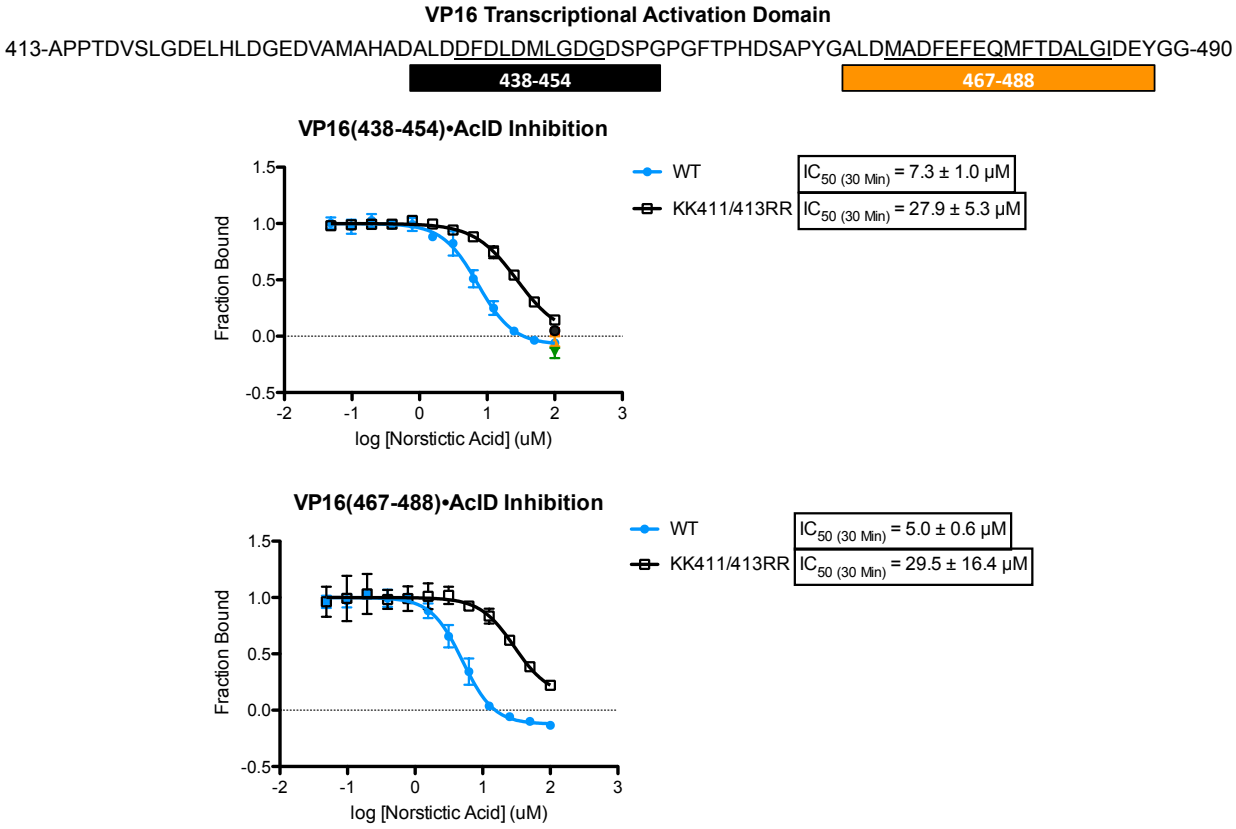


***Figure 3.29*- Effect of H1 Lysine Mutations on Norstictic Acid Inhibition of VP16•AcID.** Competition experiments of norstictic acid for the VP16(438-454)•AcID and VP16(467-488)•AcID interactions with the H1 double mutant and wild type AcID. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the relative fraction of tracer bound in the presence of the indicated concentration of norstictic acid. Curves were fit using GraphPad Prism 5.

Table 3.4- Effect of H1 Binding Site Lysine to Arginine Mutations on Norstictic Acid Inhibition

	VP16(438-454)		*VP16(467-488)*	
	IC50 (30 Min)	**Fold Change**	**IC50 (30 Min)**	**Fold Change**
WT	7.3 ± 1.0	---	5.0 ± 0.6	---
K411R/K413R	27.9 ± 5.3	3.82	29.5 ± 16.4	5.90

Consistent with the hypothesis that K411 and K413 are sites of covalent modification by norstictic acid, mutation of these residues to arginine significantly perturbed the ability of norstictic acid to inhibit VP16(438-454) and VP16(467-488) interactions with AcID. The degree of perturbation was similar to the results observed for the H2 site mutations, with the potency of the molecule decreasing four to six fold

following mutation of the lysine residues to arginine. Interestingly, the K411 and K413 mutations affected the ability of the molecule to disrupt interactions at both the H1 and H2 binding sites, similar to the effects observed following the mutation of lysine residues within the H2 site. This finding further supports the hypothesis that the H1 and H2 binding sites are in allosteric communication and suggests that lysine residues within both binding sites may play a critical role in these structural changes, possibly through interaction with helix α2.

In order to further support K411 and K413R as sites of covalent modification by norstictic acid, the ability of the molecule to label the mutant protein was assessed by mass spectrometry, as shown in Figure 3.30.

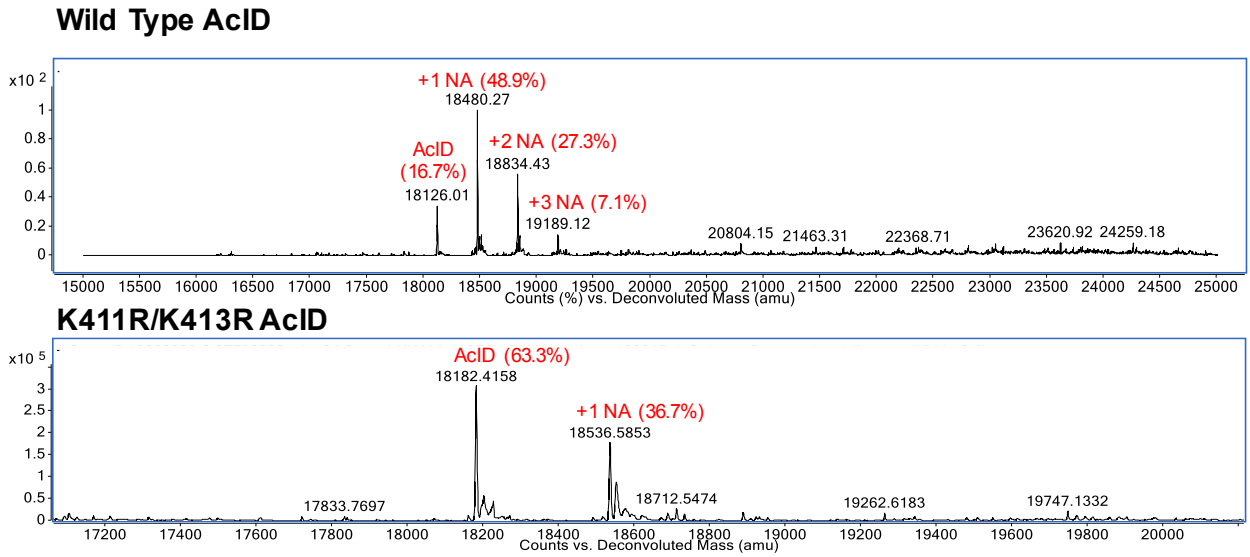


***Figure 3.30*- Effect of H1 Lysine Mutations on the Covalent Labeling of AcID by Norstictic Acid.** The H1 double mutant AcID protein was incubated with four equivalents of norstictic acid for two hours at room temperature and then subjected to analysis by mass spectrometry. Peaks of interest are labeled in the spectra with 'AcID' referring to the mass of the unlabeled protein and "+X NA" referring to the masses of protein labeled with the indicated number of norstictic acid molecules. Percentages refer to the approximate composition of the indicated species in the sample as determined by individual peak intensity over the total intensity of peaks of interest.

The H1 site double mutant K411R/K413R was incubated with four equivalents of norstictic acid for two hours at room temperature and the resulting complex was then analyzed by mass spectrometry and compared to the labeling observed for treatment of the wild type domain with norstictic acid. The mutations result in a significant decrease in the labeling of the domain by norstictic acid, consistent with observations that the molecule is significantly less potent in inhibiting VP16•AcID interactions with the mutated

construct. Additionally, this finding is consistent with previous observations of the effects of lysine mutations within the H2 binding site. Residual labeling of the mutated K411R/K413R AcID may result from the binding of the molecule to a residue within the H2 binding site, or to an alternative lysine residue. Taken together, these data are consistent with the hypothesis that K411 and K413 are potential sites for the covalent modification of the AcID motif.

As an additional method to support the hypothesis that K411, K413R, K518, K519, and K520 are potential sites for the covalent binding of norstictic acid to the AcID motif, a mutant AcID construct in which all five residues were mutated to arginine was expressed and purified, which we refer to as the 'Quint. Mutant' in the data below. Following the expression and purification of the quintuple mutant K411R/K413R/K518R/K519R/K520R, the ability of the mutant to bind VP16(438-454) and the ability of norstictic acid to inhibit the interaction of VP16(438-454) with the mutant AcID construct was determined, as shown below in **Figure 3.31**.

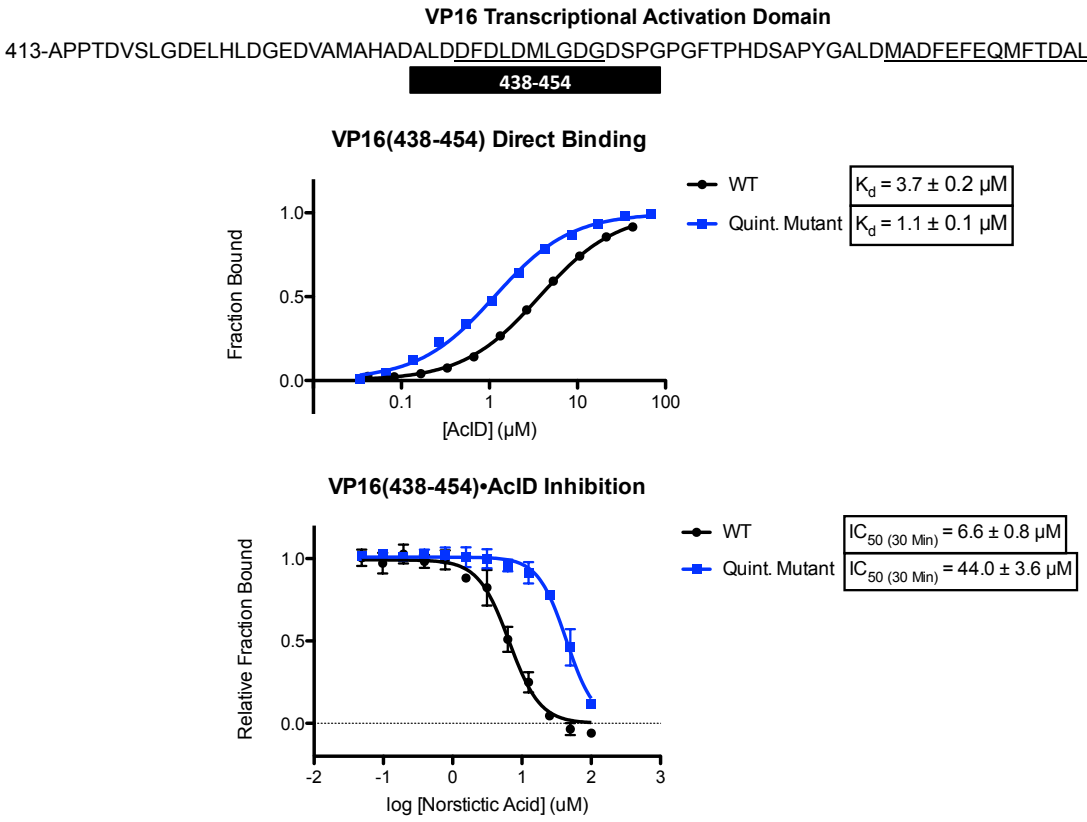


***Figure 3.31*- Effect of H1 and H2 Lysine Mutations on VP16(438-454)•AcID and Norstictic Acid Inhibition.** Direct binding of VP16(438-454) to the quintuple lysine to arginine AcID mutant is shown in the top figure. An inhibition experiment of norstictic acid on the interaction between VP16(438-454) is shown in the bottom figure. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the fraction of tracer bound at the indicated concentration of AcID or norstictic acid. Curves were fit using GraphPad Prism 5.

The direct binding experiments indicate that the mutation of the five lysine residues to arginine does not significantly perturb the ability of the construct to bind VP16(438-454), consistent with expectations and previous observations of the effects of mutation on the lysine residues within the H1 and H2 binding site individually. Furthermore, the potency of norstictic acid is decreased approximately seven fold by the mutation of the five lysine residues, which is a greater perturbation than was observed for the mutation of lysine residues within the H1 or H2 binding site individually, further supporting the hypothesis that both sites may simultaneously be labeled by norstictic acid. The residual effects of norstictic acid for inhibition of VP16(438-454)•AcID may be a result of the competition of the molecule for TAD binding sites in the absence of covalent linkage to

the domain, or may be the result of the molecule binding covalently to alternative lysine residues that have yet to be identified.

As further support of these five lysine residues as potential sites for covalent binding of norstictic acid, the ability of the molecule to covalently label the mutant AcID construct was determined by mass spectrometry and compared to the labeling observed with wild type AcID protein, as shown below in Figure 3.32.

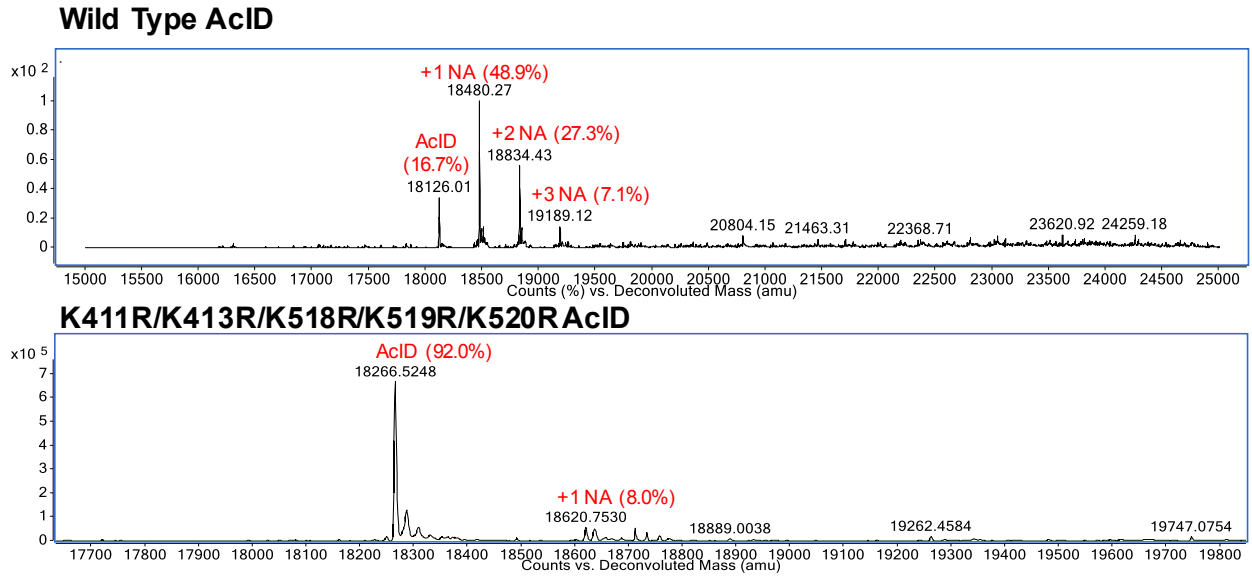


***Figure 3.32*- Effect of H1 and H2 Mutations on the Covalent Labeling of AcID by Norstictic Acid.** The quintuple mutant AcID protein was incubated with four equivalents of norstictic acid for two hours at room temperature and then subjected to analysis by mass spectrometry. Peaks of interest are labeled in the spectra with 'AcID' referring to the mass of the unlabeled protein and "+X NA" referring to the masses of protein labeled with the indicated number of norstictic acid molecules. Percentages refer to the approximate composition of the indicated species in the sample as determined by individual peak intensity over the total intensity of peaks of interest.

Mass spectrometric results of the AcID quintuple mutant demonstrate that the mutation of lysine residues in both the H1 and H2 binding sites severely attenuates the ability of norstictic acid to covalently label AcID, with only a very small fraction, approximately 8%, of the protein bearing a single norstictic acid modification. These data suggest that K411, K413, K518, K519, and K520 represent the most significant binding sites on the domain for norstictic acid, as their mutation to arginine significant reduces the potency of the small molecule in the inhibition of VP16•AcID interactions. The trace amounts of protein that are still labeled in this sample are likely the result of nonspecific reactions between norstictic acid and the AcID motif. Thus, the observed IC_{50} of 44 μM for norstictic acid against the interaction between VP16 and the quintuple mutant is likely

the activity of the molecule in the absence of the formation of a covalent adduct. Therefore, because the molecule was in equilibrium under these conditions, the K_i of the inhibitor for the AcID motif was calculated and determined to be 22.1 μM.

Overall, the above mutational analysis experiments confirm that norstictic acid covalently binds to the AcID motif through the formation of an imine between the reactive amine of lysine residues within the domain and an aldehyde present on norstictic acid. Lysine residues within the H1 (K411 and K413) and H2 (K518, K519, and K520) binding sites on the domain appear to be critical sites of covalent modification as demonstrated by a series of inhibition assays and mass spectrometric analysis. $^{1}H,^{15}N$-HSQC chemical shift perturbation experiments indicate that these lysine residues are within close proximity to perturbations induced by the binding of norstictic acid and are located within the VP16 H1 and H2 binding sites on the domain. Functionally, the formation of these imines may interfere with critical electrostatic contacts between the VP16 TAD and AcID by relieving the positive charge of the targeted lysine residues. Additionally, the binding of norstictic acid at these sites may add considerable steric bulk to surfaces important for the binding of the VP16 TAD, thereby further inhibiting the ability of the TAD to bind to the domain. In order to validate this hypothesis, additional AcID mutants in which the lysine residues in question are mutated to alanine or glutamate will be prepared. The mutation of the lysine residues to alanine will remove positively charged contacts from the surface of the protein without significantly altering local secondary structure[55], thereby supporting the role of these residues as critical points of electrostatic interaction. As a result, the expectation would be for the binding of the VP16 TAD to be compromised and norstictic acid to be far less potent as it may no longer be electrostatically attracted to a critical surface of AcID. Additional mutation of these residues to glutamate results in a charge inversion, which one would expect to abrogate the binding of VP16 and the activity of norstictic acid to an even more significant extent than observed for the mutation of these sites to alanine as a result of significant repulsion between the negatively charged TAD or small molecule and the negative surface of the AcID protein.

In addition, the mutational analyses and $^{1}H,^{15}N$-HSQC data suggest that the molecules can induce allosteric changes within the domain. Thus, the binding of a single norstictic acid molecule to AcID may inhibit the binding of activators at both interaction

surfaces by orthosterically inhibiting interactions at the binding site to which the molecule is bound and allosterically inhibiting interactions at the secondary site through conformational modulation of the domain, consistent with the conformational plasticity following activator binding discussed in Chapter 2. The mutational analysis data suggests that norstictic acid does not have a particular preference for either the H1 or H2 binding site, suggesting that the allosteric modulation of the domain is bidirectional, with binding at either site inducing allosteric shifts in the alternative site. This hypothesis is demonstrated below in Figure 3.33.

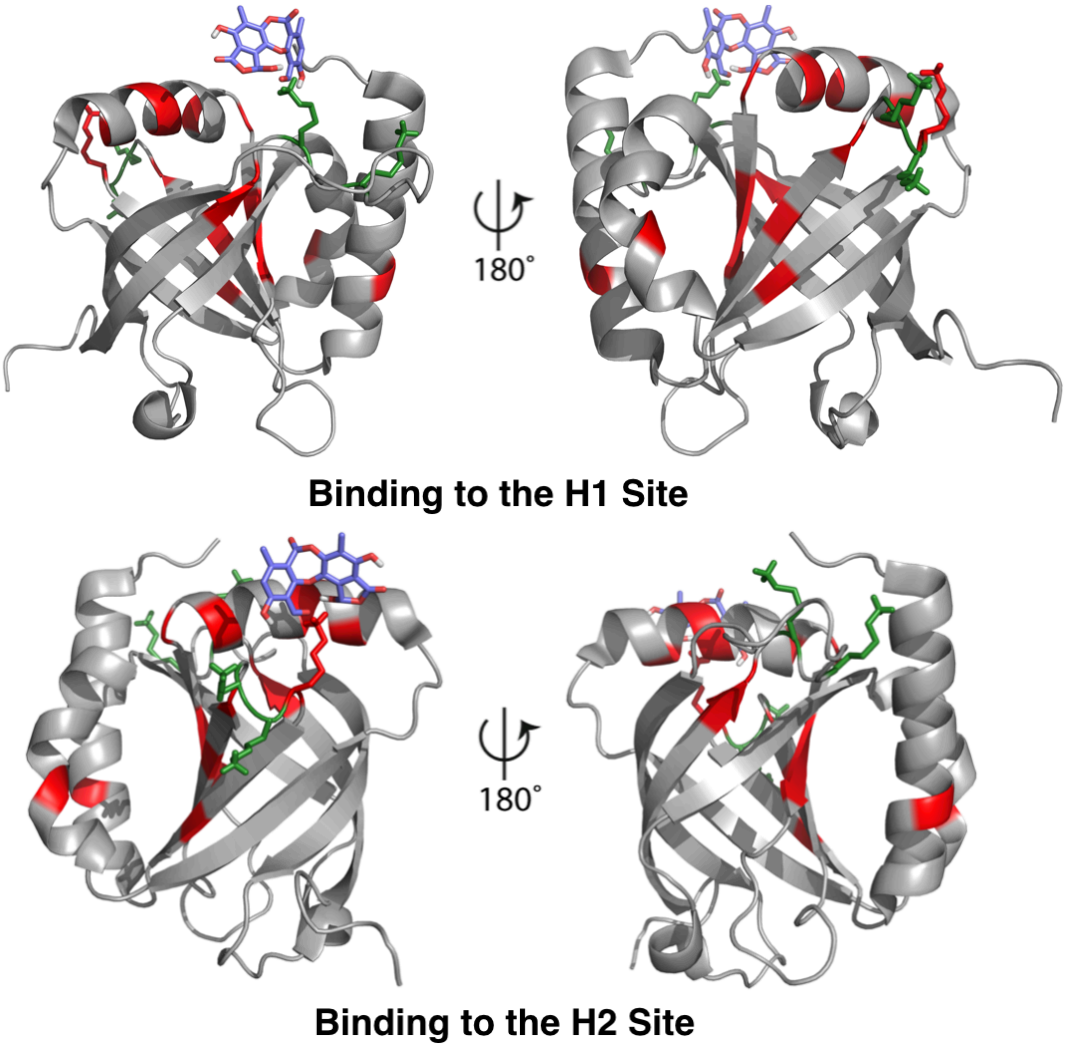


Figure 3.33- Norstictic Acid is a Mixed Orthosteric/Allosteric Inhibitor of the AcID Motif The binding of norstictic acid at either the H1 or H2 binding site orthosterically inhibits interactions at that binding surface and induces conformational changes within the alternative binding site that inhibits activator interactions.

Attempts at Identifying Sites of Covalent Labeling of AcID by Norstictic Acid Using Protein Digestion and Mass Spectrometric Analysis

In order to attempt to further support that K411, K413, K518, K519, and K520 are sites for the covalent binding of norstictic acid to the AcID motif, we next sought to use a combination of protein digestion and mass spectrometric analysis to determine which specific lysine residues were covalently modified by the small molecule. Towards this end, a sample of wild type AcID protein was incubated with four equivalents of norstictic acid for two hours at room temperature and subsequently reduced with sodium borohydride for an hour. The samples were then checked by mass spectrometry to confirm that the protein had been covalently labeled by the inhibitor as shown in Figure 3.34.

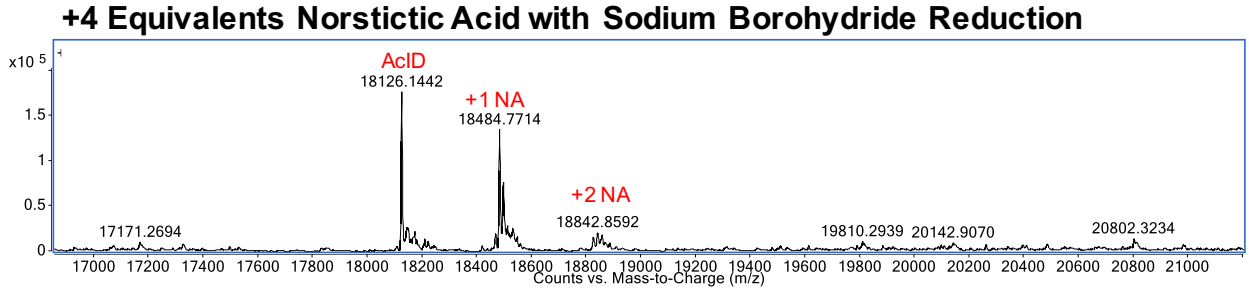


Figure 3.34- AcID Labeled with Norstictic Acid Submitted for Proteomics Analysis Wild type acid protein was incubated with four equivalents of norstictic acid for two hours and then reduced with sodium borohydride for one hour prior to analysis by mass spectrometry.

After confirming that the AcID protein had been labeled by norstictic acid, it was subsequently purified by SDS-PAGE and the protein bands of interest were stained with Coomassie blue and excised. The excised gel slices were then submitted to the Proteomics and Peptide Synthesis Core at the University of Michigan for digestion and proteomic analysis. The labeled AcID protein was then digested in gel overnight using chymotrypsin and the resulting peptide fragments were extracted and analyzed by mass spectrometry for covalent modification of lysine residues within the resulting peptides. The mass spectrometric analysis of the fragments generated following chymotrypsin digestion resulted in the observation of fragments that covered greater than 95% of the AcID protein including all eleven of the lysine residues. Unfortunately, the analysis failed to reveal the norstictic acid modification (+358.06 Da) on of any of the lysine residues

within the resulting peptide fragments, despite clear evidence that the intact protein was labeled with the small molecule.

One possibility for the lack of observed norstictic acid modifications is that the MS/MS conditions under which the peptidic fragments are analyzed are a significantly harsher environment that results in the fragmentation or degradation of the covalently bound norstictic acid molecule. This phenomenon has been previously reported in the literature during an MS/MS characterization of lichen-derived natural products, including norstictic acid.[56] The norstictic acid fragmentation products observed under MS/MS conditions reported in the literature are shown in panel A of Figure 3.35. Notably, the modifications shown in Figure 3.35 include the reduction of the molecule that occurs during sample preparation prior to proteomic analysis and thus are slightly different than the fragmentations reported in the literature. Additionally, storing the norstictic acid labeled protein or the peptidic fragments following digestion with chymotrypsin in aqueous buffer for extended periods of time could theoretically result in the hydrolysis of the central 7-membered lactone or the formation of a new 5-membered lactone from the reduced aldehyde and carboxylic acid if the buffer were slightly acidic. These possible modifications to the covalently bound norstictic acid molecule are shown in panel B of Figure 3.35.

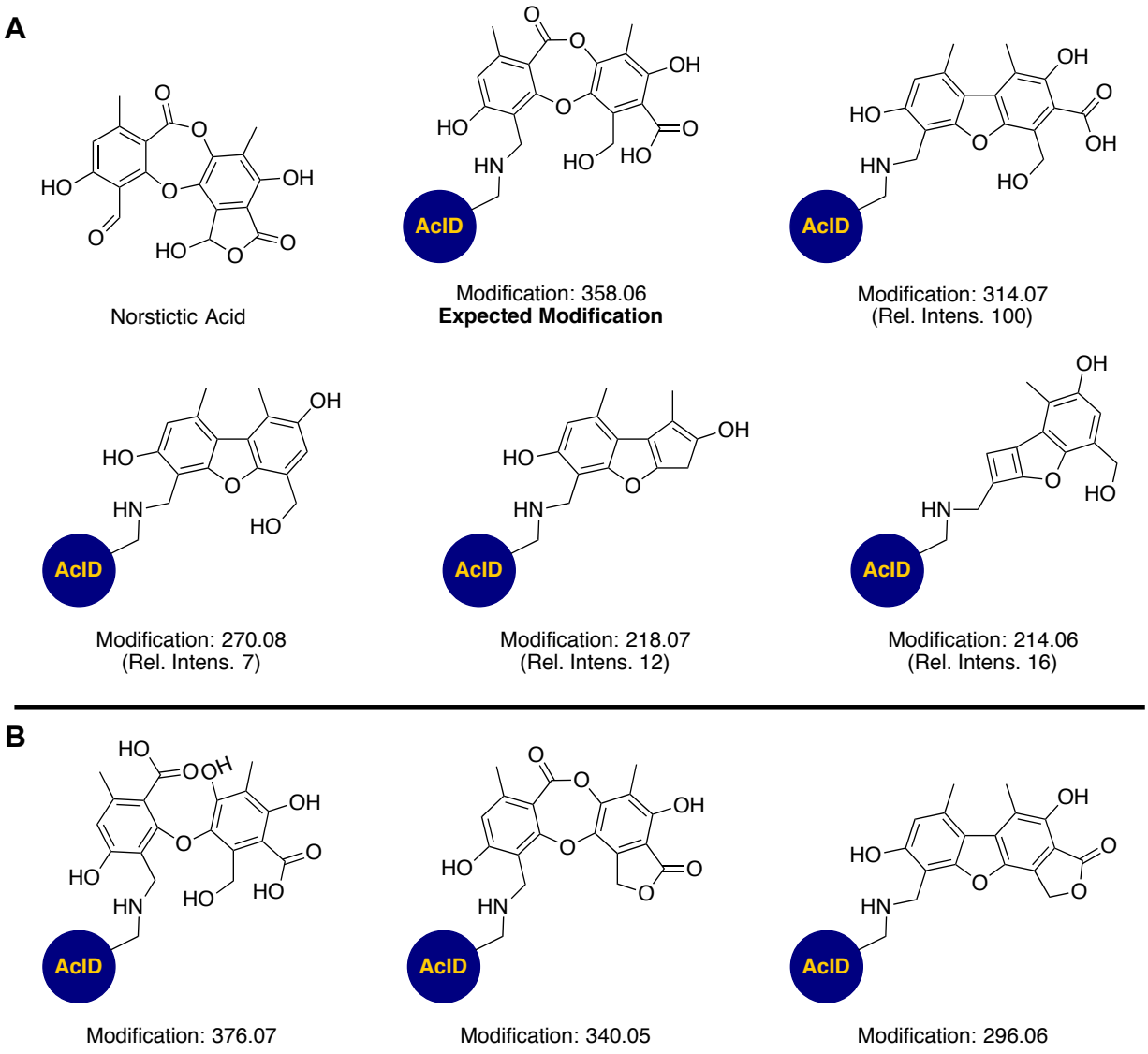


***Figure 3.35*- Possible Modifications of Norstictic Acid During MS/MS Analysis.** (A) Possible fragmentations of norstictic acid under harsh MS/MS conditions, as reported in the literature.[56] These potential modifications include the reduction of the second aldehyde, which occurs following treatment of the sample with sodium borohydride during preparation. (B) Possible modifications of the bound norstictic acid molecule that may occur under storage in aqueous environments.

The original data collected following chymotrypsin digestion and mass spectrometric analysis was analyzed a second time to search for the modifications shown in Figure 3.35. Unfortunately, this subsequent analysis again failed to identify any modification of the lysine residues on the peptidic fragments following chymotrypsin digestion. One possible explanation for this phenomenon is that the treatment of the molecule by sodium borohydride may cause the resulting reduced compound to fragment through unique mechanisms not observed for the parent molecule during MS/MS

analysis. As such, it is possible that the correct mass for the norstictic acid modifications on the lysine containing fragments has still not been identified. An additional possibility is that modification of the fragments by norstictic acid may alter the solubility of the resultant peptides or otherwise interfere with the ability to extract these fragments from the SDS-PAGE slices following digestion. In this case, the sample may simply not contain enough of the labeled fragments to be effectively observed by MS/MS analysis. Given the importance of this data in validating the binding site of norstictic acid on the AcID motif, this experiment will be repeated using a sample in which the norstictic acid•AcID complex is not treated with sodium borohydride in addition to a fresh sample that has been reduced. In the event that these samples still fail to generate conclusive evidence of the norstictic acid binding sites on AcID, a more thorough analysis of the effects of MS/MS conditions on covalently bound norstictic acid will be undertaken in order to determine the most abundant fragmentation products before repeating the analysis of the digested material.

Preliminary Cellular Activity of Norstictic Acid and Psoromic Acid

We next sought to determine if norstictic acid and the closely related lead molecule psoromic acid could perturb AcID-dependent transcriptional processes within a cellular context. The first assay by which this hypothesis was tested involved examining the expression of endogenous HSPA5 by Real Time quantitative Polymerase Chain Reaction (RT-qPCR) of reverse transcribed mRNA transcripts. HSPA5 is a canonical target gene of the endoplasmic reticulum stress response transcription factor activating transcription factor 6 (ATF6α), which is dependent upon interaction with the AcID motif of Med25 for full transcriptional activation.[42] This interaction occurs through an eight amino acid sequence within the TAD of ATF6α, DFDLDLMP, which bares striking similarity to the VP16 sequence, DFDLDMLG, which is termed VN8.[57] Free VN8 peptide is capable of competing with the interaction between ATF6α and Med25, indicating that the ATF6α TAD binds to a similar surface within the domain as VP16. Furthermore, siRNA mediated knockdown of Med25 expression has been shown to significantly abrogate the expression of HSPA5 following the induction of oxidative stress.[42] Schematically, DNA-bound ATF6α interacts with the AcID motif of Med25, thereby recruiting the Mediator complex to the

promoters of ATF6α dependent genes. Mediator then in turn recruits other required coactivator proteins and the RNAPII holocomplex, thereby upregulating expression of the target gene. Thus, blocking the critical interaction between the ATF6α TAD and Med25 through the use of small molecule inhibitors of AcID should result in a decrease in the expression of ATF6α target genes such as HSPA5, similar to observations following knockdown of Med25 expression. The results of this experiment are shown below in Figure 3.36.

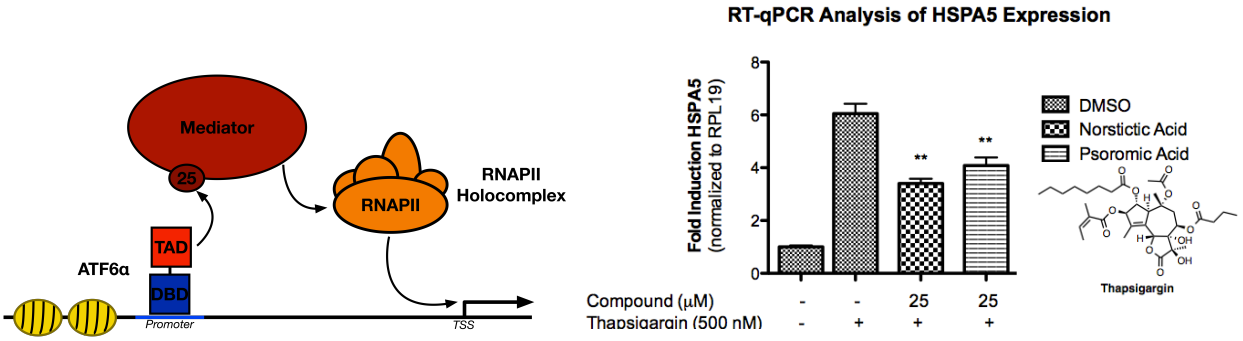


Figure 3.36- Inhibition of ATF6α Driven Genes by Norstictic and Psoromic Acid On the left is a schematic representation of Med25-dependent activation of ATF6α target genes. On the right is a graph showing the effects of Norstictic and psoromic acid on the expression of HSPA5 following induction of ER stress. Cells were incubated with DMSO (negative control) or 25 μM of norstictic acid or psoromic acid for 12 hours. Cells were then treated with thapsigargin at a final concentration of 500 nM to induce ER stress and incubated an additional three hours. Cells were then lysed and isolated mRNA was quantified for HSPA5 activation. All cells treated with control (DMSO) or compounds contained 0.5% v/v DMSO. All signals are the mean and standard deviation of 3 technical replicates. **, $P < 0.01$, $n = 3$. *(Completed by Paul A. Bruno)*

Treatment of cells with norstictic acid or psoromic acid for twelve hours prior to the induction of ER stress through treatment with thapsigargin resulted in inhibition of the expression of HSPA5, relative to untreated cells, as shown in Figure 3.36. This observation is consistent with the hypothesis that norstictic acid and psoromic acid may be capable of perturbing activator•AcID interactions within a cellular context.

Another transcriptional activator that interacts with Med25 in order to recruit the Mediator complex is retinoic acid receptor α (RARα), as demonstrated by Lee and colleagues.[58] RARα does not interact with Med25 at the AcID motif, but instead contacts the subunit at the nuclear receptor (NR) box located near its C-terminus, as demonstrated by a series of yeast two-hybrid assays and co-immunoprecipitation experiments. siRNA mediated knockdown of Med25 demonstrates a more significant loss of RARα transcriptional activity in reporter assays than knockdown of Med1, another well

characterized nuclear receptor binding subunit within the Mediator complex.[59] In turn, the master coactivator and histone acetyltransferase (HAT) CBP is recruited to the DNA-bound RARα•Mediator complex through an interaction between a sequence within the N-terminal domain of CBP and the AcID motif of Med25. The recruitment of CBP and its requisite HAT activity is a requirement for full transcriptional activity of RARα.[60-62] Thus, Med25 acts as an important coactivator target that allows for the simultaneous recruitment of the HAT CBP and the Mediator complex, which in turn recruits the transcriptional machinery, by the nuclear receptor RARα. The simultaneous overexpression of Med25 and CBP results in the significantly enhanced transcriptional activation of RARα, relative to the individual overexpression of CBP or Med25, suggesting that both coactivators are important for RARα function. Therefore, it was hypothesized that norstictic acid or psoromic acid may be capable of inhibiting the recruitment of CBP to the promoter of a RARα reporter plasmid by blocking its interaction with the AcID motif of Med25, resulting in a decrease in RARα transcriptional activity within the reporter assay. The results of this experiment are shown below in Figure 3.37.

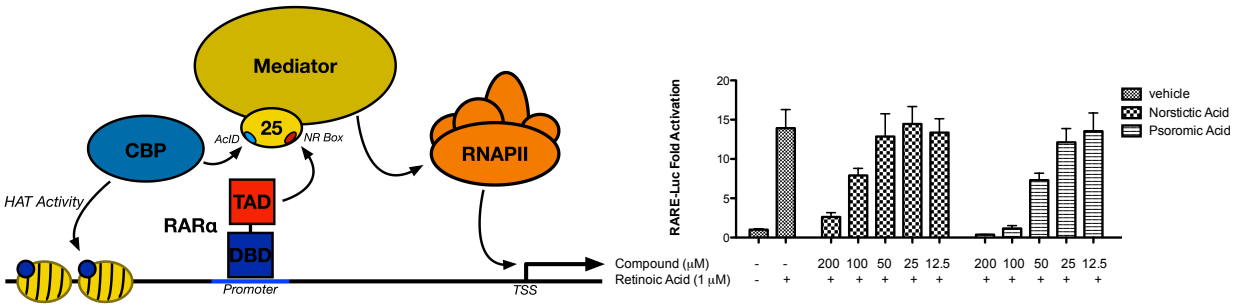


Figure 3.37- Inhibition of a RARα Transcriptional Reporter Assay Using Norstictic Acid and Psoromic Acid On the left is a schematic representation of the RARα reporter assay. RARα binds to a response element on the reporter plasmid following treatment with retinoic acid and recruits the Mediator complex through an interaction with the NR box of Med25. CBP is then recruited to the promoter through an interaction with the AcID motif and its HAT activity acetylates adjacent histones, resulting in upregulation of the reporter gene. On the right is a graph showing the effects of various concentrations of norstictic acid and psoromic acid on the expression of the luciferase reporter gene. Cells transfected with pRARE-luciferase reporter and pCMV-β-Gal were dosed with retinoic acid and either DMSO or compound as noted above. All DMSO levels were kept below 1% v/v for cellular dosing. All signals were normalized to β-Gal activity and represent the mean and standard deviation of 4 biological replicates. *(Completed by Paul A. Bruno)*

Both norstictic acid and psoromic acid demonstrate a dose dependent inhibition of the expression of a luciferase reporter gene by RARα, consistent with the hypothesis that the molecules may interfere with the recruitment of CBP through its interaction with AcID. These data demonstrating that norstictic acid and psoromic acid may be capable of inhibiting ATF6α and RARα transcriptional programs is encouraging and suggests that these small molecules warrant further investigation as potential mechanistic probes for Med25-dependent transcriptional processes.

3.4 Conclusions

At the outset of the work discussed in this chapter we sought to determine if we could identify small molecule inhibitors of the interaction between the VP16 transcriptional activation domain and one of its coactivator binding partners, Med25. Towards this end, a fluorescence polarization based assay using the VP16(465-490)•AcID interaction was developed and optimized for high-throughput screening. This assay was then used to screen a library containing 4,000 biologically active molecules and led to the identification of a number of lead molecules belonging to the depside and depsidone classes, supporting the hypothesis that activator•AcID interactions could be effectively targeted by small molecule inhibitors. Interestingly, our lab has previously identified the depside sekikaic acid and depsidone lobaric acid as inhibitors of another activator•coactivator interaction, suggesting that the depside and depsidone core scaffold may be privileged for the inhibition of activator•coactivator interactions.

Importantly, norstictic acid and the closely related depsidone psoromic acid appear to be selective for the AcID motif as both molecules failed to inhibit alternative activator•coactivator interactions. Specifically, these inhibitors failed to perturb the unrelated activator•coactivator interactions between MLL or CREB and the KIX domain of CBP/p300, suggesting that these molecules do not merely mimic transcriptional activation domains or generic amphipathic helices. More impressively, both molecules also failed to comparably inhibit the VP16•Med15 interaction, demonstrating that these inhibitors are not VP16 mimetics as they fail to inhibit other VP16•coactivator interactions. Thus, both norstictic acid and psoromic acid are actually more selective for the AcID motif

than one of its native binding partners, indicating that they may be particularly useful probes for studying AcID-dependent transcriptional processes.

The salt-concentration dependent activity of norstictic acid against the VP16•AcID interaction is consistent with the hypothesis that the molecules form electrostatic contacts with the domain, similar to activator binding partners as described in Chapter 2. Additionally, a reactive ortho-phenolic aldehyde functionality on these molecules is capable of forming covalent adducts with the domain following reaction with the ε-amine of reactive lysine residues within the protein. The formation of this covalent adduct relieves a positive charge on the surface of the domain, which may contribute to the inhibition of the VP16•AcID interaction as electrostatic contacts have been previously shown to be of particular importance for the interactions. Furthermore, chemical shift perturbation analysis of AcID in complex with norstictic acid reveals that the molecule induces shifts within the H1 and H2 binding sites, suggesting that the molecules are competing for surfaces critical for the VP16 interaction and further supporting the hypothesis that the molecules interfere with electrostatic contacts within these interaction surfaces in addition to sterically blocking activator binding.

Interestingly, the binding of the molecule also induces significant chemical shifts in regions of the domain that are not perturbed following activator binding, specifically at helix α2, which is located at the interface of the H1 and H2 binding sites. These data are consistent with a model in which norstictic acid is capable of inducting conformational shifts of the domain that interferes with the binding of activators to the motif. Mutational analysis in which specific lysine residues were mutated to arginine in order to abrogate the formation of the covalent adduct while maintaining critical electrostatic contacts identified specific lysine residues within the H1 and H2 binding surfaces as the sites for adduct formation. Mutation of lysine residues within the H2 binding site affected the ability of norstictic acid to inhibit interactions at the H2 interface, as expected, but also affected the ability of the molecule to inhibit interactions at the H1 binding site, further supporting the hypothesis that norstictic acid can function as an allosteric modulator in addition to orthosterically blocking activator binding surfaces. The converse was also demonstrated as well, where mutation of critical lysine residues within the H1 binding site perturbed the activity of the molecule at the H2 interface as well as the H1 surface. Mass spectrometric

analysis of AcID in complex with norstictic acid demonstrates that the major population of the sample contains a single norstictic acid modification. Taken together, these results are consistent with a mechanism of action in which norstictic acid functions as a mixed orthosteric/allosteric inhibitor of VP16•AcID interactions, capable of simultaneously perturbing binding events at both binding sites regardless of which site the molecule is bound to. Thus, this finding further supports the advantage of identifying inhibitors that function through allosteric modulation and target unique coactivator conformations.

Finally, norstictic acid and psoromic acid were tested for their ability to perturb AcID-dependent transcriptional processes within a cellular context. RT-qPCR analysis of the expression of HSPA5, a canonical target gene of the ER stress response transcription factor ATF6α, reveals that expression of the gene is inhibited by treatment with norstictic acid and psoromic acid prior to induction of ER stress. Similarly, norstictic acid and psoromic acid are capable of inhibiting the transcriptional activity of RARα in a reporter assay, likely by inhibiting the recruitment of the histone acetyl transferase CBP through its interaction with the AcID motif. These results are consistent with the hypothesis that norstictic acid and psoromic acid are cell permeable small molecules that can perturb AcID-dependent transcriptional processes. Thus, this preliminary evidence suggests that norstictic acid and psoromic acid warrant further analysis and validation as molecular probes that target the AcID motif in order to further elucidate the role of activator•AcID interactions in a variety of important cellular processes.

3.5 Materials and Methods

Plasmids

Plasmid pET21b-Med25(394-543)-His_6, henceforth referred to as pAcID-His_6, was a generous gift from Patrick Cramer.[36] AcID mutants were generated by site directed mutagenesis of the as previously described using the primers indicated below.[63] pAcID-(KK518RR)-His_6 was generated by site directed mutagenesis of pAcID-(K518R)-His_6. pAcID-(KKK518RRR)-His_6 was generated by site directed mutagenesis of pAcID-(KK518RR)-His_6. pAcID-(K411R/K413R/K518R/K519R/K520R)-His_6 was generated by site directed mutagenesis of pAcID-(KKK518RRR)-His_6 using the pAcID-(K411R/K413R)-His_6 primer set. pAcID-(K411R/K413R)-His_6 was prepared by Paul A. Bruno. pGL3-RARE-luc was purchased from Addgene and as previously described.[64] pCMV-β-Gal and pBSSK (non-coding plasmid) were kind gifts from Jorge Iñiguez-Lluhí.

pAcID-(K518R)-His6	F Pr.	TCATGCTCCTGTACTCGTCCAGGAAGAAGATCTTCATGGGCCTCATCCC
	R Pr.	GGGATGAGGCCCATGAAGATCTTCTTCCTGGACGAGTACAGGAGCATGA
pAcID-(K519R)-His6	F Pr.	TCATGCTCCTGTACTCGTCCAAGAGGAAGATCTTCATGGGCCTCATCCC
	R Pr.	GGGATGAGGCCCATGAAGATCTTCCTCTTGGACGAGTACAGGAGCATGA
pAcID-(K520R)-His6	F Pr.	TCATGCTCCTGTACTCGTCCAAGAAGAGGATCTTCATGGGCCTCATCCC
	R Pr.	GGGATGAGGCCCATGAAGATCCTCTTCTTGGACGAGTACAGGAGCATGA
pAcID-(KK518RR)-His6	F Pr.	TCATGCTCCTGTACTCGTCCAGGAGGAAGATCTTCATGGGCCTCATCCC
	R Pr.	GGGATGAGGCCCATGAAGATCTTCCTCCTGGACGAGTACAGGAGCATGA
pAcID-(KKK518RRR)-His6	F Pr.	TCATGCTCCTGTACTCGTCCAGGAGGAGGATCTTCATGGGCCTCATCCC
	R Pr.	GGGATGAGGCCCATGAAGATCCTCCTCCTGGACGAGTACAGGAGCATGA
pAcID-(K411R/K413R)-His6	F Pr.	GGGGTCCTGGAGTGGCAAGAGAGACCCAGACCTGCCTCAGTGGATGCCAAC
	R Pr.	GTTGGCATCCACTGAGGCAGGTCTGGGTCTCTCTTGCCACTCCAGGACCCC

Protein expression and purification

AcID ($Med25_{394-543}$) protein was expressed and purified together with Paul A. Bruno. Plasmid pAcID-His_6 was transformed into heat-shock competent Rosetta pLysS cells (Novagen), streaked onto LB Agar plates containing ampicillin and chloramphenicol, and incubated at 37 °C overnight. The next evening, a 25 mL Terrific Broth (TB) starter culture with 0.1 mg/mL ampicillin and 0.034 mg/mL chloramphenicol was then inoculated with a colony selected from the LB Agar plate and incubated at 37 °C overnight. The next morning, 5 mL from the starter culture was added to 1L TB containing ampicillin and

bacteria were grown at 37 °C to an OD_{600} of 0.8. Temperature was reduced to 18 °C and protein expression was induced upon addition of IPTG to a final concentration of 0.5 mM. Cells were incubated overnight at 18 °C. The 1 L cultures were then collected and centrifuged at 6000xg for 20 mins at 4 °C. Cell pellets were stored at -80 °C prior to purification. The harvested pellet was thawed on ice and resuspended in 20 mL of lysis buffer (50 mM phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 6.8). Cells were then lysed by sonication on ice and cellular lysates were cleared by centrifugation at 9500 rpm for 20 min at 4 °C. The supernatant lysate was then added to Ni-NTA beads (Qiagen) and incubated for 1 hour at 4 °C. The resin was pelleted by centrifugation at 2500 rpm for 2 min at 4 °C and washed with wash buffer (50 mM phosphate, 300 mM sodium chloride, 30 mM imidazole, pH 6.8) a total of five times. Protein was then eluted with 2 mL of elution buffer (50 mM phosphate, 300 mM sodium chloride, 400 mM imidazole, pH 6.8) a total of three times. Eluent was then pooled and purified by cation exchange FPLC (Source 15S, GE Healthcare) using a gradient of Buffer B (50 mM phosphate, 100 mM NaCl, 1 mM DTT, pH 6.8) in Buffer A (50 mM phosphate, 1 mM DTT). The FPLC purified protein was then dialyzed into storage buffer (10 mM phosphate, 50 mM NaCl, 10% v/v glycerol, 0.001% v/v NP-40, pH 6.8) overnight, concentrated, aliquoted, and stored at -20 °C. Final protein was greater than 90% pure as determined by coomassie stained polyacrylamide gel. Protein concentration was determined by UV-Vis spectroscopy using an extinction coefficient, ε = 22,460 $M^{-1}cm^{-1}$.

^{15}N-labeled AcID protein was expressed and purified together with Paul A. Bruno. Rosetta pLys cells transformed with pAcID-His_6 were used to inoculate a 25 mL LB starter culture with 0.1 mg/mL ampicillin and 0.034 mg/mL chloramphenicol. The starter culture was pelleted by centrifugation at 2500 RPM and washed twice with 20 mL of M9 minimal media. Following the second wash, cells were resuspended in 20 mL of M9 minimal media and two 500 mL cultures of M9 minimal mediated supplemented with ampicillin and 3 mL 10x BioXpress media (Cambridge Isotope Laboratories) were inoculated with 5 mL of the resuspended starter culture. Bacteria were grown at 37 °C to an OD_{600} of 0.8, at which point the temperature was reduced to 18 °C and expression was induced by addition of IPTG to a final concentration of 0.5 mM. Cells were incubated at 18 °C overnight and

pelleted at 6000 xg the following morning. Protein purification was as described for the unlabeled AcID protein.

Med25 AcID (K518R) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (K519R) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (K520R) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (KK518RR) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (KKK518RRR) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (K411R/K413R) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med25 AcID (K518R/K519R/K520R/K411R/K413R) was expressed and purified as described for the WT Med25 AcID protein. The presence of the desired mutation was confirmed by mass spectrometric analysis of the purified protein on an Agilent Q-ToF LC/MS.

Med15 (1-345) was expressed and purified by Paul A. Bruno as previously described.[65]

CBP/p300 KIX protein was expressed and purified together with Paul A. Bruno as previously described.[30]

Gal4 (1-100) was expressed and purified by Paul A. Bruno as previously described.[51]

Peptides

All peptides used in the studies described in this chapter have been previously reported in **Chapter 2**.

Fluorescence polarization direct binding assays

Direct binding assays were performed in triplicate with a final sample volume of 20 µL in a low volume, non-binding, 384-well black plate (Corning). FITC-labeled peptides were diluted in assay buffer (10 mM PBS, 100 mM NaCl, 10 % glycerol, 0.001% NP-40 pH 6.8) to a centration of 40 nM. 10 µL of AcID protein was serially diluted two-fold on the 384-well plate for the number of data points indicated for each experiment using assay buffer. 10 µL of the diluted fluorescent peptide stock was then added to each well of diluted protein for a final tracer concentration of 20 nM. An additional well containing 10 µL buffer and 10 µL fluorescent peptide was prepared for use as a 'tracer only control' to determine optimal gain settings on the plate reader. Samples were then incubated for 30 minutes at room temperature before fluorescence polarization was measured on a Pherastar plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm. A binding isotherm that accounts for ligand depletion (assuming a 1:1 binding model of peptide to ACID) was fit to the observed polarization values as a function of ACID to obtain the apparent equilibrium dissociation, K_d:

$$y = c + (b - c) \times \frac{(K_d + a + x) - \sqrt{(K_d + a + x)^2 - 4ax}}{2a}$$

Where “a” and “x” are the total concentrations of fluorescent peptide and AcID, respectively, “y” is the observed anisotropy at a given AcID concentration, “b” is the maximum observed anisotropy value, and “c” is the minimum observed anisotropy value. Each data point is an average of three independent experiments with the indicated error representing the standard deviation of the three replicates. All curves and calculations were generated using GraphPad Prism 5.

Primary screen of bioactive compounds

The primary screen was completed together with Paul A. Bruno at the Center for Chemical Genomics (University of Michigan) in collaboration with Martha Larsen and Steve Vander Roest. Assays were performed in a final volume of 20 µL in a low volume, non-binding, black 384- well plate (Corning) and read by plate reader (Pherastar) with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicularly polarized 535 nm filters. Final concentration of AcID protein was 850 nM, final concentration of Flo-VP16(465-490) was 20 nM, and compounds were assayed at a concentration of 20 µM with a final DMSO concentration of 1% v/v. 10 µL of AcID protein at a concentration of 1.7 µM was added to columns 1-22 of the assay plate by Multidrop dispenser (Thermo Fisher Scientific). Next, compounds were added to columns 3-22 of the assay plate and DMSO was added to columns 1-2 (negative control, AcID•VP16 complex) and 23-24 (positive control, VP16 tracer only) by pin tool. Finally, Flo-VP16(465-490) was added to all wells at a concentration of 40 nM. Plates were incubated for thirty minutes at room temperature and read by plate reader as described above with gain settings determined based on a well from columns 23-24 (tracer only). Data was published to and analyzed using MScreen (http://mscreen.lsi.umich.edu).

Substructure search of CCG libraries for compounds containing Depside/Depsidone core

All CCG compound libraries were searched for the depside and depsidone core scaffold described in panel A of Figure 3.12 above using cross-reference tools included in the MScreen software package. The substructure search was completed together with Paul A. Bruno.

Fluorescence polarization dose-response inhibition assays

Inhibition assays were performed in triplicate with a final sample volume of 20 µL in a low volume, non-binding, 384-well black plate (Corning). A complex of the indicated fluorescent tracer and protein was prepared at a 2x concentration such that 50% of the tracer was bound following dilution onto the assay plate. Small molecule inhibitors were diluted in assay buffer (10 mM PBS, 100 mM NaCl, 10 % glycerol, 0.001% NP-40 pH 6.8) to the desired concentration and serially diluted two-fold on the assay plate to a final volume of 10 µL. 10 µL of the pre-formed fluorescent tracer-protein complex was then added to each well for a final volume of 20 µL. An additional well containing tracer only was prepared and used to determine optimal gain settings on the plate reader. Samples were incubated for thirty minutes at room temperature before fluorescence polarization was measured on a Pherastar plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm. Polarization values were converted to relative fraction bound and plotted opposite the log of inhibitor concentration using GraphPad Prism 5 and curves were fit with a non-linear regression using the built-in equation “log(inhibitor) vs response – variable slope” from which the IC_{50} value was calculated.

Salt titration effects on Norstictic Acid inhibition

The fluorescence polarization inhibition experiments were performed as noted above with one modification. The NaCl concentration of assay buffer was adjusted to the indicated concentrations by adding an appropriate volume of assay buffer containing 5 M NaCl prior to preparing assay components.

Mass spectrometric analysis of covalent adducts reduced with $NaBH_4$

AcID protein was diluted to a concentration of 20 µM using storage buffer (10 mM phosphate, 50 mM NaCl, 10% v/v glycerol, 0.001% v/v NP-40, pH 6.8). Norstictic acid was added to the diluted protein to a final concentration of 80 µM (4 equivalents). Samples were incubated for two hours at room temperature with gentle mixing on an orbital shaker. Samples were subsequently reduced by addition of a 100 mM freshly prepared stock of $NaBH_4$ in storage buffer to a final concentration of 1 mM $NaBH_4$. Samples were incubated

with the reducing agent at room temperature for one hour with gentle mixing on an orbital shaker. Samples were subsequently diluted five-fold and buffer exchanged into fresh storage buffer. Samples were the analyzed by mass spectrometry using an Agilent Q-ToF LC/MS equipped with a Poroshell 300SB C_8 reverse-phased column using a gradient of 5-100% acetonitrile with 0.1% formic acid in water with 0.1% formic acid over five minutes. Analysis of data was completed using the Agilent Qualitative Analysis Program with background subtraction and deconvolution settings for an intact protein of 10,000-30,000 Da.

Mass spectrometric analysis of non-reduced covalent adducts

Mass spectrometric analysis of AcID•norstictic acid covalent adducts without sodium borohydride reduction were completed as described above without the reduction step.

Norstictic Acid inhibition time course

The norstictic acid inhibition time course was completed as described above with several modifications. Small molecule inhibitor was serially diluted onto the assay plate and the preformed AcID•Flo-VP16(465-490) was added rapidly to the wells using a multi-dropping multi-channel pipette. Samples were incubated five-minutes before being read by plate reader as described. Additional measurements were made of the same plate at the indicated time points.

^{1}H,^{15}N-HSQC NMR chemical shift perturbation experiments of covalent adducts

Sample preparation was completed by Paul A. Bruno. ^{15}N-labeled AcID protein was expressed and purified as described above. Protein was diluted to a concentration of 30 μM and separate samples containing DMSO (negative control), 90 μM norstictic acid, or 150 μM norstictic acid were prepared. Samples were incubated for two hours at room temperature and analyzed by mass spectrometry as described above. Following confirmation, samples were immediately submitted for ^{1}H,^{15}N-HSQC analysis by Felicia Gray (Cierpicki Lab, University of Michigan). Peak assignment and chemical shift perturbation assessment was completed by Felicia Gray.

Covalent adduct digestion and proteomic analysis

A 500 µL sample of AcID protein in complex with 4 equivalents of norstictic acid was reduced with sodium borohydride as described above. Labeling was verified by mass spectrometric analysis. Samples were then purified by SDS-polyacrylamide gel electrophoresis using a 4-12% bis-tris gel (Roche). Gels were stained with coomassie blue and bands of interest were excised and submitted to the Peptide Synthesis and Proteomics Core (University of Michigan) for in-gel chymotrypsin digestion and subsequent mass spectrometric analysis.

RT-qPCR analysis of HSPA5 gene expression

RT-qPCR analysis of endogenous HSPA5 gene expression levels was completed by Paul A. Bruno. For endogenous gene expression analysis, $1x10^5$ HeLa cells were seeded into a 24:well plate and allowed to adhere overnight. Media was removed and replaced with OptiMem media containing vehicle or compound delivered in DMSO (0.5% v/v) at the indicated concentrations. After incubating for 12 h, cells were treated with thapsigargin at a final concentration of 500 nM. After 3 h, the media was removed and total RNA was isolated using RNeasy Plus RNA isolation kits (Qiagen) according to manufacturer's instructions. Each RNA sample was used to synthesize cDNA using iScript cDNA synthesis kits (Bio-Rad). RT-qPCR reactions were carried out in triplicate in an Applied Biosystems StepPlusOne instrument using SYBR green master mix and primers for: human RPL19 F pr. 5':ATGTATCACAGCCTGTACCTG:3'; R Pr., 5':TTCTT GGTCTCTCTTCCTCCTTG:3') and HSPA5 (F Pr., 5':CTGGGTACATTTGATCTGACTG G:3'; R Pr., 5': CTTACCGACCTTTCGGTGGTCCTACG:3'). RT:qPCR analysis was carried out using the comparative C_T Method ($\Delta\Delta C_T$ Method) to estimate HSPA5 mRNA levels relative to the reference RPL19 mRNA levels.

RARα luciferase reporter assay

The RARα luciferase reporter assay experiments were completed by Paul A. Bruno. The RARα luciferase reporter containing 3 tandem RARα sites (RARα-luc) was obtained from Addgene. CMV-β-Gal, and pBSSK were generously provided by Dr. Jorge Iñigues-Lluhí. All cells were maintained in 5% CO_2 at 37°C. HeLa cells were grown in Dulbecco's

modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% FBS. For luciferase assays, $4x10^5$ cells were seeded in a 6:well dish and allowed to adhere overnight. The media was removed and cells were transfected in OptiMem (Invitrogen) with 1 μg RARα:luc, 200 ng CMV:β:Gal, and 800 ng pBSSK using Lipofectamine 2000 (Life Technologies) according to manufacturer's instructions. After 4.5 h, transfection solution was removed and replaced with DMEM containing 10% FBS. At 24 h after transfection, cells were trypsinized and resuspended in DMEM supplemented with 10% FBS and seeded into a 96:well plate at a density of $8x10^3$ cells per well. After an additional 16 h, media was removed and replaced with OptiMem containing vehicle or small molecule at the indicated concentration as a solution in DMSO co-dosed with retinoic acid (1 μM). Cells were incubated with either vehicle or compound and retinoic acid for 16 h, media was removed and cells were lysed with 60 μL of passive lysis buffer. Luciferase and β:Galactosidase activities were determined as previously described. RARα luciferase activity was plotted using GraphPad Prism 5.

3.6 References

1. Gann, M. P. A. A. Genes and Signals. 1–209 (2005).

2. Giordano, A. & Avantaggiati, M. L. p300 and CBP: partners for life and death. *J. Cell. Physiol.* **181,** 218–230 (1999).

3. Davidson, B. *et al.* PEA3 is the second Ets family transcription factor involved in tumor progression in ovarian carcinoma. *Clin. Cancer Res.* **9,** 1412–1419 (2003).

4. Heinlein, C. A. & Chang, C. Androgen receptor in prostate cancer. *Endocr. Rev.* **25,** 276–308 (2004).

5. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **11,** 891–894 (2015).

6. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chemical Biology* **7,** 1311–1320 (2012).

7. Cesa, L. C., Mapp, A. K. & Gestwicki, J. E. Direct and Propagated Effects of Small Molecules on Protein-Protein Interaction Networks. *Front Bioeng Biotechnol* **3,** 119 (2015).

8. Ng, H. H. & Bird, A. Histone deacetylases: silencers for hire. *Trends Biochem. Sci.* **25,** 121–126 (2000).

9. Peterson, C. L. & Laniel, M.-A. Histones and histone modifications. *Curr. Biol.* **14,** R546–51 (2004).

10. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30,** 235–239 (2005).

11. Conaway, R. C. & Conaway, J. W. Function and regulation of the Mediator complex. *Curr. Opin. Genet. Dev.* **21,** 225–230 (2011).

12. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14,** 1553–1577 (2000).

13. Tosovská, P. & Arora, P. S. Oligooxopiperazines as nonpeptidic alpha-helix mimetics. *Org. Lett.* **12,** 1588–1591 (2010).

14. Lee, L. W. *et al.* Inhibition of ErbB2(Her2) expression with small molecule transcription factor mimics. *Bioorganic & Medicinal Chemistry Letters* **19,** 6233–6236 (2009).

15. Yin, H. *et al.* Terphenyl-based helical mimetics that disrupt the p53/HDM2 interaction. *Angew. Chem. Int. Ed. Engl.* **44,** 2704–2707 (2005).

16. Kushal, S. *et al.* Protein domain mimetics as in vivo modulators of hypoxia-inducible factor signaling. *Proceedings of the National Academy of Sciences* **110,** 15602–15607 (2013).

17. Minter, A. R., Brennan, B. B. & Mapp, A. K. A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* **126,** 10504–10505 (2004).

18. Buhrlage, S. J. *et al.* Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chemical Biology* **4,** 335–344 (2009).

19. Kristensen, V. N. *et al.* Gene expression profiling of breast cancer in relation to estrogen receptor status and estrogen-metabolizing enzymes: clinical implications. *Clin. Cancer Res.* **11,** 878s–83s (2005).

20. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747–752 (2000).

21. Pupa, S. M., Tagliabue, E., Ménard, S. & Anichini, A. HER-2: a biomarker at the crossroads of breast cancer immunotherapy and molecular medicine. *J. Cell. Physiol.* **205,** 10–18 (2005).

22. Asada, S. *et al.* External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc Natl Acad Sci USA* **99,** 12747–12752 (2002).

23. Benz, C. C. *et al.* HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene* **15,** 1513–1525 (1997).

24. Shimogawa, H. *et al.* A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction. *J. Am. Chem. Soc.* **126,** 3461–3471 (2004).

25. Zhang, M. *et al.* Genetic and chemical targeting of epithelial-restricted with serine box reduces EGF receptor and potentiates the efficacy of afatinib. *Mol. Cancer Ther.* **12,** 1515–1525 (2013).

26. Lao, B. B. *et al.* In vivo modulation of hypoxia-inducible signaling by topographical helix mimetics. *Proceedings of the National Academy of Sciences* **111,** 7531–7536 (2014).

27. Hirota, K. & Semenza, G. L. Regulation of angiogenesis by hypoxia-inducible factor 1. *Crit. Rev. Oncol. Hematol.* **59,** 15–26 (2006).

28. Freedman, S. J. *et al.* Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc Natl Acad Sci USA* **99,** 5367–5372 (2002).

29. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci USA* **99,** 5271–5276 (2002).

30. Wang, N. *et al.* Ordering a dynamic protein via a small-molecule stabilizer. *J. Am. Chem. Soc.* **135,** 3363–3366 (2013).

31. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *Journal of Biological Chemistry* **277,** 43168–43174 (2002).

32. De Guzman, R. N., Goto, N. K., Dyson, H. J. & Wright, P. E. Structural basis for cooperative transcription factor binding to the CBP coactivator. *J. Mol. Biol.* **355,** 1005–1013 (2006).

33. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. *Proceedings of the National Academy of Sciences* **111,** 12061–12066 (2014).

34. Majmudar, C. Y. *et al.* Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300. *Angew. Chem. Int. Ed.* **51,** 11258–11262 (2012).

35. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **18,** 410–415 (2011).

36. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18,** 404–409 (2011).

37. Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *Journal of Structural Biology* **174,** 245–251 (2011).

38. Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* **20,** 1455–1464 (2001).

39. Verger, A. *et al.* The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41,** 4847–4859 (2013).

40. Kinoshita, J. *et al.* Clinical significance of PEA3 in human breast cancer. *Surgery* **131,** S222–5 (2002).

41. Aytes, A. *et al.* ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer. *Proceedings of the National Academy of Sciences* **110,** E3506–15 (2013).

42. Sela, D. *et al.* Role for human mediator subunit MED25 in recruitment of mediator to promoters by endoplasmic reticulum stress-responsive transcription factor ATF6α. *J. Biol. Chem.* **288,** 26179–26187 (2013).

43. Wysocka, J. & Herr, W. The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. *Trends Biochem. Sci.* **28,** 294–304 (2003).

44. Yang, F., DeBeaumont, R., Zhou, S. & Näär, A. M. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc Natl Acad Sci USA* **101,** 2339–2344 (2004).

45. Johnson, P. F., Sterneck, E. & Williams, S. C. Activation domains of transcriptional regulatory proteins. *The Journal of Nutritional Biochemistry* **4,** 386–398 (1993).

46. Jonker, H. R. A., Wechselberger, R. W., Boelens, R., Folkers, G. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44,** 827–839 (2005).

47. McGovern, S. L., Helfand, B. T., Feng, B. & Shoichet, B. K. A specific mechanism of nonspecific inhibition. *J. Med. Chem* **46,** 4265–4272 (2003).

48. Zhang, J., Chung, T. & Oldenburg, K. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4,** 67–73 (1999).

49. Fischer, E. SYNTHESIS OF DEPSIDES, LICHEN-SUBSTANCES AND TANNINS. *J. Am. Chem. Soc.* **36,** 1170–1201 (1914).

50. Parker, D. *et al.* Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular cell* **2,** 353–359 (1998).

51. Wands, A. M. *et al.* Transient-state kinetic analysis of transcriptional activator·DNA complexes interacting with a key coactivator. *Journal of Biological Chemistry* **286,** 16238–16245 (2011).

52. Majmudar, C. Y., Wang, B., Lum, J. K., Håkansson, K. & Mapp, A. K. A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew. Chem. Int. Ed.* **48,** 7021–7024 (2009).

53. McMillen, D. A. *et al.* Identifying regions of membrane proteins in contact with phospholipid head groups: covalent attachment of a new class of aldehyde lipid labels to cytochrome c oxidase. *Biochemistry* **25,** 182–193 (1986).

54. Sakaguchi, K., Miyakawa, T. & Takeuchi, S. Interaction of benzaldehyde to the membrane protein of Escherichia coli. *Agricultural and …* (1979). doi:10.1080/00021369.1979.10863704

55. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2,** 718–729 (1988).

56. Parrot, D., Jan, S., Baert, N., Guyot, S. & Tomasi, S. Comparative metabolite profiling and chemical study of Ramalina siliquosa complex using LC-ESI-MS/MS approach. *Phytochemistry* **89,** 114–124 (2013).

57. Thuerauf, D. J., Morrison, L. E., Hoover, H. & Glembotski, C. C. Coordination of ATF6-mediated transcription and ATF6 degradation by a domain that is shared with the viral transcription factor, VP16. *Journal of Biological Chemistry* **277,** 20734–20739 (2002).

58. Lee, H.-K. H., Park, U.-H. U., Kim, E.-J. E. & Um, S.-J. S. MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* **26,** 3545–3557 (2007).

59. Poss, Z. C., Ebmeier, C. C. & Taatjes, D. J. The Mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **48,** 575–608 (2013).

60. Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85,** 403–414 (1996).

61. Torchia, J. *et al.* The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387,** 677–684 (1997).

62. Kawasaki, H. *et al.* Distinct roles of the co-activators p300 and CBP in retinoic-acid-induced F9-cell differentiation. *Nature* **393,** 284–289 (1998).

63. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chemical Biology* **7,** 1345–1350 (2012).

64. Hoffman, L. M. *et al.* BMP action in skeletogenesis involves attenuation of retinoid signaling. *J. Cell Biol.* **174,** 101–113 (2006).

65. Majmudar, C. Y. *et al.* Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300. **51,** 11258–11262 (2012).

CHAPTER 4

Natural Product Extract Screening to Identify Inhibitors of the PEA3 Transcriptional Activator Subfamily[4]

4.1 Abstract

The PEA3 subfamily of ETS transcriptional activators, comprised of three distinct proteins with highly similar transcriptional activation domains, have been reported as activator binding partners of the Activator Interaction Domain of Mediator subunit 25. Furthermore, this activator•coactivator interaction has been shown to be necessary for the full transcriptional activity of the subfamily. Transcriptional programs mediated by these activators have been linked to a variety of malignant cellular processes, including tumorigenesis and metastasis and as a result represent attractive targets for the development of small molecule inhibitors. Given the previously demonstrated success in targeting protein-protein interactions between transcriptional activators and the AcID motif of Med25 using small molecule natural products in Chapter 3, it was next sought to identify novel natural products capable of disrupting the interaction between PEA3 subfamily members and AcID. Inhibitors that are not reliant on the formation of covalent adducts with the domain for their activity, unlike norstictic acid and psoromic acid, are particularly attractive as it is hypothesized that non-covalent inhibitors may have enhanced selectivity for AcID in a complex cellular context and may offer an alternative mechanism of action for the inhibition of AcID-dependent transcriptional processes.

[4] The work completed in this chapter was a collaborative effort. The primary screen, Gal4•DBD counter screen, and MLL•KIX counter screen were completed by Steven M. Sturlis and Paul A. Bruno. Extract triage and final hit selection was completed in collaboration with Martha Larsen and Steve Vander Roest of the Center for Chemical Genomics at the University of Michigan. The selection of strains for regrowth was completed by Prof. David H. Sherman, Pamela Schultz, and Matthew S. Beyersdorf. Strains were regrown, fresh extracts prepared, confirmation of activity against the ERM•AcID assay, and initial fractionation were completed by Matthew S. Beyersdorf. The RARα reporter assay was performed by Paul A. Bruno.

Towards this goal, a primary screen of the full Natural Product Extract Library held by the Center for Chemical Genomics at the University of Michigan using an optimized ERM•AcID fluorescence polarization assay was completed to identify extracts with inhibitory activity against this interaction. Ultimately, 332 extracts with potent activity against the ERM•AcID interaction were identified, with thirty-one selected for further analysis. Fractionation and structural elucidation of the most active extract is currently ongoing and will yield potent inhibitors of the ERM•AcID interaction.

4.2 Introduction

The PEA3 Subfamily of ETS Transcriptional Activators Has Been Implicated in the Progression of Cancer

The ETS superfamily of transcriptional activators contains over twenty-five unique proteins that are central drivers of a multitude of cellular processes including cell cycle regulation and proliferation, tissue remodeling during development, and cellular differentiation. [1,2] Given their central importance in cellular processes, the dysregulation of members of this superfamily have been implicated in a variety of cancers.[3,4] One subfamily of ETS transcription factors of particular interest is the PEA3 subfamily, composed of three transcriptional activators ETV1/ER81, ETV4/PEA3, and ETV5/ERM, which have been reported as important in the progression of a variety of cancers to a metastatic phenotype. [5-9] As metastasis is the ultimate cause of mortality in more than 90% of cancers, the disruption of this process would be an attractive therapeutic strategy as it could keep cancers from spreading beyond the primary tumor, allowing for more effective eradication of the disease by surgical, chemotherapeutic, or radiological means.[10]

For example, in breast cancer the overexpression of one or multiple members of the PEA3 subfamily results in an invasive phenotype and subsequent knockdown of the overexpressed PEA3 activators can abrogate this phenotype. [9,11] The activators within the PEA3 subfamily govern the expression of genes that are important for cellular invasion and migration, such as the matrix metalloproteinases (MMP). [12,13] Studies have demonstrated that in triple negative breast cancer shRNA mediated knockdown of these activators results in reduced expression of genes such as the MMPs, resulting in a less

invasive phenotype. Importantly, the most effective inhibition of cellular invasion required that all three activators be knocked down, suggesting that there may be significant functional overlap between the members of this subfamily in driving metastatic processes.[14]

Based on this observation, small molecule inhibitors that are capable of disrupting the transcriptional activity of all three activators simultaneously will be more effective than designing inhibitors that target the activators individually. One manner by which this could be achieved would be by targeting a common activator•coactivator interaction required by the three members of the subfamily for transcriptional activity. The N-terminal transactivation domains of all three proteins are highly conserved, as shown in Figure 4.2, suggesting that the coactivators that are recruited during the course of transcriptional activation are also conserved.[12,13] Consistent with this model, recent studies reveal that the three PEA3 activators require the recruitment of the Mediator complex through an interaction with the Activator Interaction Domain of Mediator subunit Med25 for full transcriptional activity.[13,15] The common requirement of this interaction for all three activators suggests that this is an attractive point of intervention for the development of small molecule inhibitors. Such molecules would be useful as mechanistic probes in the further elucidation of the role of the subfamily in the progression of cancer to a metastatic phenotype, as well as other cancer progression pathways such as tumorigenesis, and may also serve as potential lead compounds in the development of novel therapeutic strategies.

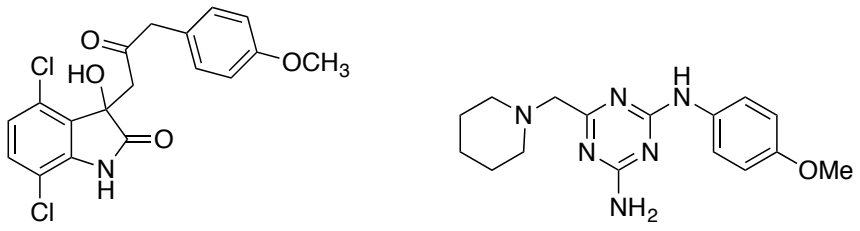
Previous Success in the Disruption of PEA3 Mediated Transcription by Small Molecule Inhibitors

Despite the demonstrated role of the PEA3 subfamily in the disease progression of cancers including breast cancer, prostate cancer, and pancreatic cancer there have been few successful attempts at inhibiting the activity of these activators using small molecule inhibitors. [5,16-19] Specifically, there are no published reports of inhibitors of ETV4 nor ETV5 mediated transcriptional processes and only two reports of small molecules that directly inhibit the activity of ETV1, via the targeting of the DNA binding moiety.

The first of these inhibitors, known as YK-4-279, was reported in 2009 as an inhibitor of the fusion protein EWS-FLI1 in Ewing's Sarcoma.[20] FLI1, ETV1, and ERG are

all ETS transcription factors that share 60% identity and 80% homology in their primary structures, suggesting that YK-4-279 may also function as an inhibitor of ETV1 and ERG.[21] The compound was tested within the prostate cancer derived cell lines LNCaP and VCaP, which are FLI1 negative but ERG and ETV1 positive, and shown to be capable of inhibiting the expression of ETV1 and ERG dependent genes to levels comparable to siRNA knockdowns of these activators.[22] The mechanism by which this inhibition is achieved has not yet been elucidated, but surface plasmon resonance (SPR) studies demonstrate that YK-4-279 interacts with ETV1 and ERG directly, though it is unclear where on the protein this interaction occurs. As a consequence of the inhibition of ETV1 and ERG dependent transcription, treated LNCaP and VCaP cells demonstrated reduced invasive potential and cellular motility, underscoring the role of ETV1 and ERG in cellular processes that result in a metastatic phenotype. More recent reports have demonstrated that this inhibitor is also capable of inhibiting metastatic phenotypes *in vivo* in mouse xenograft models of prostate cancer.[23] The differential effects of YK-4-279 against ETV1 and ERG are currently unclear, which, coupled with the currently uncharacterized mechanism of action, suggests that this inhibitor may not be particularly useful as a mechanistic probe for PEA3 dependent transcriptional processes.

Reported ETV1 Small Molecule Inhibitors



YK-4-279 *BRD32048*

***Figure 4.1*- Reported Small Molecule Inhibitors of ETV1** YK-4-279 and BRD32048 are the only reported inhibitors of one of the PEA3 subfamily members.

The second reported ETV1 inhibitor is BRD32048, which was first reported in 2014 by Pop and colleagues.[24] This inhibitor was identified using a small molecule microarray to screen 45,000 small molecules against ETV1 from cellular lysates. Subsequent SPR and pull-down experiments using modified BRD32048 demonstrated that the inhibitor interacts with ETV1 directly. Genome wide microarray profiling of ETV1 regulated genes

demonstrated that BRD32048 was capable of modulating a subset of ETV1 dependent genes compared to shRNA mediated knockdown of ETV1 gene expression. As a functional consequence of this transcriptional modulation, BRD32048 was shown to inhibit cellular invasion in ETV1-dependent prostate cancer model cell lines. Initial studies to determine the mechanism by which inhibition of ETV1 is achieved suggest that the small molecule blocks acetylation of critical lysine residues responsible for maintaining the stability of the protein. As a result of the inhibition of this acetylation event, the protein is destabilized and subsequently degraded. Thus, the molecule achieves its inhibitory activity by decreasing the half-life of ETV1 within the cell.

While these molecules are relatively potent inhibitors of ETV1 capable of downregulating the expression of ETV1 dependent genes and inhibiting cellular invasion in a number of prostate cancer cell lines, they are specifically targeted to ETV1 and thus do not inhibit the other member of the PEA3 subfamily. Given the potential for functional overlap between the three members of the subfamily and their differential expression in a variety of cancers, the inhibition of a common activator•coactivator interaction required by all three should result in more pronounced inhibition of cellular invasion across a number of cancer phenotypes.

PEA3 Transcriptional Activators Interact with AcID

A recent report in the literature from the laboratory of Alexis Verger demonstrated that the PEA3 subfamily interacts with the AcID motif of Med25 through its N-terminal TAD (residues 1-72).[13] Notably, the N-terminal TADs of the three PEA3 subfamily members are highly conserved, as shown in Figure 4.2, suggesting that they may bind to the same coactivator targets. Furthermore, this highly conserved TAD sequence is homologous to the VP16 TAD, particularly the C-terminal elements of the VP16 H1 subdomain. Secondary structure prediction suggested that the N-terminal TAD of ETV5/ERM is capable of adopting an α-helical secondary structure, specifically between residues 42 and 57 of the TAD, which was later supported by NMR solution structures of the TAD in complex with AcID.[15] Thus, the PEA3 subfamily was hypothesized to bind to the AcID motif in a similar fashion as VP16 given the common structural features between the two TADs.

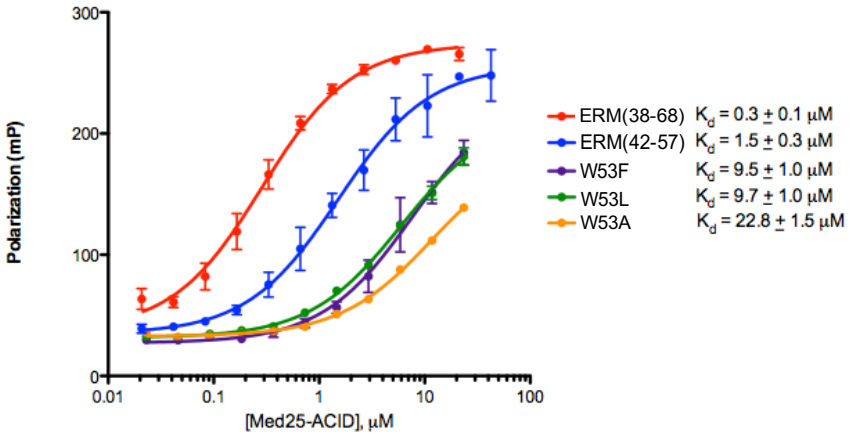
***Figure 4.2*- Structural Features of the N-terminal TAD of the PEA3 Subfamily** (A) The three activators that comprise the subfamily have N-terminal TADs that are highly conserved. (B) The N-terminal TAD of the PEA3 subfamily is homologous to the TAD of VP16, specifically the C-terminal portion of the VP16 H1 subdomain as shown in this sequence alignment. Identical residues are highlighted in gray and conserved residues are indicated by a black box. (C) Secondary structure prediction of the conserved PEA3 TAD suggests it is capable of adopting an α-helical secondary structure.

In order to demonstrate that this interaction actually occurs within the cell, Verger and colleagues first performed a GST pull-down assay and demonstrated that the N-terminal TAD of ERM was capable of precipitating Med25 from cellular lysates and the GST-tagged AcID was capable of precipitating ERM from cellular lysates. Subsequent ITC experiments using ERM(38-68) and purified AcID protein demonstrated that the two interact with an apparent K_d value of approximately 540 nM. Additional precipitation experiments using Halo-tagged ERM(1-72) demonstrated that this interaction did not disrupt the interaction between Med25 and the rest of the Mediator complex, indicating that this interaction would result in the recruitment of the full Mediator complex to ERM target genes. In support of this conclusion, reporter assays using a Gal4(DBD)-ERM(1-72) fusion protein demonstrated that the interaction between the ERM TAD and Med25 was required for full transcriptional activity. The repressive effect of overexpression of AcID on the activity in the reporter assay demonstrates that the TAD interacts with AcID, while the repressive effect of overexpression of the VWA domain of Med25 demonstrates that the recruitment of the Mediator complex through the AcID interaction is required for full activity. This hypothesis was further supported by the observation that the siRNA mediated knockdown of Med25 and overexpression of a ΔAcID Med25 mutant did not recover transcriptional activity. Finally, this interaction was shown to relevant in the context of endogenous gene expression as the siRNA mediated knockdown of Med25 resulted in decreased expression of the ETV4 target gene MMP-1 that could not be recovered by overexpression of ETV4 as determined by RT-qPCR. The same level of

decreased expression required that all three PEA3 subfamily members be knocked down, suggesting that there may be significant functional overlap between the subfamily members.

Mutational experiments of critical residues within the VP16 H1 and H2 binding sites suggest that the PEA3 TADs bind to a surface within the H1 binding site, which was corroborated later by $^{1}H,^{15}N$-HSQC chemical shift perturbation studies of the domain in complex with the TAD.[15] The degree of overlap between the binding site for VP16 H1 and the PEA3 TAD is still unclear and it is possible that the two TADs bind to similar but distinct surfaces of the domain.

Building from the results first published by Verger and colleagues, we developed a fluorescently labeled peptide of ERM(38-68) for use in fluorescence polarization assays and found that it bound to AcID with comparable affinity to the values reported by Verger, as shown in Figure 4.3.[13] Truncation of this peptide to include only the purported α-helical sequence resulted in an approximately five-fold decrease in affinity, suggesting that the α-helix likely contributes significantly to the interaction with the AcID motif, consistent with previous observations of the interaction between the VP16 TAD and AcID. Furthermore, the N-terminal TADs of PEA3 family members contain a tryptophan residue (W53) that was suspected to function as a hot-spot within the interaction as tryptophan residues account for 21% of all characterized hot-spots.[25] Mutation of this tryptophan to bulky hydrophobic amino acids such as phenylalanine or leucine, relatively conservative mutations, resulted in an approximately six fold decrease in affinity. Mutation of the tryptophan to alanine, a less conservative mutation, resulted in greater than fifteen-fold decrease in affinity, underscoring the importance of this residue in the interaction with AcID. With a $\Delta\Delta G$ of 1.6 kCal/mol, the W53A mutation does not qualify this tryptophan residue as a hotspot, but nevertheless suggests that the ERM•AcID interaction may utilize more focused contacts than the previously discussed VP16•AcID interaction.



ERM(38-68)	38-	DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ	-68
ERM(42-57):	42-	DSEELFQDLSQLQEAWLAEA	-57
ERM(42-57) W53F:	42-	DSEELFQDLSQLQEAFLAEA	-57
ERM(42-57) W53L:	42-	DSEELFQDLSQLQEALLAEA	-57
ERM(42-57) W53A:	42-	DSEELFQDLSQLQEAALAEA	-57

***Figure 4.3*- Truncation and Mutation of the ERM TAD** The VP16 TAD was truncated to produce a peptide containing the putative α-helix alone, which bound with only moderately attenuated affinity. An important tryptophan residue was mutated to generate three distinct peptides and their affinity for the domain was assessed. (*Completed by Nicholas Foster*)

Thus, given the demonstrated functional importance of the PEA3•AcID interactions for the full transcriptional activity of this subfamily of activators and the similarity in molecular features of the interactions to those of VP16•AcID, we hypothesized that this interface represents an attractive point of intervention for small molecule inhibitors. We further hypothesize that the inhibition of this interaction will result in a decrease in the transcriptional activation of PEA3 target genes, as demonstrated for siRNA mediated knockdown of Med25, and a concomitant decrease in metastatic potential in a variety of cancer-derived cell lines. Furthermore, the inhibition of this interaction using small molecule inhibitors will also allow for greater perspective of the role of these activators in other cancer-related cellular processes including tumorigenesis.

Natural Products as Small Molecule Inhibitors

In order to identify inhibitors of the PEA3•AcID interaction, we performed a high-throughput screen of the Natural Product Extract Library held by the Center for Chemical Genomics at the University of Michigan against a fluorescence polarization based assay

of the interaction between the TAD of the PEA3 subfamily member ERM and the AcID motif. This library contains natural product extracts (NPE) from remarkably diverse organisms including terrestrial and marine microorganisms, plants, and lichens amongst other organisms of interest and has been successfully screened to identify natural product inhibitors against a variety of targets. [26-29] The genetic diversity of the organisms from which these NPEs have been collected indicates that the library likely contains small molecules with structural diversity that is unmatched by more traditional small molecule screening libraries which generally contain flat, aromatic, and relatively simple chemical scaffolds. Thus, the structural diversity contained within the NPE library dramatically increases the opportunity to find potent inhibitors of activator•coactivator interactions as the more complex and three-dimensional structures typical of natural products are more likely to match the complex topography common to most protein-protein interaction interfaces. [30,31] The identification of the depsidones norstictic acid and psoromic acid as inhibitors of AcID-dependent activator•coactivator interactions in Chapter 3 supports this hypothesis and the decision to screen the NPE library.

The NPE library was screened using an iterative screening strategy that our group has previously utilized to identify selective allosteric inhibitors of KIX-dependent activator•coactivator interactions.[28] A generalized version of this protocol is shown below in Figure 4.4. Following primary screening of the NPE library, false positives and extracts with reactive or overly fluorescent components are removed from follow up study using standard screening techniques. Subsequently, active extracts are tested for inhibition against a series of functionally related protein-protein interactions, such as the interaction between activators and the KIX or CH3 domains of CBP/p300 or activator interactions with TBP. These selectivity filters remove extracts that are not selective for the coactivator target of interest and eliminates inhibitors that merely function as TAD or general amphipathic α-helix mimetics. By selecting only extracts that are selective for the coactivator of interest, the preliminary hits are enriched for extracts that likely contain allosteric inhibitors of the target.

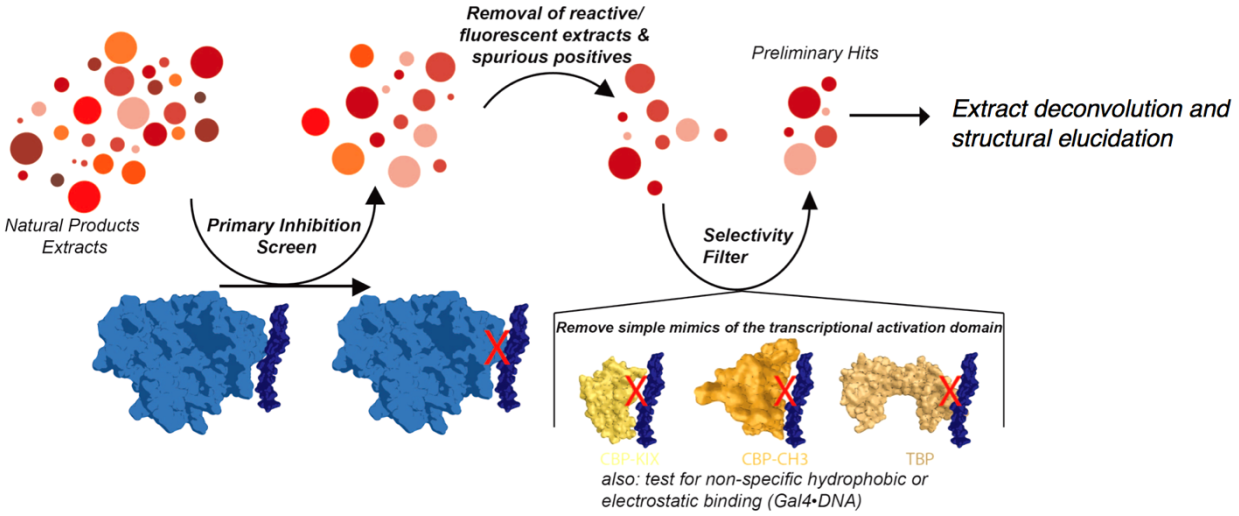


Figure 4.4- Generalized Iterative Screening Workflow for Allosteric Inhibitors of Activator-Coactivator Interactions Following primary screening and the removal of false positives or fluorescent/reactive extracts, initial hits are tested against a series of functionally related protein-protein interactions to identify inhibitors that are selective for the target of interest. Preliminary hits are subsequently fractionated and the structures of the active components elucidated.

In screening the ERM•AcID interaction, counter screens that test initial hits against Gal4•DNA and MLL•KIX interactions were utilized in order to identify selective inhibitors of ERM•AcID. We hypothesize that these inhibitors will likely be allosteric modulators of the AcID domain, allowing for the disruption of the broad ERM•AcID interaction surface.

4.3 Results and Discussion

Norstictic Acid and Psoromic Acid Inhibit ERM•AcID Interactions in vitro

Given the functional importance of the PEA3 subfamily of ETS transcription factors in a variety of cancers and the apparent importance of the interaction between the TADs of these activators and the AcID motif of Med25 for their activity, we sought to identify small molecule inhibitors that are capable of disrupting this interaction. These inhibitors will be useful mechanistic probes in exploring the role of the PEA3 subfamily in the establishment of metastatic phenotypes, the role of these activators in tumorigenesis, and could ultimately lead to the development of novel therapeutic strategies. In order to demonstrate that this interaction is targetable, the ability of norstictic acid and psoromic acid to inhibit the interaction between the ERM TAD and AcID was tested *in vitro*. The results of these experiments are shown below in Figure 4.5. It was hypothesized that

these molecules would be effective in disrupting ERM•AcID interactions as well as VP16•AcID interactions because of the similarity in the sequences of the two activators and the fact that both TADs contain α-helices that are responsible for the majority of their affinity, suggesting that they likely bind to similar surfaces of AcID using similar features for molecular recognition. Furthermore, this hypothesis is consistent with published NMR solution structures of the domain in complex with ERM and VP16. [15,32-34]

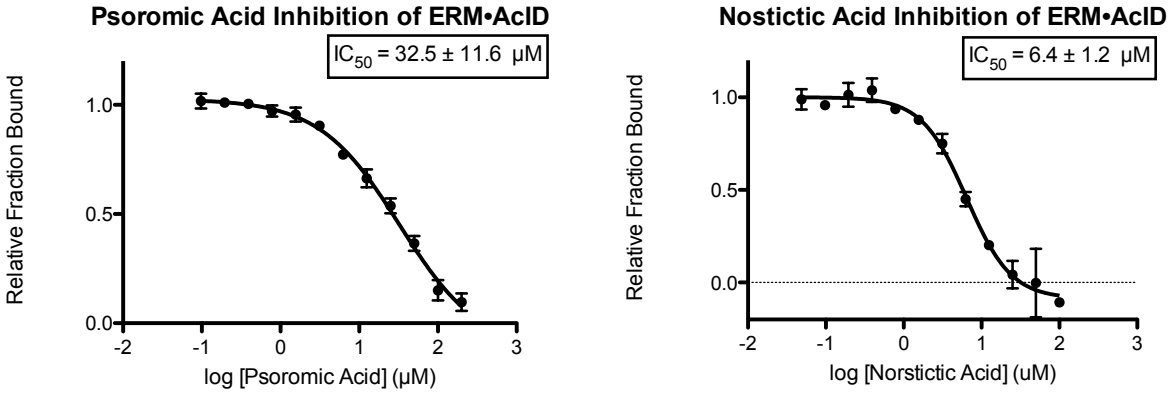


Figure 4.5- **Inhibition of ERM•AcID Interaction by Norstictic Acid and Psoromic Acid** Competitive inhibition experiments of the ERM(38-68)•AcID interaction by psoromic acid and norstictic acid. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the relative fraction of tracer bound at the indicated concentration of small molecule. Curves were fit using GraphPad Prism 5.

Consistent with this hypothesis, the ERM•AcID interaction is effectively inhibited by both psoromic acid and norstictic acid with IC_{50} values of 32.5 μM and 6.4 μm, respectively. Interestingly, psoromic acid, which inhibits the VP16(438-464)•AcID and VP16(465-490)•AcID interactions with IC_{50} values of 1.8 μM and 3.9 μM, is significantly less potent against the ERM•AcID interaction than VP16•AcID interactions. Norstictic acid, conversely, inhibits the ERM•AcID interaction with approximately the same potency as the VP16•AcID interactions. This observation suggests that the two molecules may bind to distinct surfaces of the AcID motif, given their differential activity against ERM and VP16 interactions. Alternatively, these data may suggest that norstictic acid induces more significant structural changes within the activator binding surfaces of AcID, resulting in the enhanced potency of this inhibitor against a number of activators. More importantly, these results suggest that, like the VP16•AcID interaction, the ERM•AcID interaction can be effectively targeted by small molecule inhibitors.

Subsequent to these observations, we next sought to demonstrate that norstictic acid has a similar mode of action against the ERM•AcID interaction as the VP16•AcID interaction by exploring the effects of lysine to arginine mutations on the activity of the inhibitor. These mutations, first described in Chapter 3, remove the ability of the inhibitor to bind covalently to specific surfaces on the AcID motif within the purported VP16 H1 and H2 binding sites and resulted in four to six fold decreases in potency when multiple residues were mutated within the same binding surface.

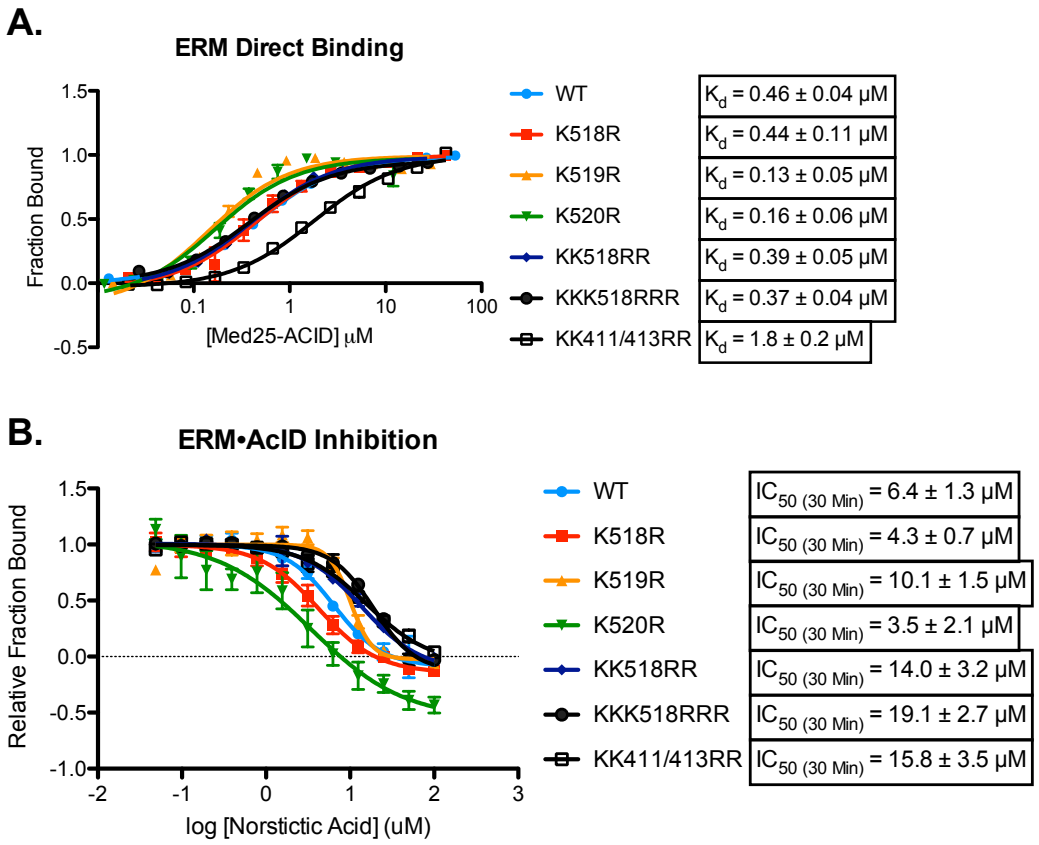


***Figure 4.6*- Effect of AcID Lysine to Arginine Mutations on Norstictic Acid Inhibition of ERM•AcID** (A) Direct binding of the fluorescent ERM(38-68) tracer to a variety of AcID constructs. (B) norstictic acid inhibition of the fluorescent ERM(38-68) interaction with a variety of AcID constructs. K518, K519, and K520 are located within the H2 binding site, while K411 and K413 are located within the H1 binding site. KK518RR refers to the double mutant K518R/K519R and KKK518RRR refers to the triple mutant K518R/K519R/K520R. Curves represent the mean values of three independent experiments with vertical error bars indicating the standard deviation of the fraction of tracer bound at the indicated protein or small molecule concentration. Curves were fit with GraphPad Prism 5.

The lysine to arginine mutations generally had little impact on the ability of ERM(38-68) to bind to AcID, with the exception of K411R/K413R, which resulted in an approximately four-fold decrease in affinity of the tracer for the domain. This mutation did

not perturb the binding of VP16(438-454), which is hypothesized to bind to a similar surface as ERM on AcID within the H1 binding site, but perturbed the binding of VP16(467-488) more strongly than ERM. Thus, these data may suggest that ERM binds to an AcID surface that is unique from either of the surfaces used by VP16(438-454) or VP16(467-488). The mutations K519R and K520R moderately enhanced the affinity of the ERM tracer for the domain, and this indicate allosteric communication between the binding sites within the domain that has been discussed previously in Section 3.3.

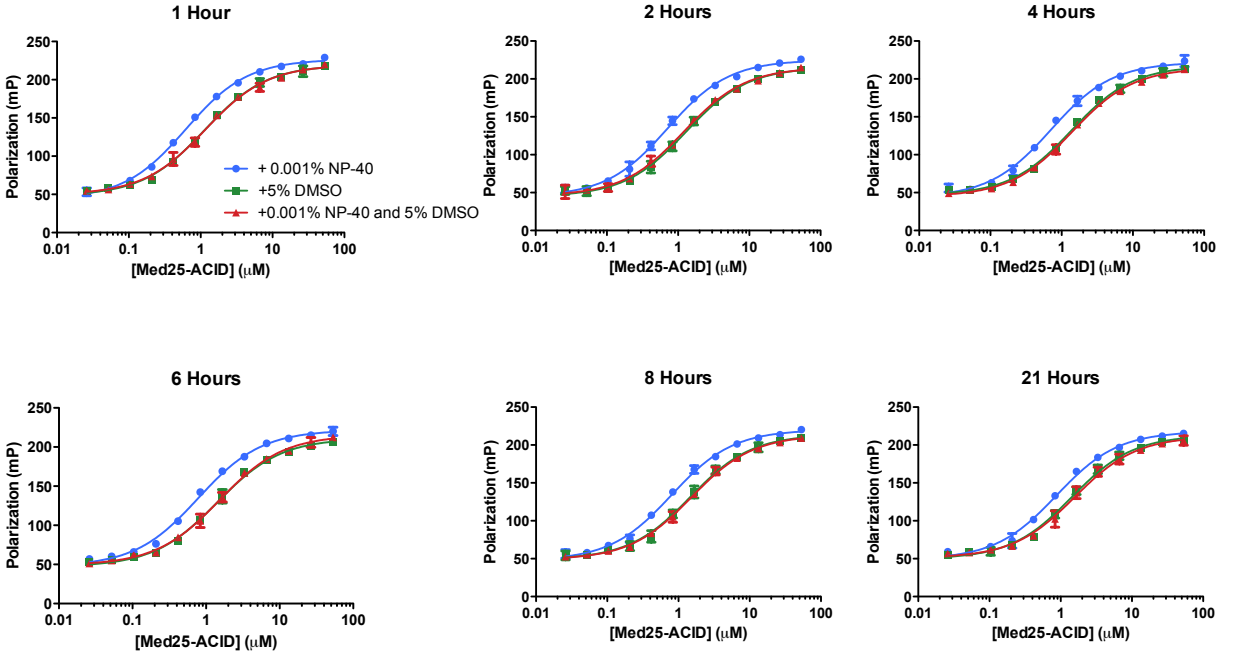
The lysine to arginine mutations moderately attenuated the inhibitory activity of norstictic acid against the ERM•AcID interaction, consistent with previous observations for the effects of the mutations on the inhibition of VP16•AcID interactions. This attenuation ranges from approximately two to three fold decreases in potency and requires that multiple lysine residues be mutated within the same interaction surface similar to the results obtained against the VP16•AcID interactions. The effects on potency against the ERM•AcID interaction were not as pronounced as the effects on potency against VP16•AcID interactions, indicating that the mechanism of action against the ERM interaction may be distinct from that of VP16. For example, the ERM•AcID interaction may be less dependent upon electrostatic contacts or the ERM binding site may be far enough removed from the mutated lysine residues that the loss of covalent labeling at those sites does not as adversely effect the inhibition of ERM.

ERM•AcID High-Throughput Screening Assay Development

Following the encouraging observation that small molecule inhibitors of the VP16•AcID interaction also demonstrated activity against the ERM•AcID interaction *in vitro*, we decided to complete a full high-throughput screen against the ERM•AcID interaction. In order to accomplish this, a fluorescence polarization based assay adapted for a high-throughput screening format using the ERM(38-68) fluorescent tracer and purified AcID protein was developed and optimized. The data shown in Figure 4.2 C demonstrates that this tracer binds to AcID over a broad dynamic range spanning approximately 160 mP. This span is slightly less than the approximately 210 mP observed for the VP16(465-490) tracer used in the screen discussed in Chapter 3, but more than sufficient. Building from this result, the dependence of the dynamic range of the binding

assay on the fluorescent tracer concentration was assessed by comparing direct binding assays completed with 20 nM fluorescent ERM(38-68) tracer and 50 nM tracer. Both assays gave identical dynamic ranges and K_d values, indicating that the assay is not dependent upon tracer concentration and that 20 nM tracer is sufficient for the screen.

Subsequently, the effects of DMSO and low concentrations of NP-40 were tested on the behavior of the assay, as was previously completed in the development of the VP16(465-490)•AcID HTS assay reported in Chapter 3. Compounds are added to the assay plates from concentrated DMSO stocks and NP-40 at concentrations of 0.001% v/v are added to assay buffer in order to minimize protein loss from non-specific interactions with the liquid-handling equipment used to prepare the assay plates and also minimizes protein or small molecule aggregation in solution. Furthermore, the behavior of the ERM(38-68)•AcID assay was tested over time in order to demonstrate that it is stable over the time period that a typical HTS assay requires. These results are shown below in Figure 4.7.



	K_d of ERM(38-68)•AcID Interaction					
	1 hour	2 hours	4 hours	6 hours	8 hours	20 hours
+ NP-40	0.61 μM	0.71 μM	0.70 μM	0.77 μM	0.79 μM	0.86 μM
+DMSO	1.1 μM	1.3 μM	1.4 μM	1.4 μM	1.5 μM	1.4 μM
+ NP-40 and DMSO	1.2 μM	1.2 μM	1.4 μM	1.5 μM	1.5 μM	1.6 μM

***Figure 4.7*- Assay Stability Time Course, DMSO Effects, and NP-40 Effects on Tracer Affinity**
The effects of NP-40 and DMSO on the ERM(38-68)•AcID interaction was assessed over time. Curves

represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the polarization at the indicated protein concentration. Curves were fit and analyzed using GraphPad Prism 5. *(With Paul A. Bruno)*

NP-40 at a concentration of 0.001% v/v has a minimal effect on the binding of ERM(38-68) to the AcID motif as the K_d value determined in the presence of NP-40 is within error of the value in the absence of the detergent. DMSO at a concentration of 5% v/v has a slightly stronger effect as the K_d value in the presence of DMSO at this concentration approximately doubles the value in the absence of the solvent. However, this effect is not so strong as to preclude the use of this assay in a HTS format and actual concentrations of DMSO used in the primary screen will be 1% v/v, unless higher concentrations of extracts are required to identify lead molecules. Finally, the assay is highly stable at room temperature over extended periods of time as the K_d values were within error of one another at one hour and twenty hours. Thus, these results indicate that an FP assay based upon the interaction between a fluorescent tracer of the ERM TAD and AcID is well suited to a high-throughput screening format.

The final experimental parameter to optimize prior to beginning the screen was the concentration of AcID protein to use in the assay. It was hypothesized that a concentration of 850 nM AcID protein would be appropriate given that this is the concentration of protein used in the VP16•AcID screen and the similarity in the affinities of VP16(465-490) and ERM(38-68) for AcID. In order to demonstrate that this concentration of protein was optimal for use in an HTS assay the Z' score was calculated using Equation 3.1 for a 384-well plate containing 182 wells each of the positive and negative controls. In this assay format, the positive control is the polarization of a sample containing the fluorescent ERM tracer only and the negative control is the polarization of a sample containing the fluorescent ERM tracer in complex with AcID. This experiment resulted in a calculated Z' score of 0.69 for 850 nM AcID, above the 0.6 threshold for an assay to be defined as high quality.[35] Thus, these experiments suggest that the ERM•AcID assay is well adapted to HTS and suitably optimized.

Given the demonstrated efficacy of natural products, specifically depsides and depsidones, against the VP16•AcID interaction and prior observations of molecules of this class inhibiting CBP/p300 KIX domain-dependent activator•coactivator interactions, we decided to screen the Natural Product Extract (NPE) library held by the Center for

Chemical Genomics (CCG) at the University of Michigan rather than more traditional small molecule libraries.[28] It was hypothesized that the greater structural diversity afforded by natural products may produce a greater number of lead molecules than the more 'flat' and drug-like structures typical of small-molecule compound libraries. As a proof of principle to support this hypothesis, a plate (Plate #2041) from the NPE library rich in extracts collected from lichens was tested for inhibition against the ERM•AcID FP assay. This plate was selected as lichens contain a significant number of depsides and depsidones and thus, based on the evidence presented in Chapter 3 that compounds of these classes may be privileged for the inhibition of activator•coactivator interactions, would presumably yield a relatively high hit rate. [36-40]

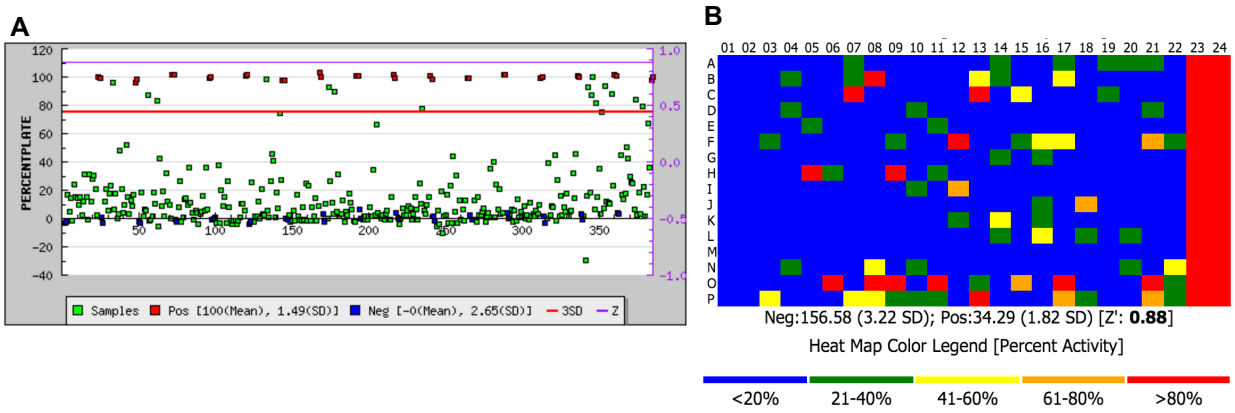


***Figure 4.8*- Lichen Rich Natural Product Extracts Tested Against ERM•AcID** (A) Campaign view of NPE Library Plate 2041 showing percent inhibition against the ERM•AcID interaction on the y-axis. (B) Heat map of NPE Library Plate 2041 showing inhibition against the ERM•AcID interaction. *(With Paul A. Bruno)*

The ERM•AcID FP assay performed well in testing against a representative segment of the NPE library with a Z' score of 0.88. Using the least stringent definition of a hit, a percent inhibition that is three standard deviations above the negative control (approximately 8% inhibition in this case), the lichen rich extracts produced 151 hits, for a hit rate of 47.2%. This rate is significantly higher than would be desired given the necessary work to validate extracts as hits and to elucidate active components from the extracts. Therefore, if such a rate were to manifest itself in the full NPE Library screen, alternative criteria to select lead compounds would be necessary. However, it was hypothesized at the time that this hit rate was most likely inflated and not necessarily indicative of expected performance in a full primary screen of the NPE library on the basis

that this plate is rich in lichen extracts and thus likely contains numerous extracts abundant in depsides and depsidones. Additionally, these results serve as further evidence consistent with the hypothesis that depsides and depsidones are privileged for the inhibition of activator•coactivator interactions.

Taken together, the above data demonstrate that we have developed a fluorescence polarization based assay of the interaction between the TAD of the transcriptional activator ERM and the AcID motif of Med25. Furthermore, this assay was tested against a small portion of the NPE library held by the CCG and found to perform well in a high-throughput format, though it was hypothesized at the time that alternative criteria for selecting hits might ultimately be necessary.

Primary Screen of NPE Library Against the ERM•AcID Interaction

Following the successful development of an assay of the ERM•AcID interaction that is well adapted to high-throughput screening and the encouraging results of a limited pilot screen of one NPE library plate against the target, a primary screen of the full NPE library against the ERM•AcID interaction was completed. In total, 33,400 unique extracts were tested against the assay, providing a significant number of potential lead extracts for further validation and analysis. The results of the primary screen, which had an average Z' score of 0.78, are presented as a campaign view in **Figure 4.9** below.

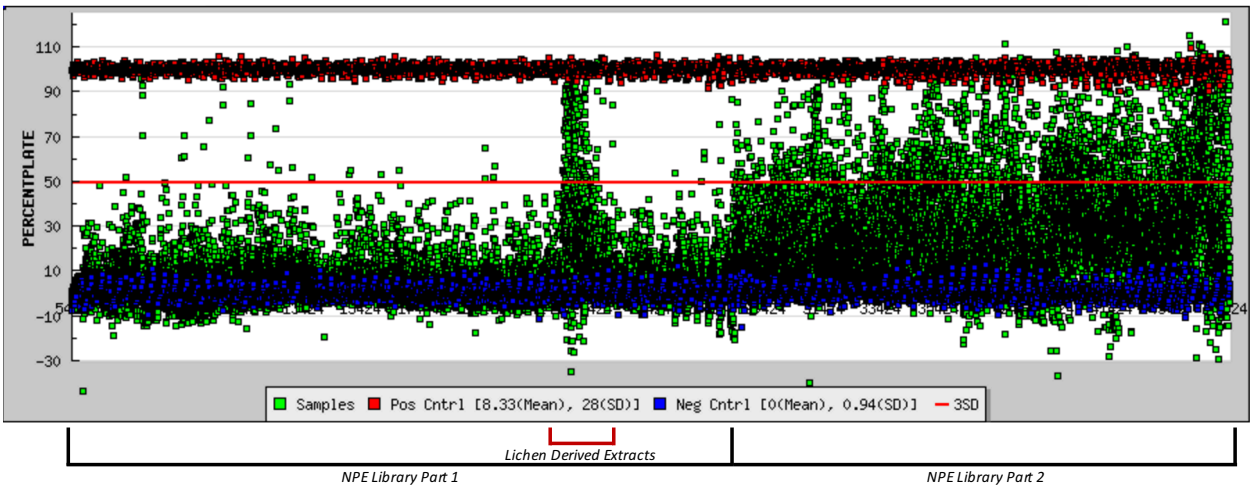


***Figure 4.9*- Campaign View of the ERM•AcID Primary Screen** A full campaign view of the primary screen of the NPE library against the ERM•AcID interaction. Positive control well measurements (tracer and protein, treated with DMSO) are shown as red points, negative control well measurements

(tracer only, treated with DMSO) are shown as blue points, and wells treated with 200 nL of extract are shown as green points. The x-axis represents the sample number, and the y-axis indicates the percent inhibition relative to positive and negative controls. The library has been broken into two parts given the dramatic shit in hit rates midway through the library. NPE library plates 2041 and 2042 are believed to be rich in extracts derived from lichens and are marked as such beneath the campaign view. *(With Paul A. Bruno)*

A striking feature of the primary screen campaign view is the dramatic shift in hit rates and extract efficacy midway through the screen. This is not a result of experimental error in preparing the assay or a failure of the assay design, but is instead a result of different methods for extract preparation. As the NPE library collection grew at a rapid pace, protocols for preparing extracts were changed from performing crude extraction with three solvents to using only two solvents. Furthermore, growth and media conditions were modified for extracts sourced from some bacterial strains, which may have had an effect on the production of certain natural products within these organisms. As a result, the extracts from the second part of the library now contain a greater number of component molecules at higher concentrations, which results in significantly inflated hit rates, as has been observed in other screens of the NPE library according to CCG personnel. Thus, when the initial hit filtering to select extracts for follow up study required the establishment of two distinct sets of criteria for the two subsets of the NPE library.

Beyond the dramatic difference in hit rates for the two methods of extract preparation within the library, there was also a significant increase in the hit rate near the end of the first part of the library. This segment of the library contains library plates 2041 and 2042, which are rich in extracts prepared from lichens and thus are likely abundant in depsides and depsidones. Thus, given the demonstrated efficacy of molecules belonging to these classes at inhibiting activator•coactivator interactions and the previously observed hit rate of plate 2041 against the ERM•AcID interaction, this high hit rate is not unexpected.

Overall, the least stringent definition of a hit as being an extract that demonstrated a percent inhibition greater than three standard deviations above the negative control, would result in 14,575 hits against the ERM•AcID interaction with 3,636 of them coming from the first portion of the library and 10,939 coming from the second portion. This would result in an overall hit rate of 43.6%, with individual hit rates of 20.8% and 68.4% for the first and second portions of the library, respectively. Obviously, such a vast number of

hits could not be sufficiently analyzed in follow up studies, requiring that alternative criteria be developed to select leads for further analysis.

For the first portion of the library, a hit was defined having greater than or equal to 30% inhibition of the ERM•AcID interaction relative to controls, identifying 557 extracts, or inhibition that was greater than or equal to three standard deviations above *all* wells on the plate, identifying an additional 166 extracts. Including extracts that are above three standard deviations of all wells on the plate ensured that a mix of extracts from a variety of organisms was selected by advancing the best extracts from each plate, even if the activity was below the 30% threshold. Of the 723 total extracts identified from the first portion of the library, eleven were eliminated on the basis of excess fluorescence within the well and 43 were eliminated on the basis of promiscuity, with a promiscuous compound defined as being active in more than 30% of screens it is tested against. Thus, the first portion of the NPE library ultimately provided 669 compounds for further analysis and a significantly more manageable hit rate of 3.4%.

For the second portion of the library, an extract was defined as a hit if it had greater than 60% inhibition or extracts that displayed inhibition greater than three standard deviations above all wells on the plate, again to provide extracts from a mix of plates and not just those compounds with the greatest apparent efficacy. These criteria identified 1240 extracts, 265 of which were eliminated on the basis of excess fluorescence within the well and 42 of which were eliminated on the basis of promiscuity. Thus, the second portion of the library identified 933 extracts for further analysis and a hit rate of 5.8%. Combined, the two portions of the NPE library provided 1,602 extracts, with an overall hit rate of 4.8%, for follow up study that spanned a range of efficacies and were derived from a number of genetically diverse organisms. Overall, the identification of 1,602 hits using stringent selection criteria suggests that the AcID motif may be significantly more targetable than typical coactivator proteins. For example, a high-throughput screen of the NPE library against the interaction of the transcription factor MLL with the KIX domain of CBP/p300 identified 280 hits for follow up studies, where a hit was defined as greater than 20% inhibition of the interaction.[28]

The next step in the screening strategy was to further filter the 1,602 extracts identified in the primary screen by eliminating those samples that demonstrated

significant native fluorescence or quenched the fluorescent ERM tracer, as these phenomena could interfere with the read-out of the FP assay. Additionally, the primary extracts were retested against the ERM•AcID interaction and eliminated if they did not inhibit the assay greater than 50% in two out of the three wells, which served as a method to eliminate false positives and also functioned as an additional potency filter. The results of these three confirmation experiments are presented below in Figure 4.10.

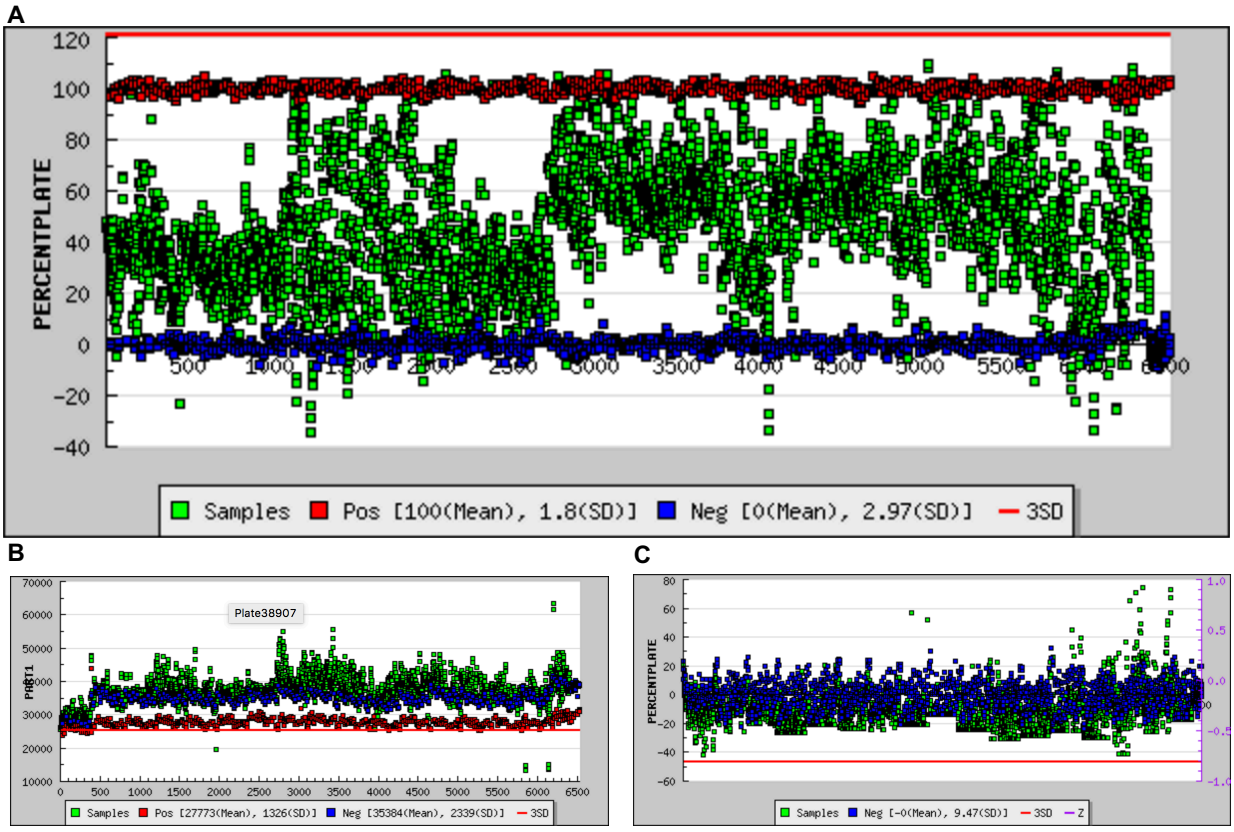


***Figure 4.10*- Confirmation of Hits from the ERM•AcID Primary Screen** (A) The top 1,602 extracts identified in the primary screen against the ERM•AcID interaction were retested in triplicate to remove false positives from final hit selection. (B) Intrinsic fluorescence of lead extracts identified in the primary screen against the ERM•AcID interaction. (C) Fluorescence quenching of the ERM tracer by lead extracts identified in the primary screen against the ERM•AcID interaction. *(With Paul A. Bruno)*

In order to complete these confirmation experiments, the order of addition of assay components was reversed. Specifically, extracts were first added to buffer and intrinsic fluorescence was measured, the fluorescent ERM tracer was added second and the plate was measured for fluorescence quenching, and AcID protein was added last at which point samples were incubated for thirty minutes and the polarization of the samples were

measured. Following these assays, the 1,602 lead extracts were filtered to reduce the number of hits carried on to final analysis.

The first criterion used in this filtering step was potency, wherein only extracts that achieved a greater than 50% inhibition in all three of the retest samples were selected as hits, based on the results shown in Figure 4.10 A. The potency filter eliminated 888 of the initial lead extracts, leaving 714 extracts for further consideration.

The second criterion used for initial hit filtering was significant intrinsic fluorescence of compounds within the extracts. Extracts that resulted in a fluorescence signal of greater than 40,000 mAU in the parallel channel in two of the three retest samples were eliminated from further consideration, a value which represents a 40% increase over the fluorescence of the tracer alone or a 15% increase over the fluorescence of the tracer in complex with the AcID protein. Though, fluorescent compounds may still be capable of inhibition, the potential for false positives as a result of interference in the FP results requires that these extracts be removed from follow up study.[41] Using this criterion for significant intrinsic fluorescence, 406 extracts were flagged as fluorescent from the full sample of 1,602 extracts; of the 714 extracts that passed the potency filter, an additional 254 were removed following the intrinsic fluorescence filter, resulting in 460 extracts selected for further analysis.

The final criterion used to further filter the additional hits identified in the primary screen was the ability of the extracts to quench the fluorescence of the tracer, could result in the identification of false positives as hits.[41] Specifically, extracts were added to sample wells that contained tracer only and were eliminated from further consideration if their ability to quench the fluorescein tag of the tracer resulted in greater than 30% inhibition in two of the three retest samples, corresponding to three standard deviations above the negative control. On the basis of this criterion, twenty-seven of the 1,602 extracts identified in the primary screen were found to significantly quench the fluorescent tracer, though these twenty-seven extracts had been eliminated from further consideration by the previous potency and intrinsic fluorescence filtering steps. Ultimately, the initial hit filtering completed on the 1,602 extracts identified for follow up study from the primary screen resulted in 460 extracts that were validated as inhibitors of the ERM•AcID FP assay.

Counter-Screens to Eliminate Non-selective ERM•AcID Inhibitors

In order to increase the likelihood that extracts with natural products that selectively target the ERM•AcID interaction were identified and to further refine the number of extracts identified as hits against this target, a series of counter screens to eliminate inhibitors that do not specifically inhibit AcID-dependent interactions were completed. The first counter screen utilized to further filter the selected extracts was an FP assay based upon the interaction between the DNA binding domain of the yeast transcriptional activator Gal4, spanning residues 1-100, and a fluorescently tagged DNA oligomer. This assay, developed by our laboratory, uses a fluorescein-tagged twenty base pair DNA oligomer containing a consensus Gal4 DBD binding sequence that has been annealed to an unlabeled complementary oligonucleotide to simulate the binding of the DBD to double-stranded DNA.[42] Furthermore, our lab has previously employed this assay in a high-throughput format as a counter-screen in the discovery of natural product inhibitors of the interaction between transcriptional activators and the KIX domain of CBP/p300.[28] In particular, it was hypothesized that this was an attractive assay to use as a counter screen in order to eliminate natural products that may interfere with the ability of transcriptional activators to bind to their requisite DNA sequences. Furthermore, DNA-protein interactions are partially driven by electrostatic interactions, similar to the interaction between activators and AcID.[43] Thus, a counter screen against an electrostatically driven interaction would also be useful in eliminating NPEs that indiscriminately interfere with electrostatic contacts. Prior to completing the counter screen, we first ran a direct binding experiment of the fluorescent DNA oligomer to Gal4(1-100) in order to ensure that the assay performed as expected in our hands, as shown in Figure 4.11.

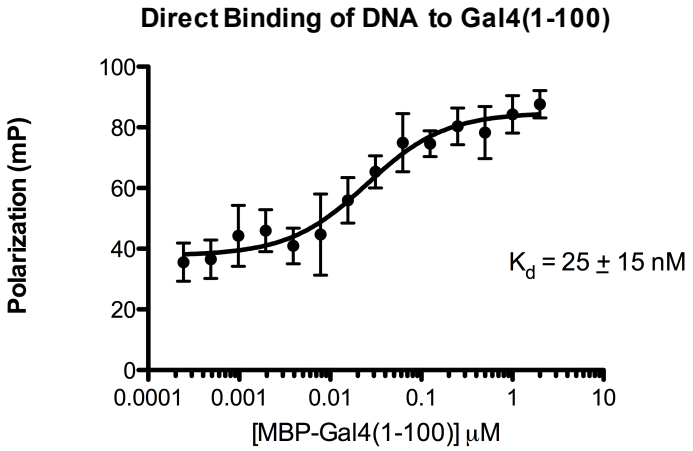


***Figure 4.11*- Direct Binding of Fluorescent DNA Oligomer to Gal4(1-100)** The Gal4(1-100)•DNA FP assay was tested prior to use as a counter screen by performing a direct binding experiment of the oligomer to the Gal4 DBD. The curve represents the mean values of three independent experiments with vertical error bars representing the standard deviation of the polarization in the presence of the indicated concentration of protein. Curves were fit using GraphPad Prism 5. *(Completed by Paul A. Bruno)*

The assay performed as expected with a calculated K_d value of 25 ± 15 nM, which was within error of previously reported values for the affinity of this interaction.[42] Subsequently, this FP assay was utilized in a high-throughput format as a counter screen against the 1,602 extracts identified as hits from the primary screen of the NPE library against the ERM•AclD interaction, as shown in the campaign view of **Figure 4.12**. The screen was completed using 50 nM Gal4(1-100) and 10 nM fluorescent DNA tracer with extracts tested in triplicate, per conditions previously utilized in counter screening applications.[28]

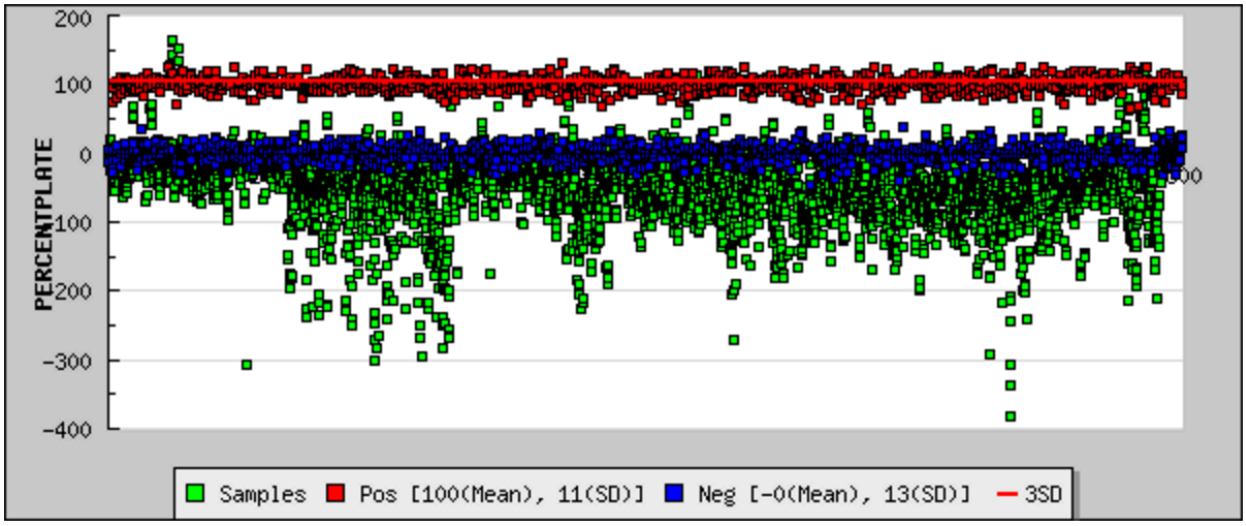


***Figure 4.12*- Gal4(1-100)•DNA Counter Screen of Lead Extracts from ERM•AclD Primary Screen** Campaign view of Gal4(1-100)•DNA counter screen against the 1,602 extracts selected from the primary screen of the ERM•AclD interaction. Red points represent positive controls (fluorescent tracer only), blue points represent negative controls (tracer in complex with Gal4(1-100)), and green points represent samples treated with NPEs. The percent inhibition of the assay is shown on the y-axis. *(Completed with Paul A. Bruno)*

Given the relatively large standard deviation within the negative controls of 13% inhibition, a threshold of 50% inhibition in at least two of the three samples was set for the elimination of an extract from further consideration. Using this criterion, 291 extracts were identified as inhibitors of the Gal4(1-100)•DNA interaction. Of those extracts, only ten had not been previously eliminated from further consideration by the initial hit filtering steps, leaving 450 extracts for further analysis.

One striking result observed in the campaign view of this counter screen is the number of extracts that demonstrated a significant negative percent inhibition. One potential explanation for this phenomenon is that the dynamic range of this assay spans approximately 35 mP, which is significantly smaller than the dynamic range of the primary assay. Furthermore, the standard deviation for the polarization values of the positive and negative controls were 4.23 mP and 4.82 mP, respectively, further obfuscating the analysis by decreasing the effective dynamic range. A final potentially confounding factor is that the concentration of fluorescent tracer in this assay was only 10 nM, as opposed to the 20 nM concentration used in the primary screen. Thus, even though the fluorophore was identical between the two assays, the higher gain setting of the plate reader for the Gal4(1-100)•DNA interaction likely increased the noise within the measurements as the

intrinsic fluorescence of compounds within the extracts would be elevated relative to the fluorescence of the tracer at higher gain settings.

Subsequently, a counter screen of the 1,602 extracts identified in the primary screen against the interaction between the transcription factor MLL and the KIX domain of CBP/p300 was completed. This assay was previously used as a method for demonstrating the selectivity of norstictic acid and psoromic acid against activator•AcID interactions in Chapter 3 and has also been previously utilized in a high-throughput screening application by our laboratory.[28] This assay is an attractive counter screen as our lab and others have extensively studied the interaction and demonstrated that it relies upon α-helical secondary structure within the TAD of MLL and that the binding of the TAD occurs through a two-step mechanism wherein the unstructured TAD makes preliminary contact to KIX before properly folding into an α-helix, similar to several of the features common to activator•AcID interactions.[42,44] Thus, this counter screen was useful for the elimination of generic TAD mimetics and enriched the preliminary hits for inhibitors selective for the AcID motif. Prior to completing the counter screen, a direct binding experiment of the fluorescent MLL tracer to the KIX domain, as well as a direct binding experiment of a fluorescent tracer derived from the TAD of the transcriptional activator CREB (pKID), were completed in order to ensure that the assay performed well in our hands, as shown in Figure 4.13.

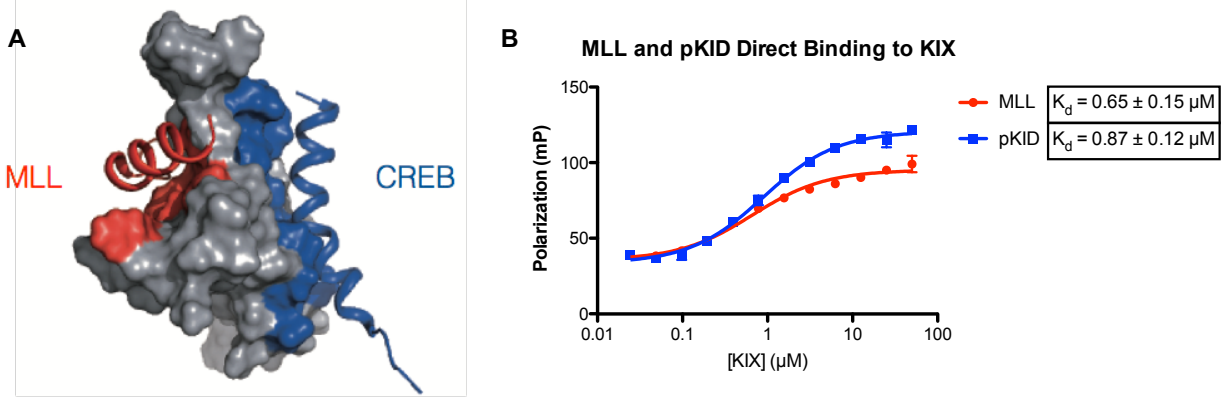


***Figure 4.13*- Direct Binding of MLL and pKID to the KIX Domain of CBP/p300** (A) Structure of the KIX domain in complex with the TADs of MLL and CREB indicating their preferred binding sites on the domain. (B) Direct Binding of the MLL and pKID fluorescent tracers to the KIX protein. Curves represent the mean values of three independent experiments with vertical error bars representing the standard deviation of the polarization value at the indicated concentration of protein. Curves were fit with GraphPad Prism 5.

The direct binding assays performed as expected with the calculated K_d values of both tracers within error of previously reported values, indicating that the assay performs well in our hands with the prepared fluorescent tracer and purified protein. [42,45,46] The MLL•KIX FP assay was then used as a counter screen against the 1,602 extracts identified in the primary screen of the ERM•AcID interaction, as shown in Figure 4.14. The counter screen was completed with 10 μM KIX protein and 25 nM fluorescent MLL tracer with extracts tested in triplicate, consistent with conditions utilized in a previous screening application.[28]

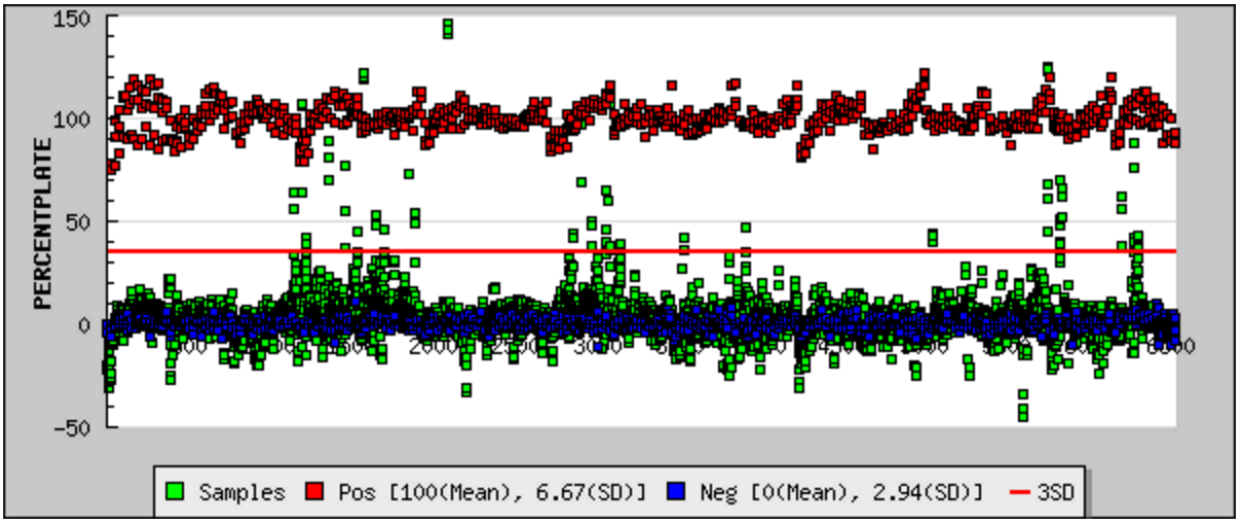


***Figure 4.14*- MLL•KIX Counter Screen of Lead Extracts from ERM•AcID Primary Screen** Campaign view of the MLL•KIX counter screen against the 1,602 extracts selected from the primary screen of the ERM•AcID interaction. Red points represent positive controls (fluorescent tracer only), blue points represent negative controls (tracer in complex with KIX), and green points represent samples treated with NPEs. The percent inhibition of the assay is shown on the y-axis. *(Completed with Paul A. Bruno)*

The MLL•KIX counter screen assay performed as expected with a Z' score of 0.72, indicating a high quality assay. A threshold of 35% inhibition in at least two of the three replicates was set in order to eliminate an extract from further consideration. This value was selected in order to eliminate extracts with relatively potent inhibition of KIX-dependent activator•coactivator interactions. On the basis of this criterion, 85 extracts were identified as inhibitors of the MLL•KIX interaction, with 33 of these extracts not having been eliminated by previous filtering steps, leaving 417 extracts for further analysis with demonstrated selectivity for AcID-dependent interactions. The relatively low hit rate

of the extracts identified from the primary screen is consistent with previous observations that targeting the KIX domain of CBP/p300 is challenging.[28]

As a last filter prior to determining the final list of confirmed hits, extracts that had been active in greater than ten screens of the NPE library were eliminated in order to remove overly promiscuous natural products. Using this criterion, an additional eighty-five extracts that had not been removed by previous filtering steps were eliminated for a final total of 332 unique natural product extracts that harbor activity against the ERM•AcID interaction.

Summary of Hit Filtering and Selection of Extracts for Strain Regrowth

Following the filtering of lead extracts identified in the primary screen and the completion of counter screens to identify selective inhibitors of the ERM•AcID interaction, we identified 332 unique extracts as containing potential inhibitors of the interaction between the ERM TAD and the AcID motif of Med25. The full process that identified the final extracts for fractionation and active component elucidation in summarized below in Figure 4.15. A full listing, with relevant screening statistics, of the 332 extracts identified as inhibitors of the ERM•AcID interaction is presented in Appendix B.

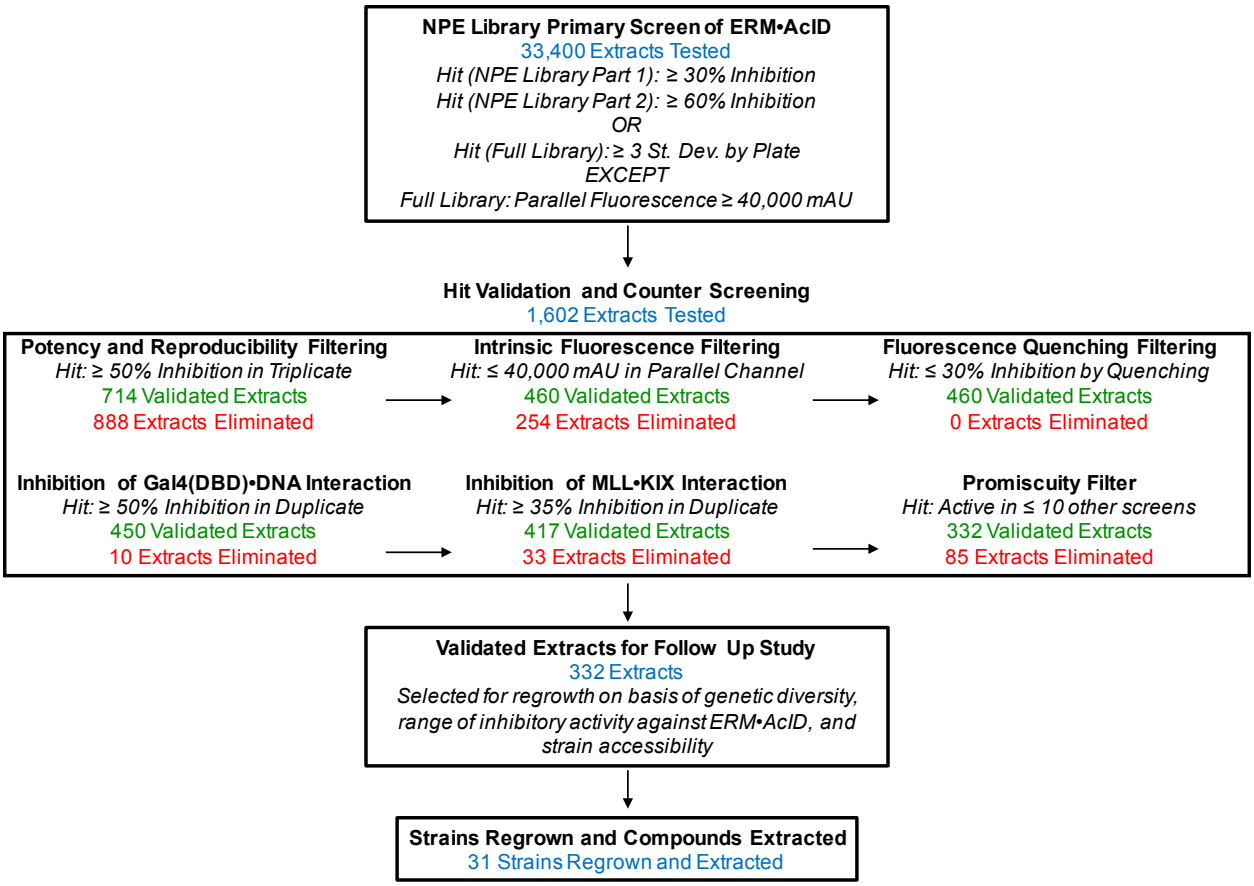


***Figure 4.15*- Schematic Representation of Final Hit Selection** The full battery of screens and counter screens utilized to identify natural product extracts containing small molecule inhibitors of the ERM•AcID interaction is represented schematically.

Subsequent to the final hit selection, a subset of extracts was selected to characterize further. This process requires extracting new samples from the original organisms, either by growing new cultures of microorganisms or obtaining intact samples of macroscopic organisms depending on the extract, retesting the crude extracts for activity to confirm that the active component in the NPE library is not a result of degradation of the sample, fractionating the extract to isolate and purify the active components, and finally elucidating the structure of the active molecule. Subsequent to this characterization, the genome of the organism that provided the natural product is submitted for sequencing in order to definitively identify the organism and to potentially identify the biosynthetic pathways responsible for the production of the molecule.[47] In collaboration with the laboratory of Prof. David Sherman at the University of Michigan, who are responsible for preparation of the extracts for the NPE library and have significant experience with the isolation and structural characterization of natural products, thirty-one

extracts were selected for further characterization. These extracts were selected on the basis of: (1) genetic diversity by selecting extracts from across the NPE library collected from a variety of locales across the globe, (2) a range of inhibitory activity against the ERM•AcID interaction to identify a variety of molecules with distinct modes of action that occupy distinct chemical space, and (3) accessibility of the organism that yielded the initial extract held within the NPE library. These thirty-one extracts were all derived from microorganisms, with the majority believed to belong to the *Actinomyces* genus of the actinobacteria class, though the specific species of the organisms has not yet been determined. This observation is encouraging given that marine actinomycetes have been demonstrated as a remarkable source of diverse biologically active natural products with unique anti-microbial and anti-cancer activities.[48] The Sherman lab was in possession of spore stocks of all thirty-one organisms and was able to reconstitute pure cultures following standard protocols. [28,29,49,50] The thirty-one strains regrown for further analysis, along with their activities in the various screens and counter screens, are shown below in **Table 4.1**.

***Table 4.1*- Strains Selected for Regrowth with Filtering Statistics**

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	% Fluorescence Quenching	% Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
A	06131R	SID-142998	56.3	35430	-6.1	-47.7	-7.9	2
B	12587-2Z	SID-143065	83.2	37498	3	-32.8	9.1	5
C	32294-H1Z	SID-142868	82.1	35766	3.1	-103.6	13	5
D	34908-2Z	SID-142867	83.2	35799	-3.7	-108.6	8.2	2
E	54916-1C	SID-141597	75	38438	-26.3	-101.4	7.9	4
F	78874R	SID-143056	65.5	36958	-6.5	-10.1	9.9	3
G	91085R	SID-142997	85.2	35887	-8.6	30.6	7.1	4
H	PAN101-7I	SID-32305	74.1	38049	-20.4	-88.2	9.1	5
I	41445-N3I	SID-31523	80.4	39056	-13.9	-28.7	13.3	8
J	44293-N1N	SID-33128	77.8	38368	-21.9	-90.4	-0.9	8
K	71961N	SID-69043	70	38711	-14.3	-53.4	3.4	5
L	41429-N1I	SID-68148	82.3	39516	-28.6	-85.4	4.1	4
M	39040-1I	SID-27645	64.4	39972	-15.1	-68.2	1.7	6
N	68157N	SID-32195	55.1	36047	-3.1	-20.7	-1	3
O	36180-2N	SID-31830	57.4	37403	-10	-16.4	3.9	4
P	49546-N4I	SID-68085	52	39949	-14.7	-52.7	-4.3	8
Q	52245-N2N	SID-33188	101.7	35559	-18.2	11.9	20	5
R	58236N	SID-140604	79.9	33193	-0.3	-7.9	9.3	5
S	65440-N1N	SID-68121	70	36498	-10.3	-34.5	6.5	3
T	71747Z	SID-142274	81.2	38212	-16.8	-74	10.8	4
U	86815-N2Z	SID-143008	102.5	38773	-8.9	-19.8	40.1	4
V	5538-A2N	SID-33057	86.2	38027	-19	-54.9	7.1	7
W	18163-N13Z	SID-142848	96.6	33242	3.9	-66.3	27.3	4
X	24815-H2Z	SID-142871	90.4	33750	32.7	-33	24.1	4
Y	55270-N4N	SID-68851	73.7	39860	-18.4	-60.5	2.8	4
Z	74604-N1I	SID-140915	74	38202	-18.3	-45.9	6.1	3
AA	82349-N4I	SID-140914	73.6	37721	-13.3	-39.7	7.2	2
BB	20731-H2I	SID-18195	77	34656	-5.9	-4.5	2.2	10
CC	73401N	SID-140641	57	36386	-10.6	-46.6	5.1	2
DD	67360-N9N	SID-68685	70.1	35517	-3.6	-18	11.4	4
EE	84215-1I	SID-141865	100.5	35290	-4.5	-33.1	16.7	3

For the purpose of ease in identifying these strains and their extracts in subsequent experiments and discussions, they have each been assigned a letter ranging from A to EE, as indicated in Table 4.1.

Confirmation of Crude Natural Product Extract Activity from Regrown Organisms

Following the growth of pure samples of the organisms selected for further analysis, the cultures were homogenized and crude natural product extracts were prepared using a variety of solvents including acetone (solvent **1**), methanol (solvent **2**), dichloromethane (solvent **3**), and a one to one mixture of acetone and methanol (solvent **4**). The extracts from each organism were thus named by the letter corresponding to the strain from which the extract was derived and the number of the solvent used to prepare the extract. In total, sixty-nine crude extracts were prepared, and the ability of these extracts to inhibit the ERM•AcID interaction was determined by testing their ability to inhibit the ERM•AcID FP assay used in the primary screen in duplicate. The results of this experiment are presented as a heat map in Figure 4.16.

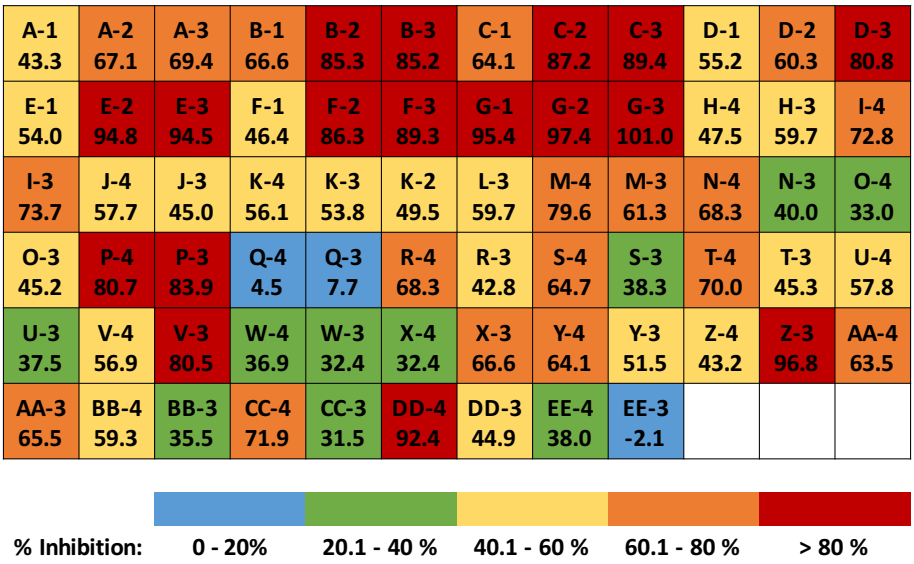


Figure 4.16*- Heat Map of Inhibition by Extracts from Regrown Organisms Against the ERM•AcID Interaction** The letter in the extract identifier at the top of each well refers to the 31 strains from ***Table 4.1, while the number after the dash in the identifier refers to the solvent used to extract compounds from the homogenized organism. Specifically, **1** refers to extraction with acetone, **2** refers to extraction with methanol, **3** refers to extraction with dichloromethane, and **4** refers to extraction with a one to one mixture of acetone and methanol. The percent inhibition of the ERM•AcID FP assay by the crude extracts are given below the extract identifier. All measurements were made in duplicate and percent inhibitions were calculated using positive (ERM tracer only) and negative (ERM•AcID treated with DMSO) controls. *(Completed by Matthew S. Beyersdorf)*

With the exception of strains **O**, **Q**, **W**, and **EE**, all strains produced at least one extract with greater than fifty percent inhibition of the ERM•AcID interaction, the previously used threshold in selecting the final hits from the screen. In particular, strain **Q** appears to have been a false positive given that the two extracts prepared from this organism inhibited the interaction by 4.5% and 7.7%. A possible explanation for this false positive result could be that the compounds within the extract degraded into more active molecules over time. Generally, the freshly prepared crude extracts were able to reconstitute the activities against the ERM•AcID interaction that were observed during the original screen of the interaction. Of particular note, strain **G** appears to be especially promising as all three extracts prepared from this organism resulted in greater than 95% inhibition when tested against the ERM•AcID FP assay. Overall, these data demonstrate that a majority of the strains selected for further analysis likely contain active natural products capable of inhibiting the interaction between the ERM TAD and AcID and that additional filtering will be necessary prior to selecting an extract for fractionation and structural elucidation of the active components.

In addition to testing the extracts from the regrown strains against the ERM•AcID FP assay, they were also tested against the retinoic acid receptor α (RARα) reporter assay first described in Chapter 3. This assay was performed in order to determine which extracts contained component molecules that were cell permeable and capable of disrupting AcID-dependent interactions in a complex cellular context. Med25 acts as an important coactivator contact for RARα that allows for the simultaneous recruitment of the Mediator complex as well as the HAT CBP.[51,52] The recruitment of CBP and its requisite HAT activity is a requirement for full transcriptional activity of RARα.[53-55] Therefore, it was hypothesized that, similar to results obtained with norstictic acid and psoromic acid in Chapter 3, that the freshly prepared extracts would be capable of decreasing RARα activity in a luciferase reporter assay if they contained component small molecules that were cell permeable and capable of inhibiting the interaction between the N-terminus of CBP and the AcID motif of Med25. The results of testing the extracts prepared from regrown strains against the RARα reporter assay are shown below in Figure 4.17.

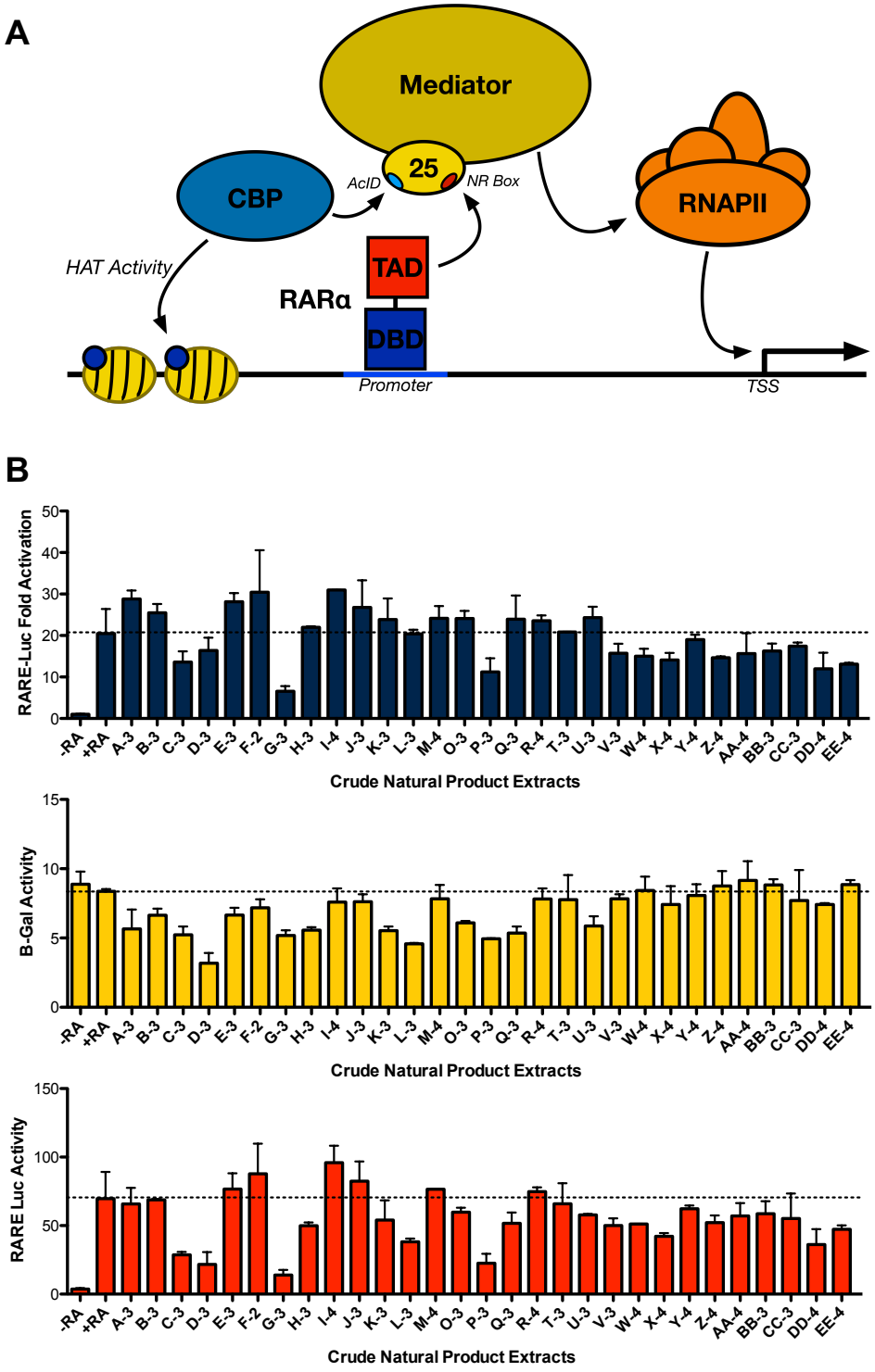


***Figure 4.17*- Inhibition by Crude Natural Product Extracts from Regrown Organisms Against a RARα Luciferase Reporter Assay** (A) Schematic of the reporter assay and the assembly of the PPI network at the Retinoic Acid Response Element. NPEs are hypothesized to inhibit the interaction between CBP and the AcID motif of Med25. (B) Inhibition of luciferase expression by the indicated crude NPE. The dashed lines represent the negative control, which is activity following induction by retinoic acid in the absence of inhibitor. All values represent biological duplicates. *(Completed by Paul A. Bruno)*

A number of the extracts tested against the RARα reporter assay demonstrated significant inhibition of RARα activation. In particular, the dichloromethane extracts of strains **G** and **P** demonstrated significant inhibition in the assay, suggesting that the active molecules within these extracts are cell permeable and capable of inhibiting AcID-dependent interactions in a cellular context. This activity is consistent with the results of these extracts against the ERM•AcID FP assay. Beyond these samples, a number of other extracts also demonstrated inhibition of the reporter assay, though not as effectively as extracts from strains **G** and **P**. Additionally, a number of other extracts showed no inhibition of the reporter assay, such as extracts from strains **A**, **B**, **E**, **F**, **I**, **J**, **K**, **L**, **M**, **O**, **Q**, **R**, **T**, **U**, and **Y** suggesting that the active molecules within these extracts may not be cell permeable. Another possibility for the lack of activity in these extracts is that little is known about the binding surfaces required for the interaction between the N-terminus of CBP and AcID and it is possible that this interaction occurs at a site on AcID unique from the binding site for ERM. Thus, these extracts may be selective for inhibition of the ERM•AcID interaction. Additional experiments will be completed to further explore the activity of these extracts in disrupting ERM•AcID interactions in a cellular context. Despite these caveats, this experiment is useful in selecting an initial extract for fractionation and structural characterization. Specifically, the dichloromethane extract of strain **G**, sample **G-3**, was selected for fractionation in order to identify and structurally characterize the active small molecule inhibitor, which is currently ongoing.

Fractionation of the extract by sequential preparative HPLC purification has identified three fractions that harbor significant activity against the ERM•AcID interaction (>75% inhibition). Additional HPLC purification of these fractions is being completed to ensure that they have been purified to homogeneity, at which point the material will be analyzed and structurally characterized using a combination of mass spectrometry, NMR, and x-ray crystallography if the chemical properties of the purified natural products permit. Furthermore, genomic DNA has been isolated from cultures of strain 91085R and submitted for sequencing in the hopes of identifying the specific species of the strain and to potentially identify the biosynthetic pathway by which the natural product of interest is produced. Following the final purification and structural elucidation, a full battery of *in vitro* experiments will be completed to more thoroughly characterize the activity of these

molecules against the ERM•AcID interaction and to identify the mechanism of action for the observed inhibition. Subsequent experiments will explore the use of the natural products as chemical probes in a cellular context in order to explore the role of Med25 in PEA3 transcriptional programs and the effect of this interaction on specific functional and phenotypic responses within a variety of cancer-derived cell lines. These experiments are described in greater detail in the Future Directions section of Chapter 5.

4.4 Conclusions

The driving hypothesis for the work completed within this chapter was that novel natural products capable of inhibiting the interaction between the transcriptional activation domains of the PEA3 subfamily of ETS transcription factors and the AcID motif of Med25 could be identified by screening a library containing approximately 33,000 natural product extracts against this interaction. The discovery of such molecules is an attractive goal as these inhibitors will be useful molecular probes for the elucidation of the role of the PEA3 subfamily in cellular processes related to cancer progression, specifically metastasis and possibly tumorigenesis, and could potentially serve as lead molecules in the development of novel therapeutic strategies. Though the depsidones norstictic acid and psoromic acid have been previously identified as covalent inhibitors of AcID-dependent activator•coactivator interactions, it is hypothesized that alternative non-covalent inhibitors of AcID-dependent interactions may allow for greater selectivity in a complex cellular environment and may be capable of inhibiting activator interactions with the domain through an alternative mechanism of action. In order to identify these inhibitors, an iterative screening strategy was utilized that eliminates generic TAD mimetics and identifies molecules that selectively target AcID-dependent interactions, thereby increasing the likelihood of identifying conformationally selective inhibitors of the domain.

Towards this goal, the ability of norstictic acid and psoromic acid to inhibit the interaction between the TAD of the PEA3 subfamily member ERM and AcID was tested using an FP assay to demonstrate that the interaction could be effectively targeted and inhibited by small molecules. Both molecules were capable of inhibiting the ERM•AcID interaction, albeit with reduced potency in the case of psoromic acid relative to the inhibition of the VP16•AcID interaction as shown in Chapter 3. These data suggest that

the ERM•AcID interaction can be inhibited by small molecule natural products and that the TADs of PEA3 subfamily members may interact with AcID at a surface unique from the VP16 interaction surface. Thus, in addition to inhibiting activator•AcID interactions through unique mechanisms of action, novel natural product derived inhibitors identified in a screen may also be significantly more potent for inhibition of ERM•AcID interactions by more directly interfering with the AcID surfaces required for interaction with the PEA3 subfamily. By targeting the AcID motif instead of the activators themselves inhibitors will perturb the activity of all three PEA3 subfamily members, mitigating functional redundancy within the subfamily in the regulation of transcriptional programs related to the progression of several cancers.

Following the observation that the ERM•AcID interaction could be effectivity inhibited by small molecules, a fluorescence polarization based assay of this interaction was developed and optimized for high-throughput screening. Subsequently, a primary screen of the NPE library against the optimized ERM(38-68)•AcID assay resulted in 1,602 hits that were selected for further filtering steps. These extracts were selected using criteria significantly more stringent than is typically used in screens at the CCG in order to reduce the number of lead extracts to a manageable quantity. The fact that such measures were necessary suggests that the ERM•AcID interaction is likely far more amenable to small molecule inhibitors than is typical for protein-protein interactions in general and activator•coactivator interactions in particular. For comparison, a screen of the interaction between the transcription factor MLL and the KIX domain of CBP/p300 identified only 280 lead extracts using far laxer criteria. Had the same criteria been used to select lead extracts against the ERM•AcID interaction, there would have been in excess of 10,000 extracts for follow up study. Interestingly, a number of extracts resulted in enhanced binding of the ERM peptide to the AcID motif, suggesting that it may be possible to identify positive modulators of the interaction that enhance the affinity of the activator for AcID. The mechanism by which this occurs is unclear, but significant possibilities are that they stabilize the domain in a conformation that is preordered for interaction with the ERM TAD or stabilize the ERM•AcID complex and reduce the rate of dissociation. Such modulators would be useful as probes to further explore the role of ERM and the other PEA3 activators in malignant processes.

These initial hits were then filtered to remove false positives and extracts containing molecules with significant intrinsic fluorescence or the ability to quench the fluorescent tracer were removed from further consideration. The hits from the primary screen were then tested for inhibition of the interaction between the DNA binding domain of Gal4 and a fluorescently labeled DNA oligonucleotide in order to eliminate extracts that may interfere nonspecifically with the binding of transcriptional activators to DNA. Furthermore, this counter screen also likely eliminated extracts that non-specifically interfere with electrostatic interactions, as electrostatics play a central role in protein-DNA interactions, which was advantageous given the previously demonstrated importance of electrostatic contacts in activator•AcID interactions. This counter screen eliminated only a small number of extracts from further consideration, indicating that they may be selective for the AcID motif.

The primary screen hits were then screened against the interaction between the TAD of the transcription factor MLL and the KIX domain of CBP/p300 in order to remove extracts containing natural products that non-specifically inhibit activator•coactivator interactions or mimic amphipathic TADs, resulting in extracts that are selective for the AcID motif and likely exert their activity through allosteric mechanisms of action. Additional selectivity experiments will be completed with the final natural products in order to further confirm their selectivity against the AcID motif. In particular, testing the inhibition against other ERM•coactivator interactions will demonstrate that the molecules are not functioning as mimics of the ERM TAD.

As a final filter, overly promiscuous extracts, specifically those that were active in greater than ten screens of the NPE library, were also precluded from further analysis in an effort to identify selective inhibitors of the ERM•AcID interaction. Following these filtering steps, we were ultimately left with 332 extracts that were identified as inhibitors of the ERM•AcID FP assay. This large number of confirmed hits, which were selected using relatively stringent criteria, further underscores the hypothesis that AcID may be a significantly more targetable coactivator domain than is typical.

In collaboration with the laboratory of Prof. David Sherman at the University of Michigan, thirty-one strains primarily belonging to the bacterial genus *Actinomyces* were regrown and freshly extracted material was prepared from each. These strains were

selected from the list of 332 hits on the basis of genetic diversity by selecting extracts from across the NPE library collected from different geographic locales, a range of inhibitory activities against the ERM•AcID interaction in order to identify inhibitors with distinct mechanisms of action, and the availability of pure spore stocks of the strain. The freshly prepared extracts were then tested for their ability to inhibit the ERM•AcID FP assay, with all but one extract demonstrating strong inhibition of the interaction, suggesting a low false positive hit rate for the extracts identified in the screen. Subsequently, the extracts were tested for the ability to inhibit the activity of an AcID-dependent reporter assay in order to identify extracts with active cell-permeable natural products capable of inhibiting AcID-dependent interactions in a cellular environment. Strain 91085R demonstrated excellent inhibition of the ERM•AcID FP assay and the AcID-dependent reporter assay, resulting in the selection of the extracts from this strain for initial fractionation and structural elucidation of the active natural products. These natural products will then be further validated and employed as mechanistic probes to further study the role of the PEA3 subfamily in malignant process such as metastasis and tumorigenesis, as discussed in the Future Directions section of Chapter 5.

4.5 Materials and Methods

Plasmids

Plasmid pET21b-Med25(394-543)-His_6, henceforth referred to as pAcID-His_6, was a generous gift from Patrick Cramer.[33] Plasmids encoding mutant AcID variants used in the mutational analysis experiments described in Figure 4.6 were described previously in Chapter 3. pGL3-RARE-luc was purchased from Addgene and was as previously described.[56] pCMV-β-Gal and pBSSK (non-coding plasmid) were kind gifts from Jorge Iñiguez-Lluhí.

Protein expression and purification

AcID ($Med25_{394-543}$) protein was expressed and purified together with Paul A. Bruno. Plasmid pAcID-His_6 was transformed into heat-shock competent Rosetta pLysS cells (Novagen), streaked onto LB Agar plates containing ampicillin and chloramphenicol, and incubated at 37 °C overnight. The next evening, a 25 mL Terrific Broth (TB) starter culture with 0.1 mg/mL ampicillin and 0.034 mg/mL chloramphenicol was then inoculated with a colony selected from the LB Agar plate and incubated at 37 °C overnight. The next morning, 5 mL from the starter culture was added to 1L TB containing ampicillin and bacteria were grown at 37 °C to an OD_{600} of 0.8. Temperature was reduced to 18 °C and protein expression was induced upon addition of IPTG to a final concentration of 0.5 mM. Cells were incubated overnight at 18 °C. The 1 L cultures were then collected and centrifuged at 6000xg for 20 mins at 4 °C. Cell pellets were stored at -80 °C prior to purification. The harvested pellet was thawed on ice and resuspended in 20 mL of lysis buffer (50 mM phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 6.8). Cells were then lysed by sonication on ice and cellular lysates were cleared by centrifugation at 9500 rpm for 20 min at 4 °C. The supernatant lysate was then added to Ni-NTA beads (Qiagen) and incubated for 1 hour at 4 °C. The resin was pelleted by centrifugation at 2500 rpm for 2 min at 4 °C and washed with wash buffer (50 mM phosphate, 300 mM sodium chloride, 30 mM imidazole, pH 6.8) a total of five times. Protein was then eluted with 2 mL of elution buffer (50 mM phosphate, 300 mM sodium chloride, 400 mM imidazole, pH 6.8) a total of three times. Eluent was then pooled and purified by cation exchange FPLC (Source 15S, GE Healthcare) using a gradient of Buffer B (50 mM phosphate, 100 mM NaCl, 1 mM

DTT, pH 6.8) in Buffer A (50 mM phosphate, 1 mM DTT). The FPLC purified protein was then dialyzed into storage buffer (10 mM phosphate, 50 mM NaCl, 10% v/v glycerol, 0.001% v/v NP-40, pH 6.8) overnight, concentrated, aliquoted, and stored at -20 °C. Final protein was greater than 90% pure as determined by coomassie stained polyacrylamide gel. Protein concentration was determined by UV-Vis spectroscopy using an extinction coefficient, ε = 22,460 $M^{-1}cm^{-1}$.

Expression and purification of the AcID mutants used in Figure 4.6 was described previously in Chapter 3.

Peptides

All peptides used in the experiments presented within this chapter have been described previously in Chapter 2.

Fluorescence polarization dose-response inhibition assays

Inhibition assays were performed in triplicate with a final sample volume of 20 µL in a low volume, non-binding, 384-well black plate (Corning). A complex of the indicated fluorescent tracer and protein was prepared at a 2x concentration such that 50% of the tracer was bound following dilution onto the assay plate. Small molecule inhibitors were diluted in assay buffer (10 mM PBS, 100 mM NaCl, 10 % glycerol, 0.001% NP-40 pH 6.8) to the desired concentration and serially diluted two-fold on the assay plate to a final volume of 10 µL. 10 µL of the pre-formed fluorescent tracer-protein complex was then added to each well for a final volume of 20 µL. An additional well containing tracer only was prepared and used to determine optimal gain settings on the plate reader. Samples were incubated for thirty minutes at room temperature before fluorescence polarization was measured on a Pherastar plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm. Polarization values were converted to relative fraction bound and plotted opposite the log of inhibitor concentration using GraphPad Prism 5 and curves were fit with a non-linear regression using the built-in equation “log(inhibitor) vs response – variable slope” from which the IC_{50} value was calculated.

Fluorescence polarization direct binding assays

Direct binding assays were performed in triplicate with a final sample volume of 20 µL in a low volume, non-binding, 384-well black plate (Corning). FITC-labeled peptides were diluted in assay buffer (10 mM PBS, 100 mM NaCl, 10 % glycerol, 0.001% NP-40 pH 6.8) to a centration of 40 nM. 10 µL of AcID protein was serially diluted two-fold on the 384-well plate for the number of data points indicated for each experiment using assay buffer. 10 µL of the diluted fluorescent peptide stock was then added to each well of diluted protein for a final tracer concentration of 20 nM. An additional well containing 10 µL buffer and 10 µL fluorescent peptide was prepared for use as a 'tracer only control' to determine optimal gain settings on the plate reader. Samples were then incubated for 30 minutes at room temperature before fluorescence polarization was measured on a Pherastar plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm. A binding isotherm that accounts for ligand depletion (assuming a 1:1 binding model of peptide to ACID) was fit to the observed polarization values as a function of ACID to obtain the apparent equilibrium dissociation, K_d:

$$y = c + (b - c) \times \frac{(K_d + a + x) - \sqrt{(K_d + a + x)^2 - 4ax}}{2a}$$

Where "a" and "x" are the total concentrations of fluorescent peptide and AcID, respectively, "y" is the observed anisotropy at a given AcID concentration, "b" is the maximum observed anisotropy value, and "c" is the minimum observed anisotropy value. Each data point is an average of three independent experiments with the indicated error representing the standard deviation of the three replicates. All curves and calculations were generated using GraphPad Prism 5.

Primary screen of NPE library against ERM•AcID

The primary screen was completed together with Paul A. Bruno at the Center for Chemical Genomics (University of Michigan) in collaboration with Martha Larsen and Steve Vander

Roest. Assays were performed in a final volume of 20 µL in a low volume, non-binding, black 384- well plate (Corning) and read by plate reader (Pherastar) with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicularly polarized 535 nm filters. Final concentration of AcID protein was 850 nM, final concentration of Flo-ERM(38-68) was 20 nM, and 200 nL of NPEs were assayed with a final DMSO concentration of 1% v/v. 10 µL of AcID protein at a concentration of 1.7 µM was added to columns 1-22 of the assay plate by Multidrop dispenser (Thermo Fisher Scientific). Next, compounds were added to columns 3-22 of the assay plate and DMSO was added to columns 1-2 (negative control, AcID•ERM complex) and 23-24 (positive control, VP16 tracer only) by pin tool. Finally, Flo-ERM(38-68) was added to all wells at a concentration of 40 nM. Plates were incubated for thirty minutes at room temperature and read by plate reader as described above with gain settings determined based on a well from columns 23-24 (tracer only). Data was published to and analyzed using MScreen (http://mscreen.lsi.umich.edu).

Primary screen hit filtering

Assays to validate extracts identified in the primary screen and initial hit filtering was completed with Paul A. Bruno at the Center for Chemical Genomics (University of Michigan). Hits selected for further study following the primary screen were assayed in triplicate using a slightly modified protocol from the primary screen. Assays were performed in a final volume of 20 µL in a low volume, non-binding, black 384- well plate (Corning) and read by plate reader (Pherastar) with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicularly polarized 535 nm filters. Final concentration of AcID protein was 850 nM, final concentration of Flo-ERM(38-68) was 20 nM, and 200 nL of NPEs were assayed with a final DMSO concentration of 1% v/v. Rather than adding compounds to the plate by pin tool as described previously, 200 nL of the desired extracts were pre-plated with the cherry-picker on the assay plates in columns 3-22, while 1-2 and 23-24 contained 200 nL of DMSO. 5 µL of buffer was added to all 24 columns and the plate was read to assess the intrinsic fluorescence of compounds within the extracts. Subsequently, 5 µL of Flo-ERM(38-68) at a concentration of 80 nM was added to all 24 columns. The plate was incubated for 30 minutes at room

temperature and the polarization of the plate was read to assess the ability of the extracts to quench the fluorescence of the tracer. Finally, 10 µL of AcID protein at a concentration of 1.7 µM was added to columns 1-22, while 10 µL of buffer was added to columns 23-24. Plates were incubated an additional thirty minutes at room temperature and polarization was assessed as described previously. Data was published to and analyzed using MScreen.

Gal4-DNA direct binding assay

The Gal4-DNA direct binding assay was completed following the direct binding protocol described above and as previously reported.[42]

Gal4(1-100)•DNA counter screen

The Gal4 counter screen was completed with Paul A. Bruno at the Center for Chemical Genomics (University of Michigan). Assays were performed in a final volume of 20 µL in a low volume, non-binding, black 384- well plate (Corning) and read by plate reader (Pherastar) with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicularly polarized 535 nm filters. Final concentration of Gal4(1-100) protein was 50 nM, final concentration of Flo-DNA(TCCGGAGGACTGTCCTCCGG) was 10 nM, and 200 nL of NPEs were assayed with a final DMSO concentration of 1% v/v. Rather than adding compounds to the plate by pin tool as described previously, 200 nL of the desired extracts were pre-plated with the cherry-picker on the assay plates in columns 3-22, while 1-2 and 23-24 contained 200 nL of DMSO. 10 µL of 100 nM Gal4(1-100) was added to columns 1-22 and 10 µL of buffer was added to columns 23-24. Subsequently, 10 µL of 20 nM Flo-DNA was added to all 24 columns and plates were incubated at room temperature for thirty minutes prior to measuring the polarization of the samples. Data was published to and analyzed using MScreen.

KIX direct binding assays

The MLL•KIX and pKID•KIX direct binding assays were completed following the direct binding protocol described above and as previously reported.[46]

MLL•KIX counter screen

The MLL•KIX counter screen was completed with Paul A. Bruno at the Center for Chemical Genomics (University of Michigan). Assays were performed in a final volume of 20 µL in a low volume, non-binding, black 384- well plate (Corning) and read by plate reader (Pherastar) with polarized excitation at 485 nm and emission intensity measured through parallel and perpendicularly polarized 535 nm filters. Final concentration of KIX protein was 10 µM, final concentration of Flo-MLL was 25 nM, and 200 nL of NPEs were assayed with a final DMSO concentration of 1% v/v. Rather than adding compounds to the plate by pin tool as described previously, 200 nL of the desired extracts were pre-plated with the cherry-picker on the assay plates in columns 3-22, while 1-2 and 23-24 contained 200 nL of DMSO. 10 µL of KIX protein at a concentration of 20 µM was added to columns 1-22, while 10 µL of buffer was added to columns 23-24. Subsequently, 10 µL of Flo-MLL at a concentration of 50 nM to all 24 columns and plates were incubated at room temperature for thirty minutes prior to measuring the polarization of the samples. Data was published to and analyzed using MScreen.

Regrowth of strains identified in screen

Strains were regrown by Matthew S. Beyersdorf (Mapp/Sherman Labs, University of Michigan). Strains were selected for regrowth following consultation with Pamela Schultz (University of Michigan) and completed following standard procedures.

Retesting extracts from regrown strain

Fresh extracts were prepared from the regrown strain following standard procedures by Matthew S. Beyersdorf. The freshly prepared extracts were then tested against the ERM•AcID FP assay following conditions analogous to the conditions used in the primary screen (20 µL total volume, 850 nM AcID, 20 nM ERM tracer) at concentrations of 0.3 mg/mL.

RARα luciferase reporter assay

The RARα luciferase reporter assay was performed as described in **Chapter 3** by Paul A. Bruno. Natural products were assayed at a final concentration of 0.3 mg/mL, co-dosed with retinoic acid at a final concentration of 1 μM.

HPLC fractionation of NPEs

NPE fractionation was completed by Matthew S. Beyersdorf. The crude extract of strain 91085R was fractionated by preparative HPLC on a C_{18} column using a gradient of 10% to 30% acetonitrile in water over fifteen minutes followed by a gradient of 30% to 45% acetonitrile in water over thirty minutes. Active material was then further sub-fractionated on a preparative C_{18} column using a gradient of 15% to 17.5% acetonitrile in water over twenty minutes followed by a gradient of 17.5% to 35% acetonitrile in water over ten minutes. Isolated fractions were then lyophilized to a fine powder, resuspended in DMSO at a concentration of 0.75 mg/mL, and assayed for inhibition of the ERM•AcID FP assay as described above.

4.6 References

1. Oikawa, T. ETS transcription factors: possible targets for cancer therapy. *Cancer Sci.* **95,** 626–633 (2004).

2. Gilliland, D. G. The diverse role of the ETS family of transcription factors in cancer. *Clin. Cancer Res.* **7,** 451–453 (2001).

3. Tomlins, S. A. *et al.* Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia* **10,** 177–188 (2008).

4. Kita, D. *et al.* Expression of dominant-negative form of Ets-1 suppresses fibronectin-stimulated cell adhesion and migration through down-regulation of integrin alpha5 expression in U251 glioma cell line. *Cancer Res.* **61,** 7985–7991 (2001).

5. Aytes, A. *et al.* ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer. *Proceedings of the National Academy of Sciences* **110,** E3506–15 (2013).

6. Pellecchia, A. *et al.* Overexpression of ETV4 is oncogenic in prostate cells through promotion of both cell proliferation and epithelial to mesenchymal transition. *Oncogenesis* **1,** e20 (2012).

7. Hollenhorst, P. C., Paul, L., Ferris, M. W. & Graves, B. J. The ETS gene ETV4 is required for anchorage-independent growth and a cell proliferation gene expression program in PC3 prostate cells. *Genes Cancer* **1,** 1044–1052 (2011).

8. Benz, C. C. *et al.* HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene* **15,** 1513–1525 (1997).

9. de Launoit, Y. *et al.* The PEA3 group of ETS-related transcription factors. Role in breast cancer metastasis. *Adv. Exp. Med. Biol.* **480,** 107–116 (2000).

10. Mehlen, P. & Puisieux, A. Metastasis: a question of life or death. *Nature Reviews Cancer* **6,** 449–458 (2006).

11. Firlej, V. *et al.* Reduced tumorigenesis in mouse mammary cancer cells following inhibition of Pea3- or Erm-dependent transcription. *J. Cell. Sci.* **121,** 3393–3402 (2008).

12. de Launoit, Y. *et al.* The Ets transcription factors of the PEA3 group: transcriptional regulators in metastasis. *Biochim. Biophys. Acta* **1766,** 79–87 (2006).

13. Verger, A. *et al.* The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41,** 4847–4859 (2013).

14. Fauquette, V. *et al.* The antagonistic regulation of human MUC4 and ErbB-2 genes by the Ets protein PEA3 in pancreatic cancer cells: implications for the proliferation/differentiation balance in the cells. *Biochem. J.* **386,** 35–45 (2005).

15. Landrieu, I. *et al.* Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **43,** 7110–7121 (2015).

16. Myers, E. *et al.* A positive role for PEA3 in HER2-mediated breast tumour progression. *Br. J. Cancer* **95,** 1404–1409 (2006).

17. Kinoshita, J. *et al.* Clinical significance of PEA3 in human breast cancer. *Surgery* **131,** S222–5 (2002).

18. Oikawa, T. & Yamada, T. Molecular biology of the Ets family of transcription factors. *Gene* **303,** 11–34 (2003).

19. Fauquette, V. *et al.* The antagonistic regulation of human MUC4 and ErbB-2 genes by the Ets protein PEA3 in pancreatic cancer cells: implications for the proliferation/differentiation balance in the cells. *Biochem. J.* **386,** 35–45 (2005).

20. Erkizan, H. V. *et al.* A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nature medicine* **15,** 750–756 (2009).

21. Wei, G.-H. *et al.* Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* **29,** 2147–2160 (2010).

22. Rahim, S. *et al.* YK-4-279 inhibits ERG and ETV1 mediated prostate cancer cell invasion. *PLoS ONE* **6,** e19343 (2011).

23. Rahim, S. *et al.* A small molecule inhibitor of ETV1, YK-4-279, prevents prostate cancer growth and metastasis in a mouse xenograft model. *PLoS ONE* **9,** e114260 (2014).

24. Pop, M. S. *et al.* A small molecule that binds and inhibits the ETV1 transcription factor oncoprotein. *Mol. Cancer Ther.* **13,** 1492–1502 (2014).

25. Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **68,** 803–812 (2007).

26. Larsen, M. J. *et al.* The role of HTS in drug discovery at the University of Michigan. *Comb. Chem. High Throughput Screen.* **17,** 210–230 (2014).

27. Raveh, A. *et al.* Discovery of potent broad spectrum antivirals derived from marine actinobacteria. *PLoS ONE* **8,** e82318 (2013).

28. Majmudar, C. Y. *et al.* Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300. **51,** 11258–11262 (2012).

29. Tripathi, A. *et al.* Baulamycins A and B, broad-spectrum antibiotics identified as inhibitors of siderophore biosynthesis in Staphylococcus aureus and Bacillus anthracis. *J. Am. Chem. Soc.* **136,** 1579–1586 (2014).

30. Cesa, L. C., Mapp, A. K. & Gestwicki, J. E. Direct and Propagated Effects of Small Molecules on Protein-Protein Interaction Networks. *Front Bioeng Biotechnol* **3,** 119 (2015).

31. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chemical Biology* **7,** 1311–1320 (2012).

32. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **18,** 410–415 (2011).

33. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18,** 404–409 (2011).

34. Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *Journal of Structural Biology* **174,** 245–251 (2011).

35. Zhang, J., Chung, T. & Oldenburg, K. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **4,** 67–73 (1999).

36. Fischer, E. SYNTHESIS OF DEPSIDES, LICHEN-SUBSTANCES AND TANNINS. *J. Am. Chem. Soc.* **36,** 1170–1201 (1914).

37. Parrot, D., Jan, S., Baert, N., Guyot, S. & Tomasi, S. Comparative metabolite profiling and chemical study of Ramalina siliquosa complex using LC-ESI-MS/MS approach. *Phytochemistry* **89,** 114–124 (2013).

38. Goel, M., Dureja, P., Rani, A., Uniyal, P. L. & Laatsch, H. Isolation, characterization and antifungal activity of major constituents of the Himalayan lichen Parmelia reticulata Tayl. *J. Agric. Food Chem* **59,** 2299–2307 (2011).

39. Behera, B. C., Mahadik, N. & Morey, M. Antioxidative and cardiovascular-protective activities of metabolite usnic acid and psoromic acid produced by lichen species Usnea complanata under submerged fermentation. *Pharm Biol* **50,** 968–979 (2012).

40. Sultana, N. & Afolayan, A. J. A new depsidone and antibacterial activities of compounds from Usnea undulata Stirton. *J Asian Nat Prod Res* **13,** 1158–1164 (2011).

41. Lea, W. A. & Simeonov, A. Fluorescence polarization assays in small molecule screening. *Expert Opin Drug Discov* **6,** 17–32 (2011).

42. Wands, A. M. *et al.* Transient-state kinetic analysis of transcriptional activator·DNA complexes interacting with a key coactivator. *Journal of Biological Chemistry* **286,** 16238–16245 (2011).

43. Ohlendorf, D. H. & Matthew, J. B. Electrostatics and flexibility in protein-DNA interactions. *Adv. Biophys.* **20,** 137–151 (1985).

44. Parker, D. *et al.* Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular cell* **2,** 353–359 (1998).

45. Wang, N. *et al.* Ordering a dynamic protein via a small-molecule stabilizer. *J. Am. Chem. Soc.* **135,** 3363–3366 (2013).

46. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. *Proceedings of the National Academy of Sciences* **111,** 12061–12066 (2014).

47. Kleigrewe, K. *et al.* Combining Mass Spectrometric Metabolic Profiling with Genomic Analysis: A Powerful Approach for Discovering Natural Products from Cyanobacteria. *J. Nat. Prod.* **78,** 1671–1682 (2015).

48. Subramani, R. & Aalbersberg, W. Marine actinomycetes: an ongoing source of novel bioactive metabolites. *Microbiol. Res.* **167,** 571–580 (2012).

49. Fribley, A. M. *et al.* Complementary cell-based high-throughput screens identify novel modulators of the unfolded protein response. *Journal of Biomolecular Screening* **16,** 825–835 (2011).

50. Walter, G. M. *et al.* High-throughput screen of natural product extracts in a yeast model of polyglutamine proteotoxicity. *Chem Biol Drug Des* **83,** 440–449 (2014).

51. Lee, H.-K. H., Park, U.-H. U., Kim, E.-J. E. & Um, S.-J. S. MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* **26,** 3545–3557 (2007).

52. Poss, Z. C., Ebmeier, C. C. & Taatjes, D. J. The Mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* **48,** 575–608 (2013).

53. Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85,** 403–414 (1996).

54. Torchia, J. *et al.* The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387,** 677–684 (1997).

55. Kawasaki, H. *et al.* Distinct roles of the co-activators p300 and CBP in retinoic-acid-induced F9-cell differentiation. *Nature* **393,** 284–289 (1998).

56. Hoffman, L. M. *et al.* BMP action in skeletogenesis involves attenuation of retinoid signaling. *J. Cell Biol.* **174,** 101–113 (2006).

CHAPTER 5
Conclusions and Future Directions

5.1 Conclusions

The guiding hypothesis for the work described in this thesis was that we could identify selective small molecule inhibitors that perturbed the interactions of various transcriptional activators with the Activator Interaction Domain (AcID) of Med25. This hypothesis was made on the basis that AcID is comprised of a unique fold not previously identified in any known coactivator proteins and found in only one other protein of unknown function. Towards this goal, we first explored the molecular underpinnings that define the interaction of transcriptional activators with this unique coactivator domain in order to determine whether activators interact with AcID in a unique fashion from typical activator•coactivator interactions and to define features that may be effectively exploited in an inhibitor discovery strategy. We next completed a small high-throughput screen of known biologically active molecules and natural products and identified several molecules belonging to the depside and depsidone classes of natural products. Experiments validating these molecules as AcID-targeted inhibitors and preliminary elucidation of their mechanism of action were also completed. We finally described a screen of the full Natural Product Extracts Library held by the Center for Chemical Genomics at the University of Michigan against the interaction of a family of transcriptional activators linked to metastatic processes in cancer and the AcID motif in order to identify potent small molecule inhibitors that will be useful as mechanistic probes in exploring these processes.

In order to better define the molecular underpinnings of activator interactions with Med25 AcID, we began with a serial truncation of the VP16 TAD that ultimately revealed that the majority of the binding affinity could be recapitulated solely by the purported α-helices within the TAD, consistent with most other activator•coactivator interactions. This finding was interesting as the defining structural feature of the AcID motif is a β-barrel, which is highly uncommon for activator-targeted coactivator domains. The importance of

α-helical secondary structure in the TADs of activator binding partners was further supported by the ERM•AcID interaction, suggesting that this may be a common phenomenon in all AcID binding partners.

We next completed a hotspot analysis using alanine scanning mutagenesis studies of the putative α-helices within the VP16 TAD in order to determine whether activator•AcID interactions require a few critical contacts, or broad interactions over large surfaces of the domain. The alanine scanning mutagenesis studies failed to identify specific hotspot residues, suggesting that VP16•AcID interactions rely upon a broad interaction surface wherein each residue contributes weakly to the affinity of the domain. Thus, small molecule inhibitors capable of inducing allosteric changes to the domain at surfaces required for activator binding were expected to offer the greatest opportunity for an effective inhibition strategy.

The highly acidic nature of most activator binding partners for AcID and the highly basic character of the domain itself suggested that electrostatic contacts may be of particular importance in activator•AcID interactions. Determination of the binding affinity of a VP16 derived peptide in the presence of various concentrations of three different salts demonstrated that the affinity for AcID decreased with increasing salt concentration, supporting the hypothesis that electrostatic contacts are critical to activator•AcID interactions. This hypothesis was further supported by a mutation within the AcID motif that caused a charge inversion near a surface important for the binding of transcriptional activators that resulted in a significant decrease in affinity of the VP16 TAD for the domain. Interestingly, this mutation affected affinity of peptide ligands for both of the activator binding sites on the AcID motif, suggesting that this residue is not merely an electrostatic contact, but that the charge inversion may also result in structural changes to the secondary or tertiary structure of the domain. This observation thus supports the hypothesis that AcID motif is relatively plastic and capable of a variety of unique conformations.

We next developed a fluorescence polarization based assay of the interaction between the VP16 TAD and AcID and optimized the assay for screening applications in a high-throughput format. This assay was then screened against a small library rich in known bioactive compounds, natural products, and FDA approved drugs, leading to the

identification of a number of potential lead molecules belonging to the depside and depsidone classes of small molecules. This observation was particularly significant as a previous screening project to identify inhibitors of an α-helix dependent activator•coactivator interaction completed in our laboratory also identified molecules belonging to these classes as effective inhibitors, suggesting that perhaps these core scaffolds may be privileged for the inhibition of α-helix dependent activator•coactivator interactions. The depsidones Norstictic Acid and Psoromic Acid were selected as lead molecules for further study based upon their activity in competitive inhibition assays and their structural similarities.

These molecules were then confirmed as inhibitors of the VP16•AcID interaction by obtaining samples from commercial sources and testing their ability to inhibit the VP16 interaction. Psoromic acid and norstictic acid were then tested for their ability to inhibit two other activator•coactivator interactions demonstrating that they were selective for AcID-dependent interactions. One of the interactions assayed was an alternative VP16•coactivator interaction, suggesting that they do not function merely as VP16 mimetics and are in fact more selective for the AcID domain than one of its native binding partners. Subsequent studies, namely molecular dynamics simulations and inhibition studies in the presence of various salt concentrations, indicate that Norstictic and Psoromic Acid are localized to the AcID domain through electrostatic interactions between the acidic functionalities harbored by the molecules and positively charged surfaces of the domain.

The presence of an aldehyde on all of the lead molecules identified in the primary screen led us to question whether they may be covalent inhibitors of AcID interactions. Subsequent mass spectrometric analyses of reduced covalent adducts between the AcID domain and small molecules indicated that they are in fact covalent and that this attachment occurs through the formation of a Schiff base between the aldehyde and nucleophilic lysine residues on the surface of the AcID domain. In order to identify potential sites of covalent labeling, we completed $^{1}H,^{15}N$-HSQC chemical shift perturbation studies of the AcID domain in complex with norstictic acid. The strongest chemical shifts were induced in regions required for the binding of the VP16 TAD to the AcID domain, suggesting that Norstictic Acid directly perturbs surfaces required for

interaction with VP16. Furthermore, a subset of the chemical shifts occurred at residues at the interface of the two activator binding sites on the AcID motif, which suggests that norstictic acid may be capable of inducing allosteric changes in the domain that affects activator binding at both sites

Subsequent mutational analysis studies in which lysine residues in close proximity to critical binding surfaces were mutated to arginine were undertaken in order to validate the role that they may play in the inhibitory activity of norstictic acid. These mutations did not alter the affinity of VP16 for the domain, but did significantly alter the inhibitory potency of norstictic acid and resulted in decreased labeling of the domain by the inhibitor, suggesting that these residues are involved in the covalent attachment of Norstictic Acid to the domain. Additionally, lysine to arginine mutations in the H1 or H2 binding sites individually was able to perturb the activity of norstictic acid against interactions at both sites, supporting the hypothesis that the two sites may be in allosteric communication. Taken together, these experiments suggest that norstictic acid may exert its inhibitory effects by interfering with critical electrostatic contacts between the domain and the TADs of activator binding partners by relieving the positive charge of specific lysine residues through the formation of an imine. The covalent attachment of norstictic acid to these lysine residues additionally adds steric bulk to regions important for activator binding, providing an additional mechanism by which norstictic acid inhibits the binding of activators to AcID.

Finally, we tested the ability of norstictic acid and psoromic acid to perturb AcID-dependent transcriptional processes within a cellular context. RT-qPCR analysis of the expression of a canonical target gene of an activator binding partner of AcID, ATF6α, revealed that expression of the gene was inhibited by treatment with norstictic acid and psoromic acid, likely through inhibition of the interaction between the activator and the AcID motif. Similarly, norstictic acid and psoromic acid are capable of inhibiting the transcriptional activity of RARα in a reporter assay, likely by inhibiting the recruitment of the histone acetyl transferase CBP through its interaction with the AcID domain. These results are consistent with the hypothesis that Norstictic Acid and Psoromic Acid are cell permeable small molecules that can perturb AcID-dependent transcriptional processes.

Following the observation that AcID-dependent interactions could be effectively targeted by small molecule inhibitors, we next sought to identify novel natural products capable of inhibiting the interaction between the transcriptional activation domains of the PEA3 subfamily of ETS transcription factors and the AcID domain of Med25. The discovery of such molecules was an attractive goal as these inhibitors will be useful molecular probes for the elucidation of the role of the PEA3 subfamily in cellular processes related to cancer progression, specifically metastasis and possibly tumorigenesis, and could potentially serve as lead molecules in the development of novel therapeutic strategies. Though norstictic acid and psoromic acid were identified as covalent inhibitors of AcID-dependent activator•coactivator interactions, we sought to identify novel inhibitors with the hypothesis that alternative non-covalent inhibitors of the AcID domain may allow for greater selectivity in a complex cellular environment and may be capable of inhibiting activator interactions with the domain through an alternative mechanism of action.

Towards this goal, we first tested the ability of norstictic acid and psoromic acid to disrupt the interaction between the TAD of the PEA3 subfamily member ERM and AcID in order to demonstrate that the interaction was amenable to inhibition and found that both molecules were capable of perturbing the interaction. We then optimized an FP assay based on the ERM•AcID interaction for application in a high-throughput screen. Prior to screening the full NPE library, the optimized assay was tested against one sample plate of the library rich in natural products derived from lichens. This plate produced a number of hits, likely due to the presence of multiple extracts abundant in depsides and depsidones, indicating that the assay was well adapted to screening of the full library in a high-throughput format.

Subsequently, a primary screen of the NPE library against the optimized assay resulted in 1,602 hits that were selected for further filtering steps. These initial hits were tested in triplicate to remove false positives and extracts containing molecules with significant intrinsic fluorescence or the ability to quench the fluorescent tracer were removed from further consideration. The hits from the primary screen were then tested for inhibition of alternative interactions, namely the interaction between the DBD of Gal4 and DNA, as well as the interaction between the KIX domain of CBP/p300 and MLL in an

effort to remove non-specific inhibitors from further analysis. As a final filter overly promiscuous extracts, specifically those that were active in greater than ten screens of the NPE library, were also precluded from further analysis in an effort to identify selective inhibitors of the ERM•AcID interaction. Following these filtering steps, we were ultimately left with 332 extracts that were validated inhibitors of the ERM•AcID FP assay, suggesting that AcID may be a significantly more targetable coactivator domain than is typical.

In collaboration with the laboratory of Prof. David Sherman at the University of Michigan, thirty-one strains from the list of 332 validated hits were regrown and fresh extracts prepared on the basis of genetic diversity, a range of inhibitory activities against the ERM•AcID interaction, and the availability of pure spore stocks of the strain. The freshly prepared extracts were then tested for their ability to inhibit the ERM•AcID FP assay, with all but one extract demonstrating strong inhibition of the interaction, suggesting a low false positive hit rate for the extracts identified in the screen. Subsequently, the extracts were tested for the ability to inhibit the activity of an AcID-dependent reporter assay in order to identify extracts with active cell-permeable natural products capable of inhibiting AcID-dependent interactions in a cellular environment. Strain 91085R demonstrated excellent inhibition of the ERM•AcID FP assay and the AcID-dependent reporter assay, resulting in the extracts from this strain being the first selected for fractionation and structural elucidation of the active natural products.

The extract from strain 91085R has undergone multiple fractionation step by preparative HPLC, resulting in the isolation of three compounds within the extract that are capable of inhibiting the ERM•AcID interaction *in vitro*. This fractionated material is currently undergoing final purification steps by HPLC prior to elucidating their structures using various mass spectrometric, NMR, and crystallographic techniques.

Through the experiments summarized above and described in this thesis, we have demonstrated that the Activator Interaction Domain of Med25 can be effectively targeted by small molecule inhibitors, despite the fact that activator binding partners bind to the domain over broad surfaces with moderate affinity. Furthermore, the evidence presented suggests that this inhibition can be achieved with selectivity, likely as a result of the unique protein fold that comprises the domain and the ability of the inhibitors to induce allosteric

changes to the structure of the domain due to the inherent conformation plasticity of the motif.

5.2 Future Directions

Using Natural Products as Molecular Probes for PEA3 Dependent Transcriptional Processes

Following the final deconvolution, purification, and structural elucidation of the active natural products within the extract of 91085R described in Chapter 4 we will use those molecules as molecular probes to gain a better understanding of PEA3 dependent transcriptional processes and their functional consequences. Initially, the affinity of these molecules and their binding mode will be assessed. Full inhibition curves will be generated against the ERM(38-68)•AcID FP assay in order to determine IC_{50} and K_i values, provided the molecules bind reversibly with a 1:1 stoichiometry, of the inhibitor for the AcID motif. Additionally, the K_d of the inhibitors for AcID will be assessed by isothermal titration calorimetry, provided that the affinity is between 100 nM and 10 μM. In the event that the affinity of the molecules for the domain is outside of that range, affinity will be alternatively assessed by transient kinetic analyses or with Protein Observed Fluorine NMR, as previously described. [1-3] In addition to determining the affinity of the molecules for AcID, their selectivity for the domain will also be assessed. In order to accomplish this, full inhibition curves will be generated against FP assays of the MLL•KIX (as validation of the results obtained during counter-screening), pKID•KIX, VP16•Med15, and HIF1-α•CH1 interactions. [4-6] In the event that the isolated natural product does not exhibit at least five-fold selectivity for AcID, an alternative extract will be fractionated and the active component identified.

Subsequently, we will attempt to map the binding site of the inhibitors on the AcID domain using a combination of NMR techniques including $^{1}H,^{15}N$-HSQC chemical shift perturbations, Protein Observed Fluorine (PrOF) NMR spectroscopy of an AcID mutant containing fluorinated amino acids in suspected binding sites, and ^{13}C-CACO NMR spectroscopy. PrOF NMR leverages the incredible sensitivity of fluorine to changes in its chemical environment in order to identify subtle shifts in protein structure, allowing for a thorough accounting of shifts induced by binding of the molecule or allosteric changes in

the structure of the protein.[1] ^{13}C-CACO NMR observes ^{13}C directly, allowing for elucidation of structural changes in highly flexible protein structures that cannot be easily visualized using more conventional techniques such as HSQC.[7] The information regarding putative natural product binding sites on AcID will then be used to generate AcID mutants that perturb the binding of the inhibitor, while leaving PEA3 activator interactions intact. These mutants will be particularly useful in studies designed to demonstrate target engagement within the complex environment of the cell.

Following *in vitro* experiments designed to validate the activity of the natural products against the ERM•AcID interaction and elucidation of the mechanism of action by which this inhibition is achieved, we will next complete a series of experiments designed to demonstrate that the molecules engage the AcID domain within a complex cellular environment. Initially, a reporter assay using a fusion of the ERM TAD to the DBD of Gal4 to drive expression of a luciferase gene regulated by the Gal4 promoter will be developed and the ability of the natural products to inhibit the expression of this reporter gene will be assessed following transient transfection into a suitable host. In addition to the luciferase reporter plasmid, cells will also be transfected with a plasmid that constitutively expresses a β-galactosidase reporter gene. Thus, monitoring the activity of β-galactosidase in treated cells will allow for initial determination of the cytotoxicity of the identified natural products. In addition to a Gal4(DBD)-ERM(TAD) fusion, the ability of the molecule to inhibit analogous reporter assays using the TADs of transcriptional activators such as CREB or c-Jun, which are not dependent upon AcID for their activation, fused to the Gal4 DBD will be completed as an additional metric by which we can assess selectivity of the natural products against the AcID motif. Subsequently, the Gal4-ERM driven reporter assay will also be tested in cells that have been transfected with a plasmid expressing full length Med25 with the previously identified AcID mutations that abrogate inhibition by the natural products. In this case, loss of potency in the inhibition of the assay by the inhibitors would serve as further evidence of target engagement within the cell. Additionally, the ability of the small molecules to inhibit the co-immunoprecipitation (co-IP) of the PEA3 family members (ER81/ETV1, PEA3/ETV4, and ERM/ETV5) and Med25 will be assessed with variable doses as a further assessment of target engagement. Finally, if the natural products are synthetically accessible or contain non-critical reactive

handles, biotinylated variants will be synthesized in order to complete Med25 pull-down experiments.

Following successful demonstration of target engagement within the cell we will next assess the ability of the natural products to perturb the expression of PEA3-dependent genes. A number of genes with distinct functional effects, including the matrix metalloproteinases MMP-1, MMP-2, MMP-3, and MMP-9[8,9], FAK (cell motility and migration) [9,10], CXCR4 (metastasis biomarker)[11], and Cyclin D2 (proliferation and migration)[12] have been identified as PEA3-dependent through a combination of shRNA knockdown studies and 'squelching' experiments in which overexpression of AcID reduces expression of the gene. Thus, the effects of the natural products on the expression of these genes over time will be assessed using RT-qPCR and western blotting at distinct time points. In support of this, a partially purified extract from the 91085R strain has already been tested for its ability to inhibit the expression of MMP-2 and MMP-9 at a concentration of 0.1 mg/mL and demonstrated to significantly attenuate the expression of these genes as shown in Figure 5.1 in a preliminary experiment.

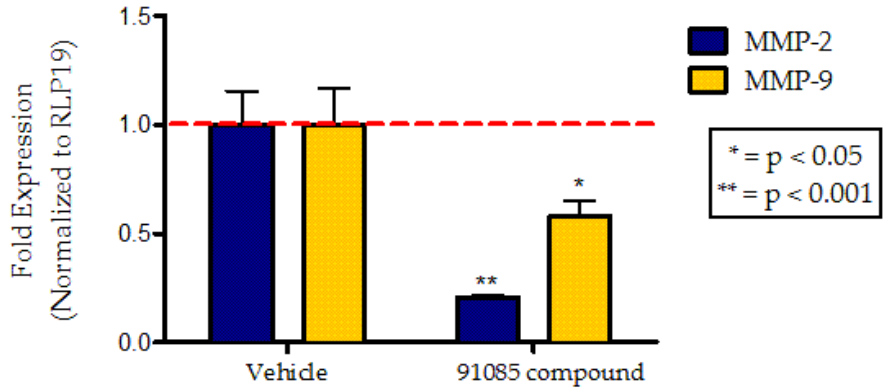


***Figure 5.1*- Inhibition of MMP-2 and MMP-9 Expression by a Partially Purified Extract of Strain 91085R** The expression of MMP-2 and MMP-9 in MDA-MB-231 cells treated with a partially purified extract of strain 91085R at a concentration of 0.1 mg/mL was assessed by RT-qPCR and compared to a control in which cells were treated with ethanol vehicle only. Experiments represent the mean of technical triplicates with error indicating the standard deviation of the fold expression.

Subsequent to analysis by RT-qPCR, analysis of global gene expression by RNA-seq will be considered in order to identify the full subset of PEA3-dependent genes that are attenuated by treatment with the natural product inhibitors.[13]

Finally, we will use the natural products to probe the effects of inhibiting PEA3-dependent transcriptional programs on cellular processes related to the tumorigenesis

cascade, with a particular focus on breast cancer. The PEA3 subfamily has been shown to be critical for cellular proliferation, invasion, and migration in a number of breast cancer model systems. [14-20] Given these functional consequences associated with PEA3-dependent transcription we will complete a number of phenotypic assays including assessing proliferative capacity using MTT, invasive potential using Matrigel invasion assays, cellular migration using wound healing assays, anchorage independent survival, and colony formation assays. These experiments will be completed using various concentrations of the natural products in order to define a dose response. Furthermore, these experiments will be completed using MB-231, SK-BR-3, and MCF-7 cell lines which overexpress, normally express, and underexpress the PEA3 subfamily, respectively. [18,21] Analogous experiments will be additionally be completed in the non-tumorigenic HMEC and MCF-10A cell lines in order to assess off-target effects and cytotoxicity. The effects of PEA3 inhibition on these phenotypic responses will further clarify the importance of PEA3 subfamily members in the tumorigenesis cascade and the conversion to a metastatic phenotype.

Following the successful completion of these experiments, a potential avenue of research that we may investigate will be to test the effects of the natural product PEA3 inhibitors in combination with inhibitors of other cellular pathways linked to PEA3 transcription that are implicated in breast cancer, such as the Her2, Ras, or PI3K signaling axes.[20,22] We have previously demonstrated that combination therapies incorporating transcriptional inhibitors can produce significant enhancement over the effects of either therapy in isolation, as in the case of the transcriptional activator ESX and Her2/EGFR.[23,24]

5.3 References

1. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chemical Biology* **7,** 1345–1350 (2012).

2. Wands, A. M. *et al.* Transient-state kinetic analysis of transcriptional activator·DNA complexes interacting with a key coactivator. *Journal of Biological Chemistry* **286,** 16238–16245 (2011).

3. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. *Proceedings of the National Academy of Sciences* **111,** 12061–12066 (2014).

4. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *Journal of Biological Chemistry* **277,** 43168–43174 (2002).

5. Majmudar, C. Y., Wang, B., Lum, J. K., Håkansson, K. & Mapp, A. K. A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew. Chem. Int. Ed.* **48,** 7021–7024 (2009).

6. Henchey, L. K. *et al.* Inhibition of hypoxia inducible factor 1-transcription coactivator interaction by a hydrogen bond surrogate alpha-helix. *J. Am. Chem. Soc.* **132,** 941–943 (2010).

7. Gray, F. L. V., Murai, M. J., Grembecka, J. & Cierpicki, T. Detection of disordered regions in globular proteins using ^{13}C-detected NMR. *Protein Sci.* **21,** 1954–1960 (2012).

8. Verger, A. *et al.* The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41,** 4847–4859 (2013).

9. Firlej, V. *et al.* Reduced tumorigenesis in mouse mammary cancer cells following inhibition of Pea3- or Erm-dependent transcription. *J. Cell. Sci.* **121,** 3393–3402 (2008).

10. Li, S. *et al.* Requirement of PEA3 for transcriptional activation of FAK gene in tumor metastasis. *PLoS ONE* **8,** e79336 (2013).

11. Gu, S. *et al.* PEA3 activates CXCR4 transcription in MDA-MB-231 and MCF7 breast cancer cells. *Acta Biochim. Biophys. Sin. (Shanghai)* **43,** 771–778 (2011).

12. Ladam, F. *et al.* Loss of a negative feedback loop involving pea3 and cyclin d2 is required for pea3-induced migration in transformed mammary epithelial cells. *Molecular Cancer Research* **11,** 1412–1424 (2013).

13. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5,** 621–628 (2008).

14. Davidson, B. *et al.* The clinical role of the PEA3 transcription factor in ovarian and breast carcinoma in effusions. *Clin. Exp. Metastasis* **21,** 191–199 (2004).

15. Benz, C. C. *et al.* HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene* **15,** 1513–1525 (1997).

16. de Launoit, Y. *et al.* The Ets transcription factors of the PEA3 group: transcriptional regulators in metastasis. *Biochim. Biophys. Acta* **1766,** 79–87 (2006).

17. Myers, E. *et al.* A positive role for PEA3 in HER2-mediated breast tumour progression. *Br. J. Cancer* **95,** 1404–1409 (2006).

18. Bièche, I. *et al.* Expression of PEA3/E1AF/ETV4, an Ets-related transcription factor, in breast tumors: positive links to MMP2, NRG1 and CGB expression. *Carcinogenesis* **25,** 405–411 (2004).

19. Kinoshita, J. *et al.* Clinical significance of PEA3 in human breast cancer. *Surgery* **131,** S222–5 (2002).

20. O'Hagan, R. C. & Hassell, J. A. The PEA3 Ets transcription factor is a downstream target of the HER2/Neu receptor tyrosine kinase. *Oncogene* **16,** 301–310 (1998).

21. Baert, J. L. *et al.* Expression of the PEA3 group of ETS-related transcription factors in human breast-cancer cells. *Int. J. Cancer* **70,** 590–597 (1997).

22. Aytes, A. *et al.* ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer. *Proceedings of the National Academy of Sciences* **110,** E3506–15 (2013).

23. Zhang, M. *et al.* Genetic and chemical targeting of epithelial-restricted with serine box reduces EGF receptor and potentiates the efficacy of afatinib. *Mol. Cancer Ther.* **12,** 1515–1525 (2013).

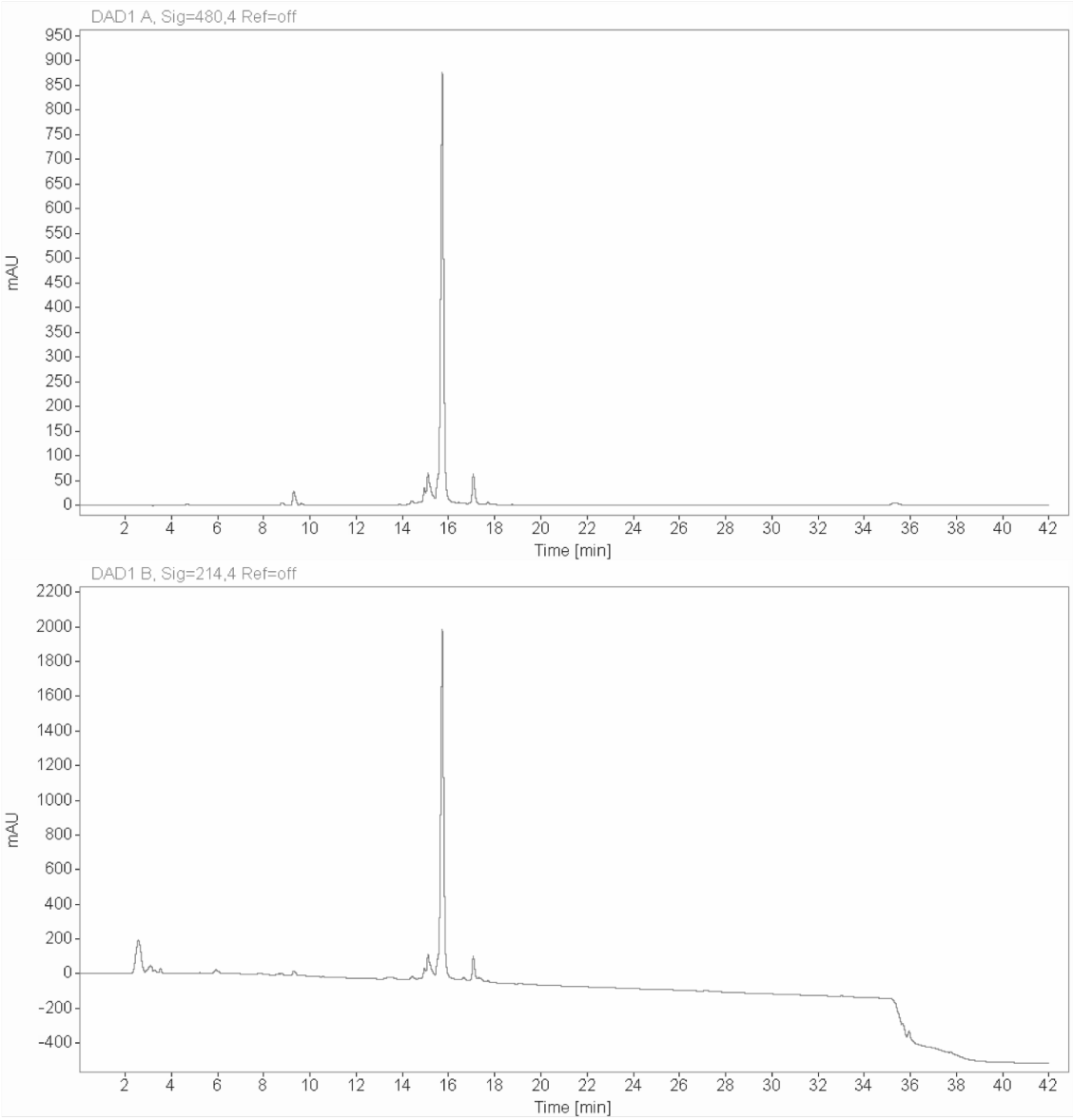
24. Taylor, C. E., Pan, Q. & Mapp, A. K. Synergistic enhancement of the potency and selectivity of small molecule transcriptional inhibitors. *ACS Med. Chem. Lett.* **3,** 30–34 (2012).

APPENDIX A

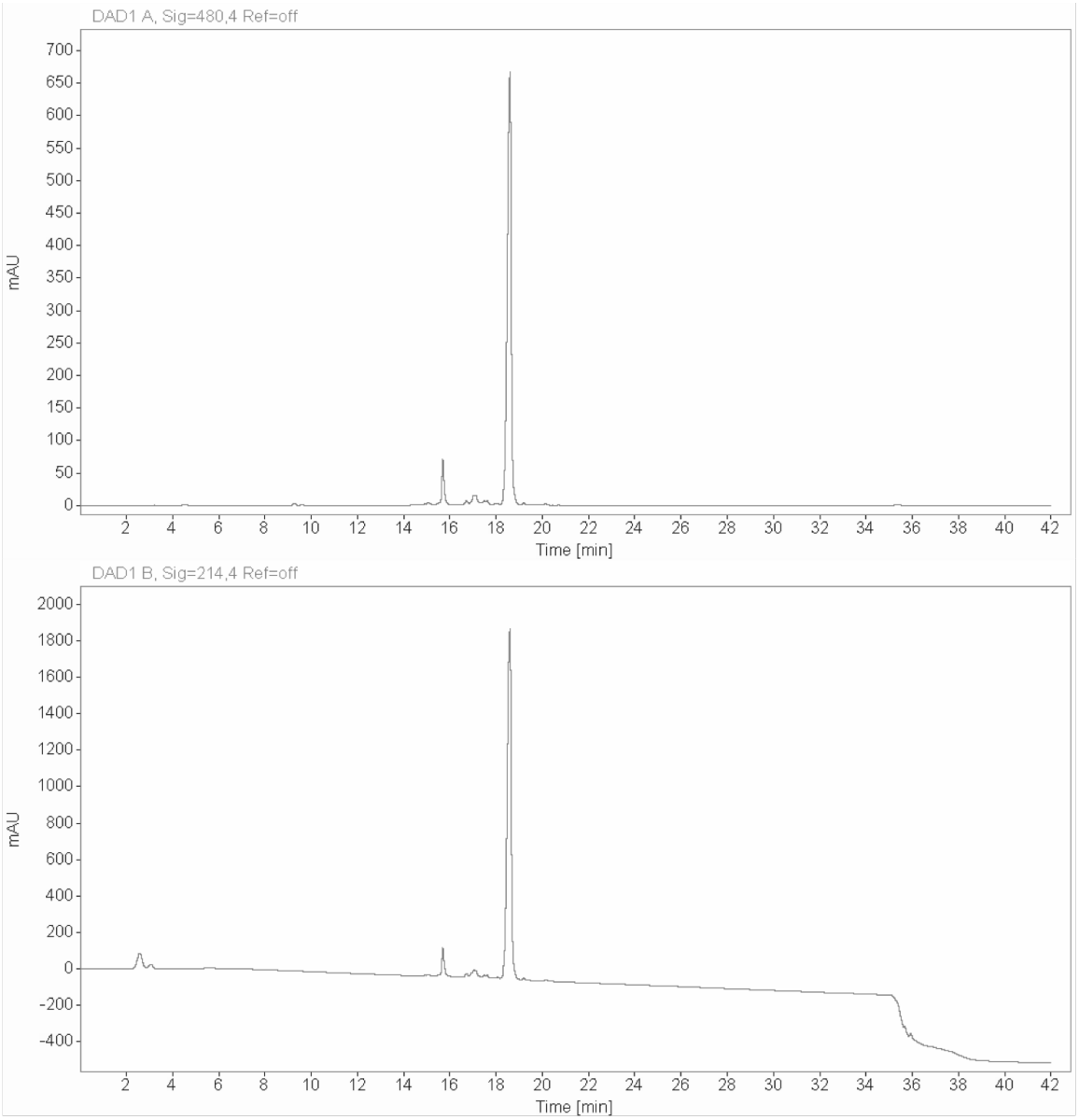
Characterization of Synthesized Peptides

This appendix contains the analytical HPLC chromatograms of fluorescent tracers derived from transcriptional activators that have been used in the various fluorescence polarization based assays throughout the work described in this thesis.

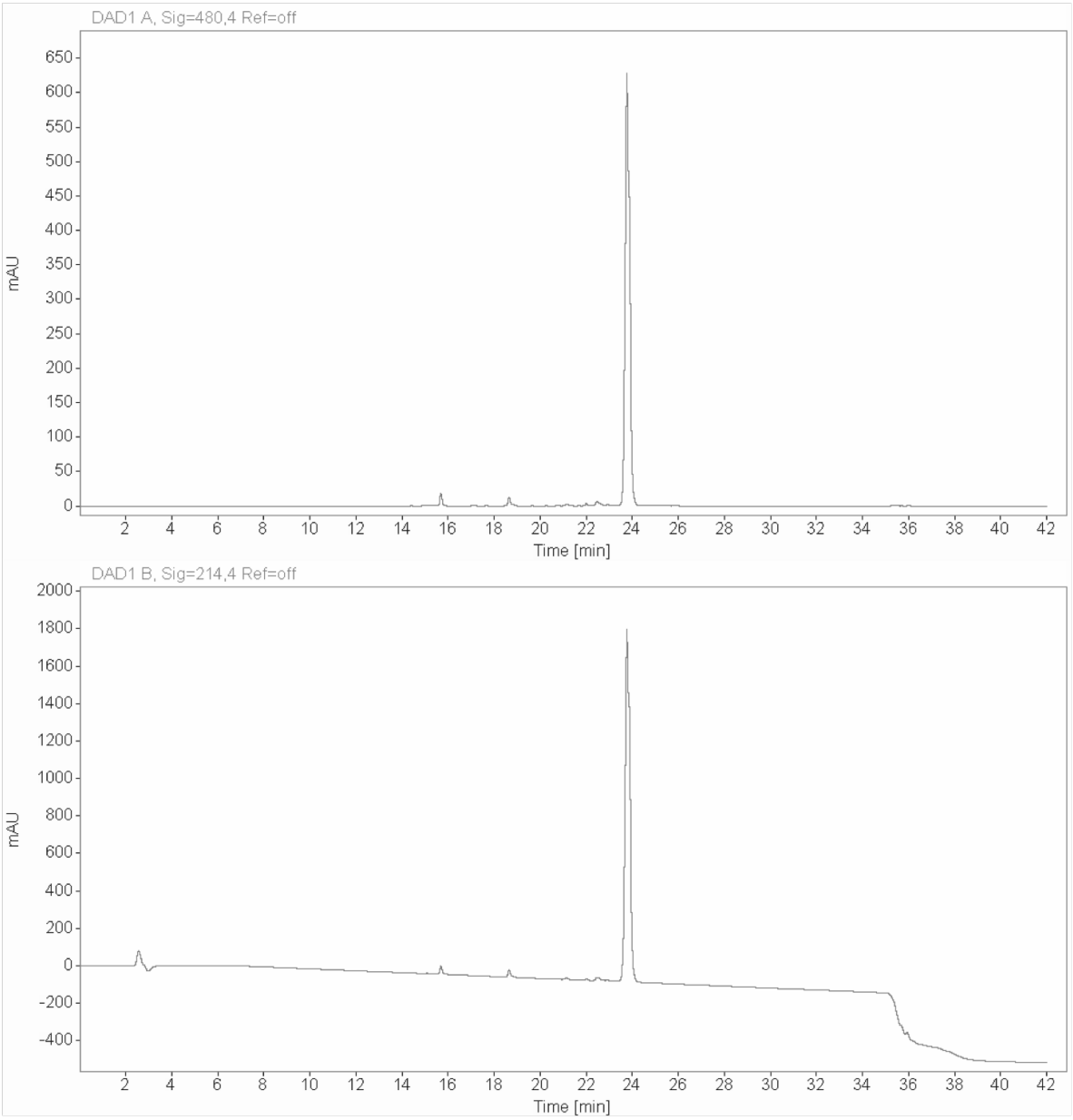
Analytical HPLC Chromatogram of Flo-VP16(413-437)



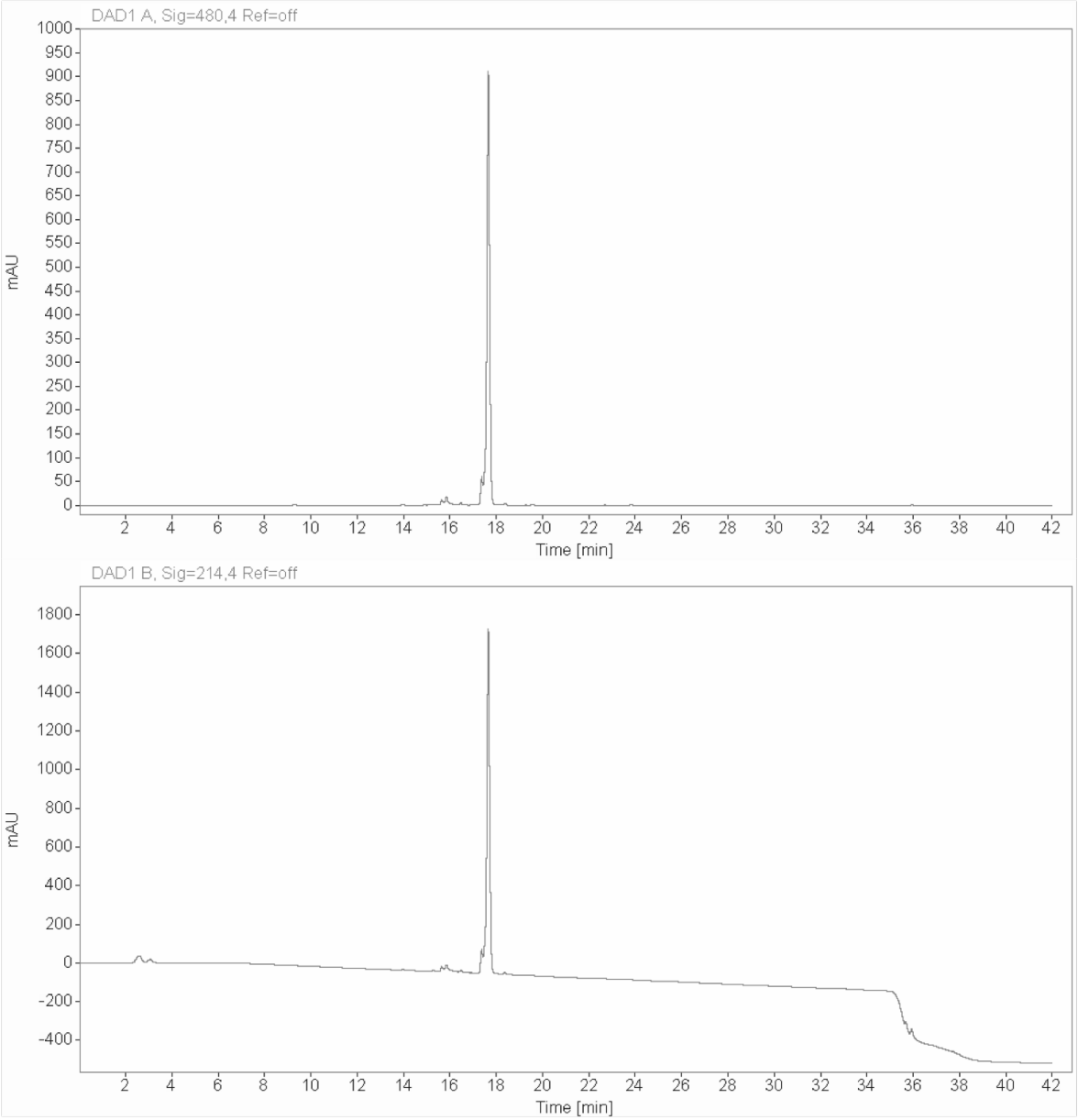
Analytical HPLC Chromatogram of Flo-VP16(438-464)



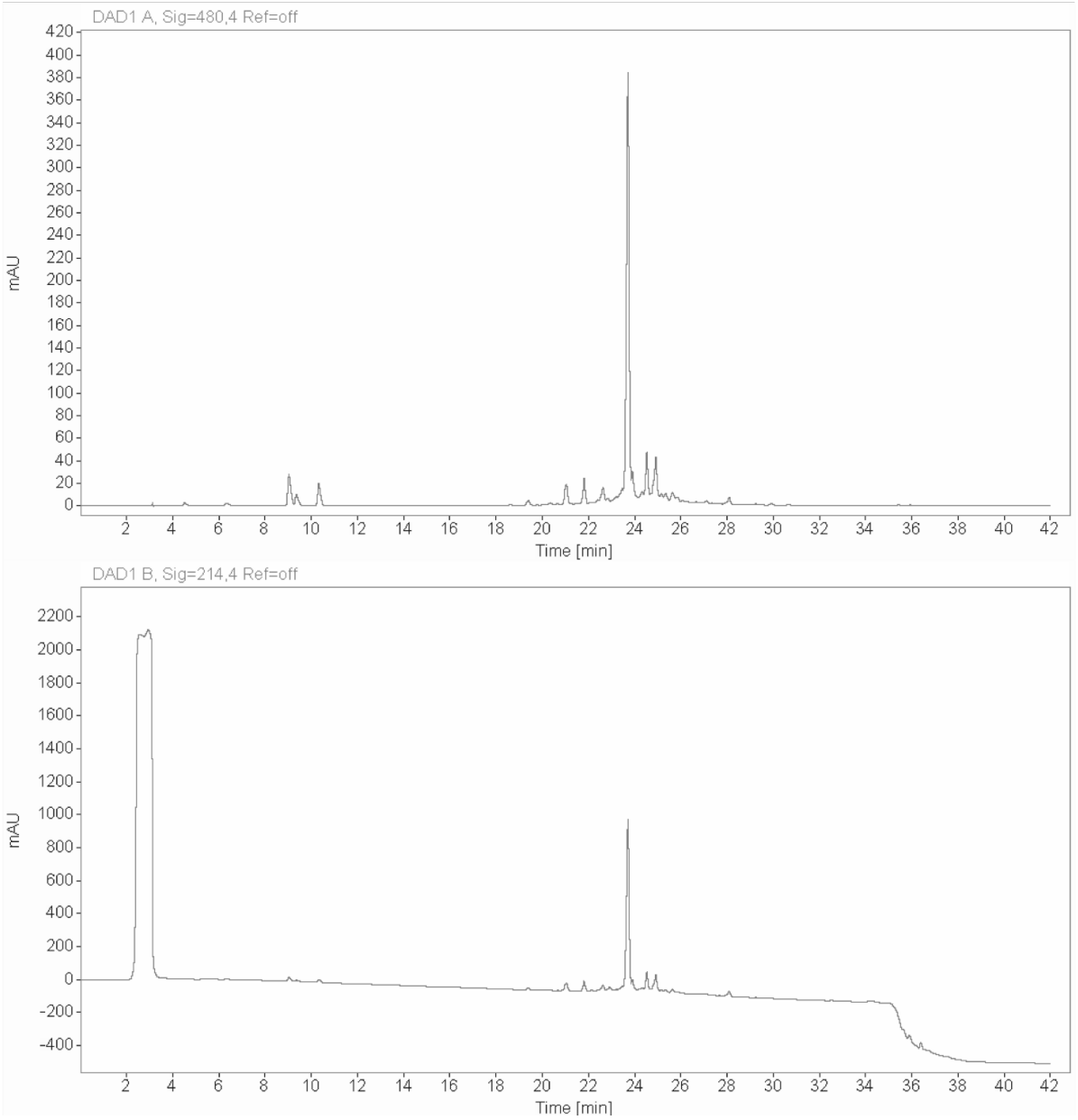
Analytical HPLC Chromatogram of Flo-VP16(465-490)



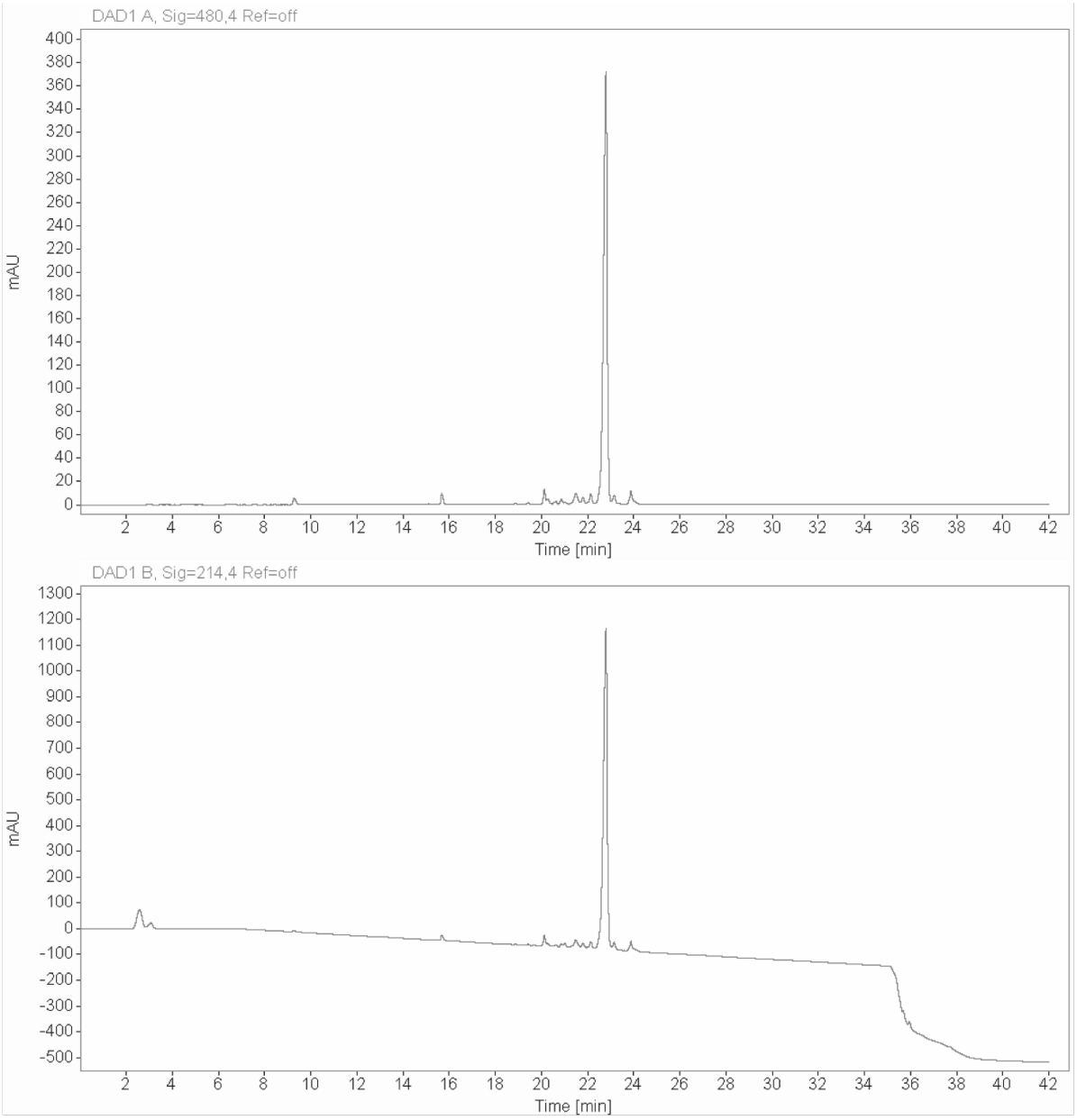
Analytical HPLC Chromatogram of Flo-VP16(438-454)



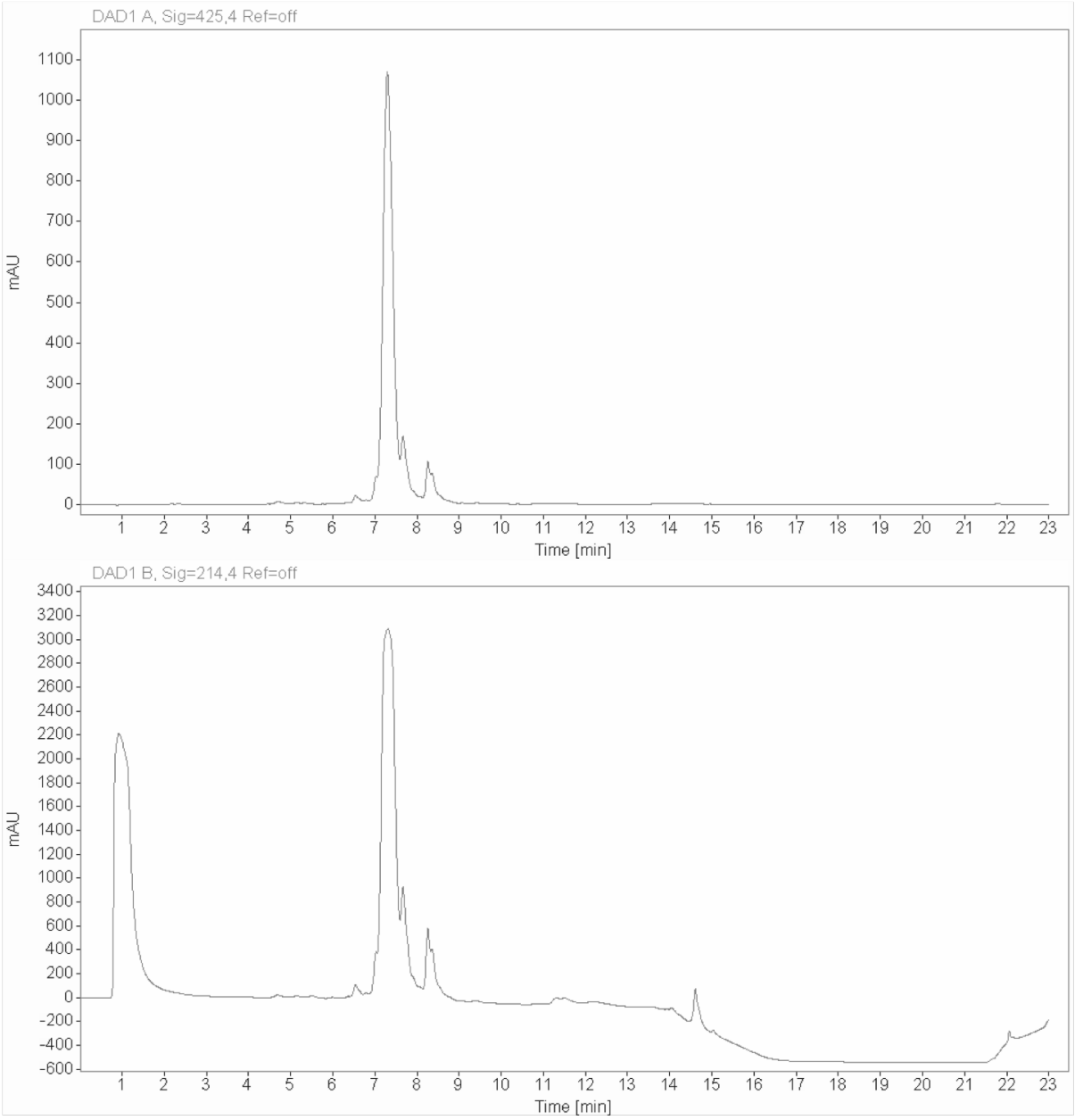
Analytical HPLC Chromatogram of Flo-VP16(467-488)



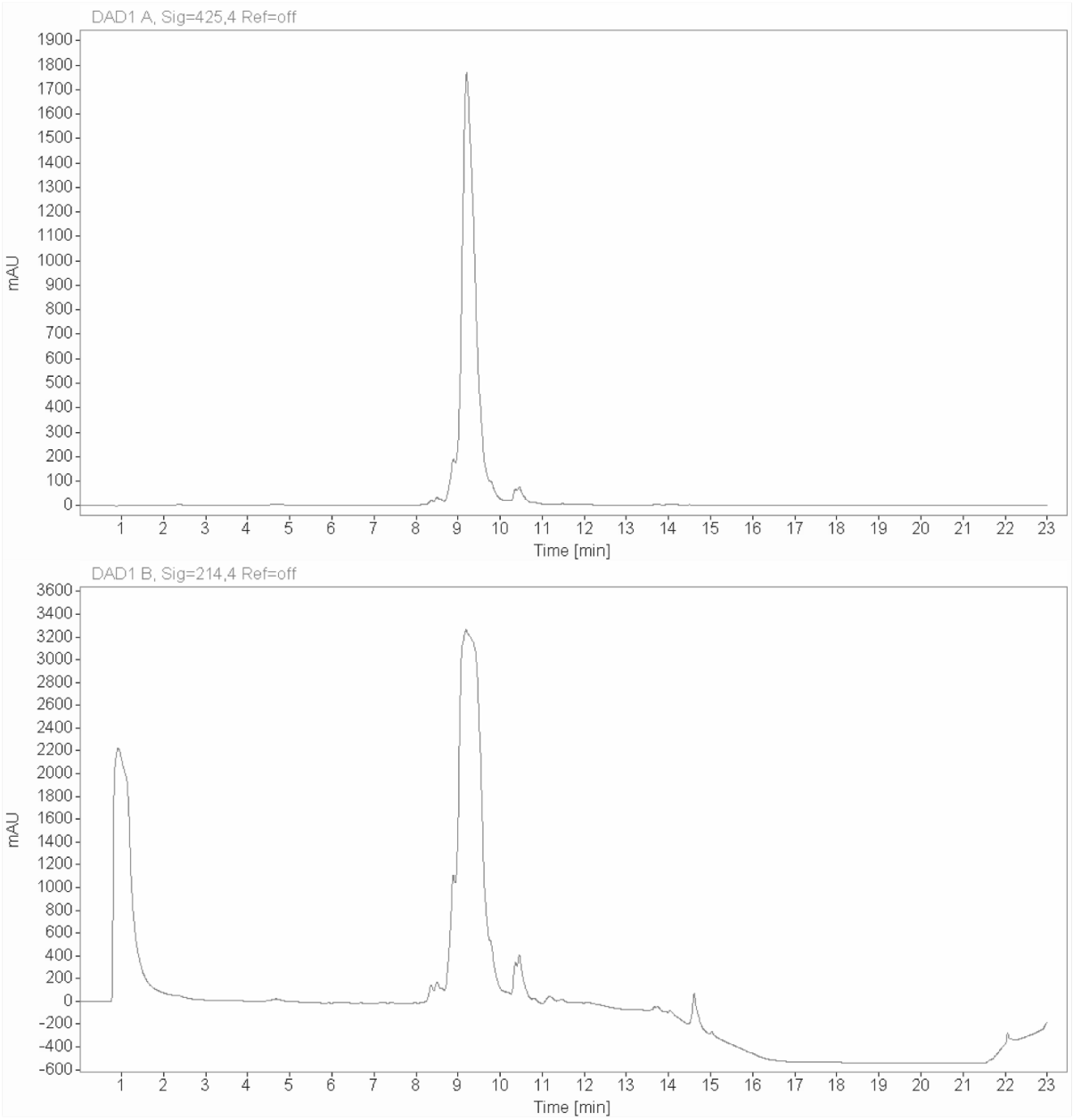
Analytical HPLC Chromatogram of Flo-ERM(38-68)



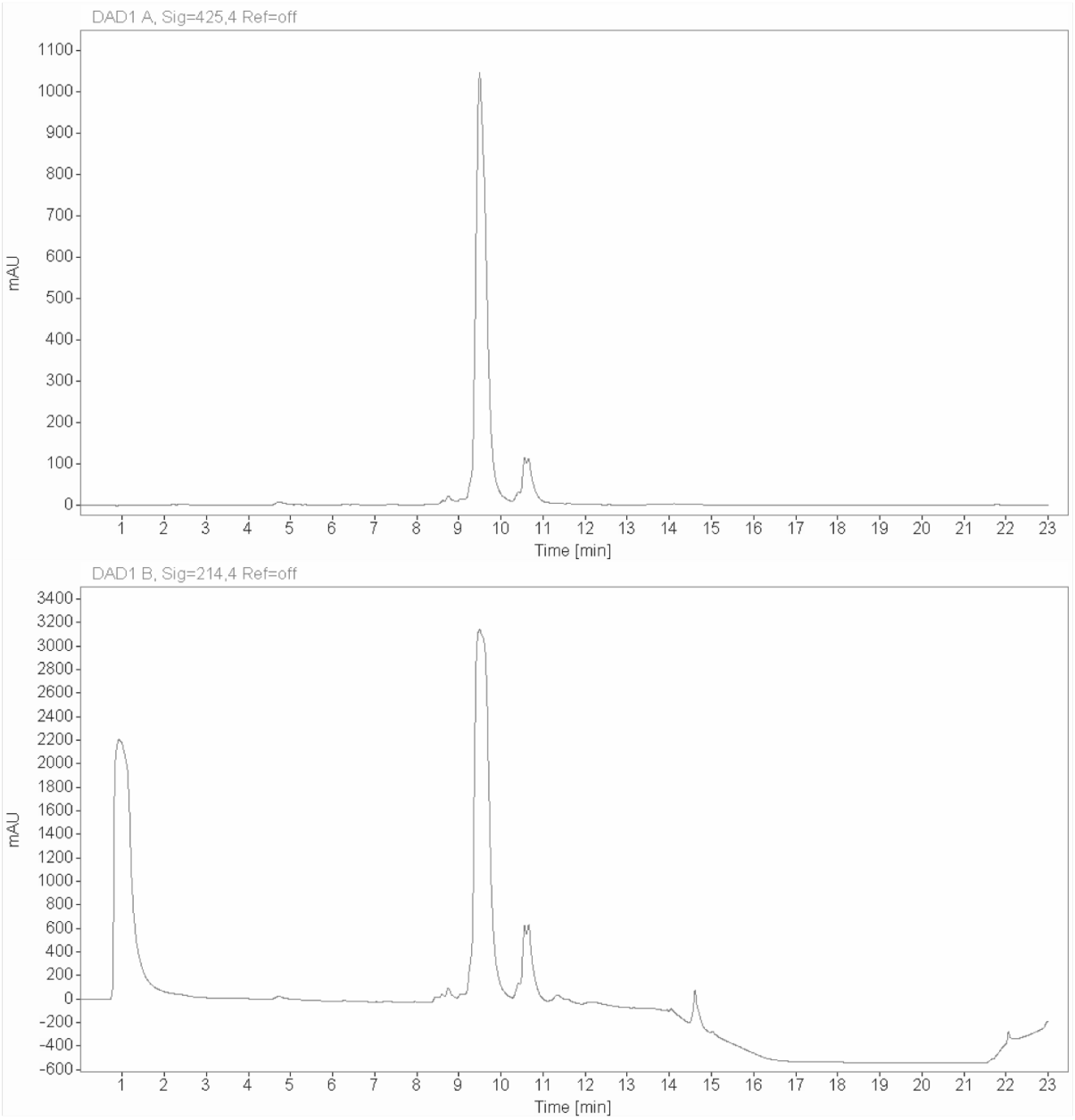
Analytical HPLC Chromatogram of Flo-VP16(438-454) L439A



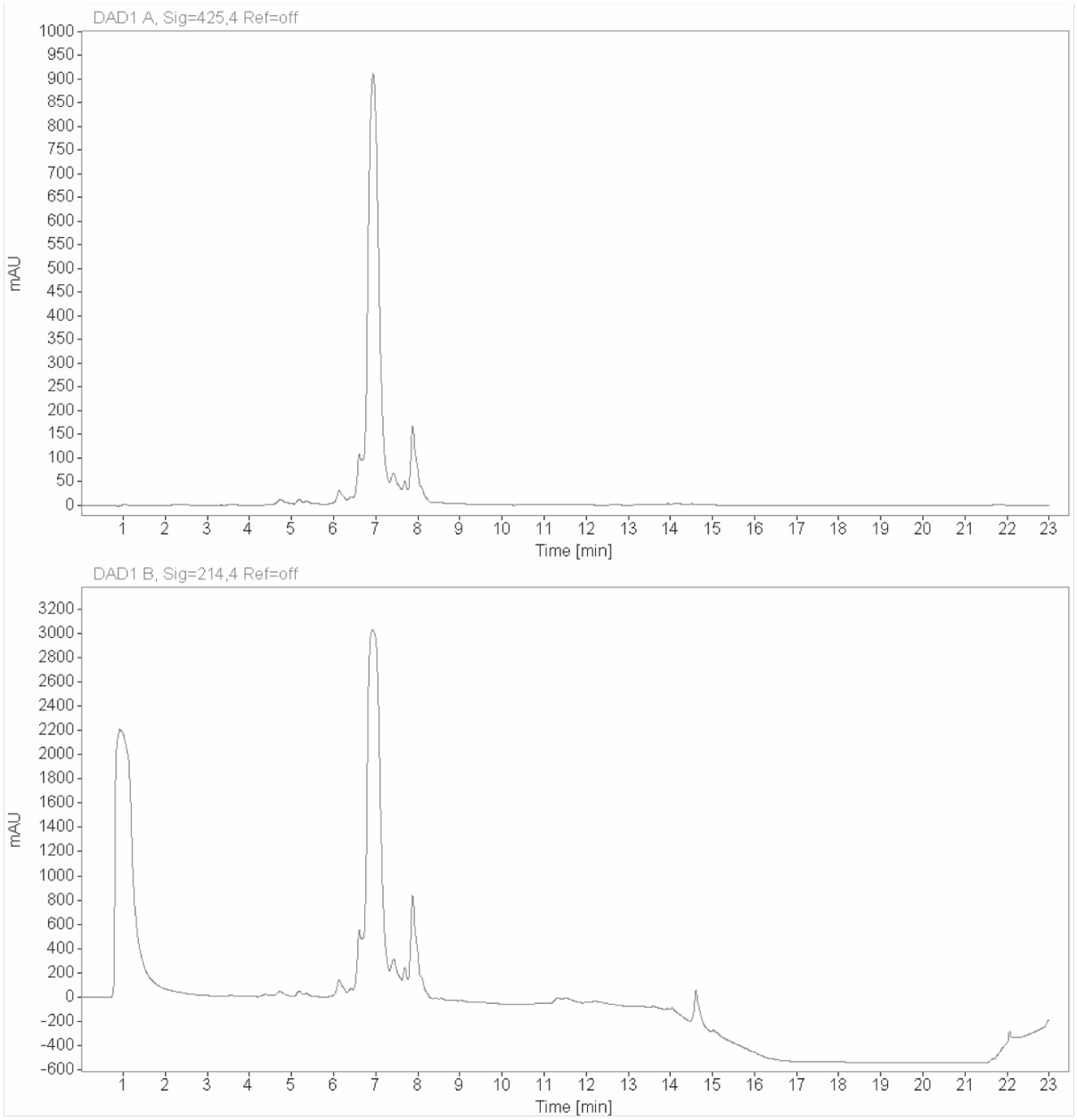
Analytical HPLC Chromatogram of Flo-VP16(438-454) D440A



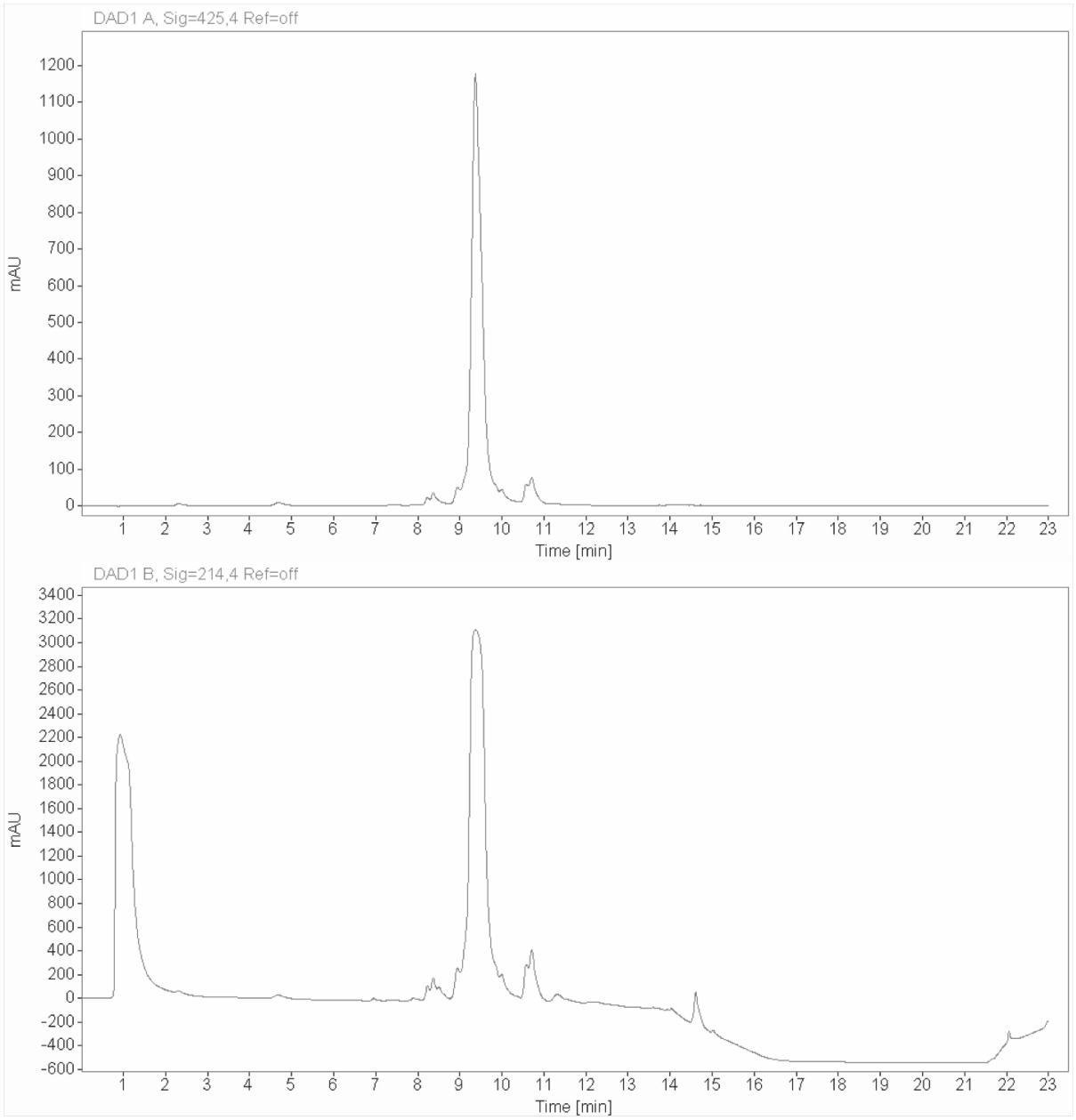
Analytical HPLC Chromatogram of Flo-VP16(438-454) D441A



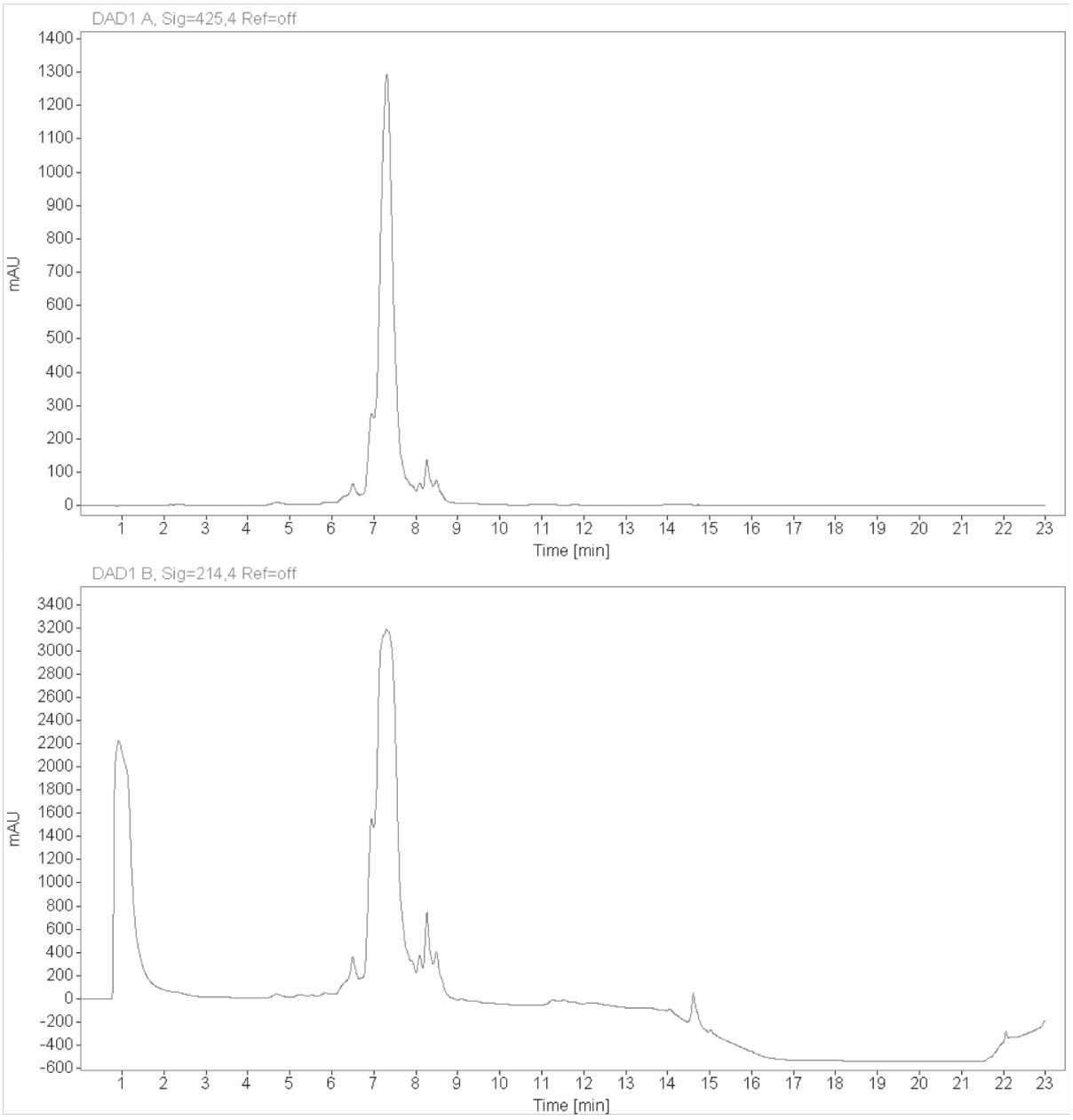
Analytical HPLC Chromatogram of Flo-VP16(438-454) F442A



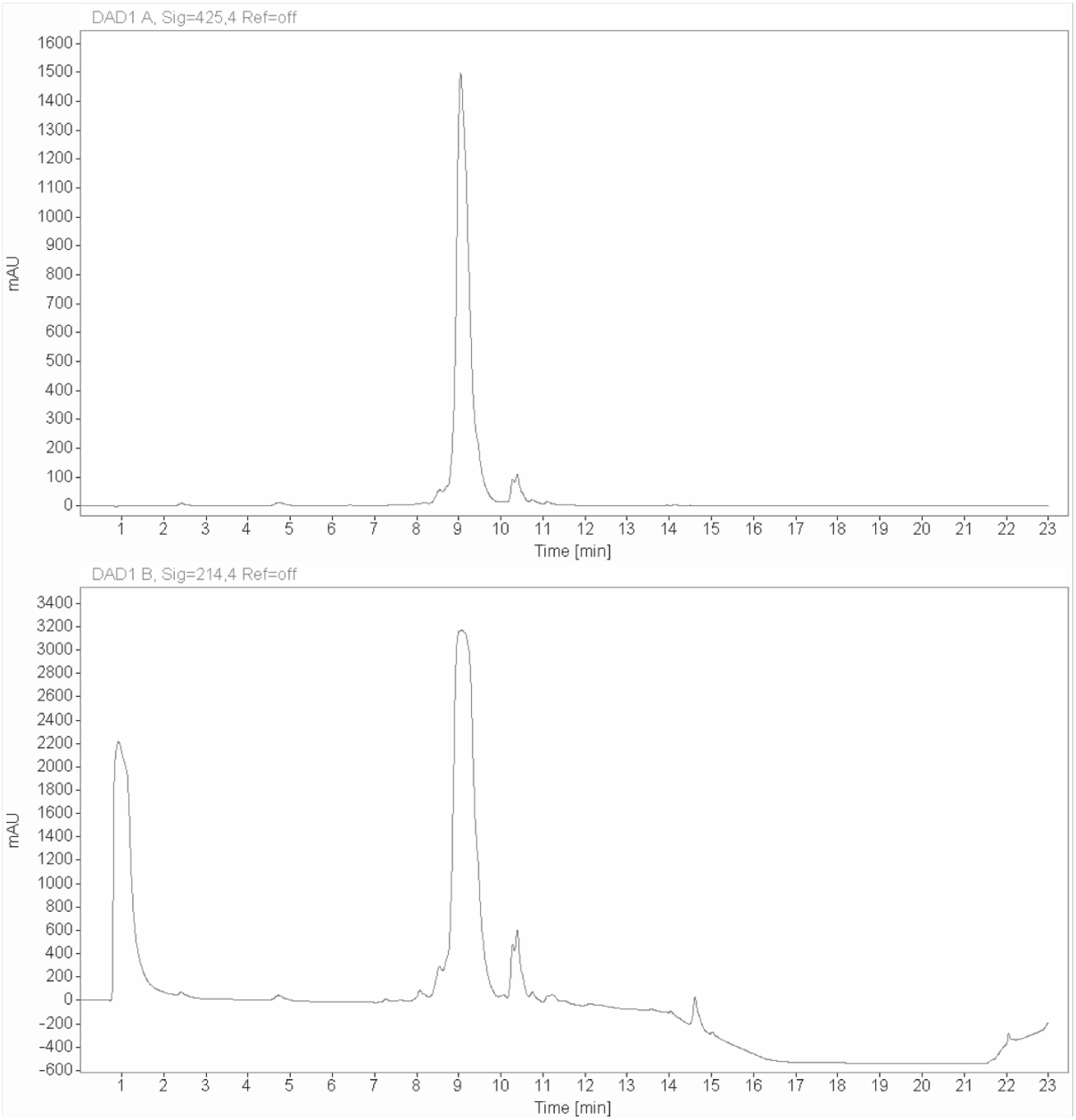
Analytical HPLC Chromatogram of Flo-VP16(438-454) D443A



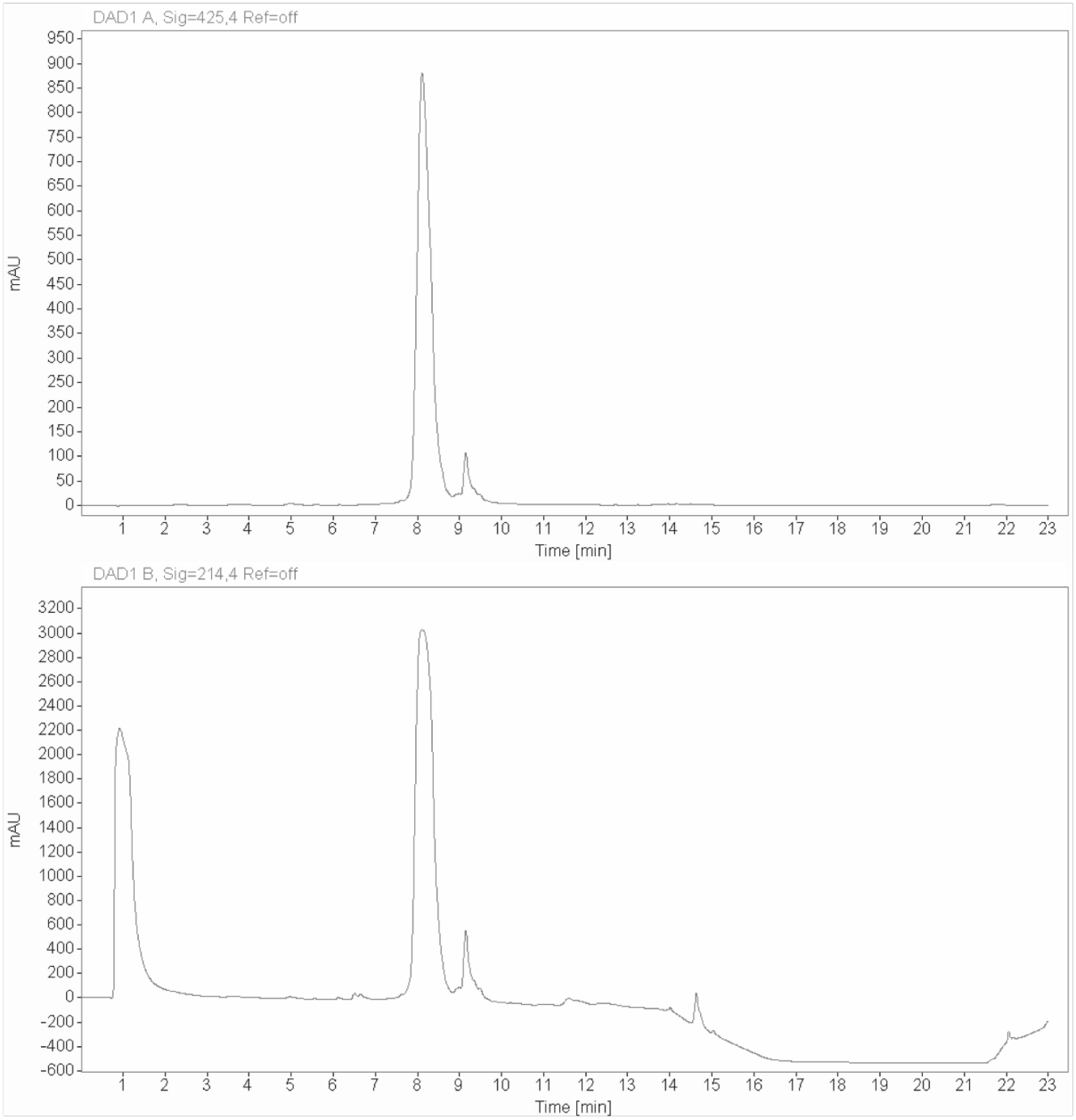
Analytical HPLC Chromatogram of Flo-VP16(438-454) L444A



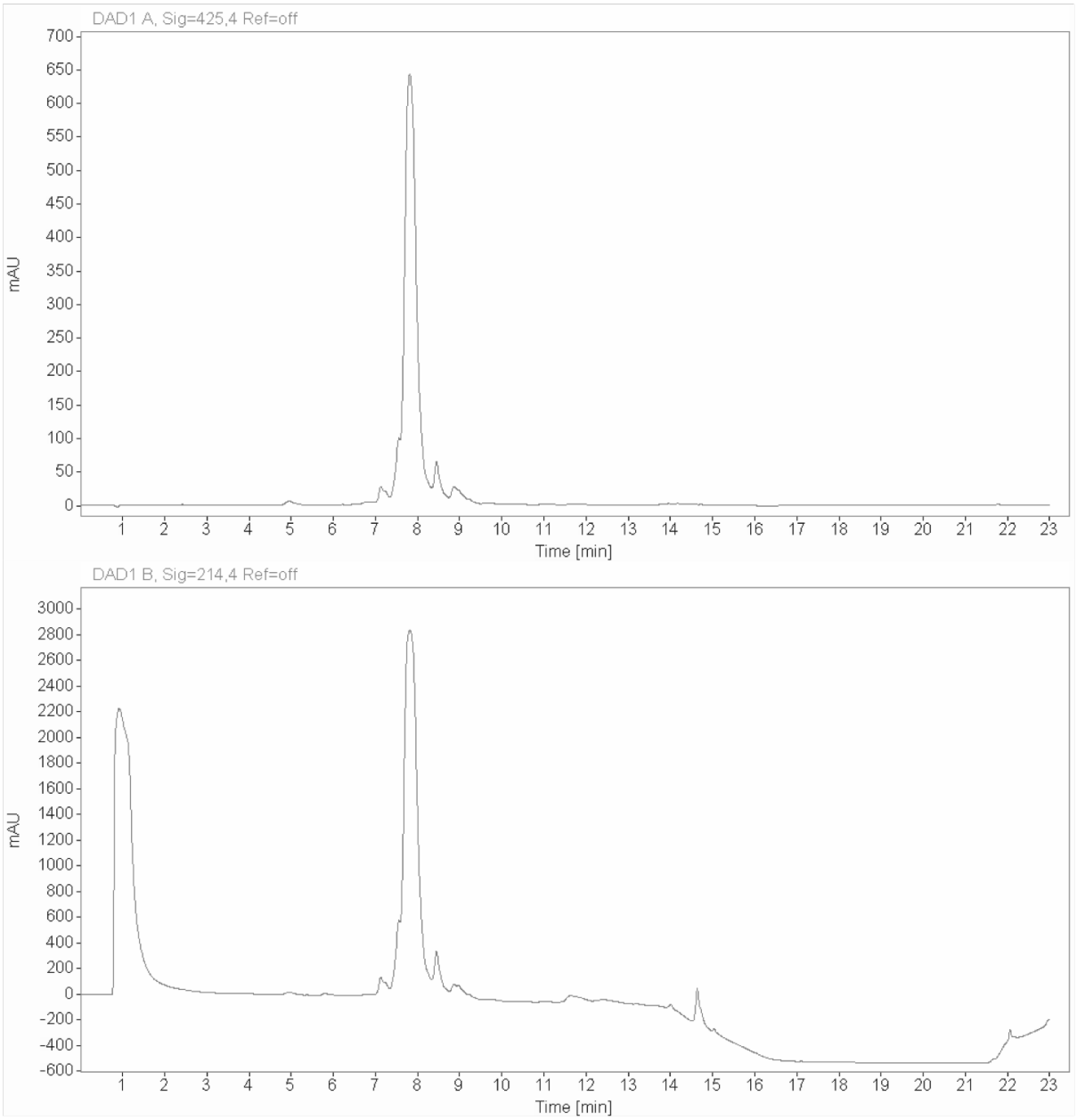
Analytical HPLC Chromatogram of Flo-VP16(438-454) D445A



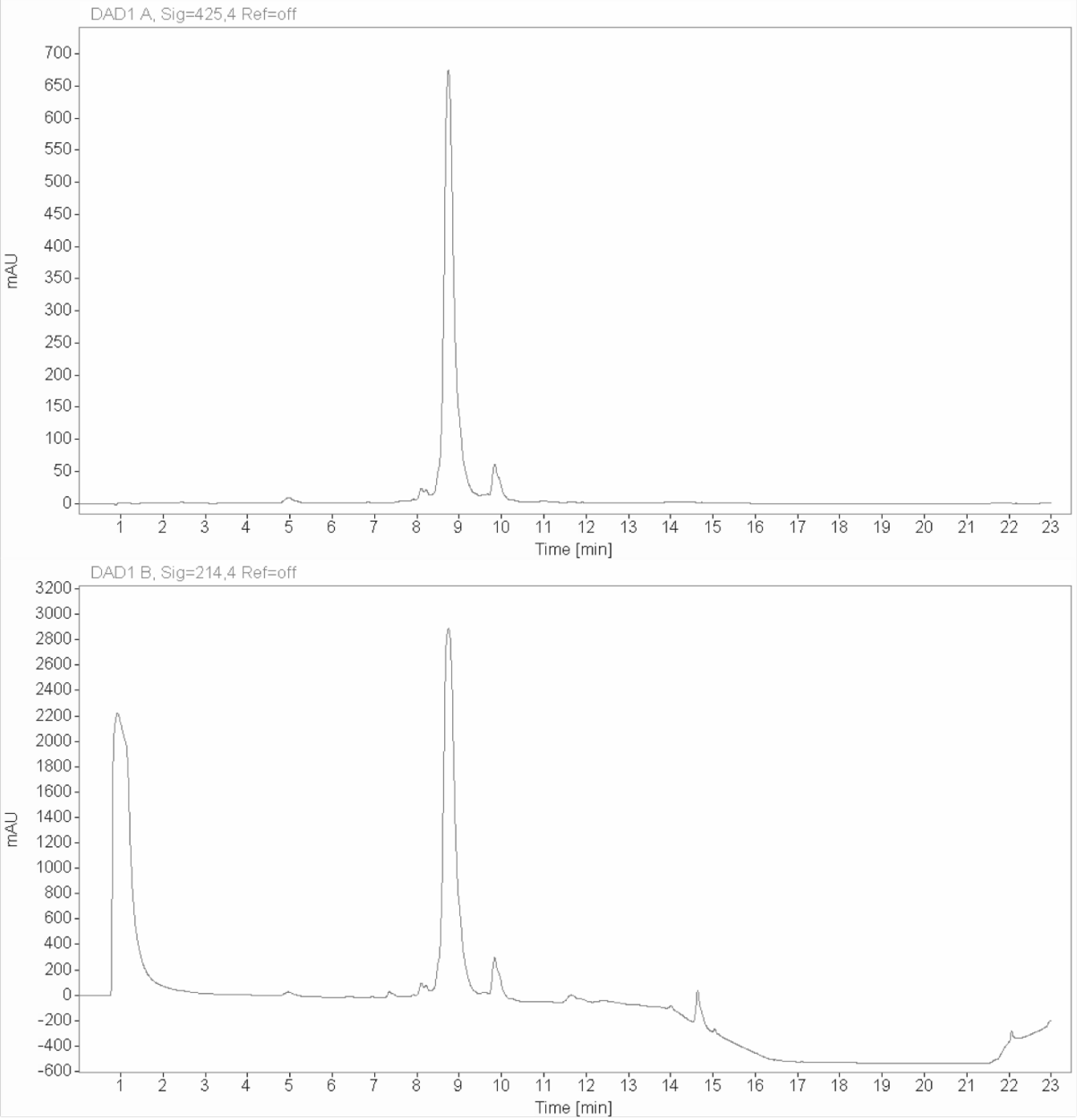
Analytical HPLC Chromatogram of Flo-VP16(438-454) M446A



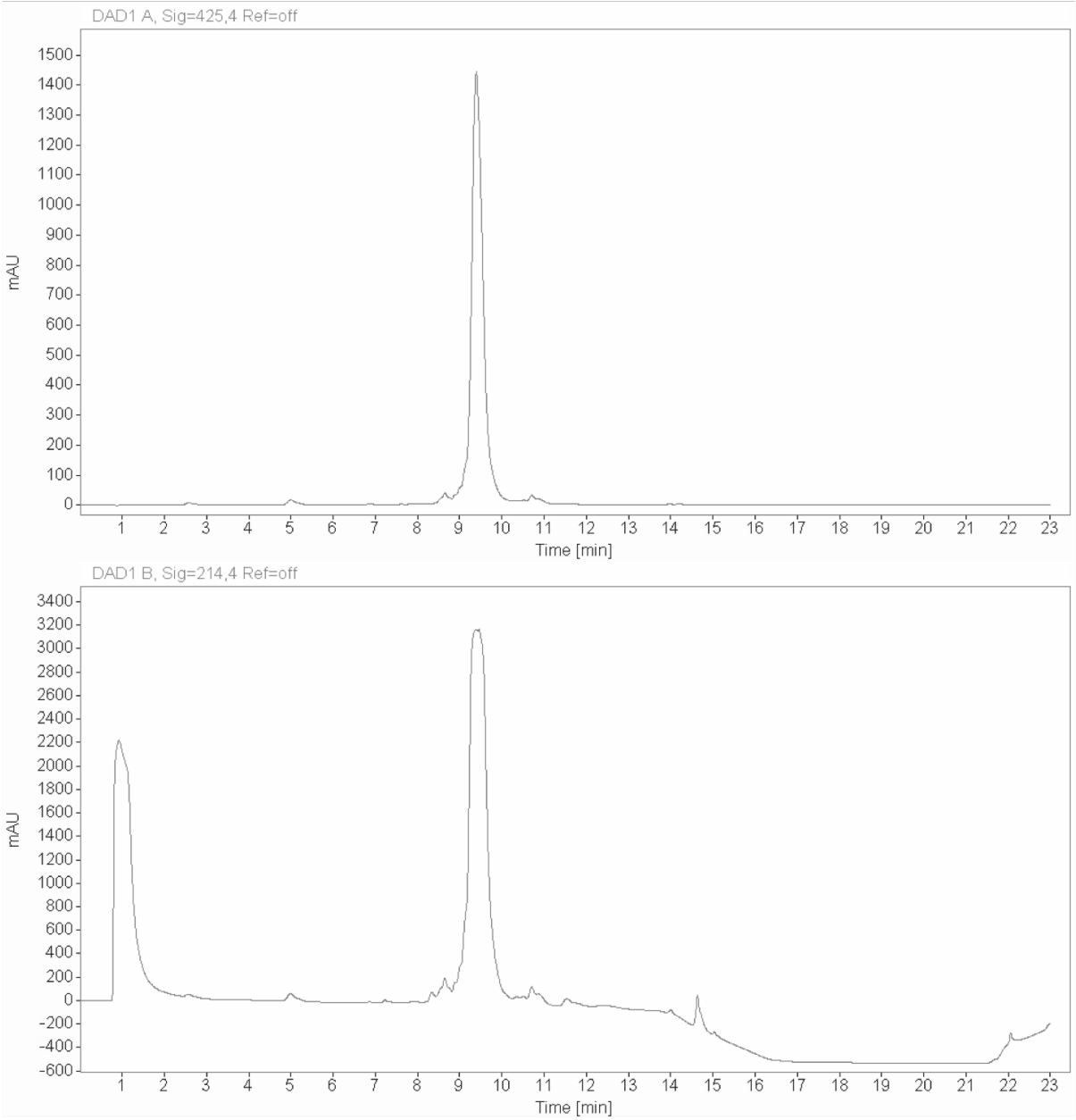
Analytical HPLC Chromatogram of Flo-VP16(438-454) L447A



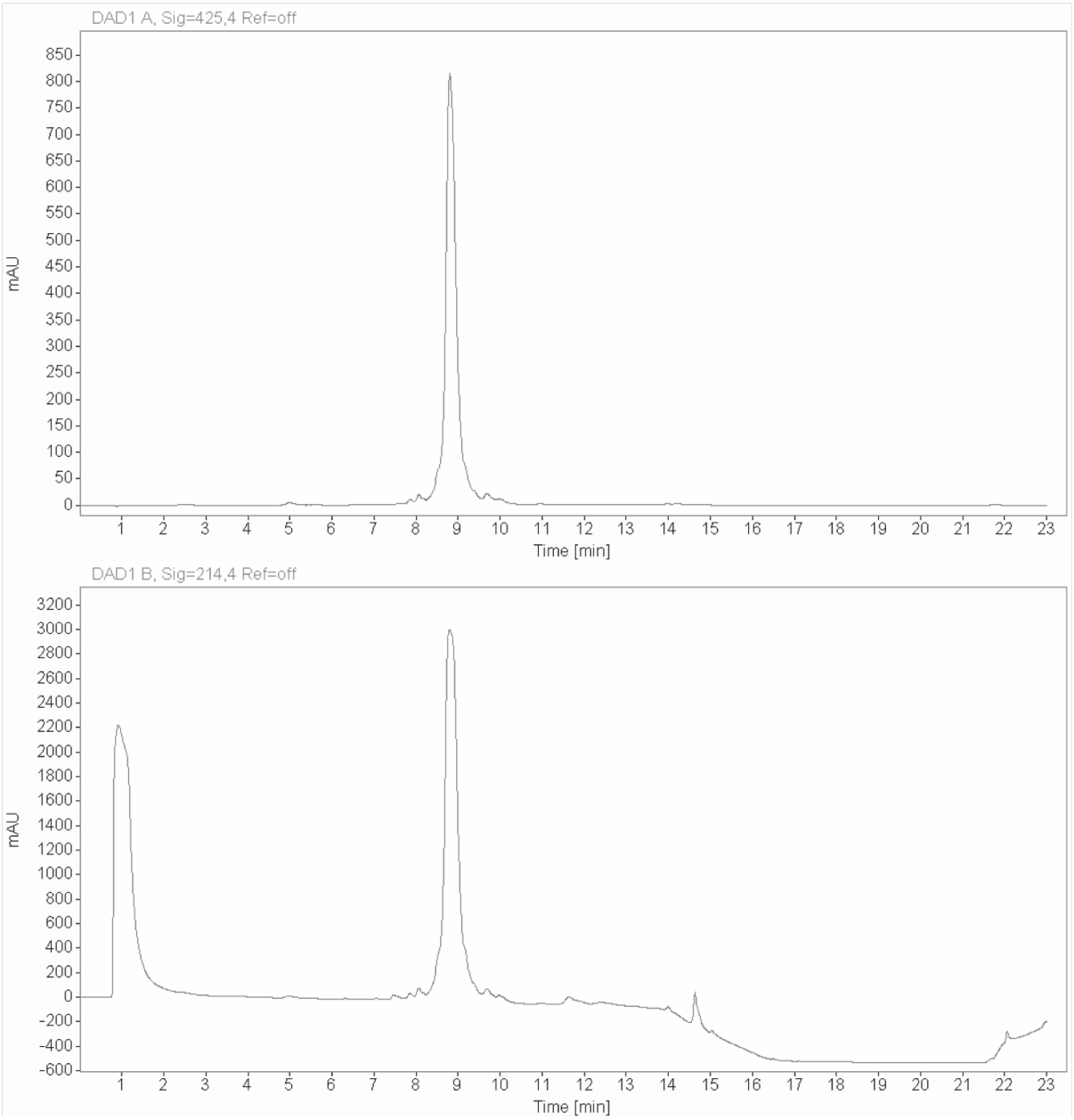
Analytical HPLC Chromatogram of Flo-VP16(438-454) G448A



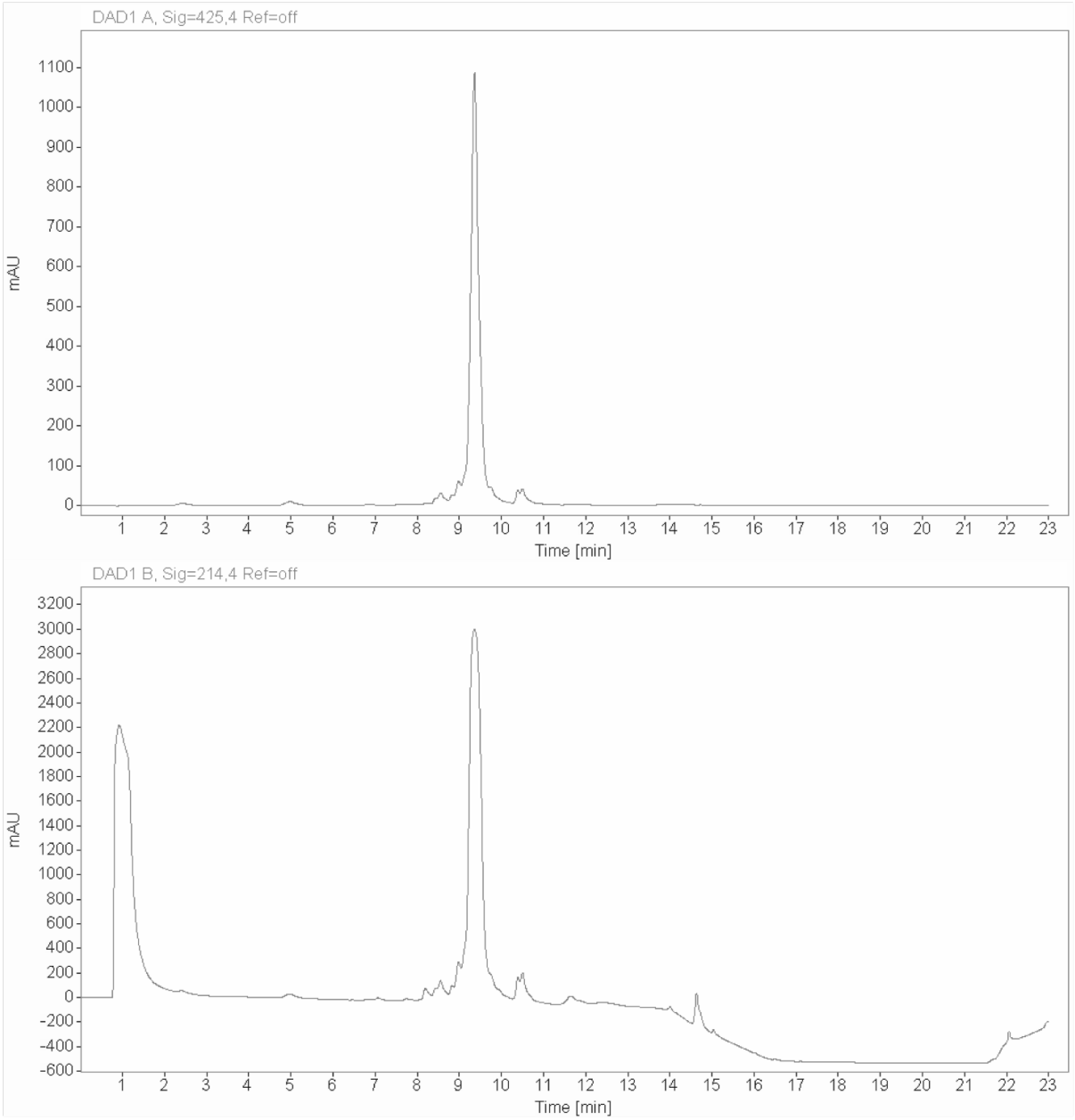
Analytical HPLC Chromatogram of Flo-VP16(438-454) D449A



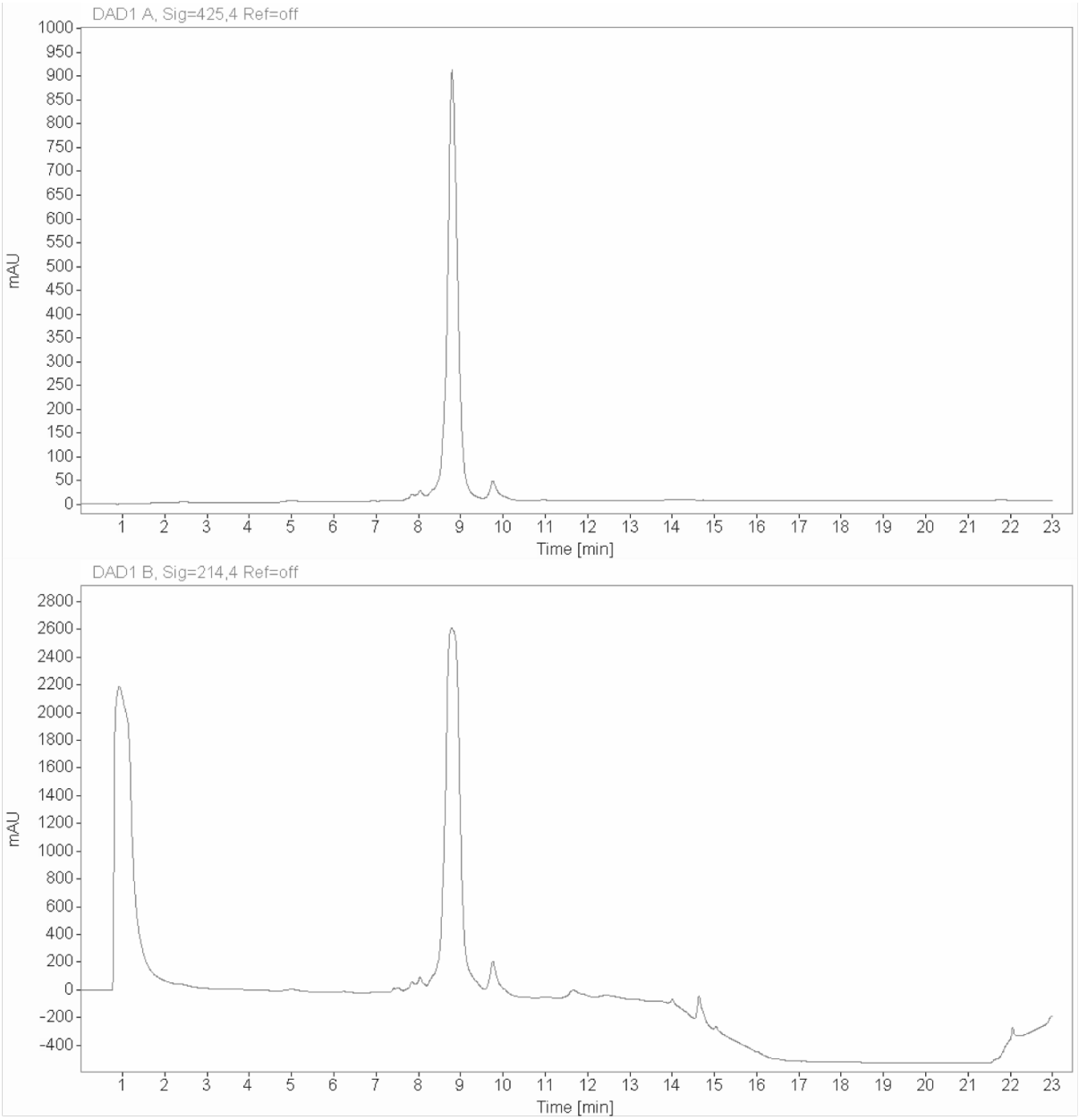
Analytical HPLC Chromatogram of Flo-VP16(438-454) G450A



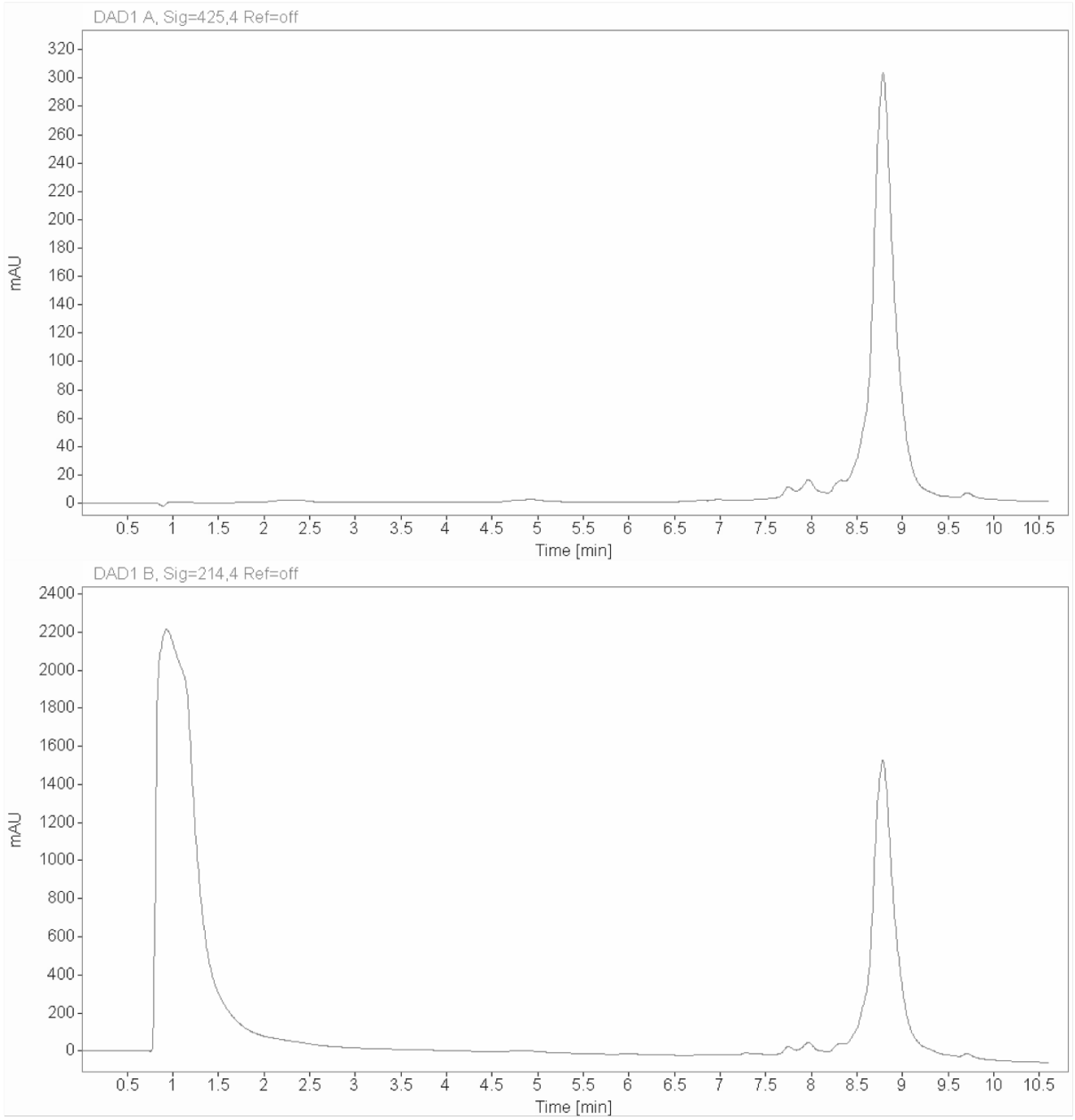
Analytical HPLC Chromatogram of Flo-VP16(438-454) D451A



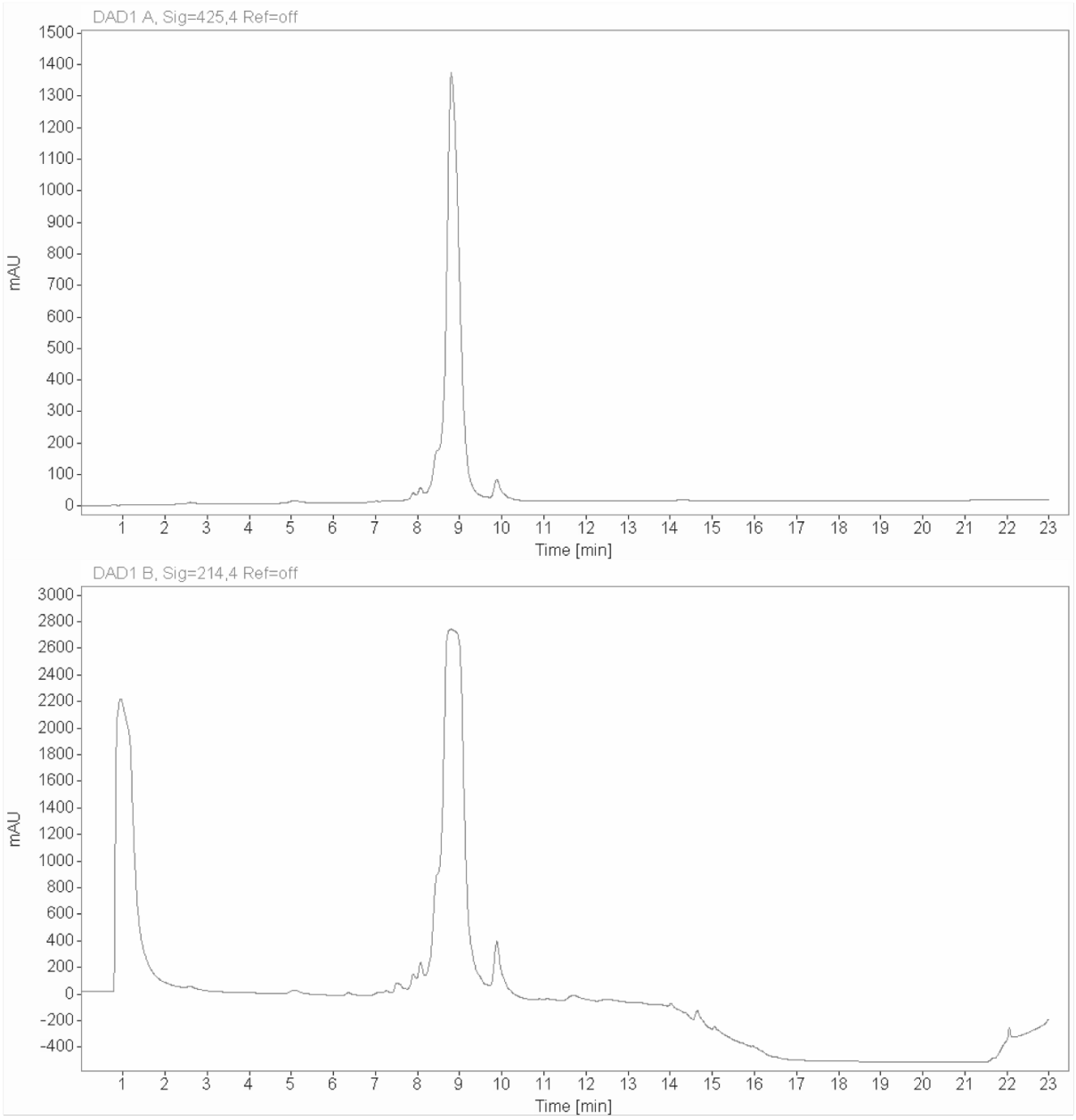
Analytical HPLC Chromatogram of Flo-VP16(438-454) S452A



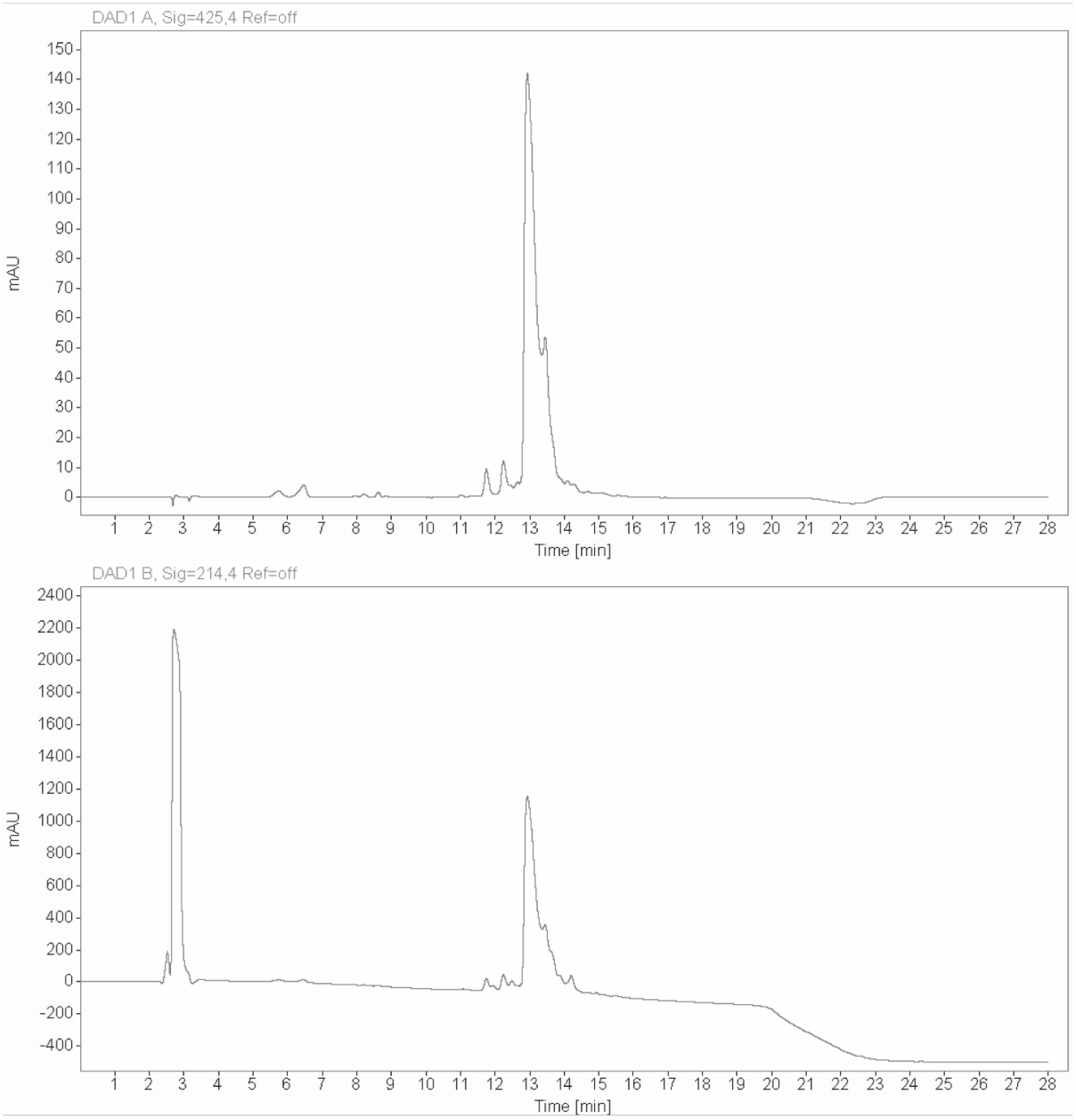
Analytical HPLC Chromatogram of Flo-VP16(438-454) P453A



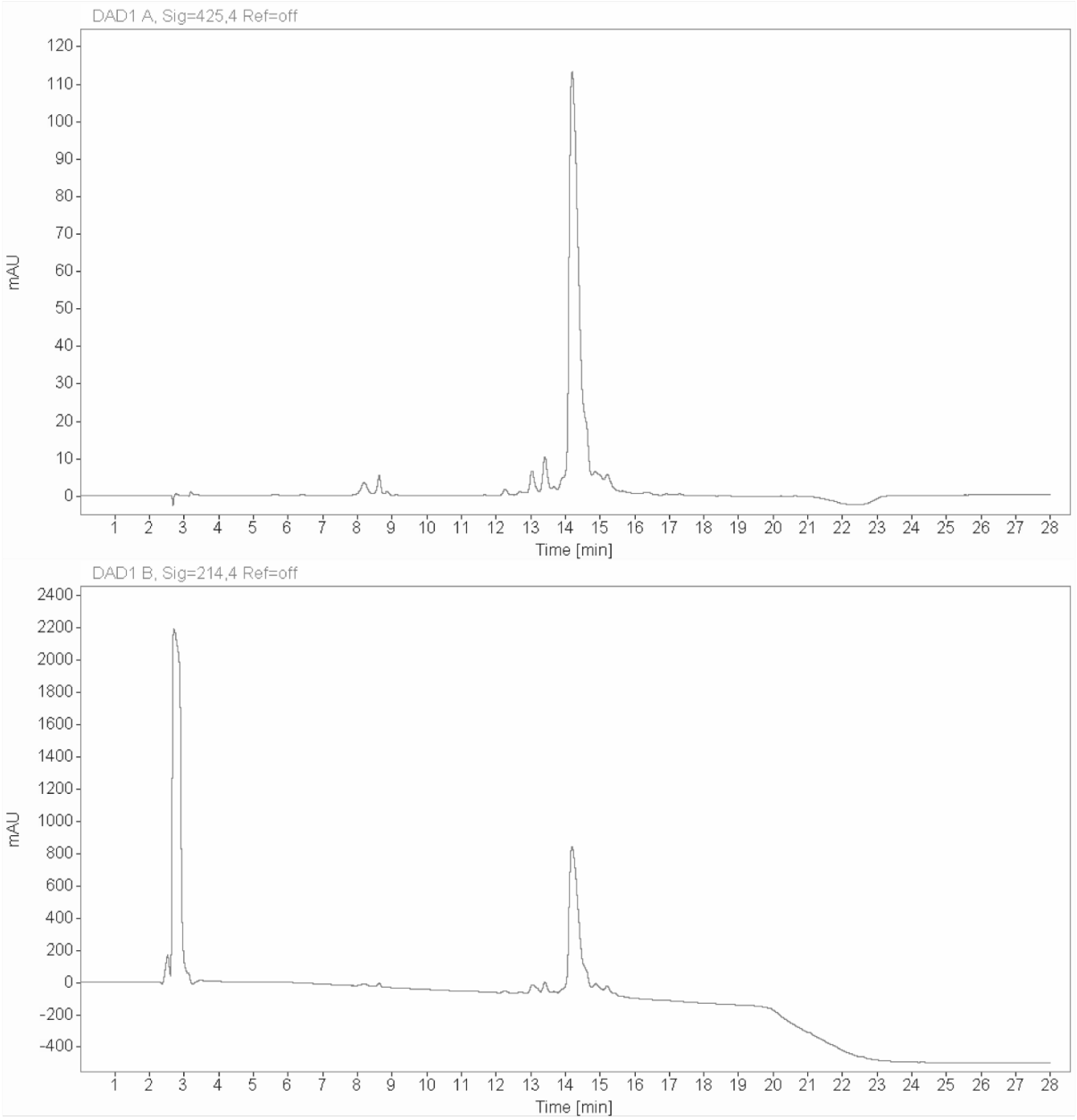
Analytical HPLC Chromatogram of Flo-VP16(438-454) G454A



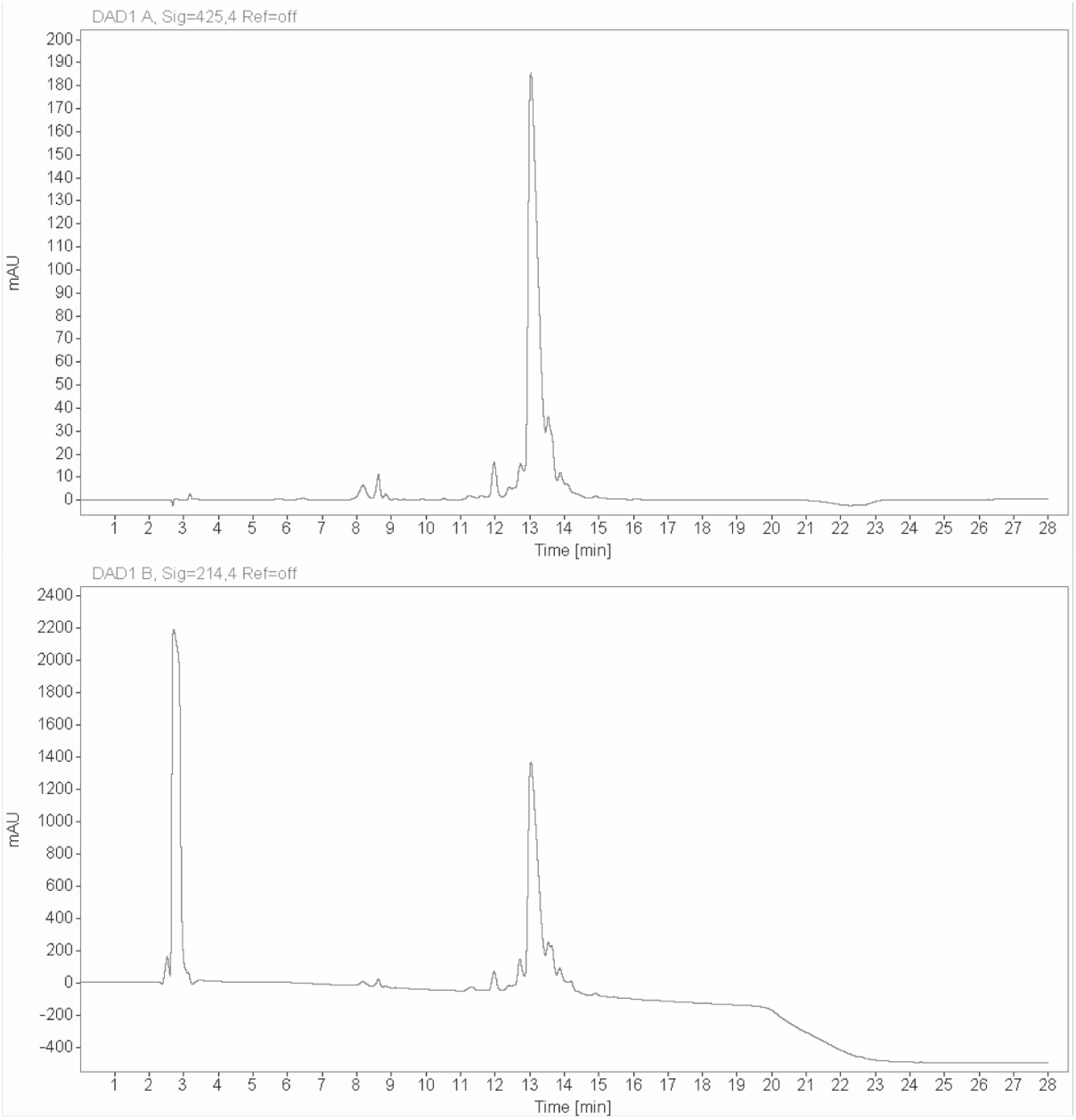
Analytical HPLC Chromatogram of Flo-VP16(467-488) L468A



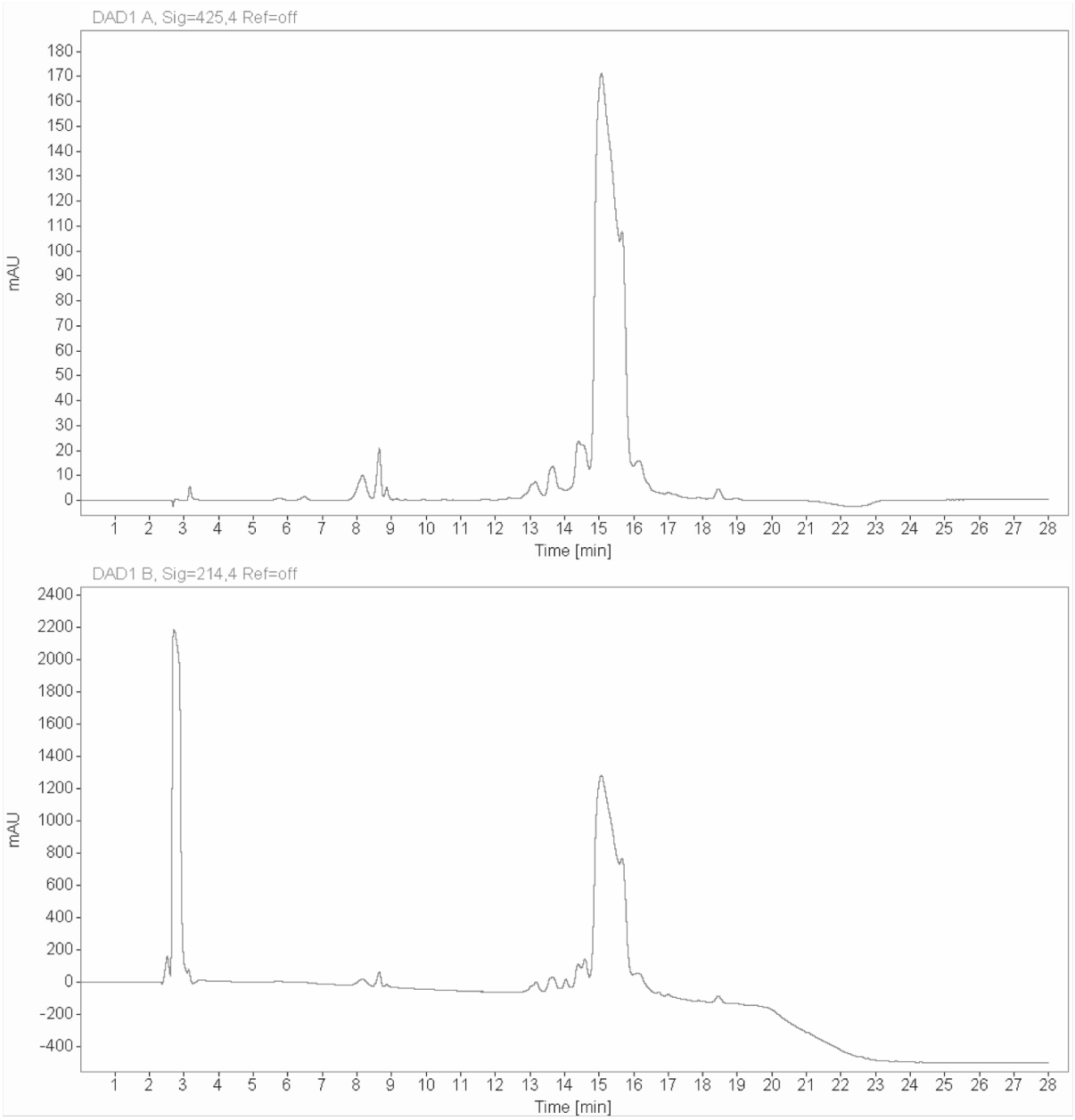
Analytical HPLC Chromatogram of Flo-VP16(467-488) D469A



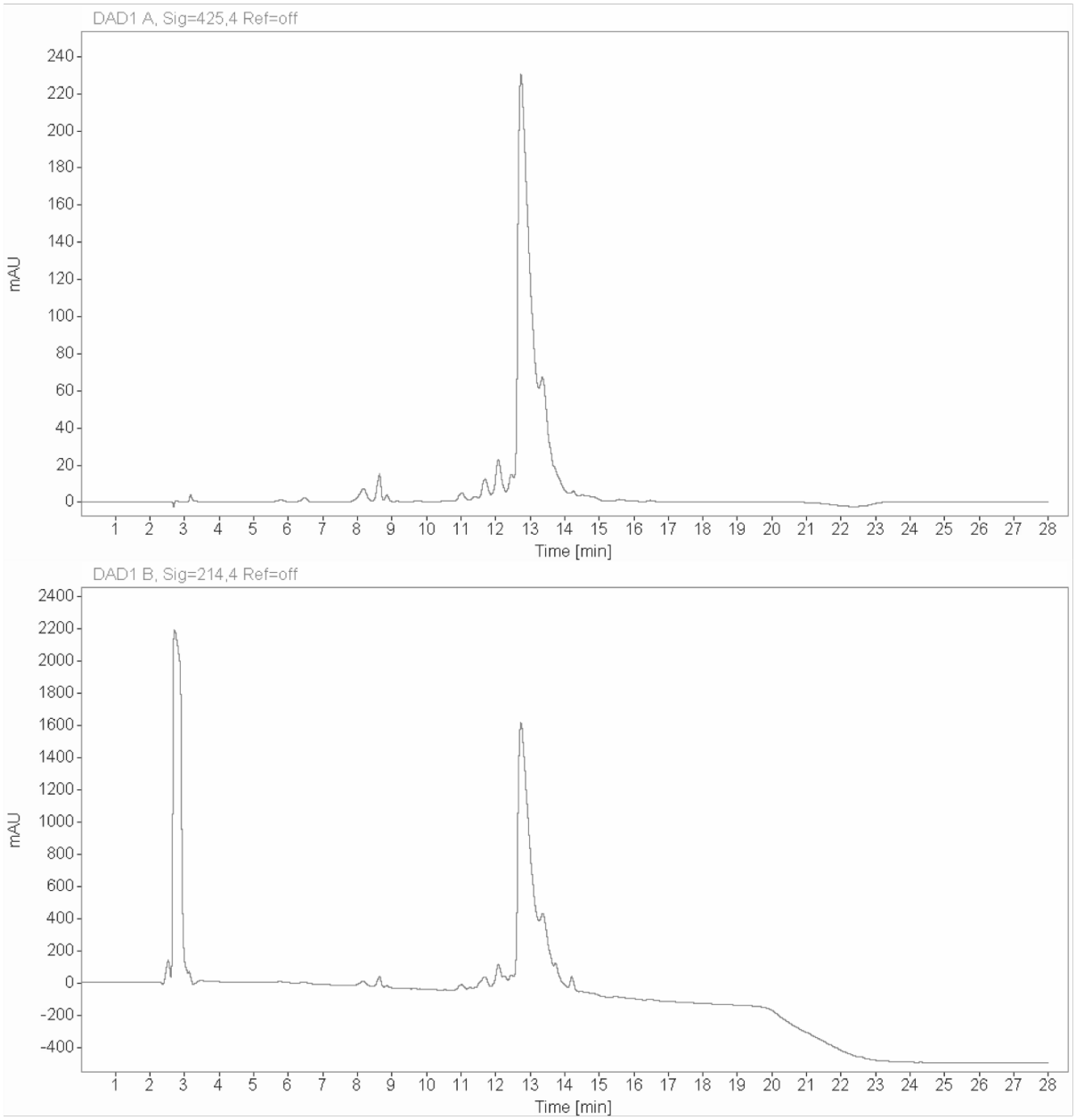
Analytical HPLC Chromatogram of Flo-VP16(467-488) M470A



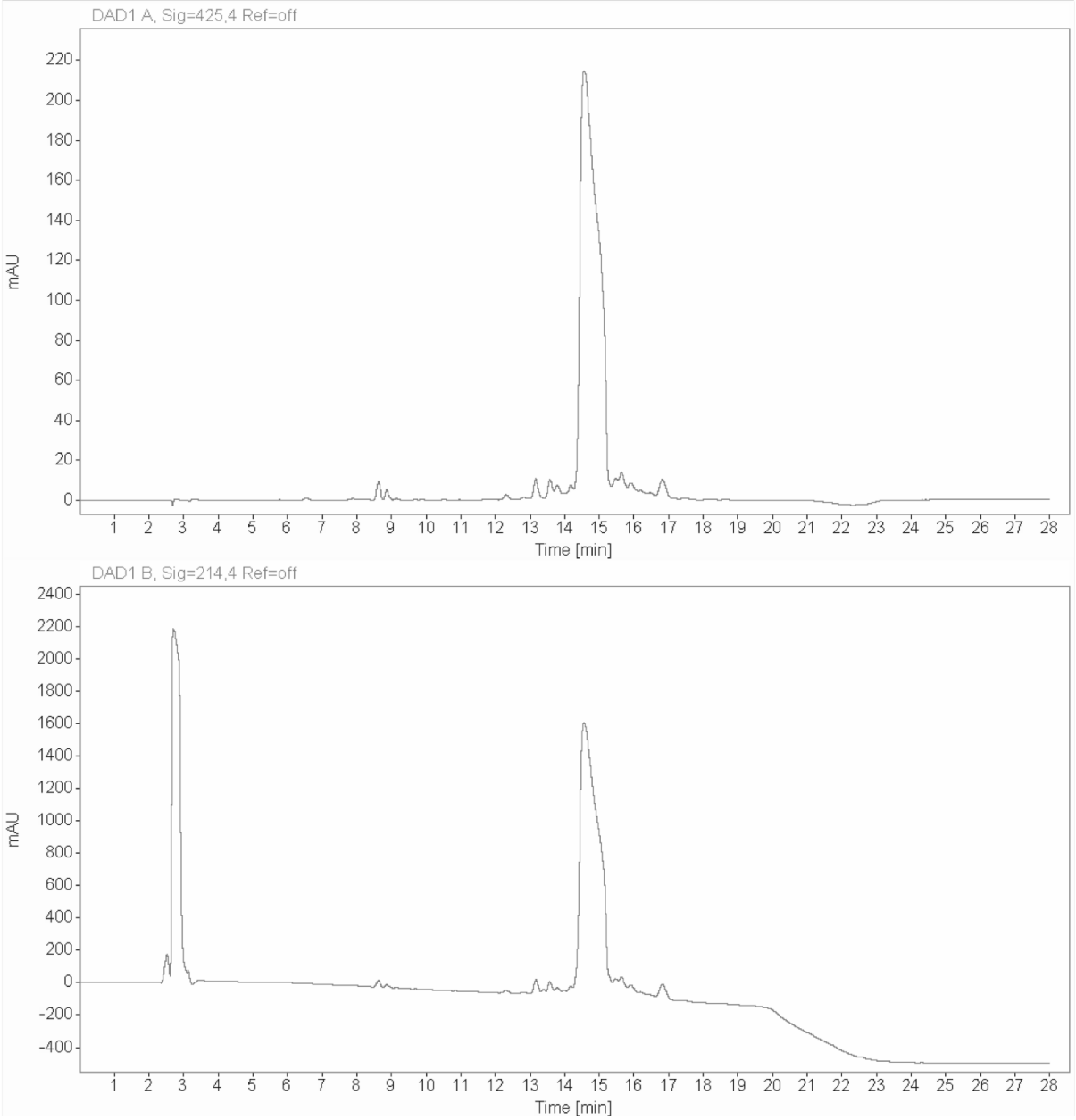
Analytical HPLC Chromatogram of Flo-VP16(467-488) D472A



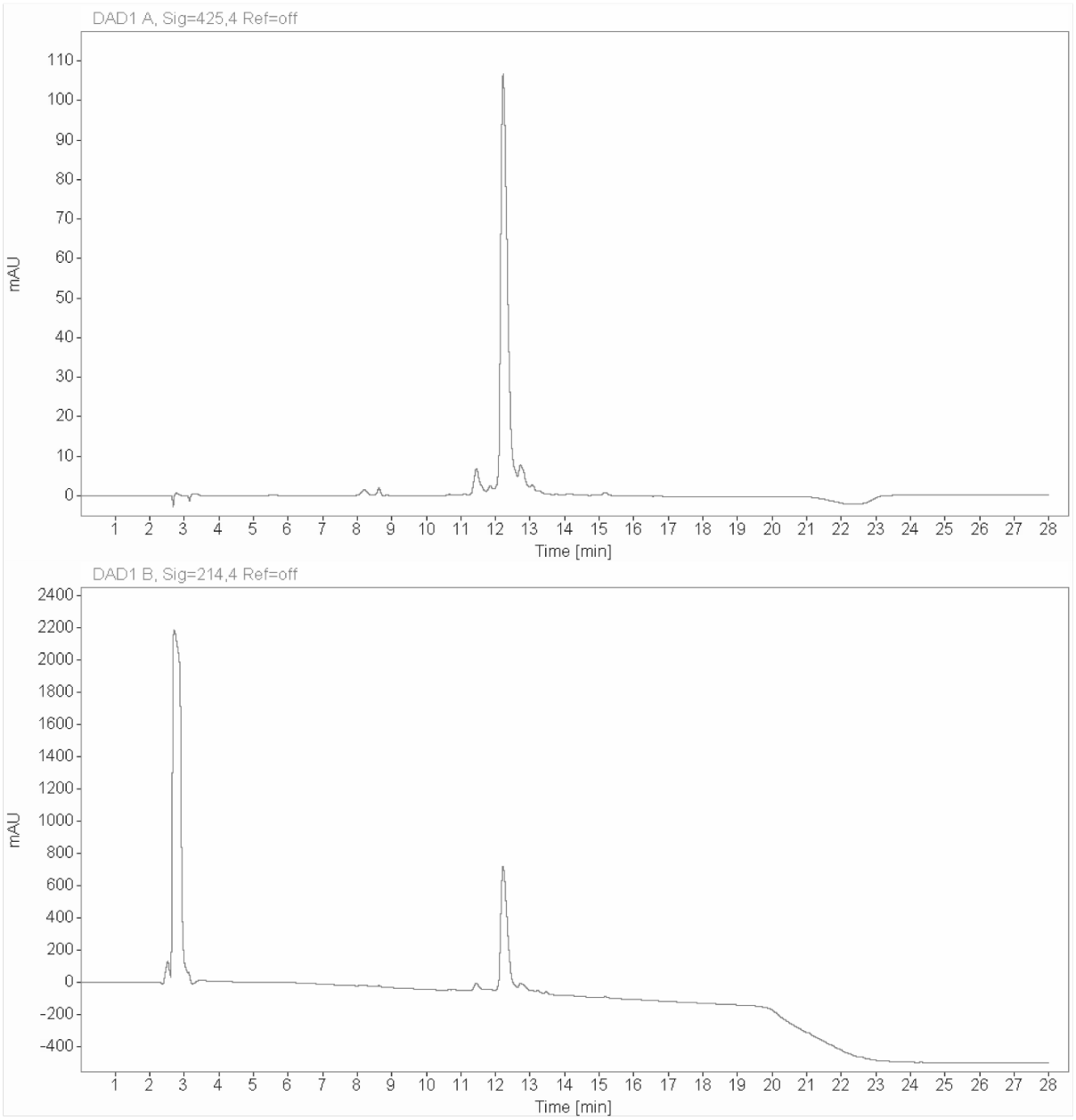
Analytical HPLC Chromatogram of Flo-VP16(467-488) F473A



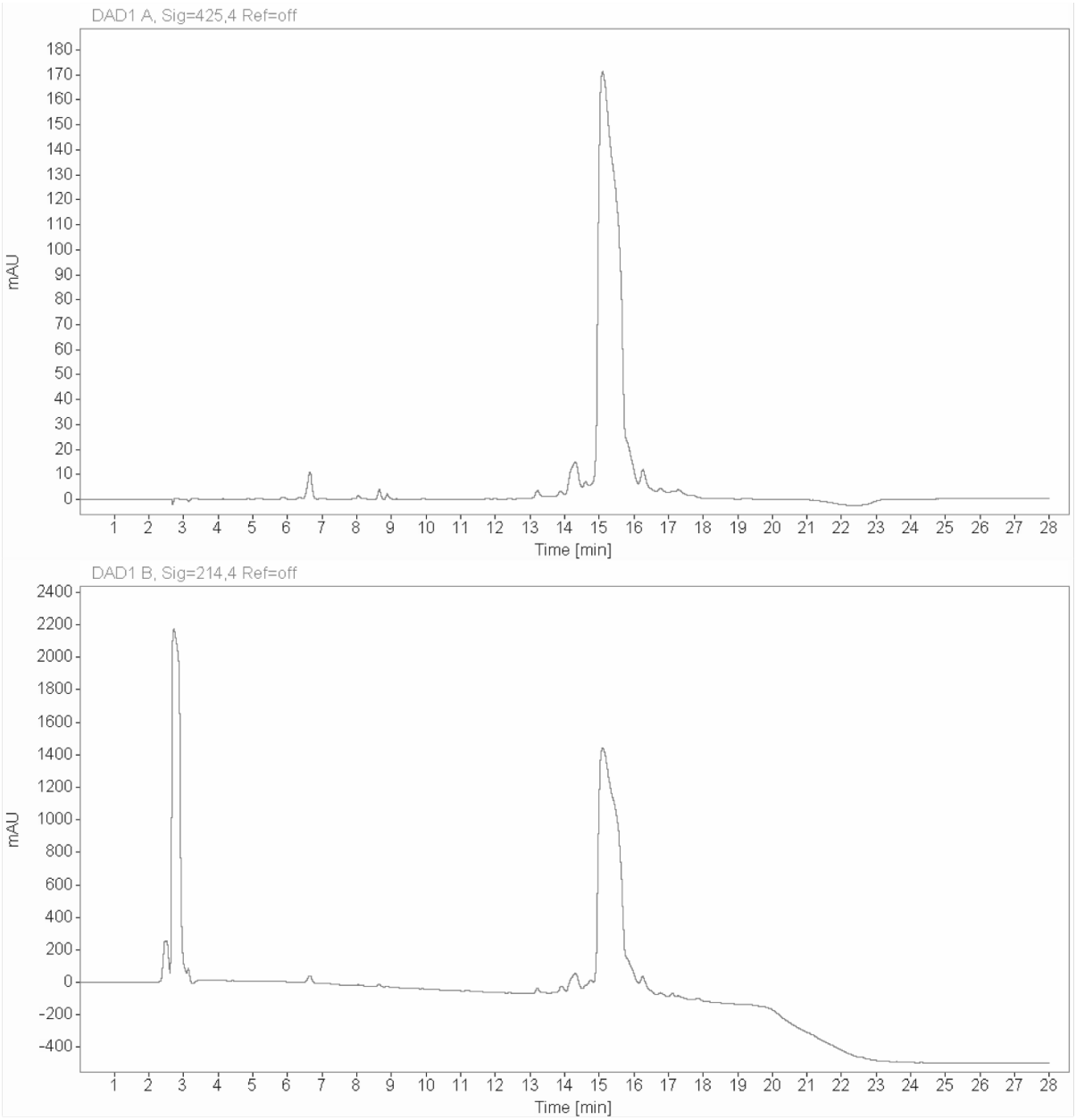
Analytical HPLC Chromatogram of Flo-VP16(467-488) E474A



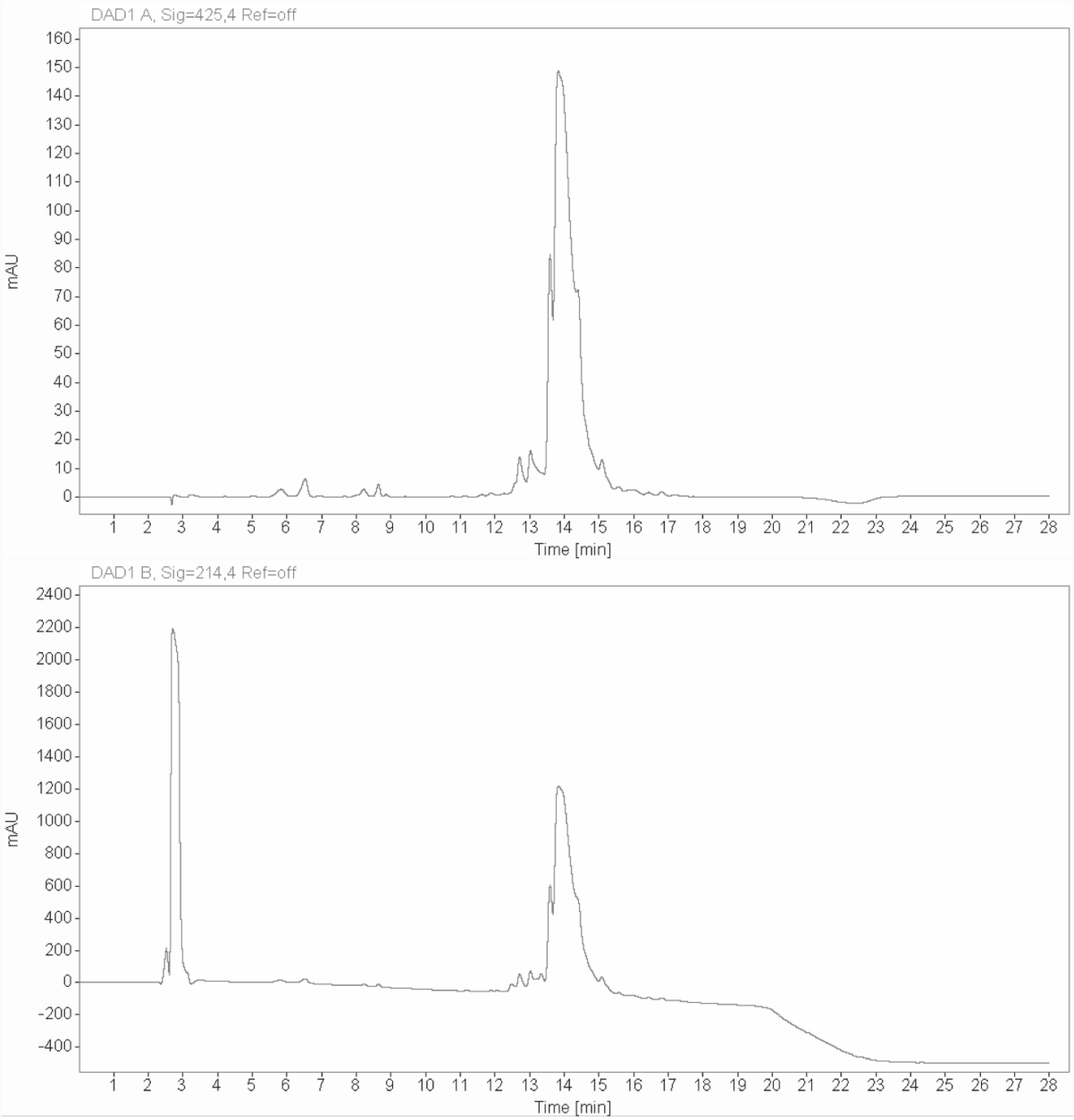
Analytical HPLC Chromatogram of Flo-VP16(467-488) F475A



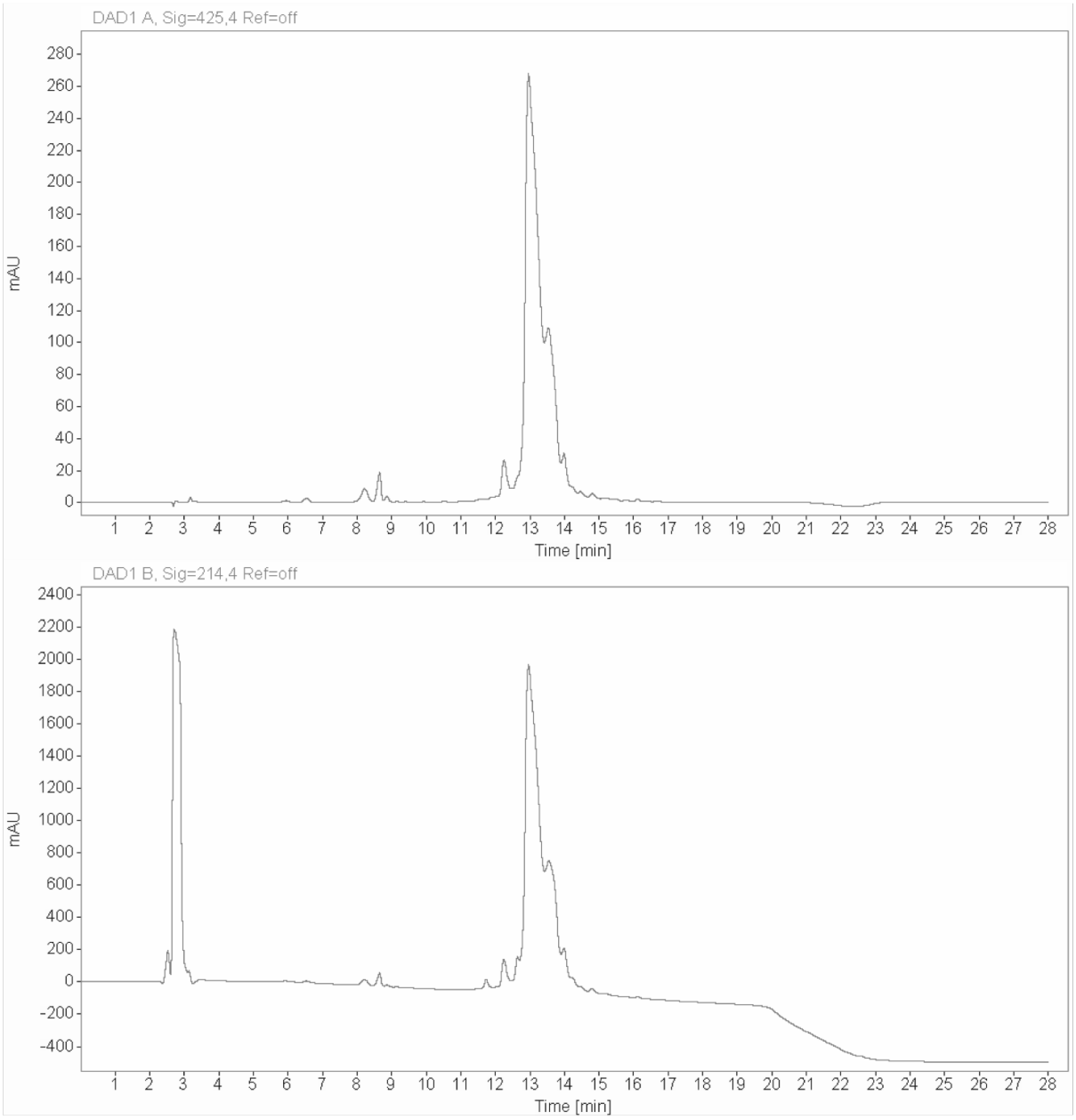
Analytical HPLC Chromatogram of Flo-VP16(467-488) E476A



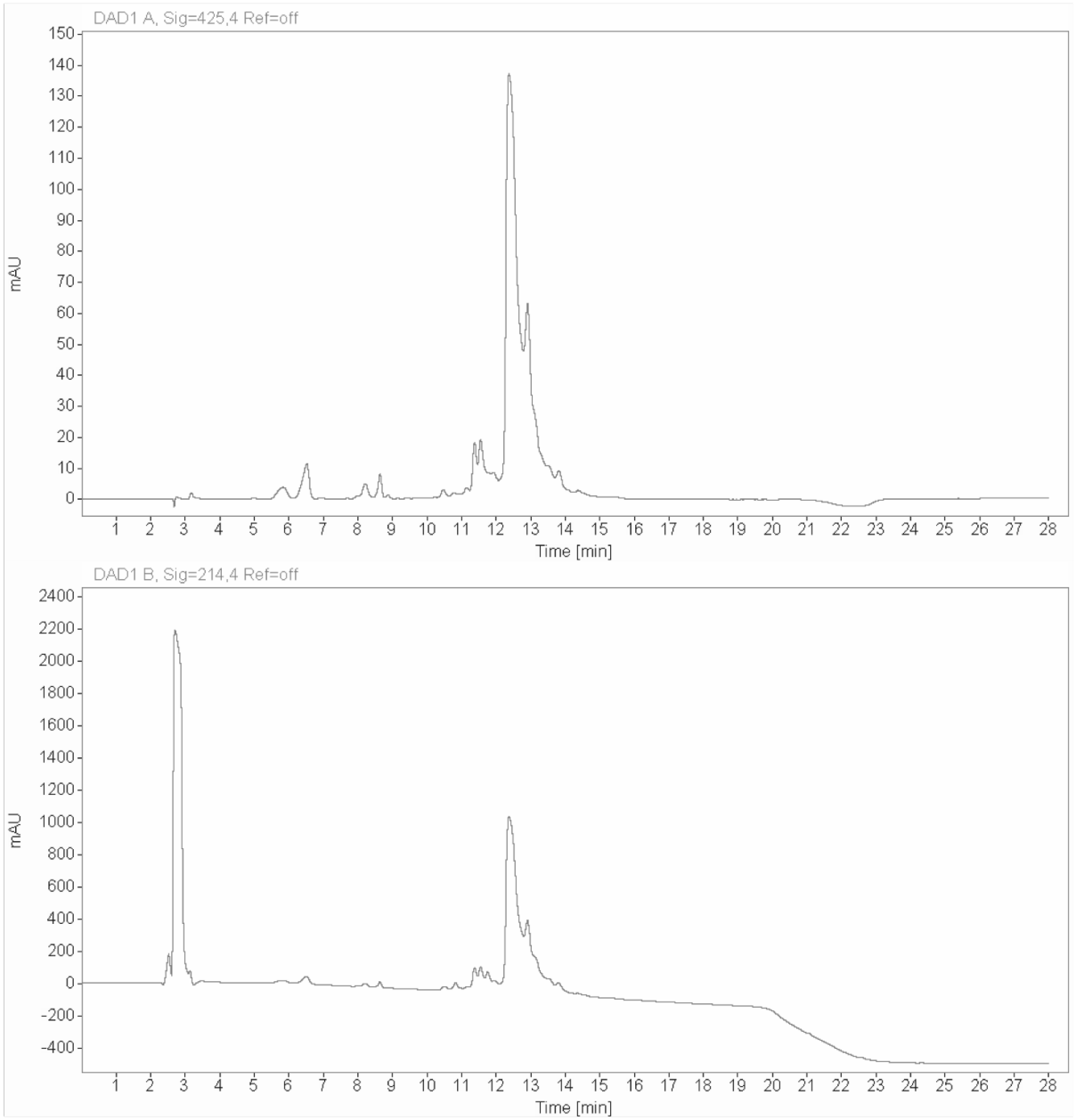
Analytical HPLC Chromatogram of Flo-VP16(467-488) Q477A



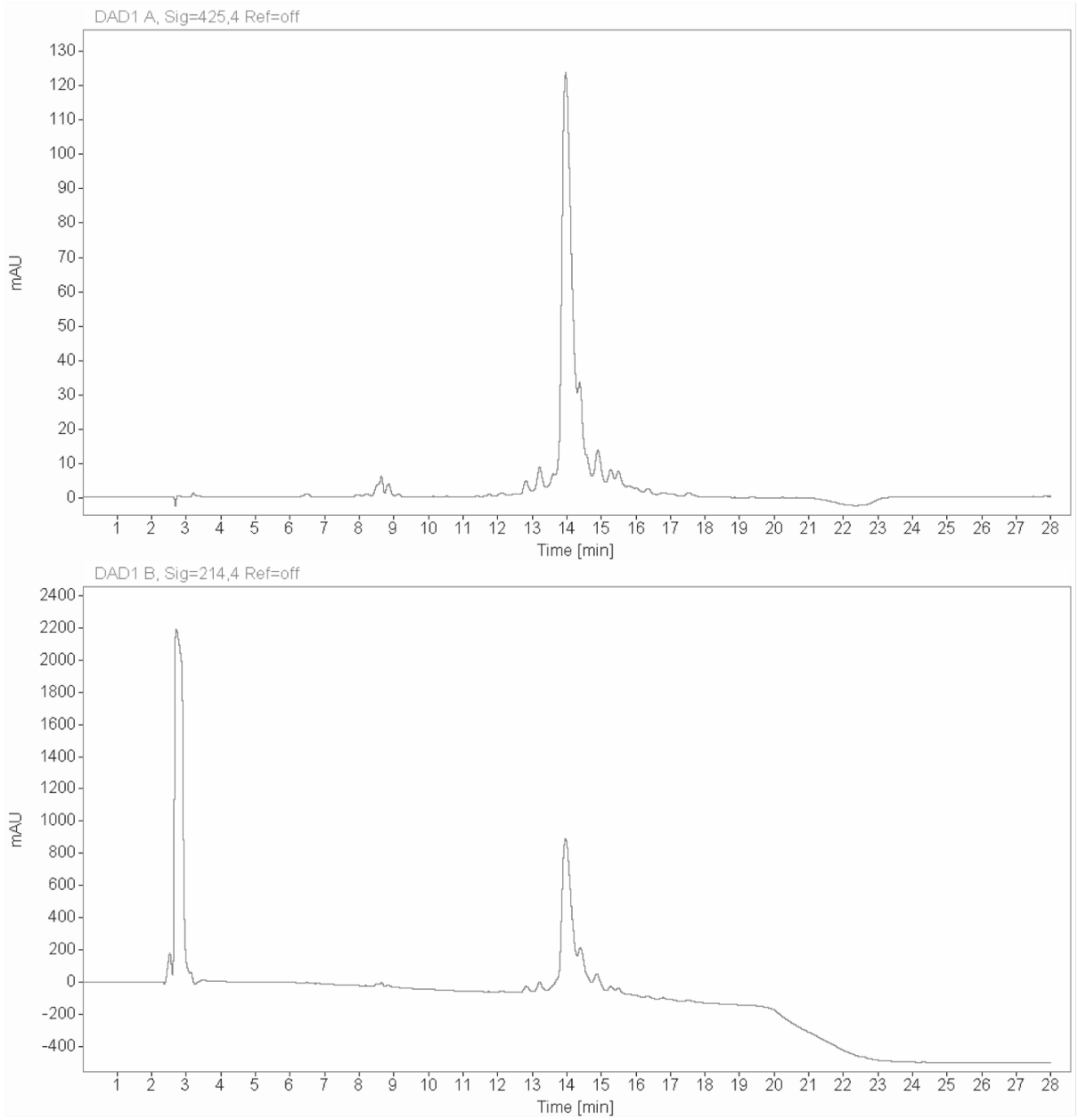
Analytical HPLC Chromatogram of Flo-VP16(467-488) M478A



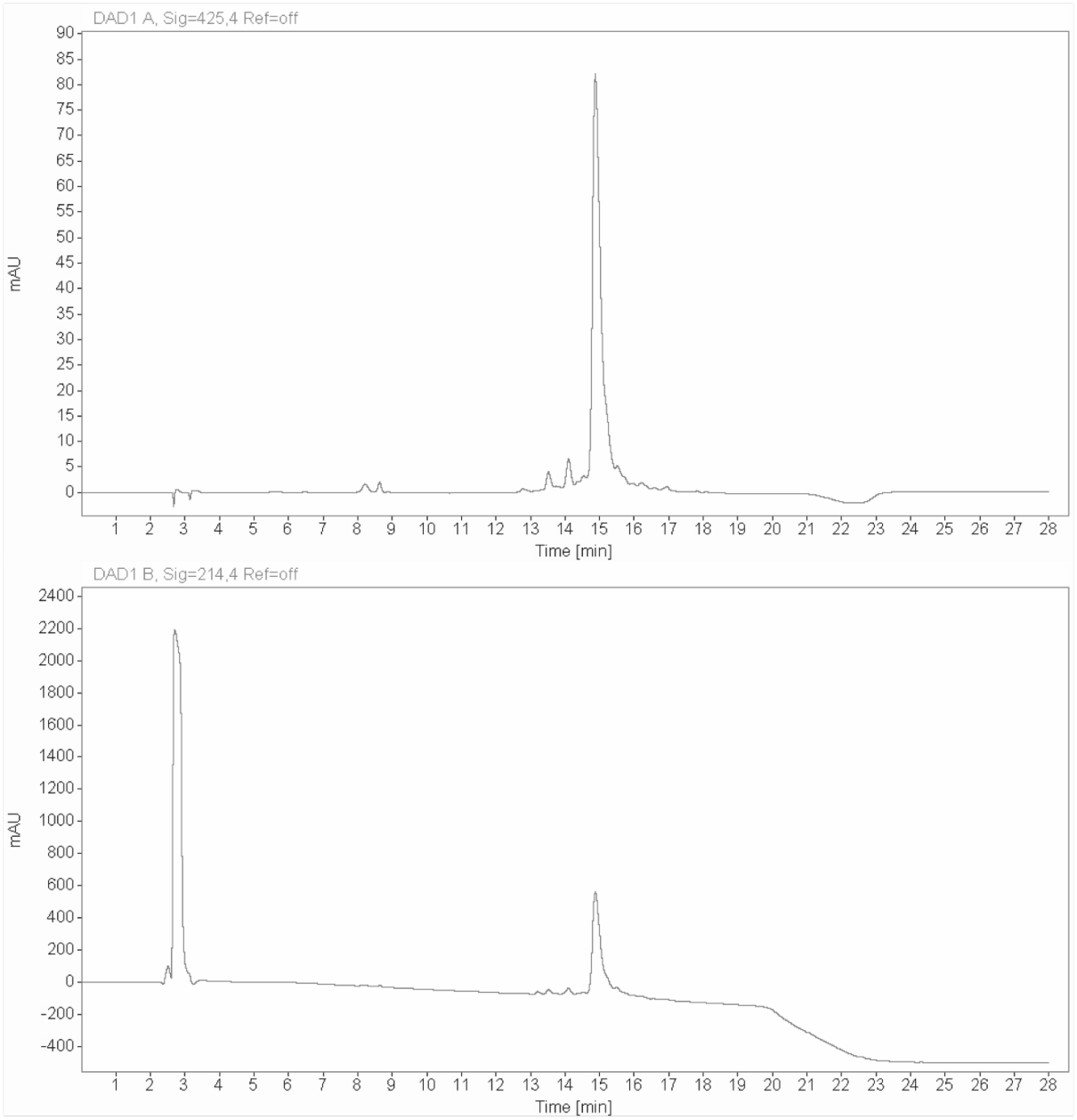
Analytical HPLC Chromatogram of Flo-VP16(467-488) F479A



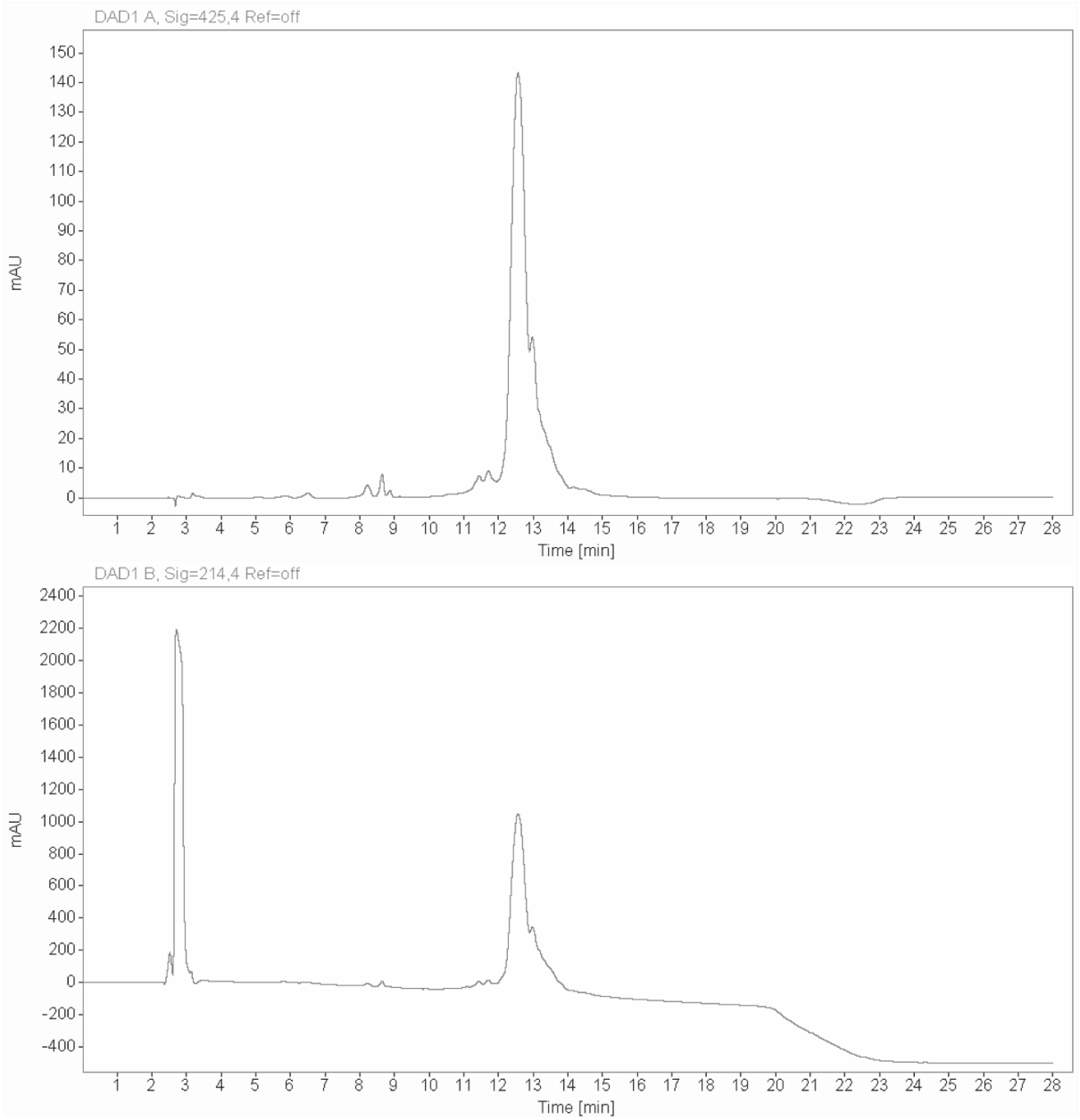
Analytical HPLC Chromatogram of Flo-VP16(467-488) T480A



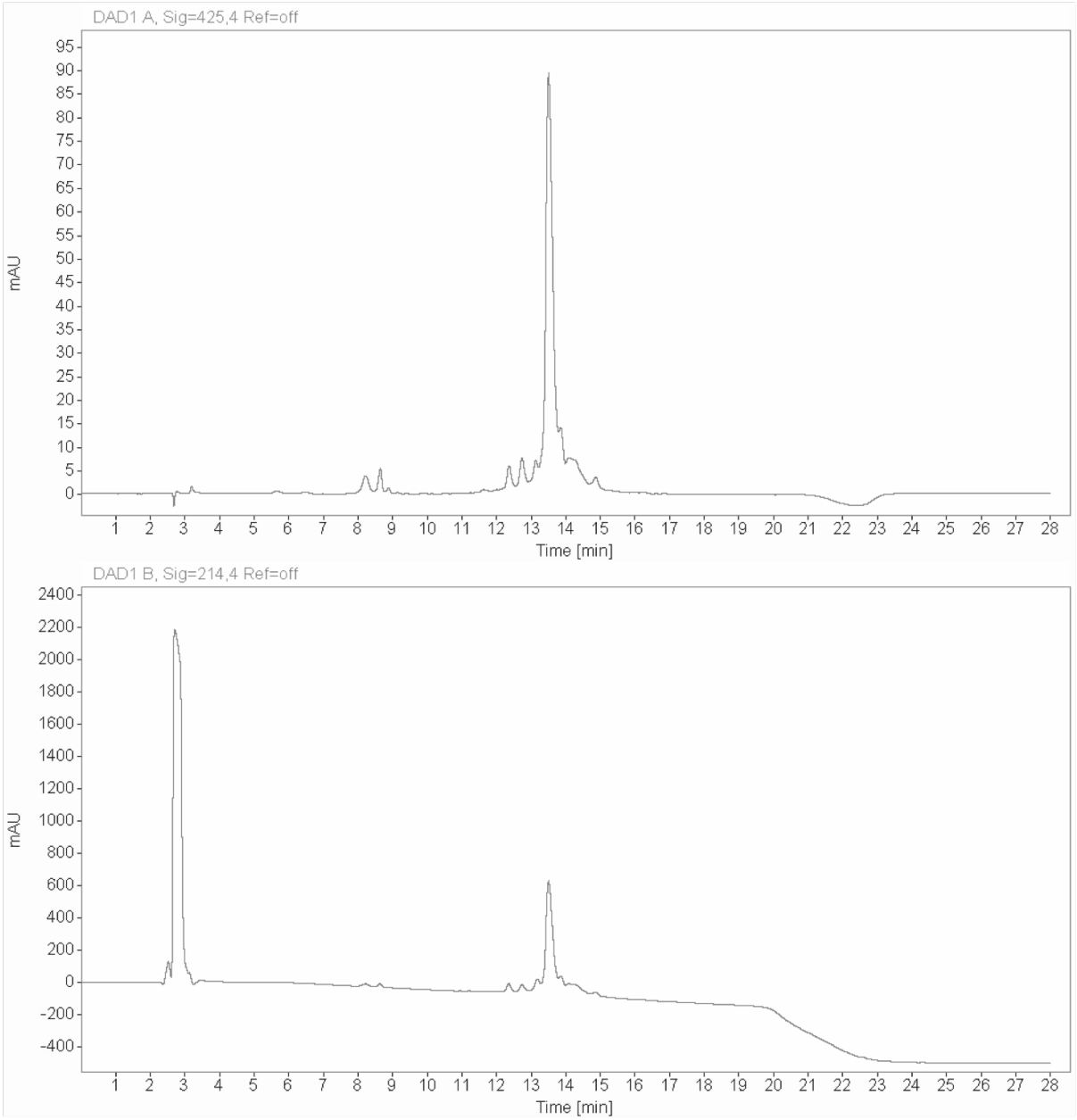
Analytical HPLC Chromatogram of Flo-VP16(467-488) D481A



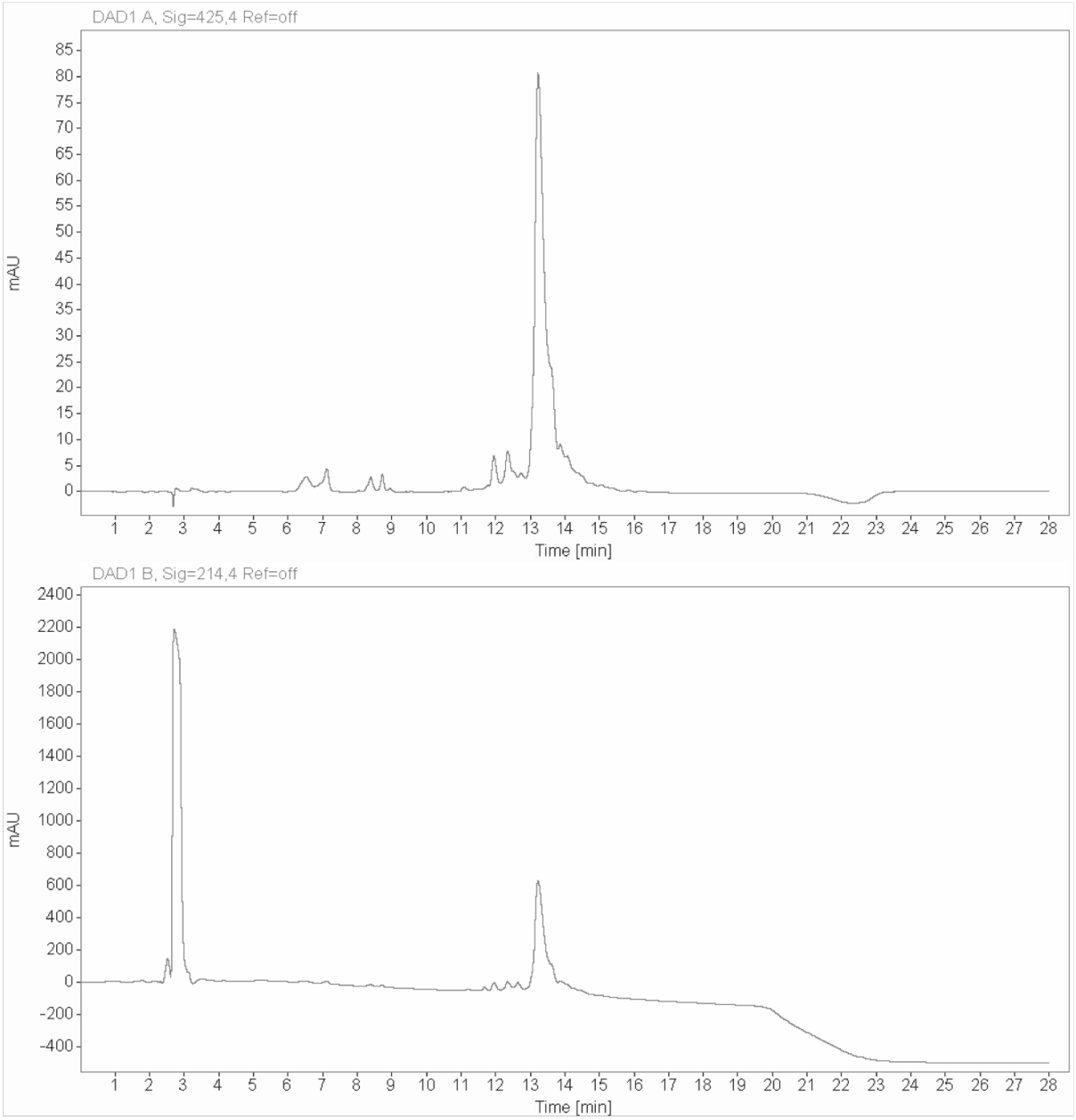
Analytical HPLC Chromatogram of Flo-VP16(467-488) L483A



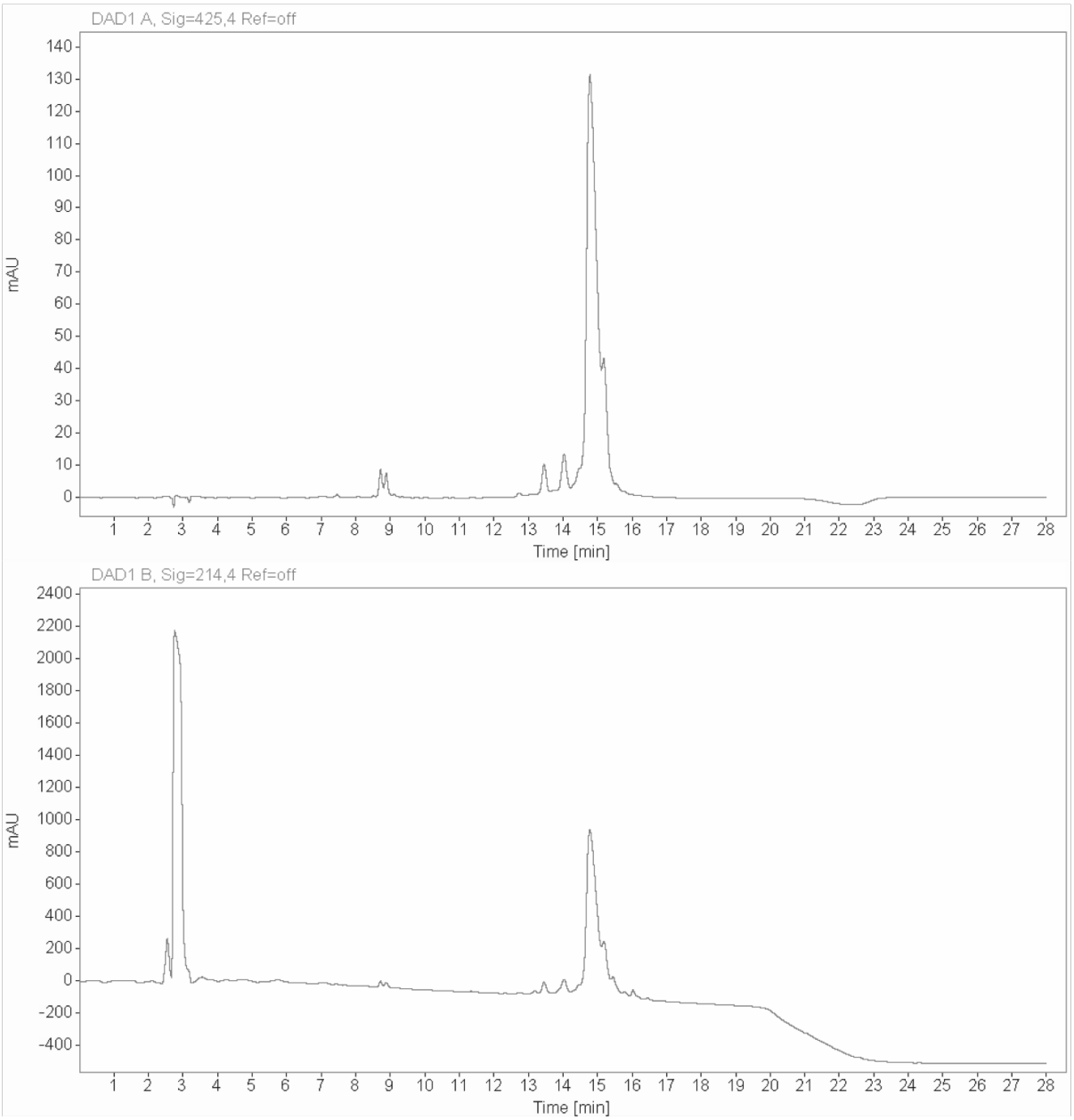
Analytical HPLC Chromatogram of Flo-VP16(467-488) G484A



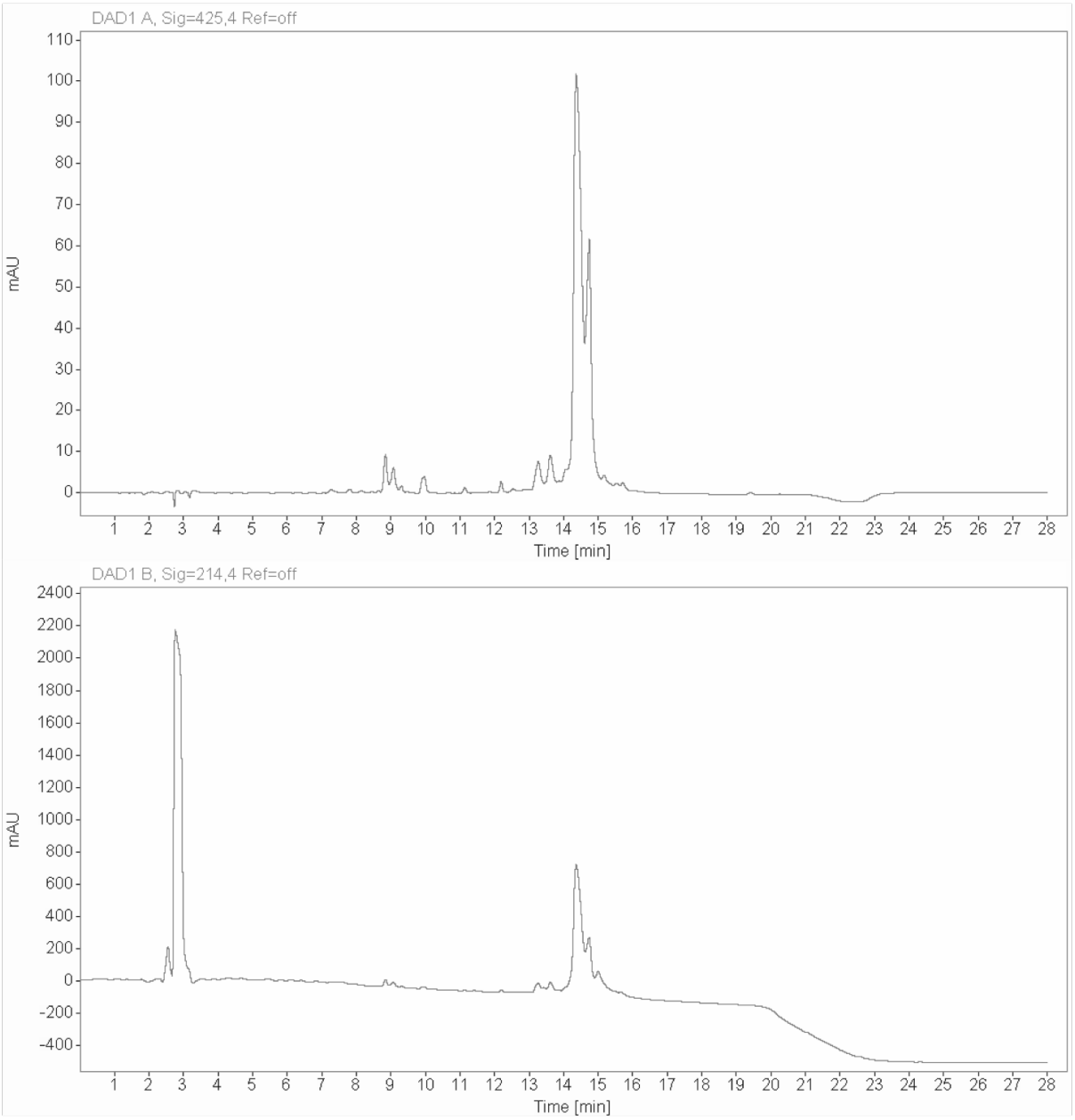
Analytical HPLC Chromatogram of Flo-VP16(467-488) I485A



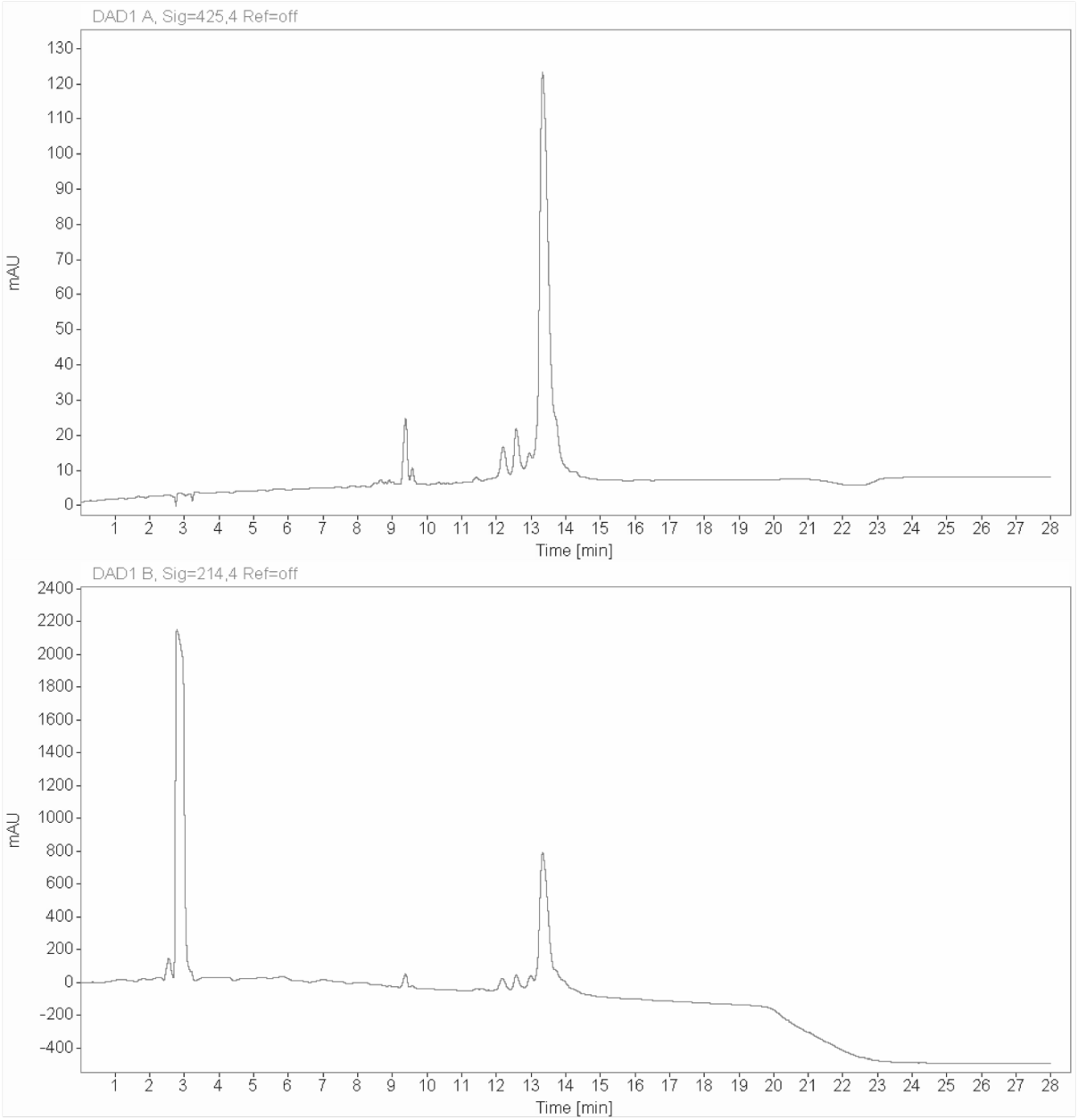
Analytical HPLC Chromatogram of Flo-VP16(467-488) D486A



Analytical HPLC Chromatogram of Flo-VP16(467-488) E487A



Analytical HPLC Chromatogram of Flo-VP16(467-488) Y488A



APPENDIX B

List of Confirmed Extracts from NPE Library Screen of ERM•AcID Interaction

The full list of 332 extracts identified as validated hits of the ERM•AcID interaction following hit filtering and counter screens is presented below. The criteria used to select these hits is reviewed in **Figure B.1**. Those strains producing extracts with activity against the interaction that were selected for regrowth are highlighted in green in the table of hits.

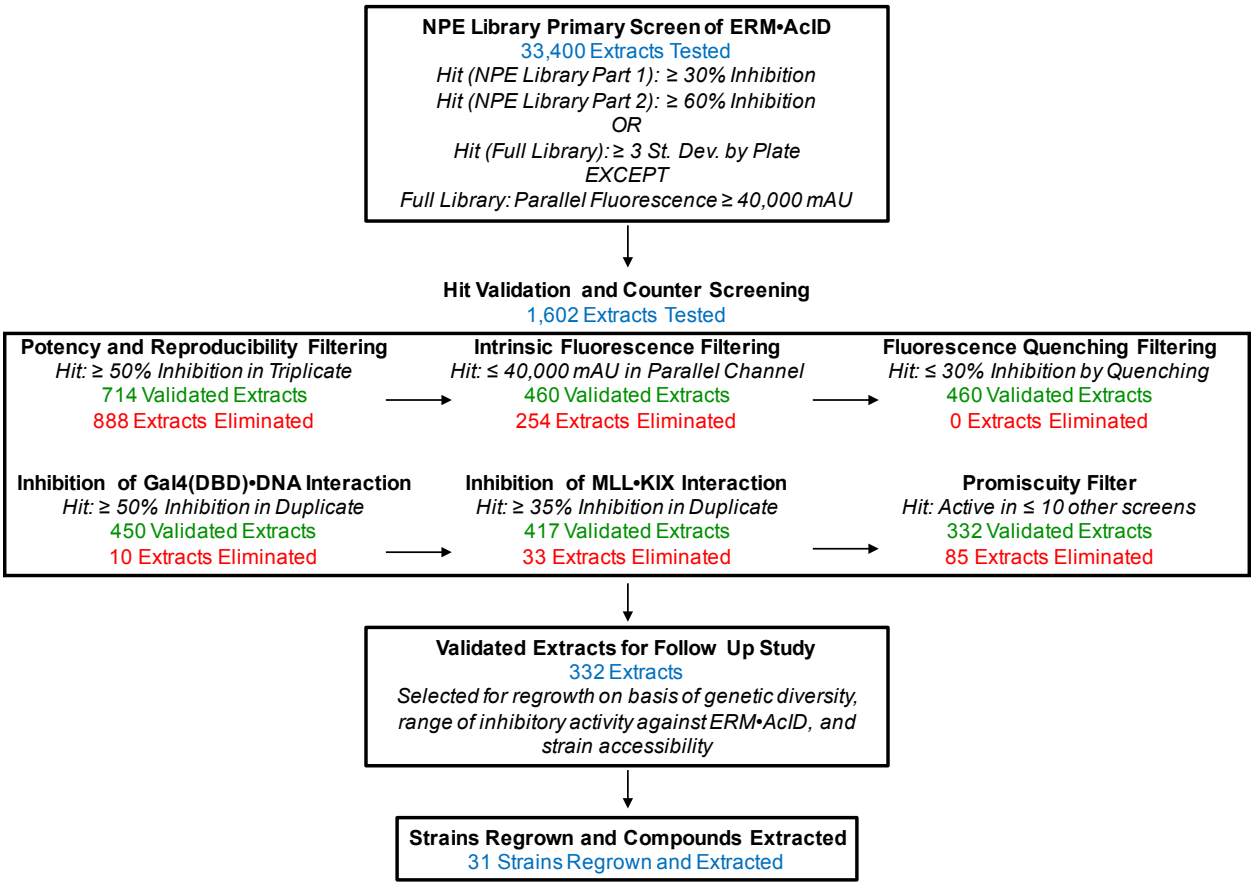


***Figure B.1*- Schematic Representation of Final Hit Selection** The full battery of screens and counter screens utilized to identify natural product extracts containing small molecule inhibitors of the ERM•AcID interaction is represented schematically.

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	Fluorescence Quenching	Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
1	7719-N10I	SID-9822	70.8	29645	-10.8	-65.2	7.9	9
2	82389-N2N	SID-69812	73.6	35359	-3.8	-45	0.9	7
3	82389-N1Z	SID-69741	52.5	34149	15.8	-69.7	4.3	5
4	65370-N4Z	SID-69717	58	37181	-11.3	-79.5	7	6
5	74494-N1I	SID-69497	98.8	37394	-20.2	-7.7	12.4	6
6	72479-N1I	SID-69475	95.7	37398	-16.2	0.5	10.2	6
7	56390-N1N	SID-69452	65.2	37029	-15.2	-20.9	9.3	4
8	73332N	SID-69427	57.2	36476	-11.2	-30.7	5	3
9	78816I	SID-69354	68.1	40306	-25.3	-84.8	5.8	4
10	72509-N1N	SID-69284	73.3	35954	-11.7	-26.4	2.9	8
11	74424-N1N	SID-69252	72.8	36858	-16	-4.4	7	7
12	74499-N1N	SID-69195	59.5	36919	-11.2	-27.7	7.1	4
13	73260N	SID-69113	80.9	38557	-17.9	-86	2.8	5
14	**71961N**	**SID-69043**	**70**	**38711**	**-14.3**	**-53.4**	**3.4**	**5**
15	78799I	SID-68915	58.4	39716	-17	-58.5	3.2	5
16	78921N	SID-68879	67.9	39940	-21.4	-70.9	3.2	8
17	65435-N4N	SID-68856	67.7	37994	-10.5	-21.4	3.1	8
18	**55270-N4N**	**SID-68851**	**73.7**	**39860**	**-18.4**	**-60.5**	**2.8**	**4**
19	65395-N6I	SID-68834	62.6	38301	-7.7	-80.4	-3.5	4
20	65409-1N	SID-68805	95.6	29690	14.9	59.3	12.9	8
21	55270-N4I	SID-68785	66.3	35206	22.9	-76.3	-0.7	4
22	67313-1I	SID-68780	55.6	38657	-6.6	-82.6	-7.4	6
23	49385-2I	SID-68717	74	38133	-13.8	-69.4	9.4	6
24	**67360-N9N**	**SID-68685**	**70.1**	**35517**	**-3.6**	**-18**	**11.4**	**4**
25	65405-N9I	SID-68681	64.5	38184	-12	10.4	14.9	7
26	49595-N4I	SID-68651	62.6	38391	-21	-24.4	2.8	4
27	71839N	SID-68620	65.9	39590	-22.5	1.9	11.5	9
28	78761I	SID-68593	61.6	38808	-22.7	-78.2	-1.7	6
29	73293N	SID-68571	65	34465	-3	-56.7	7.8	5
30	78928N	SID-68570	71	34014	-3.3	-10.7	8	6
31	78935I	SID-68568	71.5	34390	-7.7	-7.4	15.1	7
32	53116I	SID-68527	65.7	37550	-20	-63.9	0.1	4
33	71885N	SID-68498	55.8	35231	-9.7	-21.1	4.4	4
34	71777N	SID-68492	58.6	37670	-14.7	19.9	12.6	8
35	73413I	SID-68482	58.7	36678	-9.7	-70.3	-0.3	5
36	71533I	SID-68473	59.8	37616	-17.5	-16.6	0.4	4
37	73397N	SID-68466	57.7	34244	-2.9	-41.2	6.9	5
38	71885I	SID-68465	80.5	35714	-6.7	-57.2	10.5	4
39	73239I	SID-68460	57.3	38188	-14.9	-47.3	0.5	6
40	73370N	SID-68455	77.7	35096	-1.8	-63.3	7.2	4
41	71953I	SID-68435	62.9	32984	-6.9	-5.9	5.9	6
42	73273I	SID-68420	64.1	38588	-18	-62.3	2.6	6
43	73276I	SID-68419	73.7	39264	-23.7	-109.3	-0.5	5
44	53086N	SID-68373	74	38543	-23.9	-96.3	0.5	7
45	53086I	SID-68364	78	36184	-13	-58.4	1.6	6
46	71899I	SID-68355	56.2	39465	-20.5	16.9	9.8	5
47	53088I	SID-68336	82	34562	0.6	-34.8	-16.1	5
48	71757I	SID-68315	56.9	39540	-26.8	-64.3	2	4
49	53072I	SID-68311	59.4	40129	-28.6	-73.1	1.7	3
50	58140N	SID-68308	62.8	38897	-24.2	-28.1	4.7	4
51	73393N	SID-68300	64.3	37647	-17.4	-38.2	5.1	4
52	71835I	SID-68294	53.5	36700	-15.1	-14	1.4	4
53	58146I	SID-68255	72.6	32108	-0.8	-3	10.7	10
54	65455-N2I	SID-68244	86.8	36981	-18.9	-15.1	8	9
55	65465-N1I	SID-68220	61.4	40258	-28.6	-45.6	0.1	5
56	52249-1I	SID-68214	95.1	34082	-4.6	14.1	18.3	10
57	64665-N3I	SID-68213	85.6	37601	-17.6	-58.6	4.2	9
58	34395-1N	SID-68209	51.8	33675	-1.3	-21.3	9.3	8
59	67325-N7I	SID-68206	62.4	39962	-28.6	-26.7	2.1	6
60	34376-2I	SID-68198	75.4	35631	-12.2	-49	7.1	10
61	49830-N4I	SID-68174	58.9	39213	-24.1	-86.4	-3.7	4
62	49546-N4N	SID-68163	61.5	36647	-4	-37.3	2.5	4
63	41429-N1I	SID-68149	70.7	39078	-18.1	-42.7	8.5	4
64	**41429-N1I**	**SID-68148**	**82.3**	**39516**	**-28.6**	**-85.4**	**4.1**	**4**
65	65440-N1I	SID-68128	79.1	37997	-18.1	5.8	16.9	9
66	**65440-N1N**	**SID-68121**	**70**	**36498**	**-10.3**	**-34.5**	**6.5**	**3**
67	64820-N5I	SID-68101	62.8	39742	-17.4	-34.1	3.4	5

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	% Fluorescence Quenching	% Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
68	**49546-N4I**	**SID-68085**	**52**	**39949**	**-14.7**	**-52.7**	**-4.3**	**8**
69	71689N	SID-68050	51.4	38786	-18	-3.3	5.4	4
70	58132N	SID-67958	59.8	36237	-9.7	-17.2	0	5
71	58186I	SID-67946	57.1	37825	-12.8	-99.7	-13.7	3
72	58092I	SID-67943	54.8	39334	-16.2	-92.1	-13.5	4
73	71813I	SID-67893	57.7	37950	-13.9	-70.3	-12.3	3
74	71579I	SID-67887	57	38948	-10.3	-75.9	-11.1	3
75	71603I	SID-67857	62	39216	-13.6	-84.6	-8.7	4
76	71865I	SID-67795	56	38314	-9.4	-92.9	-11.7	7
77	71577I	SID-67700	52.4	38773	-12	-80.6	-10	6
78	AM82677	SID-67401	65.1	34493	7.8	13.2	9.3	9
79	AM82597	SID-67356	62	36839	-8.5	9	2	4
80	AM82597	SID-67355	73	36443	-20.9	-66.8	7.6	5
81	AM82597	SID-67348	66.2	37678	-30.9	-117.7	8.8	6
82	AM77835	SID-35025	55.1	36011	-13	-150.7	0.7	6
83	AM77833	SID-35023	65.2	37123	-15.8	-140.4	-1.7	7
84	AM77794	SID-34984	54.2	38736	-27.5	-24.6	7.4	6
85	AM77733	SID-34923	72.5	37076	-4.8	-117.3	1.6	6
86	AM77732	SID-34922	68.5	36303	-1.4	-69.3	5.6	5
87	MS77698	SID-34888	63.7	34401	-11.1	-80.1	7.7	8
88	AM77658	SID-34848	82.6	37541	-22	16.8	13.5	10
89	AM77491	SID-34681	78.6	34422	-12.2	-5.7	19.8	6
90	AM77419	SID-34609	55.5	33375	-4.6	2	4.3	6
91	AM77414	SID-34604	83.3	34437	-2.5	-67	11.2	9
92	AM77412	SID-34602	65.9	35794	22.8	-78.8	7.5	9
93	AM77348	SID-34538	69.3	35310	-16.9	-17.6	14.4	7
94	MS75409	SID-34518	92.5	39420	-30.9	-30.5	15.8	8
95	MS75388	SID-34497	50.6	33638	10.9	-104.1	-0.4	7
96	MS75258	SID-34367	65.6	33327	-1.2	-14.6	10.3	8
97	MS75180	SID-34289	93.5	32737	3.6	-5.9	19.9	5
98	MS72144	SID-34022	60.3	38120	-16.2	-1.8	8.8	7
99	MS72040	SID-33918	59.3	38568	-19.3	-106.2	3.1	7
100	MS72036	SID-33914	52	38734	-22.6	7.2	11.3	6
101	MS72018	SID-33896	72.5	37295	-24.6	-41	-0.9	7
102	58222N	SID-33845	77.3	35635	-6.9	-39.9	4.7	7
103	68179N	SID-33795	54.8	33240	-2.9	-14.3	-10.4	4
104	58236I	SID-33783	74.1	37049	-16.7	-69.1	2.1	9
105	58236I	SID-33782	72.3	38062	-12.1	-74.1	8.3	6
106	58238N	SID-33753	62.7	37595	-9.2	-59.1	0.5	7
107	68161I	SID-33736	61.5	39012	-16.9	-74.8	-2.2	6
108	54925-N1I	SID-33698	85.7	35780	-12.3	-88.4	4.8	9
109	54925-N1N	SID-33685	81.7	38238	-14.9	-118.8	3.2	10
110	70189-C1I	SID-33682	70.4	38462	-18	-86.8	1	5
111	41361-1I	SID-33611	58.8	39205	-22.9	-78.2	-0.2	5
112	49800-N2N	SID-33599	90.7	39067	-18	-37.5	13.2	9
113	49800-N2I	SID-33594	69	38885	-17.4	-29.7	7.5	9
114	70189-C1N	SID-33544	53.8	35037	-2.6	-23.1	5.1	6
115	70189-C1N	SID-33543	65.8	35907	-8.1	-11.1	4.8	7
116	41445-N2I	SID-33360	64.1	35115	19.5	-14.1	8.5	8
117	41445-N2I	SID-33359	79.6	39192	-24.6	-50.2	4.8	8
118	41445-N2N	SID-33326	74.8	37234	-18.4	-20.4	5.3	6
119	41445-N3N	SID-33291	62.7	36068	-8.5	-12.9	4.9	5
120	52315-N6N	SID-33285	65.6	36158	-8.1	-24.6	3.4	7
121	64825-N5AIA	SID-33215	74.7	38527	-22.4	-29.1	9	9
122	**52245-N2N**	**SID-33188**	**101.7**	**35559**	**-18.2**	**11.9**	**20**	**5**
123	52245-N2N	SID-33187	65.5	38902	-20.8	-26.6	6.2	5
124	50247-1I	SID-33181	63.7	39824	-19.8	-93.1	-3.4	4
125	52315-N6I	SID-33179	66.2	34796	8	-10.2	4	10
126	52235-N2N	SID-33146	58	35093	-9.9	-20.7	-1.6	8
127	44293-N1I	SID-33138	86.1	37951	-23.3	-38	9.4	9
128	**44293-N1N**	**SID-33128**	**77.8**	**38368**	**-21.9**	**-90.4**	**-0.9**	**8**
129	41454-AC3N	SID-33066	67	34710	-8.6	-17.2	-6.7	4
130	5538-A2N	SID-33058	75.9	35501	-12.5	-66	6	8
131	**5538-A2N**	**SID-33057**	**86.2**	**38027**	**-19**	**-54.9**	**7.1**	**7**
132	54875-N1N	SID-33042	59.7	38680	-22.7	-45.3	-1.3	8
133	65430-N16N	SID-33027	69.1	38758	-20.6	3.7	6.5	8
134	34365-A1N	SID-33026	60.6	34899	-8.4	-45.4	4.1	6

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	% Fluorescence Quenching	% Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
135	65435-N4N	SID-32992	65.7	36589	-16.2	-8.6	1.9	8
136	41450-N4N	SID-32936	73.2	32329	-1.1	-23.8	-3.5	10
137	52260-N4AN	SID-32929	76.5	32279	5.8	-29.7	-2	6
138	64825-N5N	SID-32926	77.6	35565	-5.2	-34	1	10
139	54915-N1I	SID-32858	71	37749	-8.5	-96.2	-7	9
140	41450-N4I	SID-32808	84.9	39633	-13.9	7.5	6.6	8
141	64825-N3N	SID-32737	56.9	38832	-13.6	-45.4	6.6	4
142	64825-N3I	SID-32667	60.4	40523	-13.9	-63.5	4.2	7
143	41435-N1I	SID-32595	69.2	39075	-13.9	-34.6	6.2	8
144	34389-1I	SID-32497	68.2	39920	-13.9	-83.8	6.9	8
145	**PAN101-7I**	**SID-32305**	**74.1**	**38049**	**-20.4**	**-88.2**	**9.1**	**5**
146	PAN91-1I	SID-32295	61.3	37471	-8.7	-48.3	4.7	6
147	**68157N**	**SID-32195**	**55.1**	**36047**	**-3.1**	**-20.7**	**-1**	**3**
148	68217I	SID-32183	62.3	39238	-13.9	-51.4	-6.6	7
149	68175N	SID-32061	61.9	38570	-13.9	-7.2	4.8	6
150	18112-N8N	SID-31842	53.6	38368	-13.3	-50.6	-2.6	4
151	**36180-2N**	**SID-31830**	**57.4**	**37403**	**-10**	**-16.4**	**3.9**	**4**
152	41374-AC5N	SID-31827	61.7	38496	-13.9	-39.1	2.2	6
153	52328-4N	SID-31753	85.3	35035	-4.1	25.3	10.8	10
154	32240-1I	SID-31550	64.6	38317	-7.9	-14.1	6.6	6
155	34948-A3I	SID-31544	68.2	35328	-8.2	-7.7	9.3	8
156	41422-AC3I	SID-31541	65.1	38596	-12.5	-20.1	3.6	8
157	**41445-N3I**	**SID-31523**	**80.4**	**39056**	**-13.9**	**-28.7**	**13.3**	**8**
158	32424-H2I	SID-31484	53.1	38247	-13.9	-39.1	5.4	6
159	06-284-3I	SID-31395	63.8	35946	-13.6	-44.5	4.1	5
160	06-284-3N	SID-31377	52.3	35944	-10.5	12.1	2.4	5
161	06-282-1N	SID-31351	56.6	32033	-6.3	-12.6	8.9	10
162	06-392-3I	SID-31335	63.9	39812	-21.9	-47.3	0.3	4
163	06-226-H2I	SID-31323	55.9	38013	-15.5	-71.9	-0.9	5
164	06-269-H2N	SID-31303	62.8	35566	-13	-72.7	-0.6	4
165	52295-N2I	SID-28739	67.5	36661	-13.7	-27.4	5.2	10
166	58069N	SID-28433	59	38961	-10.9	-3.3	9.2	7
167	58241I	SID-28408	54.7	39178	-13.9	10.7	6.7	8
168	58241N	SID-28400	57.3	35823	-10.6	3.6	2.1	8
169	58235I	SID-28392	67.2	36500	-4.1	28.1	7.6	9
170	58235I	SID-28391	68.7	39028	-13.9	34.8	6.2	10
171	58195I	SID-28073	84.5	38044	-4.1	-22.9	11.7	8
172	**39040-1I**	**SID-27645**	**64.4**	**39972**	**-15.1**	**-68.2**	**1.7**	**6**
173	41392-AC5I	SID-27633	57.8	37222	-5	-34	7.5	10
174	5746-A9I	SID-27504	65.5	36663	-7.2	-53.5	4	7
175	MS60767	SID-27424	64.5	36975	-11.3	-19	8.1	7
176	MS60665	SID-27322	68.5	38728	-13.7	-104.1	6.8	10
177	MS60652	SID-27309	55.2	40536	-16.7	-66	6.3	9
178	MS60592	SID-27249	69.4	38904	-11.7	-64.9	13.2	9
179	36284-N18N	SID-26726	56.1	37554	-8.3	-11.6	6.1	10
180	41426-A4I	SID-26608	64.7	39542	-17.8	-51.3	11.3	8
181	18141-3N	SID-26421	85.7	39419	-17.1	-10.4	13.1	9
182	32350-1I	SID-24794	69.6	37778	-8.3	-55.8	5.6	9
183	34946-N13I	SID-24274	57.9	39147	-16.8	-65.8	8.3	10
184	41392-MH5I	SID-24167	52.9	36158	-6.5	-13.6	-8	5
185	MS38947	SID-20463	57.1	35924	-10.5	-52.7	7.1	10
186	MS38908	SID-20424	76.4	36876	-19.6	-139.4	-0.5	10
187	MS38768	SID-20284	53	34145	-2	-19.8	8.3	10
188	6425-L1I	SID-18508	51.5	35979	-9.7	1.9	2.6	10
189	**20731-H2I**	**SID-18195**	**77**	**34656**	**-5.9**	**-4.5**	**2.2**	**10**
190	AM86311	SID-143669	75.2	36327	-9.2	-126	8.6	5
191	AM86308	SID-143666	56.3	36334	-9.4	-64.6	5.4	3
192	AM88563	SID-143573	102.3	37526	4.8	39.3	23.4	6
193	AM88558	SID-143568	99.7	37519	5.6	-42	-8.5	3
194	AM96229	SID-143450	68	36764	4.8	-61.9	1.4	6
195	AM96213	SID-143434	53.9	37930	-4.4	-50.7	-2.3	3
196	AM96190	SID-143411	51.6	30442	-0.1	32.7	-0.1	3
197	MS96181	SID-143402	76.5	34958	-4	-17.5	10.3	5
198	FU96134	SID-143355	58.3	34196	13.9	-38.5	8.7	3
199	AM96065	SID-143286	54	32424	-8.1	6.1	9	2
200	FU95982	SID-143203	69.3	38172	-18.4	-37.9	3.3	4
201	FU95864	SID-143085	55.9	37467	-12.2	-37.3	8	3

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	% Fluorescence Quenching	% Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
202	78950C	SID-143079	62.5	35298	6.5	-37.9	8.2	3
203	5644-N7C	SID-143077	62.6	36207	-12.1	-53.3	5.7	3
204	9762-N2R	SID-143067	59.7	36968	-9	-42.4	8.3	6
205	**12587-2Z**	**SID-143065**	**83.2**	**37498**	**3**	**-32.8**	**9.1**	**5**
206	**78874R**	**SID-143056**	**65.5**	**36958**	**-6.5**	**-10.1**	**9.9**	**3**
207	9848-1R	SID-143048	81.9	37493	-10.3	-77.8	10.5	5
208	**86815-N2Z**	**SID-143008**	**102.5**	**38773**	**-8.9**	**-19.8**	**40.1**	**4**
209	**06131R**	**SID-142998**	**56.3**	**35430**	**-6.1**	**-47.7**	**-7.9**	**2**
210	**91085R**	**SID-142997**	**85.2**	**35887**	**-8.6**	**30.6**	**7.1**	**4**
211	9814-N17R	SID-142991	92.1	34878	11.3	-52.6	17.3	3
212	87774-1Z	SID-142942	63	34446	20.4	-67.7	2.4	3
213	34318-1R	SID-142926	71	34592	5.6	-116.5	1.8	5
214	65395-N2Z	SID-142900	64.2	37401	-1.2	-105.8	2.6	5
215	74389-N6Z	SID-142885	93	36458	25.7	-78.4	21.6	5
216	82284-N4Z	SID-142882	89.6	37218	10.4	-111.1	12.5	3
217	82354-N1Z	SID-142881	99.3	33466	36.9	-18	25.4	7
218	74393-3R	SID-142875	94	38353	-1	-42.8	25.5	6
219	65392-5Z	SID-142874	88.1	33963	71.1	-25.8	14.6	5
220	15538-H2Z	SID-142872	57.7	34343	4.7	-119.3	3	3
221	**24815-H2Z**	**SID-142871**	**90.4**	**33750**	**32.7**	**-33**	**24.1**	**4**
222	65371-3Z	SID-142870	73.6	34733	25.7	-17.2	8	3
223	24889-H2Z	SID-142869	59.3	37627	5.1	-126.4	-1.1	2
224	**32294-H1Z**	**SID-142868**	**82.1**	**35766**	**3.1**	**-103.6**	**13**	**5**
225	**34908-2Z**	**SID-142867**	**83.2**	**35799**	**-3.7**	**-108.6**	**8.2**	**2**
226	34318-1Z	SID-142864	79.1	33932	2.7	-136	11.3	4
227	68950-2R	SID-142854	69.2	36581	-2.5	-77.9	-2.9	5
228	**18163-N13Z**	**SID-142848**	**96.6**	**33242**	**3.9**	**-66.3**	**27.3**	**4**
229	18035-H1Z	SID-142847	59.8	33301	7.7	-80.7	4	3
230	87797-1N	SID-142833	96.3	31359	10.1	58.4	11.1	7
231	86791-1I	SID-142790	54.9	38588	-19.2	-40	-1	2
232	86870-N3N	SID-142781	64.1	33906	-13	-46	-12.7	3
233	86840-1I	SID-142768	56.7	33814	65.4	-54.4	3.9	4
234	86840-N1N	SID-142763	78.5	34899	-5.2	-20.1	6.8	3
235	65392-5N	SID-142733	55.2	34596	-7.6	-53.6	5.4	5
236	83115-N12N	SID-142728	56.2	34526	-10.7	-40	9.8	4
237	74474-N2N	SID-142649	77.9	37802	-16.8	-92.5	9.5	4
238	87767-1C	SID-142602	77.2	34710	4.7	-58.5	10.1	3
239	86895-1Z	SID-142590	79.5	35987	-13.1	-69.6	15.4	5
240	82294-N3R	SID-142565	60	39764	-11.1	-40.2	0.8	6
241	84329-1Z	SID-142557	54.8	35313	-9.8	-47.5	6.7	3
242	83011-1C	SID-142553	56.2	35648	-9.5	-85.1	14.4	4
243	84131-2R	SID-142545	56.6	39171	-6.1	-142.3	-3.6	5
244	54913-2I	SID-142487	65.5	36171	-19.2	-20.7	14	8
245	74650-N2C	SID-142438	72.6	37274	-17.8	-71.9	5.6	3
246	74389-1C	SID-142430	52.4	39471	-23.8	-85.4	0.5	6
247	83615-N3C	SID-142421	67.6	38904	-14.1	-43.9	5.3	5
248	64647-1R	SID-142403	61.3	34704	-6.6	-81.2	9.3	5
249	64647-1Z	SID-142401	71	34064	-3.6	-72.6	7.9	6
250	82402-N2R	SID-142397	72.3	35467	-10.2	-63	6.9	4
251	69071-5R	SID-142394	58.2	38148	-10.6	-105.3	1	5
252	86825-1R	SID-142393	57.8	37873	-11.5	-94.3	1.6	4
253	72365-3BC	SID-142392	79.5	37205	-15.1	-207.1	2.6	4
254	83138-2C	SID-142384	69.7	35980	29	-85.7	5.9	3
255	74393-2C	SID-142383	60.5	35480	-6.9	-62.6	5.5	5
256	82414-N10C	SID-142382	58.5	34942	3.5	-36	7.8	4
257	54913-2Z	SID-142372	89.4	37181	-9.7	-67.6	24.5	7
258	84243-1Z	SID-142364	52.4	36120	-5.1	-48.5	3.3	5
259	82294-N3Z	SID-142358	63	37871	-12.9	-57.3	-0.7	5
260	83105-N3I	SID-142344	67.8	38433	-15.5	-56.9	0.9	2
261	69074-6I	SID-142285	71.5	36024	-4.2	-65.7	4.3	4
262	78930Z	SID-142279	63	37280	-14.4	-84.8	-4.2	4
263	**71747Z**	**SID-142274**	**81.2**	**38212**	**-16.8**	**-74**	**10.8**	**4**
264	73371R	SID-142257	61.3	34254	14.1	-135.7	-5.4	3
265	78670R	SID-142256	55.4	33281	2.8	-38.7	-4.9	3
266	78699Z	SID-142250	56.2	33550	-0.3	-64.8	-2.5	5
267	78699R	SID-142237	62.8	34251	2.6	-93.3	-2.3	3
268	82362-1C	SID-142174	57.8	34053	-0.5	-124.4	3.9	4

	Strain	CCG Identifier	% ERM•AcID Inhibition	Parallel Fluorescence	% Fluorescence Quenching	% Gal4(DBD)•DNA Inhibition	% MLL•KIX Inhibition	Promiscuity Filter
269	82339-N7Z	SID-142171	52.4	35302	1.4	-15.2	5.1	4
270	65392-4C	SID-142164	53	35946	13.9	-110.6	5.3	4
271	83115-N8Z	SID-142128	71.4	36689	25.1	-179.1	7.1	3
272	83115-N8I	SID-142114	77.8	38711	-4.7	-102.8	6.7	3
273	83017-2N	SID-142098	94	38395	1.8	-60.1	26.5	3
274	41433-AC5N	SID-142094	56.6	36110	-2.9	-27.5	4.5	2
275	84381-1I	SID-142061	81.4	39925	-21	-67.2	4.9	4
276	78685N	SID-142022	60.5	37913	-9.5	-56.7	5.5	5
277	78832I	SID-141973	67.8	35497	-17.2	-44.2	7.7	3
278	71747I	SID-141966	66.4	39419	-18.4	-129.1	0.4	5
279	34381-1R	SID-141950	81.5	34965	-8.1	-73.4	5.1	6
280	83115-N8C	SID-141939	62.5	35364	11.5	-97.9	1.3	5
281	82417-N10Z	SID-141909	58.8	38694	-6.1	-65.2	0.7	4
282	82354-N2Z	SID-141887	59.4	36311	-3.9	-119.7	-2.3	4
283	84250-3N	SID-141875	68.8	39868	-21.6	-49.1	5.6	2
284	**84215-1I**	**SID-141865**	**100.5**	**35290**	**-4.5**	**-33.1**	**16.7**	**3**
285	86810-N2N	SID-141852	75.5	33365	1.8	-12.5	6.7	6
286	87690-N1N	SID-141726	59.5	34599	-8.4	-33.7	-17.1	2
287	72365-3BI	SID-141718	88	35136	-5.5	-58.3	9.9	5
288	82319-N8I	SID-141677	70.7	35620	-6.7	-37.2	4.6	4
289	82319-N4I	SID-141651	99.7	35423	-11.8	26.6	28.3	5
290	68923-4R	SID-141639	94.4	35026	3.2	-27.6	11.5	4
291	68948-2C	SID-141623	70.4	35067	-2.4	-47.7	8.2	4
292	**54916-1C**	**SID-141597**	**75**	**38438**	**-26.3**	**-101.4**	**7.9**	**4**
293	67325-N7C	SID-141528	87.9	36762	-10.8	-50	15.9	10
294	65387-4Z	SID-141491	57.4	35908	-3.8	-58.7	1.4	5
295	84260-1C	SID-141434	69.5	35393	-1.2	-30.9	6	3
296	69078-1BR	SID-141381	54.8	36086	3.8	-75	6	3
297	84215-1Z	SID-141266	77	38332	-14.1	-94.8	7.5	4
298	82319-N2I	SID-141243	82.2	37700	-18.6	-43.1	6.9	4
299	74434-N2I	SID-141235	60.6	36384	-8.1	-76.5	4.1	4
300	83115-N5N	SID-141219	60.3	35582	-13.4	-43.7	6.2	2
301	82414-N2R	SID-141113	76.2	36444	-14.6	-128.4	10.5	4
302	83665-1I	SID-141073	65.8	35382	-13.9	-69.4	10.7	5
303	82344-N2AN	SID-141041	56.6	34276	-5.5	-81.7	7.5	4
304	82422-N8I	SID-140995	64.4	39088	-27.4	-98.5	2.7	2
305	83120-N3I	SID-140985	91.9	39179	-26.1	-8	15.3	5
306	82399-N1I	SID-140980	56.6	38292	-26	-101.6	5.7	3
307	86853-2I	SID-140964	54.6	36399	-12.2	-54.6	3.7	4
308	49701-1I	SID-140963	62.2	40415	-24.2	-104.5	0.5	3
309	87790-1N	SID-140956	69.7	34124	-4.3	-38.7	1.7	4
310	7736-N9AI	SID-140934	59.1	37010	-15.5	-44.1	4.8	5
311	82344-N7I	SID-140926	69.7	37812	-14.7	-43.9	7	3
312	82299-N4I	SID-140924	65.1	39637	-18.3	-33	-5.8	3
313	82344-N5I	SID-140920	100.8	36759	-14.5	34.9	14.9	5
314	82344-N2AI	SID-140919	85.9	37865	-24.6	-61.8	7.8	5
315	82344-N6I	SID-140916	96	37214	-14.6	27.4	9.7	6
316	**74604-N1I**	**SID-140915**	**74**	**38202**	**-18.3**	**-45.9**	**6.1**	**3**
317	**82349-N4I**	**SID-140914**	**73.6**	**37721**	**-13.3**	**-39.7**	**7.2**	**2**
318	82299-N7I	SID-140912	69.2	39813	-30.2	-64.7	7.2	6
319	82399-N3I	SID-140908	91.8	39665	-25.3	11.7	16.2	6
320	82319-N1AI	SID-140895	80.4	38345	-20.6	-20.9	8	4
321	82349-N1AI	SID-140894	72.9	37550	-14.2	-39.5	0.5	3
322	83665-1N	SID-140794	70.9	30390	9.9	-2.6	3	6
323	82284-N11I	SID-140785	68	35749	-11.9	-58.1	-2.2	6
324	74474-N1R	SID-140742	55.3	34234	12.3	-91.8	1.9	3
325	68926-2Z	SID-140739	51.7	34394	10	-80.2	2.3	3
326	34951-A6R	SID-140714	55	37486	-2.9	-54	6.4	3
327	68965-N5R	SID-140703	78.2	38654	-10.9	-132.7	-0.4	4
328	78776N	SID-140665	54.9	32604	-1.4	-36.4	-19.8	2
329	78796I	SID-140657	52.8	38920	-13.7	-183	-4.7	5
330	**73401N**	**SID-140641**	**57**	**36386**	**-10.6**	**-46.6**	**5.1**	**2**
331	**58236N**	**SID-140604**	**79.9**	**33193**	**-0.3**	**-7.9**	**9.3**	**5**
332	32381-H1I	SID-140525	52.9	36797	-10.5	-67.1	3.3	7