Dissecting the Roles of Dynamic Substructures in Beta-Barrel Containing Coactivators

by

Samantha De Salle

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2021

Doctoral Committee:

Professor Anna K. Mapp, Chair
Associate Professor Alison Narayan
Associate Professor Patrick O'Brien
Assistant Professor Wenjing Wang

Samantha De Salle

sdesalle@umich.edu

ORCID iD: 0000-0002-1571-8503

©Samantha De Salle 2021

Acknowledgements

The end of my Ph.D. journey is far from how I pictured this to go with COVID-19 completely changing how life operates. Starting my last year in quarantine and not getting to see my lab mates has been incredibly difficult. It is hard to grasp how life has changed so much over a year's time, but here we are. Light at the end of the tunnel, and while it has been difficult, I am incredibly thankful for all the opportunities my time at Michigan has provided.

I want to start by thanking Dr. Cherie Dotson for getting my feet wet with grad school. Completing my REU going into my senior year of college helped me figure out what I wanted to go to grad school for. You taught me to never doubt myself, encouraged me to reach for things I would have never thought possible, and stood by all of us ducklings. I am forever thankful for you Mama Duck.

Of course, none of this would be possible without my mentor, Anna. You took me in during my REU as an undergrad, helped me through my grad school applications, and gave me direction these last 5 years. You taught me how to be an independent scientist and encouraged me to push out of my comfort zone. It has been an invaluable experience working under your guidance, and I know I am far better for it. Your empathy and compassion as a mentor continue to astound me, and I strive to achieve that in my career.

I would like to next thank my committee members- Professors Pat O'Brien, Alison Narayan, and Wenjing Wang. Rotating in Prof. O'Brien and Prof. Narayan's labs taught me skills that I have carried forward during my Ph.D. Prof. Wang, it has been incredible to watch your lab

grow these last 3 years. All three of you have been in my corner since day one, and your advice has encouraged me to be a better researcher.

To my lab mates, I would have never survived these last five years without you. Each and every single one of you has impacted my life for the better. You all have pushed me and inspired me to be a better scientist, mentor, and friend. It truly breaks my heart I have not been able to see all of you recently, but know I am forever grateful for our friendships.

I am forever indebted to Cassie Joiner, who taught me as a naïve REU student. Your patience, compassion, and encouragement are qualities I strive to emulate. Not only did you treat me as a peer when I was just a little undergrad, you genuinely made me feel like I belonged in science. Thank you. I also owe a HUGE shout out to Meg Breen. Working with you during my rotation taught me how to survive in the Mapp lab. You gave me independence, confidence, and helped me when I barely knew how to be an adult.

I am especially thankful to have the opportunity to have worked with Stephen Joy, lab uncle turned lab uncle-manager. Due to your proximity, you got to see me through all my lows and highs in grad school-thank you. I think you've seen me cry more than anyone else, but you always knew just what to say. I know I am not alone in saying you are incredibly compassionate andamazing lab mate. Thanks for letting me drag you to an aquarium, trying to get lab mates to laugh with crazy stories, and just being an invaluable resource, even if you'll never get coffee at one particular Starbucks with me ever again.

I next need to thank the person that had to deal with me day in and day out, Brennan. Meeting you at the start of grad school, I would never in a million years have thought we'd have built the life we did, and I cannot be more grateful. You have been my safety net, my personal cheerleader, and my rock. Thank you for taking me to get ice cream in the freezing cold, not getting

annoyed every time I see a turtle, and for loving me for me. Without you and the cats, I would never have survived these last five years. Coming home to you, Phox, and Nora is the best cure for a bad day.

Last, and certainly not least, I have to thank my family. My grandparents have helped shape me to the become the woman I am today. Grandpa, you have been my biggest advocate, and taught me so many things, but most importantly you let me be me, even if that meant getting your hair badly styled. Grad school has been hard, and I unfortunately lost my grandmother and grandfather while in grad school. There isn't a day that goes by that I don't think about them, and I truly hope I am continuing to make them proud. Grandma, you taught me how to sew, how to cook, and how to have fun. Nothing brought you more joy than being with your family, and I will always be your little gypsy child. Nunu, I will forever miss our lunch dates, watching Big Brother with you, and Nestle Crunch Bars in the back seat when you picked me up from school. No one could pull off a bucket hat like you. Thanks for taking the heat for me when I took a bite out of every single tomato. I miss you both every single day.

Mom and Dad, there are no words. From an early age, you always made me believe I could do anything. You put up with my never-ending questions, and always encouraged my passion for science. Mom, bringing me into your office piqued my interest in biology. Dad, your science talks taught me to think like a scientist. Thank you both for not killing me in high school. When I didn't believe in myself, you always did. You've always been in my corner and my biggest supporters. Thank you for teaching me to be a strong and independent woman, for letting me go off into the world. Take comfort in raising me to be able on my own, and for being there when I need help. I love you both endlessly, and I hope I continue to make you both proud.

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

ABD	Activator binding domain
AcID	Activator interaction domain
ATF6α	Activation transcription factor 6α
CBP	CREB binding protein
CD	Circular dichroism
DBD	DNA binding domain
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DREB2a	Drought response element binding protein 2a
DTT	Dithiothreitol
4-DMN	4-*N.N*-dimethylamino-1,8-napthalimide
ETV	ETV translocation variant
FITC	Fluorescein isothiocyanate
FP	Fluorescence polarization
FPLC	Fast protein liquid chromatography
HIF1α	Hypoxia inducible factor 1^{α}
HPLC	High performance liquid
HSQC	Heteronuclear single quantum coherence spectroscopy
IC_{50}	Half-maximal inhibitory concentration
IPTG	Isopropyl-β-D-1-thiogalactopyranoside

KIX	Kinase inducible domain interaction domain
LC-MS	Liquid chromatography mass spectrometry
Med1	Mediator subunit 1
Med15	Mediator subunit 15
Med25	Mediator subunit 25
MLL	Multi-lineage leukemia
mRNA	Messenger RNA
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance spectroscopy
NR	Nuclear receptor
PDB	Protein Data Bank
PEA3	Polyoma enhancer activator 3
PPI	Protein-protein interaction
PrOF	Protein-observed fluorine NMR
PTOV1	Protein overexpressed variant 1
RA	Retinoic acid
RAR	Retinoic acid receptor
RNA	Ribonucleic acid
RNA Pol II	RNA Polymerase II
TAD	Transcriptional activation domain
TAZ	Transcriptional adapter zinc finger
TF	Transcription factor
VP16	Herpes simplex viral protein 16

VWA	von Willebrand factor type A

Abstract

The process of transcription is integral in all of life. Transcription is dependent upon many players and moving parts, such as transcriptional activators and coactivators that assemble and constitute the transcriptional machine, respectively. Despite the fundamental role of transcriptional activator and coactivators, a mechanistic understanding of these binding partners recognizes one another is lacking, largely due to their challenging biophysical characteristics. Both activators and coactivators are highly dynamic and the protein-protein interactions they form are transient, making them difficult study mechanistically and structurally. Given the role that activator-coactivator complexes play in healthy and diseased organisms, however, it is vital to understand their function.

Historically molecular recognition models of activator-coactivator complexes have characterized the interactions as largely nonspecific, dictated by unstructured and amphipathic transcriptional activators interacting with hydrophobic surfaces within coactivators. However, this model does not accurately represent the critical role of activator-coactivator interactions. Recent data from our lab and others has shifted the model to one that supports specific molecular intermolecular contacts underpinning activator-coactivator PPIs. In two structurally different coactivators, highly dynamic regions adjacent to the hydrophobic binding surfaces, for example, play a role in recognition of activators and in allosteric communication between binding sites.

The goal of the work in this thesis was to test this model further through comparing several coactivators bearing structurally similar motifs, an Activator Interaction Domain (AcID). Specifically, we studied the AcID domains from the Mediator subunit Med25 and a second human

protein, Prostate Tumor Overexpressed Variant 1 (PTOV1), which contains two tandem AcIDs. The key differences between these domains reside in the dynamic regions flanking the primary binding surfaces. We show that the AcIDs are capable of recognizing overlapping binding partners *in vitro*, with modest differences in equilibrium binding affinities (up to 4-fold). However, using transient kinetics, binding mechanisms of the different AcID motifs revealed some more significant differences. Specifically, it was found that each AcID paralog exhibit a distinct conformational signature upon binding to a given activator, suggesting that the sequence differences in the dynamic substructures of each AcID was indeed playing a role in the recognition.

A second role of the dynamic substructures is in allosteric communication. In the context of the AcID paralogs examined here, allosteric communication between two binding sites was observed via transient kinetics. Consistent with this model, we show that a covalent allosteric modulator attenuates binding the PTOV1 AcID more potently than Med25 AcID, demonstrating the changes in the dynamic regions alters small molecule binding. These dynamic substructures can be exploited as hotspots for targeting, as these allosteric regions are not as highly conserved in paralogs. We demonstrate that identified allosteric modulators can be used as chemical probes to perturb the dynamic hotspots, providing an opportunity to target homologous proteins with high selectivity. Further, we show that even highly related activators are able to induce differential conformations in activator binding domains, highlighting that these interactions are specific.

Chapter One

Molecular Recognition Models and Allostery of Transcriptional Coactivators

1.1 Abstract

Transcription is the central process that helps maintain normal cell function in all living organisms. This process is tightly regulated through a series of protein-protein interactions at the genomic loci. Recruitment to the genomic loci is mediated via DNA-bound activators that can bind the activator binding domain (ABDs) of coactivators. These ABDs are central hubs that relay information to the rest of the transcriptional machinery. The nature of these interactions has long been debated, with early models of molecular recognition arguing for a nonspecific model driven by electrostatics and intrinsic disorder of the activator. However, emerging evidence shifts this argument towards a more specific model; that is dynamic regions in coactivators undergoing unique conformational shifts upon binding to different activators. Moreover, these dynamic substructures can dictate allosteric communication within the ABD.

In this chapter, we explore the models of molecular recognition and allostery that underpin coactivator binding. The transient nature of these interactions are discussed in detail, specifically how the transcriptional machinery can be organized. Lastly, we explore how dynamic substructures of ABDs can dictate allostery and ways this has been modulated.

1.2 Introduction

Transcriptional protein-protein interactions (PPIs) are highly regulated due to their integral role in gene expression. These PPIs are responsible for the assembly of transcriptional machinery in response to developmental and environmental signals. These signals prompt transcriptional activators to localize to enhancer or promoter regions, and bind to coactivators.[1–3] Once the complex is formed, the coactivator acts as a bridge, relaying information from the genomic loci to the rest of the transcriptional machinery, thus initiating transcription (Fig. 1.1).[3,4] Dysregulation of transcriptional PPIs results in a variety of diseases, such as cancer metastasis, hypoxia, and inflammation.[5–8] Despite interest in pursuing transcriptional PPIs as a therapeutic target, these interactions are challenging to study as they are not amenable to traditional structural and biophysical techniques due to intrinsic disorder and highly dynamic regions.[9–11]

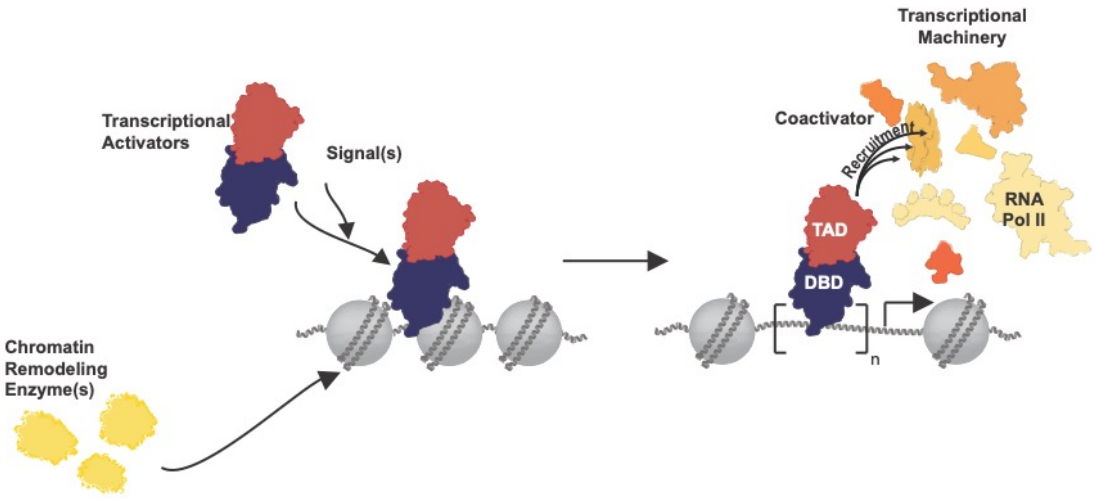


Figure 1.1: General mechanism of transcriptional activation. Upon signal transduction and localization of the activator to the promoter/enhancer site through the DBD, chromatin remodeling enzyme unwind histone bound DNA, revealing the genomic loci. The TAD interacts with the ABD of the coactivator. This complex then recruits the rest of the transcriptional machinery, including RNA polymerase II and the general transcription factors to initiate transcription.[2,12]

Recruitment to the genomic loci is mediated by transcriptional activators, binding to DNA directly through a DNA binding domain (DBD). Localized to the promoter and enhancer regions, activator•coactivator PPIs are formed from amphipathic regions within the activator referred to as transcriptional activation domains (TADs) and conformationally dynamic regions of the coactivator called activator binding domains (ABDs).[1,3] Because of their dynamic nature, coactivators act as central hubs at the site of transcription, binding to multiple partners in order to relay information to recruit the rest of the transcriptional machinery.[4] For example, the master coactivator complex CREB binding protein (CBP) consists of seven individual coactivator motifs that are linked dynamic intrinsically disordered regions, allowing for a highly flexible protein. The domains within CBP interact with hundreds of activators, aided by the overall malleability of the protein (Fig. 1.2a).[13,14]

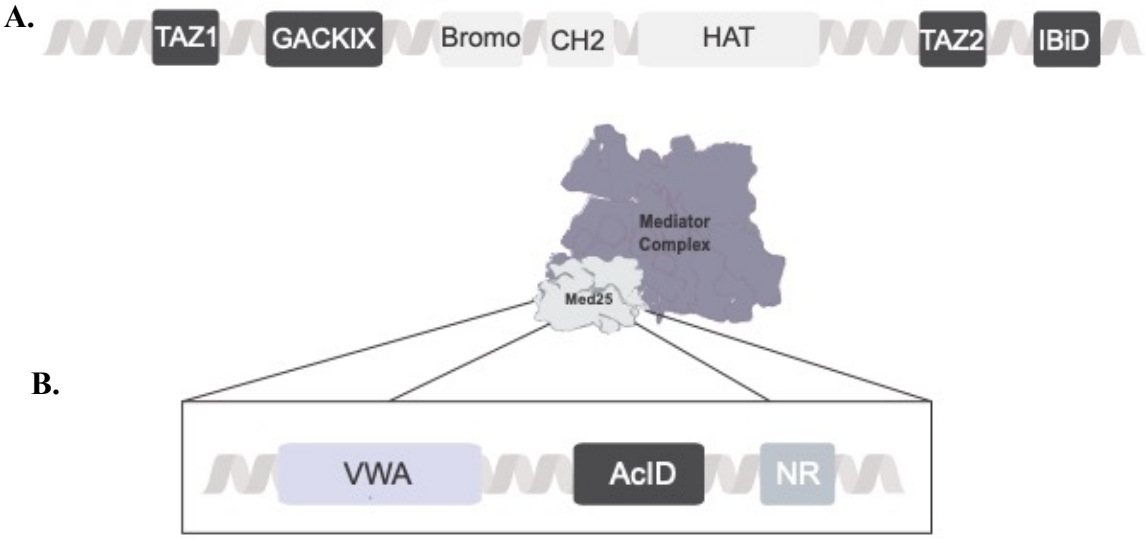


Figure 1.2: Coactivator complexes and their ABDs. A). Gene layout of the master coactivator CBP. The individual domains in CBP are linked by long, flexible linker regions, resulting in high levels of conformational plasticity. Highlighted in dark gray are coactivator ABDs that bind directly to transcriptional activators. In light gray are coactivators that act as chromatin remodelers, such as a histone acetyltransferase domain (HAT) and a bromodomain.[13,15] B). Mediator complex and its Med25 subunit. The Mediator complex is another excellent example of a coactivator complex. Within the tail module of Mediator are dozens of coactivators, such as Med25. Within Med25 is its VWA domain that links it to the rest of the Mediator complex, its ABD domain termed AcID and a nuclear receptor that can bind to hormones such as retinoic acid.[16–18]

Coactivators act as hub proteins, and they are critical in connecting all the key players of transcription.[4,16] One such example is the coactivator complex named Mediator. Mediator is comprised of four units, the head, middle and tail modules, and a removable kinase domain, acting in concert as a coactivator complex.[12,19] Mediator itself binds to RNA Pol II, acting as an intermediary between transcriptional activators bound to DNA and the rest of the general transcription machinery.[4] Located within the tail region are coactivator subunits such as Med25. Med25, connected to the Mediator complex through its von Willebrand factor A domain (VWA), is able to interact with other coactivators, such as CBP, and a multitude of transcription factors (Fig. 1.2b).[12,18] It is through these interactions that information is relayed through the preinitiation complex.

1.3 Transcriptional activators in activator-coactivator complexes

The mechanisms by which activators and coactivators recognize each other have long been argued.[18,20–22] Early models were largely nonspecific, arguing that the gene-specific function of activators was due to the DNA binding domain recognizing particular DNA sequences. Biochemical and biophysical studies indicated that activator TADs were unstructured when unbound.[20–22] Additionally, the only identifiable sequence patterns among TADs were the presence of acidic and hydrophobic residues; further, virtually any amphipathic peptide could function as a TAD when localized to DNA (Fig. 1.3).[1,9,23,24] One leading example is the promiscuity of activators such as the herpes simplex viral protein 16 (VP16) which was found to bind to a variety of different coactivators and also to function in all eukaryotes.[25,26] Structural studies then revealed that TADs, once termed “acidic blobs and negative noodles” undergo a coupled folding and binding, often taking on an amphipathic helical structure when bound to a coactivator, changing the viewpoint of models to argue for specificity.[20,27,28]

Unfortunately, the discovery that not all TADs form ordered complexes upon binding caused the molecular recognition model to shift back to one of non-specificity.[10,29] It has been shown that some TADs can bind to the ABD binding face in multiple distinct orientations, or even to different binding sites all together, thus supporting the argument that there are no specific recognition motifs within TADs themselves.[30] A well-cited example of this is the coactivator Med15's binding partner Gcn4. The Gcn4 TAD was shown to bind to all of Med15's ABDs in multiple orientations, in what has been described as a TAD "hydrophobic cloud." None of these models take into consideration, however, any role the coactivator may serve in molecular recognition.[31,32]

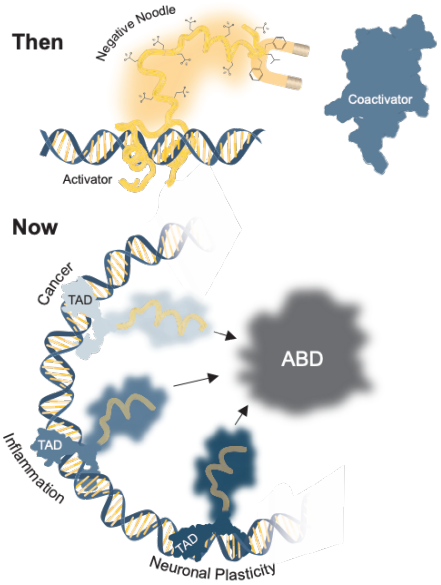


Figure 1.3: Models of molecular recognition. The model for the mechanism of which activators•coactivator interactions occur is largely contested. The original hypothesis suggested one of largely nonspecific interactions, with intrinsically disordered, negatively charged TADs binding to coactivators. Recent biophysical and structural data suggests this model is oversimplified. In addition to TADs existing in conformational ensembles, often experiencing coupled folding upon binding, but the ABD domain also undergoes conformational changes due to internal malleability. The result is "fuzzy" complexes that can undergo conformational changes.

While it has been well-accepted that TADs undergo a coupled folding and binding, it has largely been ignored that the ABD can also undergo significant conformational changes. As described in more detail in section 1.4, biophysical studies have demonstrated that coactivators can adopt unique conformations themselves when binding to different activators, suggesting that ABDs can adapt to each TAD (Fig. 1.3).[27,33] The kinase-inducible domain interaction domain (KIX) of CBP is a prime example of an ABD undergoing a significant remodeling event. Upon binding of MLL, KIX is shown to undergo repacking of its hydrophobic core, stabilizing the loop region of the MLL binding site.[34] This ultimately stabilizes interactions with TADs to the other binding site, such as cMyb, by almost two-fold. In fact, binding of one TAD to KIX can significantly influence which binding partners interacts with the other site, thus demonstrating that conformational changes upon binding can influence function and selectivity.[35–37]

1.4 Phase separation of transcriptional components

It is critical that the process of transcriptional machinery assembly is rapid and reversible, and thus it is underpinned by transient, short-lived interactions. The transient nature of these interactions allows for expression to be turned on and off in a temporally and spatially appropriate manner.[38] One model proposed that this process is mediated is through phase separation. At its core, phase separation is a physiochemical process by which molecules separate into a dense phase and a depleted phase, allowing for rapid phase transitioning. Phase separation is a well-studied phenomenon in cells given their colloidal nature (Fig. 1.4).[39] Cells use phase separation as an organization tool, creating biomolecular condensates that can compartmentalize and concentrate biochemical reactions as well as allow for rapid movement of compartments into and within in the dense phase.[39,40]

Proteins can undergo phase separation, with a variety of factors leading to their propensity to undergo phase transitioning, including concentration, multivalency and solubility. Intrinsically disordered proteins, like TADs, have been found to have a high propensity to undergo phase separation.[40,41] Moreover, phase separation has been observed at the site of super-enhancers, long stretches of enhancers that are often responsible for regulating genes controlling cell differentiation. Specifically, super-enhancers are known to concentrate and compartmentalize transcriptional machinery, with the multivalency of transcription factors enabling to supposed crosslinking and ultimately leading to phase separation.[42,43]

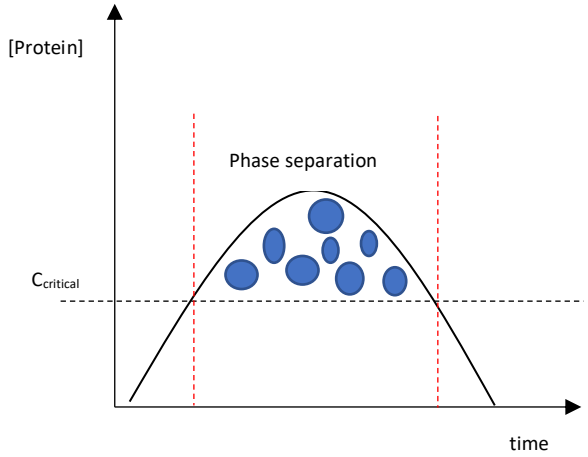


Figure 1.4: Phase separation as an organization tool for transcription. Phase separation is a proposed method for organization and the rapid assembly of biomolecules. As protein concentration increases, rising above the critical concentration, solubilized protein can form into liquid droplets, called biomolecular condensates. These droplet compartments allow for rapid diffusion of protein within the condensate and promotes dynamic exchange with the dilute phase.[39]

One early example for the evidence of phase separation at transcription is interaction between coactivator Med1 and activator BRD4, both of which contain intrinsically disordered regions. Med1 and BRD4 were found to localize to puncta within the nucleus, and these puncta were traceable to genes known to be regulated by super-enhancers. Extensive biophysical and biochemical analysis demonstrated that these puncta are not membrane-bound, suggesting they exist as biomolecular condensates close to the site of transcription. Further, Med1 and BRD4 were found to phase separate *in vivo* and *in cellulo,* and the phase separated Med1 intrinsically

disordered region concentrates were found to activate transcriptional machinery from nuclear extracts.[44]

The phase separation model for transcriptional regulation suggest that phase separated liquid droplets are highly concentrated in transcriptional machinery. Transcription initiation would greatly benefit from highly concentrated biomolecular condensates. For example, CBP/p300 is known to be nonabundant in the nucleus, and if found to exist in phase separated droplets, could allow for the quick localization to the site of transcription. Essentially, the physical consequences of phase separation lead to hyper-cooperativity, with a sort of switch-like behavior, and can cause high frequency transcriptional bursting. This behavior can help accommodate the short-lived, transient nature of activator•coactivator interactions, allowing for rapid assembly and disassembly at the site of transcription.[43,45] However, there has not been direct evidence of phase separation in a biological system, and has yet to be observed at non-super-enhancer regions.

1.5 ABDs are the molecular recognition units

While phase separation can accommodate the short-lived nature of these interactions, it does not account for molecular recognition. Moreover, the molecular recognition model that activator•coactivator binding mechanism is non-specific does not account for the critical role these PPIs play in gene regulation. While TAD promiscuity is often argued as a driving force, as seen with VP16 or p53, promiscuity is often studied in isolation on individual domains.[25,26] This is partially due to the fact that is often technically impossible (or nearly so) to study full length systems *in vitro*, resulting in the use of truncated domains and therefore not accurately recapitulating biological systems.[46,47] Moreover, the nonspecific model does not take into consideration other components of the transcriptional machinery ensemble, specifically coactivators. Coactivators bind to tens of different binding partners, each resulting in different

gene expression. The KIX domain of CBP binds to dozens of different partners, (Fig. 1.5), initiating genes responsible for a diverse array of different pathways.[14] Co-localization does not explain ABD molecular recognition, as these dynamic motifs need to be able to quickly bind, recognize, and initiate gene-specific transcription, then advance to the next gene.

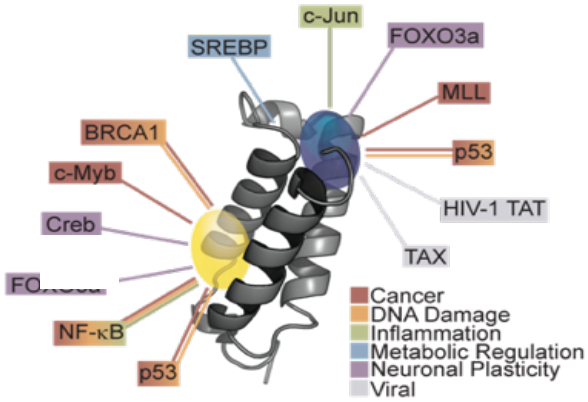


Figure 1.5: The two binding faces of CBP GACKIX and its interaction partners. GACKIX consists of two binding faces, termed the cMyb binding face and MLL binding face. As illustrated, GACKIX binds to a suite of different binding partners implicated in a variety of networks and diseases.[13,14]

It is becoming increasingly accepted that ABDs within the coactivators serve as molecular recognition units. The KIX domain, for example, is one of the most important molecular recognition sites for PPIs at gene regulation. The KIX motif has been identified in a multitude of coactivators, such as CBP/p300, the human activator recruited cofactor Arc105, and Med15 of the yeast Mediator complex.[14,48,49] KIX motifs are responsible for recognizing a range of transcription factors, with each interaction playing a critical role in a variety of biological pathways, including long term memory storage in the hippocampus, lipid homeostasis in mammals, and even processing of HIV-dependent binding partners.[14,49–51]

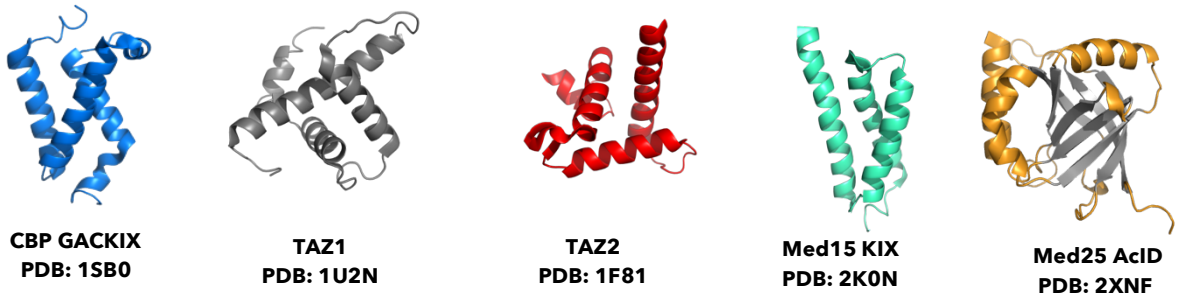


Figure 1.6: Structures of transcriptional coactivators ABDs. Typically, ABDs are helical bundles connected by dynamic lo0p regions. Excellent examples of this are the KIX domains found in CBP/p300, and Med15 subunit of the yeast Mediator complex. TAZ1 and TAZ2 are two other coactivators of the CBP complex. Lastly, the Med25 subunit of the human Mediator complex contains a seven-stranded β-barrel core, shown in gray, flanked by more common structural features: α-helices and loops.[15,18,35]

These ABDs often are helical motifs, as observed with the individual coactivator domains found within CBP/p300 (Fig. 1.6). These helical motifs are linked by loop regions, thus allowing for a highly dynamic protein capable of adopting many different conformations. In fact, the dynamic, helical nature of ABDs plays a crucial role in the physical binding of activators and coactivators.[5,13,52] While ABDs typically lack topography with obvious binding sites, the malleable nature of ABDs allows for them to adopt conformations specific to each binding partner with the interaction stabilized by hydrophobic and electrostatic interactions.[53] In contrast to this, and of importance to the work in this thesis, is the structurally distinct ABD of Mediator subunit Med25, termed the Activator Interaction Domain (AcID), possessing a seven-stranded β-barrel serving as the core of the domain. However, similar to prototypical ABD structures, Med25 AcID is flanked by three α-helices and flexible loop regions (Fig. 1.6).[18,54]

1.6 Hotspots and allostery within ABDs

While the activator•coactivator PPI interface is rather featureless, lacking any defined binding pocket and binding taking place over large surface areas, adjacent dynamic substructures also play an important role in molecular recognition.[46] These hotspots can help regulate internal

allosteric networks upon when a conformational change is induced upon at non-overlapping sites.[46,55] This is observed in ABDs and is a direct result of ABD malleability.[56] Proteins in their native state exist in essentially a conformational ensemble with pre-existing populations of different states become redistributed upon binding and a resulting allosteric perturbation.[57,58]

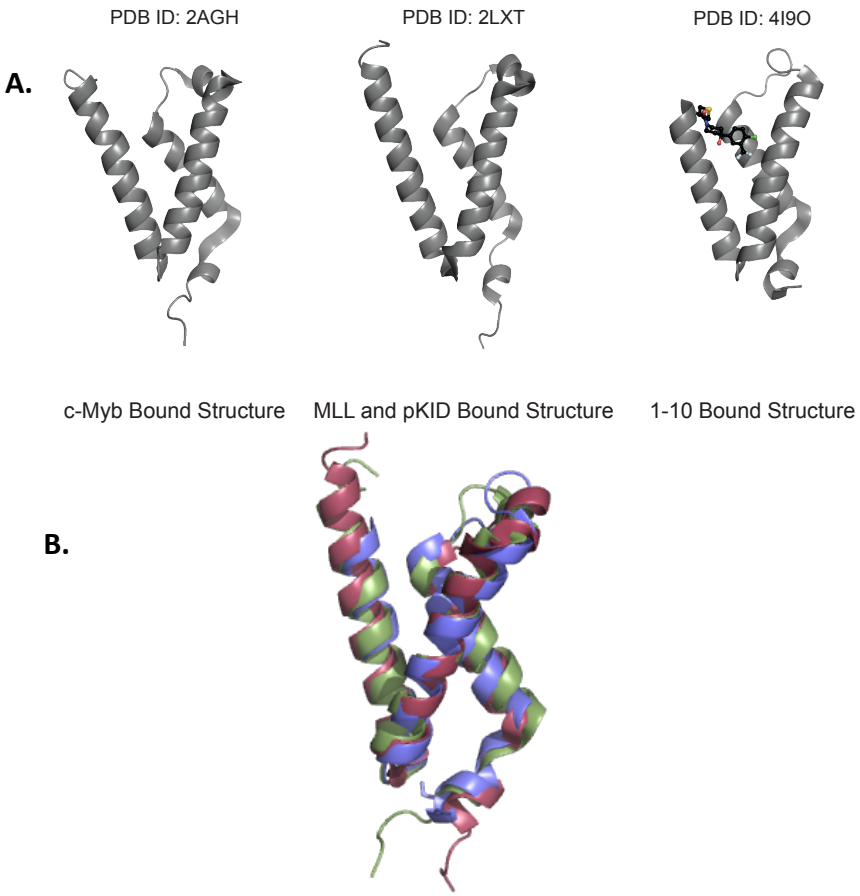


Figure 1.7: Different conformations of CBP GACKIX. A) Different binding partners of KIX result in different conformation and can even result in overall stability of the protein. cMyb, mLL and pKID are all TAD binding partners, and can form ternary complexes via cooperative binding. Fragment **1-10** is an allosteric modulator that tethers to N627C, changing cooperativity of KIX binding. B) Structural overlay of the different KIX conformations. In red is MLL and pKID bound, green is c-Myb bound and purple is **1-10** bound. The biggest changes in conformation can be observed in the loop regions.

The presence of dynamic regions allows for communication within coactivators. Upon binding at one interface, a conformational change is induced, which can change binding at secondary sites.[56,57] The previously described CBP KIX is a key example of this (Fig. 1.7). Not

only does binding to one face result in cooperative binding to the other face but binding to both faces causes an overall stabilization of the ABD. In its native state, KIX, and other ABDs, essentially exists in a conformational ensemble. Once KIX binds to another protein or peptide, it is shown to undergo a winnowing of conformational ensembles, resulting in a stabilization.[33,37,48] Another example of ABD restructuring is observed with another CBP ABD, TAZ1. Binding of CITED2, a negative regulator of hypoxic stress response, to a TAZ1 to a secondary site results in a conformational change in TAZ1 that inhibits binding of a second TAD, HIF1α. Put another way, through a forced competitive mechanism, CITED2 binding to the ABD results in a restructuring that destabilizes HIF1α binding (Fig. 1.8).[22] These observations demonstrate that ABD remodeling is critical component that needs to be considered in molecular recognition, and a guiding factor in the work presented in this thesis.

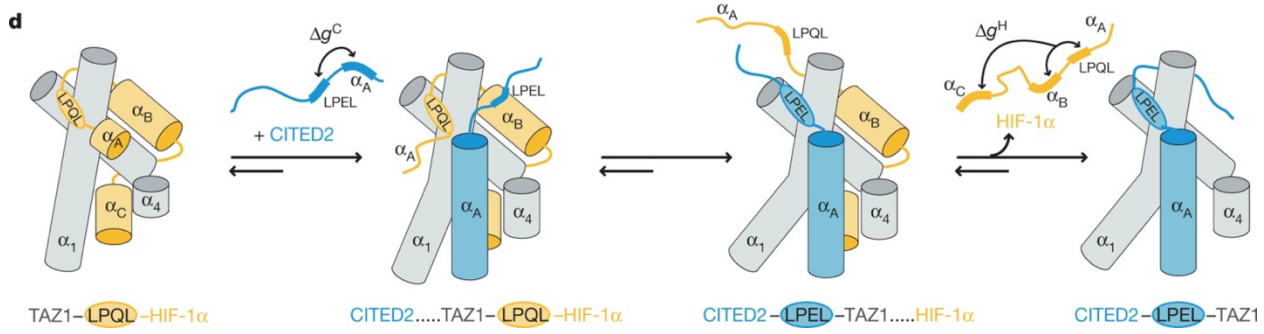


Figure 1.8: Schematic for the displacement of HIF1α. TAZ1 is shown in gray, HIF1α is in orange, and CITED2 is in blue. Rather than trying to outcompete the tight interaction of HIF1α, CITED2 binds to an allosteric site. Binding to the second site induces a change in the dissociation rate of HIFα, enhancing it by an order of magnitude.[22]

Due to their role in allosteric regulation of ABDs, dynamic substructures serving as hot spots can be exploited for targeting.[46,55] Given the inherently difficult nature of targeting PPIs, use of allosteric modulation is promising.[11] Typically, PPI interfaces are highly hydrophobic and often conserved, making it difficult to target orthosterically while achieving selectivity and specificity. Use of allosteric modulators that can bind to dynamic hot spots can induce conformational changes

that can inhibit or enhance binding.[59] Allosteric modulators that target dynamic hot spots can act as chemical probes that can be used to dissect binding mechanisms.[11,47,60] Additionally, given that dysregulation of these PPIs often results in disease, it is critical to understand binding mechanisms as well as be able to selectively and specifically target.

KIX is a prime example of an ABD present in multiple systems, therefore increasing the difficulty in selective modulation. Using a tethering screen approach, the fragment **1-10** was found to target CBP GACKIX N627C with high affinity (Fig. 1.8).[59] This allowed for the development of irreversible chemical cochaperones that are specific to CBP KIX over its paralog p300 KIX. Not only was **1-10** was found to cause allosteric enhancement of a pKID binding, **1-10** derivatives were able to induce conformational changes that either positively or negatively impacted cooperativity.[36]

Over the last decade, there has been a critical reevaluation of the mechanism models of activator•coactivator interactions. Once proposed to be non-specific and driven by electrostatics, the paradigm is shifting to one of more specificity. Conformational changes of the ABD upon TAD binding can result in a winnowing of conformational ensembles from the native to bound state. The conformational rearrangement is driven by dynamic substructures. Additionally, there exists an allosteric network, influenced by these conformational changes. However, there is still much to be discovered about this process. The work outlined in this thesis aims to expand upon the model of ABD conformational changes by demonstrating that coactivators are linked via a common mechanism. Specifically, the malleable nature of ABDs can induce conformational changes upon TAD binding, with each TAD demonstrating a unique conformation.

1.7 Dissertation summary

The goal of this dissertation was to explore the binding mechanisms the activator binding domain AcID utilize. Deviating from prototypical coactivator structures, the AcID motif contains a seven-stranded β-barrel core in addition to its loop and helical substructures. Recent work has indicated that despite the rigid core, AcID interactions with activators can be dynamic due to the presence of dynamic loops and helices.[61] Using three different AcID containing systems, Med25, PTOV1A and PTOV1B, each with key residue differences in their dynamic substructures, we set out to determine how these changes affect ABD plasticity and their role in molecular recognition.

In chapter two, we use a biophysical approach to determine how different AcID motifs recognize activator binding partners. We start establishing the changes in sequences and how this affects predicted structures. Using this information, we set out to determine if activator binding is conserved, and how this affects conformation and binding mechanisms.

In chapter three, we examine how hotspots can be exploited in a two-pronged approach. To begin with, we use the highly similar ETV/PEA3 family of activators to examine selectivity. Specifically, we set out to see how slight changes in sequence of the activator can affect conformational changes, driven by in coactivators. We would predict that such small changes would not significantly alter mechanisms. We then explore allosteric communication in coactivators, and how changes sequence of these dynamic hotspots affects modulator binding.

1.8 References

1. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
2. Vernimmen, D. & Bickmore, W. A. The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends in Genetics* **31**, 696–708 (2015).
3. Krasnov, A. N., Mazina, M. Y., Nikolenko, J. V. & Vorobyeva, N. E. On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell Biosci* **6**, 15 (2016).
4. Transcriptional Coactivator Complexes | Annual Review of Biochemistry. https://www.annualreviews.org/doi/10.1146/annurev.biochem.70.1.475?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub++0pubmed.
5. Garlick, J. M. & Mapp, A. K. Selective Modulation of Dynamic Protein Complexes. *Cell Chem Biol* **27**, 986–997 (2020).
6. 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).
7. 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).
8. 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).
9. Brent, R. & Ptashne, M. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729–736 (1985).

10. Brzovic, P. S. *et al.* The acidic transcription activator Gcn4 binds the Mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell* **44**, 942–953 (2011).
11. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat Chem Biol* **11**, 891–894 (2015).
12. Poss, Z. C., Ebmeier, C. C. & Taatjes, D. J. The Mediator complex and transcription regulation. *Crit Rev Biochem Mol Biol* **48**, 575–608 (2013).
13. Dyson, H. J. & Wright, P. E. Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. *J Biol Chem* **291**, 6714–6722 (2016).
14. 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).
15. 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).
16. Conaway, R. C. & Conaway, J. W. Origins and Activity of the Mediator Complex. *Semin Cell Dev Biol* **22**, 729–734 (2011).
17. Lee, H.-K., Park, U.-H., Kim, E.-J. & Um, S.-J. MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J* **26**, 3545–3557 (2007).
18. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature Structural and Molecular Biology* **18**, nsmb.1997 (2011).
19. Ries, D. & Meisterernst, M. Control of gene transcription by Mediator in chromatin. *Seminars in Cell & Developmental Biology* **22**, 735–740 (2011).

20. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).
21. Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* **45**, 6873–6888 (2006).
22. Berlow, R. B., Dyson, H. J. & Wright, P. E. Hypersensitive termination of the hypoxic response by a disordered protein switch. *Nature* **543**, 447–451 (2017).
23. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113–119 (1987).
24. 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 U S A* **90**, 883–887 (1993).
25. 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).
26. Fadda, E. & Nixon, M. G. The transient manifold structure of the p53 extreme C-terminal domain: insight into disorder, recognition, and binding promiscuity by molecular dynamics simulations. *Phys Chem Chem Phys* **19**, 21287–21296 (2017).
27. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
28. Robustelli, P., Piana, S. & Shaw, D. E. Mechanism of Coupled Folding-upon-Binding of an Intrinsically Disordered Protein. *J Am Chem Soc* **142**, 11092–11101 (2020).
29. Warfield, L., Tuttle, L. M., Pacheco, D., Klevit, R. E. & Hahn, S. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc Natl Acad Sci U S A* **111**, E3506-3513 (2014).
30. Arnold, C. D. *et al.* A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J* **37**, (2018).

31. Brzovic, P. S. *et al.* The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell* **44**, 942–953 (2011).
32. Tuttle, L. M. *et al.* Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep* **22**, 3251–3264 (2018).
33. Shammas, S. L., Travis, A. J. & Clarke, J. Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc Natl Acad Sci U S A* **111**, 12055–12060 (2014).
34. 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).
35. 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).
36. Wang, N. *et al.* Ordering a Dynamic Protein Via a Small-Molecule Stabilizer. *J. Am. Chem. Soc.* **135**, 3363–3366 (2013).
37. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator–coactivator complexes using a covalent small molecule ligand. *Proc Natl Acad Sci U S A* **111**, 12061–12066 (2014).
38. Mir, M. *et al.* Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. (2018) doi:10.1101/377812.
39. Alberti, S. Phase separation in biology. *Curr Biol* **27**, R1097–R1102 (2017).
40. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285–298 (2017).

41. Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**, 420–435 (2018).
42. El Khattabi, L. *et al.* A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell* **178**, 1145-1158.e20 (2019).
43. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13–23 (2017).
44. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, (2018).
45. Hahn, S. Phase Separation, Protein Disorder, and Enhancer Function. *Cell* **175**, 1723–1725 (2018).
46. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the reality. *Chem Biol* **21**, 1102–1114 (2014).
47. Cesa, L. C. *et al.* Inhibitors of Difficult Protein-Protein Interactions Identified by High Throughput Screening of Multi-Protein Complexes. *ACS Chem Biol* **8**, 1988–1997 (2013).
48. Law, S. M., Gagnon, J. K., Mapp, A. K. & Brooks, C. L. Prepaying the entropic cost for allosteric regulation in KIX. *PNAS* **111**, 12067–12072 (2014).
49. Wang, F. *et al.* Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc Natl Acad Sci U S A* **109**, 6078–6083 (2012).
50. Odoux, A. *et al.* Experimental and molecular dynamics studies showed that CBP KIX mutation affects the stability of CBP:c-Myb complex. *Comput Biol Chem* **62**, 47–59 (2016).
51. Lodge, J. M., Majmudar, C. Y., Clayton, J. & Mapp, A. K. Covalent Chemical Cochaperones of the p300/CBP GACKIX Domain. *ChemBioChem* **19**, 1907–1912 (2018).

52. Breen, M. E. & Mapp, A. K. Modulating the masters: chemical tools to dissect CBP and p300 function. *Curr Opin Chem Biol* **45**, 195–203 (2018).
53. Smith, M. C. & Gestwicki, J. E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **14**, e16 (2012).
54. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nature Structural and Molecular Biology* **18**, nsmb.1999 (2011).
55. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* **3**, 301–317 (2004).
56. Papaleo, E. *et al.* The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem Rev* **116**, 6391–6423 (2016).
57. Tsai, C.-J., Sol, A. del & Nussinov, R. Protein allostery , signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Molecular BioSystems* **5**, 207–216 (2009).
58. Abrusán, G. & Marsh, J. A. Ligand-Binding-Site Structure Shapes Allosteric Signal Transduction and the Evolution of Allostery in Protein Complexes. *Mol Biol Evol* **36**, 1711–1727 (2019).
59. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**,.
60. Garbaccio, R. M. & Parmee, E. R. The Impact of Chemical Probes in Drug Discovery: A Pharmaceutical Industry Perspective. *Cell Chem Biol* **23**, 10–17 (2016).
61. Henderson, A. R. *et al.* Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *PNAS* 201806202 (2018) doi:10.1073/pnas.1806202115.

Chapter Two

Sequence Differences within Dynamic Regions of AcID Motifs Alter Conformation of Complexes with Activators *

2.1 Abstract

Activator binding domains (ABDs) are the molecular recognition units of transcriptional coactivators, allowing for a single protein to interact with a diverse array of transcriptional activators. While most ABDs are helical bundles, one ABD stands out with a distinct structure-the Activator Interaction Domain (AcID). AcID, found in Med25 and PTOV1, consists of a seven-stranded β-barrel flanked by dynamic substructures including helices and loops. Recent work in our group has demonstrated the role of these dynamic substructures in molecular recognition of distinct activator binding partners. That is, the dynamic loops and helices allow for Med25 AcID to access unique conformations in response to each binding partner. Our objective was therefore to determine how sequence changes in the dynamic regions of three AcID paralogs ultimately affects binding. Using a biophysical approach, we determine how sequence deviations alter the binding affinities, mechanisms and conformations of the AcID motif. We demonstrate that, despite high sequence homology, there is a loss in secondary structure of the PTOV1 AcIDs, as well as a decrease in thermal stability. Despite the changes in sequence of the dynamic regions and structural instability, however, all three AcIDs are capable of recognizing the same binding partners with only moderate changes in affinity. Further, using transient stopped-flow kinetics, we observed the binding modes of the different AcID motifs. Specifically, despite being paralogs, we show

different binding modes for each AcID motif, suggesting that changes in the overall dynamic regions are specific to each AcID without changing overall binding affinities.

2.2 Introduction

The transient protein-protein interactions (PPIs) formed between transcriptional activators and coactivators plays an essential role in the initiation of transcription.[1,2] Coactivators generally act as hubs in transcriptional initiation, interacting with DNA-bound transcriptional activators and with other coactivators to assemble the transcriptional machine.[3,4] A single coactivator is capable of recognizing a variety of transcriptional activators using its activator binding domain (ABD). [4,5] Dysregulation of these PPIs can result in a variety of disease states, including oncogenesis, viral infection, and tumor metastasis.[6–9] Although there have been many biophysical studies of ABD PPIs, general molecular recognition principles for complex formation have only recently begun to emerge.[10–12] These principles are critical for both understanding transcription and for the discovery of small molecules that modulate PPI networks.

Most often, ABDs are comprised of helices connected by dynamic loops.[11] In contrast, the coactivator Med25 contains an ABD that deviates from the helical bundle model, instead containing a core seven-stranded β-barrel flanked by three alpha-helices (Fig. 2.1). This ABD is the Activator Interaction Domain (AcID). Previous work by others suggested that the β-barrel is where most of the interactions with activators occur.[14,15] In contrast, recent work in the Mapp lab indicates that the β-barrel may serve as a central scaffold, but the dynamic loops and helices flanking the barrel make critical and specific contacts with activators (Fig. 2.1).[11]

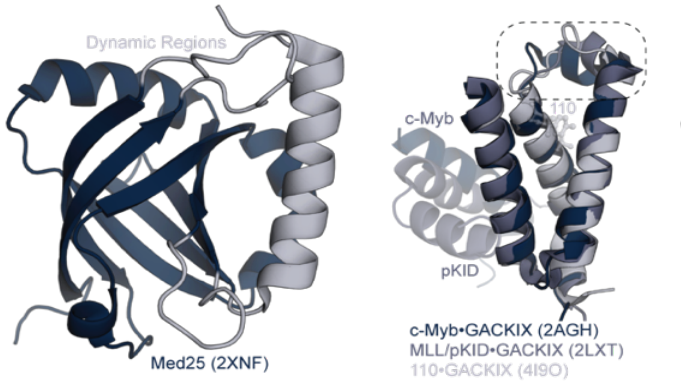


Figure 2.1: Dynamic substructures of ABDs make critical contacts with binding partners. On the left is Med25 AcID, comprised of a seven-stranded β-barrel flanked by three α-helices and loops. While having a β-barrel scaffold, it still contains dynamic substructures. On the right are the different conformations of CBP GACKIX when bound to different binding partners. The biggest changes in shape are observed in the top loop region.[16,17]

The work in this chapter aims to dissect how dynamic substructures affects molecular recognition within AcID motifs. Specifically, we look at paralogs of AcID, found in Med25 and PTOV1, a second protein in humans that contains two tandem AcID motifs.[18,19] These proteins share high sequence identity, 81% and 71% respectively to Med25 AcID, but deviate in sequence in their dynamic regions (Fig. 2.2). Using a biophysical approach, we explore how these sequence deviations alter the binding mechanisms in complexing with transcriptional activators. The guiding hypothesis was that the sequences differences would influence binding selectivity and also the conformational changes of the resulting complexes. In the former case, this can be tested by measuring the equilibrium binding affinities for a panel of activators. In the latter, transient kinetic studies can provide insight into the conformational changes that might occur with complex formation. We show that while overall the binding affinities and mechanisms are conserved, we see a loss in conformation states. While not the original hypothesis, these results demonstrate that each AcID paralog demonstrates a unique set of conformations, suggesting there is a degree of specificity in binding.

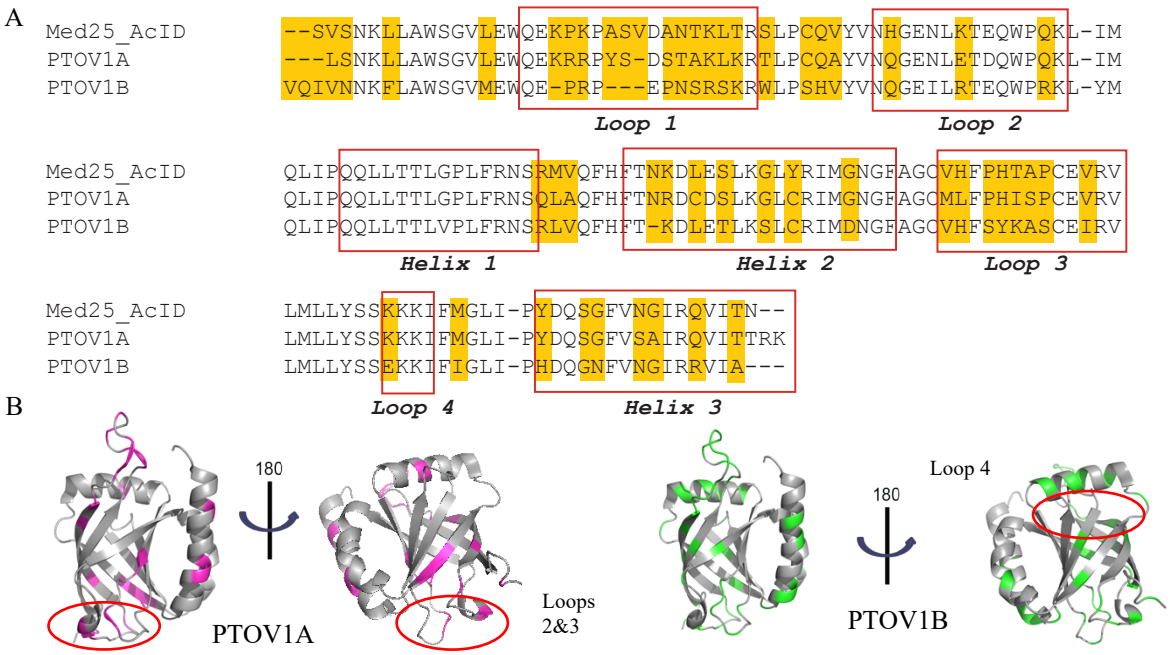


Figure 2.2: Sequence alignment of the AcID paralogs. A) Shown in yellow are the sequences differences. Boxed are the dynamic substructure regions. Sequence alignments of the AcID homologs reveals that PTOV1A shares 81% sequence identity to Med25 AcID, and PTOV1B shares 71% identity. Specifically, while there are some deviations, the β-barrel remains largely conserved. In contrast, residue deviations, such as charge flips and loss of charges, are found in the dynamic loop and helical regions. B) Sequence deviations of PTOV1A (right) and PTOV1B (left) mapped onto Med25 AcID's two binding faces. Circled in red are the loops and helices that are implicated in binding.

2.3 Results and Discussion

2.3.1 Structural characterization of the AcID paralogs reveals loss in secondary structure

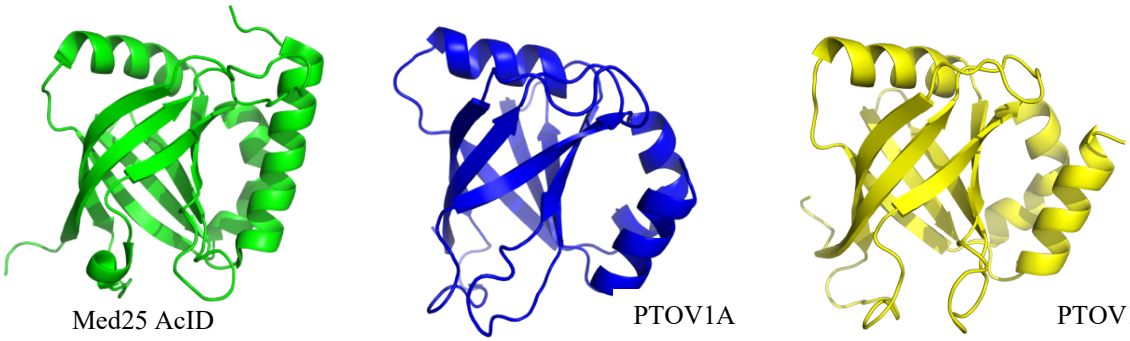


Figure 2.3: Structures of the AcID motifs. Med25 AcID's structure was solved via NMR and is shown to have a β-barrel flanked by three α-helices (PDB 2XNF).[14] Threading models predict the structure of the two PTOV1 AcIDs. The overall barrel region is conserved, while there appears to be a loss in secondary structure, specifically with the helices, resulting in more dynamic loops.[20]

Previous work has determined that Med25 AcID is comprised of a core seven-stranded β-barrel flanked by three α-helices and loops. While there are no structures of the PTOV1 AcID

motifs, threading models created by *I-TASSER* demonstrate an overall structural conservation in the β-barrel, with a slight loss in structure in the helices (Fig. 2.3).[20] However, circular dichroism (CD) scans suggests loss in the barrel region for the A and B domains at 208 nm, as well as a slight loss in other secondary structures at 220 nm, consistent with threading models (Fig. 2.4).

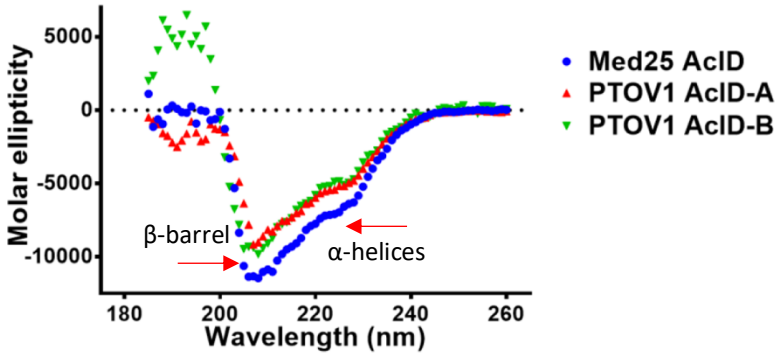


Figure 2.4: CD scans of the AcID motifs. Illustrated above are the collected CD spectra of the A (Red) and B (Green) AcID domains in PTOV1 and Med25 AcID (Blue). The β-barrel, appearing at a wavelength of 209 nm, appears more conserved than the helical reigons, measured at 222 nm. Data was obtained by Olivia Pattelli and Dr. Matthew Beyersdorf.

In order to determine protein stability of the AcID homologs, CD thermal scans were conducted to determine the melting temperature. Thermal melts were conducted and monitored at 222 nm and 208 nm. Med25 AcID has a melting temperature of approximately 70 ˚C. Despite high sequence identity with Med25 AcID, PTOV1A and PTOV1B had melting temperatures of 61 ˚C and 66 ˚C respectively. Med25 AcID's thermal melt displayed two phases, with the first phase corresponding to the fraying of the helices, and the second phase corresponding to the barrel unraveling. Only one phase was observed for PTOV1A and PTOV1B. Taken together, these data suggest there is a loss in secondary structure in the PTOV1 AcID motifs comparatively to Med25 AcID. Given that PTOV1 appears to be more unstructured than Med25, and has a lower melting temperature overall, we predicted that it would be difficult to observe conformational changes of PTOV1 in complex with activators, due to the likely very low energy differences between conformations.

2.3.2 Activators preferentially bind to individual activator paralogs

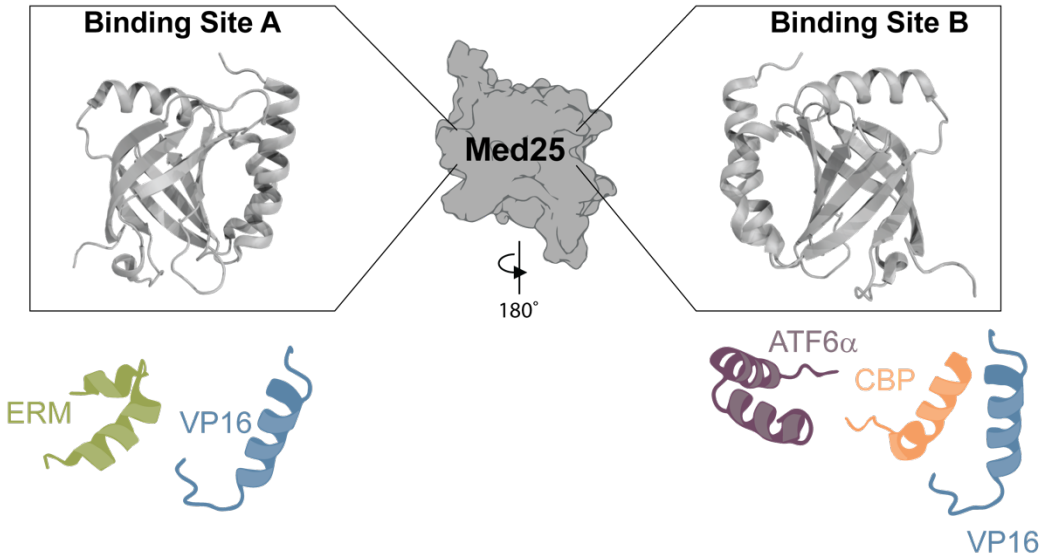


Figure 2.5: The two binding faces of AcID interact with different binding partners. Binding site A, termed the H1 face, interacts with ERM from the ETV/PEA3 family of transcription factors as well as one of the two N-terminal TADs of VP16.[14,21] Binding site B, termed the H2 face, binds to the second VP16 TAD as well as ATF6α, responsible for hypoxic stress response.[14,22]

We wanted to test if the different AcID paralogs bound to the same binding partners, and if so, with what affinity. PTOV1A and PTOV1B were tested against Med25 AcID's TAD binding partners in an equilibrium binding experiment (Fig. 2.5). More specifically, we used the TADs VP16, ERM, and ATF6α. ERM belongs to the ETV/PEA3 family of transcription factors and is implicated in cancer metastasis.[7,21,23,24] The herpes simplex protein VP16, which contains two tandem TADs that bind to both faces of AcID is responsible for activating transcription of viral genes during an infection.[8] Lastly, ATF6α is an endoplasmic reticulum stress response transcription factor involved in the unfolded protein response under hypoxic conditions.[22]

The-FITC-labeled TADs were tested against the PTOV1 AcID motifs via fluorescence polarization (FP) assays (Table 2.1). The activators were all observed to bind to PTOV1A/B, which is not surprising due to the high degree of sequence conservation in the β-barrel. Prior work in our lab suggests that the β-barrel plays a significant role in binding partner affinity. However, there were differences in the exact values. Of note is the difference in ATF6α's affinity for

PTOV1A and PTOV1B. There is a significant difference in affinity, with PTOV1A having almost 6-fold affinity over PTOV1B. This is likely due to the Lys-rich loop located on the H2 binding face, the binding site of ATF6α. This loop has been demonstrated to be highly critical for activator binding. In Med25 AcID, there are three consecutive Lys residues, conserved in PTOVA, located on the H2 binding face, the binding site of ATF6α. This Lys-rich loop has been demonstrated to be highly critical for activator binding. In Med25 AcID, there are three consecutive Lys residues, conserved in PTOV1A. In PTOV1B, however, the last Lys is a Glu residue. This alters the surrounding electrostatics with the change of charge in the loop. As discussed below, electrostatics play a critical role in binding affinity (Fig. 2.6).

Table 2.1: Apparent equilibrium K_d's. Kd values were measured using fluorescence polarization assay. Serial diluted protein was complexed with 20 nM FITC-labeled peptide. Data is in triplicate and the error is reported as SDOM.

	MED25	PTOV1A	PTOV1B
ATF6A	0.52 ± 0.6 μM	0.20 ± 0.01μM	1.2 ± 0.2 μM
ERM	0.59 ± 0.02 μM	0.42 ± 0.04 μM	0.43 ± 0.03 μM
VP16	0.06 ± 0.004 μM	0.03 ± 0.002 μM	0.02 ± 0.005 μM

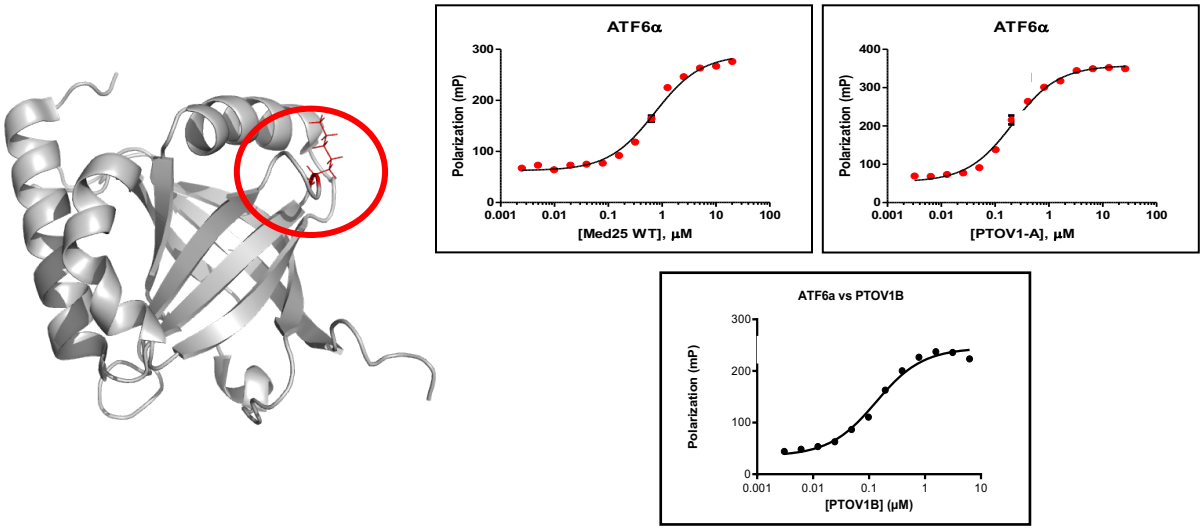


Figure 2.6: Key differences and binding curves of the AcID motifs with ATF6α. On the left is Med25 AcID with the triple Lys loop circled in red. K520 is a Glu residue in PTOV1B (shown as the red side chain). On the right are the three binding curves for Med25 AcID, PTOV1A and PTOV1B. PTOV1B has almost 6-fold affinity for ATF6α over Med25 and PTOV1A, which is likely due to the residue changes in the Lys-rich loop on the H2 face.

Additionally, VP16 demonstrates a higher affinity for PTOV1A and PTOV1B over Med25 AcID. VP16 binds to both the H1 and H2 faces, using its two N-terminal TAD regions with an intrinsically disordered linker that wraps around the β-barrel to contact both faces (Fig. 2.7). The loss in affinity may be due to the loss in secondary structure, as shown by the predicted structures and CD data (Fig. 2.3 and 2.4). With less secondary structure, VP16 can wrap and contact both sites easier and create more points of contact, resulting in a higher binding affinity. Additionally, there is an increase in hydrophobic residues, particularly on the H1 face for PTOV1A and PTOV1B. Residue M470 is a Leu, thus increasing the overall hydrophobicity in the groove VP16 binds to on the H1 face.

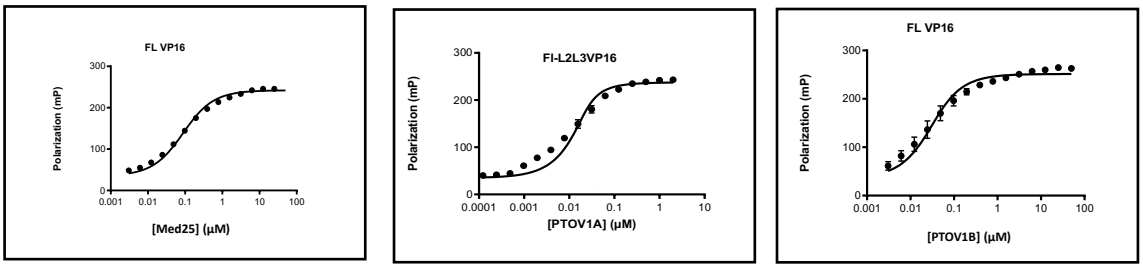


Figure 2.7: Binding curves of the AcID motifs with VP16. VP16 binds to the PTOV1 AcID motifs with a tighter binding affinity than to Med25 AcID. This is likely due to the loss in secondary structure, allowing VP16 to wrap around the barrel region and have more contact, thus increasing the binding affinity.

2.3.3 Transient stopped-flow kinetics offer insight into PTOV1A complex formation

Electrostatics play an important role in initial binding of TAD•ABD interactions (Fig 2.8). Data generated by Dr. Nicholas Foster in the Mapp lab demonstrates that increasing NaCl concentrations attenuates TAD affinity for Med25 AcID. Additionally, salt ions can play a critical role in binding affinity. Specifically, the use of Mg^{2+} cation was shown to critically affect activator binding to Med25 AcID, showing a 13-fold decrease in activator binding over Na^{+}. However, electrostatics alone do not dictate binding, as recent data has shown that conformational changes within the coactivators driven through dynamic regions can modulate molecular recognition.[10,11]

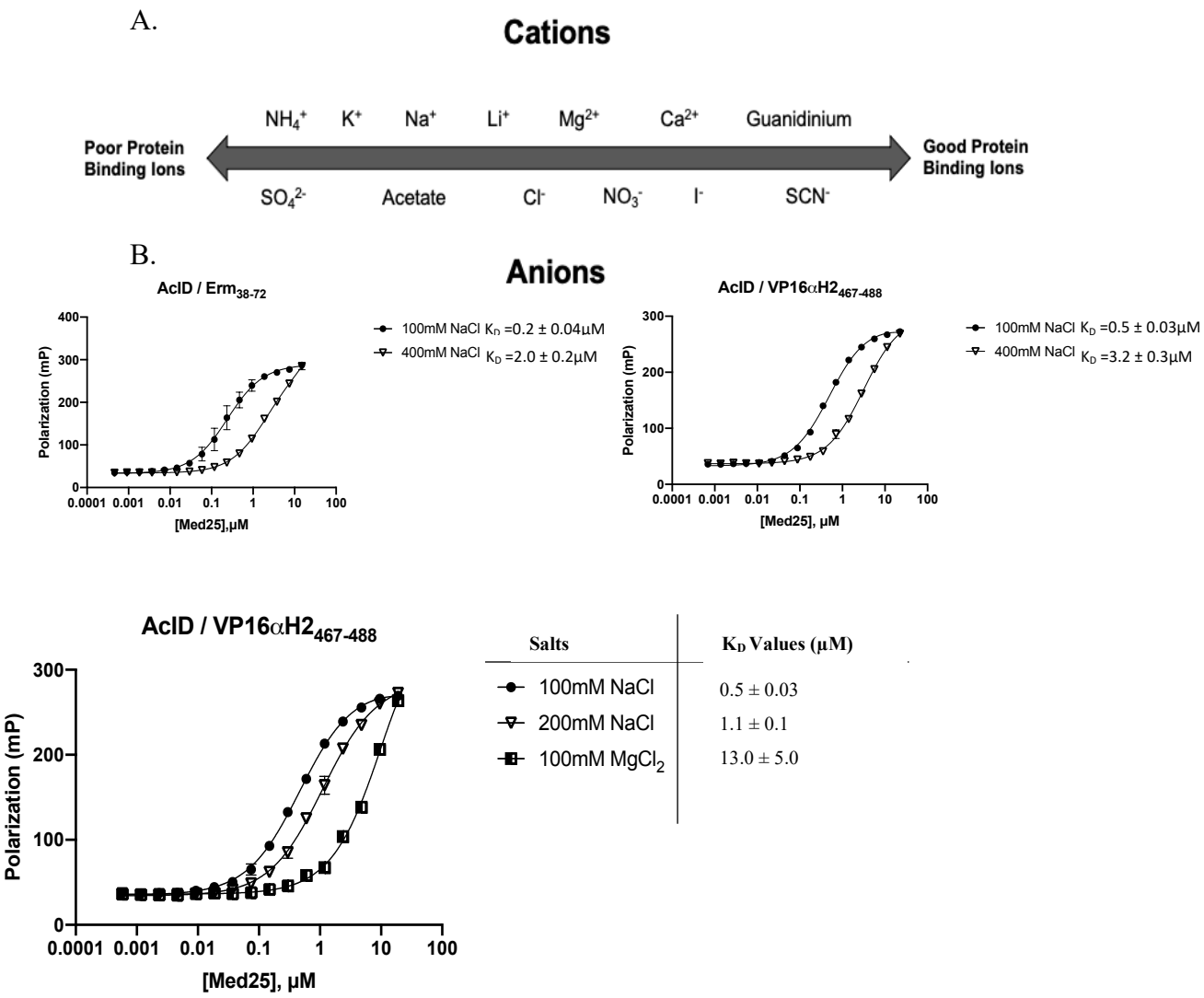


Figure 2.8: The role of electrostatics in activator binding. A) The Hofmeister Series is a classification of how different ions affect protein interactions. B) FP traces demonstrating how different salt ions affect binding. Not only does higher NaCl concentrations attenuate binding, but the type of ion can greatly alter binding as well. Data was collected and analyzed by Dr. Nicholas Foster.

To test this hypothesis, transient stopped-flow fluorescence spectroscopy was utilized to examine the underlying mechanistic features of PTOV1A in comparison to Med25 AcID using the VP16, ERM and ATF6α TADs (see Methods for details). By determining the association mechanism of PTOV1A•activators, we can calculate the dissociation constant, microscopic rate constants, and rate and equilibrium constants associated with conformational changes. The system was originally developed by Dr. Matthew Henley for Med25 AcID. In order detect activator-coactivator interactions, the environmentally sensitive fluorophore 4-*N,N*-dimethylamino-1,8-napththalimide (4-DMN) was used. 4-DMN, conjugated to β-alanine, was incorporated to the N-terminus of the TAD activators.[11] For the work discussed this thesis, the system was further optimized for detection of PTOV1A. Figure 2.9 depicts the overall experimental layout.

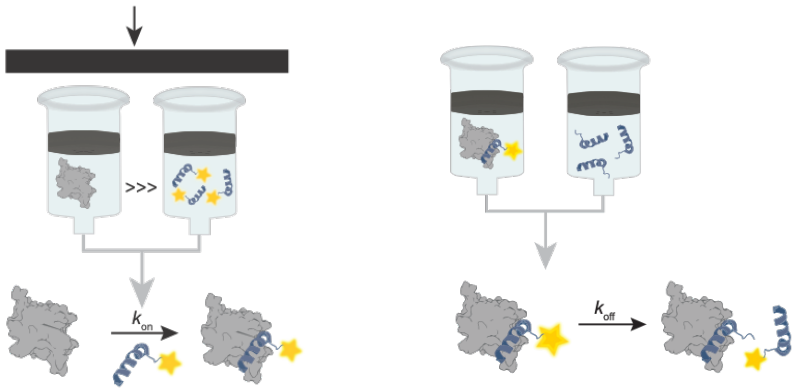


Figure 2.9: Design of the transient stopped flow experiments. On the left depicts how we measured k_{on}. Increasing concentrations of protein were rapidly mixed with DMN-labeled tracer and fit using the equation outlined in the methods section below. For the k_{off} experiments (right), protein was precomplexed with DMN-labeled peptide and competed off using unlabeled peptide.

ERM binds to the H1 face of PTOV1A with a distinct conformational ensemble

Analysis of HSQC NMR data has revealed that ERM, a member of the ETV/PEA3 transcriptional activators binds to the H1 binding face of Med25 AcID. Specifically, ERM was found to localize to a groove on the H1 binding face, with predominate perturbations on α3 and β5 (Fig. 2.10).[21] Mutations to residues R538, Q451, and K411 inhibit ERM binding, but left H2

binding largely unaffected.[11] Interestingly, residues R538 and Q451, located on α3 and β3 respectively, are conserved in PTOV1A. K411, located in the loop between β1 and β2, is changed to an Arg residue in PTOV1A, thus maintaining a positive charge, but the terminal amino is removed. Additionally, changes in neighboring residues, such as K440 to a Glu residue cause a change in charge (Fig. 2.2). We hypothesized that due to the deviations of the dynamic regions in PTOV1A, there will be distinct changes to the conformations observed.

ERM•Med25 rate constants were determined using global fitting in Kintek Explorer and an "inverted" association experiment in order to validate the three-step mechanism observed. PTOV1A was, on the other hand, shown to undergo complex formation within the timescale of the instrument. Figure 2.11 shows the association and dissociation experiments. PTOV1A had a k_{on} two-fold faster than Med25 AcID, 780 $\mu M^{-1}s^{-1}$ and 300 $\mu M^{-1}s^{-1}$ respectively, but showed very similar dissociation constants, 360 s^{-1} for PTOV1A and 380 $s^{-1}$for Med25 AcID (Fig 2.11a-e). The fast on rate is most likely the result of enhanced electrostatic interactions.

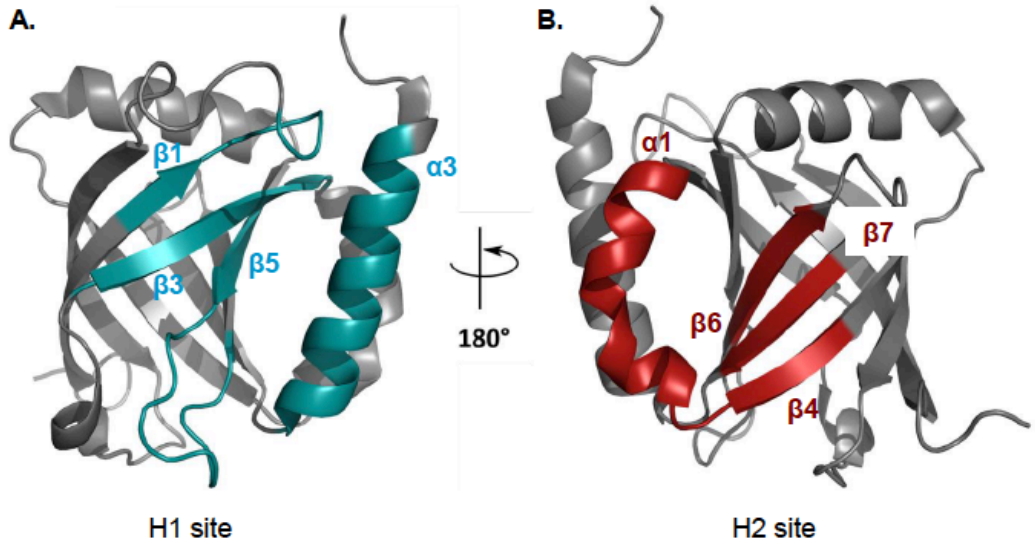


Figure 2.10: The binding faces of AcID. A) The H1 binding site is formed by β1-β3-β5 and α2 (shown in teal). Activators such as ERM and one of the TADs of VP16 bind to this site.[14,15] B) The H2 binding site is formed by β6-β7-β4 and α1 (shown in red). ATF6α and the other VP16 TAD bind to this face.[11,14]

Only one conformation state is observed, with 100% population state observed after binding. This is quite different from Med25 AcID, which has three very clear conformational changes. While it is highly likely that PTOV1A is undergoing additionally conformational states, the low energy differences between the different conformers is masking the identifiable changes between states. This is consistent with the observation that PTOV1A is more unstructured.

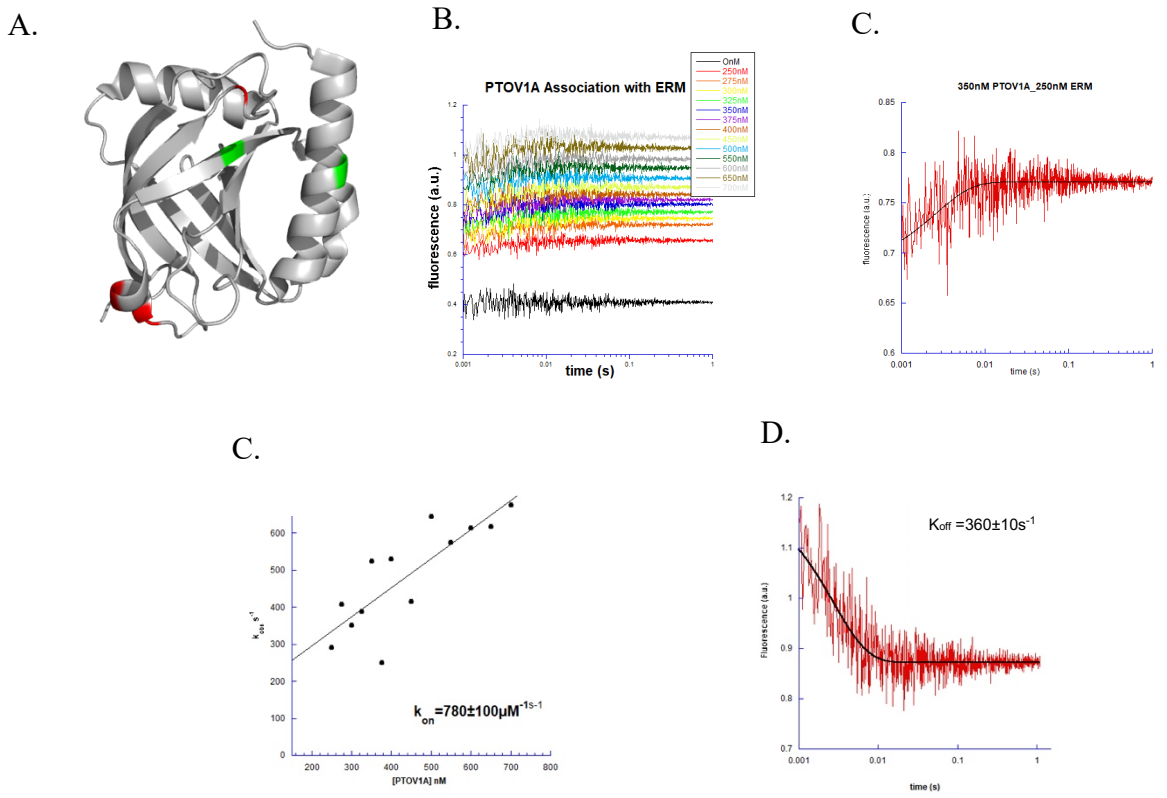


Figure 2.11: Experiments used to define PTOV1A•ERM binding parameters. A) The H1 binding face of Med25 AcID. In green are the key perturbed residues that are conserved in PTOV1A, and in red are the differences, such as K411R, K440E and Y486, which is a Cys in PTOV1A [15,21] B) Plotted association traces of PTOV1A•ERM. C)Representative association trace of 350 nM PTOV1A and 250 nM ERM, fit to a single exponential. D) Plotted k_{obs} to determine $k_{on.}$ D) Dissociation trace of PTOV1A•4-DMN-ERM. 2 μM PTOV1A was precomplexed with 0.5 μM DMN-labeled ERM, and then rapidly mixed with 25 μM unlabeled ERM. The trace was fit to a single exponential. For reference, Med25 AcID•ERM associationa and dissociation traces were both fit to triple exponentials.[11]

Binding at the H2 face suggests stabilization of energetically favorable conformer upon binding

The transcriptional activator ATF6α, responsible for initiating the unfolded protein response pathway under hypoxic stress, binds to Med25 AcID on the H2 binding face, opposite

that of ERM (Fig. 2.10).[9,11] Key residues R466 and M523, located on α2 and β7 respectively, were shown to undergo significant chemical shift perturbations in HSQC NMR experiments, and mutations to these residues disrupted ATF6α binding. Both residues are conserved in PTOV1A, with neighboring residues. However, given the proposed increase in malleability due to the loss of structure, we propose there will be a loss in observed conformational states (Fig. 2.12a).

PTOV1A had a k_{on} half that of Med25 AcID when bound to ATF6α, 310 $\mu M^{-1}s^{-1}$ and 610 $\mu M^{-1}s^{-1}$ respectively. Additionally, the dissociation rate constant was 2-fold slower, 140 s^{-1} for Med25 and 49 s^{-1} for PTOV1A, leading to a similar K_d for both Med25 and PTOV1A (Fig. 2.12). Med25 AcID underwent three distinct conformational changes, with the first conformation state being consisting of the highest population (>80%). Only one conformation state was detected for PTOV1A in the stopped flow kinetics experiments. However, this may be due to the other populations being too small to detect with current methods, or that the timescale of exchange was too fast to observe. This is likely due to the loss in secondary structure, as described above.

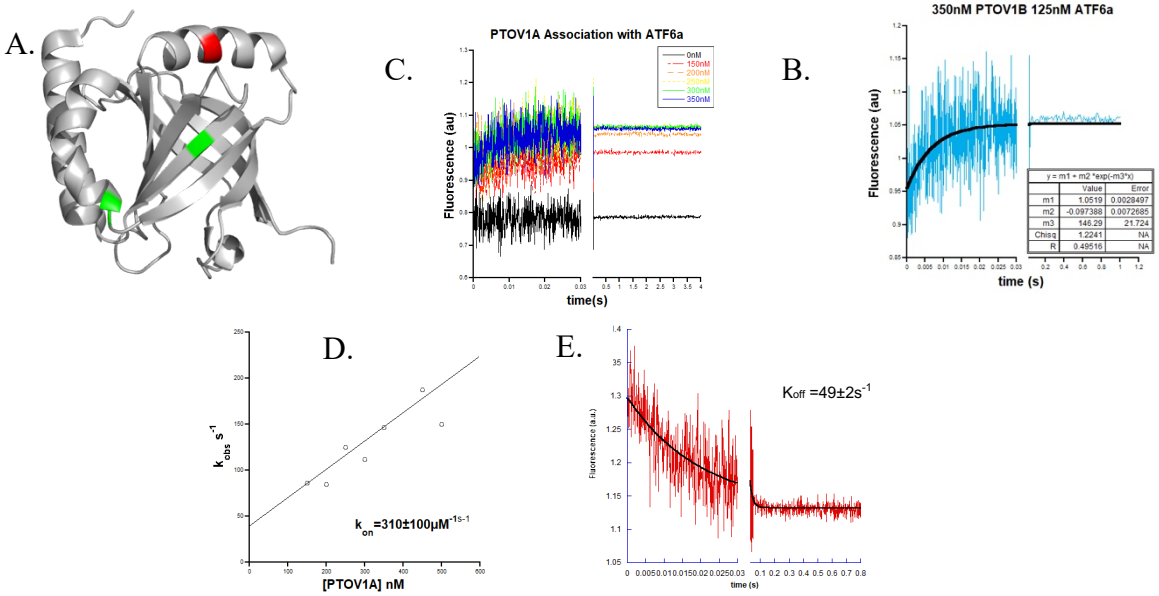


Figure 2.12: Experiments used to define PTOV1A•ATF6α binding parameters. A) The H2 binding face of Med25 AcID. In green are the key perturbed residues that are conserved in PTOV1A, and in red is Y486, which is a Cys in PTOV1A [13,16,21] B) Plotted association traces of PTOV1A• ATF6α. C) Representative association trace of 350 nM

PTOV1A and 125 nM ATF6α, fit to a single exponential. D) Plotted k_{obs} to determine $k_{on.}$ D) Dissociation trace of PTOV1A•4-DMN- ATF6α. 1 μM PTOV1A was precomplexed with 0.25 μM DMN-labeled ATF6α, and then rapidly mixed with 25 μM unlabeled ATF6α. The trace was fit to a single exponential. For reference, Med25 AcID• ATF6α assciation and dissociation traces were both fit to triple exponentials.[11]

A TAD that binds to both faces reveals complex conformational changes

The two amino-terminal TADs of VP16, termed H1 and H2, bind to both faces of VP16, overlapping with the binding regions of ERM and ATF6α.[8,14] NMR titration experiments revealed that the H1 TAD of VP16 binds to a groove in Med25 formed by β1-β3-β5 and α2 on the H1 binding face. The H2 TAD was found to bind to a groove on the opposite H2 binding face formed by β6-β7-β4 and α1 (Fig. 2.10).[14] It is hypothesized that the loss in structural order of PTOV1A compared to Med25 AcID, in addition to VP16 binding to both faces over a greater area, will result in significant changes in rate constants.

Unlike Med25 AcID, PTOV1A complexed with VP16 was shown to undergo additional kinetic phases. An 'inverted experiment' was used to the support the mechanism of conformational changes after binding to Med25 AcID. Additionally, global fitting was completed using the calculated rate constants from the standard experiments in order to further validate the model. Figure 2.13 illustrates the results of PTOV1A association and dissociation complex formation with VP16. PTOV1A's k_{on} was approximately half that of Med25, 1,100 $\mu M^{-1}s^{-1}$ for Med25 and 650 $\mu M^{-1}s^{-1}$ for PTOV1A. Interestingly, PTOV1A•VP16 had a K_d that was 4.6-fold tighter than Med25 AcID.

The PTOV1A and Med25 were shown to undergo two conformational changes. The fast phase is linearly dependent upon protein concentration. However, due to the amplitudes of change being too small to accurately derive rate parameters, we were unable to calculate population distributions. Interestingly, these data suggest that the PTOV1A•VP16 system is the most similar to that of Med25 AcID in terms of conformational ensembles. Despite a slower $k_{on,}$ we observed a

k_{off} slower for PTOV1A than Med25 AcID, 7.1 s^{-1} versus 102 s^{-1}, leading to a significantly tighter K_d. Moreover, there is a +2 charge difference between PTOV1A (+9) and Med25 AcID (+7), which is likely resulting in the slower association rate given the critical role of electrostatics.

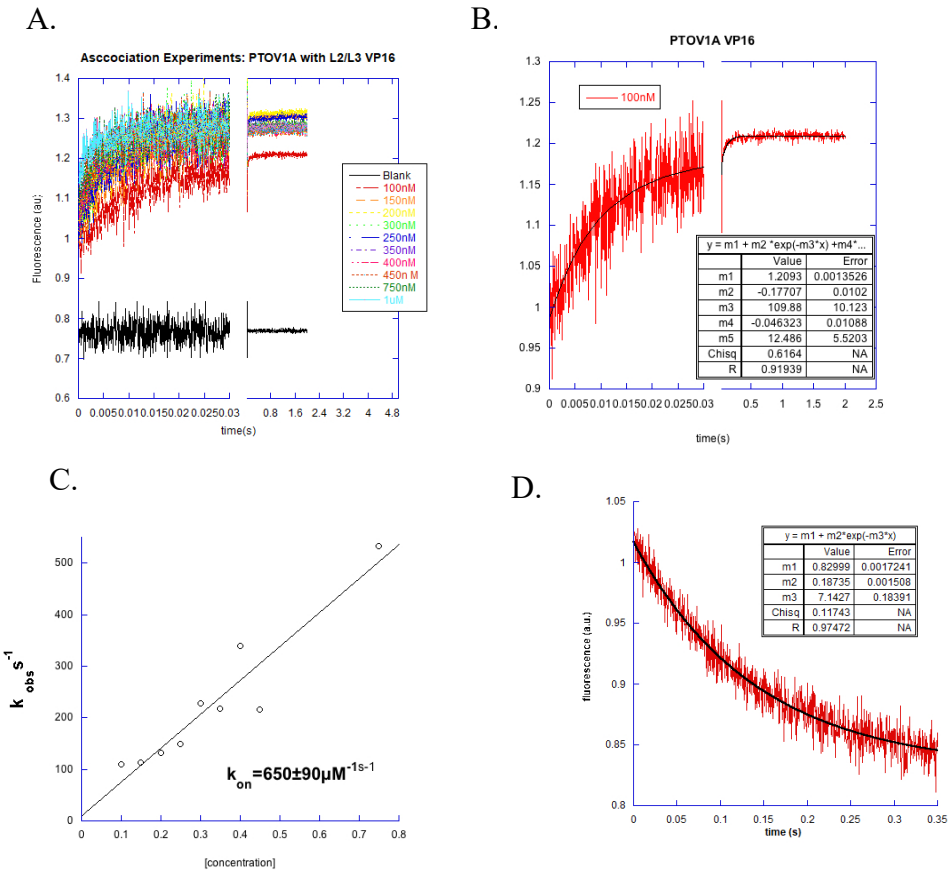


Figure 2.13: Experiments used to define PTOV1A•VP16 binding parameters. A) Plotted association traces of PTOV1A• VP16. B) Representative association trace of 100nM PTOV1A and 100 nM VP16, fit to a double exponential. C) Plotted k_{obs} to determine k_{on}. D) Dissociation trace of PTOV1A•4-DMN- ATF6α. 250 nM PTOV1A was precomplexed with 100 nM DMN-labeled VP16, and then rapidly mixed with 15 µM unlabeled VP16. The trace was fit to a single exponential. For reference, Med25 AcID• VP16 association and dissociation traces were both fit to double exponentials.

Complex formation with helical coactivators and activators typically proceeds by fast association and dissociation rate constants. Previous work by Dr. Matt Henley and the Mapp lab demonstrates that Med25 AcID forms complexes consistent with this observation. Med25 AcID had k_{on} values that were 1-2 orders of magnitude faster than previously observed with most other activator•coactivator systems, likely due to the role of electrostatics (Fig. 2.8). PTOV1A complex

formation demonstrated similar trends, with noticeable differences in complex dissociation, often forming tighter complexes. While Med25 AcID complexes showed k_{on} values ranging from 300 to 1,100$\mu M^{-1}s^{-1}$, and k_{off} values between 100 and 400s^{-1}, PTOV1A complex formation was often slower than Med25, but had k_{off} and K_D values that were for some interactions, over 4-fold tighter than that of Med25 AcID. Taken together, these data suggest that while PTOV1A does not demonstrate significantly different binding mechanisms, the changes in sequence alter the conformational ensembles observed. This is likely due to the loss in structure and thermal stability, which in turn can lower the energy differences between conformation states and thus masking changes between conformation states.

Table 2.2: Calculated rate constants for PTOV1A and Med25 AcID. Values were measured using transient stopped flow kinetics, and are the average of 2-3 traces. Reported error is SDOM.

Activators	PTOV1A k_{on} ($\mu M^{-1}s^{-1}$)	PTOV1A k_{off} (s^{-1})	Med25 k_{on} ($\mu M^{-1}s^{-1}$)	Med25 k_{off} (s^{-1})
ERM	780 ± 100	360 ± 10	~300	380 ± 40
ATF6α	310 ± 100	49 ± 2	610 ± 40	140 ± 7
VP16	650 ± 90	7 ± 1	1,100 ± 100	102 ± 6

2.4 Conclusions and Assessment

Coactivators must be able to quickly bind transcriptional activators in a specific manner to initiate transcription. Adding to the complexity of seemingly featureless binding interfaces, coactivators must be able to bind a multitude of binding partners, each with a different outcome. These ‘fuzzy interactions’ have long been termed undruggable due to the evasive nature of their While it was originally proposed that binding was driven through electrostatic and hydrophobic interactions, this model does not account for the intricate specificity and selectivity required for such a highly regulated process as transcription initiation. Towards this end, recent work in the Mapp lab has demonstrated that molecular recognition is driven through conformational changes

driven through dynamic regions. In some cases, such as that observed with CBP GACKIX, upon activator binding to the ABD, there is a winnowing of conformers, allowing a unique state specific to each interaction.[10,25]

Using AcID paralogs, we set out to determine how sequence deviations in the dynamic substructures affected binding mechanisms. In this chapter we demonstrate that there is a predicted loss in structure for the AcID motifs. The guiding hypothesis was that the sequences differences would influence binding selectivity and also the conformational changes of the resulting complexes. Using equilibrium binding experiments, we observed that there were only modest changes in binding affinity. Transient stopped-flow kinetics provided insight into the binding mechanisms of PTOV1A and Med25 AcID. We observe that while there is not a significant change in mechanism, the sequence deviations result in change of conformational plasticity. This is directly observed by PTOV1A undergoing less conformation transitions upon activator binding. While there are most likely some conformation states that cannot be detected by current available methods, it suggests that there is a differential response upon binding to PTOV1A, thus indicating that these changes in secondary structure alter the binding modes for AcID•activator interactions.

Taken together, the work in this chapter demonstrates that sequences changes in the dynamic substructures alter the conformations. Given that PTOV1 has been observed to outcompete Med25 for binding to CBP when overexpressed in metastatic prostate cancer, these changes in sequence are highly implicated in function. Specifically, the changes in conformational plasticity as a result of the sequence deviations may alter transcriptional output by altering the AcID•activator complex formation timescale. Due to its large role in many cancer states, it would

be critical moving forward to determine PTOV1's role in modulating transcription, as well as further dissect its antagonistic behavior against Med25.

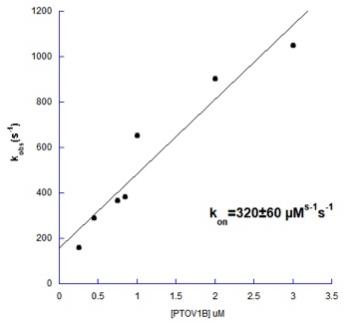


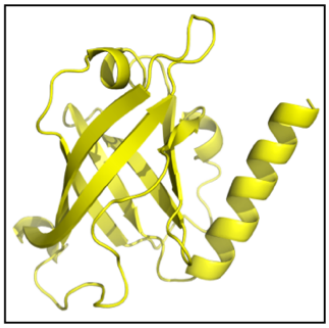
Figure 2.14: Plotted k_{obs} of PTOV1B•ERM. Above is the plotted, fast linear phase PTOV1B associating to ERM. ERM was fit to a double exponential, compared to Med25 AcID, which was fit to a triple exponential.

Moving forward, it would be critical to study how other AcID motifs are implicated in binding. For example, association experiments were conducted with PTOV1B•ERM, we see a k_{on} almost identical to that of Med25 AcID (~300 $\mu M^{-1}s^{-1}$) (Fig. 2.14). PTOV1B exhibited two conformational changes, while PTOV1A exhibited one and Med25 AcID exhibited three. What is even more interesting is that PTOV1B has higher sequence deviation to Med25 as well, with an overall increase in hydrophobic residues, but a loss in negative residues (Fig. 2.2). It is therefore critical to continue to study PTOV1B, as this can help elucidate the paradigm of how sequence deviations in dynamic regions can alter binding mechanisms.

Additionally, we can expand our understanding of molecular recognition within AcID motifs by studying an AcID motif that has equal sequence deviations in both the barrel and helical regions. Med25 AcID has only been identified in metazoans, including *Arabidopsis thaliana,* a model plant organism.[26] Similar to human Med25, plant Med25 was found to associate to the Mediator complex (Fig. 2.15).[26,27] Despite the general structure being conserved, plant Med25 AcID only shares 16% sequence identity, with sequence deviations spread equally across the β-barrel and dynamic substructures. Interestingly, there appears to be a loss in one of the helices.

Experiments to determine the apparent K_d at equilibrium show that Med25's cognate binding partners bind to plant Med25 with similar affinity (Table 2.3). Further biophysical studies using the plant Med25 AcID motif coupled with the PTOV1 AcIDs can help elucidate how structural features play a role in dictating molecular recognition.

Figure 2.15: Predicted structure of *Arabidopsis thaliana's* Med25 AcID. Sharing only 16% sequence identity, plant



Med25 AcID is an excellent addition to help elucidate the binding mechanisms of coactivators. There is a loss predicted loss in secondary structure, specifically with the loss in the top helix, α2.

Table 2.3: Apparent K_d's of the four AcID motifs. Despite only sharing 16% sequence identity, plant Med25 AcID is capable of recognizing Med25 AcID's binding partners with similar affinity. Data is in triplicate and the error is reported as SDOM.

	PLANT MED25	MED25	PTOV1A	PTOV1B
ATF6A	0.16 ± 0.2 μM	0.52 ± 0.6 μM	0.20 ± 0.01μM	1.2 ± 0.2 μM
ERM	0.59 ± 0.02 μM	0.59 ± 0.02 μM	0.42 ± 0.04 μM	0.24 ± 0.03 μM
VP16	0.06 ± 0.004 μM	0.06 ± 0.004 μM	0.03 ± 0.002 μM	0.02 ± 0.005 μM

2.5. Materials and Methods

Plasmids for protein expression

Prof. Patrick Cramer generously provided the Med25 expression plasmid pET21b-Med25 (394-543)-His_6. Plasmid sequence identity was confirmed via standard Sanger sequencing methods on an Applied Biosystems 3730xl DNA Analyzer at the University of Michigan DNA Sequencing Core and analyzed using SeqMan Pro from the Lasergene DNASTAR software suite.

pET21b plamids for PTOV1A (88-235) and PTOV1B (253-398) were purchased from Genescript. Both plasmids contain a C-terminal 6x His tag. Plasmid sequence identity was confirmed via standard Sanger Sequencing as described above.

Plant Med25 AcID was designed using IDT gene designer and placed in a pmcsg7 expression vector.

Expression of Med25 AcID and plant Med25 AcID

Med25 AcID and was expressed as follows. Plasmids were transformed into chemically competent Bl21-AI cells (Novagen), plated onto LB/ampicillin/streptomycin agar, and incubated at 37˚C overnight. The next day, plates were stored at 4˚C until further use. In the evening, a single colony from the plate was selected and placed in 50 mL of Terrific Broth (TB) with 0.1 mg/mL ampicillin and 0.05 mg/mL streptomycin and incubated at 37˚C at 250 RPM. The following morning, 5 mLs from the starter culture was added to 1L TB with 0.1 mg/mL ampicillin and 0.05 mg/mL streptomycin and was grown to an OD_{600} of 0.6-0.8 at 37˚C, 250 RPM. Once OD_{600} was reached, temperature was reduced to 20˚C for a minimum of thirty minutes before induction. Cells were induced with 0.250 mM IPTG and 2% arabinose. Bacteria were shaken overnight at 20˚C, 250 RPM. The next morning, cultures were centrifuged at 7000 x g for 20 minutes at 10˚C. Cell pellets were stored at -80˚C until purification.

Expression of PTOV1A and PTOV1B

pET21b-PTOV1A(88-235)-His_6 and pET21b-PTOV1B(253-398)-His_6 were both expressed using the same method. Plasmid was transformed into chemically competent Bl21-AI cells, plated onto LB/ampicillin/streptomycin agar and incubated overnight at 37˚C. The next day, plates were stored at 4˚C until needed. In the evening, a single colony was selected and added to 50 mL of TB with 0.1 mg/mL ampicillin and 0.05 mg/mL streptomycin and incubated overnight at 37˚C, 250 RPM. The following morning, 7 mLs from the starter were added to 1 L TB with the same concentrations of antibiotics used for the overnight and grown to an OD_{600} of 1.0 at 37˚C, 250 RPM. Once OD_{600} was reached, the temperature was reduced to 30˚C for at least 30 min before induction. Cells were induced with 0.300 mM IPTG and 2% arabinose. Bacteria were shaken overnight at 30˚C, 250 RPM. The next morning, cultures were centrifuged at 7000 x g for 20 minutes at 10˚C. Cell pellets were stored at -80˚C until purification.

Purification of expressed proteins.

All proteins were purified using the same FPLC purification methods and were subjected to affinity and ion exchange chromatography. Frozen cell pellets were thawed and resuspended in lysis buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 6.8, 0.7 µL/mL β-ME, 1 protease inhibitor tablet) and lysed via sonication. Cellular debris was pelleted by centrifugation for 20 min at 9500 RPM. The clear cell lysate was then subjected to one more round of sonication to help breakup DNA, and then centrifuged again. Supernatant was filtered using 0.45 µM syringe filter (CellTreat) and loaded directly onto an AKTA Pure FPLC with a 5 mL Ni-NTA column (HisTrap HP, GE Healthcare) using a gradient of 100% Buffer A (50 mM sodium phosphate, 300 mM NaCl, and 30 mM imidazole, pH 6.8) to 100% Buffer B (50 mM sodium phosphate, 300 mM NaCl, and 500 mM imidazole, pH 6.8) over ten column volumes. AcID-containing fractions were

polled and underwent a second purification via cation exchange (HiTrap SP HP, GE Healthcare) with a gradient of 100% Buffer A (50 mM sodium phosphate, 1mM DTT, pH 6.8) to 100% Buffer B (50 mM sodium phosphate, 1 M NaCl 1mM DTT, pH 6.8) over fifteen column volumes. Fractions containing purified protein were pooled and dialyzed into storage buffer (10 mM sodium phosphate, 100 mM NaCl, 0.001% NP-40, 1% glycerol, 1mM DTT, pH6.8). Following dialysis, protein concentration was determined via UV/Vis spectroscopy on a NanoDrop at 280 nm using an extinction coefficient, ε=22,460 $M^{-1}cm^{-1}$. Protein identity was confirmed via mass spectrometry (Agilent QToF) and purity was assed via SDS-PAGE on a 4-12% bis-tris gel stained using Quick Coomassie (Anatrace). Purified protein was aliquoted, snap frozen, and stored at -80˚C.

Direct binding experiments

Direct binding experiments measured by fluorescence polarization were performed in triplicate. Final sample volume was 20 µL in a low volume, round bottom, non-stick, 384-black well plate (Corning). FITC labeled peptides were diluted in binding buffer (5 mM sodium phosphate, 100 mM NaCl, 10% glycerol, 0.001% NP-40, pH 6.8) to 40 nM. 10 µL Protein was serially diluted two-fold with assay buffer going down the plate, allowing for eight total concentration readings per experiment. The final well was a negative control, peptide only. 10 µL of diluted peptide was then added to each well for a final concentration of 20 nM. The plate was then incubated for 30 min at room temperate before fluorescence polarization was read on PHERAstar plate reader (polarized excitation at 485 nM and emission intensity measured through a parallel and perpendiculary polarized 535 nM filter.

GraphPad Prism was used to calculate the apparent K_D and indicated error (SDOM) via the equation below. A binding isotherm that accounts for ligand depletion, assuming a 1:1 binding model), was fit to the observed values as a function of protein in order to obtain the apparent

equilibrium constant, where 'a' and 'x' are the total concentrations of fluorescent peptide and protein, respectively. 'b' is the maximum observed polarization value, and 'c' is the minimum value.

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

$$Fold\ change = \frac{K_D for\ Med25\ AcID\ peptide\ interaction}{K_D for\ PTOV1\ peptide\ interaction} \times 100$$

Transient kinetic experiments

Stopped-flow kinetic assays were performed on a Kintek SF-2001 stopped-flow equipped with a 100 W xenon arc lamp in two syringe mode. The 4-DMN fluorophore was excited at 440 nm and its emission was monitored at wavelengths greater than 510 nm, using a long-pass filter (Corion). All experiments were run at 10˚C in buffer (10 mM sodium phosphate, 100 mM NaCl, 1% glycerol, 0.001% NP-40, pH 6.8), and solutions were allowed to equilibrate on the instrument for at least five minutes before the experiment was run. Concentrations reported are after rapid mixing. All kinetic traces are an average of 20-80 individual traces, and in triplicate. A series of exponential equations were fit to the transient kinetic time courses, $F(t)$ as described in the below equation, to obtain fluorescence amplitude (F_n) and the observed rate constants (k_{obs}) for each exponential phase, where $F(0)$ is the initial fluorescence intensity and t=time.

$$F(t) = F(0) + \Sigma F_n \times (1 - e^{-k_{obs,n} \times t})$$

Association experiments were performed by mixing excess protein with a constant concentration of labeled activator. Dissociation experiments were performed by precomplexing protein with labeled activator and mixing with excess of the corresponding unlabeled activator. The concentrations of the association experiments are as follows: 50 nM VP16, 125 nM ATF6α, and 250 nM ERM. The concentrations for the dissociation experiments are as follows: 50 nM

labeled VP16, 100 nM protein, 10 μM unlabeled VP16; 250 nM labeled ATF6α, 500 nM protein, 10 μM unlabeled ATF6α; 500 nM labeled ERM, 1μM protein, 50 μM unlabeled ERM. The labeled activator concentrations in both association and dissociation experiments were chosen to maximize S/N and minimize ligand depletion effects in the association experiments. The plotted k_{obs} is the rate parameter from the fast phase from the first amplitude in the association experiments and fit to a linear equation using Kaleidagraph.

$$k_{obs} = k_{on}\, x\, [Protien] + k_{off}$$

Solid-phase peptide synthesis and purification

Table 2.4 Sequence of peptides used in this study

Entry	*Peptide*	*Sequence*
1	VP16(438-490)	AcALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGI DEYGG
2	ATF6α(38-75)	AcFTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
3	ERM(38-68)	AcDLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
4	ETV4(45-76)	AcLPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
5	FITC-VP16(438-490)	Fl-βAla-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGI DEYGG
6	FITC-ATF6α(38-75)	Fl-βAla-FTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
7	FITC-ERM(38-68)	Fl-βAla-LAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
8	FITC-ETV4(45-76)	Fl-βAla-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
9	4-DMN-VP16(438-490)	4-DMN- βAla-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGI

		DEYGG
10	4-DMN-ATF6α(38-75)	4-DMN- βAla-FTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
11	4-DMN-ERM(38-68)	4-DMN- βAla-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
12	4-DMN-ETV4(45-76)	4-DMN- βAla- LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ

The peptides listed in Table 2.3 were made using standard FMOC solid-phase synthesis methods on a Liberty Blue Microwave Synthesizer (CEM). Deprotection of the FMOC groups were done in 20% piperidine (ChemImpex) in DMF that was sumplemented with 0.2 M Oxyma Pure (CEM) and irradiating under variable power to maintain a temperature of 90˚C for one minute. For coupling reactions, 5 eq of amino acids relative to resin (CEM, ChemImpex, and NovaBiochem) were added to diisopropylcarbodiimide (7 eq, ChemImpex) and Oxyma Pure (5 eq) in DMF and irradiated at variable power to maintain a temperature of 90˚C for 4 minutes. Between all couplings and deprotections, the resin was rinsed four times with excess DMF.

Following synthesis, peptides 1-4 were deprotected one final time and acetylated at the amino terminus using a cocktail of acetic anhydride and triethylamine (Fisher Scientific) in dichloromethane. Peptides 5-8 were deprotected one final time following synthesis, and then treated with fluorescein isothiocyanate (FITC, ThermoFisher) in the presence of N,N-diisopropylethylamine (Sigma Aldrich). The remaining peptides were deprotected and then subjected to one final coupling N-terminally to 4-DMN linked to a β-alanine prior to cleavage. All peptides were cleaved in 95% trifluoroacetic acid (Sigma Aldrich), 2.5% H_20 and 2.5% triisopropylsilane (Sigma Aldrich) for four hours, then filtered. The remaining solution was concentrated under nitrogen and precipitated in cold diethyl ether.

The precipitated peptide was dissolved in 50/50 0.1% TFA in water and acetonitrile with minimal ammonium hydroxide to help with solubility, syringe filtered using a 0.45µn, and purified via HPLC purification. The peptide was purified using reversed phase HPLC on an Agilent 1260 Series instrument with a 250 x 10 mm Luna Omega 5 µm PS C18 column (Phenomenex) using a gradient elution of acetonitrile in 0.1% TFA in water. Purified peptide were tested for purity using analytical HPLC, and identity was determined via mass spectrometry under negative ion mode (Agilent ToF). Once determined to be pure, peptides were dissolved in minimal DMSO and quantified via UV/Vis spectroscopy using either Tyr absorbance at 280 for acetylated peptides (ε_{280}=1,280$M^{-1}cm^{-1}$), FITC absorbance at 495 nm (ε_{495}=72,000 $M^{-1}cm^{-1}$), or 4-DMN absorbance at 440 nm (ε_{440}=10,800 $M^{-1}cm^{-1}$).

Circular dichroism of proteins

Circular dichroism spectra of Med25, PTOV1A/B and plant Med25 AcID were obtained using a J-715 spectropolarimeter (Jasco Inc) using a 1 mM quartz cuvette. Protein was dialyzed into CD buffer (5 mM sodium phosphate, 100 mM NaF, pH 6.8) to remove any trace of NaCl. Final protein concentrations of 25 µM were used. A background scan was performed using buffer only. Data was collected from 260-180 nm in 1 nm increments at a scanning speed of 100 nm/min. Background subtracted data was converted to mean residue ellipticity using the equation below. Ψ is the CD signal in degrees, n is the number of amides, l is the pathlength in centimeters, and c is the concentration in decimoles/cm^3. Each spectra reported is the average of eight scans.

$$[\Theta] = \Psi/(1000 \times \mathrm{n} \times \mathrm{l} \times \mathrm{c})$$

Circular dichroism for observed thermal melts

Following the collection of CD spectra, CD-observed thermal melts were conducted using the ‘variable temperature’ module (see above for instrument, buffer, concentrations, etc. Protein

was heated from 20-100˚C at 1˚C/min, and monitoring the molar ellipticity at 208 nm and 222 nm. Data was collected at every degree point. The molar ellipticity values were then converted to 'fraction unfolded' and the Tm was determined by fitting the data to Prism's 'log(inhibitor) vs response – variable slope' equation.

2.6 References

1. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
2. Ma, J. Transcriptional activators and activation mechanisms. *Protein Cell* **2**, 879–888 (2011).
3. Mir, M. *et al.* Dynamic multifactor hubs interact transiently with sites of active transcription in Drosophila embryos. (2018) doi:10.1101/377812.
4. Krasnov, A. N., Mazina, M. Y., Nikolenko, J. V. & Vorobyeva, N. E. On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell Biosci* **6**, 15 (2016).
5. Spiegelman, B. M. & Heinrich, R. Biological Control through Regulated Transcriptional Coactivators. *Cell* **119**, 157–167 (2004).
6. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat Chem Biol* **11**, 891–894 (2015).
7. 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. *PNAS* **110**, E3506–E3515 (2013).
8. 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 to Cells* **7**, 49–58 (2002).
9. Sela, D. *et al.* Endoplasmic Reticulum Stress-responsive Transcription Factor ATF6α Directs Recruitment of the Mediator of RNA Polymerase II Transcription and Multiple Histone Acetyltransferase Complexes♦. *J Biol Chem* **287**, 23035–23045 (2012).
10. 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).

11. Henderson, A. R. *et al.* Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *PNAS* 201806202 (2018) doi:10.1073/pnas.1806202115.
12. Henley, M. J. *et al.* Unexpected specificity within dynamic transcriptional protein–protein complexes. *PNAS* **117**, 27346–27353 (2020).
13. Vo, N. & Goodman, R. H. CREB-binding Protein and p300 in Transcriptional Regulation. *J. Biol. Chem.* **276**, 13505–13508 (2001).
14. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature Structural and Molecular Biology* **18**, nsmb.1997 (2011).
15. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nature Structural and Molecular Biology* **18**, nsmb.1999 (2011).
16. Lodge, J. M., Majmudar, C. Y., Clayton, J. & Mapp, A. K. Covalent Chemical Cochaperones of the p300/CBP GACKIX Domain. *ChemBioChem* **19**, 1907–1912 (2018).
17. Wang, N., Lodge, J. M., Fierke, C. A. & Mapp, A. K. Dissecting allosteric effects of activator–coactivator complexes using a covalent small molecule ligand. *Proc Natl Acad Sci U S A* **111**, 12061–12066 (2014).
18. Youn, H.-S., Park, U.-H., Kim, E.-J. & Um, S.-J. PTOV1 antagonizes MED25 in RAR transcriptional activation. *Biochemical and Biophysical Research Communications* **404**, 239–244 (2011).
19. Cánovas, V., Lleonart, M., Morote, J. & Paciucci, R. The role of prostate tumor overexpressed 1 in cancer progression. *Oncotarget* **8**, 12451–12471 (2016).
20. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725 (2010).

21. 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).
22. 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).
23. 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).
24. 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).
25. Odoux, A. *et al.* Experimental and molecular dynamics studies showed that CBP KIX mutation affects the stability of CBP:c-Myb complex. *Comput Biol Chem* **62**, 47–59 (2016).
26. Bäckström, S., Elfving, N., Nilsson, R., Wingsle, G. & Björklund, S. Purification of a Plant Mediator from Arabidopsis thaliana Identifies PFT1 as the Med25 Subunit. *Molecular Cell* **26**, 717–729 (2007).
27. Elfving, N. *et al.* The Arabidopsis thaliana Med25 mediator subunit integrates environmental cues to control plant development. *Proc Natl Acad Sci U S A* **108**, 8245–8250 (2011).
28. Aguilar, X. *et al.* Interaction Studies of the Human and Arabidopsis thaliana Med25-ACID Proteins with the Herpes Simplex Virus VP16- and Plant-Specific Dreb2a Transcription Factors. *PLoS One* **9**, (2014).

Chapter Three

Dynamic Substructures in Homologous Coactivators are 'Hotspots' for Dictating Allosteric Communication

3.1 Abstract

The activation of gene transcription in eukaryotic organisms is regulated by the formation of protein-protein interactions. These interactions are between DNA-bound transcription factors and coactivators. These coactivators act as central hubs, binding to a multitude of different binding partners, and their activator binding domains (ABDs) are the molecular recognition units. ABDs contain dynamic substructures, loops and helices, that offer a high degree of malleability. These dynamic substructures allow for allosteric communication within the ABDs. In this chapter, we use coactivator paralogs, termed AcID, that deviate in the sequences of their dynamic substructures to determine how these changes alter allostery. Using a biophysical approach, we demonstrate that the two binding faces of AcID are in allosteric communications. Moreover, nature of allostery is different among the paralogs. We use this information to test how an identified allosteric modulator alters binding amongst the AcID motifs, demonstrating that changes in the dynamic regions alter binding. Lastly, using transient stopped-flow kinetics, we show how a family of transcriptional activators, containing minor deviations in sequence, can induce different conformations amongst the AcID paralogs. Taken together, the work in this chapter demonstrates the importance of dynamic substructures in coactivators, with even minor deviations in sequence resulting in changes to allostery and binding modes.

3.2 Introduction

Protein-protein interactions (PPIs) between activators and coactivators are critical for gene regulation.[1,2] Despite their critical role, the mechanism by which binding occurs remains poorly understood. Early biochemical and biophysical studies suggested that molecular recognition was largely nonspecific, occurring through hydrophobic and electrostatic interactions.[3] However, as outlined in Chapter two, recent work in the Mapp lab has demonstrated that this nonspecific view of recognition is incomplete, and that specific interactions between dynamic substructures in coactivators and activators underpins complex formation.[4,5] Another role of the dynamic substructures is in allosteric communication between binding sites of coactivators.[4]

Often, ABDs are comprised of helices connected by dynamic loops.[11] The ABD of the Mediator complex subunit Med25, termed Activator Interaction Domain (AcID) is structurally unique in that it contains a seven stranded β-barrel, flanked by helix and loop domains. While it was originally proposed that the barrel region is critical for molecular recognition, recent work in the Mapp lab indicates that the dynamic loops and helices undergo significant conformational changes upon binding (Fig. 3.1) Further, these dynamic regions can help regulate cooperativity and selectivity in binding.[6–9]

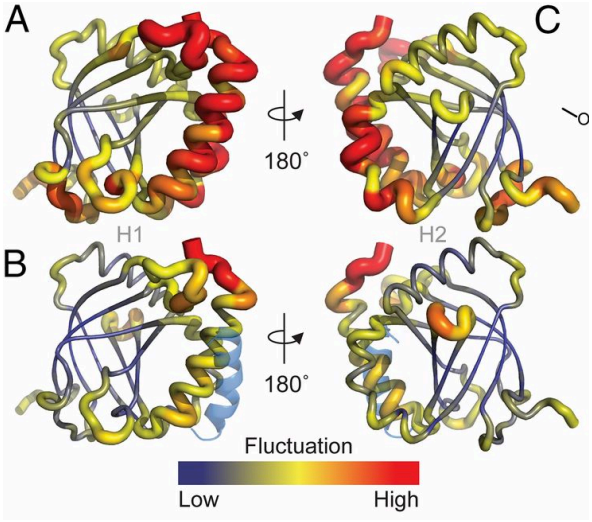


Figure 3.1: Putty figure of Med25 AcID illustrates dynamic substructures. Above is depiction showing the emerging model for AcID-activator complex formation. Using the NMR coordinates for Med25 AcID (PDB ID code 2xnf), the initial structure of Med25 was built in CHARMM using the Multiscale Modeling Tools for Structural Biology (shown in A). B) VP16 $(438\text{-}454)_{G450C}$, a peptide tethered to Med25 AcID via C506, was constructed in CHARMM as a helical peptide. This was then patched in CHARMM to Med25 through the formation of a disulfide bond at C506 (shown as a transparent blue helix). Coloring correlates to the degree of dynamical behavior, with the highest degree observed in the loops and helices.[4] Data and figure were generated by Amanda Peiffer.

In the previous chapter, I demonstrated that the sequence differences in the dynamic regions of AcID paralogs correlated with changes in affinity for activators and significant differences in conformational dynamics of the complexes. Taken together, these data support a role in for the dynamic substructures in molecular recognition. Here we use the AcID paralogs to determine how sequence difference in the dynamic regions will alter allosteric communication (Fig. 3.2). We establish that allosteric communication is conserved in the AcID motifs. Using this information, we look at how changes in the dynamic substructures alter allosteric modulator binding. Additionally, using a biophysical approach, we examine how slight changes in TAD sequence can result in differential changes in coactivator conformations. Taken together, these data demonstrate that sequence changes in the dynamic regions are highly implicated in allosteric communication.

Figure 3.2: Sequence alignment of the AcID paralogs. Shown in yellow are the sequences differences. Boxed are the dynamic substructure regions. Sequence alignments of the AcID homologs reveals that PTOV1A shares 81% sequence identity to Med25 AcID, and PTOV1B shares 71% identity. Specifically, while there are some deviations, the β-barrel remains largely conserved. In contrast, residue deviations, such as charge flips and loss of charges, are found in the dynamic loop and helical regions.

3.3 Results and Discussion

3.3.1 PTOV1A's binding faces are in allosteric communication

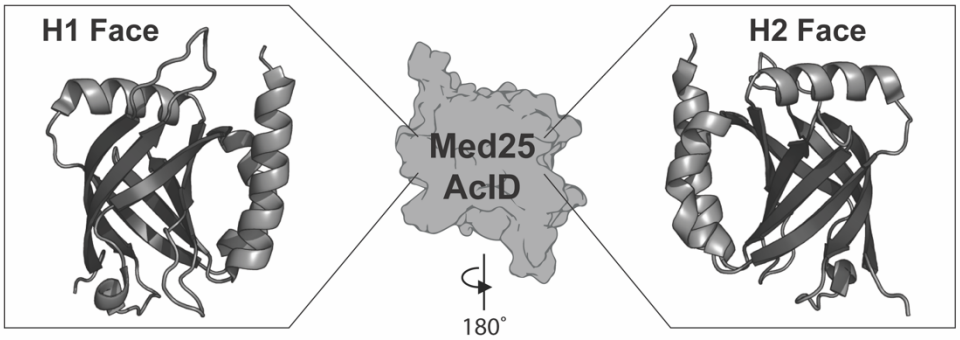


Figure 3.3: The two binding faces of AcID. AcID contains two binding faces termed H1 and H2, named for the two TAD domains of VP16.[10] For reference, ETV5/ERM binds to the H1 face, and ATF6α binds to the H2 face.[4,11]

Through NMR studies and mutational analysis, it has been demonstrated that Med25 AcID contains two distinct binding faces, termed H1 and H2 (Fig. 3.3).[4,12,13] This raised the question if these two faces were in allosteric communication. By measuring the dissociation rate constants of Med25 AcID with various combinations of H1 and H2 binding face partners, Dr. Matthew Henley was able to demonstrate that the two faces of Med25 AcID are indeed in allosteric communication. Fluorescently labeled probes that would interact with the H2 binding face, ATF6α (38-75) and VP16 (467-488), were used to asses changes in the dissociation rate constant upon ternary complex

formation when the H1 face became occupied. When ERM was bound to the H1 site, VP16 had a 20% reduction in k_{off}, while ATF6α had a 25% reduction (Fig. 3.5).[4]

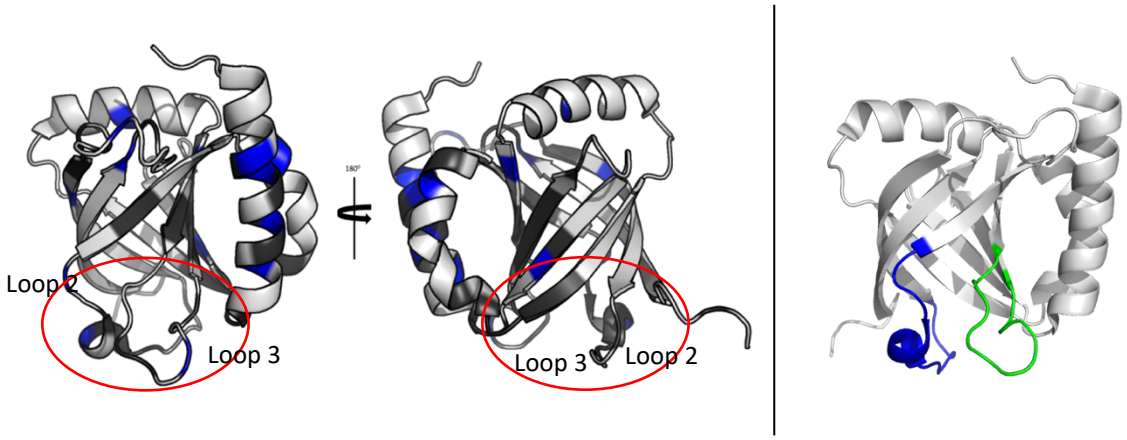


Figure 3.4: Chemical shift perturbations induced by VP16. Mapped onto Med25 AcID in (PDB 2xnf) are the chemical shift perturbations induced by VP16 are shown in blue. In black are the residues that displayed significant line broadening or were unassignable due to large shifts. As depicted, many of these perturbations are found on dynamic substructure Experiments and assignments were completed by Dr. Andrew Henderson and Dr. Brian Linhares.[4] Circled in red are the loop regions that undergo significant perturbations upon VP16 binding. Another view of the loops are shown on the right. In blue is loop 2 and in green is loop 3.

We next wanted to determine if this internal allosteric network extended to PTOV1A. The bottom loop regions, loops 2 and 3 show significant peturbations in Med25 AcID (Fig. 3.4). In PTOV1A, these regions in have quite a bit of sequence deviation. Specifically, the charge in loop 2 has an overall negative charge compared to the same loop in Med25 AcID, which has an overall positive. (Fig. 3.4). These bottom loops have been shown to be critical for allostery in Med25 AcID, and even the use of a covalent chaperone tethered to loop 3 has successfully recapitulatd allosteric changes.[4] We therefore hypothesize that these sequence deviations will alter the allosteric network in PTOV1A, directly measurable by monitoring the k_{off}.[14]

To test this hypothesis, a similar experiment was set up with PTOV1A. PTOV1A's binding faces were 100% occupied by precomplexing with DMN-labeled ATF6α, an H2 binder, and unlabeled ERM, an H1 binder, and unlabeled ATF6α served as the competitor. We monitored changes to the H2 face through changes in k_{off}. When the H1 site is occupied, we observe an 18% increase in k_{off} of the H2 face, an opposite trend to what we observe for Med25 AcID in the same

experiment (Fig. 3.5). This suggests that the lifetime of the ternary complex for PTOV1A is shorter than the lifetime of the binary complex with ATF6α, an unusual trend for coactivators. These data support that PTOV1A binding faces are in allosteric communication. Given the overall lack of structure, as described in Chapter two, this reduction in binding is likely facilitated by the increase in dynamic loops of PTOV1A. By demonstrating that PTOV1A's two binding faces are in allosteric communication, we can theoretically modulate binding by targeting one of these dynamic regions.

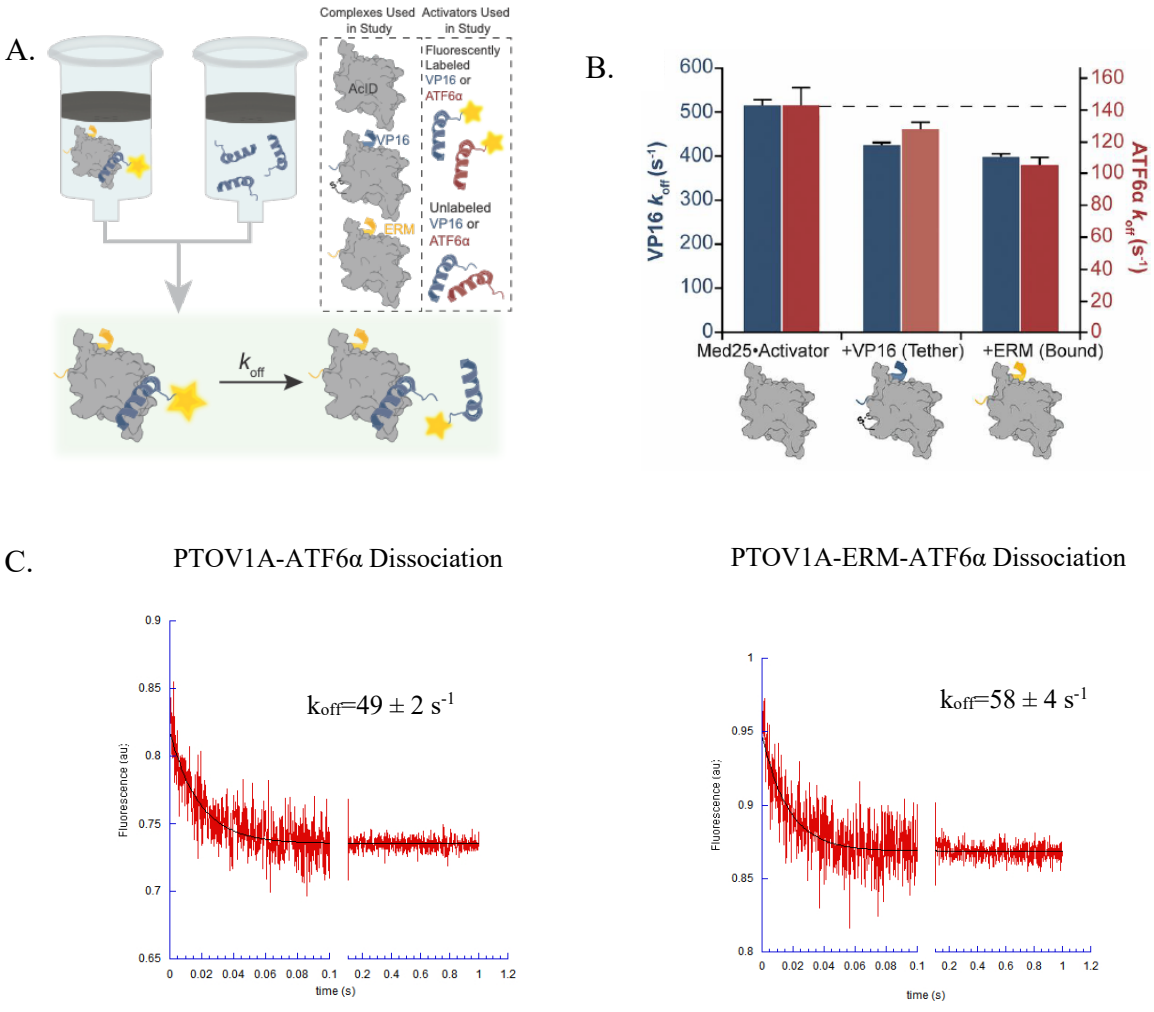


Figure 3.5: Kinetic experiments show allosteric within AcID. A) Schematic of the experiment. B) Comparison of k_{off} for VP16 (H2) (blue bars) for Med25 AcID, Med25 AcID with VP16 $(H1)_{G450C}$ covalently tethered, and Med25 AcID with ERM prebound; the red bars summarize data from an analogous experiments with ATF6α. C) Dissociation experiment of (left), and the allosteric complex (right). 250 nM protein was prebound to 125 nM DMN labeled ATF6α

and 7.5 µM unlabeled ERM. The complex was then rapidly mixed with unlabeled ATF6α. Both traces were fit to single exponentials. k_{off} values are the average from three traces and the reported error is the SDOM.[15]

3.3.2 A natural product allosterically inhibits AcID interactions

The xanthone natural product garcinolic acid (GA) was identified in natural product screen conducted by Dr. Julie Garlick as an inhibitor of Med25 AcID. With an IC_{50} of 9.1 µM against ERM, GA with modest potency (Fig. 3.6). In order to determine where GA was binding to binding to Med25 AcID, HSQC NMR experiments led by Dr. Matthew Beyersdorf were conducted (Fig. 3.7). Interestingly, GA was shown to bind to the H2 face of Med25 AcID, suggesting that GA is an allosteric inhibitor of the Med25 AcID•ERM interaction.

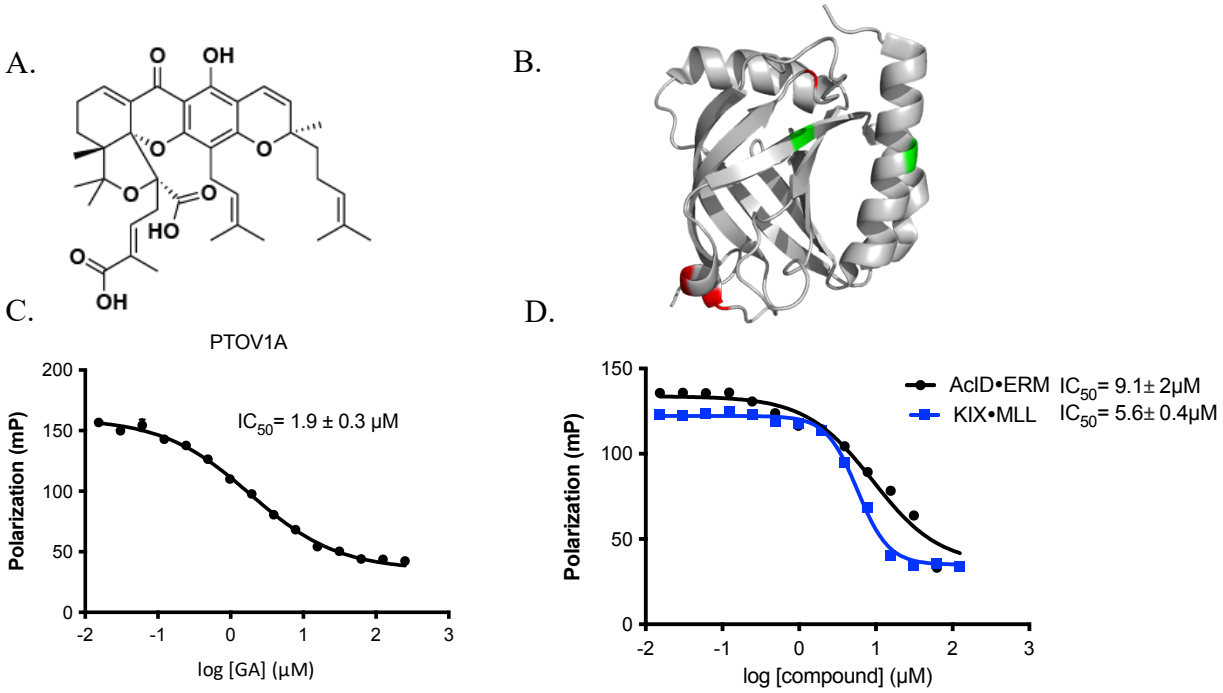


Figure 3.6: Garcinolic acid is a natural product that inhibits TAD binding to coactivators A) Structure of garcinolic acid, a xanthone natural product identified in a library screen. B) FP based competition assays of GA with CBP GACKIX•MLL (see Appendix 1 for more information) and Med25•ERM. Protein was incubated with 20 nM final FITC-labeled peptide at 50% occupancy. The max concentration of GA was 200 µM. C) FP based competition assays of GA with PTOV1A•ERM. Protein was incubated with 20 nM final FITC-labeled peptide at 50% occupancy. The max concentration of GA was 200 µM. D) he H1 binding face of Med25 AcID. In green are the key perturbed residues that are conserved in PTOV1A, and in red are the differences, such as K411R, K440E and , which is a Cys in PTOV1A [11,12]

Having shown that PTOV1A's two binding faces are in allosteric communication, we wanted to determine how an allosteric modulator such as GA would affect TAD binding. We hypothesized that due to the changes in the allosteric network, this would be reflected by an allosteric modulator. Using a fluorescence polarization-based competition assay, we measured the IC_{50} value starting with a maximum concentration of 200 μM GA, we titrated in decreasing concentrations of GA. PTOV1A was precomplexed with FITC-labeled ERM at 50% occupancy. We show that GA has an IC_{50} of 1.9 μM against ERM, exhibiting a six-fold tighter inhibition than Med25 (Fig. 3.6). This is likely due to a loss in surrounding electrostatics on PTOV1A in the proposed ERM binding site coupled with the changes to the bottom loops in PTOV1A. Taken together, these data support overarching hypothesis that changes in the sequences of the dynamic substructures alters binding of an allosteric modulator.

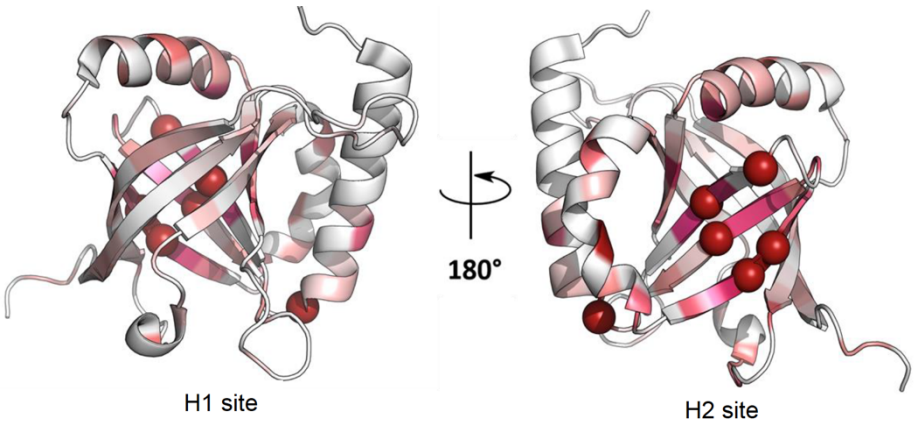


Figure 3.7: NMR experiments show the binding site of GA. Using ^{1}H-^{15}N HSQC NMR experiments, GA caused the biggest perturbations on the H2 face. Increasing shades of red indicate more chemical shift perturbations, with the red balls depicting residues that underwent complete peak broadening. Data and figures made by Dr. Matthew Beyersdorf.

3.3.3 AcID motifs can discern even small changes in sequence identity

```
ETV1/ER81 (38-69)   DLAHDSEELFQDLSQLQETWLAEAQVPDNDEQ
ETV4/PEA3 (45-76)   LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
ETV5/ERM  (38-68)   DLAHDSEELFQDLSQLQEAWLAEAQVPD-DEQ
                    ∿∿∿∿∿∿∿∿∿∿∿∿∿
```

Figure 3.8: Sequence alignment of the ETV/PEA3 TADs. The middle portion of the TAD sequence has been predicted to have an increased propensity to take on a helical structure, as depicted. The four N-terminal residues are more likely to be disordered.

Recent work in the Mapp lab by Dr. Matthew Henley demonstrated surprising specificity of TAD recognition using the ETV/PEA3 family of ETS transcriptional activators, including ETV1, ETV4 and ETV5/ERM (Fig. 3.8).[11,16,17] The ETV•Med25 AcID interactions appeared to behave as a prototypical TAD•ABD interaction, occurring over a shallow binding surface and drive by hydrophobic and electrostatic interactions. For all three ETV/PEA3 TADs, there is a rapid binding step followed by two conformational changes steps, with each step occurring with similar rates of exchange. However, upon further investigation using transient stopped-flow kinetics and structural approaches, it was revealed that the slight sequence deviations among the ETV/PEA3 TADs resulted in significant changes within the conformational ensembles of Med25•TAD interactions. Additionally, NMR analysis demonstrated that while the ETV TADs bind to the same face of Med25 AcID, they do so in distinct orientations, making contacts with dynamic substructures. Additionally, using ^{15}N labeled ETV1 and ETV4 TADs, N-terminal Leu residues in proposed disordered regions underwent chemical shifts upon binding to Med25 AcID with L39 of ETV1 undergoing a more dramatic shift than L48 of ETV4, showing different engagement of the ETV TADs (Fig 3.9).[5]

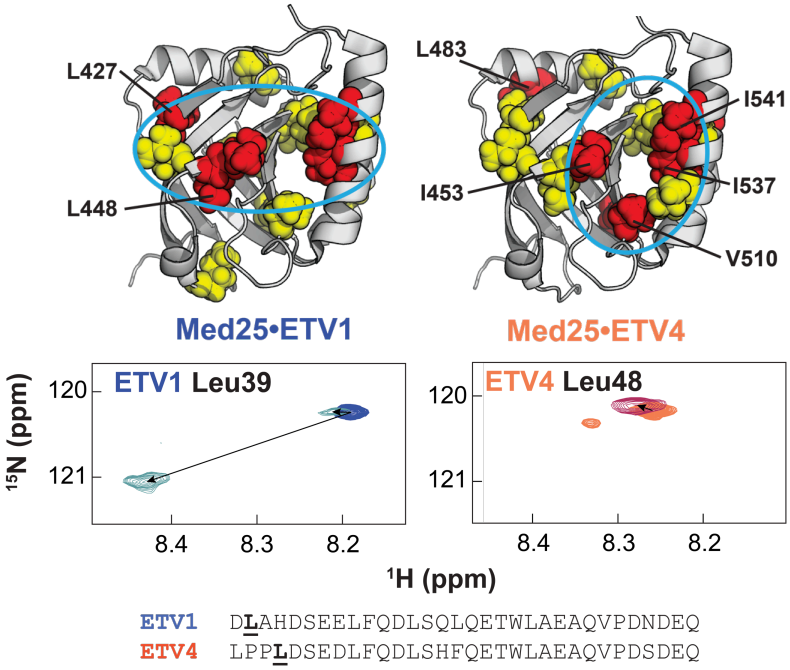


Figure 3.9: The dynamic regions of ETV/PEA3 TADs bind to Med25 AcID's dynamic regions in distinct orientations. Top) Chemical shift perturbations induced by ETV/PEA3 binding plotted onto the structure of Med25 AcID (PDB 2xnf). Circled in cyan is the general orientation of the ETV1/ETV5 and ETV4. Bottom) Chemical shift perturbations of ^{15}N ETV1/ETV5 (left) and ETV4 (right) TADs in the absence (blue and orange respectively) and presence (light blue and maroon respectively) of unlabeled Med25 AcID. Taken together, these data demonstrate that not only do the two TADs engage in unique orientations, but that the implicated dynamic regions of the TADs themselves are directly engaging with the dynamic substructures of Med25 AcID. *Data and figure generated by Dr. Matthew Henley.

We wanted to test how slight changes in TAD sequences alter the different AcID motifs conformational ensembles upon binding due to the changes in sequence in their dynamic regions. As described in the previous chapter, ETV5 is an H1 face binder, with key perturbations occurring on α3 and β5. Mutations to key residues at these sites attenuated ETV5 binding.[11] In Chapter two, we showed that while binding affinities and mechanisms are not altered, there is a change in the conformational ensembles observed via transient stopped-flow kinetics. Specifically, we showed that there was a loss in conformation transitions, with only one observed conformer versus the three observed for Med25 AcID. Given the high sequence identity, we hypothesize that a similar

trend will be observed with PTOV1A binding to the other ETV/PEA3 TADs; that is, there will be a reduction in the number of conformers observed, but with similar affinities.

We tested PTOV1A against the ETV/PEA3 TADS using transient stopped-flow kinetics. As ETV1 behaved similarly to ETV5, we only tested ETV4 and ETV5. Similar K_d's were observed for each system (Table 3.1). Interestingly, PTOV1A when complexed with ERM had only one phase, where with ETV4, a second phase was observed. For Med25 AcID, both ERM and ETV4 observed 3 phases (Figure 3.10). Moreover, each complex demonstrated a unique distribution of conformation populations, showing clear differences in the engagement modes of ETV/PEA3 family members (Table 3.1) For ETV4, PTOV1A shows two distinct conformations, with 90 % of the population being observed in the second conformation. Med25, on the other hand, has > 60 % of the population in the third conformation (Fig. 3.10). Taken together, these data suggest that when comparing the AcID motifs, the sequence changes in the dynamic regions result alters how the ABDs respond to TAD binding in a similar trend observed in in Chapter two.

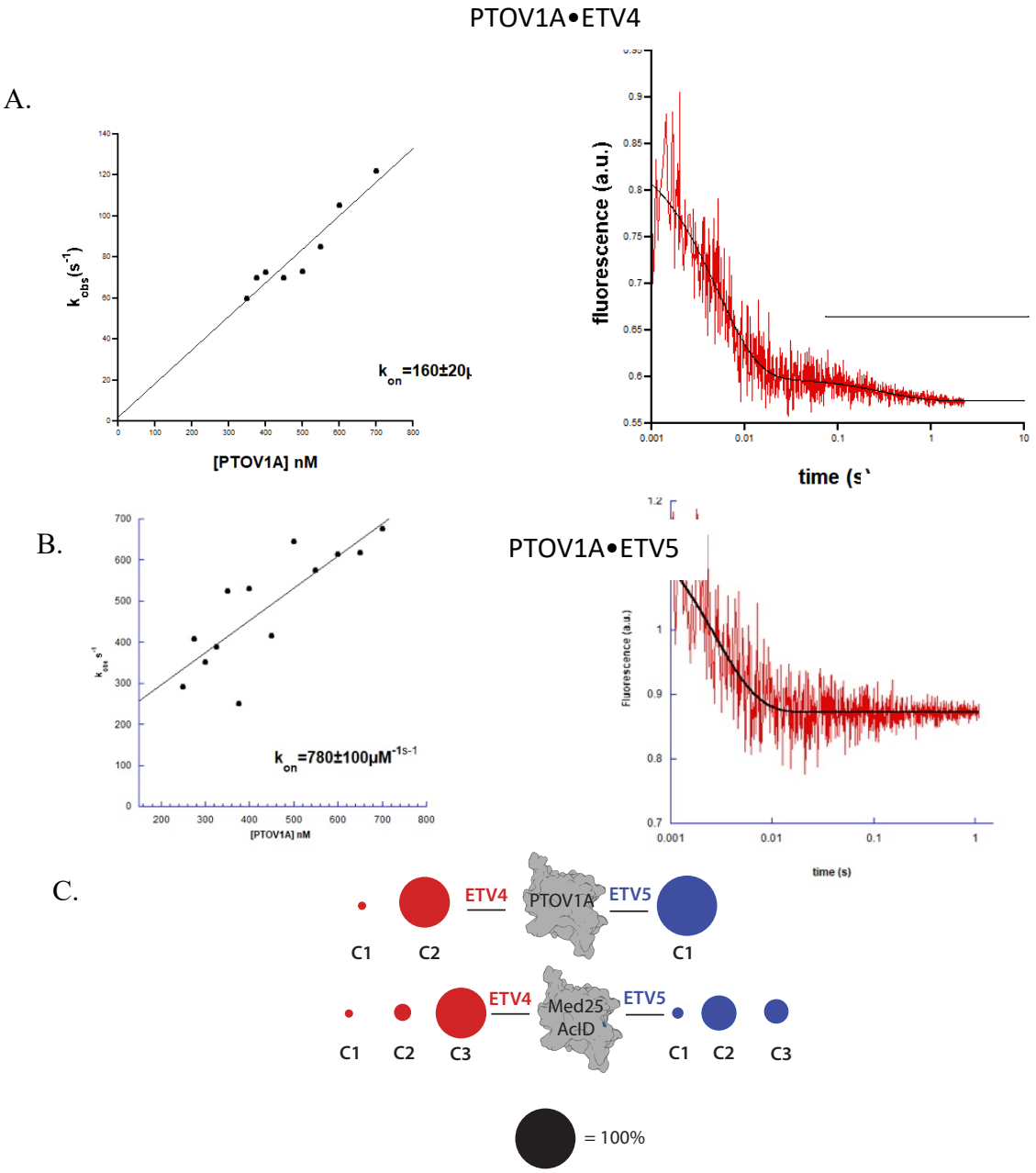


Figure 3.10: Experiments used to define ETV/PEA3 binding parameters. A) Plotted k_{obs} of the fast phase (left) and dissociation trace (right) of PTOV1A•ETV4. The association trace (not pictured) was fit to a double exponential. For the dissociation trace, 0.5 µM DMN-labeled ETV4 was complex with 2µM PTOV1A, then rapidly mixed with 25 µM unlabeled ETV4. The trace was fit to a double exponential. For reference, Med25 AcID was fit to a triple exponential for both association and dissociation. B) Plotted k_{obs} of the fast phase (left) and dissociation trace (right) of PTOV1A•ETV5. The association trace (not pictured) was fit to a single exponential. For the dissociation trace, 0.5 µM DMN-labeled ETV5 was complex with 2µM PTOV1A, then rapidly mixed with 25 µM unlabeled ETV5. The trace was fit to a single exponential. For reference, Med25 AcID was fit to a triple exponential for both association and dissociation. C)Equilibruim populations of the AcID•ETV conformations calculated from microscopic equilibrium constants.

Table 3.1: Table of apparent K_d and population distributions for against ETV4 and ETV5

Complex	$K_{d\ (\mu M)}$	C1(%)	C2 (%)	C3 (%)
Med25•ETV4	0.7 ± 0.2	9 ± 1	28 ± 3	63 ± 4
PTOV1A•ETV4	0.25 ± 0.2	10 ± 2	90 ± 8	0
Med25 ETV5	0.9 ± 0.1	16 ± 3	53 ± 9	31 ± 9
PTOV1A•ETV5	0.6 ± 0.2	100	0	0

3.3.4 Dynamic regions of TADs and ABDs directly engage with one another

An emerging hypothesis from the observation that members of the ETV/PEA3 family of activators bind to Med25 in unique conformations is that individual ETV•Med25 interactions may be differentially affected as a result of conformational shifts in Med25.[5] Dr. Amanda Peiffer in the Mapp lab showed using molecular dynamic (MD) simulations that the modulation of a loop on the Med25 AcID H2 face, shown to be critical for allosteric regulation, induces changes in selectivity of ETV/PEA3 transcription factors. Specifically, the MD simulations demonstrated a conformational shift in the loop region, Figure 3.11, that altered binding affinity. Mutational analysis demonstrated that binding of ETV transcription factors to Med25 in a conformationally-specific manner. Specifically, the mutant M523E weakened ETV1 and ETV5 binding, while ETV4 binding was enhanced by approximately 3-fold. These data suggest that this mutation results in preferential binding of Med25 to ETV4 over ETV1 and ETV5. Additionally, these data demonstrate the direct engagement of TADs to the dynamic regions of ABDs.

To study how sequence deviations and changes in allosteric modulation observed in PTOV1A will affect ETV binding to this same loop regions, as well identity direct interactions of dynamic regions, mutational analysis was also conducted for PTOV1A. The residues identified in the MD simulations are conserved in PTOV1A. The mutants S517N and M523E were generated

and tested against FITC-labeled ETV4 and ETV5. S517N, located directly in the loop region, showed no significant changes in binding affinity. However, M523E weakened binding affinity for both ERM and ETV4 by 2-fold. Given that there is an overall loss in structure, this is not too surprising. These changes in binding affinity, as seen with the equivalent M523E mutant made to PTOV1A demonstrate that changes in conformations can alter binding, and that dynamic substructures play a role in binding.

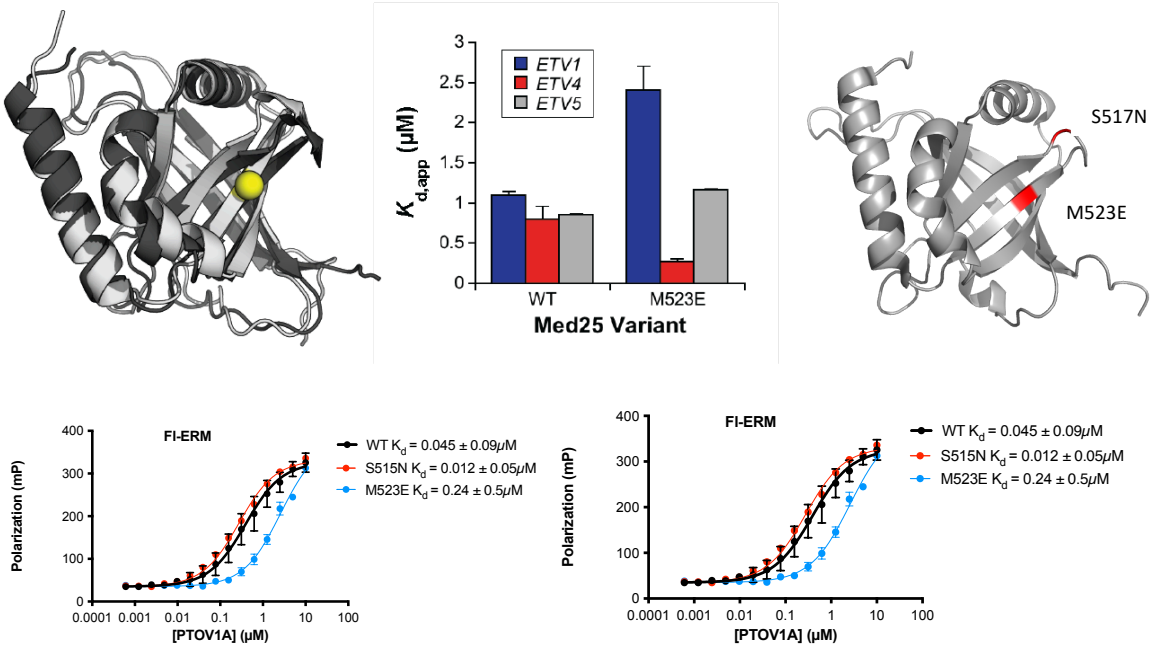


Figure 3.11: Direct engagement of dynamic regions between ABDs and TADs. Top left demonstrates the molecular dynamic simulations that show changes in conformation in the allosteric loop (shown above the yellow sphere that represents the Med25^{M523E} mutant). These data were generated by Dr. Amanda Peiffer. The top middle graph shows the changes in affinity induced by the mutant. These data were generated by Dr. Matthew Henley. The top right figure shows the location of the PTOV1A mutants generated. S517N is located in the proposed allosteric loop, and by adding a bulkier residue, the loop may pulled down. M523E is located below, with the introduction of a charged residue proposed to pull down the loop region. The bottom figures show the binding curves, with error reported as the standard deviation.

3.4 Conclusions and Assessment

These 'fuzzy interactions' have long been termed undruggable due to the evasive nature in studying by traditional means. While it was originally proposed that binding was driven through electrostatic and hydrophobic interactions, this model does not account for the intricate specificity

and selectivity required for such a highly regulated process as transcription initiation. Towards this end, recent work in the Mapp lab has demonstrated that molecular recognition is driven through conformational changes driven through dynamic regions. Upon activator binding to the ABD, there is a winnowing of conformers, allowing a unique state specific to each interaction.

In this chapter, we investigate how hotspots can be used to differentiate highly similar TAD sequences. Using two TADs from the ETV/PEA3 family of transcription factors, we tested how slight sequence deviations can lead to changes in conformation of AcID motifs. Given that these TADs are highly related, this system served as an opportunity to demonstrate mechanism specificity. Using in depth transient kinetic analysis, we demonstrated that not only is there a high degree of specificity of TAD recognition within one coactivator, but that the different AcIDs exhibit different conformational ensembles. Small changes in TAD sequence resulted in a high degree of specificity. These changes in conformations is mediated by the high degree of plasticity in the AcIDs. Moving forward, testing PTOV1B against the ETV/PEA3 TADs can provide further insight into each AcID motif responds. Additionally, HSQC NMR experiments to explore how each TAD interacts with the different AcID motifs can provide powerful insight molecular recognition of these complexes.

Additionally, we tested how modulators can be used to target hotspot residues. GA, a natural product, while not a covalent labeler, was found to show a range of different IC_{50} values, with PTOV1A showing a tighter IC_{50}. These data suggest that the changes in dynamic structures can be used as an asset, allowing for selective targeting. Testing against PTOV1B, as well as an in-depth kinetic analysis to determine how k_{off} is affected will further help exploit hotspots. Taken together, the work in this chapter demonstrates that the nonspecific model of recognition is outdated. Through this work, we begin to demonstrate that the AcID motifs utilize their differences in

dynamic substructures to recognize different binding partners. Given their overall plasticity, we can use these less conserved allosteric regions to selectively and specifically target these ABDs.

Typically, PPI interfaces are highly polar and often conserved, making it difficult to target orthosterically while achieving selectivity and specificity. Use of allosteric modulators that can bind to dynamic hot spots can induce conformational changes that can inhibit or enhance binding.[18] Allosteric modulators that target dynamic hot spots can act as chemical probes that can be used to dissect binding mechanisms. Moving forward, it is critical to continue testing allosteric modulators. One such modulator is the covalent inhibitor norsticitic acid (NA) (Fig. 3.12). NA is a depsidone containing orthophenolic aldehyde moiety that modulates Med25 AcID via a covalent mechanism via imine formation with Lys side chains. Specifically, NA covalently labels a Lys rich loop on the H2 binding face. Extensive work led by Dr. Julie Garlick and Dr. Steven Sturlis demonstrate an IC_{50} against ERM of 2.1 μM. Binding of NA thus induces an allosteric change by inducing a slight unfolding of the two dynamic loops on the H1 face, as shown through molecular dynamic simulations led by Amanda Peiffer. Work is currently underway to determine how sequence deviations of the dynamic regions amongst the AcID motifs are thus affected by NA binding.

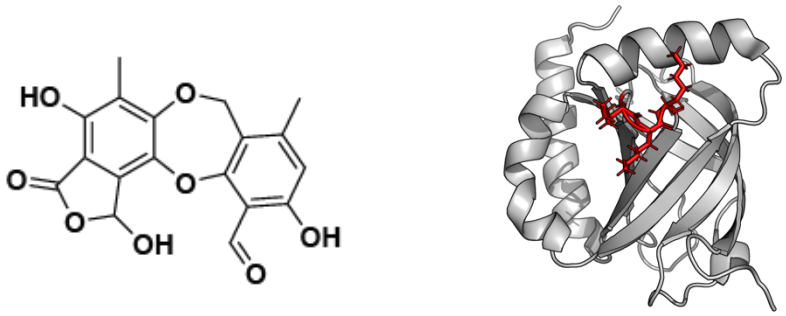


Figure 3.12: Norstictic acid is a covalent modulator of AcID motifs. On the left structure of norstictic acid, an orthophenolic aldehyde containing depsidone. On the right structure of Med25 AcID highlighting the triple Lys loop that is preferential labeled. Specifically, NA labels K519.

Transcriptional activators and coactivators have long been considered "undruggable" due to the difficulty with which they are targeted by small molecules.[19] For example, AcID is a

challenging system to target.[4,15] Given its rather featureless binding interface coupled with binding to multiple binding partners over transient, short-lived interactions, it is perhaps not surprising that there has only been one reported successful discovery of a small-molecule ligand.[20] To compound this issue, there contains a second AcID containing protein in humans, PTOV1, with high sequence conservation, although there is differentiation in the dynamic regions (Fig. 3.2).[21,22] Dysregulation of both Med25 AcID and PTOV1 result in disease states, such as prostate cancer, and it is thus critical to selectively target the AcID motifs.[21,22] Like other coactivators, the presence of dynamic loops and helices in the AcID motif have been hypothesized by our group to present better biding sites for small molecules.[4]

3.5 Materials and methods

Plasmids for protein expression

Prof. Patrick Cramer generously provided the Med25 expression plasmid pET21b-Med25 (394-543)-His_6. Plasmid sequence identity was confirmed via standard Sanger sequencing methods on an Applied Biosystems 3730xl DNA Analyzer at the University of Michigan DNA Sequencing Core and analyzed using SeqMan Pro from the Lasergene DNASTAR software suite.

pET21b plasmids for PTOV1A (88-235) and PTOV1B (253-398) were purchased from Genescript. Both plasmids contain a C-terminal 6x His tag. Plasmid sequence identity was confirmed via standard Sanger Sequencing as described above.

Quick-change site-directed mutagenesis was used to create the PTOV1A mutants. Identity was determined by Sanger Sequencing. The primers for SDM are as shown in Table 3.2

Table 3.2 Primers used for site-directed mutagenesis of PTOV1A

Plasmid	**Primer Sequence**
pET21b-PTOV1A S517N-His_6	F: CTGCTGTACAGCAACAAGAAAAAGATCTTC R: GAAGATCTTTTTCTTGTTGC GTACAGCAG
pET21b-PTOV1A M523E-His_6	F: AAGAAAAAGATCTTCGAGGGTCTGATTCCG R: CGGAATCAGACCCTCGAAGATCTTTTTCTT

Expression and purification of CBP GACKIX

CBP GACKIX was expressed and purified using previously described protocols.[23]

Expression of Med25 AcID

Med25 AcID and was expressed as follows. Plasmids were transformed into chemically competent Bl21-AI cells (Novagen), plated onto LB/ampicillin/streptomycin agar, and incubated at 37˚C overnight. The next day, plates were stored at 4˚C until further use. In the evening, a single colony from the plate was selected and placed in 50 mL of Terrific Broth (TB) with 0.1 mg/mL ampicillin and 0.05 mg/mL streptomycin and incubated at 37˚C at 250 RPM. The following morning, 5 mLs from the starter culture was added to 1L TB with 0.1 mg/mL ampicillin and 0.05

mg/mL streptomycin and was grown to an OD_{600} of 0.6-0.8 at 37˚C, 250 RPM. Once OD_{600} was reached, temperature was reduced to 20˚C for a minimum of thirty minutes before induction. Cells were induced with 0.250 mM IPTG and 2% arabinose. Bacteria were shaken overnight at 20˚C, 250 RPM. The next morning, cultures were centrifuged at 7000 x g for 20 minutes at 10˚C. Cell pellets were stored at -80˚C until purification.

Expression of PTOV1A and PTOV1B and mutants

pET21b-PTOV1A(88-235)-His_6 (including $PTOV1A^{S517N}$ and $PTOV1A^{M523E}$) and pET21b-PTOV1B(253-398)-His_6 were expressed using the same method. Plasmid was transformed into chemically competent Bl21-AI cells, plated onto LB/ampicillin/streptomycin agar and incubated overnight at 37˚C. The next day, plates were stored at 4˚C until needed. In the evening, a single colony was selected and added to 50 mL of TB with 0.1 mg/mL ampicillin and 0.05 mg/mL streptomycin and incubated overnight at 37˚C, 250 RPM. The following morning, 7 mLs from the starter were added to 1 L TB with the same concentrations of antibiotics used for the overnight and grown to an OD_{600} of 1.0 at 37˚C, 250 RPM. Once OD_{600} was reached, the temperature was reduced to 30˚C for at least 30 min before induction. Cells were induced with 0.300 mM IPTG and 2% arabinose. Bacteria were shaken overnight at 30˚C, 250 RPM. The next morning, cultures were centrifuged at 7000 x g for 20 minutes at 10˚C. Cell pellets were stored at -80˚C until purification.

Purification of expressed proteins.

All proteins were purified using the same FPLC purification methods and were subjected to affinity and ion exchange chromatography. Frozen cell pellets were thawed and resuspended in lysis buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 6.8, 0.7 µL/mL β-ME, 1 protease inhibitor tablet) and lysed via sonication. Cellular debris was pelleted by centrifugation

for 20 min at 9500 RPM. The clear cell lysate was then subjected to one more round of sonication to help breakup DNA, and then centrifuged again. Supernatant was filtered using 0.45 μM syringe filter (CellTreat) and loaded directly onto an AKTA Pure FPLC with a 5 mL Ni-NTA column (HisTrap HP, GE Healthcare) using a gradient of 100% Buffer A (50 mM sodium phosphate, 300 mM NaCl, and 30 mM imidazole, pH 6.8) to 100% Buffer B (50 mM sodium phosphate, 300 mM NaCl, and 500 mM imidazole, pH 6.8) over ten column volumes. AcID-containing fractions were polled and underwent a second purification via cation exchange (HiTrap SP HP, GE Healthcare) with a gradient of 100% Buffer A (50 mM sodium phosphate, 1mM DTT, pH 6.8) to 100% Buffer B (50 mM sodium phosphate, 1 M NaCl 1mM DTT, pH 6.8) over fifteen column volumes. Fractions containing purified protein were pooled and dialyzed into storage buffer (10 mM sodium phosphate, 100 mM NaCl, 0.001% NP-40, 1% glycerol, 1mM DTT, pH6.8). Following dialysis, protein concentration was determined via UV/Vis spectroscopy on a NanoDrop at 280 nm using an extinction coefficient, ε=22,460 $M^{-1}cm^{-1}$. Protein identity was confirmed via mass spectrometry (Agilent QToF) and purity was assed via SDS-PAGE on a 4-12% bis-tris gel stained using Quick Coomassie (Anatrace). Purified protein was aliquoted, snap frozen, and stored at -80˚C.

Expression and purification of CBP GACKIX

CBP GACKIX was expressed and purified using previously described protocols.[23]

Direct binding experiments

Direct binding experiments measured by fluorescence polarization were performed in triplicate. Final sample volume was 20 μL in a low volume, round bottom, non-stick, 384-black well plate (Corning). FITC labeled peptides were diluted in binding buffer (5 mM sodium phosphate, 100 mM NaCl, 10% glycerol, 0.001% NP-40, pH 6.8) to 40 nM. 10 μL Protein was serially diluted two-fold with assay buffer going down the plate, allowing for eight total

concentration readings per experiment. The final well was a negative control, peptide only. 10 μL of diluted peptide was then added to each well for a final concentration of 20 nM. The plate was then incubated for 30 min at room temperate before fluorescence polarization was read on PHERAstar plate reader (polarized excitation at 485 nM and emission intensity measured through a parallel and perpendiculary polarized 535 nM filter.

GraphPad Prism was used to calculate the apparent K_D and indicated error (SDOM) via the equation below. A binding isotherm that accounts for ligand depletion, assuming a 1:1 binding model), was fit to the observed values as a function of protein in order to obtain the apparent equilibrium constant, where 'a' and 'x' are the total concentrations of fluorescent peptide and protein, respectively. 'b' is the maximum observed polarization value, and 'c' is the minimum value.

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

Fluorescence polarization competition 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). Protein was precomplexed with fluorescent (FITC) peptide using two-fold the concentrations of protein and peptide in order to reach 50% of the tracer bound in binding buffer (5 mM sodium phosphate, 100 mM NaCl, 10% glycerol, pH 6.8). Small molecule, dissolved in DMSO and diluted in assay buffer was serially diluted down two-fold with assay buffer going down the plate. 10 μL of the pre-complexed peptide-protein complex was then added to each well for a final volume of 20 μL. The final well, row P, served as a negative control, tracer only. The plate was incubated for thirty minutes at room temperature before fluorescence polarization was measured on a PHERAstar plate reader. The experiment was conducted using polarized excitation at 485 nm and emission intensity was measured through a parallel and

perpendicularly polarized 535 nm filter. Polarization values were converted to relative fraction bound and plotted against log[inhibitor]. Inhibition curves were then fit to a non-linear regression using Prism's 'log(inhibitor) vs response – variable slope' equation from which the IC_{50} values were calculated.

$$Fold\ change = \frac{K_D for\ Med25\ AcID\ peptide\ interaction}{K_D for\ PTOV1\ peptide\ interaction} \times 100$$

Transient kinetic experiments

Stopped-flow kinetic assays were performed on a Kintek SF-2001 stopped-flow equipped with a 100 W xenon arc lamp in two syringe mode. The 4-DMN fluorophore was excited at 440 nm and its emission was monitored at wavelengths greater than 510 nm, using a long-pass filter (Corion). All experiments were run at 10˚C in buffer (10 mM sodium phosphate, 100 mM NaCl, 1% glycerol, 0.001% NP-40, pH 6.8), and solutions were allowed to equilibrate on the instrument for at least five minutes before the experiment was run. Concentrations reported are after rapid mixing. All kinetic traces are an average of 40-80 individual traces, and in triplicate. Error reported is standard deviations. A series of exponential equations were fit to the transient kinetic time courses, F(t) as described in the below equation, to obtain fluorescence amplitude (F_n) and the observed rate constants (k_{obs}) for each exponential phase, where F_∞ is the final fluorescence intensity and t=time.

$$F(t) = F_\infty + \Sigma F_n \times (1 - e^{-k_{obs,n} \times t})$$

Association experiments were performed by mixing excess protein with a constant concentration of labeled activator. Dissociation experiments were performed by precomplexing protein with labeled activator and mixing with excess of the corresponding unlabeled activator. The concentrations of the association experiments are as follows: 50 nM VP16, 125 nM ATF6α,

and 250 nM ERM. The concentrations for the dissociation experiments are as follows: 50 nM labeled VP16, 100 nM protein, 10 μM unlabeled VP16; 250 nM labeled ATF6α, 500 nM protein, 10 μM unlabeled ATF6α; 500 nM labeled ERM, 1μM protein, 50 μM unlabeled ERM. The labeled activator concentrations in both association and dissociation experiments were chosen to maximize S/N and minimize ligand depletion effects in the association experiments.

Calculation of Rate Constants

The individual $k_{obs,n}$ values were plotted as a function of concentration and fit to a square hyperbola (equation below) in order to determine the maximal observed rate constant ($k_{obs,n,max}$) and the half maximal concentration ($K_{1/2,n}$). The value of $k_{obs,n,min}$ was included for the fit, but not for calculations as it is defined by the corresponding $k_{obs,n,off.}$ The microscopic rate constants were calculated using a combined rapid equilibrium and steady-state approximation.

$$k_{obs,n} = \frac{k_{obs,mas} \times [protein]}{[protein] + K_{1/2,n}} + k_{obs,n.min}$$

A rapid equilibrium and steady-state approach was used to determine rate parameters. For simplicity sake, the method for calculating the microscopic rate constants is split into two sections: the initial binding step in which C1 transitions into C2, and the transition of C2 to C3. The first two steps, we can determine the maximal observed rate constant of the first conformational change ($k_{obs,2,max}$) using the following equation. In all cases, $k_{obs,1,off}$ is well defined for PTOV1A.

$$k_{obs,2,max} = k_{F,1} + k_{R,1}$$

Using steady-state approximation, the corresponding observed rate constant for the dissociation ($k_{obs,2,off}$) is as follows:

$$k_{obs,2,off} = k_{R,1} \times \frac{k_{off}}{k_{off} + k_{F,1}}$$

K_{off} can be calculated using the observed rate constants from dissociation experiments ($k_{obs,n,off}$) as follows:

$$k_{off} = k_{obs,1,off} + k_{obs,2,off} - k_{obs,2,max}$$

Since we only observed two conformations, we do not need to consider the transition from C2 to C3. Relative populations C1 and C2 (for PTOV1A) at equilibrium were then determined by the conformational equilibrium constants ($K,c,n = k_{F,n} / k_{R,n}$), which are ratios between the conformational states, by definition. Below are the equations used to calculate the populations of each states. C1 is serving as the references state. Only the relative conformational populations of the bound state were considered, thus all values are concentration dependent.

In all cases, $k_{obs,1,off}$ is well defined for PTOV1A.

$$Population\ C1 = \frac{1}{1 + K_{C,1} + \left(K_{C,1} \times K_{C,2}\right)} \times 100\%$$

$$Population\ C2 = \frac{K_{C,1}}{1 + K_{C,1} + \left(K_{C,1} \times K_{C,2}\right)} \times 100\%$$

Solid-phase peptide synthesis and purification

Table 3.3 Sequence of peptides used in this study

Entry	*Peptide*	*Sequence*
1	VP16(438-490)	AcALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGI DEYGG
2	ATF6α(38-75)	AcFTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
3	ERM(38-68)	AcDLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
4	ETV4(45-76)	AcLPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
5	FITC-VP16(438-490)	Fl-βAla-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGI DEYGG
6	FITC-ATF6α(38-75)	Fl-βAla-FTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
7	FITC-ERM(38-68)	Fl-βAla-LAHDSEELFQDLSQLQEAWLAEAQVPDDEQ

8	FITC-ETV4(45-76)	Fl-βAla-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
9	4-DMN-VP16(438-490)	4-DMN- βAla-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
10	4-DMN-ATF6α(38-75)	4-DMN- βAla-FTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
11	4-DMN-ERM(38-68)	4-DMN- βAla-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
12	4-DMN-ETV4(45-76)	4-DMN- βAla- LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ

The peptides listed in Table 3.2 were made using standard FMOC solid-phase synthesis methods on a Liberty Blue Microwave Synthesizer (CEM). Deprotection of the FMOC groups were done in 20% piperidine (ChemImpex) in DMF that was sumplemented with 0.2 M Oxyma Pure (CEM) and irradiating under variable power to maintain a temperature of 90˚C for one minute. For coupling reactions, 5 eq of amino acids relative to resin (CEM, ChemImpex, and NovaBiochem) were added to diisopropylcarbodiimide (7 eq, ChemImpex) and Oxyma Pure (5 eq) in DMF and irradiated at variable power to maintain a temperature of 90˚C for 4 minutes. Between all couplings and deprotections, the resin was rinsed four times with excess DMF.

Following synthesis, peptides 1-4 were deprotected one final time and acetylated at the amino terminus using a cocktail of acetic anhydride and triethylamine (Fisher Scientific) in dichloromethane. Peptides 5-8 were deprotected one final time following synthesis, and then treated with fluorescein isothiocyanate (FITC, ThermoFisher) in the presence of N,N-diisopropylethylamine (Sigma Aldrich). The remaining peptides were deprotected and then subjected to one final coupling N-terminally to 4-DMN linked to a β-alanine prior to cleavage. All peptides were cleaved in 95% trifluoroacetic acid (Sigma Aldrich), 2.5% H_20 and 2.5% triisopropylsilane (Sigma Aldrich) for four hours, then filtered. The remaining solution was concentrated under nitrogen and precipitated in cold diethyl ether.

The precipitated peptide was dissolved in 50/50 0.1% TFA in water and acetonitrile with minimal ammonium hydroxide to help with solubility, syringe filtered using a 0.45µn, and purified via HPLC purification. The peptide was purified using reversed phase HPLC on an Agilent 1260 Series instrument with a 250 x 10 mm Luna Omega 5 µm PS C18 column (Phenomenex) using a gradient elution of acetonitrile in 0.1% TFA in water. Purified peptide were tested for purity using analytical HPLC, and identity was determined via mass spectrometry under negative ion mode (Agilent ToF). Once determined to be pure, peptides were dissolved in minimal DMSO and quantified via UV/Vis spectroscopy using either Tyr absorbance at 280 for acetylated peptides (ε_{280}=1,280$M^{-1}cm^{-1}$), FITC absorbance at 495 nm (ε_{495}=72,000 $M^{-1}cm^{-1}$), or 4-DMN absorbance at 440 nm (ε_{440}=10,800 $M^{-1}cm^{-1}$).

3.6 References

1. Krasnov, A. N., Mazina, M. Y., Nikolenko, J. V. & Vorobyeva, N. E. On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell Biosci* **6**, 15 (2016).
2. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
3. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).
4. Henderson, A. R. *et al.* Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *PNAS* 201806202 (2018) doi:10.1073/pnas.1806202115.
5. Henley, M. J. *et al.* Unexpected specificity within dynamic transcriptional protein–protein complexes. *PNAS* **117**, 27346–27353 (2020).
6. Vo, N. & Goodman, R. H. CREB-binding Protein and p300 in Transcriptional Regulation. *J. Biol. Chem.* **276**, 13505–13508 (2001).
7. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the reality. *Chem Biol* **21**, 1102–1114 (2014).
8. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* **3**, 301–317 (2004).
9. Papaleo, E. *et al.* The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem Rev* **116**, 6391–6423 (2016).
10. 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 to Cells* **7**, 49–58 (2002).

11. 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).
12. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nature Structural and Molecular Biology* **18**, nsmb.1999 (2011).
13. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature Structural and Molecular Biology* **18**, nsmb.1997 (2011).
14. Shammas, S. L., Travis, A. J. & Clarke, J. Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *PNAS* **111**, 12055–12060 (2014).
15. Garlick, J. M. *et al. Norstictic acid is a selective allosteric transcriptional regulator*. http://biorxiv.org/lookup/doi/10.1101/2021.03.26.437253 (2021) doi:10.1101/2021.03.26.437253.
16. 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. *PNAS* **110**, E3506–E3515 (2013).
17. 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).
18. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**,.
19. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat Chem Biol* **11**, 891–894 (2015).

20. Smith, M. C. & Gestwicki, J. E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **14**, e16 (2012).
21. Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* **20**, 1455 (2001).
22. Youn, H.-S., Park, U.-H., Kim, E.-J. & Um, S.-J. PTOV1 antagonizes MED25 in RAR transcriptional activation. *Biochemical and Biophysical Research Communications* **404**, 239–244 (2011).
23. Buhrlage, S. J. *et al.* Amphipathic Small Molecules Mimic the Binding Mode and Function of Endogenous Transcription Factors. *ACS Chem. Biol.* **4**, 335–344 (2009).

Chapter 4

Conclusions and Future Directions

4.1 Summary

The work presented in this thesis aims to expand the view of binding mechanisms and molecular recognition of coactivator activator binding domains (ABDs). Specifically, we look to understand the mechanisms by which β-barrel containing ABD homologs form protein-protein interactions (PPIs), with transcriptional activators. This work sheds light on how sequence deviations in the dynamic substructures of a unique coactivator motif affect conformational changes and allostery.

Historically, it has been proposed that molecular recognition occurs over nonspecific interactions, largely driven by electrostatics.[1] Activators were proposed to dictate molecular recognition by recognizing DNA sequences through their DNA binding domains.[2,3] Early data suggested that the transcriptional activation domain (TADs) of activators were largely unstructured in their free state, supporting the notion that there is a lack of specific interactions with coactivator binding partners.[4] Moreover, the only identifiable sequences among TADs were the presence of acidic and hydrophobic residues.[4,5] It was hypothesized that any specificity of recognition came from activator localization to the DNA, or even co-localization events. Once bound to the genomic loci, the DNA-bound activators could then recruit the other components to then initiation transcription.[4,6] This model thus suggests that ABDs can bind to different TADs as a result of the TADs promiscuous behavior.[7–9]

A competing theory to this model suggests more specificity. Often largely ignored, coactivator ABDs play a critical role in molecular recognition. In the model arguing for more specificity, TADs recognition occurs through flexible, intrinsically disordered, nature of the TAD allowing for it to adapt to the surface of its cognate binding partner. More specifically, this would mean that TADs are adapting the topology of ABDs, fitting into the hydrophobic grooves and charged regions of the ABD.[10,11] Even in nonspecific models, ABDs are highly implicated for the binding mechanisms, with posits that ABDs are nothing more than featureless surfaces, largely hydrophobic binding interfaces that interact with the negatively charged sidechains of the TAD.[9]

Deviating from this model of non-specificity is one that highly implicates the role of ABDs. It is well accepted that TADs undergo a coupled folding and binding, but ABD's ability to undergo significant conformational changes has largely been ignored. Structural studies of the ABD KIX have demonstrated that KIX can adopt unique conformations themselves when binding to different activators, suggesting that ABDs can adapt to each TAD. [11,12,13]

Recently the Mapp lab has thoroughly examined the molecular recognition model of transcriptional PPIs using the structurally distinct ABD from the Mediator subunit Med25.[14,15] It was originally proposed that the barrel region, an unusual feature in ABDs, was critical for activator recognition.[16] Extensive biophysical and structural work has shown, however, that it is the dynamic regions of the ABD, similar to that of KIX, that are undergoing conformational changes to accommodate TAD binding. Furthermore, by using a family of highly similar TADs, surprising specificity was observed in ABD recognition, with slight sequence deviations result in a significant degree of conformational sensitivity of AcID.

In this thesis, we examine the binding mechanisms of a unique coactivator ABD, termed Activator Interaction Domain (AcID). This ABD consists of a seven stranded β-barrel flanked by

α-helices and loops. Specifically, we examine three AcID paralogs that share high sequence identity but deviate in their dynamic substructure. Using a biophysical approach, we explore how changes in sequence of these dynamic regions affect activator binding and the concurrent conformational change of the ABD upon binding.

Sequence deviations of dynamic substructures alter conformations upon binding

In chapter two, we dissect how dynamic substructures affects molecular recognition within AcID motifs using AcID paralogs found in Med25 and PTOV1, a second protein in humans that contains two tandem AcID motifs.[17,18] These proteins share high sequence identity but deviate in sequence in their dynamic regions. We use a biophysical approach to explore how these sequence deviations alter the binding mechanisms in complexing with transcriptional activators. Our guiding hypothesis was sequences deviations would alter binding selectivity and the conformational changes of the resulting complexes. We demonstrate that while overall the binding affinities and mechanisms are conserved, we see a loss in conformation states. This is likely due to the loss in secondary structure of the PTOV1 AcIDs, resulting in lower energy between conformations and therefore making it difficult to discern different conformations using current methods. Together, these data suggest there is a degree of specificity in binding amongst the AcID motifs.

We show that sequences changes in the dynamic substructures alter the conformations. PTOV1 has been observed to outcompete Med25 for binding to CBP when overexpressed in metastatic prostate cancer, these changes in sequence are highly implicated in function. Specifically, the changes in conformational plasticity as a result of the sequence deviations may alter transcriptional output by altering the AcID•activator complex formation timescale. Due to its large role in many cancer states, it is crucial to determine PTOV1's role in modulating transcription, as well as further dissect its antagonistic behavior.

Dynamic substructures dictate allostery and recognition

In chapter three, we use the AcID paralogs to determine how sequence deviations in the dynamic regions will alter allosteric communication. We establish that allosteric communication is conserved in the AcID motifs, but that PTOV1A has a shorter-lived ternary complex compared to Med25 AcID. Using this information, we look at how changes in the dynamic substructures alter allosteric modulator binding. Using the natural product garcinolic acid, we demonstrate it is capable of allosterically inhibiting ERM binding and is more potent against the PTOV1A-ERM than Med25's interaction with ERM. Additionally, we investigate how hotspots can be used to differentiate highly similar TAD sequences by using a biophysical approach to examine how slight changes in TAD sequence can result in differential changes in coactivator conformations. Specifically, we tested how slight sequence deviations can lead to changes in conformation of AcID motifs. Given that these TADs are highly related, this system served as an opportunity to demonstrate mechanism specificity.

The work in this chapter demonstrates that the nonspecific model of recognition is outdated. Through this work, we begin to demonstrate that the AcID motifs utilize their differences in dynamic substructures to recognize different binding partners. Given their overall plasticity, we can use these less conserved allosteric regions to selectively and specifically target these ABDs. Protein-protein interactions between activators and coactivators have long been termed undruggable due to the difficult nature in studying by traditional means. While it was originally proposed that binding was driven through electrostatic and hydrophobic interactions, this model does not account for the intricate specificity and selectivity required for such a highly regulated process as transcription initiation. Towards this end, recent work in the Mapp lab has demonstrated that molecular recognition is driven through conformational changes driven through dynamic

regions. Given their role in disease it is critical to identify small molecule inhibitors that can selectively target the different AcID motifs.[18–20]

4.2 Future Directions

The role of PTOV1 in modulating transcription

The AcID motif has only been identified in higher eukaryotic organisms. In fact, it has only been identified in one other protein aside from Med25, Prostate Tumor Overexpressed Variant 1 (PTOV1). PTOV1 was identified using differential display screening methods. Work by the Um group identified PTOV1 as having two tandem AcID motifs, termed PTOV1A and PTOV1B (Fig. 4.1). Sequence alignments to Med25's AcID showed PTOV1A and PTOV1B share 81% and 71% sequence identify, respectively. This raised the question of the role of PTOV1's two AcID motifs. PTOV1 was not found to associate with the Mediator complex, and was found originally to be overexpressed in primary tumors and early prostate cancer patients. In most healthy cells, PTOV1 is undetectable, but has since been shown to be overexpressed in a variety of early and late stage cancers.

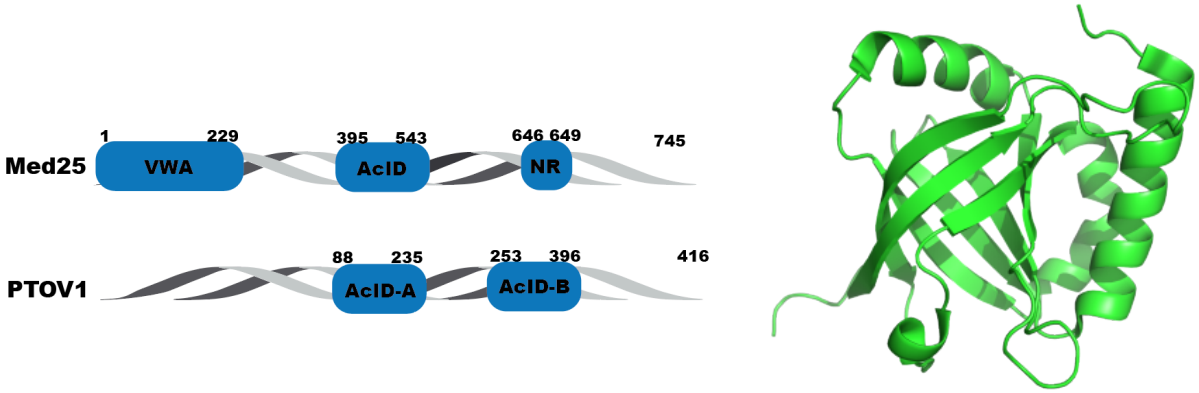


Figure 4.1: Domain architect and structure of the AcID containing proteins. On the left is the domain layout of Med25 and PTOV1. On the right is the structure of Med25 AcID (PDB 2xnf).

Extensive research has demonstrated that PTOV1 plays a variety of roles, acting as an 'adaptor protein,' that can regulate gene expression at both the transcriptional and translational levels.[18,21,23] One proposed role, and a role relevant to the work in this thesis, is that PTOV1 can

act as a negative regulator of Med25 AcID function. When overexpressed, PTOV1 was found to block the Med25•CBP interaction. Specifically, PTOV1 was found to bind to outcompete Med25 binding to CBP, thus preventing Med25 from binding to CBP at the retinoic acid receptor (RAR) in the presence of retinoic acid (Fig 4.2).[17,24] Moreover, reports suggest that PTOV1, while undetectable in most normal cell tissues, is overexpressed in a wide-range of early and late stage cancers.[18,21,25] These studies highlight that it is critical to study PTOV1's biological function. Not only would it provide insight into a variety of disease states, but it can lead to further discovery of Med25 function and regulation.

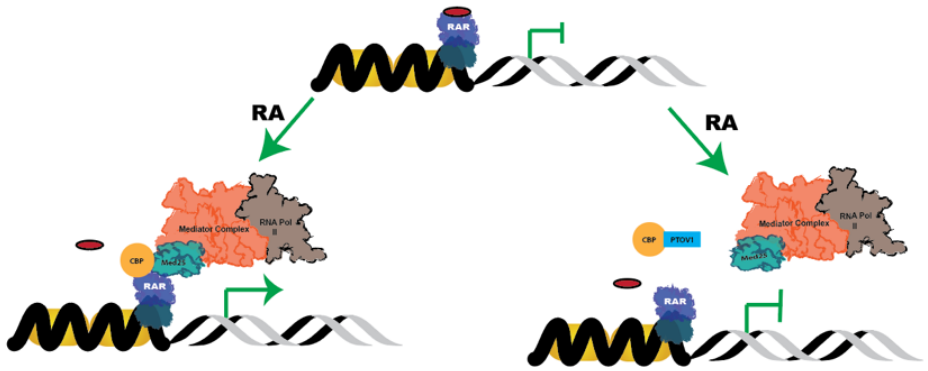


Figure 4.2: Binding events of Med25/CBP activity in the presence and absence of PTOV1. In healthy cells, the presence of retinoic acid allows for the repressor to be removed from RAR, allowing for Med25 to bind to CBP through its AcID domain and to RAR through its NR. CBP also binds to RAR. However, in metastatic prostate cancer, overexpressed PTOV1 binds to CBP, blocking it from binding to Med25.

Recently it has been discovered that both AcID domains in PTOV1 contain nuclear localization sequences.[25] This information, coupled with the emergent data from this thesis, poses the question of PTOV1's role in modulating transcription. Specifically, it would be critical to look at the role of the two AcID motifs, as we have demonstrated that the sequence deviations in the dynamic regions results in different conformational signatures upon TAD binding. One proposed hypothesis is that PTOV1 is using one AcID motif to localize to the genomic loci, either directly to the DNA or to other transcription factors, and the other is binding to the rest of the transcriptional components, thus bypassing the need for the Mediator complex.

One way to test this hypothesis is to utilize a mammalian two-hybrid system using the prostate cancer cell PC-3, which has high mRNA levels of PTOV1.[26,27] Another option would to conduct an *in vitro* transcriptional assay. Specifically, by using nuclear extracts from Med25 knockouts derived from patient derived VARI068 cells, we can determine not only if PTOV1 can activate transcription, but further probe the proposed squelching mechanism PTOV1 utilizes to outcompete Med25.[28–30] Given PTOV1's role in cancer, it is critical to find inhibitors that can selectively target one AcID motif over the other. One potentially powerful method is to utilize allosteric modulators as these can exploit the less conserved dynamic substructures.

4.3 References

1. Sigler, P. B. Acid blobs and negative noodles. *Nature* 333, 210–212 (1988).
2. Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* 45, 6873–6888 (2006).
3. Berlow, R. B., Dyson, H. J. & Wright, P. E. Expanding the Paradigm: Intrinsically Disordered Proteins and Allosteric Regulation. *J Mol Biol* 430, 2309–2320 (2018).
4. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* 386, 569–577 (1997).
5. 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 U S A* 90, 883–887 (1993).
6. Brent, R. & Ptashne, M. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43, 729–736 (1985).
7. Brzovic, P. S. *et al.* The acidic transcription activator Gcn4 binds the Mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell* 44, 942–953 (2011).
8. Warfield, L., Tuttle, L. M., Pacheco, D., Klevit, R. E. & Hahn, S. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc Natl Acad Sci U S A* 111, E3506-3513 (2014).
9. Tuttle, L. M. *et al.* Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep* 22, 3251–3264 (2018).
10. Dyson, H. J. & Wright, P. E. Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. *J Biol Chem* 291, 6714–6722 (2016).

11. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021–1025 (2007).
12. Shammas, S. L., Travis, A. J. & Clarke, J. Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc Natl Acad Sci U S A* 111, 12055–12060 (2014).
13. 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).
14. Henderson, A. R. *et al.* Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *PNAS* 201806202 (2018) doi:10.1073/pnas.1806202115.
15. Henley, M. J. *et al.* Unexpected specificity within dynamic transcriptional protein–protein complexes. *PNAS* 117, 27346–27353 (2020).
16. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nature Structural and Molecular Biology* 18, nsmb.1997 (2011).
17. Youn, H.-S., Park, U.-H., Kim, E.-J. & Um, S.-J. PTOV1 antagonizes MED25 in RAR transcriptional activation. *Biochemical and Biophysical Research Communications* 404, 239–244 (2011).
18. Cánovas, V., Lleonart, M., Morote, J. & Paciucci, R. The role of prostate tumor overexpressed 1 in cancer progression. *Oncotarget* 8, 12451–12471 (2016).
19. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the reality. *Chem Biol* 21, 1102–1114 (2014).
20. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat Chem Biol* 11, 891–894 (2015).

21. Santamaría, A. *et al.* PTOV-1, a Novel Protein Overexpressed in Prostate Cancer, Shuttles between the Cytoplasm and the Nucleus and Promotes Entry into the S Phase of the Cell Division Cycle. *Am J Pathol* 162, 897–905 (2003).
22. 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. *PNAS* 101, 2339–2344 (2004).
23. Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* 20, 1455 (2001).
24. Youn, H., Kim, E.-J. & Um, S.-J. Zyxin cooperates with PTOV1 to confer retinoic acid resistance by repressing RAR activity. *Cancer Letters* 331, 192–199 (2013).
25. Maggio, V. *et al.* A novel DNA-binding motif in prostate tumor overexpressed-1 (PTOV1) required for the expression of ALDH1A1 and CCNG2 in cancer cells. *Cancer Letters* 452, 158–167 (2019).
26. Luo, Y., Batalao, A., Zhou, H. & Zhu, L. Mammalian two-hybrid system: a complementary approach to the yeast two-hybrid system. *BioTechniques* 22, 350–352 (1997).
27. Stynen, B., Tournu, H., Tavernier, J. & Dijck, P. V. Diversity in Genetic In Vivo Methods for Protein-Protein Interaction Studies: from the Yeast Two-Hybrid System to the Mammalian Split-Luciferase System. *Microbiol. Mol. Biol. Rev.* 76, 331–382 (2012).
28. Losick, R. In Vitro Transcription. *Annual Review of Biochemistry* 41, 409–446 (1972).
29. Baugh, L. R., Hill, A. A., Brown, E. L. & Hunter, C. P. Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res* 29, e29–e29 (2001).
30. Aw Yong, K. M. *et al.* Heterogeneity at the invasion front of triple negative breast cancer cells. *Scientific Reports* 10, 5781 (2020).

Appendix I

Initial Discovery and Characterization of Garcinolic Acid[1]

Garcinolic acid (GA), a xanthone natural product isolated from *Garcinia*, was originally identified as an inhibitor of CBP KIX (Fig. A.1). Working with Dr. Meghan Breen, GA emerged from library screen of more than 5900 compounds comprised of natural products, FDA-approved drugs and bioactive molecules using a competitive-inhibition binding assay. The screen was originally conducted using Med15 KIX from *Saccharomyces cerevisiae* (sc) and *Candida galbrata* (cg). However, when we conducted follow-up screening using a fluorescence polarization-based competition assay and found that GA is highly selective for the CBP GACKIX motif, only modestly inhibiting Gal11/Med15 KIX complexes, and not interacting at all with Arc105 KIX (Table A.1).

[1]The work in this section was completed in collaboration with Dr. Meghan Breen and Dr. Matthew Beyeresdor.

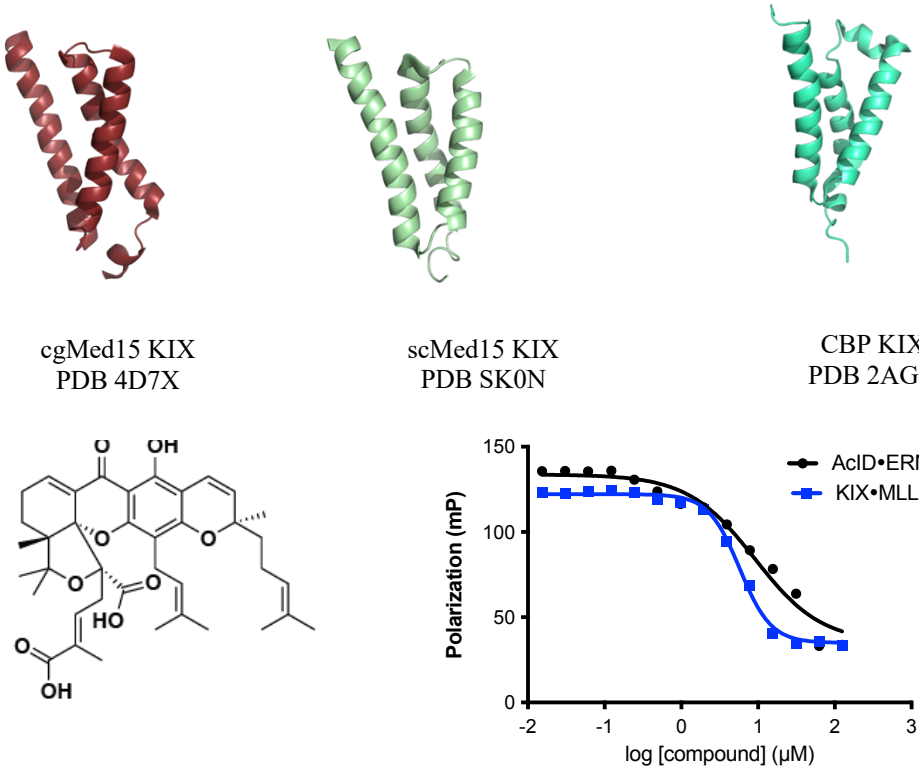


Figure A.1: Garcinolic acid and the various KIX motifs. Pictured above are the structures of KIX from various organisms, including yeast, as well as garcinolic acid. Additionally, we demonstrate that GA potently inhibits CBP KIX-MLL interactions.

We demonstrate that GA is a potent inhibitor and shows preferential inhibition to CBP/p300 KIX. Several lines of evidence indicated that GA was a reversible, noncovalent inhibitor of KIX, despite the presence of at least two potentially reactive site. No detectable covalent adducts were observed via mass spectrometry, even after extended incubation. Moreover, transient stopped-flow experiments conducted by Dr. Matthew Henley reveals a reversible binding mechanism.

Table A.1: Summary of the inhibitory effects of GA against various KIX domains. Half maximal inhibitory constants of GA for tested TAD•ABD PPIs. Data is in triplicate and reported error is the standard deviation. (Data obtained with the help of Dr. Meghan Breen).

Coactivator Domain	TAD	IC_{50} (μM)
scMed15 KIX	Pdr1	27
cgMed15 KIX	Pdr1	59

CBP KIX	MLL	3.3 ± 0.9
CBP KIX	c-Myb	4.5 ± 1
p300 KIX	MLL	4.6 ± 1
p300	c-Myb	7.2 ± 2

GA inhibited MLL and cMyb TADs with similar IC_{50} values despite their binding to opposite binding faces. As stated in earlier chapters, there exists a well characterized allosteric network between these two binding faces. Due to the observation GA inhibiting both TADs, this raised the question of whether GA orthosterically bound to one face and induced inhibition at the second site via allostery. Through extensive analysis using HSQC and Protein-Observed ^{19}F (PrOF) NRM completed by Dr. Matthew Beyersdorf, GA was shown to engage at the MLL binding site, thus blocking the binding of this TAD (Fig. A.2). Additionally, perturbations were observed at the conformationally dynamic loops that flank the MLL binding site. Therefore, it is not surprising that GA engagement to the MLL site also inhibits c-Myb binding. In fact, the conformation created from GA-KIX engagement resembles a previously reported conformation that is shown to disfavor binding to the c-Myb site. Taken together, these initial data demonstrate that GA has the potential to serve as a powerful chemical probe to explore the binding mechanisms of CBP/p300 KIX. We show that GA is capable of orthosteric and allosteric inhibition, and this can be useful in the development of chemical scaffolds to selectively modulate CBP KIX interactions.

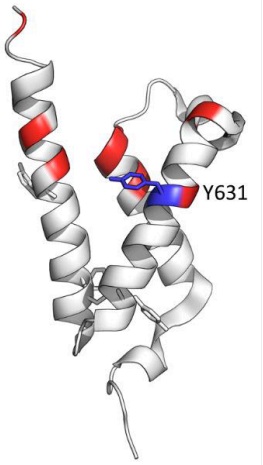


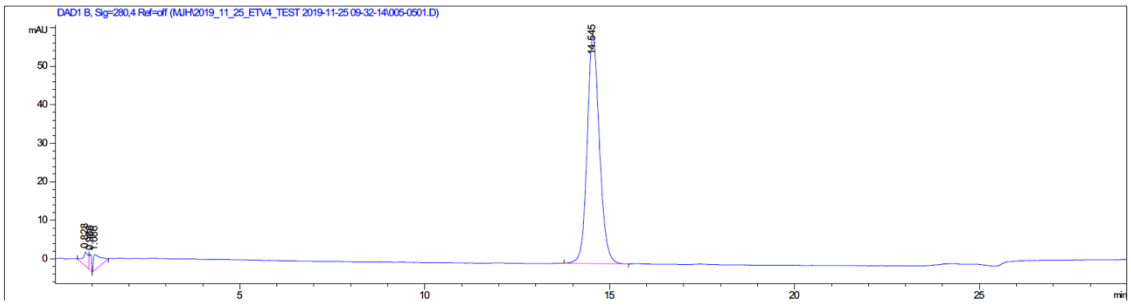
Figure A.2: GA binds near the MLL site and allosterically inhibits c-Myb binding. NMR experiments conducted by Dr. Matthew Beyersdorf reveal that GA causes significant chemical shift perturbations near the MLL site. HSQC NMR perturbations are shown in red, and the only 3FY residue that was perturbed in PrOF NMR was Y631 on the MLL site (depicted as a blue stick), consistent with the HSQC data.

Materials and Methods

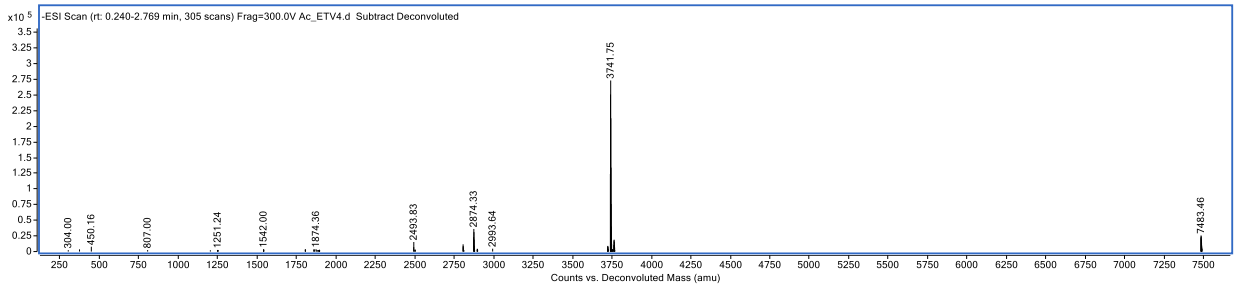
All relative materials and methods have been previously described in chapter 3.

Appendix II

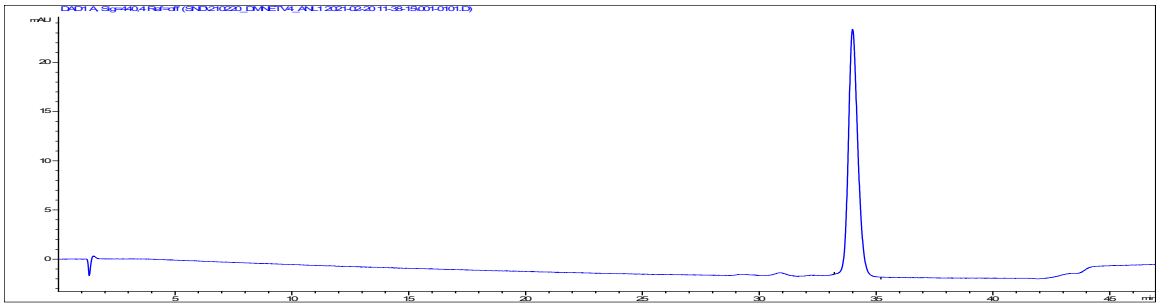
Peptide Characterization



Analytical HPLC trace of **Ac-ETV4 (45-76)** monitored at 280 nM. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increase from 10-35% acetonitrile over 20 mins.

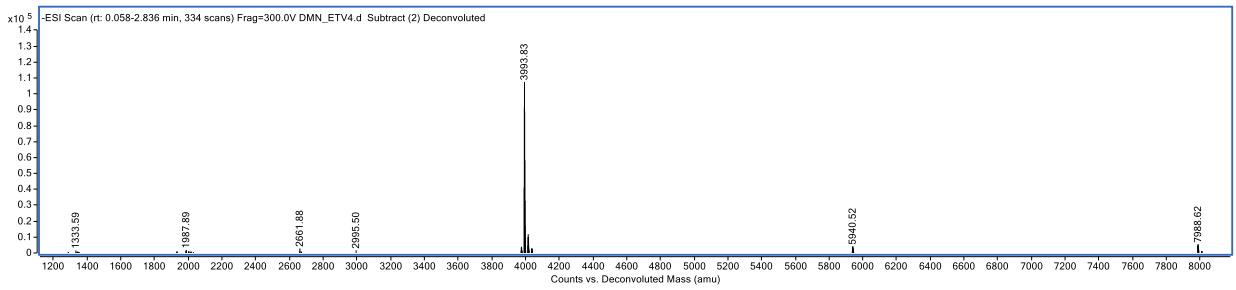


LC-MS of **Ac-ETV4 (45-76)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.

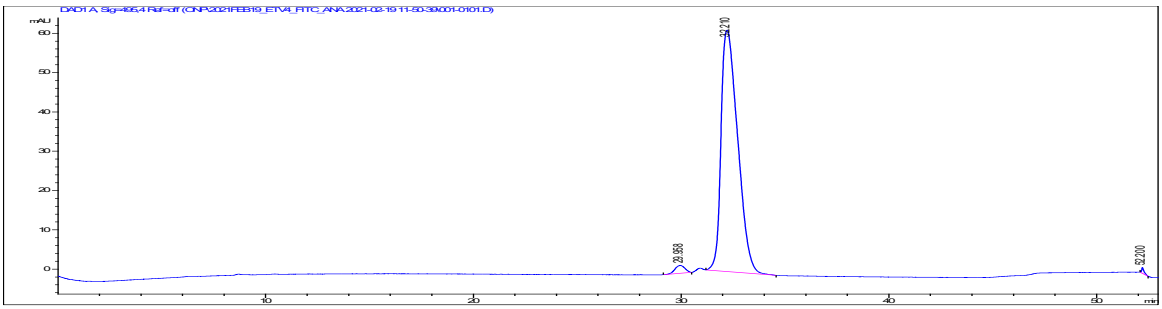


Analytical HPLC trace of **DMN-ETV4 (45-76)** monitored at 440 nM. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with

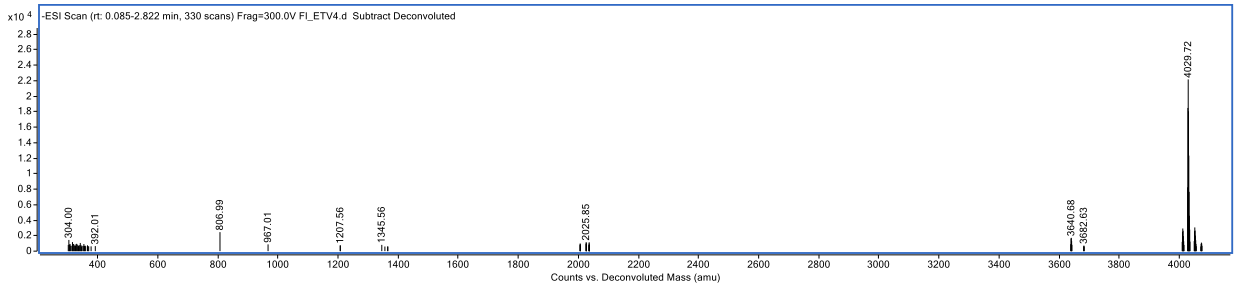
an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



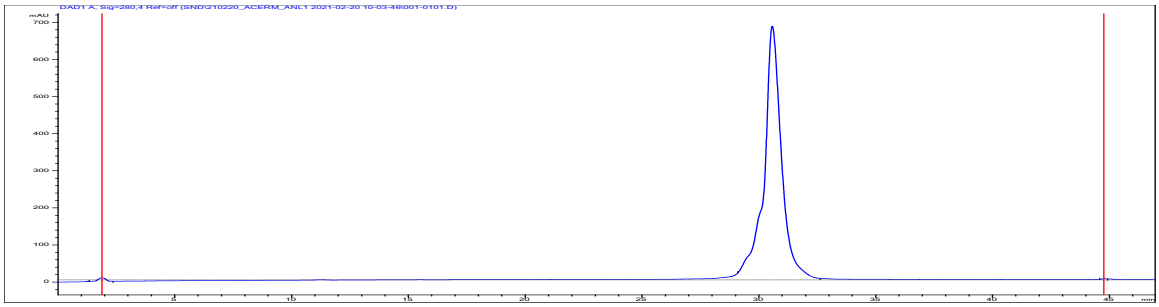
LC-MS of **4-DMN-ETV4 (45-76)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



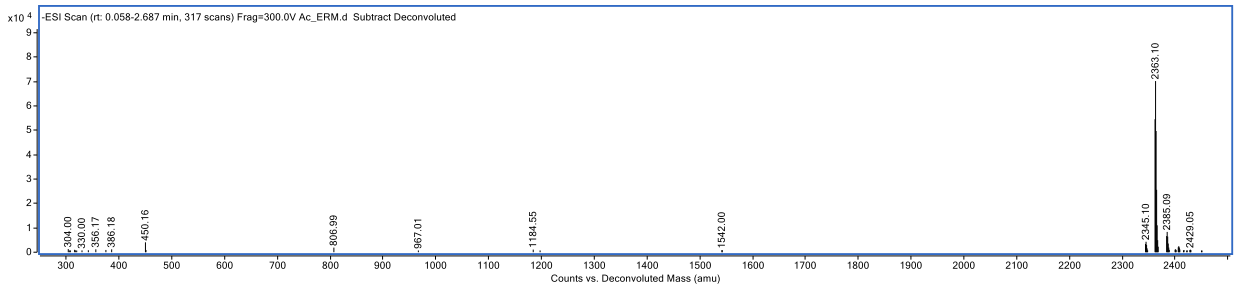
Analytical HPLC trace of **FITC-ETV4 (45-76)** monitored at 495 nM. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



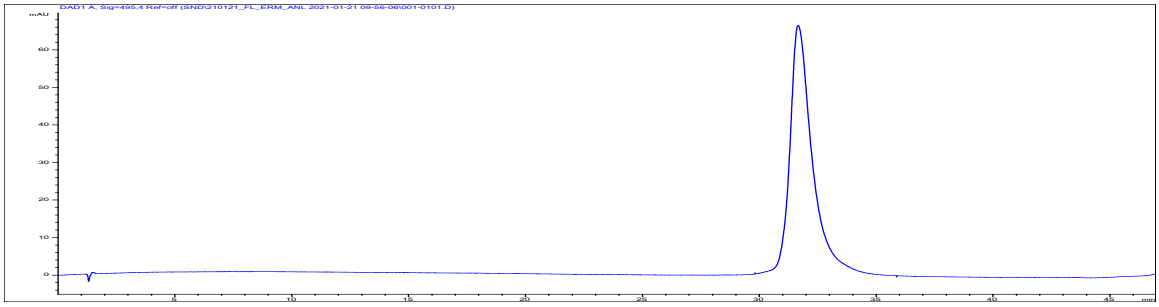
LC-MS of **FITC-ETV4 (45-76)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



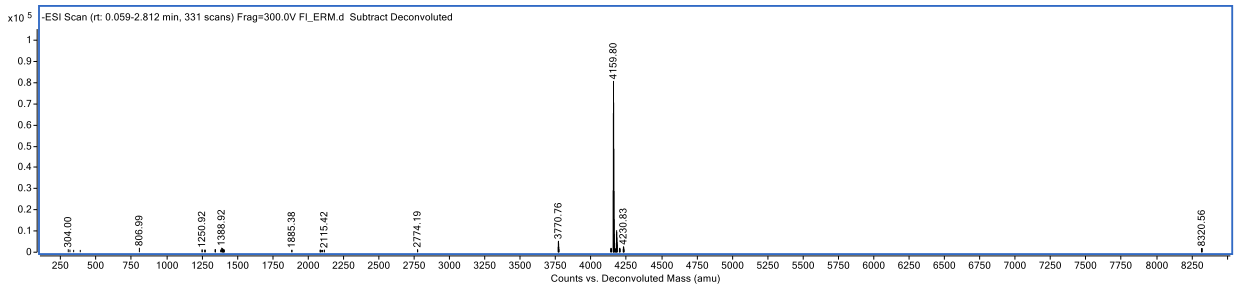
Analytical HPLC trace of **Ac-ERM (38-68)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



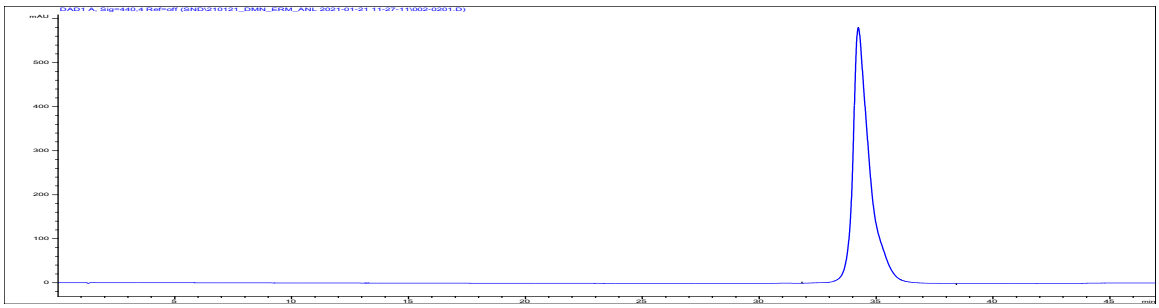
LC-MS of **Ac-ERM (38-68)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



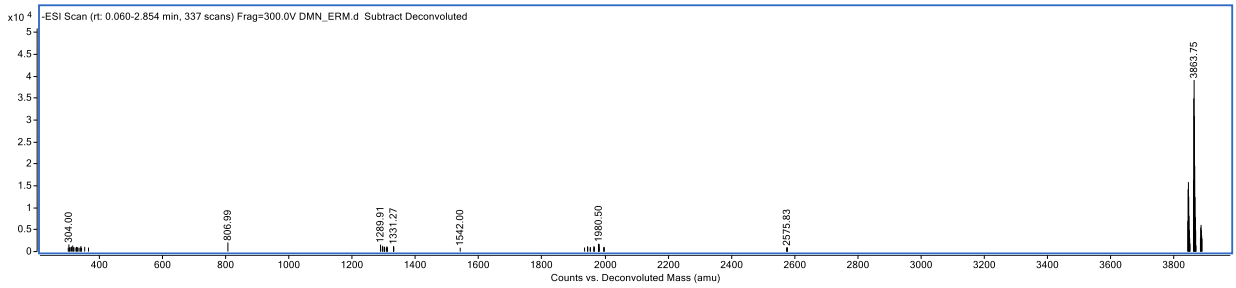
Analytical HPLC trace of **FITC-ERM (38-68)** monitored at 440 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



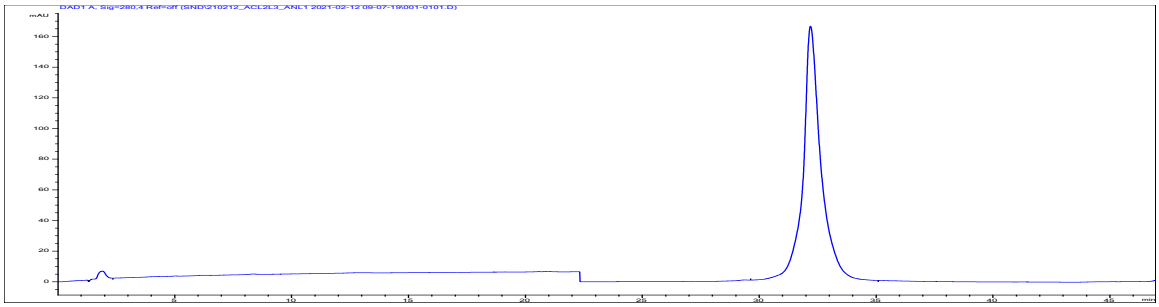
LC-MS of **FITC-ERM (38-68)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



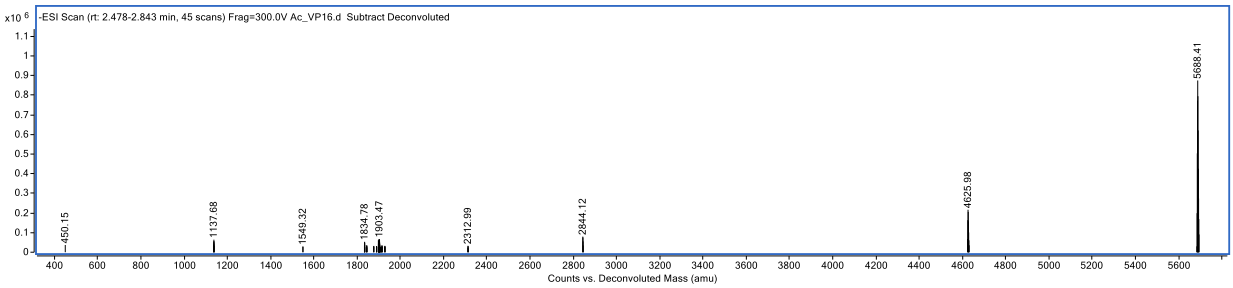
Analytical HPLC trace of **4-DMN-ERM (38-68)** monitored at 440 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



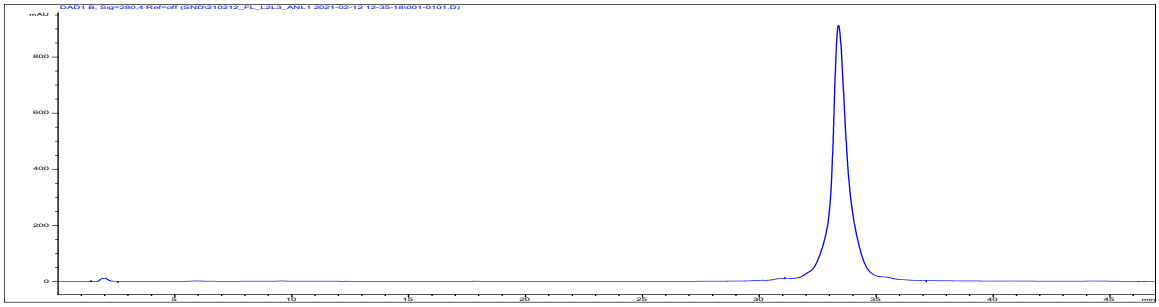
LC-MS of **4-DMN-ERM (38-68)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



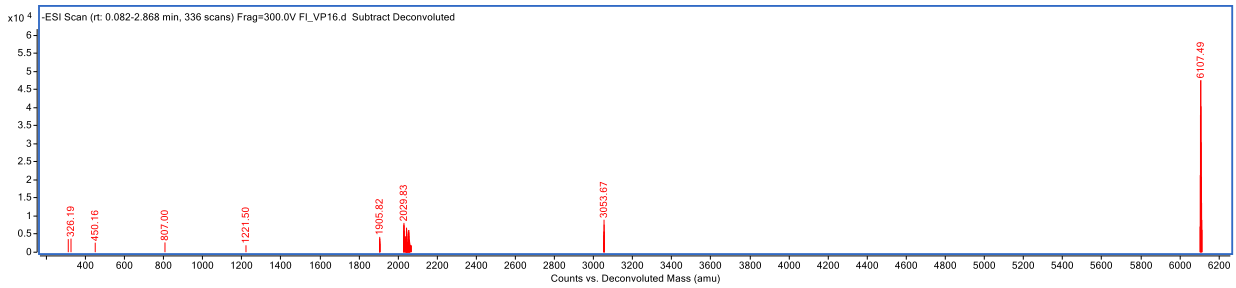
Analytical HPLC trace of **Ac-L2L3 VP16 (438-490)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



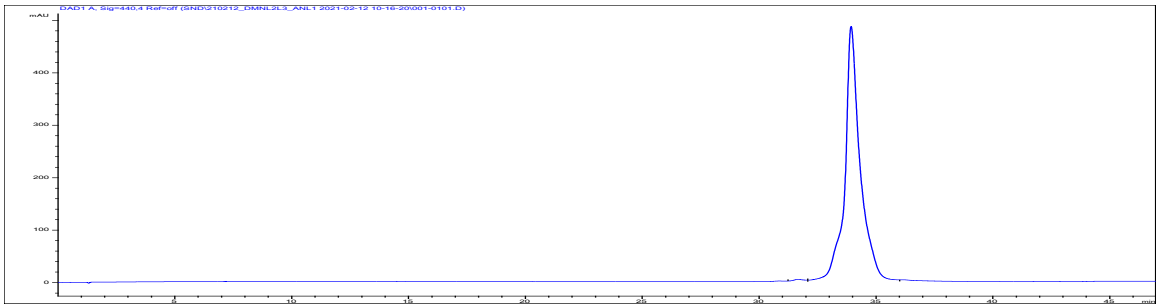
LC-MS of **Ac-L2L3 VP16 (438-490)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



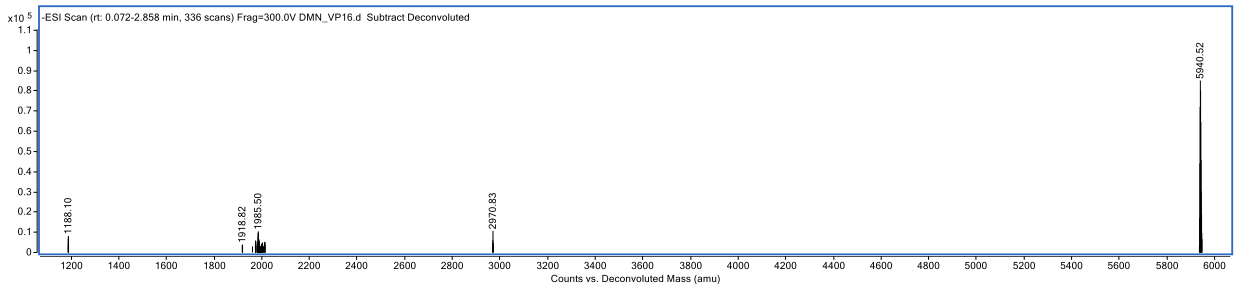
Analytical HPLC trace of **FITC-L2L3 VP16 (438-490)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



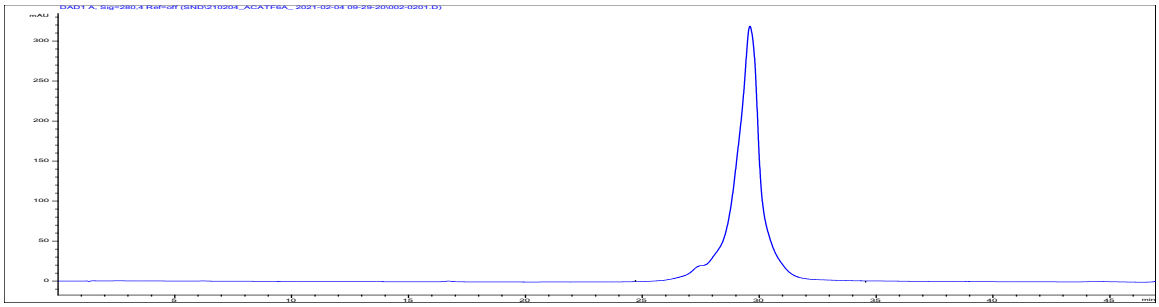
LC-MS of **FITC-L2L3 VP16 (438-490)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



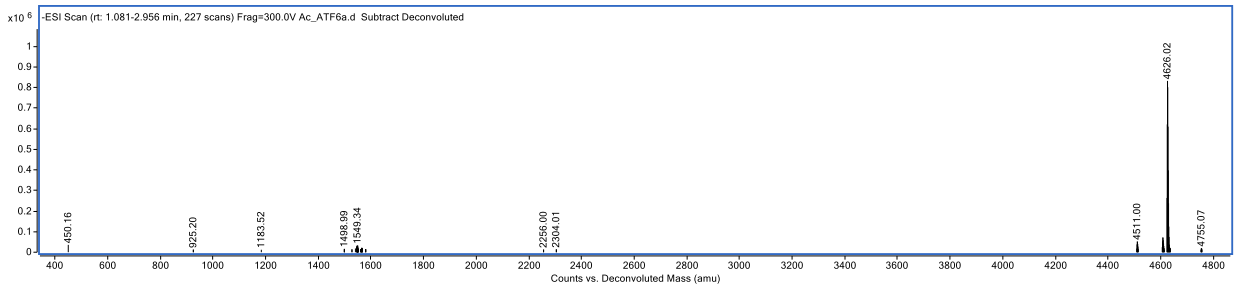
Analytical HPLC trace of **4-DMN-L2L3 VP16 (438-490)** monitored at 440 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



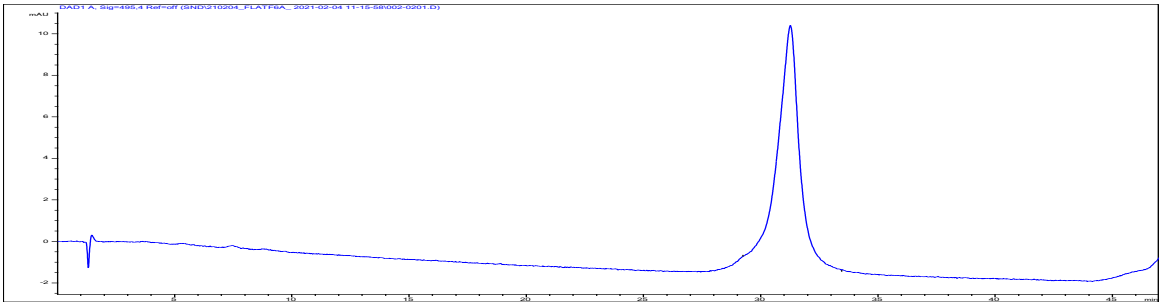
LC-MS of **4-DMN-L2L3 VP16 (438-490)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



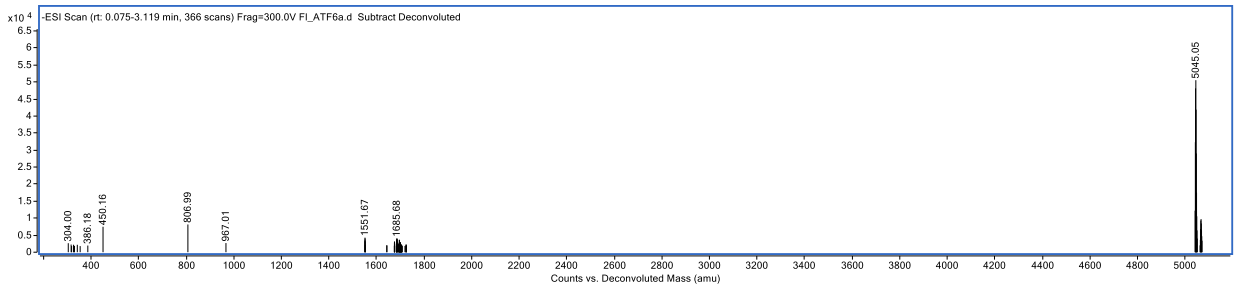
Analytical HPLC trace of **Ac-ATF6α (38-75)** monitored at 280 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



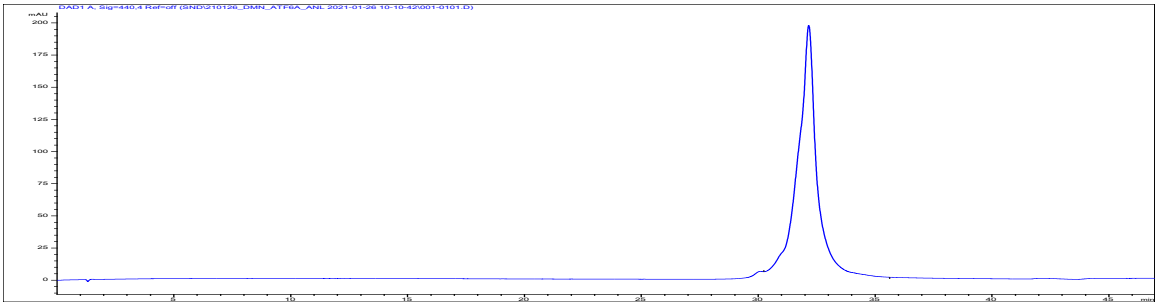
LC-MS of **Ac-ATF6α (38-75)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



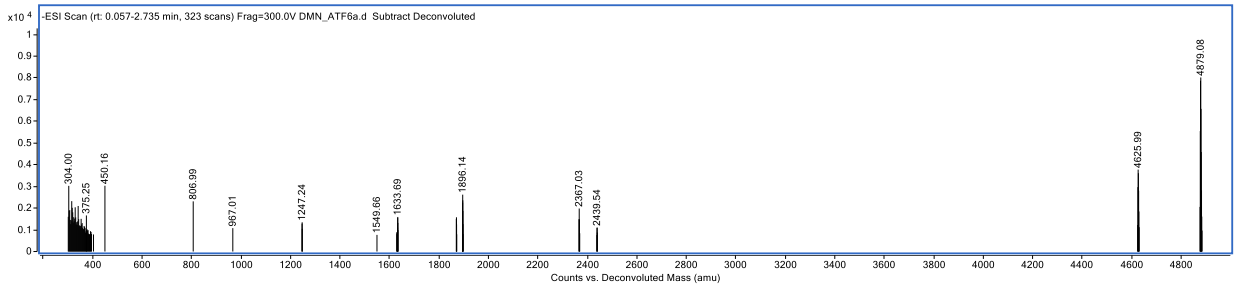
Analytical HPLC trace of **FITC-ATF6α (38-75)** monitored at 495 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



LC-MS of **FITC-ATF6α (38-75)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.



Analytical HPLC trace of **4-DMN-ATF6α (38-75)** monitored at 440 nM. Analytical sample was run in a water (with 0.1% TFA)/ acetonitrile system. The sample was injected with an isocratic flow of 50% water (with 0.1% TFA) and 50% acetonitrile. After 2 mins, the solvent gradient was increased from 10-50% acetonitrile over 40 mins.



LC-MS of **4-DMN-ATF6α (38-75)** using an Agilent TOF. Samples were run in 50/50 0.1% TFA in water, and acetonitrile. Samples were injected onto a C8 column with a C4 guard. Identity was confirmed under negative mode ionization conditions.