Mechanistic Analysis of Dynamic Transcriptional Protein-Protein Interactions

by

Matthew Henley

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2020

Doctoral Committee:

Professor Anna K. Mapp, Chair
Associate Professor Tomasz Cierpicki
Assistant Professor Kristin S. Koutmou
Associate Professor Patrick J. O'Brien

Matthew Henley

mjhenley@umich.edu

ORCID iD: 0000-0002-0709-6186

Acknowledgements

In these last few months, I have reflected a considerable amount on how I made it to where I am today. Growing up in rural Minnesota, I never imagined I would be getting a Ph.D. from a place like the University of Michigan, let alone in a field called "Chemical Biology." I owe a huge debt to the people who have inspired me along this path, who have enabled me to push past my doubts and take challenges head on.

I firstly need to thank my advisor, Anna, for giving me the trust and space I needed to develop as an independent scientist, and for always pushing me to think past my current level of understanding and think about my science on a broader scale. I have grown immensely over these last four years in your lab, and I was so lucky to have an advisor willing to let me go in my own direction.

I also would like to thank my committee: Pat, Tomek, and Kristin, for their feedback and guidance over these years. It is exhilarating to me to talk science with you and to hear your perspectives on my work. I deeply appreciated my collaboration with Tomek and his graduate student Brian, learning about different NMR experiments and approaches with you was a major scientific highlight of my time in Michigan. I need to thank Carol Fierke as well; when she was on my thesis committee, she was an outspoken advocate for me and constantly reminded me that I had the potential to do great things if I pushed myself. I also distinctly remember that when I was having trouble communicating a concept or idea, Carol would know exactly the right way to ask a question to help me communicate. This had a profound impact on me.

I also have to thank the Mapp Lab for being such an energetic group of scientists to work with over the past years. I especially enjoyed the opportunities I had to mentor younger students, including Mitchell, Yujin, and Natalie. It was elating to see you take such challenging projects and put your own spin on them.

I am especially glad that I had the privilege to work with Dr. Brittany Morgan over the past two years. I can attribute a significant portion of my scientific growth in the last

me smile and laugh. By far the most difficult day of graduate school was not any failed experiment or project but losing Baby to cancer at the end of my fourth year.

And finally, I need to say thank you to my partner, Gray. Your resolve, your desire to improve yourself and persevere in the face of all life's challenges has been a constant source of inspiration to me. I am reminded every day how lucky I am to know you. I love you so much, and I can't wait to begin our next adventure.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABD	Activator binding domain
ATF6α	Activation transcription factor 6α
CBP	CREB binding protein
CSP	Chemical shift perturbation
DBD	DNA binding domain
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
4-DMN	4-*N,N*-dimethylamino-1,8-naphthalimide
Ets	E-twenty six
ETV	Ets translocation variant
GTF	General transcription factor
HIF1α	Hypoxia inducible factor 1 alpha
HSQC	Heteronuclear single quantum coherence spectroscopy
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
KIX	Kinase inducible domain interaction domain
LBD	Ligand binding domain
Med15	Mediator subunit 15
Med25	Mediator subunit 25
NCBD	Nuclear coactivator binding domain
NMR	Nuclear magnetic resonance spectroscopy
NOE	Nuclear Overhauser effect
NR	Nuclear receptor
PEA3	Polyoma enhancer activator 3

PIC   Pre-initiation complex

PPI   Protein-protein interaction

PPM   Parts per million

RNA   Ribonucleic acid

RNA Pol II   RNA Polymerase II

TAD   Transcriptional activation domain

TAZ   Transcriptional adapter zinc finger

TEAD   Transcriptional enhanced associate domain

TF   Transcription factor

TOCSY   Total correlation spectroscopy

TROSY   Transverse relaxation optimized spectroscopy

VP16   Herpes simplex virus protein 16

YAP   Yes-associated protein

List of Appendices

Abstract

Transcription, the basal process by which all genes are regulated, is one of the most important tasks a cell undertakes. The molecular details of this process are well-studied, from the enzymatic action of RNA Polymerase II, the assembly of factors that load RNA Polymerase II onto a gene promoter, to the specific sets of transcriptional regulators that control the expression of individual genes. However, despite decades of study, the key regulatory interactions occurring between DNA-bound transcriptional activators and the transcriptional machinery are poorly understood. This is largely due to the preponderance of intrinsically disordered proteins involved in the process of transcriptional activation, leading to dynamic interactions difficult to characterize from a structural and mechanistic standpoint. Thus, a key challenge is to develop mechanistic models that explain how transcriptional activators are recognized by transcriptional machinery.

Currently, molecular recognition models of transcriptional activators are based on analyses of a relatively small set of structurally related binding partners. The motivation behind this dissertation was therefore to dissect these recognition models by examining how a structurally divergent binding partner, namely the activator binding domain of Mediator subunit Med25, recognizes its activator binding partners.

In chapter two, we dissect how Med25 forms binary and ternary complexes with a set of unrelated transcriptional activators. Using NMR and transient kinetic analysis, we find that Med25 uses conformational rearrangements along with two distinct binding interfaces to recognize partners with diverse sequences. Furthermore, kinetics experiments show that this mechanism of molecular recognition enables cooperative formation of ternary

complexes with activators that bind to distinct sites. Molecular dynamics simulations demonstrate that, similar to the mechanisms of well-studied activator binding domains, conformational changes and allosteric communication are mediated by dynamic substructures in the activator binding domain. We establish the applicability of this observation to small molecule discovery by using disulfide-Tethering technology to discover a small molecule that covalently targets one of these dynamic substructures and induces allosteric effects that mimic natural activators.

In chapter three, we examine the mechanisms by which Med25 forms complexes with a set of highly related activators. The common molecular recognition models of highly dynamic activator•coactivator interactions dictate that these "fuzzy" complexes are formed through entirely nonspecific mechanisms. In contrast, transient kinetics experiments demonstrate that small changes in the activator sequence result in redistribution of the activator•Med25 conformational ensemble, suggesting that the ensemble is controlled by specific intermolecular interactions. NMR analysis indicates the sensitivity of activator•Med25 conformational ensembles originates from specific interactions formed between the activator and the Med25 surface. Furthermore, this specific activator•Med25 recognition mechanism is enabled by the ability of the Med25 binding surface to remodel its conformation to complement the activator.

This dissertation examines in detail how the structurally divergent activator binding domain of Med25 recognizes transcriptional activators and contributes to how activator molecular recognition is understood broadly. The work in this dissertation advances the conserved mechanistic role that coactivator conformational plasticity plays in molecular recognition and highlights a general framework for the development of small molecule modulators that exploit this mechanism. Finally, this work demonstrates that conformational remodeling of coactivators can play a key role even in the formation of exceptionally dynamic activator•coactivator complexes, suggesting that plasticity of folded binding partners is a critical element underlying molecular recognition of other dynamic protein-protein interactions.

CHAPTER 1.

Role and Molecular Recognition of Intrinsically Disordered Proteins in Transcriptional Activation

1.1 Introduction

Transcription, the process by which all gene expression is regulated, is one of the most critical molecular processes in all biology. Transcription is carried out through the enzyme RNA Polymerase II (RNA Pol II) reading the DNA sequence of a gene and synthesizing a complementary strand of messenger RNA, which is then translated into a polypeptide by the ribosome. While transcription is an enzymatic process, it is controlled almost entirely by noncovalent protein-DNA and protein-protein interactions (PPIs).[1–4] This is due to the fact that RNA Pol II cannot act on DNA sequences without being loaded by over 30 factors that form the pre-initiation complex,[5–8] which themselves are controlled by the action of transcription factors (TFs) such as transcriptional activators and coactivators.[1]

As a class of proteins, TFs are unique in their high propensity to contain sequences that are enriched in intrinsically disordered regions (IDRs).[9] Indeed, 80-90% of TFs are predicted to contain extended IDRs, significantly more than other protein classes.[9] While several decades ago it was commonly believed that IDRs were functionally irrelevant because they formed no specific structure,[10] studies of TFs in the 1980s and 1990s indicated that TF IDRs are often critical to function.[11–13] For example, early biochemical studies showed that disordered regions of transcriptional activators such as VP16,[14–17] Gal4,[18] and Gcn4[12,19] are *required* for the function of the transcriptional activator. IDRs of TFs typically mediate function through the formation of PPIs with other members of the transcriptional machinery,[1,9] and the intrinsic plasticity of IDRs is thought to be critical for these regions to utilize the same sequence to mediate PPIs with multiple distinct binding partners.[9,20–22] Because of the early recognition of functional IDRs in several TFs, many

biophysical and structural studies of IDRs used TF PPIs as model systems.[23–33] Thus, TFs have significantly shaped how functional protein disorder is understood at a broad level.

In addition to being a biophysical marvel, the activity of TFs often play a direct role in disease and are thus attractive targets for therapeutic intervention.[34–36] For example, the oncogenic transcriptional activator cMyc is involved in 40% of all cancers and loss of cMyc results in cancer regression.[37] Many transcriptional PPIs are also specific to subsets of genes, making them attractive targets for selective therapeutics with minimal on-target toxicity. However, despite their promise as drug targets, there are relatively few successes of targeting transcriptional PPIs. This is especially true for the PPIs of transcriptional activators, which are some of the most therapeutically appealing TFs due to the direct role they play in driving disease phenotypes.[37,38] There are several reasons for this outcome, including the challenge of targeting PPIs in general,[39,40] but a significant contributor is that disorder in one or both partners of a given PPI can make structural characterization of the resulting complex exceptionally challenging.[41] Thus, outside of the select transcriptional PPIs that have been heavily characterized as model systems, the basis for molecular recognition of transcriptional PPIs involving IDRs is poorly understood. Because medicinal chemistry campaigns significantly benefit from structural and mechanistic information about the target, this serves as an impediment for the development of small molecule chemical probes and/or therapeutics that target transcriptional PPIs.[42,43]

IDRs contribute to a multitude of transcriptional functions including DNA binding,[44] epigenetic regulation,[45,46] TF trafficking,[47,48] and assembly of the transcriptional machinery.[1] Of these, the IDRs that mediate recruitment of the transcriptional apparatus to promoters and enhancer regions are some of the least understood from a functional and mechanistic standpoint, yet they hold exceptional therapeutic promise as targets for selective modulation of gene expression.[42] In this chapter, we review the mechanisms by which IDRs in transcriptional activators facilitate transcriptional upregulation through PPIs, and how these IDRs are recognized by their binding partners.

1.2 Transcriptional Activators as Drug Targets

The therapeutic potential of transcriptional activators is well-documented, especially as cancer targets.[34–36,43] A significant advantage of targeting these TFs is that it represents a direct approach: the activity of specific activators is often required for the production of genes that enable cancer cells to continually grow and spread.[34,37] In contrast, while modulation of signaling processes has been a successful therapeutic approach to indirectly target gene expression, significant redundancy in signaling pathways can lead to compensatory effects and resistance.[34] Because activation of a transcriptional activator often represents the endpoint of a signaling event, direct modulation of the activator or its binding partners by a small molecule has less potential for compensatory behavior by the cell. Furthermore, many cancers are driven by gene translocations in activators that could be modulated with large therapeutic windows due to the fact that these oncogenic TFs don't exist in normal cells.[49,50] Outside of cancer, activation or inhibition of transcription could also have significant potential in genetic diseases where a disease is caused by transcriptional defects, such as Friedreich's ataxia.[51]

In contrast to their high potential, transcriptional activators are traditionally exceptionally challenging to target with small molecules and are often considered "undruggable." The fact that most activators contain no enzymatic activity and function entirely through PPIs may be a contributor to this outcome. However, in the past 20 years, PPIs in general have begun to yield to medicinal chemistry efforts and by and large are no longer considered "undruggable".[40] Perhaps a more significant contribution to the challenge of targeting transcriptional PPIs is the high fraction of activator PPIs mediated by relatively large IDRs in one or both partners. This leads to PPIs that often are of weak affinity, occur over broad and shallow interfaces, and retain more disorder in the protein complex than "tractable" PPIs.[9,42,43] Together, these issues result in topologically challenging binding interfaces for the development of small molecules.

A key to many successful medicinal chemistry campaigns against PPIs has been the judicious use of structural and mechanistic information of the target protein-protein complex.[36,40,52] The former defines the interface to be inhibited, and can enable the

rational design of molecules that mimic the native binding mode of one of the partners.[52] The latter can help define strategies to enhance success: for example, hotspot analysis can define the regions of the interface that are energetically indispensable for binding, thus enabling more focused design of molecules.[39,53] Together, structural and mechanistic information defines the mechanism of molecular recognition in the target PPI. However, the mechanisms of molecular recognition for the PPIs of transcriptional activators are poorly defined due to the higher level of disorder in these complexes.[42] Thus, it is critical to develop a wholistic understanding of molecular recognition in these critical interactions and develop techniques and approaches to define the mechanisms by which these PPIs form.

1.3 Transcriptional Activators

The "Blueprint" of a Functional Transcriptional Activator

For essentially all eukaryotic transcriptional programs, recruitment of the transcriptional machinery to specific genomic loci is mediated by transcriptional activators.[2] A functional transcriptional activator contains at least two domains: a DNA-binding Domain (DBD) and an interaction domain (ID) that binds to transcriptional coactivators and other transcriptional machinery (Fig. 1.1). In most transcriptional activators, the ID is an IDR typically referred to as a transcriptional activation domain, or TAD.[2,54] TADs are thought to function by binding to activator binding domains (ABDs) in the transcriptional machinery, thereby localizing transcriptional apparatus to promoter and enhancer regions and upregulating transcription at nearby genes.[1] There is also some evidence that TADs play other roles in this process, such as inducing conformational rearrangements in key coactivators to enable the transition of the preinitiation complex into a functional state.[55–58]

4
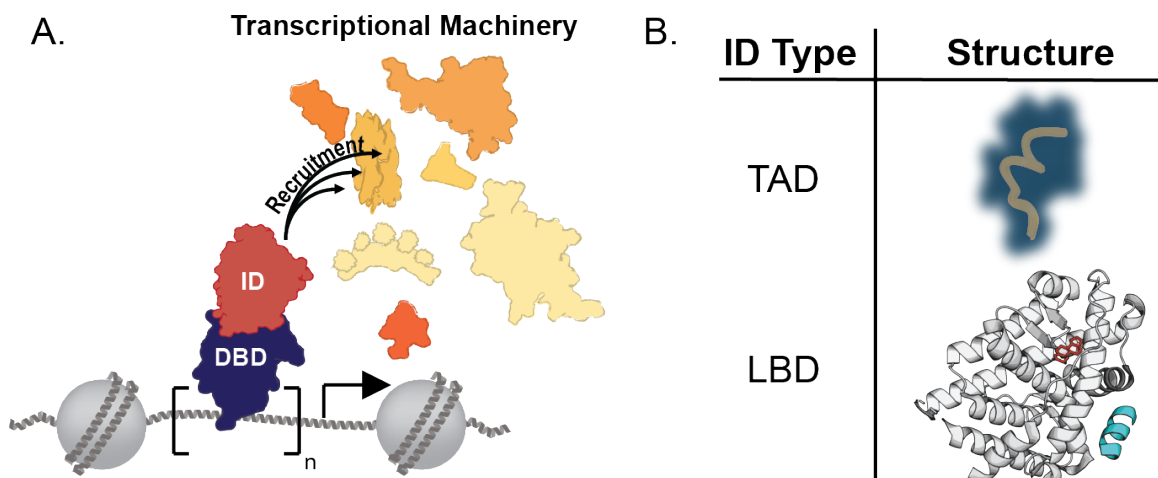
Figure 1.1. Blueprint of a transcriptional activator. A. General mechanism of recruitment by transcriptional activators. B. The two classes of interaction domain and their general structures. For the LBD, the structure of the Androgen Receptor bound to testosterone (red) and a coactivator peptide (cyan) was used (PDH ID 2Q7I).[59]

In contrast, the Nuclear Receptor (NR) class of transcriptional activators contains a folded ID known as a ligand binding domain (LBD).[60,61] NRs are regulated quite distinctly from other transcriptional activators, as the functional state of the LBD is controlled by the direct binding of small molecules or peptide ligands.[62,63] Upon binding to a native ligand, LBDs undergo a conformational switch to expose an allosteric hydrophobic groove, which binds to the disordered LXXLL NR-binding motif in transcriptional coactivators.[62] In the unliganded state, the LBD is typically in a repressed state and does bind to coactivators. Because LBDs of NRs have evolved to bind to small molecules, it is relatively straightforward to develop small molecule modulators of NR function and several NR-targeted drugs have been approved for use in cancer therapies.[61,63,64]

Composite Transcriptional Activators

While in the simplest form of a transcriptional activator, the TAD and DBD are contained in a single peptide sequence, there are several TFs that alone contain only one of the requisite domains to function as an activator.[2] These TFs must bind to cofactors that contain the missing domain to activate transcription, with these PPIs functioning as

"adaptors" to enable function as a transcriptional activator. Interestingly, several examples of these adaptor PPIs are mediated by IDRs. For example, the TEAD family of transcriptional activators contain a DBD but not a TAD.[65] TEAD transcription factors instead use a folded motif to interact with an IDR of the cofactor YAP, which itself contains a TAD.[66,67] Thus, the YAP•TEAD complex is a functional transcriptional activator. This type of architecture is also utilized by viral activators, such as VP16, which use DNA-binding capabilities of endogenous TFs to upregulate transcription of herpes simplex virus immediate-early genes.[68]

Composite activators are often observed in signaling pathways, where one of the partners is localized in the cytoplasm in the absence of a signaling event and is translocated into the nucleus upon activation of the pathway. An example of this is β-catenin in the Wnt signaling pathway.[69] The protein β-catenin is normally bound to axin in the cytoplasm, where it is continuously marked for degradation by phosphorylation by the kinase GSK3, bound to an adjacent site on axin. However, upon activation of the Wnt signaling, β-catenin is excluded from its binding sites on axin and accumulates in the cytoplasm. Upon sufficient accumulation, β-catenin translocates to the nucleus where it interacts with DNA-bound Tcf/LEF transcription factors through an IDR in Tcf.[70] The β-catenin•Tcf complex can thus act as a transcriptional activator and recruit coactivators through the TAD of β-catenin.[71,72]

The Transcriptional Activation Domain

One of the most important classes of functional IDRs in transcription are transcriptional activation domains.[2] TADs are typically relatively short (<30 amino acids) and are separated into distinct classes based on amino acid content: acidic, glutamine-rich, and proline-rich motifs are the most common.[73] Because of their function as assemblers, TADs use the same short sequence to interact with multiple unique members of the transcriptional machinery, and thus have often been characterized as promiscuous binders.[20,26,74–76] However, this is characterization is misleading because while TADs do indeed function through formation of several distinct PPIs, not all TADs function by binding the same partners. Furthermore, several TADs require specific ABDs to function,[77–80] which suggests a level of specificity in the transcriptional activation process

that is typically overlooked. Thus, an open question is how to define functional interactomes of TADs,[81,82] which could enable more judicious selection of therapeutic targets.[43] Furthermore, defining the mechanisms by which TADs are recognized by their partners is a largely open question where only the general details are known.

Activation by Post-translational Modification

Several transcriptional activators function as the endpoint of signaling cascades, and thus often represent molecular "switches" that turn on gene expression in response to a signaling event.[83] Typically the activators that function as molecular switches are latent TFs: they reside outside the nucleus and/or are otherwise prevented from acting as a TF, and upon the signaling event they translocate into the nucleus and activate transcription. These TFs are typically activated by one or more post-translational modifications that induce a change in the binding partners, localization, and/or stability of the activator that enable it to bind to DNA and upregulate transcription. In many cases, the PTMs occur in IDRs of TFs and function by changing the PPIs made by the IDR. Several examples are discussed below.

One of the most common PTMs that TFs undergo is phosphorylation at serine or threonine residues, and phosphorylation of the TADs of transcriptional activators often results in activation by changing the preferred binding partners of the TAD. For example, one of the most characterized TFs is p53, which contains two tandem TADs at its *N*-terminus.[84] Typically, the p53 TAD is bound to the repressor HDM2, which is an E3 ubiquitin ligase that functions to inactivate p53 via degradation by the ubiquitin-proteosomal machinery.[85] Upon phosphorylation at several Ser/Thr residues in the TAD, p53 loses affinity for HDM2 by ~50-fold and gains 2-30-fold for each individual ABD of the coactivator CBP/p300, thereby phosphorylation acts as a functional "rheostat".[30] In contrast, single-site phosphorylation of the TAD of CREB is sufficient to enhance its affinity for CBP/p300 by ~300-fold.[86]

Activating phosphorylation events also need not directly enhance binding to coactivators. For example, the TF STAT3 is phosphorylated at residue Tyr705 in its *C*-terminal TAD, which causes the protein to dimerize through interactions of the phosphotyrosine residue of one monomer with the SH2 domain of the other STAT3

monomer.[87] This STAT3 homodimer then translocates to the nucleus and activates transcription. Interestingly, multisite Ser phosphorylation of the TAD of the family member STAT6 degrades its DNA-binding capabilities and inactivates the transcription factor.[88]

In a rather unique case, the hypoxia inducible factor 1a (HIF1α) is a latent TF that is activated by the *lack* of a unique hydroxylation PTM in one of its IDRs.[89,90] HIF1α is a master regulator of the hypoxic response, and under normoxic conditions it is continually hydroxylated by prolyl hydroxylase enzymes (PHD) and degraded by the ubiquition proteasome pathway. However, upon low-oxygen conditions (hypoxia) the PHD enzymes—which use molecular oxygen as a substrate—are inactivated, leading to buildup of HIF1α. Subsequently HIF1α translocates to the nucleus and activates hypoxic response genes through PPIs with the coactivator CBP/p300.

Other types of PTMs are also known to stimulate transcriptional activators. For example, the activator p53 undergoes multisite acetylation by CBP/p300 and PCAF in its *C*-terminal regulatory domain, which in its unacetylated state inactivates the DBD.[91] Methylation of TF lysine residues is also known, and the mechanisms by which the function is not clear.[92]

1.4 Protein-Protein Interactions of Transcriptional Activators

Transcriptional activators, as mentioned previously, function through assembly of the transcriptional apparatus. To carry out this role, the TADs of transcriptional activators contact several distinct protein classes in the transcriptional machinery, including coactivators, general transcription factors, and chromatin remodeling enzymes (Fig. 1.2).[2] Here, we review the molecular basis for these interactions and how they contribute to the function of a transcriptional activator.
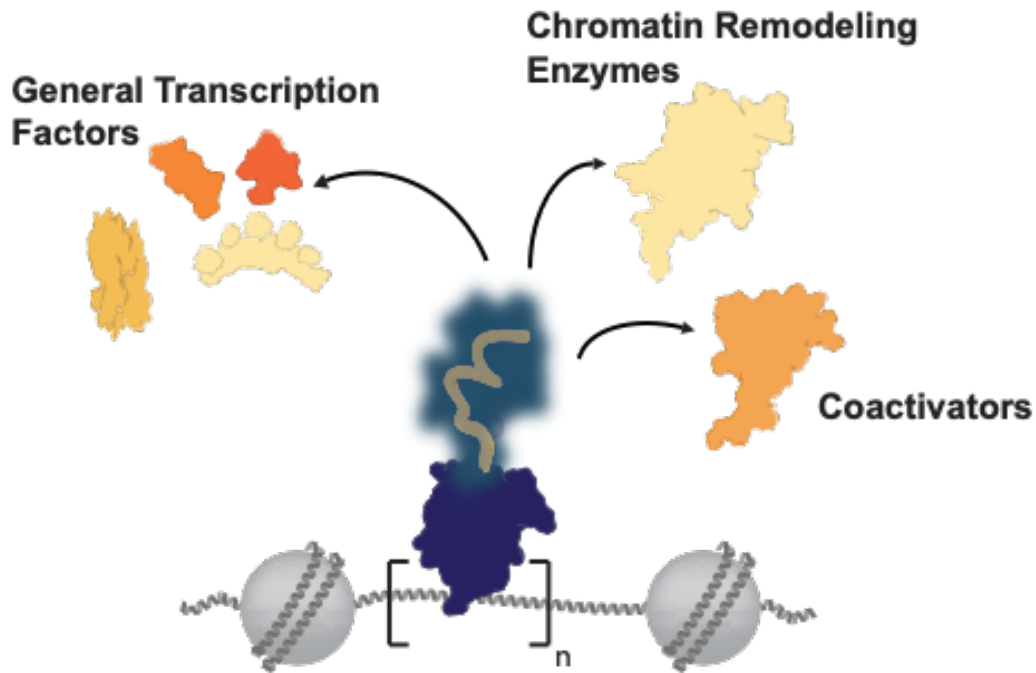
Figure 1.2. Contacts made by transcriptional activation domains with the transcriptional machinery.

Coactivators

One of the most important contacts made by transcriptional activators are with transcriptional coactivators.[2] Coactivators are typically thought to be functional bridges from the activator to the pre-initiation complex,[58] and sometimes have enzymatic activity that acts synergistically alongside this function. Consequently, activator•coactivator interactions are the most well-studied class of activator PPIs from a structural and functional standpoint, especially for the "master" coactivators that are active at almost all genes, such as CBP/p300 and Mediator.[93,94] Coactivators typically contain folded ABDs to interact with the TADs of transcriptional activators, and it is often the case that a single coactivator will contain several structurally and functionally distinct ABDs.[22,95]

The general understanding of activator function has considerably benefited from the study of the coactivator paralogues CBP and p300.[22] From a functional standpoint, CBP/p300 interact with several unique activators and is thus involved in processes as unique as neurodevelopment to cancer.[96–98] To carry out such a large number distinct PPIs with activators, CBP/p300 utilizes four structurally and functionally distinct ABDs that

each recognize unique sets of transcriptional activators.[22] Similarly, several subunits of the tail module of the Mediator complex contains one or more ABDs.[93] Together, this indicates that coactivators function as general signal transducers of activators by presenting several unique surfaces to recognize structurally unique TADs.

Interestingly, the interactions between TADs and coactivators are not just recruitment. For example, several structural and biochemical studies have observed that Mediator changes in conformation upon binding to activators, which in addition to changing its other binding partners is proposed to have a role in enabling RNA Pol II promoter escape.[55–58]

General Transcription Factors

In addition to contacting coactivators, transcription factors are sometimes observed to interact with the general transcription factors (GTFs) that make up the pre-initiation complex (PIC). Specifically, activators have been observed to interact with TFIIA, TFIIB, TFIIF, TFIIH, in addition to TBP and TBP associated factors.[99,100] These interactions are thought to assist the formation of the PIC, and limited evidence appears to tie the "strength" of a transcriptional activator to its affinity for TBP.[101,102] However, in general these interactions are poorly characterized outside of *in vitro* or *in cellulo* crosslinking-based experiments,[81,82,95,103] and thus the general principles underlying the formation of activator•GTF complexes are not understood. In the few cases where the interactions have been structurally characterized,[74,104–107] such as the complex formed between VP16 with the Tfb1 subunit of TFIIH,[105,106] the TAD binds a small folded domain in an analogous manner to binding of TADs to the ABDs of coactivators. Thus, activator•GTF interactions appear to follow the general principles of activator•coactivator interactions.

Chromatin Remodeling Enzymes

The third general class of functional contacts made by transcriptional activators are with chromatin remodeling enzymes.[108] Chromatin remodeling enzymes function by using ATP to drive repackaging of nucleosomes to enable efficient transcriptional activation.[109] However, there are very few molecular details of the contacts formed between transcriptional activators and these enzymes, and until recently it was unclear whether

the interactions were direct or mediated through another partner.[81] Thus, it is currently unknown whether chromatin remodeling enzymes use ABDs to recognize activators in a conserved manner with coactivators and GTFs.

1.5 Molecular Recognition of Transcriptional Activators

The mechanisms by which transcriptional activators are recognized by their binding partners has long been an enigma of molecular recognition.[13,22,26,76] Early biochemical studies of TADs indicated that the sequence of a functional TAD followed no apparent pattern except for the general preponderance of acidic and hydrophobic residues.[11,13] Further biophysical work indicated that TADs were almost completely unstructured in the free state, which was challenging to rationalize at a time where the paradigm was that a protein's function was dictated by its structure.[10,17] Consequently, early recognition models fixated on the disordered TADs functioning through nonspecific interactions with a small subset of cofactors, including the possibility of RNA Pol II itself.[13]

Years later, many aspects of activator function and structure have become more defined. As mentioned earlier in this chapter, it is now known that activators function through mediating PPIs with several distinct factors including coactivators, general transcription factors, and chromatin remodeling enzymes. Furthermore, several structures of TAD•ABD complexes have been solved through solution NMR spectroscopy and X-ray crystallography, giving significant structural insight into the conformational changes TADs undergo upon binding their targets.[23,26,27,106,110,111] However, general principles of activator function and recognition still remain an enigma.

Models of molecular recognition that emphasize that TADs function entirely through nonspecific interactions appear to be invalidated by the discoveries of a) folded and ordered structures of TAD•ABD complexes, and b) TADs that rely on specific contacts with select members of the transcriptional machinery to function.[77–80] However, several recent systems biology studies have indicated that TADs function in a "sequence-independent" manner as long as there is a degree of hydrophobicity and negative charge to the sequence.[112–115] A limited number of biophysical studies have supported this model, finding that TADs can form exceptionally disordered complexes with ABDs that are unresponsive to sequence changes in the TAD.[26,115,116] Here, we review the models

11

of TAD molecular recognition with specific examples to attempt to bring clarity to this apparent controversy.

Coupled Folding and Binding

A common binding mechanism for disordered TADs is folding upon binding to partner proteins, which is referred to as "coupled folding and binding" (Fig. 1.3).[117] This mechanism is an exceptionally important for molecular recognition of IDRs, including TADs, as it enables them to fold into unique structures to bind to partners with unique topology. Put another way, these regions can utilize their structural plasticity to display their functional groups to complement the surface(s) of their binding partners, enabling molecular recognition to occur without necessitating the IDR to adopt a complementary structure before binding. Coupled folding and binding mechanisms enable TADs to bind to structurally unrelated domains in different sections of the transcriptional machinery, such as coactivators and GTFs.
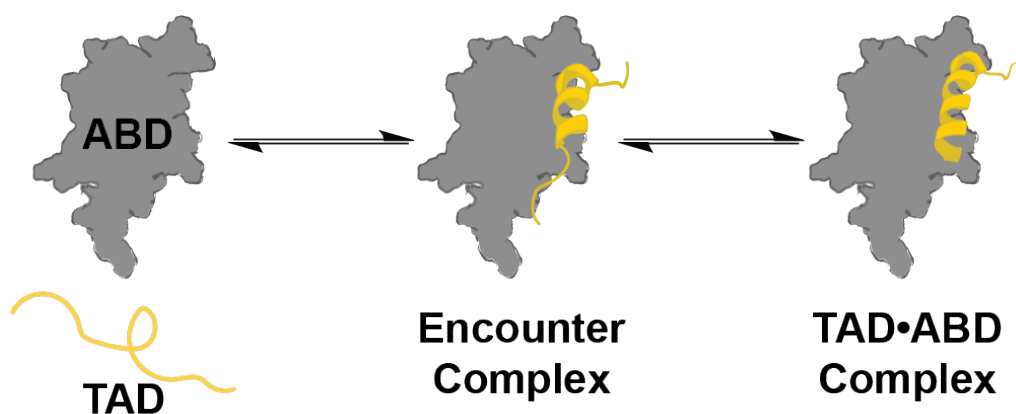
Figure 1.3. General mechanism of coupled folding and binding for TAD•ABD complex formation.

By far the most common structure for TADs to fold into are amphipathic helices, where one surface of the TAD is hydrophobic and interacts directly with the ABD, and the other is hydrophilic and exposed to solvent.[118] Differences between TAD structures are most commonly the number and orientation of the helical regions with respect to each other, and the length and identity of unstructured flanking and linking regions. For example, the p53 TAD forms structurally unique complexes with several distinct

coactivators (Fig. 1.4).[106,111,119] Across these complexes, the p53 TAD forms two helices between residues 19-25 and 47-53, but there are significant differences in the orientations and positioning of these helical regions between the complexes. Thus, disordered regions in TADs can enable repositioning of molecular recognition elements to suit the interface of the binding partner.

p53 TAD(13-61): PLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPD

p53•NCBD   **p53•TAZ1**   **p53•TAZ2**
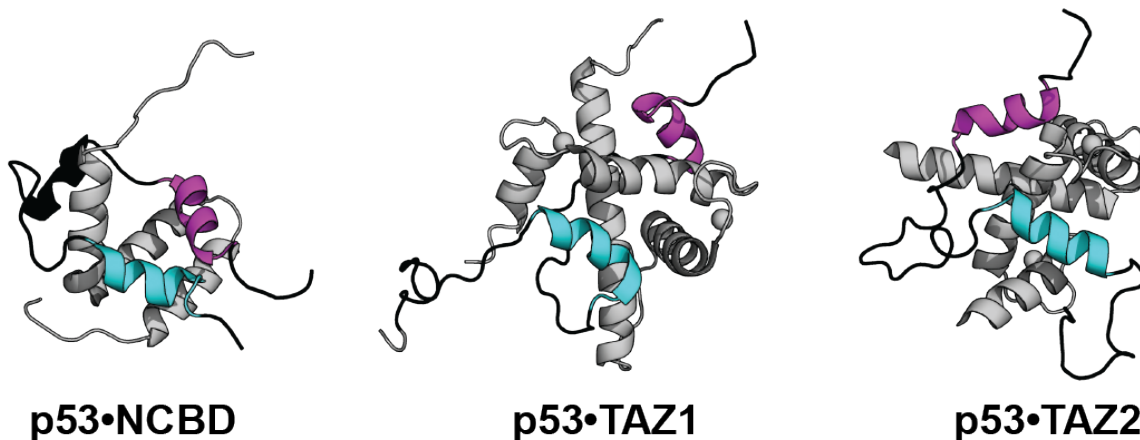
Figure 1.4. The p53 TAD binds to structurally distinct coactivators with unique tertiary conformations. Above, the sequence of the p53 TAD (black), with the two helical motifs colored in cyan and magenta. Below, the complexes formed between p53 and NCBD (PDB 2L14), TAZ1 (PDB 5HOU), and TAZ2 (PDB 5HPD).[111,119]

Contrary to initial reports of TADs lacking *any* structure in the absence of binding partners, there is considerable controversy around the degree to which TADs are folded in the unbound state. For example, a key NMR study of the interaction between the phosphorylated CREB TAD (pKID) and the KIX domain of CBP found that the folding step occurred after the formation of an unstructured "encounter complex".[25] Strikingly, the cMyb TAD—which binds to the same site of KIX—is significantly folded in the unbound state and the degree of structure formation that occurs after the binding step is reduced.[32,120] Thus, there may be a degree of conformational selection in some TAD binding mechanisms. While it is generally unknown how coupled folding and binding mechanisms are related to cellular function, a possible consequence of partial folding *before* binding could be that the TAD's interactome is more restricted due to increased conformational rigidity.

Specific Recognition

The models by which *specific* molecular recognition occurs in TAD•ABD complexes typically focus on how TADs adapt to the surfaces of partner ABD. That is, for a given ABD binding site, a TAD will optimize its tertiary structure and orientation to conform to the restrictions of the ABD surface. The molecular constraints imposed by the ABD include the size and shape of hydrophobic grooves available for the TAD to bury its hydrophobic sidechains, the relative number and positioning of grooves to each other, and the positioning of charged residues outside the grooves.[22] TADs that are unable to conform to the molecular constraints imposed by the ABD are therefore not expected to bind. This model is largely analogous to more traditional PPI interfaces, where binding partners present highly complementary surfaces to facilitate binding. A key difference, however, is that the plasticity of the TAD enables it to bind to *multiple* unrelated ABD surfaces by merely adopting distinct tertiary structures. As discussed in the previous section, the p53 TAD is an excellent example of this phenomenon.

Such a model also predicts that different TADs that bind to the same ABD will form unique tertiary structures. The PPIs formed by the TAZ1 domain of CBP/p300 represent a key example of this phenomenon (Fig. 1.5). Specifically, the TADs of HIF1α,[23] p53,[111] STAT2,[121] CITED2,[122] and NFκB[123] all engage with the TAZ1 domain in highly distinct manners. While each TAD forms multiple helices and wraps around the TAZ1 structure, a single TAD engages with only a subset of the hydrophobic grooves displayed by TAZ1. Thus, individual TADs engage different combinations of hydrophobic grooves, resulting in different engagement modes of the TAZ1 domain. ABDs that have fewer available hydrophobic binding surfaces also display this type of behavior: the TADs of cMyb and CREB interact with the same binding site of the CBP/p300 KIX domain, but form completely unique structures.[110,124]
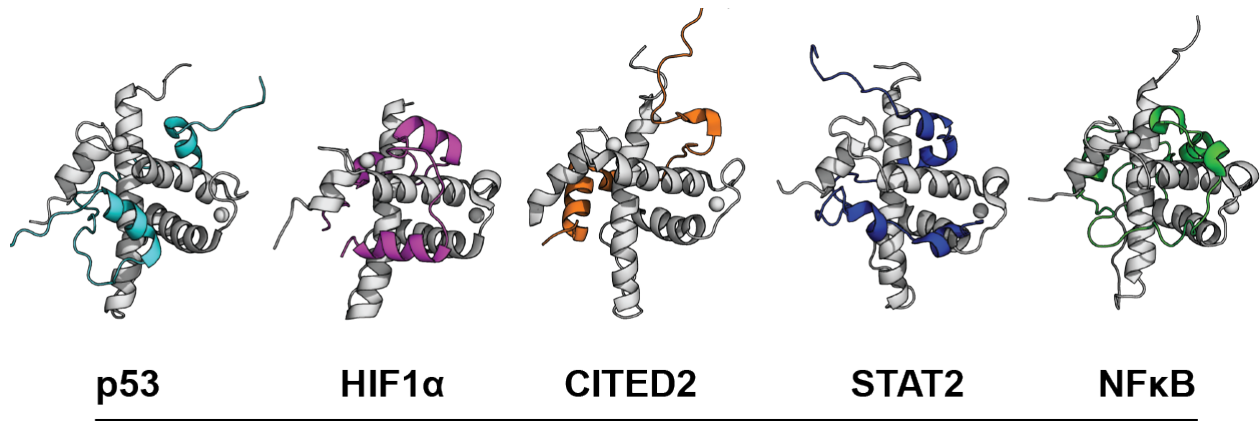
TAZ1 Complex

Figure 1.5. Molecular recognition by activator binding domain TAZ1 (gray). Shown are the structures formed between TAZ1 and p53 (cyan, PDB 5HOU), HIF1α (magenta, PDB 1L8C), CITED2 (orange, 1R8U), STAT2 (blue, 2KA4), and NFκB (green, PDB 2LWW).[23,111,121–123] The TAZ1 domain is shown from the same view in each complex.

Conformational Plasticity in Activator Binding Domains

One aspect of specific TAD•ABD molecular recognition that is often missing or downplayed from recognition models is that ABDs can significantly change in conformation upon binding to a TAD. Further, ABDs have been observed to adopt unique conformations when bound to *different* TADs, suggesting that these domains undergo conformational remodeling to adapt to the TAD. The complexes formed by the NCBD domain of CBP/p300 is a key example of ABD conformational remodeling: it folds into completely unique structures upon binding to IRF3 or ACTR (Fig. 1.6, left).[27,125] Structures of NCBD in complex with other TADs are more similar to the ACTR-bound form (Fig. 1.6, right),[119,126] but differences in the length and orientation of the three NCBD helices demonstrate that the remodeling process also involves fine-tuning of the ABD surface to fit the binding partner.
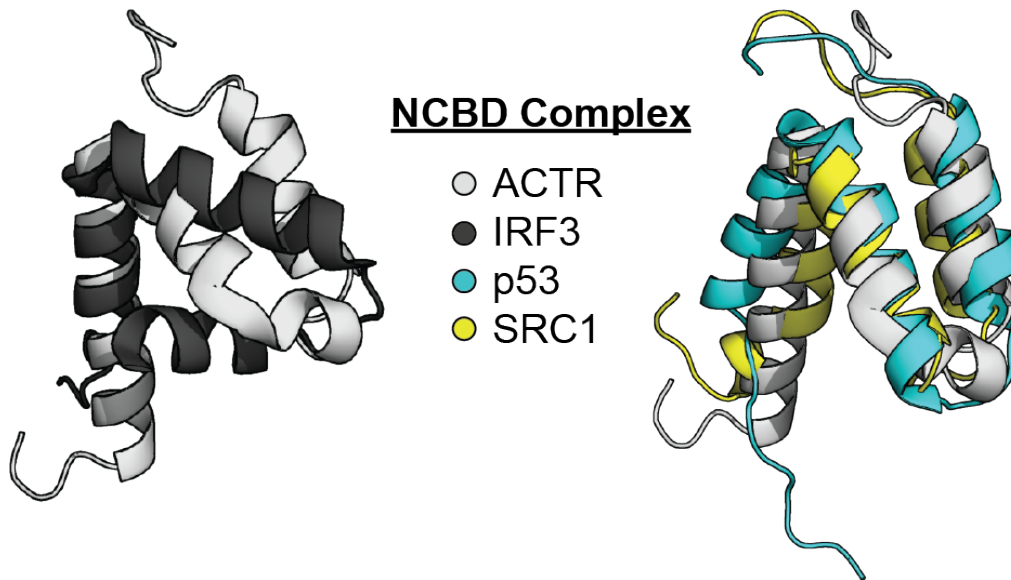
Figure 1.6. NCBD is a paradigm of ABD conformational plasticity. Show are the conformations of NCBD in complex with ACTR (gray, PDB 1KBH), IRF3 (black, PDB IZOQ), p53 (cyan, PDB 2L14), and SRC1 (yellow, PDB 2C52).[27,119,125,126] NCBD adopts unique conformations in complex with IRF3 and ACTR (left), while conformations of the domain in complex with ACTR, p53, and SRC1 are similar (right).

In other ABDs, conformational remodeling upon binding to one or more partners can have significant consequences for function. For example, binding of the MLL TAD to the CBP/p300 KIX domain induces repacking of its hydrophobic core and stabilization of a dynamic loop region in the MLL site, which together stabilize interactions with the TADs of cMyb and CREB at the allosteric binding site by ~1.5-fold.[24,28,29,127,128] Different TADs binding to each site each remodel the KIX structure in unique ways, and ternary complexes thus have cooperativity factors ranging from 1.4-18 (Fig. 1.7).[127] Thus, binding of one TAD to KIX significantly influences the binding partners that can bind at the other site, essentially rebiasing the interactome of CBP/p300.

On the other hand, ABDs can use conformational remodeling to enable functional *negative* allostery in biological feedback loops. For example, CITED2 is a negative regulator of the hypoxic response, and functions by competing for binding with the HIF1α TAD to the CBP/p300 TAZ1 domain.[129,130] However, due to the uncharacteristically tight interaction (K_d ~10 nM) of HIF1α with TAZ1, a simple competitive inhibition mechanism would likely be too inefficient for the rapid temporal response required to regulate this pathway. Instead, CITED2 binds to an allosteric site on TAZ1 and enhances the

dissociation rate of HIF1α by over an order of magnitude through a "forcing" mechanism[130] that is enabled by a conformational change in TAZ1 that destabilizes binding of the HIF1α TAD. Thus, together the available data strongly indicates that plasticity of ABDs can play a functional role that needs to be considered when assessing recognition mechanisms of TADs.

Ternary Complex	Cooperativity
○ KIX•MLL•pKID	3.1
● KIX•MLL•cMyb	1.5
○ KIX•HBZ•cMyb	18.7
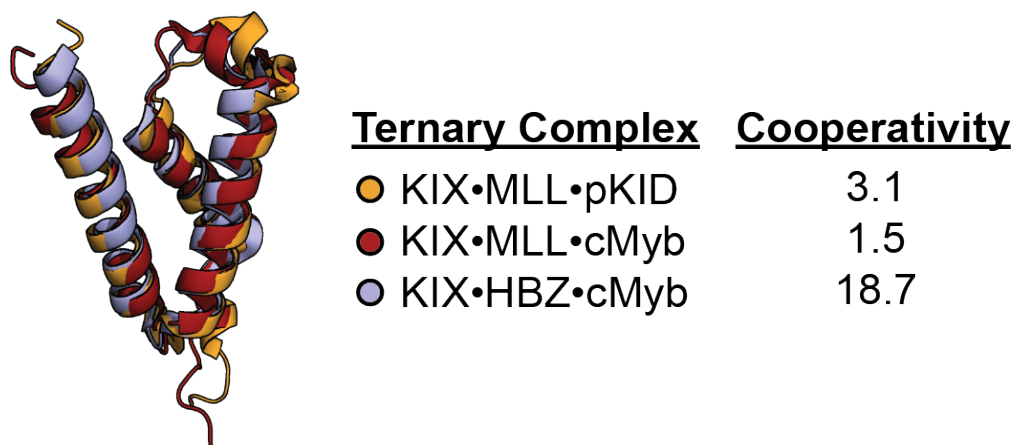
Figure 1.7. Conformational remodeling of KIX domain results in significant differences in cooperativity of ternary complex formation. Left, overlay of KIX conformations from ternary complexes with MLL/pKID (orange, PDB 2LXT), MLL/cMyb (dark red, PDB 2AGH), and HBZ/cMyb (light purple, PDB 6DMX).[24,29,131] Right, cooperativity factors for formation of ternary complexes.[127]

Nonspecific Recognition

In direct contrast to *specific* recognition TAD•ABD complexes are models of *nonspecific* recognition. These models were initially proposed in the face of perplexing functional mutagenesis studies of TADs that demonstrated that they were surprisingly resistant to structural and sequence changes.[11–13,16] However, these models were largely abandoned after structural evidence that showed many of the "acid blobs and negative noodles", such as VP16,[106] form ordered structures with specific ABDs of the transcriptional machinery.

However, not all TADs form ordered complexes with ABDs, causing many to be recalcitrant to structural determination. In the limited cases where structural information of these highly dynamic complexes was obtained, it has been found that the TADs bind to the ABD in multiple distinct orientations or in multiple binding sites.[26,132,133] Further, as

high-throughput technology has evolved, more quantitative assessments of the relationship between TAD sequence and function have been undertaken.[112–114] The results, by and large, have recapitulated early experiments[11] that demonstrated that functional TADs have a higher propensity for hydrophobic and acidic amino acids and a lower propensity for basic amino acids, but with no specific recognition motifs. This has led to nonspecific recognition models to regain traction because they can be used to rationalize these biochemical and biophysical results that appear to violate specific molecular recognition principles.
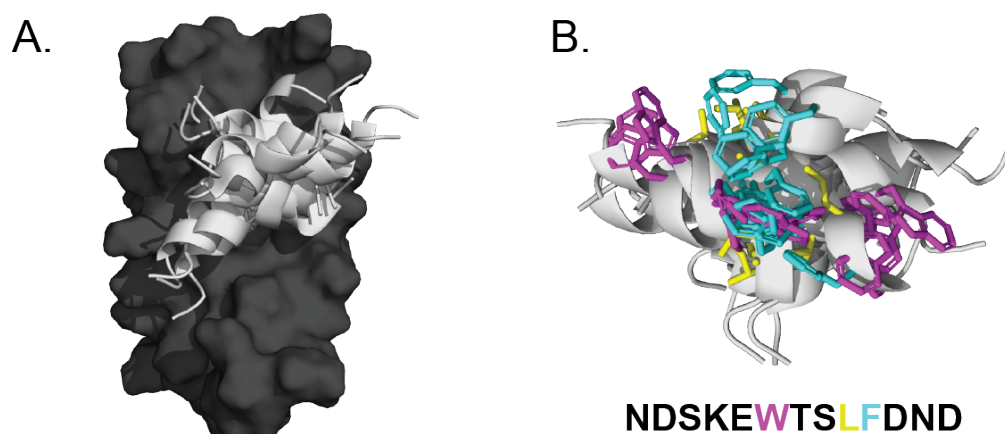
Figure 1.8. The Gcn4•ABD1 complex is highly dynamic. Left, structural model derived from NMR data.[26] The disordered flanking regions of the Gcn4 TAD were omitted for clarity. Right, view of Gcn TAD with hydrophobic sidechains highlighted to show the lack of orientational preference. Shown below is the primary sequence of the Gcn4 region shown.

A key biophysical model system that provides support for this model is the complex formed between the activator Gcn4 and the coactivator Med15.[26,115,116] Initial structural experiments, relying on NMR spin-labeling techniques, of the interaction between the *C*-terminal Gcn4 TAD and the ABD1 domain of Med15 indicated that Gcn4 binds without the formation of a specific orientation (Fig. 1.8).[26] Further studies indicated that a disordered binding mode of the Gcn4 TAD is conserved with the structurally similar ABD2 domain of Med15, and that the complex formed between full-length Med15 and Gcn4 was a disordered "free-for-all" where the Gcn4 TAD forms highly disordered interactions with all available Med15 ABDs.[116] In addition, functional mutagenesis experiments of the Gcn4 TAD demonstrated that merely increasing the degree of hydrophobicity surrounding the

key hydrophobic residues (Fig. 1.8B, colored residues) drastically increased TAD activity in addition to affinity for Med15.[115] Together, these results led to the proposal that the Gcn4 TAD functions as a "hydrophobic cloud" that binds indiscriminately to the shallow hydrophobic surfaces of the Med15 ABDs to bury its exposed hydrophobic residues.[116]

Interestingly, molecular dynamics studies of the Gcn4 TAD in complex with ABD1 appear to indicate a more restrained picture of binding.[134] In these simulations, the three critical hydrophobic groups of the Gcn4 TAD (Fig. 1.8B, colored residues) compete for three major pockets in the binding interface of ABD1, resulting in just three major orientations. Thus, it is currently controversial whether the Gcn4•Med15 interaction is truly a paradigm of nonspecific activator•coactivator complexes, or if there are underlying elements of specific complex formation that are challenging to detect with experimental biophysical approaches.

1.6 Dissertation Summary

The goal of this dissertation was to define mechanisms by which the coactivator Med25 recognizes its TAD binding partners. Med25 contains an ABD fold unique to all other known coactivators,[135,136] and initial structural and biophysical work has indicated that its complexes with activators are quite dynamic.[137] Thus, it serves as a useful model system to examine longstanding molecular recognition models of activator•coactivator interactions.

In chapter two, we use structural and biophysical techniques to investigate how Med25 recognizes diverse activator sequences. Particularly, we examined in the role of conformational changes in the ABD in recognizing disparate sequences, and whether conformational mobility of the ABD could be taken advantage of by small molecules.

In chapter three, we use the PPIs between Med25 and the PEA3 family of activators as a model system of exceptionally dynamic—or "fuzzy"—transcriptional PPIs to ask whether the dynamics of the activator•coactivator complex is sensitive to small sequence changes in the activator. This is a direct test of the "nonspecific" model of activator molecular recognition, which would predict that small changes in activator sequence would introduce no detectable differences in the mechanism of complex formation.

1.7 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. Vernimmen, D. & Bickmore, W. A. The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends Genet.* **31**, 696–708 (2015).
4. Krasnov, A. N., Mazina, M. Yu., Nikolenko, J. V. & Vorobyeva, N. E. On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell Biosci.* **6**, 15 (2016).
5. Roeder, R. G. The role of general initiation factors in transcription by RNA polymerase II. 9 (1996).
6. Orphanides, G., Lagrange, T. & Reinberg, D. The general transcription factors of RNA polymerase II. *Genes Dev.* **10**, 2657–2683 (1996).
7. Thomas, M. C. & Chiang, C.-M. The General Transcription Machinery and General Cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178 (2006).
8. Roeder, R. G. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791 (2019).
9. Liu, J. *et al.* Intrinsic Disorder in Transcription Factors [†]. *Biochemistry* **45**, 6873–6888 (2006).
10. Dyson, H. J. & Wright, P. E. Perspective: the essential role of NMR in the discovery and characterization of intrinsically disordered proteins. *J. Biomol. NMR* **73**, 651–659 (2019).
11. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113–119 (1987).
12. Hope, I. A., Mahadevan, S. & Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333**, 635–640 (1988).
13. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).
14. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. Gal4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563–564 (1988).
15. 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.* **90**, 883–887 (1993).
16. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Amp Dev.* **2**, 718–729 (1988).
17. Triezenberg, S. J. Critical Amino Acids in the Transcriptional Activation Domain of the Herpesvirus Protein VP16 Are Solvent-exposed in Highly Mobile Protein Segments. *J. Biol. Chem.* **271**, 4819–4826 (1996).
18. Kakidani, H. & Ptashne, M. Gal4 Activates Gene Expression of Mammalian Cells. *Cell* **52**, 161–167 (1988).
19. Hope, I. A. & Struhl, K. Functional Dissection of a Eukaryotic Transcriptional Activator Protein, GCN4 of Yeast. *Cell* **46**, 885–894 (1986).
20. Tompa, P., Szász, C. & Buday, L. Structural disorder throws new light on moonlighting. *Trends Biochem. Sci.* **30**, 484–489 (2005).

21.   van der Lee, R. *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

22.   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).

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

24.   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).

25.   Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

26.   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).

27.   Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).

28.   Brüschweiler, S. *et al.* Direct Observation of the Dynamic Process Underlying Allosteric Signal Transmission. *J. Am. Chem. Soc.* **131**, 3063–3068 (2009).

29.   Brüschweiler, S., Konrat, R. & Tollinger, M. Allosteric Communication in the KIX Domain Proceeds through Dynamic Repacking of the Hydrophobic Core. *ACS Chem. Biol.* **8**, 1600–1610 (2013).

30.   Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M. & Wright, P. E. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci.* **107**, 19290–19295 (2010).

31.   Dogan, J., Schmidt, T., Mu, X., Engström, Å. & Jemth, P. Fast Association and Slow Transitions in the Interaction between Two Intrinsically Disordered Protein Domains. *J. Biol. Chem.* **287**, 34316–34324 (2012).

32.   Giri, R., Morrone, A., Toto, A., Brunori, M. & Gianni, S. Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc. Natl. Acad. Sci.* **110**, 14942–14947 (2013).

33.   Wands, A. M. *et al.* Transient-state Kinetic Analysis of Transcriptional Activator·DNA Complexes Interacting with a Key Coactivator. *J. Biol. Chem.* **286**, 16238–16245 (2011).

34.   Darnell, J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749 (2002).

35.   Koehler, A. N. A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.* **14**, 331–340 (2010).

36.   Bushweller, J. H. Targeting transcription factors in cancer — from undruggable to reality. *Nat. Rev. Cancer* **19**, 611–624 (2019).

37.   Miller, D. M., Thomas, S. D., Islam, A., Muench, D. & Sedoris, K. c-Myc and Cancer Metabolism. *Clin. Cancer Res.* **18**, 5546–5553 (2012).

38.   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. *Proc. Natl. Acad. Sci.* **110**, E3506–E3515 (2013).

39.   Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).

40.   Arkin, M. R., Tang, Y. & Wells, J. A. Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing toward the Reality. *Chem. Biol.* **21**, 1102–1114 (2014).

41.   Delaforge, E. *et al.* Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *J. Am. Chem. Soc.* **140**, 1148–1158 (2018).

42.   Mapp, A. K. & Ansari, A. Z. A TAD Further: Exogenous Control of Gene Activation. *ACS Chem. Biol.* **2**, 62–75 (2007).

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

44.   Guo, X., Bulyk, M. L. & Hartemink, A. J. INTRINSIC DISORDER WITHIN AND FLANKING THE DNA-BINDING DOMAINS OF HUMAN TRANSCRIPTION FACTORS. in *Biocomputing 2012* 104–115 (WORLD SCIENTIFIC, 2011). doi:10.1142/9789814366496_0011.

45.   Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* **32**, 42–56 (2016).

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

47.   Warren, C. *et al.* Dynamic intramolecular regulation of the histone chaperone nucleoplasmin controls histone binding and release. *Nat. Commun.* **8**, 2215 (2017).

48.   Warren, C. & Shechter, D. Fly Fishing for Histones: Catch and Release by Histone Chaperone Intrinsically Disordered Regions and Acidic Stretches. *J. Mol. Biol.* **429**, 2401–2426 (2017).

49.   Nambiar, M., Kari, V. & Raghavan, S. C. Chromosomal translocations in cancer. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1786**, 139–152 (2008).

50.   Illendula, A. *et al.* A small-molecule inhibitor of the aberrant transcription factor CBFβ-SMMHC delays leukemia in mice. *Science* **347**, 779–784 (2015).

51.   Erwin, G. S. *et al.* Synthetic transcription elongation factors license transcription across repressive chromatin. *Science* **358**, 1617–1622 (2017).

52.   Scott, D. E., Bayly, A. R., Abell, C. & Skidmore, J. Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nat. Rev. Drug Discov.* **15**, 533–550 (2016).

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

54.   Triezenberg, S. Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5**, 190–196 (1995).

55.   Taatjes, D. J. Structure, Function, and Activator-Induced Conformations of the CRSP Coactivator. *Science* **295**, 1058–1062 (2002).

56.   Meyer, K. D., Lin, S., Bernecky, C., Gao, Y. & Taatjes, D. J. p53 activates transcription by directing structural shifts in Mediator. *Nat. Struct. Mol. Biol.* **17**, 753–760 (2010).

57.   Ebmeier, C. C. & Taatjes, D. J. Activator-Mediator binding regulates Mediator-cofactor interactions. *Proc. Natl. Acad. Sci.* **107**, 11283–11288 (2010).

58.   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).

59.   Askew, E. B., Gampe, R. T., Stanley, T. B., Faggart, J. L. & Wilson, E. M. Modulation of Androgen Receptor Activation Function 2 by Testosterone and Dihydrotestosterone. *J. Biol. Chem.* **282**, 25801–25816 (2007).

60.   Weatherman, R. V., Fletterick, R. J. & Scanlan, T. S. Nuclear-Receptor Ligands and Ligand-Binding Domains. *Annu. Rev. Biochem.* **68**, 559–581 (1999).

61.   McEwan, I. J. The Nuclear Receptor Superfamily at Thirty. in *The Nuclear Receptor Superfamily* (ed. McEwan, PhD, I. J.) vol. 1443 3–9 (Springer New York, 2016).

62.   Rastinejad, F., Huang, P., Chandra, V. & Khorasanizadeh, S. Understanding nuclear receptor form and function using structural biology. *J. Mol. Endocrinol.* **51**, T1–T21 (2013).

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

64.   Kojetin, D. J. & Burris, T. P. Small Molecule Modulation of Nuclear Receptor Conformational Dynamics: Implications for Function and Drug Discovery. *Mol. Pharmacol.* **83**, 1–8 (2013).

65.   Jeffrey Holden & Christian Cunningham. Targeting the Hippo Pathway and Cancer through the TEAD Family of Transcription Factors. *Cancers* **10**, 81 (2018).

66.   Vassilev, A. TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm. *Genes Dev.* **15**, 1229–1241 (2001).

67.   Li, Z. *et al.* Structural insights into the YAP and TEAD complex. *Genes Dev.* **24**, 235–240 (2010).

68.   Babb, R., Huang, C. C., Aufiero, D. J. & Herr, W. DNA Recognition by the Herpes Simplex Virus Transactivator VP16: a Novel DNA-Binding Structure. *Mol. Cell. Biol.* **21**, 4700–4712 (2001).

69.   MacDonald, B. T., Tamai, K. & He, X. Wnt/β-Catenin Signaling: Components, Mechanisms, and Diseases. *Dev. Cell* **17**, 9–26 (2009).

70.   Graham, T. A., Weaver, C., Mao, F., Kimelman, D. & Xu, W. Crystal Structure of a ᴺ𝖫-Catenin/Tcf Complex. 12.

71.   Vleminckx, K., Kemler, R. & Hecht, A. The C-terminal transactivation domain of b-catenin is necessary and sufficient for signaling by the LEF-1/b-catenin complex in Xenopus laevis. *Mech. Dev.* 10 (1999).

72.   Sun, Y. *et al.* Regulation of ᴺ𝖫-catenin transformation by the p300 transcriptional coactivator. *CELL Biol.* 6.

73.   Mitchell, P. J. & Tjian, R. Transcriptional Regulation in Mamnalian Cells by Sequence-Specific DNA Binding Proteins. *Science* **245**, 371–378 (1989).

74.   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).

75.   Hazy, E. & Tompa, P. Limitations of Induced Folding in Molecular Recognition by Intrinsically Disordered Proteins. *ChemPhysChem* **10**, 1415–1419 (2009).

76.   Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).

77.   Ruas, J. L., Poellinger, L. & Pereira, T. Functional Analysis of Hypoxia-inducible Factor-1α-mediated Transactivation: IDENTIFICATION OF AMINO ACID RESIDUES CRITICAL FOR TRANSCRIPTIONAL ACTIVATION AND/OR INTERACTION WITH CREB-BINDING PROTEIN. *J. Biol. Chem.* **277**, 38723–38730 (2002).

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

79.   Chrivia, J. C. *et al.* Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855–859 (1993).

80.   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).

81.   Krishnamurthy, M. *et al.* Caught in the Act: Covalent Cross-Linking Captures Activator–Coactivator Interactions *in Vivo*. *ACS Chem. Biol.* **6**, 1321–1326 (2011).

82.   Dugan, A. *et al.* Discovery of Enzymatic Targets of Transcriptional Activators via *in Vivo* Covalent Chemical Capture. *J. Am. Chem. Soc.* **138**, 12629–12635 (2016).

83.   Brivanlou, A. H. Signal Transduction and the Control of Gene Expression. *Science* **295**, 813–818 (2002).

84.   Raj, N. & Attardi, L. D. The Transactivation Domains of the p53 Protein. *Cold Spring Harb. Perspect. Med.* **7**, a026047 (2017).

85.   Sullivan, K. D., Galbraith, M. D., Andrysik, Z. & Espinosa, J. M. Mechanisms of transcriptional regulation by p53. *Cell Death Differ.* **25**, 133–143 (2018).

86.   Dahal, L., Shammas, S. L. & Clarke, J. Phosphorylation of the IDP KID Modulates Affinity for KIX by Increasing the Lifetime of the Complex. *Biophys. J.* **113**, 2706–2712 (2017).

87.   Sgrignani, J. *et al.* Structural Biology of STAT3 and Its Implications for Anticancer Therapies Development. *Int. J. Mol. Sci.* **19**, 1591 (2018).

88.   Maiti, N. R., Sharma, P., Harbor, P. C. & Haque, S. J. Serine Phosphorylation of Stat6 Negatively Controls Its DNA-Binding Function. *J. Interferon Cytokine Res.* **25**, 553–563 (2005).

89.   Lee, J.-W., Bae, S.-H., Jeong, J.-W., Kim, S.-H. & Kim, K.-W. Hypoxia-inducible factor (HIF-1)α: its protein stability and biological functions. *Exp. Mol. Med.* **36**, 1–12 (2004).

90.   Chowdhury, R. *et al.* Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases. *Nat. Commun.* **7**, 12673 (2016).

91.   Reed, S. & Quelle, D. p53 Acetylation: Regulation and Consequences. *Cancers* **7**, 30–69 (2014).

92.   Han, D. *et al.* Lysine methylation of transcription factors in cancer. *Cell Death Dis.* **10**, 290 (2019).

93.   Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).

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

95.   Herbig, E. *et al.* Mechanism of Mediator Recruitment by Tandem Gcn4 Activation Domains and Three Gal11 Activator-Binding Domains. *Mol. Cell. Biol.* **30**, 2376–2390 (2010).

96.   Valor, L., Viosca, J., Lopez-Atalaya, J. & Barco, A. Lysine Acetyltransferases CBP and p300 as Therapeutic Targets in Cognitive and Neurodegenerative Disorders. *Curr. Pharm. Des.* **19**, 5051–5064 (2013).

97.   Santer, F. R. *et al.* Inhibition of the Acetyltransferases p300 and CBP Reveals a Targetable Function for p300 in the Survival and Invasion Pathways of Prostate Cancer Cell Lines. *Mol. Cancer Ther.* **10**, 1644–1655 (2011).

98.   Lasko, L. M. *et al.* Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. *Nature* **550**, 128–132 (2017).

99.   Green, M. R. Eukaryotic Transcription Activation: Right on Target. *Mol. Cell* **18**, 399–402 (2005).

100.   Stargell, L. & Struhl, K. Mechanisms of transcriptional activation in vivo: two steps forward. *Trends Genet.* **12**, 311–315 (1996).

101.   Chang, J., Kim, D.-H., Lee, S. W., Choi, K. Y. & Sung, Y. C. Transactivation Ability of p53 Transcriptional Activation Domain Is Directly Related to the Binding Affinity to TATA-binding Protein. *J. Biol. Chem.* **270**, 25014–25019 (1995).

102.   Li, X.-Y., Virbasius, A., Zhu, X. & Green, M. R. Enhancement of TBP binding by activators and general transcription factors. *Nature* **399**, 605–609 (1999).

103.   Hall, D. B. & Struhl, K. The VP16 Activation Domain Interacts with Multiple Transcriptional Components as Determined by Protein-Protein Cross-linking *in Vivo*. *J. Biol. Chem.* **277**, 46043–46050 (2002).

104.   Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced α Helix in the VP16 Activation Domain upon Binding to a Human TAF. *Science* **277**, 1310–1313 (1997).

105.   Di Lello, P. *et al.* NMR Structure of the Amino-Terminal Domain from the Tfb1 Subunit of TFIIH and Characterization of Its Phosphoinositide and VP16 Binding Sites [†,‡]. *Biochemistry* **44**, 7678–7686 (2005).

106.   Langlois, C. *et al.* NMR Structure of the Complex between the Tfb1 Subunit of TFIIH and the Activation Domain of VP16: Structural Similarities between VP16 and p53. *J. Am. Chem. Soc.* **130**, 10596–10604 (2008).

107.   Wei, Y. *et al.* Multiple direct interactions of TBP with the MYC oncoprotein. *Nat. Struct. Mol. Biol.* **26**, 1035–1043 (2019).

108.   Stafford, G. A. & Morse, R. H. Chromatin Remodeling by Transcriptional Activation Domains in a Yeast Episome. *J. Biol. Chem.* **272**, 11526–11534 (1997).

109.   Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. *Cell* **154**, 490–503 (2013).

110.   Radhakrishnan, I. *et al.* Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell* **91**, 741–752 (1997).

111.   Krois, A. S., Ferreon, J. C., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc. Natl. Acad. Sci.* **113**, E1853–E1862 (2016).

112.   Ravarani, C. N. *et al.* High-throughput discovery of functional disordered regions: investigation of transactivation domains. *Mol. Syst. Biol.* **14**, (2018).

113.   Arnold, C. D. *et al.* A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J.* **37**, (2018).

114.   Staller, M. V. *et al.* A High-Throughput Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. *Cell Syst.* **6**, 444-455.e6 (2018).

115.   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.* **111**, E3506–E3513 (2014).

116.   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).

117.   Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).

118.   Piskacek, S. *et al.* Nine-amino-acid transactivation domain: Establishment and prediction utilities. *Genomics* **89**, 756–768 (2007).

119.   Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structure of the p53 Transactivation Domain in Complex with the Nuclear Receptor Coactivator Binding Domain of CREB Binding Protein. *Biochemistry* **49**, 9964–9971 (2010).

120.   Arai, M., Sugase, K., Dyson, H. J. & Wright, P. E. Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc. Natl. Acad. Sci.* **112**, 9614–9619 (2015).

121.   Wojciak, J. M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J.* **28**, 948–958 (2009).

122.   De Guzman, R. N., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Interaction of the TAZ1 Domain of the CREB-Binding Protein with the Activation Domain of CITED2: REGULATION BY COMPETITION BETWEEN INTRINSICALLY UNSTRUCTURED LIGANDS FOR NON-IDENTICAL BINDING SITES. *J. Biol. Chem.* **279**, 3042–3049 (2004).

123.   Mukherjee, S. P. *et al.* Analysis of the RelA:CBP/p300 Interaction Reveals Its Involvement in NF-κB-Driven Transcription. *PLoS Biol.* **11**, e1001647 (2013).

124.   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).

125.   Qin, B. Y. *et al.* Crystal Structure of IRF-3 in Complex with CBP. *Structure* **13**, 1269–1277 (2005).

126.   Waters, L. *et al.* Structural Diversity in p160/CREB-binding Protein Coactivator Complexes. *J. Biol. Chem.* **281**, 14787–14795 (2006).

127.   Shammas, S. L., Travis, A. J. & Clarke, J. Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc. Natl. Acad. Sci.* **111**, 12055–12060 (2014).

128.   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.* **111**, 12061–12066 (2014).

129.   Bhattacharya, S. *et al.* Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF-1. *Genes Dev.* **13**, 64–75 (1999).

130. 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).

131. Yang, K. *et al.* Structural basis for cooperative regulation of KIX-mediated transcription pathways by the HTLV-1 HBZ activation domain. *Proc. Natl. Acad. Sci.* **115**, 10040–10045 (2018).

132. 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.* **109**, 6078–6083 (2012).

133. Arai, M., Ferreon, J. C. & Wright, P. E. Quantitative Analysis of Multisite Protein–Ligand Interactions by NMR: Binding of Intrinsically Disordered p53 Transactivation Subdomains with the TAZ2 Domain of CBP. *J. Am. Chem. Soc.* **134**, 3792–3803 (2012).

134. Scholes, N. S. & Weinzierl, R. O. J. Molecular Dynamics of 'Fuzzy' Transcriptional Activator-Coactivator Interactions. *PLOS Comput. Biol.* **12**, e1004935 (2016).

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

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

137. 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).

CHAPTER 2.

Conservation of Coactivator Engagement Mechanism Enables Small-Molecule Allosteric Modulators[1]

2.1 Abstract

Transcriptional coactivators are a molecular recognition marvel because a single domain within these proteins, the activator binding domain or ABD, interacts with multiple compositionally diverse transcriptional activators. Also remarkable is the structural diversity among ABDs, which range from conformationally dynamic helical motifs to those with a stable core such as a β-barrel. A significant objective is to define conserved properties of ABDs that allow them to interact with disparate activator sequences. The ABD of the coactivator Med25 (activator interaction domain or AcID) is unique in that it contains secondary structural elements that are on both ends of the spectrum: helices and loops that display significant conformational mobility and a seven-stranded β-barrel core that is structurally rigid. Using biophysical approaches, we build a mechanistic model of how AcID forms binary and ternary complexes with three distinct activators; despite its static core, Med25 forms short-lived, conformationally mobile, and structurally distinct complexes with each of the cognate partners. Further, ternary complex formation is facilitated by allosteric communication between binding surfaces on opposing faces of the β-barrel. The model emerging suggests that the conformational shifts and cooperative

---

binding is mediated by a flexible substructure comprised of two dynamic helices and flanking loops, indicating a conserved mechanistic model of activator engagement across ABDs. Targeting a region of this substructure with a small-molecule covalent cochaperone modulates ternary complex formation. Our data support a general strategy for the identification of allosteric small-molecule modulators of ABDs, which are key targets for mechanistic studies as well as therapeutic applications.

2.2 Introduction

Biophysical studies of complexes formed between the activator binding domains (ABDs) of transcriptional coactivators and their cognate activator binding partners suggest that modulating these functionally critical protein–protein interactions (PPIs) with small molecules is a formidable task.[1] An excellent example of this is the ABD of the Mediator protein Med25, termed AcID (activator interaction domain; Fig. 1A).[2–4] As is standard for ABDs, AcID is a binding partner of a diverse array of transcriptional activators, including VP16,[2–4] ATF6α,[5] and the ETV/PEA3 activators.[6–8] Through these interactions, Med25 plays significant roles in the unfolded protein response and in oncogenesis, generating significant interest in small molecule modulators. However, data from NMR studies of AcID in complex with VP16 and ETV/PEA3 activators suggest that modulating these PPIs would not be trivial.[2–4,7] The VP16 transcriptional activation domain contacts a surface of ~1,800 Å^2 of AcID, wrapping around the topologically challenging β-barrel while also contacting two flanking helices. The transcriptional activation domain of the ETV/PEA3 member ERM (ETV5) interacts with one face of the β-barrel, a binding surface referred to as H1 that is ~900 Å^2 in area.[6,7] The β-barrel core of AcID is unusual among ABDs, with helices more commonly observed, and raises the question of the role that the barrel might play in the molecular recognition of activators relative to the other substructures within AcID.

The observation that a portion of VP16 and ERM utilize the same H1 binding surface in AcID despite their distinct sequences suggests that conformational plasticity within the ABD could play a role in its molecular recognition capabilities and, ultimately, function. We considered the following criteria essential to support this model. First, each activator AcID complex should be conformationally labile, with two or more conformation states

29

energetically accessible. Second, two functionally similar binding sites such as H1 and H2 should be allosterically connected in a conformationally plastic domain. To test this model, we first identified ATF6α as an H2 binding site-specific ligand for AcID. Transient kinetic experiments with the activators VP16, ERM, and ATF6α revealed that in each case AcID exploits conformational lability to recognize the three distinct sequences in binary complexes as well as in ternary complexes. Molecular dynamics simulations highlighted the critical role that the flexible loops and helices play in the remodeling of one PPI surface, while also suggesting how these motions relate to the larger family of ABDs. Consistent with this model, targeting one of the most dynamic regions of AcID with a small molecule cochaperone recapitulates the kinetic signatures of the native transcriptional activators. Taken together, these data suggest a conserved mechanism for transcriptional activators despite considerable structural divergence and that targeting the most dynamic regions of ABDs is a likely path forward for small molecule regulation of transcription through this important class of proteins.

2.3 Results and Discussion

ATF6α and ERM Bind to Opposite Faces of Med25 AcID

Separate NMR studies of AcID in complex with the transcriptional activation domains of VP16 and ERM suggest that the two activators both contact the H1 binding surface, with the significantly larger VP16 also interacting with the H2 surface.[2–4,7] While several lines of evidence indicate that ATF6α interacts with Med25 AcID as part of its function,[5] the binding site within the protein has not been established. We first measured the dissociation constants for each of the activators by fluorescence anisotropy experiments using fluorescein-tagged variants of VP16 (438–490), ERM (38–68), and ATF6α (40–66), and this revealed that ERM and ATF6α interact with comparable affinities (Fig. 2.1B). To provide a direct comparison of the binding modes of the three activators and identify the binding site of ATF6α, we measured the chemical shift changes in each activator–AcID complex via ^{1}H,^{15}N-HSQC NMR titration experiments, with VP16 (438–490), ERM (38–68), and ATF6α (40–66) in the presence of ^{15}N-labeled Med25 AcID.

A

H1 Face

H2 Face

Med25
AcID

↻
180°

B

	K_D (nM)
VP16	60 ± 4
ERM	590 ± 90
ATF6α	520 ± 70

Med25-Dependent Activators

VP16: 438-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG-490

ERM: 38-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ-68

ATF6α: 40-DTDELQLEAANETYENNFDNLDFDLDL-66
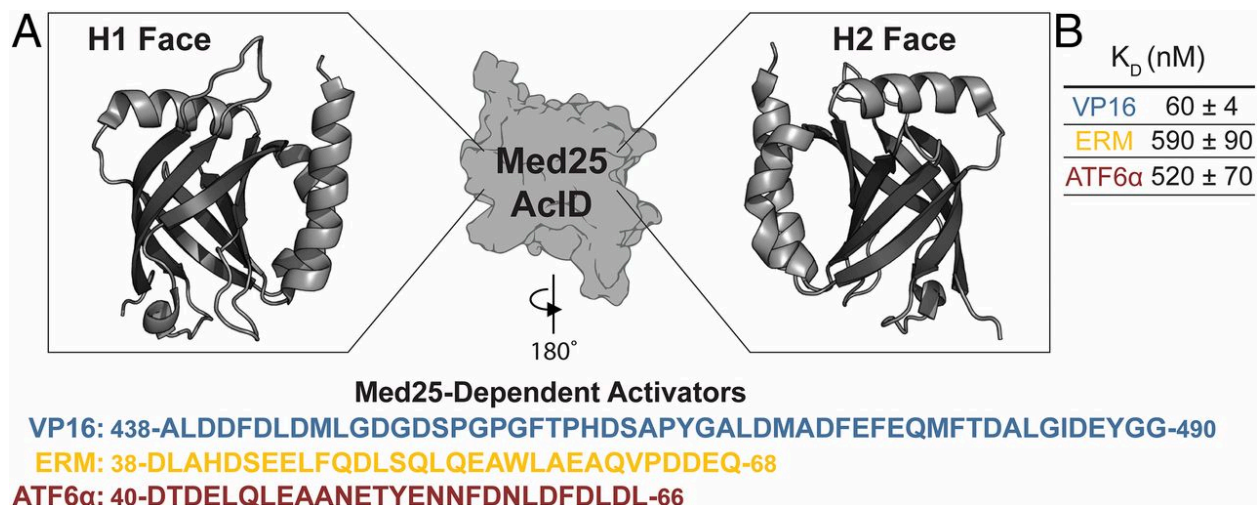
Figure 2.1 Med25 AcID forms complexes with transcriptional activators of distinct sequences. (A) The AcID is the binding partner of a growing number of transcriptional activators and contains at least two binding surfaces, termed H1 and H2. The sequences of the transcriptional activation domains of the three Med25-dependent activators used in this study are shown below the protein structure (PDB ID code 2XNF). (B) Equilibrium dissociation constants for each of Med25 AcID–activator complexes, measured through fluorescence anisotropy experiments using fluorescein-labeled peptides. These values are the average of at least three independent measurements with the error indicated (standard deviation of the mean). Binding experiments were completed in collaboration with Dr. Andrew Henderson, Nicholas Foster, Dr. Matthew Beyersdorf, Kevon Stanford, and Dr. Steven Sturlis.

The amide proton perturbation patterns measured for the activator–AcID complexes suggest a different binding mode for each of the three activators (Figs. 2.2–2.4). VP16 induced changes at both AcID binding surfaces, consistent with the tandem transcriptional activation domains within its sequence (Fig. 2.2).[2,3] ERM binding predominantly lead to perturbations at residues on the H1 surface of AcID, in agreement with the model in which it preferentially interacts at that site (Fig. 2.3).[6,7] Key changes at residues K411, R538, and Q451, for example, were seen with both VP16 and ERM (Fig. 2.4). In contrast, interaction with ATF6α lead to significant chemical shift changes on the H2 binding surface (Fig. 2.4A). ATF6α induced shifts of residues Q456, M470, and H474, which were also affected to varying degrees by VP16 and largely unaltered by ERM. Consistent with ATF6α and ERM interacting on opposing sides of AcID, mutations introduced on one or the other of the binding surfaces produced distinct effects (Fig. 2.4C). H1 mutations R538E, K411E, and Q451E inhibit ERM binding while ATF6α is largely unaffected. In contrast, H2 mutations R466D and M523E significantly inhibit ATF6α with minimal impact on ERM binding. Taken together, these data indicate that

31

ATF6α binds on the H2 binding surface of Med25 AcID, opposite the site of ERM. Further, the distinct but overlapping chemical shift patterns observed upon binding of each of the activators to Med25 suggest several unique binding modes accommodated within AcID. This is analogous to helical activator binding domains such as GACKIX of CBP/p300, a three-helix bundle that contains at least two activator binding sites.[9]

Figure 2.2 Results of VP16 chemical shift perturbation experiments superimposed upon the Med25 AcID structure (PDB 2XNF). Residues displaying chemical shift perturbation greater than 2 SD upon VP16 binding are depicted in blue spheres. Results from experiments completed by Dr. Andrew Henderson and Dr. Brian Linhares.
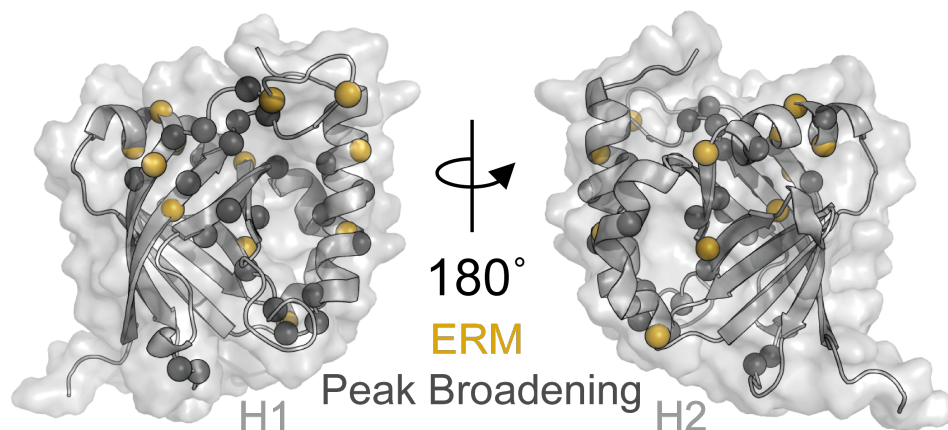
Figure 2.3 Results of ERM chemical shift perturbation experiments superimposed upon the Med25 AcID structure (PDB 2XNF). Residues displaying chemical shift perturbation greater than 2 SD upon ERM binding are depicted in maize spheres. Grey spheres indicate residues with chemical shifts that broaden upon ERM binding. Results from experiments completed by Dr. Andrew Henderson and Dr. Brian Linhares.
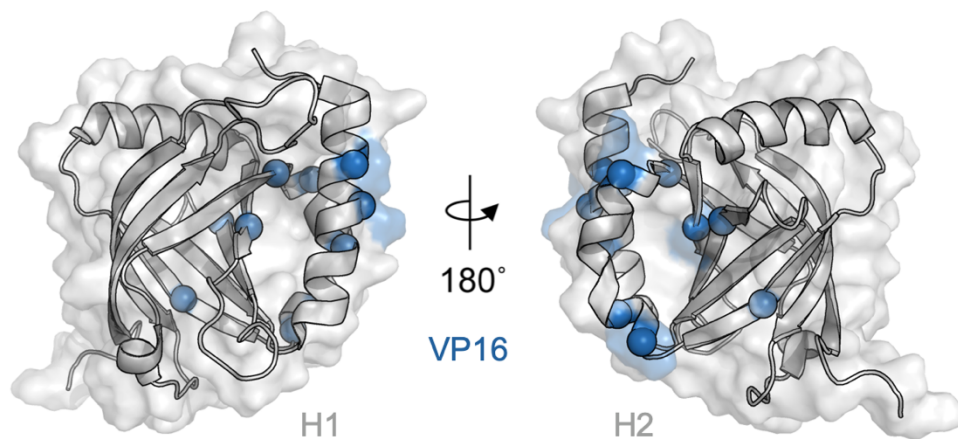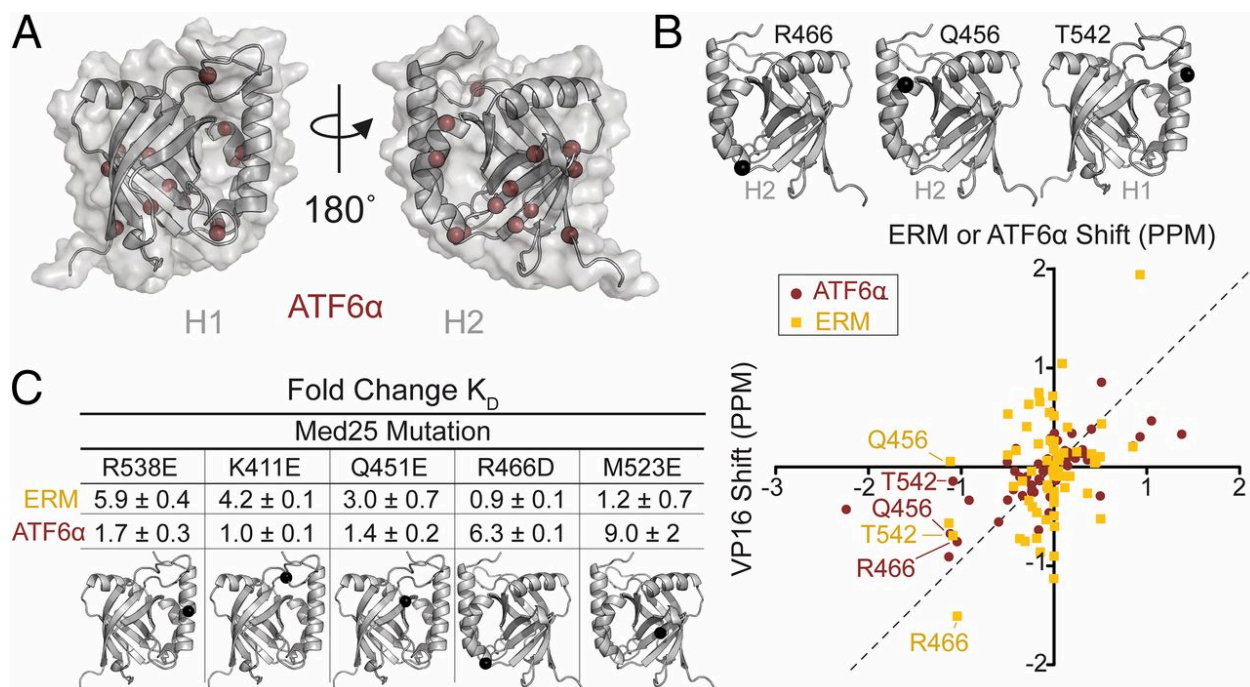
Figure 2.4 ATF6α binds to the H2 surface of Med25 AcID. (A) Results of chemical shift perturbation experiments superimposed upon the Med25 AcID structure (PDB ID code 2XNF). Residues displaying chemical shift perturbation greater than 2 SD upon ATF6α binding are depicted in rust spheres. (B) Scatter plot illustrating correlations between the chemical shift perturbations (CSPs) of individual Med25 AcID residues from HSQC experiments with ERM, ATF6α, and VP16. The position of each maize square represents the CSP of an individual residue in Med25 AcID upon binding to ERM (y-axis) and VP16 (x-axis). Thus, squares along the dotted diagonal are residues that shift similarly in both ERM–AcID and VP16–AcID complexes. The same analysis for ATF6α is shown in rust circles. Specifically labeled are the positions of three residues that are on the H1 face of AcID (T542) and H2 face of AcID (R466, Q456), highlighting the distinct pattern of correlated CSPs for ERM and ATF6α, consistent with the model in which the two activators do not interact with the same binding site. (C) Results of direct binding experiments with fluorescein-labeled activators and the indicated mutants of Med25 AcID as measured by fluorescence polarization expressed the fold change relative to the dissociation constant of each activator for the WT AcID. The indicated error is propagated from three independent dissociation constant measurements. Binding experiments were completed by Dr. Andrew Henderson, Nicholas Foster, Dr. Matthew Beyersdorf, Kevon Stanford, and Dr. Steven Sturlis. NMR experiments were completed by Dr. Andrew Henderson and Dr. Brian Linhares.

Activator•Med25 Complexes Are Conformationally Dynamic.

Next, the underlying mechanistic features of activator–AcID complex formation were examined by determining association mechanisms of AcID with the TADs of VP16, ERM, and ATF6α using stopped-flow fluorescence spectroscopy. These kinetic experiments allow calculation of microscopic rate constants for association and dissociation, as well

as forward and reverse rate constants for any conformational changes that happen during the binding process.[10,11] In practice, however, conformational changes involved in activator–coactivator interactions are often difficult to detect and quantify, and we thus chose the environmentally sensitive fluorophore 4-*N*,*N*-dimethylamino-1,8-napththalimide (4-DMN) as a fluorescence probe (Fig. 2.5A).[12,13] This fluorophore was synthesized as a conjugate with β-alanine and incorporated at the amino terminus of these activators for subsequent experiments. The overall results of these experiments are depicted in Figure 2.5.
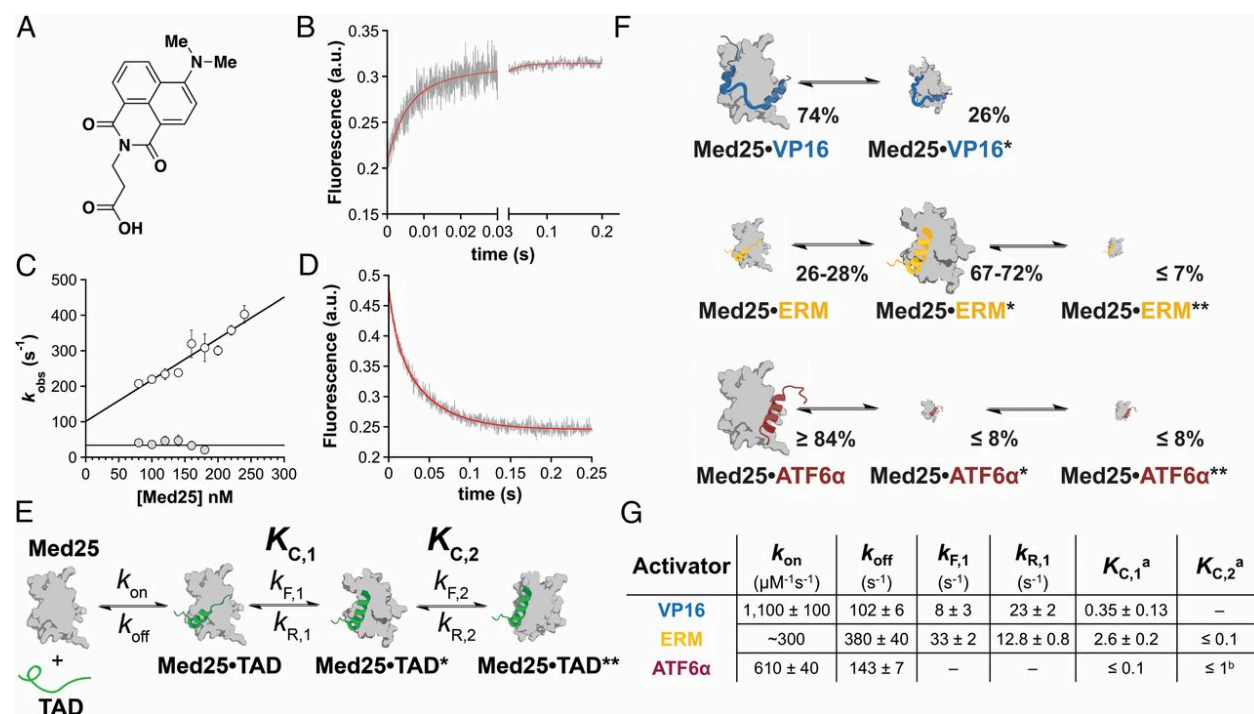
Figure 2.5 Transient kinetic experiments define minimal mechanism of activator–AcID complexation. (A) Structure of the fluorophore used in these experiments, the β-alanine conjugate of 4-*N*,*N*-dimethylamino-1,8-napththalimide (4-DMN). (B) Representative kinetic trace of association experiment with 4-DMN-VP16(438-490) and AcID. The red line is the fit to a two-step binding model. (C) Dependence of the two observed rate constants for the fast (white circles) and slow (gray circles) kinetic phases on the concentration of AcID for association experiments of VP16 with AcID. (D) Sample kinetic trace of a dissociation experiment in which 10 μM unlabeled VP16(438-490) was added in excess to a preformed complex of 50 nM 4-DMN-VP16 and 100 nM AcID. (E) General kinetic mechanism for TAD–AcID complex formation as determined by these experiments for all activators (see Figs. 2.7 and 2.8 for further details on mechanistic support). Microscopic equilibrium constants ($K_{C,n}$) are defined as the ratio of the respective forward and reverse rate constants. (F) Representation of equilibrium population distributions of bound states, calculated from equilibrium constants in 2.5G. Size of each state is scaled according to the indicated percentage population. When one equilibrium constant is too small to measure, the values are given as ranges. (G) Measured kinetic and equilibrium constants for all of the

activators. Kinetic constants $k_{F,2}$ and $k_{R,2}$ were unable to be reliably calculated from the data sets. a, The conformational change equilibrium constant is too small to be measured with precision. b, The overall equilibrium constant from Med25–ATF6a to Med25–ATF6a** is estimated to be ≤0.1 based on the limits of precision of our experiments, thus given that $K_{C,1} \leq 0.1$, $K_{C,2}$ must be ≤1.

Previous kinetic studies of helical coactivators revealed that complex formation with activators proceeds by fast association (k_{on}) and dissociation (k_{off}) rate constants.[11,14–18] Consistent with this observation, we found that activator–AcID complexes form with elevated k_{on} and k_{off} values, with k_{on} ranging between 300 and 1,100 $\mu M^{-1} \cdot s^{-1}$ and k_{off} ranging between 100 and 400 s^{-1} (Fig. 2.5G). This behavior allows for activators to form tight interactions (K_D values 50–500 nM) with Med25 AcID that are short-lived, with activator residence times less than 10 ms. The k_{on} values are 1–2 orders of magnitude faster than most other activator-coactivator systems, likely a result of significant electrostatic contributions to binding (Fig. 2.6), which can elevate association rate constants by several orders of magnitude.[15,16,19] The full mechanism of binding and all calculated rate and equilibrium constants are shown in Fig. 2.5E and 2.5G. Populations of detected conformational states at equilibrium are shown in Fig. 2.5F. At least one conformational change during the binding process was observed in all cases, with similar observed rate constants (k_{obs} = 10–40 s^{-1}) for each activator. For VP16 and ERM, the conformational change was determined to occur after the initial binding event by a combination of global fitting in Kintek Explorer and an "inverted" association experiment (Figs. 2.7, 2.8).[20,21] For ATF6α, amplitudes of conformational change phases were too small to reliably derive rate parameters, and thus an upper limit on overall conformational equilibrium constants from the first bound conformation (Med25•TAD in Fig. 2.5E) was placed at < 0.1.

Effect of NaCl on VP16(465-490)•Med25 Binding

1000 mM	$K_D = 11.5 \pm 1.0$ μM
800 mM	$K_D = 10.2 \pm 1.2$ μM
600 mM	$K_D = 5.5 \pm 0.5$ μM
400 mM	$K_D = 2.9 \pm 0.1$ μM
200 mM	$K_D = 0.91 \pm 0.08$ μM
100 mM	$K_D = 0.37 \pm 0.03$ μM
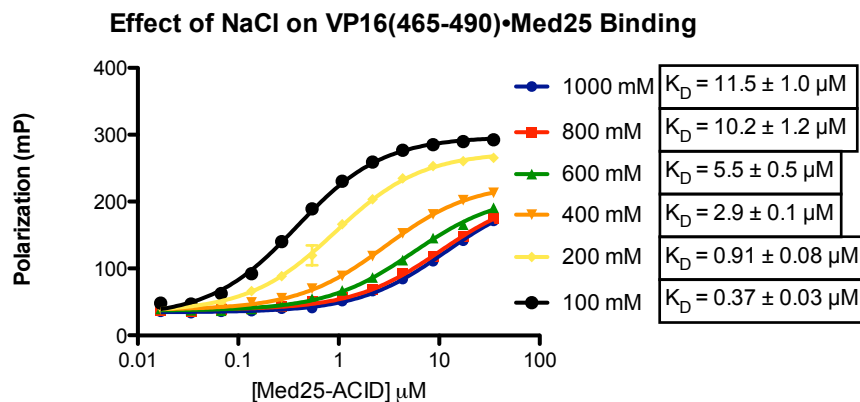
Figure 2.6 Increasing concentrations of NaCl attenuate the affinity of VP16 for Med25 AcID. Each trace is the average of three independent experiments with the indicated error (SDOM). Experiment was completed by Dr. Steven Sturlis.

Importantly, the equilibrium constant of this conformational change ($K_{C,1}$) varied significantly across activators (Fig. 2.5G). For the activator VP16, which binds both sites simultaneously, the initial bound state was somewhat more favorable ($K_{C,1} = 0.4 \pm 0.1$). In contrast, the activators ERM and ATF6α, which bind to opposite sites of AcID, had $K_{C,1}$ values that were significantly larger and smaller than VP16, respectively ($K_{C,1} = 2.6 \pm 0.2$ and < 0.1). This range of $K_{C,1}$ values is not unusual for activator-coactivator association mechanisms; we have previously observed this behavior with activator–Med15 complexes, where $K_{C,1}$ values positively correlated with transcriptional output.[11,22,23] A second conformational change for ERM and ATF6α that was slower ($k_{obs} = 2$–6 s^{-1}) was also detected in these experiments and had equilibrium constants that were too small to quantify reliably. Global fitting of the kinetic traces suggested this conformational change happens after the primary conformational change.
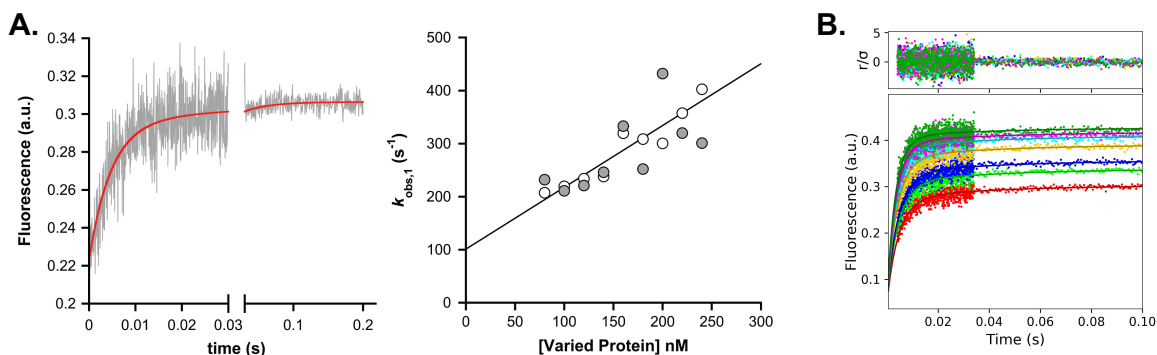
Figure 2.7 Experiments used to define and support proposed mechanism for VP16•Med25 complex formation. (A) Left: Representative association trace in 'inverted' experiment. A total of 100 nM 4-DMN-β-Ala-VP16(438-490) was rapidly mixed with 50 nM Med25. Right: concentration dependence of the fast phase of inverse concentration dependence experiment (grey circles) follows that of the "normal" concentration dependence (white circles). This behavior is expected only if the conformational change occurs after binding.[21] The inverse experiment has somewhat larger scatter due to lowered S/N. (B) Global fitting results from Kintek Explorer obtained by fitting the association traces to the proposed model, restraining the rate constants to the calculated values and fitting only fluorescence values with small offsets.
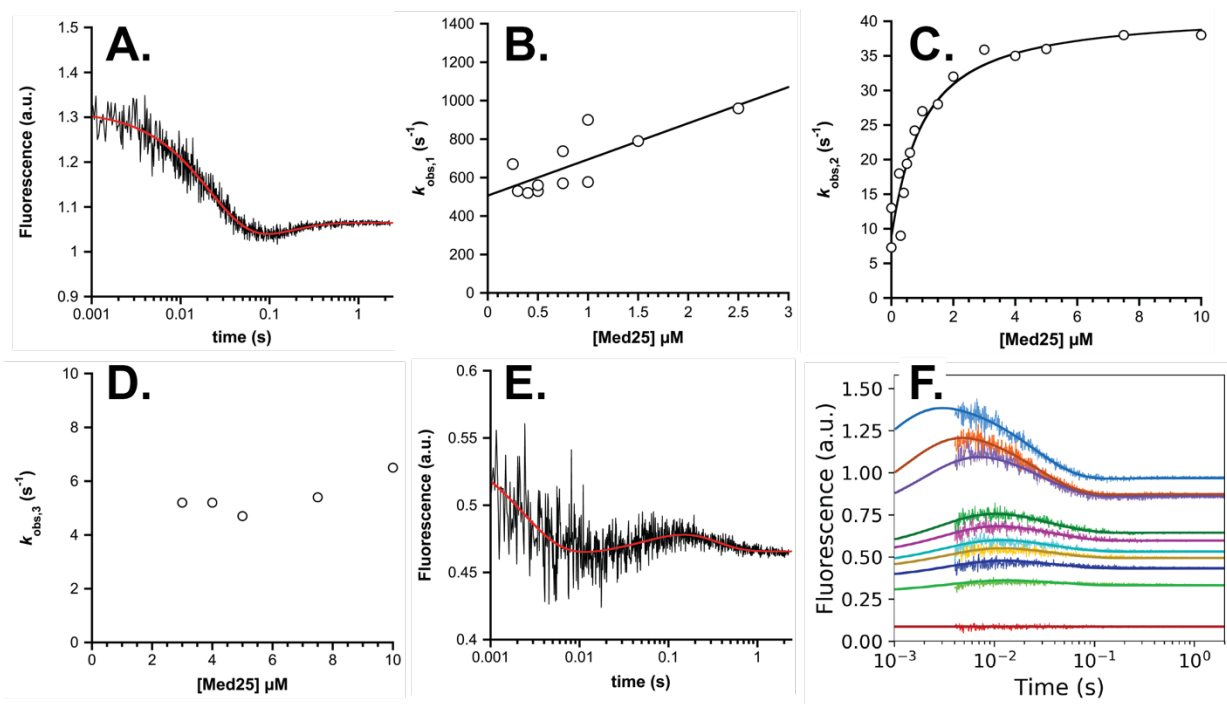
Figure 2.8 Experiments used to define the Med25•ERM(38-68) mechanism. (A) Representative association trace of Med25 with 4-DMN-ERM(38-68). 10 μM Med25 was rapidly mixed with 0.25 μM 4-DMN-ERM(38-68). Here, the initial association is complete in the dead time, and the two observed phases are from the two conformational change phases. The red line is a fit to a double exponential. (B) Concentration dependence of the fast phase in association experiments. Values

above 1000 s^{-1} were removed from the dataset before fitting. Values for B-D are from single experiments, and the data represents the combination of two independent datasets. (C) Concentration dependence of the intermediate phase in association experiments. (D) Concentration dependence of the slow phase in association experiments. (E) Representative dissociation trace of Med25•4-DMN-ERM(38-68). 2 µM Med25 precomplexed with 0.5 µM 4-DMN-ERM(38-68) was rapidly mixed with 75 µM unlabeled ERM(38-68). The red line is a fit to a triple exponential. (F) Results of the global fit. Association experiments from a single dataset were fit together in Kintek Explorer (along with the dissociation experiment) to the three step model shown in Fig. 2.4G.

Some structural insight regarding the primary conformational change of the ERM–AcID complex was gained from the ^{1}H,^{15}N-HSQC titration of ERM with AcID, as many AcID residues display slow and intermediate exchange behavior in this complex (Fig. 2.3). Specifically, the intermediate exchange behavior points to conformational exchange on the millisecond timescale ($\Delta v = 30$–180 s^{-1}), consistent with the value of the exchange rate constant measured in the kinetics experiments (45 s^{-1}). Thus, inspecting the residues that display signal broadening shows that this conformational change involves residues at both sites (Fig. 2.3), including many residues on the flexible loops and all three helices. Taken together, the NMR and kinetic data point to a mode of interaction where Med25 AcID forms conformationally unique complexes with its activator interaction partners. While these activators bind to Med25 with similar affinities, they each have different kinetic signatures characterized by a conformational change after binding that extends throughout the AcID structure. Finally, all of these complexes can be described as "fuzzy"; at equilibrium, each complex contains multiple populated conformational substates that are separated by low energy barriers.[24]

The H1 and H2 Binding Sites Are in Allosteric Communication

The presence of two binding sites that engage with distinct activators raises the question whether AcID contains an allosteric network linking the two sites to permit cooperative formation of specific activator–AcID ternary complexes. The prototypical case of allostery in ABD–activator complex formation is the GACKIX motif in CBP/p300.[14,16,25,26] In this example, the "signature" of allosteric communication is reduction of the k_{off} of an activator when another activator is bound at the allosteric site,

with the k_{on} largely unaffected. Thus, for AcID, we measured dissociation rate constants as a primary method to detect allosteric communication.

	IC_{50}	
	Competitor TAD	
Tracer	VP16(413-454)	VP16(455-490)
FITC-VP16(413-454)	4.4 ± 0.5 μM	5.6 ± 0.2 μM
FITC-VP16(455-490)	6.0 ± 0.7 μM	4.8 ± 0.4 μM
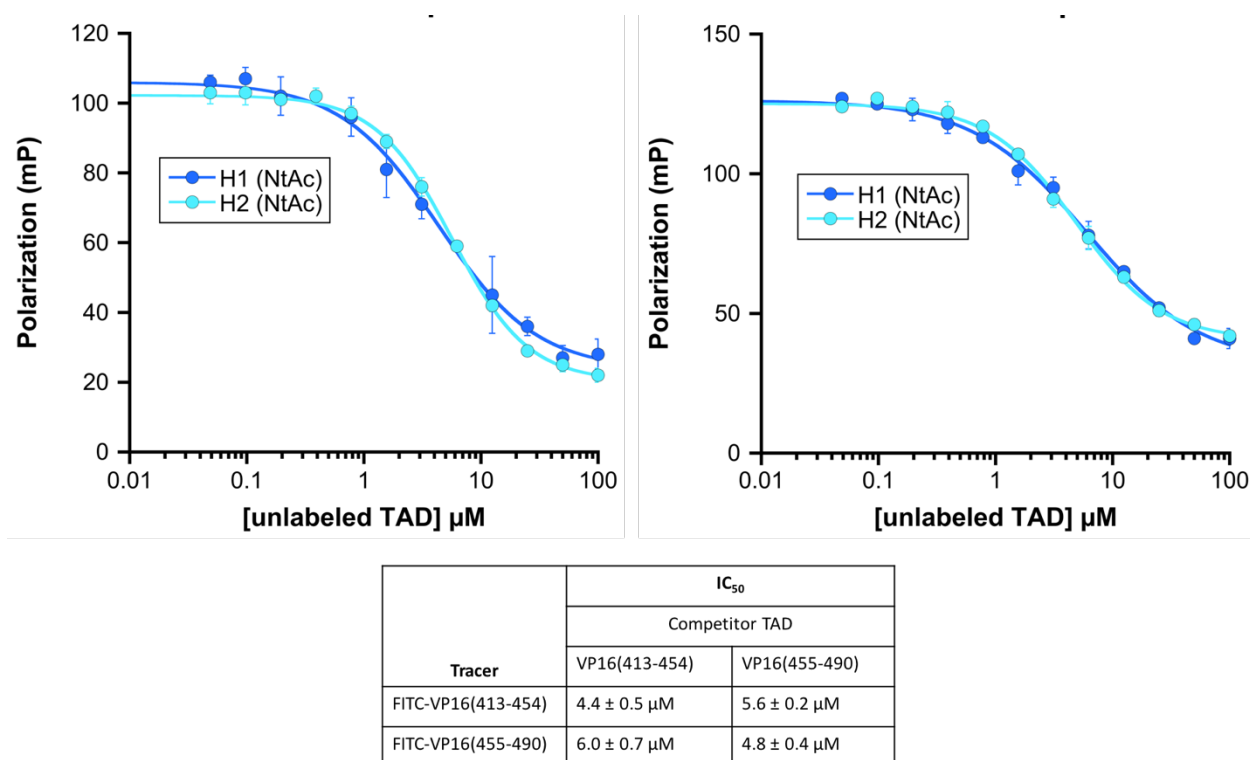
Figure 2.9 The individual halves of VP16 interact with both the H1 and H2 sites of Med25 AcID. Competition binding experiments demonstrate that *N*-terminally acetylated VP16(413-455) and VP16(455-490) can each disrupt pre-formed complexes of Med25 AcID with FITC-VP16(413-455) or FITC-VP16(455-490) with similar IC_{50} values. This indicates that these individual halves of VP16 bind at both available Med25 AcID binding surfaces.

We first looked to use the VP16 TAD as a model system; the two halves represent individual interaction motifs that bind separate sites, and when employed *in trans*, it would be expected that they could be used to dissect communication between AcID sites. However, when separated, the two VP16 halves lose 10- to 30-fold affinity and display poor selectivity for one binding site over another (Fig. 2.9), which is not surprising given the topological similarity between the two binding sites of AcID and the high sequence homology between the two VP16 activator motifs. This reduced selectivity significantly complicates data interpretation as the presence of multiple distinct ternary complexes can mask allosteric effects. To address this, we took advantage of a distinguishing feature between the AcID binding sites, the presence of two solvent-exposed cysteine residues (C497 and C506) within the H1 site, which were then employed to tether the relevant

VP16 fragment to that site via a disulfide (Fig. 2.10A). Tethering experiments with a library of disulfide-containing point mutations of the H1-targeting portion of VP16, VP16 (438–454), were carried out with AcID, and the G450C mutation led to 100% formation of a disulfide bond with the C506 residue of AcID (Fig. 2.10B). Consistent with the hypothesis of allosteric communication, multiple chemical shifts corresponding to residues in the H2 site are perturbed in the ^{1}H,^{15}N HSQC spectrum (Fig. 2.11). In contrast, the introduction of cysteine to the H2 surface at a variety of positions produced mutants with a high aggregation propensity, rendering them unsuitable for Tethering and binding studies.
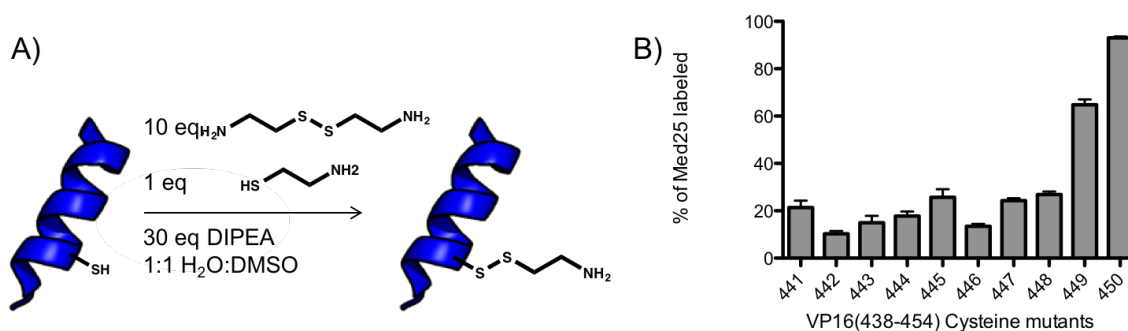
Figure 2.10 Peptide Tethering A) VP16(438-454) cysteine mutant peptides were reacted with cystamine to form disulfide-capped peptides for use in peptide Tethering experiments. B) Results from peptide Tethering experiments with Med25 AcID as measured by mass spectrometry. The labeling percentage for each peptide is the average of three independent experiments with the indicated error (SDOM). Tethering experiments were performed with Dr. Andrew Henderson.
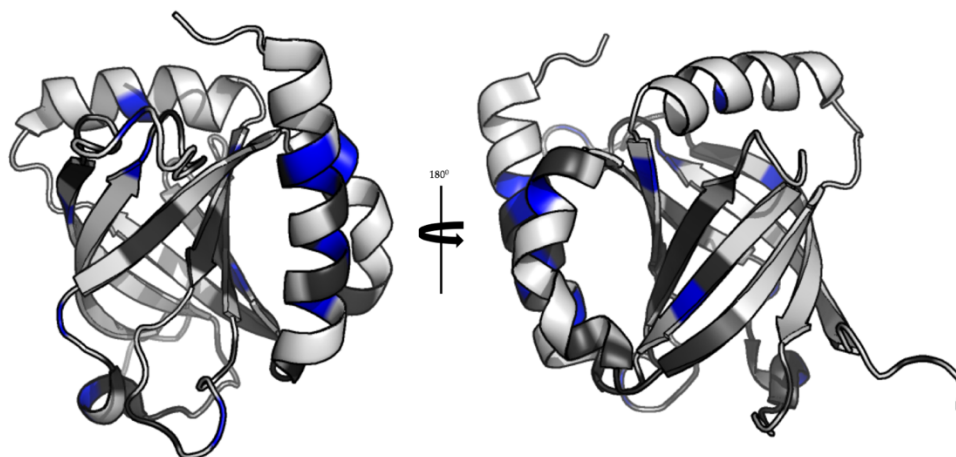
Figure 2.11 Chemical shifts induced by VP16(438-454)$_{G450C}$ mapped to the surface of Med25 are indicated in blue. Residues that displayed significant broadening or are unassignable to large shifts are highlighted in black. Experiment and assignment completed by Dr. Andrew Henderson and Dr. Brian Linhares.

With this tool in hand, two separate 4-DMN–labeled probes that interact with the H2 binding site, VP16 (467–488) and ATF6α, were used to assess changes in ternary complex formation upon occupancy of the H1 site of AcID (Fig. 2.12A). As a complement to the covalent system, ERM was also employed as a noncovalent H1 partner in a separate set of experiments due to its greater apparent selectivity for the H1 binding surface. With VP16(438–454)$_{G450C}$ tethered at the H1 site, the $\underline{k}_{off}$ value of the VP16 (467–488) ligand was reduced by 20% (Fig. 2.12B). The corresponding value for the ATF6α probe was reduced by $\sim$10%, although this did not reach statistical significance ($P = 0.1$). Consistent with our hypothesized model, k_{on} for ATF6α was unchanged (Fig. 2.13A; we were unable to measure VP16 (467–488) k_{on} due to a very high k_{off} value). When ERM was bound to the H1 site, k_{off} for VP16 (467–488) displayed a similar 20% reduction, while the ATF6α k_{off} value was reduced by 25%. In contrast to AcID covalently tethered to VP16(438-454)$_{G450C}$, when AcID is noncovalently bound to ERM the apparent k_{on} value for ATF6α is decreased by 30%. However, analysis of the raw kinetic traces strongly indicates that this change is due to a small fraction of ERM binding to the H2 site (Fig. 2.13C), and it is likely that in the absence of this masking effect k_{on} is relatively unchanged. Taken together, the data demonstrates that the two binding sites of Med25 AcID are allosterically linked, and the mechanism of allosteric communication (reduction of k_{off}) is analogous to that of the GACKIX motif.[14,16] These measured cooperativity values ($\sim$1.3) are similar to values previously measured for GACKIX, which are in the range of 1.4–2.2 for most ternary complexes, but in certain cases, are as high as 18.[16,27] We expect as more Med25 binding partners are reported and characterized these cooperativity factors will also vary significantly for different ternary complexes.[8]
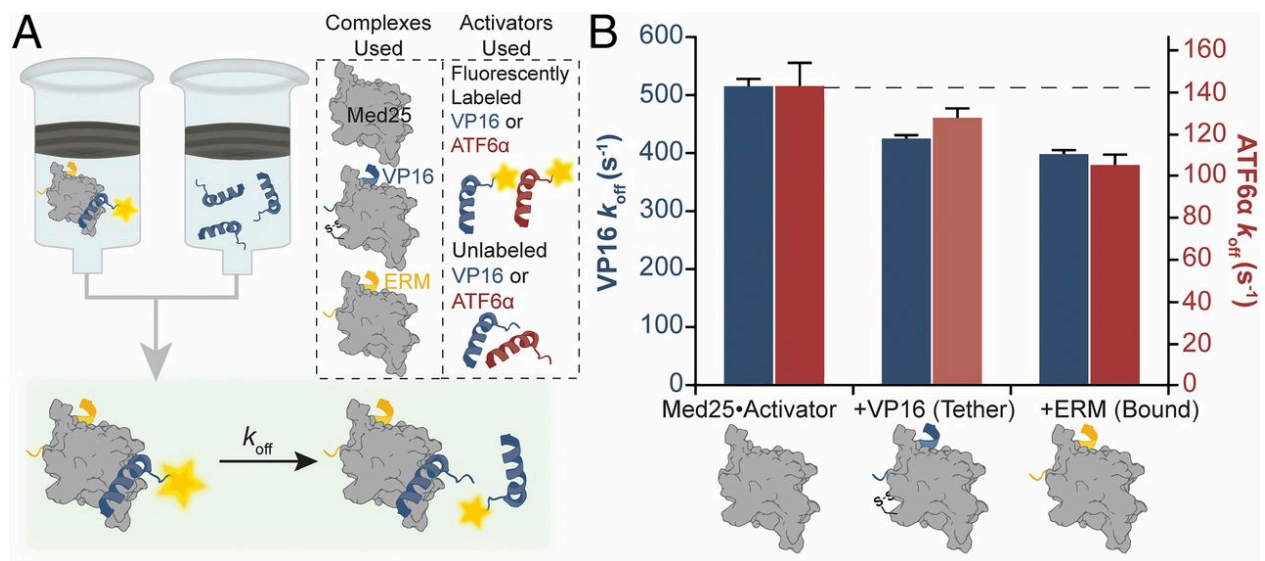
Figure 2.12 Dissociation experiments reveal allosteric communication between two binding sites. (A) Schematic of the experiment. (B) Comparison of k_{off} for 4-DMN-VP16(467–488) (blue bars) for Med25 AcID, Med25 AcID with VP16(438–454)$_{G450C}$ covalently Tethered, and Med25 AcID with ERM prebound; the red bars summarize data from analogous experiments with ATF6α. The values shown are the average of 2–3 independent experiments with the indicated errors (SD). All changes from the binary complex were statistically significant (P < 0.01), except for the Tethered complex bound to ATF6α (transparent bar).
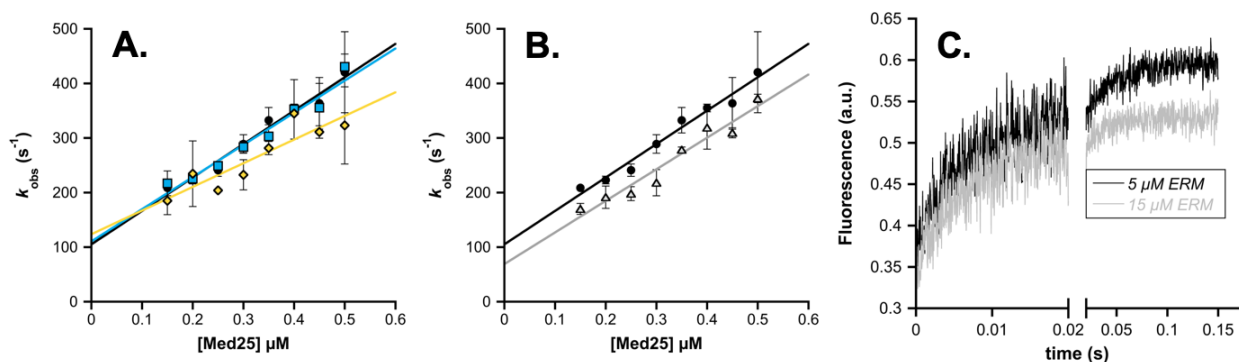
Figure 2.13 Association kinetics of ATF6α with various Med25 complexes shows no or minimal changes to association rate. The association rate k_{on} is defined by the slope of k_{obs} vs. [Med25]. (A) Comparison of free Med25 (black circles) with the Med25–VP16(438-454)G450C tethered complex (blue squares) and the Med25•ERM(38-68) noncovalent complex (yellow diamonds). (B) Comparison of free Med25 (black circles) with the covalent Med25–**22** complex (grey triangles, see next section). (C) Comparison of raw association kinetic traces at two different concentrations of unlabeled ERM, 5 µM (black trace) and 15 µM (grey trace), showing a significant decrease in signal change which can be attributed to weak binding of ERM at the H2 site. In these experiments, the final concentrations of Med25 and 4-DMN-ATF6α(38-75) were 0.35 µM and 0.125 µM, respectively. Both experiments were repeated on the same day using the same stock solutions of protein and peptide, which facilitates direct comparison.

A Covalent Cochaperone Recapitulates Allosteric Changes

We have previously demonstrated that prototypical conformationally dynamic coactivators such as GACKIX can be allosterically modulated by covalent cochaperones[14,28]; these can be rapidly identified with the covalent fragment discovery method of Tethering.[29,30] In the GACKIX case, engagement with the most dynamic sites within the coactivator lead to the most effective cochaperones. To identify such regions within Med25 AcID suitable for chemical cochaperone discovery, all-atom molecular dynamics simulations were carried out using implicit solvent models (GBSW) and with temperature replica exchange in CHARMM.[31–34] Simulations were performed on both unbound Med25 AcID and a model of this protein in which the VP16 (438–454)$_{G450C}$ is tethered at C506. To identify the substructures most stabilized upon binding, the root-mean-square fluctuations (RMSF) for each residue were calculated from the resulting trajectories (Fig. 2.14A,B). In these figures, the line color reflects the range of motion of each residue. In the unbound structure (Fig. 2.14A), the β-barrel core is relatively static whereas the loops and helices framing the two binding sites show particular mobility. The presence of VP16 (438–454)$_{G450C}$ significantly alters the extent of motion (Fig. 2.14B). Particularly notable is that the upper loop on the H1 binding surface (residues 409–424 of Med25) appears to strongly interact with VP16. Supporting this model is the effect of mutations within this region on another H1 binding activator, ERM; a K411E mutation, for example, resulted in fourfold weaker ERM binding (Fig. 2.4C). The helices flanking the H1 binding surface also undergo significant stabilization upon binding, suggesting that they also play an important role in the defining the binding site. An analysis of RMSFs of residues in Med25 AcID unbound to any ligand reveals that the most dynamical regions of the protein are indeed the loops, with significant motion in the flanking helices as well, consistent with the preliminary structural model.[2,3]
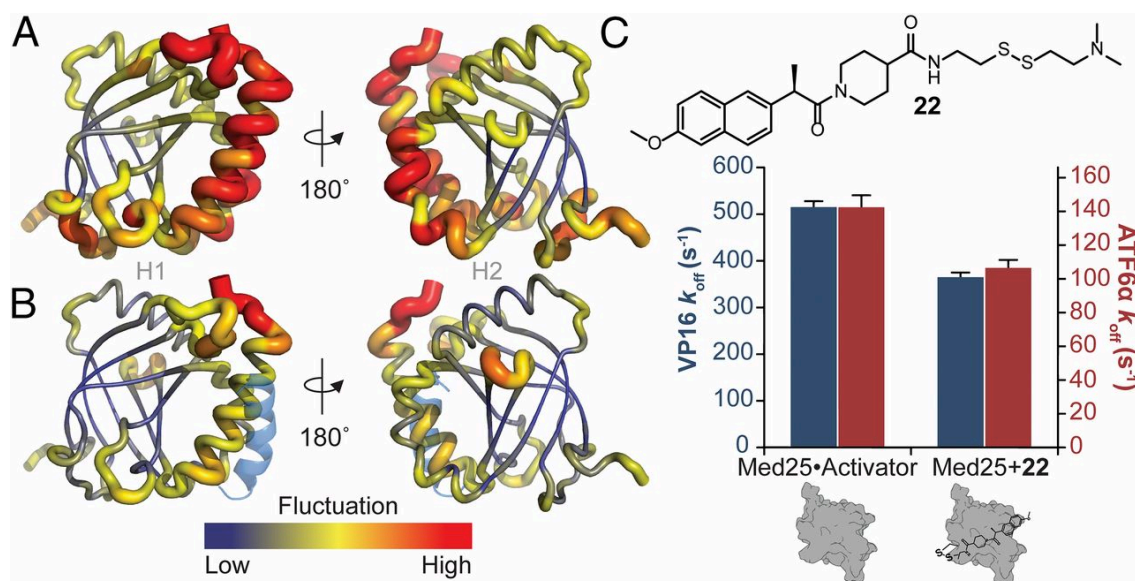
Figure 2.14 Emerging structural model for AcID–activator complex formation. (A and B) The NMR coordinates for Med25 AcID (PDB ID code 2XNF)[2] were used to construct the initial structure of Med25 in CHARMM using the Multiscale Modeling Tools for Structural Biology. For B, VP16 (438–454) G450C was constructed in CHARMM as a helical peptide, which was then patched in CHARMM to Med25 C506 through the formation of a disulfide bond at C506 (transparent blue helix). Using GBSW implicit solvent, temperature replica exchange was implemented using the CHARMM22 force field.[34] The RMSFs were calculated for each Med25 AcID residue by overlaying Cα atoms for all of the coordinate files produced from the simulations. The coloring correlates with the degree of dynamical behavior of each region. (C) Structure of chemical cochaperone 22 obtained from a Tethering screen. The bar graph is a comparison of k_{off} for VP16 (467–488) (blue bars) for Med25 AcID or Med25 AcID with **22** covalently Tethered; the red bars summarize data from analogous experiments with ATF6α. The values shown are the average of 2–3 independent experiments with the indicated errors (SD). Association data for ATF6α and Med25–22 is shown in Fig. 2.13B. MD simulations were performed by Amanda Peiffer. The Tethering Screen was performed by Dr. Zachary Hill, UCSF.

The two solvent-accessible cysteines (C497 and C506) in Med25 AcID are adjacent to regions that are predicted to be the most mobile in the preliminary structural model outlined above and that are most affected by activator binding. Thus, a Tethering screen of Med25 AcID utilizing a 1,600-member library was carried out using standard methods.[29,30] This experiment identified cochaperone **22** as a molecule that covalently labels C506 in Med25 AcID with high efficiency (Fig. 2.14C). We tested the allosteric effects of the Tethered compound **22** using transient kinetics analogous to the experiments of Fig. 2.12. Similar to the effects of natural activator ligands, the values of k_{off} for both labeled activators were reduced by 25% in the presence of **22** while k_{on} was

unaffected (Fig. 2.13B). Thus, even a fragment molecule can recapitulate the key binding features of a natural activator despite considerable differences in size. This suggests that Med25 AcID will be druggable through the targeting of its most dynamic regions, despite its large binding surfaces. Particularly given our prior success with a similarly dynamic but structurally distinct ABD from CBP/p300, this appears to be a general strategy for the discovery of small molecule modulators of transcriptional coactivators.

2.4 Conclusions

Despite having a seemingly simple function, formation of one or more protein–protein interactions with transcriptional activators, the activator binding domain of a coactivator must be able to form PPIs in binary and ternary complexes that are both specific and short-lived to facilitate appropriate assembly of the transcriptional machinery and initiation. An additional complicating factor is that a single ABD is typically the cognate binding partner for tens of different activators, requiring a significant degree of structural mobility in the ABD to accommodate this diversity. This mobility likely corresponds to local folding-like transitions; it not only allows the binding interfaces to morph into unique conformations as part of binding different activators, but underlies the allosteric interactions between different binding sites in an individual domain.[35,36] Therefore, the "important" molecular recognition elements should be the most mobile regions, which is in line with our results shown here with Med25 AcID. Despite the large surface area of the core β-barrel that is used for interacting with activators, it is changes in the flanking loops and helices that enable accommodation of the distinct cognate ligands. Further, the emerging structural model suggests that it is also these regions that are responsible for allosteric communication between the two binding surfaces. Consistent with this model, engagement of one of the most mobile regions with a covalent cochaperone indeed alters binding at the opposing sites. Importantly, this suggests that this seemingly "undruggable" protein is likely targetable by allosteric small molecules (via our targeting strategy), as should transcriptional coactivators more broadly. Given the central role that many coactivators play in human disease, this will be a critical advance. Further, since the first structural reports of AcID, the identity of activator and coactivator binding partners of Med25 has expanded,[8] and the molecular recognition model outlined here indicates that

cooperative binding of Med25 to activators and/or coactivators such as CBP may be a key regulatory mechanism.

2.5 Materials and Methods

Peptide synthesis

The peptides listed in Table 2.1 were prepared following standard FMOC solid-phase synthesis methods on a Liberty Blue Microwave Synthesizer (CEM). FMOC deprotections were completed by suspending the resin in 20% piperidine (ChemImpex) in DMF supplemented with 0.2 M Oxyma Pure (CEM) and irradiating under variable power to maintain a temperature of 90 °C for 60 seconds. Coupling reactions were completed by combining the amino acid (5 eq relative to resin; CEM, ChemImpex, and NovaBiochem), diisopropylcarbodiimide (7 eq, ChemImpex), and Oxyma Pure (5 eq) in DMF and irradiating under variable power to maintain a temperature of 90 °C for 4 minutes. The resin was rinsed four times with an excess of DMF between all deprotection and coupling steps.

At the conclusion of the syntheses, peptides listed in entry numbers 1-6 and 18-27 underwent a final FMOC deprotection and were treated with acetic anhydride (Fisher Scientific) in the presence of triethylamine (Fisher Scientific) to acetylate the amino terminus. For the peptides listed in entry numbers 7-13, at the conclusion of the synthesis a final FMOC deprotection step was performed followed by treatment with fluorescein isothiocyanate (FITC, ThermoFisher) in the presence of N,N-diisopropylethylamine (Sigma Aldrich). For peptides in entry numbers 14-17, at the conclusion of the synthesis, an additional coupling was carried out with 4-*N,N*-dimethylamino-1,8-napththalimide (4-DMN)[13] linked to β-alanine (Fig. 2.5A) prior to cleavage. Peptides were cleaved from resin with 95% trifluoroacetic acid (Sigma Aldrich), 2.5% H2O, 2.5% triisopropylsilane (Sigma Aldrich) and filtered. The resulting solution was concentrated and the peptide then precipitated by addition of cold diethyl ether. After filtration the solid was dissolved in 3:1 100 mM ammonium acetate (pH 7.0, Fisher Scientific) and acetonitrile with minimal ammonium hydroxide to fully solubilize the peptide. Purification was accomplished by 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

100 mM ammonium acetate (pH 7.0). Gradient conditions are included in the legends of the analytical traces for each peptide found at the end of this document. Final purity was assessed by analytical HPLC and peptide identity was confirmed via mass spectrometry under negative mode ionization conditions (Agilent QTOF). Purified peptides were reconstituted in a minimal volume of DMSO and quantified by UV/Vis spectroscopy using either tyrosine absorbance at 280 nm (ε_{280}=1,280 $M^{-1} \cdot cm^{-1}$), fluorescein absorbance at 495 nm (ε_{495}=72,000 $M^{-1} \cdot cm^{-1}$), or 4-DMN absorbance at 450 nm (ε_{450}=10,800 $M^{-1} \cdot cm^{-1}$). The 4-DMN extinction coefficient was measured by serial dilution of a known concentration of 4-DMN-β-Ala-OtBu in identical conditions as the measurement of peptide concentrations.

Table 2.1 Sequence of peptides used in this study

Entry	Peptide	Sequence
1	VP16 (438-490)	AcALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
2	VP16 (413-454)	AcY-βAla-APPTDVSLGDELHLDGEDVAMAHADALDDFDLDMLGDGDSPG
3	VP16 (455-490)	AcPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
4	ERM (38-68)	AcDLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
5	ATF6α (38-64)	AcFTDTDELQLEAANETYENNFDNLDFDL
6	ATF6α (38-75)	AcFTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
7	FITC-VP16 (438-490)	FITC-βAla-ALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
8	FITC-VP16 (413-437)	FITC-βAla-APPTDVSLGDELHLDGEDVAMAHAD
9	FITC-VP16 (465-490)	FITC-βAla-YGALDMADFEFEQMFTDALGIDEYG

10	FITC-VP16 (413-454)	FITC-βAla-APPTDVSLGDELHLDGEDVAMAHADALDDFDLDMLGDGDSPG
11	FITC-VP16 (455-490)	FITC-βAla-PGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
12	FITC-ERM (38-68)	FITC-βAla-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
13	FITC-ATF6α (38-64)	FITC-βAla-FTDTDELQLEAANETYENNFDNLDFDL
14	4-DMN-VP16 (438-490)	4-DMN-βAla-DALDDFDLDMLGDGDSPGPGFTPHDSAPYGALDMADFEFEQMFTDALGIDEYGG
15	4-DMN-VP16 (467-488)	4-DMN-βAla-ALDMADFEFEQMFTDALGIDEY
16	4-DMN-ATF6α (38-75)	4-DMN-βAla-FTDTDELQLEAANETYENNFDNLDFDLDLMPWESDIWD
17	4-DMN-ERM (38-68)	4-DMN-βAla-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
18	VP16 (438-454) D441C*	AcY-βAla-ALDCFDLDMLGDGDSPG
19	VP16 (438-454) F442C*	AcY-βAla-ALDDCDLDMLGDGDSPG
20	VP16 (438-454) D443C*	AcY-βAla-ALDDFCLDMLGDGDSPG
21	VP16 (438-454) L444C*	AcY-βAla-ALDDFDCDMLGDGDSPG
22	VP16 (438-454) D445C*	AcY-βAla-ALDDFDLCMLGDGDSPG

23	VP16 (438-454) M446C*	AcY-βAla-ALDDFDLDCLGDGDSPG
24	VP16 (438-454) L447C*	AcY-βAla-ALDDFDLDMCGDGDSPG
25	VP16 (438-454) G448C*	AcY-βAla-ALDDFDLDMLCDGDSPG
26	VP16 (438-454) D449C*	AcY-βAla-ALDDFDLDMLGCGDSPG
27	VP16 (438-454) G450C*	AcY-βAla-ALDDFDLDMLGDCDSPG

*The cysteine thiol has been converted to a mixed disulfide with 2-aminoethanethiol.

Preparation of disulfide-capped peptides for Tethering experiments

The VP16-derived cysteine mutants (Table S1, Entries 18-27) were cleaved from resin and following ether precipitation were dissolved in minimal 1:1 DMSO:H2O solution followed by the addition of 10 eq of cystamine•HCl (Fisher Scientific), 30 eq of diisopropylethylamine, and 1 eq of cysteamine (Fisher Scientific). The resulting solution was incubated at room temperature for 18 h on a rotating carousel. Solvent was removed by lyophilization and peptides were resuspended in 3:1 100 mM ammonium acetate (pH 7.0):acetonitrile for HPLC purification. Purified peptides were dissolved in a minimal volume of DMSO, quantified by UV/Vis spectroscopy using tyrosine and cystine absorbance at 280 nm (ε_{280}=1,440 M-1 cm-1), and stored at -20 °C.

Plasmids for protein expression

The Med25 expression plasmid pET21b-Med25(394-543)-His$_6$ was generously provided by Prof. Patrick Cramer.[2] Variants of pET21b-Med25(394-543)-His6 were prepared using site directed mutagenesis, as previously described, with the primers indicated in Table 2.2.[37] 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.

Table 2.2 Primers used for site-directed mutagenesis of Med25 AcID

Plasmid	Primer Sequence
pET21b-Med25(394-543)C506A-His$_6$	F: TCCCCCCACGGCGCCCGCCGAGGTGCGCGTGCTCATG R: CATGAGCACGCGCACCTCGGCGGGCGCCGTGGGGGGA
pET21b-Med25(394-543)C497A-His$_6$	F: GGCCAACGGCTTCGCGGGCGCCGTGCACTTCCCCACACG R: CGTGTGGGGAAGTGCACGGCGCCCGCGAAGCCGTTGCC
pET21b-Med25(394-543)C429A-His$_6$	F: ACGCGGTCACTGCCCGCCCAGGTCTAC R: ATTCACGTAGACCTGGGCGGGCAGTGA
pET21b-Med25(394-543)K411E-His$_6$	F: CTGGAGTGGCAAGAGGAGCCCAAACCTGCCTCA R: TGAGGCAGGTTTGGGCTCCTCTTGCCACTCCAG
pET21b-Med25(394-543)R538E-His$_6$	F: GGCTTCGTCAACGGCATCGAACAGGTCATCACCAACCTC R: GAGGTTGGTGATGACCTGTTCGATGCCGTTGACGAAGCC
pET21b-Med25(394-543)Q451E-His$_6$	F: CCAGAAGCTGATCATGGAACTCATCCCCCAGCAG R: CTGCTGGGGGATGAGTTCCATGATCAGCTTCTGG
pET21b-Med25(394-543)R466D-His$_6$	F: CTGGACCATCCTTGAGTTATCGAACAAAGGGCCCAG R: CTGGGCCCTTTGTTCGATAACTCAAGGATGGTCCAG
pET21b-Med25(394-543)M523E-His$_6$	F: AAGAAGAAGATCTTCGAAGGCCTCATCCCCTA R: TAGGGGATGAGGCCTTCGAAGATCTTCTTCTT

Protein expression

Wild-type Med25(394-543) and variants were expressed in E. coli Rosetta pLysS cells (EMD Millipore) using standard protocols.[2] Briefly, overnight starter cultures (50 mL) were grown in LB (Research Products International) in the presence of 100 µg/mL ampicillin (Gold Bio Technology) and 0.034 mg/mL chloramphenicol (Sigma Aldrich) at 37 °C at 150 rpm. Following overnight growth, 1 L of terrific broth (12 g tryptone, 24 g yeast extract, 4 mL glycerol, 100 mL 0.17M KH$_2$PO$_4$/0.72M K$_2$HPO$_4$, 900 mL water) supplemented with 1 mg/mL ampicillin and 0.034 mg/mL chloramphenicol was inoculated with 5 mL of the overnight culture and grown at 37 °C to an OD of 0.8, at which point the incubator temperature was lowered to 21 °C and isopropyl β-D-1-thiogalactopyranoside (IPTG, Research Products International) was added to a final concentration of 0.5 mM to induce protein expression. Agitation at 150 rpm was continued for 18 hours. Cells were harvested via centrifugation at 6500 rpm for 20 minutes, snap frozen in liquid nitrogen, and stored at -80 °C.

To isolate and purify Med25 AcID and mutants, frozen cell pellets were resuspended in lysis buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 6.5, 0.7 μL/mL β-ME) and lysed by sonication. Cellular debris was then pelleted by centrifugation for 20 min at 10,000 rpm. The supernatant was filtered using a 0.45 μm syringe filter (CellTreat) and loaded directly on an AKTA Pure FPLC equipped with a Ni column (HisTrap HP, GE Healthcare) using a gradient of Buffer A (50 mM sodium phosphate, 300 mM NaCl, and 30 mM imidazole, pH 6.8) to Buffer B (50 mM sodium phosphate, 300 mM NaCl, and 400 mM imidazole, pH 6.8) over ten column volumes. AcID-containing fractions were pooled and underwent a second purification via cation exchange (HiTrap SP HP, GE Healthcare) with a gradient of Buffer A (50 mM sodium phosphate, 1 mM DTT, pH 6.8) to Buffer B (50 mM sodium phosphate, 1 mM DTT, 1M NaCl, pH 6.8) over fifteen column volumes. Pooled fractions containing purified protein were dialyzed into storage buffer (10 mM NaPO4, 100 mM NaCl, pH 6.8). Following dialysis, protein was concentrated via Amicon 5,000 Da cutoff spin concentrator and quantified via UV/Vis spectroscopy on a NanoDrop at 280 nm using an extinction coefficient, ε_{280} = 22,460 $M^{-1} \cdot cm^{-1}$. Protein identity was confirmed via mass spectrometry (Agilent Q-TOF) and purity was assessed by SDS-PAGE on a 4-12% bis-tris gel stained using Quick Coomassie (Anatrace).

For isotopically labeled protein, Rosetta pLysS cells bearing the appropriate expression plasmid were grown in LB at 37 °C for 18 hrs. The resulting cells were pelleted via centrifugation for 10 minutes at 2500 rpm and the LB was decanted. Cell pellets were resuspended in 20 mL M9 minimal media followed by centrifugation for 10 min at 2500 RPM. M9 media was decanted and the washed pellet resuspended in 10 mL M9 media. 5 mL of suspended cells were then used to inoculate a 1 L expression flask containing M9 media supplemented with Bioexpress (6 mL/L, Cambridge Isotopes) and containing the appropriate isotopically labeled reagents (1 g/L $^{15}NH_4Cl$ or 1 g/L $^{15}NH_4Cl$ and ^{13}C D-glucose, Cambridge Isotopes). Isotopically labeled protein was isolated and purified as described above, with the exception of the storage buffer used for dialysis (20 mM sodium phosphate, 150 mM NaCl, pH 6.5).

Peptide Tethering

VP16-AcID tethered complexes were prepared by incubating 10 µM Med25 AcID with 4 eq of individual disulfide-capped VP16 peptides (Table S1, Entries 14-22) and 2 eq of β-mercaptoethanol (β-ME) for 4 hours at room temperature in protein storage buffer (10 mM sodium phosphate, 100 mM NaCl, pH 6.8). Immediately following the 4 hour incubation, labeling efficiency was determined via mass spectrometric analysis of an aliquot of each reaction on an Agilent Q-TOF and comparing the spectral peak heights of unlabeled Med25 AcID, the Med25 AcID-β-ME adduct, and the tethered Med25 AcID-VP16 complex using the following equation:

$$Percent\ labeled\ = \frac{AcID_{VP16}}{AcID_{unlabeled} + AcID_{\beta-ME} + AcID_{VP16}}\ x\ 100$$

For the preparation of covalent Med25 AcID-VP16 complexes for NMR spectroscopic analysis, Tethering was accomplished by incubating Med25 AcID with 1.5 eq of the VP16 disulfide-containing peptide of interest and 1 eq of β-ME for 6 hours at room temperature in protein storage buffer (10 mM sodium phosphate, 100 mM NaCl, pH 6.8). The complex was then purified on an AKTA Pure FPLC by cation exchange chromatography as described above. Protein complexes were subsequently buffer exchanged into either protein storage buffer (10 mM sodium phosphate, 100 mM NaCl, pH 6.8) or NMR buffer (20 mM sodium phosphate, 150 mM NaCl, pH 6.5).

Direct binding experiments

Direct binding experiments measured by fluorescence polarization were performed as previously described.[28] Binding experiments were performed in black, round-bottom 384-well plates (Corning) and read on a PHERAStar multi-mode plate reader. Data was analyzed using GraphPad Prism 5.0 and each reported dissociation constant is the average of at least three independent experiments with the indicated error (SDOM).

Transient kinetic experiments

Stopped-flow kinetic assays were performed on a Kintek SF-2001 stopped-flow apparatus equipped with a 100-W Xe arc lamp in two-syringe mode. The 4-DMN fluorophore was excited at 440 nm and its emission was monitored at wavelengths >510 nm, using a long-pass filter (Corion). All experiments were run at 10 ºC in stopped-flow buffer (10 mM sodium phosphate, 100 mM NaCl, 1% glycerol, 0.001% NP-40, pH 6.8), and the solutions were equilibrated in the instrument for at least 5 min before experiments were performed. Concentrations reported are after mixing. All kinetic traces reported are an average of 20-80 individual traces. A series of exponential equations were fit to the transient kinetic time courses, F(t) as in the equation below, to obtain the fluorescence amplitudes (F_n) and the observed rate constants (k_{obs}) for each exponential phase, where $F(0)$ is the initial fluorescence intensity and t is time.

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

Association experiments were performed by mixing excess Med25 with a constant concentration of labeled activator. Dissociation experiments were performed by precomplexing Med25 with the labeled activator and mixing with an excess of the corresponding unlabeled activator. The concentrations of labeled activators for association experiments are as follows: 50 nM VP16, 125 nM ATF6α, and 250 nM ERM. An experiment at high saturation of Med25 was used to determine $k_{obs,2,max}$ for VP16 (see next section), this experiment was carried out by mixing 500 nM labeled VP16 with 20 µM Med25. The concentrations for individual dissociation experiments are as follows: 50 nM labeled VP16, 100 nM Med25, 10 µM unlabeled VP16; 250 nM labeled ATF6α, 500 nM Med25, 10 µM unlabeled ATF6α; 500 nM labeled ERM, 2 µM Med25, 75 µM unlabeled ERM; 250 nM labeled VP16(467-488), 1.5 µM Med25 or Med25 conjugate, 75 µM unlabeled VP16(455-490). The labeled activator concentration in both association and dissociation experiments was chosen to maximize S/N and minimize ligand depletion effects in association experiments.

Calculation of Rate Constants

As described in the main text of this chapter (Fig. 2.8), the values of microscopic rate constants were determined for ERM by global fitting. For ATF6α, the fluorescence amplitudes of conformational change phases were too small to extract accurate k_{obs} and thus the overall equilibrium constants were given an upper limit of < 0.1 from the initial bound state. Thus, for ATF6α, k_{on} and k_{off} were determined from the solution to a single step binding mechanism (below), where k_{obs} is the rate parameter extract from the large amplitude association phase. The value of k_{off} was determined directly by a separate dissociation experiment, as described above and in Fig. 2.12.

$$k_{obs} = k_{on} \times [Med25] + k_{off}$$

In the case of VP16, the only one conformational change was observed and had significant amplitudes in both association and dissociation experiments. Thus, the first order microscopic rate constants were calculated from the exact expression for a conformational change after binding mechanism as follows, using the values from dissociation experiments:

$$k_{obs,off,1,2} = \frac{k_{off} + k_f + k_r \pm \sqrt{(k_{off} + k_f + k_r)^2 - 4 \times k_{off} \times k_r}}{2}$$

Because $k_{obs,2,max}$ is the sum of k_f and k_r, k_{off} is simply the sum of the two observed off-rate constants minus $k_{obs,2,max}$:

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

A value for k_r can also be easily resolved from the exact expression, by first subtracting the two observed off-rate constants and then substituting observable parameters and rearranging:

$$k_{obs,off,1} - k_{obs,off,2} = \sqrt{(k_{off} + k_f + k_r)^2 - 4 \times k_{off} \times k_r}$$

$$= \sqrt{(k_{obs,off,1} + k_{obs,off,2})^2 - 4 \times (k_{obs,off,1} + k_{obs,off,2} - k_{obs,2,max}) \times k_r}$$

$$k_r = \frac{(k_{obs,off,1} + k_{obs,off,2})^2 - (k_{obs,off,1} - k_{obs,off,2})^2}{4 \times (k_{obs,off,1} + k_{obs,off,2} - k_{obs,2,max})}$$

The value of k_f is then simply obtained by subtracting the calculated value of k_r from $k_{obs,2,max}$. The results of the calculations are shown in Table 2.3 (note in the table k_f and k_r are denoted as $k_{F,1}$ and $k_{R,1}$ to keep consistency with the mechanisms of other activators). The value of k_{on} is then obtained from the slope of $k_{obs,1}$ vs. the concentration of Med25 in Fig. 2.5C.

Table 2.3 Calculated rate and equilibrium constants from all kinetic experiments

Labeled Activator	H1 Partner	k_{on} (µM^{-1}s^{-1})	k_{off} (s^{-1})	$k_{F,1}$ (s^{-1})	$k_{R,2}$ (s^{-1})	$K_{C,1}$[a]	$K_{d,kin}$ (nM)
VP16(438-490)	–	1,100 ± 100	102 ± 6	8 ± 3	23 ± 2	0.35 ± 0.13	70 ± 30
ERM(38-68)	–	300[b]	380 ± 40[b]	33 ± 2	12.8 ± 0.8	2.6 ± 0.2	390 ± 50
ATF6α(38-75)	–	610 ± 40	143 ± 7	–	–	< 0.1	230 ± 20
ATF6α(38-75)	VP16(438-454)$_{G450C}$	590 ± 50	129 ± 8	–	–	< 0.1	220 ± 20
ATF6α(38-75)	ERM(38-68)	430 ± 80	106 ± 6	–	–	–	250 ± 50
ATF6α(38-75)	22	580 ± 60	107 ± 4	–	–	< 0.1	180 ± 20
VP16(467-488)[c]	–	–	517 ± 11	–	–	< 0.1	–
VP16(467-488)[c]	VP16(438-454)$_{G450C}$	–	427 ± 4	–	–	< 0.1	–
VP16(467-488)[c]	ERM(38-68)	–	400 ± 5	–	–	–	–
VP16(467-488)[c]	22	–	367 ± 8	–	–	< 0.1	–

[a]Values of $K_{C,1}$ that are marked < 0.1 were detectable, but due to the small amplitudes of these phases they could not be quantified reliably. $K_{C,2}$ and the corresponding $k_{F,2}$ and $k_{R,2}$ values are omitted because they were too small to measure (see supplementary discussion). [b]For ERM(38-68), k_{on} and k_{off} were not taken from global fitting (due to their dependence on the dead time of the instrument) but instead they were calculated from the fitted K_d and the fast observed dissociation phase from dissociation experiments. [c]The k_{on} values for VP16(467-488) were not measurable due to the very fast k_{off} value.

Molecular Dynamics

The objectives of our modeling efforts were to predict an ensemble of putative structures for the N-terminal predicted helical fragment of VP16(438-454) G450C tethered to the Med25 AcID domain C506 via a disulfide Tether. Our modeling was initiated from the published NMR coordinates for Med25 AcID (PDB 2xnf).[2] The protein structure of Med25 AcID was prepared for simulation in CHARMM[33,38] using the Multiscale Modeling Tools for Structural Biology (MMTSB). VP16 has been shown to form a helical conformation when in complex with Med25 AcID,[3] so VP16(438-454)$_{G450C}$ was constructed in CHARMM as a contiguous helix, which was then patched using DISU patch in CHARMM to Med25 C506 through the formation of a disulfide bond. Prior to running the implicit solvent simulations, Med25 was fixed using harmonic restraints, and the complex was minimized with 1000 steps of a steepest descent algorithm. Using GBSW implicit solvent,[31,32] temperature replica exchange[34] was implemented using the CHARMM22 force field. These simulations were run for a total of 60 ns (2 fs time steps) using 12 replicas, sampling between 280-500 K and attempting coordinate exchanges every 5000 steps. The 12 replica trajectories were sorted by their respective temperatures, and the last 40 ns of the 280 K trajectories were then parsed into 4000 coordinate files. The MMTSB tool cluster.pl was used to cluster these structures, using K-means clustering[39] with 1.5 Å RMSD cutoff for the superposed Cα backbone atoms for all of the structures. The root-mean-square fluctuations (RMSF) for the last 40 ns of the 280 K trajectory were calculated for each Med25 AcID residue by superposing Cα atoms of the coordinate files produced from the simulations. Clustering of the 4000 structures from the trajectories resulted in 20 clusters for apo Med25 AcID, with the highest populated cluster containing 40% of the structures; the Med25 AcID•VP16(438-454)$_{G450C}$ complex resulted in 5 clusters, with the highest populated cluster containing 72% of the 4000 structures.

NMR analysis of peptide-AcID complexes

NMR analysis of peptide-AcID complexes were performed via 1H, 15N HSQC on a Bruker Avance III 600 Mhz spectrometer equipped with a cryogenic probe at 30 °C. VP16(438-490) and ERM(38-68) activator AcID titrations were conducted with Med25

AcID (20 mM sodium phosphate, 150 mM NaCl, pH 6.5, 5% D2O) at 50 uM and acetylated peptides were added at 0, 0.2, 0.5, 0.8, 1.1, 2, and 3 equivalents with a 2 % final DMSO concentration. Titrations with ATF6α(38-64) were conducted analogously to VP16 and ERM except acetylated peptide was added at 0, 0.25, 0.5, and 1.1 equivalents. Control spectra were obtained with Med25 and DMSO only. Tethered activator-AcID complexes were prepared and purified as described above and spectra were obtained in the absence of DMSO. Data Processing and visualization was performed using NMR Pipe and Sparky.[40,41]

NMR Assignment

NMR assignment of Med25 AcID (395-545) was conducted with 13C,15N-labeled protein on a Bruker Avance III 600 Mhz spectrometer equipped with a cryogenic probe. HSQC, HNCA, HNCACB, HN(CO)CA, and HN(CO)CACB 3d experiments were performed to obtain sequential backbone assignment of AcID. Of the 151 residues, 117 were assigned. NMR assignment of the tethered Med25-VP16(438-454) G450C construct was also performed via HNCA and HNCACB triple-resonance experiments.

2.6 References

1. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **11**, 891–894 (2015).
2. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18**, 404–409 (2011).
3. Milbradt, A. G. *et al.* Structure of the VP16 transactivator target in the Mediator. *Nat. Struct. Mol. Biol.* **18**, 410–415 (2011).
4. Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol.* **174**, 245–251 (2011).
5. 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).
6. 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).
7. 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).

8. Currie, S. L. *et al.* ETV4 and AP1 Transcription Factors Form Multivalent Interactions with three Sites on the MED25 Activator-Interacting Domain. *J. Mol. Biol.* **429**, 2975–2995 (2017).

9. 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).

10. Johnson, K. A. [61] Rapid kinetic analysis of mechanochemical adenosinetriphosphatases. in *Methods in Enzymology* vol. 134 677–705 (Elsevier, 1986).

11. Wands, A. M. *et al.* Transient-state Kinetic Analysis of Transcriptional Activator·DNA Complexes Interacting with a Key Coactivator. *J. Biol. Chem.* **286**, 16238–16245 (2011).

12. Shammas, S. L., Crabtree, M. D., Dahal, L., Wicky, B. I. M. & Clarke, J. Insights into Coupled Folding and Binding Mechanisms from Kinetic Studies. *J. Biol. Chem.* **291**, 6689–6695 (2016).

13. Loving, G. & Imperiali, B. A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4- *N,N* -Dimethylamino-1,8-naphthalimide: A Powerful Tool for the Study of Dynamic Protein Interactions. *J. Am. Chem. Soc.* **130**, 13630–13638 (2008).

14. 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.* **111**, 12061–12066 (2014).

15. Shammas, S. L., Travis, A. J. & Clarke, J. Remarkably Fast Coupled Folding and Binding of the Intrinsically Disordered Transactivation Domain of cMyb to CBP KIX. *J. Phys. Chem. B* **117**, 13346–13356 (2013).

16. Shammas, S. L., Travis, A. J. & Clarke, J. Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc. Natl. Acad. Sci.* **111**, 12055–12060 (2014).

17. Gianni, S., Morrone, A., Giri, R. & Brunori, M. A folding-after-binding mechanism describes the recognition between the transactivation domain of c-Myb and the KIX domain of the CREB-binding protein. *Biochem. Biophys. Res. Commun.* **428**, 205–209 (2012).

18. Dogan, J., Schmidt, T., Mu, X., Engström, Å. & Jemth, P. Fast Association and Slow Transitions in the Interaction between Two Intrinsically Disordered Protein Domains. *J. Biol. Chem.* **287**, 34316–34324 (2012).

19. Schreiber, G. & Fersht, A. R. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* **3**, 427–431 (1996).

20. Johnson, K. A., Simpson, Z. B. & Blom, T. Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **387**, 20–29 (2009).

21. Gianni, S., Dogan, J. & Jemth, P. Distinguishing induced fit from conformational selection. *Biophys. Chem.* **189**, 33–39 (2014).

22. Brüschweiler, S. *et al.* Direct Observation of the Dynamic Process Underlying Allosteric Signal Transmission. *J. Am. Chem. Soc.* **131**, 3063–3068 (2009).

23. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

24.   Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).

25.   Ernst, P., Wang, J., Huang, M., Goodman, R. H. & Korsmeyer, S. J. MLL and CREB Bind Cooperatively to the Nuclear Coactivator CREB-Binding Protein. *Mol. Cell. Biol.* **21**, 2249–2258 (2001).

26.   Law, S. M., Gagnon, J. K., Mapp, A. K. & Brooks, C. L. Prepaying the entropic cost for allosteric regulation in KIX. *Proc. Natl. Acad. Sci.* **111**, 12067–12072 (2014).

27.   Yang, K. *et al.* Structural basis for cooperative regulation of KIX-mediated transcription pathways by the HTLV-1 HBZ activation domain. *Proc. Natl. Acad. Sci.* **115**, 10040–10045 (2018).

28.   Wang, N. *et al.* Ordering a Dynamic Protein Via a Small-Molecule Stabilizer. *J. Am. Chem. Soc.* **135**, 3363–3366 (2013).

29.   Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proc. Natl. Acad. Sci.* **97**, 9367–9372 (2000).

30.   Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: Fragment-Based Drug Discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).

31.   Im, W., Lee, M. S. & Brooks, C. L. Generalized born model with a simple smoothing function. *J. Comput. Chem.* **24**, 1691–1702 (2003).

32.   Chen, J., Im, W. & Brooks, C. L. Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.* **128**, 3728–3736 (2006).

33.   MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

34.   Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151 (1999).

35.   Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl. Acad. Sci.* **104**, 8311–8315 (2007).

36.   Schrank, T. P., Bolen, D. W. & Hilser, V. J. Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. *Proc. Natl. Acad. Sci.* **106**, 16984–16989 (2009).

37.   Pomerantz, W. C. *et al.* Profiling the Dynamic Interfaces of Fluorinated Transcription Complexes for Ligand Discovery and Characterization. *ACS Chem. Biol.* **7**, 1345–1350 (2012).

38.   Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).

39.   Karpen, M. E., Tobias, D. J. & Brooks, C. L. Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV. *Biochemistry* **32**, 412–420 (1993).

40.   Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, (1995).

41.   Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

CHAPTER 3

Unexpected Specificity Underlying the Formation of Dynamic Transcriptional Protein-Protein Complexes

3.1 Abstract

In eukaryotic transcriptional regulation, a key functional event for gene activation is the formation of protein-protein interactions between DNA-bound transcriptional activators and transcriptional coactivators. However, the mechanisms by which transcriptional activators are recognized by partner coactivators are poorly defined. Biochemical and biophysical studies suggest that activators use nonspecific hydrophobic and electrostatic interactions to bind to coactivators, leading to disordered protein-protein complexes that are resistant to sequence changes in the activator. Here, we mechanistically dissect a set of disordered activator•coactivator complexes between the coactivator Med25 and the PEA3 family of activators. We show that slight sequence changes between PEA3 family members, while not significantly affecting overall affinity, redistribute the conformational ensemble of the activator•coactivator complex. The PEA3•Med25 ensembles are dictated by *specific* contacts between the disordered activator and the Med25 interface, and this specificity is facilitated by structural shifts of the coactivator binding surface. Taken together, this highlights the critical role coactivator plasticity plays in recognition of disordered activators, and indicates that molecular recognition models of disordered proteins must consider the ability of the binding partners to mediate specificity.

3.2 Introduction

It is universally accepted that the protein-protein interactions (PPIs) formed between transcriptional activators and coactivators play a critical role in gene expression; these PPIs underpin co-localization of transcriptional machinery components and stimulate transcription initiation.[1–8] From a mechanistic perspective, the prevailing view of activator-coactivator PPIs is that they are non-selective and, further, that the specificity necessary for appropriate gene expression comes from other sources such as activator-DNA interactions.[3,9–14] Indeed, there is considerable data seemingly consistent with activator•coactivator complexes forming via almost entirely nonspecific intermolecular interactions, from early experiments demonstrating that a wide range of natural and non-natural amphipathic molecules interact with coactivators to more recent structural studies indicating no fixed activator•coactivator binding mode.[9–11,15–18]

Nonspecific recognition models, while attractive in their simplicity, seemingly contradict the critical functional role that activator-coactivator PPIs play in gene expression. Additionally, there are lines of evidence indicating that transcriptional activators depend on interactions with specific activator binding domains (ABDs) of coactivators for function.[19–23] It is therefore an open question whether there are other molecular recognition mechanisms at play that could account for the broad landscape of functional activators that are also more consistent with the observed selectivity of activators in vivo. Given the fact that activator•coactivator complexes often represent promising therapeutic targets, developing a more detailed understanding of the molecular recognition mechanisms of these essential PPIs is also critical for the development of small molecule modulators.[24–26] Thus, we chose to take a critical look at these general recognition models by evaluating whether a set of representative dynamic activator•coactivator complexes form by nonspecific or specific mechanisms.

Towards this end, we mechanistically dissect the PPIs formed between the ABD of Mediator subunit Med25 and the amphipathic transcriptional activation domains (TADs) of the ETV/PEA3 family of Ets transcriptional activators (ETV1, ETV4, and ETV5).[23,27–29] Previous biophysical studies indicated that the interaction of Med25 with family member ETV5 appears to be a prototypical nonspecific TAD•ABD complex[27,29]: it occurs over a shallow surface, is driven by electrostatic and hydrophobic interactions, and forms a

dynamic complex that is recalcitrant to structure determination. Herein, data from complementary biophysical and structural approaches reveal that the conformational ensembles of PEA3•Med25 PPIs are strikingly sensitive to slight changes in TAD sequence, despite being dynamic complexes with several well-populated conformational sub-states at equilibrium. Furthermore, the mechanisms by which conformational sensitivity arises are complex and intertwined, involving the ability of ordered and disordered regions of the TAD to participate in finite sets of alternative interactions with the Med25 interface, as well as conformational changes in Med25 that remodel the TAD binding site. Together, these results reveal an unappreciated degree of specificity in the formation of activator•coactivator complexes that is in direct contrast to the prevailing nonspecific recognition models of these essential PPIs.[3,9,11] Critically, the data presented here indicate alternative mechanisms by which the exceptionally broad landscape of functional TAD sequences could arise, and implicate a need to develop strategies that enable quantitative dissection of activator•coactivator molecular recognition at a broad scale.

3.3 Results and Discussion

PEA3•Med25 PPIs as a Model System for Dynamic TAD•ABD Interactions

ETV/PEA3•Med25 interactions represent an ideal system to study dynamic TAD•ABD interactions for several distinct reasons. First, previous studies indicated the interaction of PEA3 family member ETV5 with the Med25 ABD appears typical of a dynamic TAD•ABD complex[27,29]: the binding surface is shallow, both acidic and hydrophobic amino acids of the TAD determine affinity and activity, and multiple bound conformational states were detected by both NMR and kinetic analyses. Second, our previous studies of the Med25•ETV5 complex showed that the bound conformational ensemble is directly accessible by transient kinetic analysis.[29] The models of TAD•ABD molecular recognition can thus be dissected in this system without the loss of critical conformational information to equilibrium averaging. Third, the sequence variation among PEA3 family members (Fig. 3.1A) serves as an excellent natural system to test how TAD•Med25 complexes respond to minor changes in the TAD sequence. This system can therefore be used to test whether TAD•Med25 interactions are truly nonspecific and insensitive to changes in

the TAD (i.e., if these PPIs form in a "sequence-independent" manner[11]), or if these interactions are affected by TAD sequence changes and thus have a degree of specificity to formation. Finally, the Med25 ABD is ligandable by small molecules,[29] and therefore conclusions from mechanistic studies can be directly applied to guide and assess optimization of small molecule modulators of TAD•Med25 complex formation.

Small Sequence Differences Between PEA3 Family Members Lead to Conformationally Distinct PPIs with Med25

Current models of TAD function dictate that nonspecific hydrophobic and electrostatic interactions dominate the binding mechanisms of TAD•ABD complexes.[3,11] The conformational ensembles of TAD•ABD complexes are therefore expected to be insensitive to sequence changes as long as the relative arrangement of acidic and hydrophobic residues are conserved.[10] The PEA3 family of transcription factors serve as an excellent model system to test this hypothesis: they contain almost identical arrangements of acidic and hydrophobic residues across the TAD sequence, especially within the helical binding region that undergoes coupled folding and binding with Med25,[27] but the identity of specific residues varies slightly (Fig. 3.1A).

A.

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

B. Med25

≥300 s⁻¹   60-72 s⁻¹   4.8-7.3 s⁻¹

+   PEA3 TAD   C1   C2   C3

C.

C1_ETV1 C2_ETV1 C3_ETV1   ETV1
C1_ETV5 C2_ETV5 C3_ETV5   ETV5   ETV4
C1_ETV4 C2_ETV4 C3_ETV4

= 100%

D.

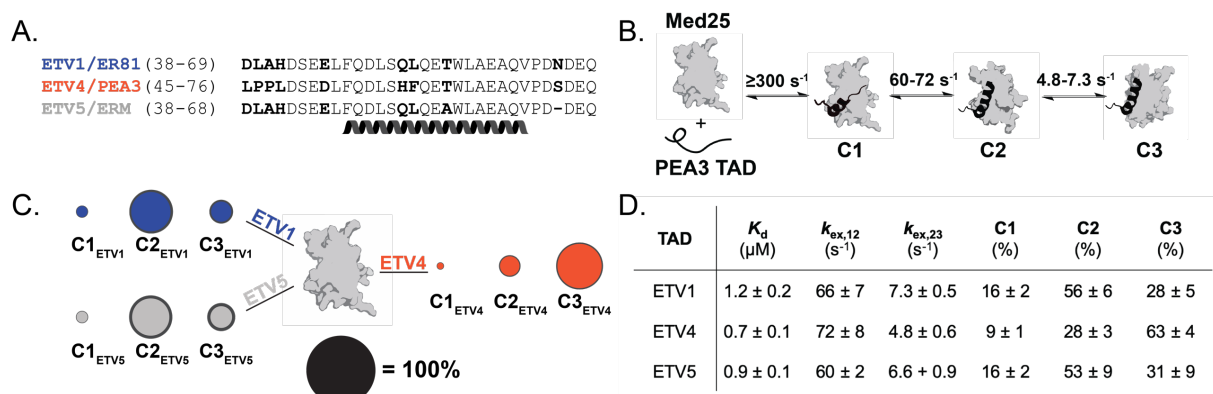| TAD	K_d (μM)	$k_{ex,12}$ (s⁻¹)	$k_{ex,23}$ (s⁻¹)	C1 (%)	C2 (%)	C3 (%)
ETV1	1.2 ± 0.2	66 ± 7	7.3 ± 0.5	16 ± 2	56 ± 6	28 ± 5
ETV4	0.7 ± 0.1	72 ± 8	4.8 ± 0.6	9 ± 1	28 ± 3	63 ± 4
ETV5	0.9 ± 0.1	60 ± 2	6.6 + 0.9	16 ± 2	53 ± 9	31 ± 9

Figure 3.1. PEA3 activators differentially engage with Med25. A) Alignment of PEA3 family activation domains. The helix denotes the residues that undergo coupled folding and binding with Med25, as determined by NMR chemical shift analysis.[27] B) Mechanism of binding of PEA3 activators to Med25,[29] determined here for ETV1 and ETV4, and previously for ETV5. The range of exchange rates between analogous steps for PEA3 TADs are shown. C) Equilibrium populations of the three Med25•ETV conformations calculated from microscopic equilibrium constants determined via kinetics experiments, scaled relative to the size of the black circle. All kinetics experiments were the average of 3-4 biological replicates, and the standard deviations of

the calculated values are shown as the dark gray outer circle. D) Table of relevant binding parameters for each PEA3 TAD, including the affinity, exchange rates between C1 and C2 ($k_{ex,12}$) and between C2 and C3 ($k_{ex,23}$), and equilibrium populations of each state. Populations were calculated as described in Materials and Methods, and equilibrium binding constants were determined from the endpoint fluorescence values of kinetics experiments.

We performed stopped-flow fluorescence experiments to directly measure the effects of these sequence differences on the conformational ensemble of individual PEA3•Med25 complexes, using TADs synthesized with the solvatochromic fluorophore 4-DMN conjugated to the N-terminus.[30] We previously demonstrated with this approach that the Med25•ETV5 complex forms in a three step linear mechanism[29]: after an initial rapid association event that occurs in the instrument dead-time (~2-4 ms), the complex undergoes two sequential conformational changes (Fig 3.1B). Results from identical experiments performed with ETV1 and ETV4 indicated that the kinetic binding mechanism is conserved: for all PEA3 TADs we observed a rapid binding step followed by two conformational change steps. Each of these individual steps occurred with similar rates of exchange for each PEA3•Med25 complex (Fig. 3.1B,D, Fig. 3.2), suggesting that they represent analogous conformational transitions in each complex. In addition, the equilibrium binding affinity between PEA3 TADs varied less than two-fold (0.7–1.2 µM, Fig. 3.1D), which is consistent with the expectation from nonspecific models that minor substitutions in the TAD will not affect the overall stability of activator•coactivator complex.[10]
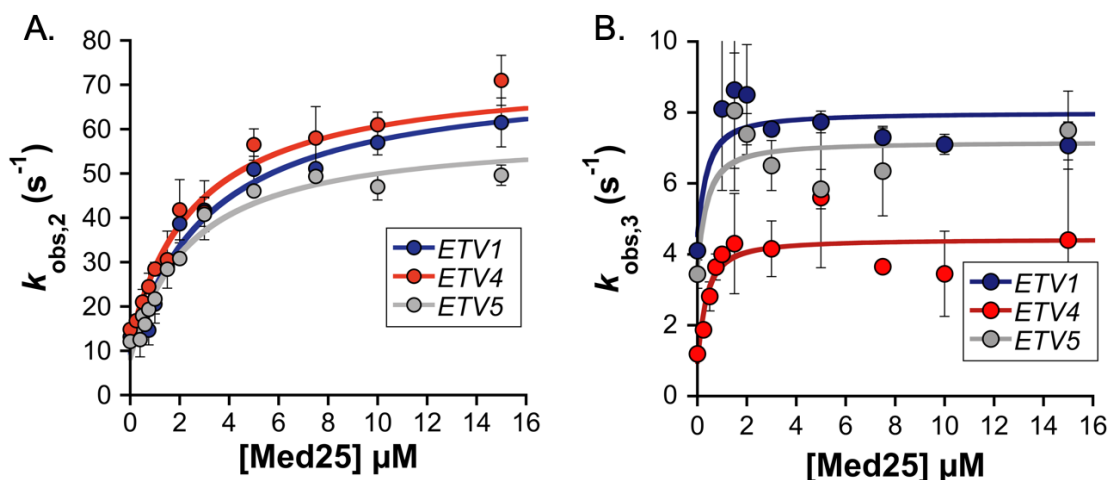
Figure 3.2. The k_{obs} values of individual conformational change phases saturate at similar values across PEA3 TADs. A) Concentration curves for $k_{obs,2}$ or the first conformational change phase. B) Concentration curves for $k_{obs,3}$ or the second conformational change phase. The values for the Y-intercept were taken from dissociation experiments. Individual values are the average of 3-4 independent experiments, with the error as standard deviation from the mean. Data are fitted to square hyperbola.

However, despite a conserved binding mechanism and similar overall affinities, calculation of equilibrium conformational populations from the kinetic data revealed clear differences between the engagement modes of PEA3 family members (Fig. 3.1C; for raw data and detailed kinetic analysis, see SI discussion). While the populations of analogous conformational states of the ETV1•Med25 and ETV5•Med25 complexes were essentially identical, a four-fold larger C3:C2 ETV4•Med25 sub-state ratio (P < 0.001, Student's t-test) leads to a distinct conformational ensemble at equilibrium (Fig. 3.1C,D). Critically, this change in conformational ensemble did not correlate with predicted structural propensity differences between the PEA3 TADs (Fig. 3.3), which suggests that variable residues between ETV1, ETV4, and ETV5 affect the resulting TAD•Med25 conformational ensemble via intermolecular interactions made in the bound state.
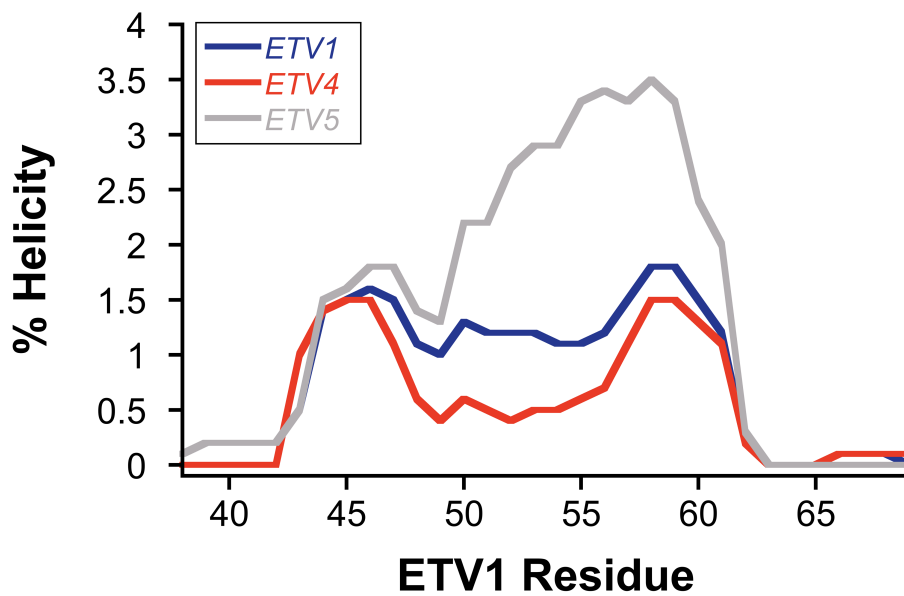
Figure 3.3. Predicted helical propensity of PEA3 activators from Agadir, shown by ETV1 residue.

We next sought to obtain structural evidence of this result via NMR spectroscopy. Side-chain methyl ^{1}H,^{13}C-HSQC experiments were used as a primary method to enable direct detection of effects on surface (engagement of TADs) and buried residues (structural shifts) of Med25 (example spectra of free Med25 and Med25 bound to ETV4 shown in Fig. 3.4A). Comparative analysis of Med25 ^{1}H,^{13}C-HSQC spectra bound to different PEA3 family members was wholly consistent with the expected differences between the PEA3 TADs: ETV1- and ETV4-bound Med25 spectra exhibited several large chemical shift differences, whereas the spectra of Med25 bound to ETV1 and ETV5 were essentially indistinguishable (Fig. 3.4B–E). Inspection of the CSP data plotted on the structure of Med25 indicated that all PEA3 family members bind to a previously identified core binding site[27–29] formed between the central β-barrel and the *C*-terminal helix (Fig. 3.4B,C), but ETV1/ETV5 and ETV4 produce unique structural signatures in the binding surface (Fig. 3.4B,C, cyan circles). In addition, several resonances representing buried and/or allosteric residues displayed significant chemical shift differences between the ETV1- and ETV4-bound complexes, suggesting that the conformation of the Med25 ABD may also be different between these complexes (Fig. 3.4E). Together, the HSQC data supports the conclusion from kinetics experiments that ETV4 has a unique Med25 engagement mode compared to ETV1 and ETV5.
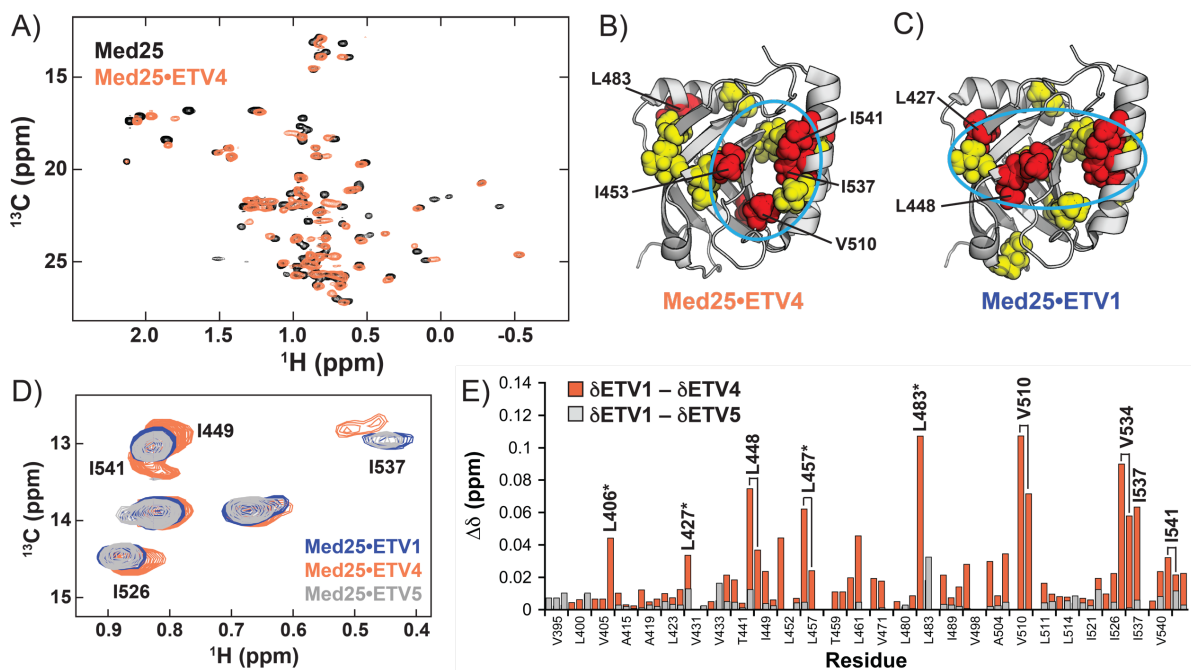
Figure 3.4 Side-chain methyl $^1H,^{13}C$-HSQC experiments demonstrate distinctions between PEA3•Med25 complexes. A. Overlay of $^1H,^{13}C$-HSQC of 60 μM free Med25 (black) and Med25 in complex with 1.1 equivalents unlabeled ETV4 (orange). Chemical Shift perturbations induced by B. ETV4 and C. ETV4 plotted on the structure of Med25 (PDB ID 2XNF)[31]. Yellow = 0.04–0.08 ppm, red > 0.08 ppm. Cyan circles highlight distinctions in perturbation patterns. D) Overlay of Ile Cδ region of Med25 $^1H,^{13}C$-HSQC spectra demonstrating differences between chemical shifts between ETV1•Med25 (blue), ETV4•Med25 (orange) and ETV5•Med25 (gray) complexes. Labeled residues have measurable chemical shift differences between ETV1/ETV5 and ETV4. Note: I449 and I541 form a single overlapped peak for the ETV1 and ETV5 complexes. E) Bar graph of chemical shift differences between Med25 bound to ETV1 vs. ETV4 (orange) or ETV1 vs. ETV5 (gray). Residues of interest with significant differences in chemical shift are labeled. When both diastereotopic methyl peaks for Leu/Val residues of interest undergo large shifts, the individual bars are connected by lines. Asterices indicates residues that are buried or partially buried in the structure of unbound Med25. HSQC experiments were completed with the assistance of Dr. Brian Linhares.

Conformational Ensembles of PEA3•Med25 PPIs are Dictated by Ordered and Disordered Regions of the PEA3 TAD

Next, we turned our attention to identifying the molecular driving forces of the distinct PEA3•Med25 conformational ensembles. We first used a mutagenesis approach to examine the precise differences between the ETV4 and ETV1/ETV5 sequences that lead to the formation of distinct conformational ensembles. This effort focused on distinct

residues that are conserved between ETV1 and ETV5, but not ETV4. Two regions of interest were indicated by these parameters (Fig. 3.5A, boxed): 1) a two amino acid 'variable motif' in the helical binding region consisting of a polar residue followed by a hydrophobic residue (QL in ETV1/ETV5 and HF in ETV4), and 2) the four amino acids at the disordered *N*-terminus of the TAD (DLAH in ETV1/ETV5 and LPPL in ETV4). A small library of mutant TADs based on the variations of the ETV4 scaffold in the two regions of interest—either the LPPL or DLAH *N*-terminus with combinations of His or Gln and Phe or Leu residues in the helical binding region—were synthesized and assessed in stopped-flow kinetic assays (Fig. 3.5B). Importantly, the equilibrium binding affinity varied only slightly across these variants (Fig. 3.5D), again demonstrating that these residue differences were not significantly affecting overall binding affinity.

Consistent with the hypothesis that one or both of these regions dictate the conformational differences between native PEA3•Med25 complexes, PEA3 variants with 'native' combinations of the *N*-terminus and the hydrophobic residue of the 'variable motif' retained conformational ensembles indistinguishable to that of the native TADs (Fig. 3.5B). That is, combinations with DLAH at the *N*-terminus and Leu in the variable motif (DLAH/L) exhibited similar conformational populations to ETV1 and ETV5, whereas combinations with the LPPL *N*-terminus and Phe in the variable motif (LPPL/F) displayed similar ensembles to ETV4 (Fig. 3.5B, left). In these and all other variants we tested, the polar residue in the variable motif had no effect on conformational populations besides in some cases affecting the rates of exchange (see Table 3.2 in Materials and Methods).
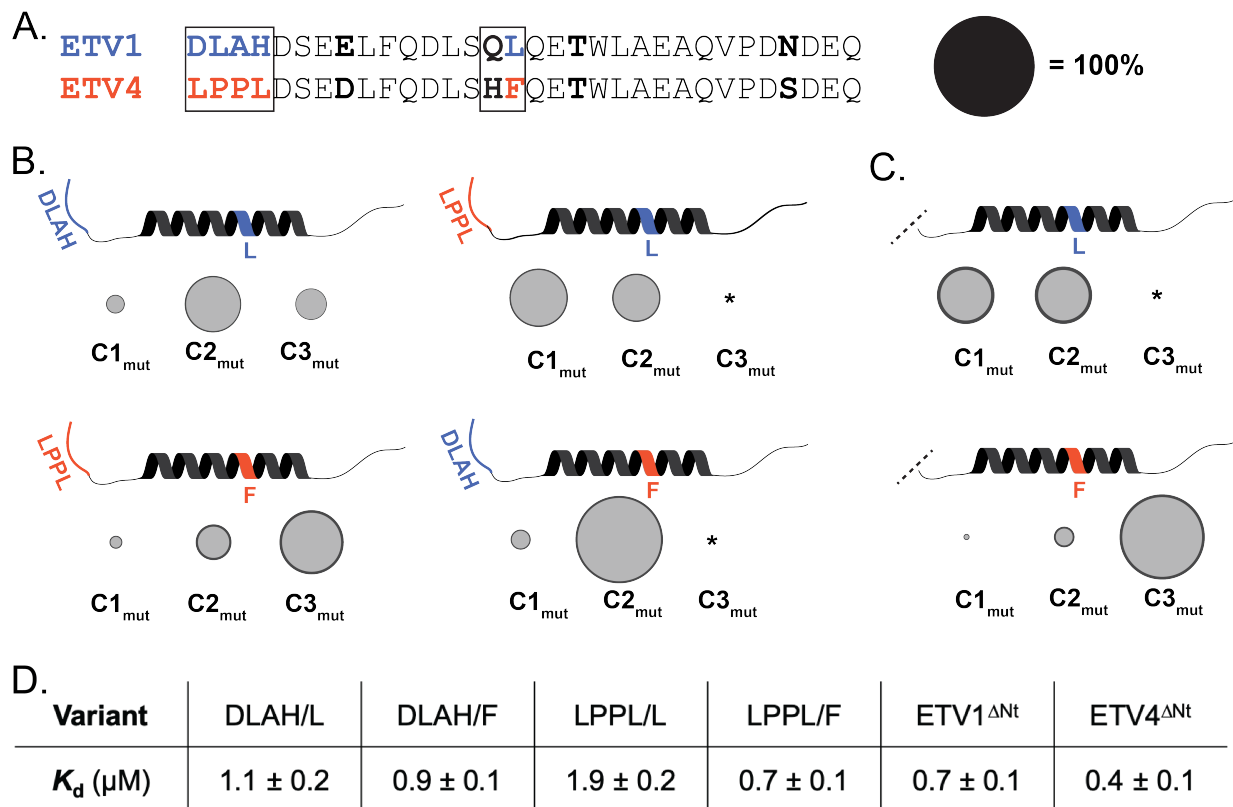
A.

ETV1 **DLAH**DSE**E**LFQDLS**QL**QE**T**WLAEAQVPD**N**DEQ
ETV4 **LPPL**DSE**D**LFQDLS**HF**QE**T**WLAEAQVPD**S**DEQ

⬤ = 100%

B.

C.

D.

Variant	DLAH/L	DLAH/F	LPPL/L	LPPL/F	ETV1$^{\Delta Nt}$	ETV4$^{\Delta Nt}$
K_d (µM)	1.1 ± 0.2	0.9 ± 0.1	1.9 ± 0.2	0.7 ± 0.1	0.7 ± 0.1	0.4 ± 0.1

Figure 3.5. Variable residues in the disordered *N*-terminus and the helical binding region mediate differences in PEA3•Med25 conformational ensembles. A) Alignment of ETV1 and ETV4 activators with regions that were selected for mutational analysis boxed. Regions/residues that affected the conformational ensemble are color coded to ETV1 (maroon) or ETV4 (blue). Populations of conformational states in B) and C) are scaled relative to the size of the black circle. B) Results from kinetics experiments of mutant TADs, for native (left) and non-native (right) combinations of variable *N*-termini and helical binding regions. Variants were made based on the ETV4 scaffold. The data shown is the average across all the variants tested from each group, with the error representing the standard deviation. C) Results from kinetics experiments with ETV1$^{\Delta Nt}$ (left) and ETV4$^{\Delta Nt}$ (right). D) Average equilibrium K_d values of variants tested. *Conformer was undetectable in kinetics experiments.

On the other hand, when nonnative combinations of the *N*-terminus and variable motif were tested, unique conformational ensembles were observed (Fig. 3.5B, right). Only two bound conformations of DLAH/F variants were detected in kinetics experiments, and displayed exchange kinetics similar to the C1–C2 transitions of the native complexes. Calculation of conformational populations indicated that the second conformation had a higher overall population (82%) than the C2 conformations of the ETV1/ETV5 (58%) or ETV4 complexes (28%). Similarly, just two bound conformations were detectable with

LPPL/L variants, but the initial bound conformation was preferentially populated (~55%) Together, these results indicate a model where both the hydrophobic residue in the variable motif and the disordered *N*-terminus dictate the conformational differences between PEA3•Med25 complexes. The latter result is particularly striking, as disordered regions of the TAD are often removed or ignored in biophysical and structural studies because they typically don't contribute to overall affinity.

To obtain further evidence for the unexpected effects of the disordered *N*-terminus on PEA3•Med25 conformational ensembles, we also tested whether removing the four variable N-terminal residues affected the conformational ensembles of ETV1 and ETV4. The resulting variants ETV1$^{\Delta Nt}$ and ETV4$^{\Delta Nt}$ both induced detectable changes from the parent ensembles (Fig. 3.5C), in addition to a slight (~1.7-fold) gain in affinity for both variants. Kinetic analysis indicated that the ETV1$^{\Delta Nt}$•Med25 complex exchanged between two equally populated conformations on a similar timescale to the C1–C2 transition of the parent ETV1•Med25 complex, similar to the LPPL/L variants in Figure 3.5B. On the other hand, the ETV4$^{\Delta Nt}$•Med25 retained a similar overall ensemble to the parent ETV4•Med25 complex that was further biased towards the C3 conformation (79% vs. 63%). Thus, our results fully support a direct role of the *N*-terminal residues of the TAD in shaping PEA3•Med25 conformational ensembles.

Variable Regions of PEA3 TADs Differentially Engage with the Med25 Surface

We next pursued an NMR strategy to probe the structural basis by which these two key regions in the PEA3 TADs modulate the bound PEA3•Med25 conformational ensembles. We utilized a "soft" mutagenesis approach,[32] where minimally perturbing mutations were first individually introduced into unlabeled TADs, and then CSP analysis of ^{1}H,^{13}C-HSQC spectra of Med25 bound to the native or mutated TAD was performed to detect Med25 methyl groups affected by the mutation. Structural analysis of these changes in the TAD•Med25 HSQC spectrum can therefore detect Med25 residues in direct proximity to the mutated site in the complex, and also has the potential to reveal allosteric connections if the effects of the soft mutation are propagated from the interaction site.[32] Furthermore, this strategy avoids the significant experimental challenge associated with NMR detection of Med25-bound PEA3 TADs, which have been observed

to undergo severe line broadening.[27] Here, soft mutations were introduced into the two key variable regions of the ETV1 and ETV4 TADs to detect differences in engagement modes that could explain how these regions cause differences in PEA3•Med25 conformational ensembles.
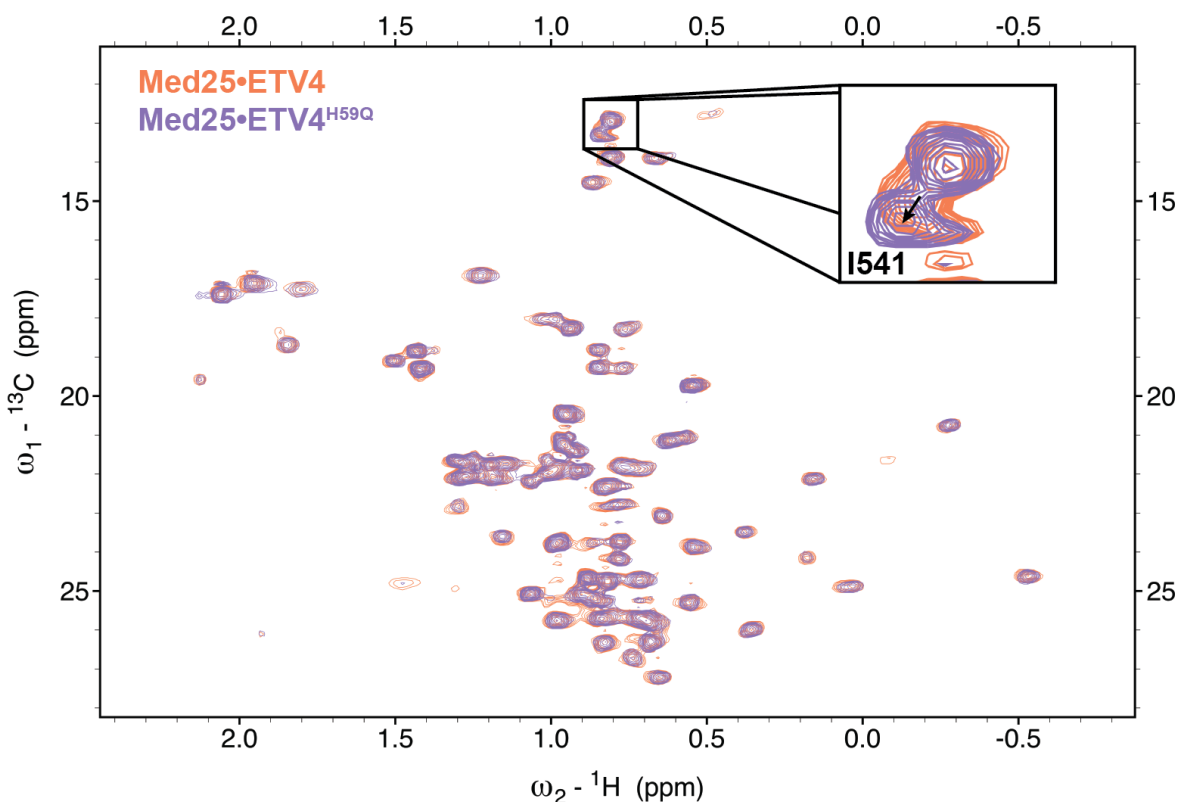

Figure 3.6. Example of the effect of soft mutation on ^{1}H,^{13}C-HSQC spectra of Med25. Orange: ETV4•Med25 spectrum; Purple: ETV4^{QF}•Med25 (H59Q) spectrum. Inset shows the large chemical shift induced for residue Ile541. The contour levels of the ETV4^{QF}•Med25 spectrum are scaled to 80% of the ETV4•Med25 spectrum. HSQC experiments were completed with the assistance of Dr. Brian Linhares.

Mutations were first made within the variable motif of the helical binding region (normally QL in ETV1 and HF in ETV4) by swapping the variable polar residue between ETV1 and ETV4 to form ETV1^{HL} and ETV4^{QF}, based on the observation that this change did not affect the conformational ensemble in kinetics experiments (Table 3.3 in Materials and Methods). Indeed, the ^{1}H,^{13}C-HSQC spectra of the mutant PEA3•Med25 complexes were almost identical to the native complexes except for shifts in single methyl peaks (Fig.

3.7A): the ETV4^{QF} variant produced a large shift (0.032 ppm) in the Med25 Ile541δ peak, whereas the analogous ETV1^{HL} variant produced a smaller perturbation (0.015 ppm) in the Ile541γ peak. These highly localized shifts are consistent with the soft mutations causing proximity-based perturbations near the mutation site in the complex. In addition, the fact that the individual mutations perturb unique methyl groups originating from the same residue suggest that the native ETV1 QL and ETV4 HF motifs are engaged in unique interactions in a similar interface.
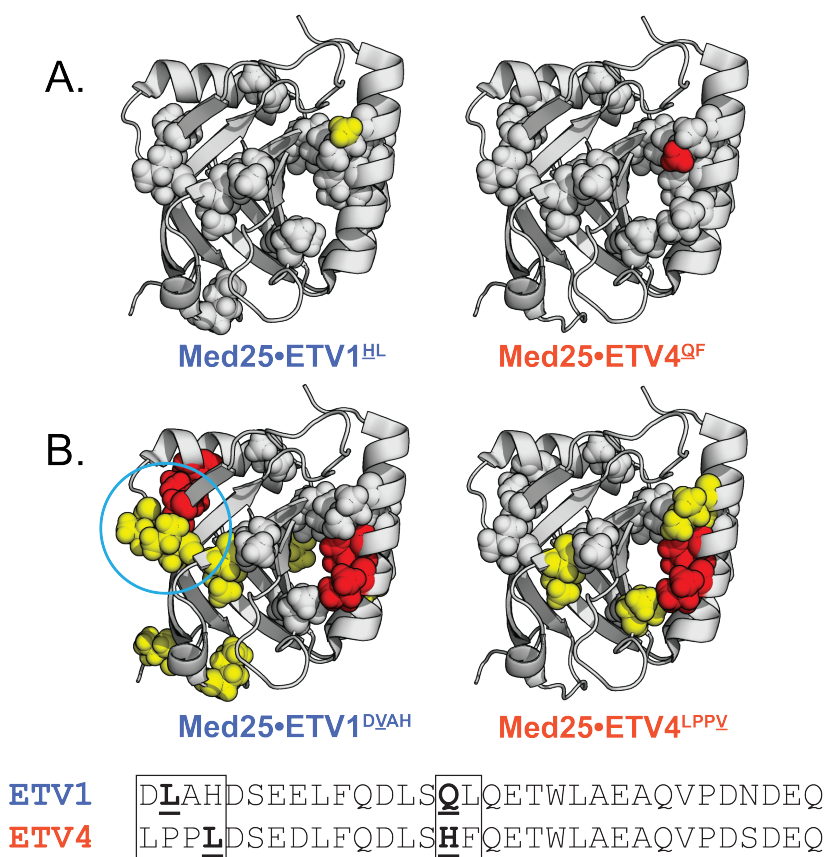
Figure 3.7. PEA3 variable regions engage in unique interactions with the Med25 surface. Effects of soft mutations in the A) helical binding region and B) *N*-termini are plotted on the structure of Med25. Yellow = 0.015 – 0.03 ppm. Red ≥ 0.03 ppm. Grey spheres denote residues that do not change from the parent complex. Below is ETV1 and ETV4 alignment where underlined are residues chosen for soft mutations presented in A. and B. HSQC experiments were completed with the assistance of Dr. Brian Linhares.

Next, we introduced soft mutations into the disordered *N*-terminus (normally DLAH in ETV1 and LPPL in ETV4). Specific point mutations in this region were not immediately

obvious from previous data, so we made several point mutations and found that minimal Leu to Val mutations ETV1^{DVAH} and ETV4^{LPPV} sufficed to produce measurable effects without significantly altering the overall bound spectra (Fig. 3.7B). In contrast to soft mutations in the helical binding region, both of these variants affected a broader distribution of Med25 residues. Both variants affected a similar sub-set of residues in the core binding site, which likely is an indirect effect because this site would be occluded by the TAD helical binding region and thus indicates an allosteric linkage between the TAD *N*-termini and the engagement mode at the Med25 core binding interface. Evidence for a direct interaction of the ETV1 *N*-terminus was also apparent from the ETV1^{DVAH}•Med25 spectrum: this variant provoked a cluster of unique shifts in a distal site including the residues Val405, Leu427, Thr476, and Leu483 (Fig. 3.7B, cyan circle). However, this site was unaffected by the ETV4^{LPPV} variant, suggesting that the unique effects on conformation from these two *N*-terminal sequences may be due to differential engagement at this site. To obtain further support for this interpretation, we synthesized ETV1 and ETV4 TADs that were selectively ^{15}N labeled at the Leu residues previously selected for soft mutations in the *N*-terminus, and analyzed CSPs upon binding to Med25. Critically, Leu39 in the ETV1 *N*-terminus underwent a relatively large (~0.2 ppm) shift whereas Leu48 in the ETV4 *N*-terminus shifted only slightly (Fig. 3.8). Interestingly, we also observed a slightly shifted minor peak in spectra of bound ETV1 Leu39, which is consistent with a secondary bound conformation where this region is weakly engaged with Med25.
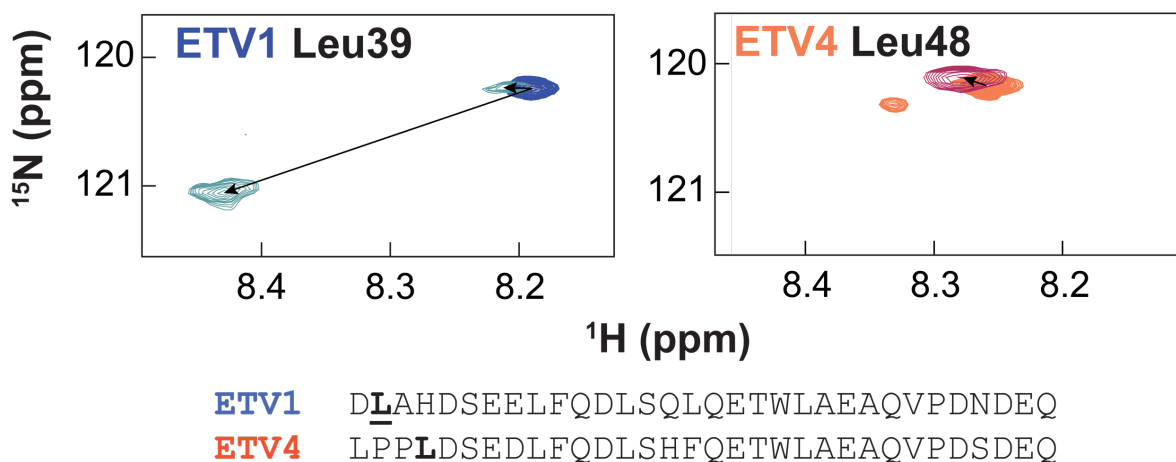
Figure 2.8. Chemical shift perturbations of 150 μM ETV1 and ETV4 TADs in the absence (blue and orange, respectively) and presence (light blue and maroon, respectively) of 280 μM Med25. TADs were selectively ^{15}N labeled at the positions underlined in the alignment. Small secondary peaks in free ETV4 spectra were observed and presumably arose to isomerization of the two tandem Pro residues in the *N*-terminal region.

PEA3•Med25 Conformational Changes Involve Shifts in Med25 Structure

Altogether, our data strongly support a model where the ensemble differences between PEA3•Med25 complexes are caused by PEA3 variable regions engaging with the Med25 surface in unique and sequence-dependent manners. Interestingly, the mode of differential engagement is not apparently due to the PEA3 helical binding regions adopting highly distinct orientations in the core binding site, as NMR analysis indicated that the ETV1 and ETV4 variable motifs localize to a similar site (Fig. 3.7A). We thus hypothesized that unique engagement in a similar bound orientation would be facilitated by remodeling of the Med25 ABD. To test this hypothesis, we used NMR to examine changes in the bound Med25 conformation that occur when an ensemble-redistributing point mutation is introduced into the partner TAD. The ETV4[F60L] point mutation (Fig. 3.5B, LPPL/L variant) was selected for this analysis because this small change in residue identity caused a drastic change in the conformational ensemble measured in kinetics experiments (compare Fig. 3.5B bottom left and bottom right) towards the initial bound state (C1) that has not undergone detectable conformational change.

Comparison of the ^{1}H,^{13}C-HSQC spectra of ETV4•Med25 and ETV4[F60L]•Med25 revealed several Med25 peaks that undergo large perturbations upon binding to native

ETV4 shift towards the unbound position in the ETV4[F60L]•Med25 spectrum (Fig 3.9A). Significantly, the ETV4[F60L] variant elicited weaker CSPs around the central binding site than native ETV4 (Fig. 3.9B), suggesting that the large CSPs in this region are tied to the conformational changes. Furthermore, this behavior was observed for peaks representing residues that are buried or in allosteric regions of the protein, including the interface between the β-barrel and the allosteric helix 2 (residues Leu427, Ile489 in Fig. 3.9B), the β-barrel core (residue Leu513), and the interface between the *C*-terminal helix and the allosteric helix 1 (residue Leu457). Thus, our evidence is fully consistent with significant remodeling of the Med25 ABD playing a role in molecular recognition, which likely serves to reveal new topology in the core binding site to enable the TAD to access handles for more specific recognition.
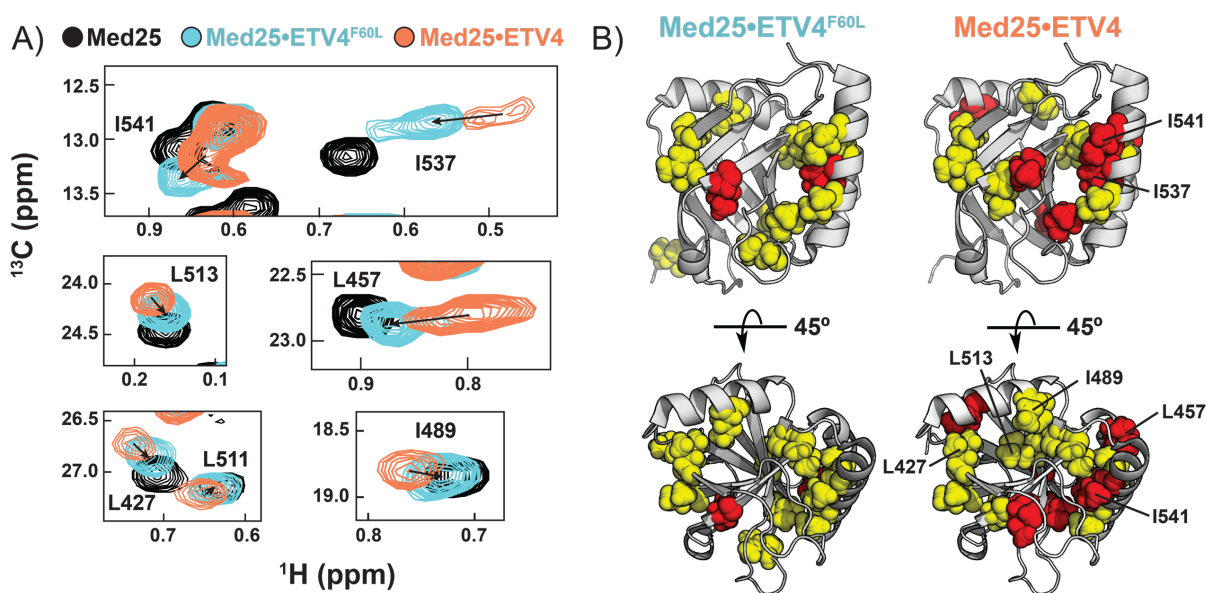
Figure 3.9. Sidechain methyl ^{1}H,^{13}C-HSQC NMR experiments demonstrate structural shifts in the Med25 ABD originating from PEA3•Med25 conformational exchange. A) Examples of resonances undergoing shifts towards unbound state upon ETV4[F60L] (cyan) mutation. B) Chemical Shift Perturbations from binding of ETV4[F60L] (left) and ETV4 (right). Residues shown in A. are labeled on the structure. Colors are as in Fig. 3.4. HSQC experiments were completed with the assistance of Dr. Brian Linhares.

3.4 Conclusions

The original functional studies of TADs conducted several decades ago represent some of the earliest realizations that intrinsically disordered proteins (IDPs) could be functional.[1,2,33,34] Initial molecular recognition models of TADs that emerged from these studies fixated on these IDPs mediating function through nonspecific or promiscuous formation of PPIs.[3] While the general understanding of the molecular factors that drive TAD function has expanded from the original "negative noodle" hypothesis, nonspecific molecular recognition continues to form the basis of how TADs and many other IDPs are thought to function.[11,33] This recognition model represents a major reason why many PPIs involving IDPs have been traditionally considered "untargetable" by small molecule modulators.[26] Thus, we chose to scrutinize the general molecular recognition model of TAD•ABD interactions by subjecting a model set of dynamic TAD•ABD interactions to detailed mechanistic and structural dissection.

Here, we investigated the PPIs formed between the PEA3 family of TADs and their binding partner Med25. On inspection of previously published data, these interactions fit the characteristics of prototypical "nonspecific" TAD•ABD PPIs: they occur over a flat surface, form multiple conformations, and rely on electrostatic and hydrophobic interactions for binding affinity.[27–29] Upon mechanistic analysis, however, we found that these interactions exhibited a striking degree of conformational sensitivity to small changes in TAD sequence, which indicated a degree of specificity to complex formation. This specificity originates in part from the plasticity of the apparently "featureless" binding interface, which remodels itself after an initial binding event to enable the TAD to access hidden molecular recognition features and form a relatively specific complex. Following binding site remodeling, the TAD transitions between two distinct conformations in the binding site, and the relative stability of these conformations is controlled by finite sets of alternative interactions of *both* the previously identified helical binding region[27] and a disordered region at the *N*-terminus of the TAD. Thus, while the interactions between Med25 and the PEA3 TADs appeared by all other means to be characteristic nonspecific TAD•ABD complexes, our mechanistic data revealed significant specificity in these PPIs that is invisible to the typical observational approaches that are employed to characterize PPIs formed by IDPs.

This structural analysis strongly suggests that the relative specificity of the formation of PEA3•Med25 complexes is related to the plasticity of the Med25 ABD. The plasticity of folded ABDs is often ignored in structural and mechanistic studies of TAD•ABD complexes,[9–11,35,36] which is likely because the conformational changes of folded partners can seem inconsequential compared to the large structural changes the TAD undergoes during the folding and binding process. Nonetheless, comparisons of the structures of the limited ABDs that have been solved in complex with multiple unique TADs have indicated that ABDs undergo substantial conformational remodeling upon binding different TADs, often having significant functional consequences.[37–40] Thus, a major question that emerges from this study is whether the alleged "sequence-independent"[11] behavior of TADs is due to the plasticity of ABDs in the transcriptional machinery, rather than wholly nonspecific TAD binding mechanisms. This prospect would significantly aid the development of small molecule modulators of TAD function: if specific ABD conformations are required for efficient TAD binding, small molecules that stabilize binding competent or incompetent conformations could affect function without needing to target the topologically challenging TAD binding interfaces of ABDs.[29,40,17,18,41]

Taken together, this study implicates a broader need to consider IDP function beyond merely the sequence and structural propensities of a given IDP, but how the IDP is recognized by its binding partners. That is, a fundamental biological question that emerges from our study is how molecular recognition mechanisms affect function. In general, it is still exceptionally challenging to predict and understand what interactions will be functional: in processes where dynamic PPIs serve as critical functional events, such as transcription and proteostasis, it is often observed that affinity does not correlate to functional activity.[42,43] Several factors play into this observation, such as subcellular localization and concentration, but a potentially significant factor is the mechanism by which the complex forms. For example, in our current study we observed that the lifetimes of individual PEA3•Med25 conformations varied up to two orders of magnitude (Table 3.2), therefore shifting of the conformational ensemble towards longer-lived sub-states (and thereby extending the average lifetime of the complex as a whole) could have significant functional outcomes without necessitating changes in affinity. Thus, there is a critical need to understand the relationship between the function of biomolecular

interactions and the mechanisms by which they are formed. This issue is especially acute for understanding the roles of dynamic biomolecules like IDPs and non-coding RNAs, which can often make numerous disparate interactions that individually are of unclear functional importance.

3.5 Materials and Methods

Protein Expression and Purification

Med25 AcID was expressed and purified from a pET21b-Med25(394-543)-His$_6$ plasmid from E. coli BL21 (DE3) cells as described previously.[29] Briefly, 50 mL starter cultures in LB were grown overnight in the presence of 0.1 mg/mL ampicillin (Gold Bio Technology) at 37 °C at 150 rpm. The next day, 5 mL of the starter culture was used to inoculate 1 L of TB media (24 g yeast extract, 12 g tryptone, 4 mL glycerol, 100 mL 0.17 M KH$_2$PO$_4$/0.72 M K$_2$HPO$_4$, 900 mL water) with 0.1 mg/mL ampicillin, which was grown at 37 °C, 250 rpm, to an OD$_{600}$ of 0.8. The incubator temperature was lowered to 21 °C and the culture was allowed to recover for 30 min, at which point isopropyl β-D1-thiogalactopyranoside (IPTG, Research Products International) was added to a final concentration of 500 µM to induce expression. The protein was allowed to express overnight (~18 hr), after which the cells were harvested via centrifugation (6000 rpm, 20 min), and then frozen and stored at –80 °C. Uniformly ^{13}C,^{15}N labeled Med25 for NMR experiments was expressed identically except using M9 minimal media supplemented with 1 g/L ^{15}NH$_4$Cl, 2 g/L ^{13}C-$_D$-glucose, and 0.5% ^{13}C,^{15}N-labelled Bioexpress media for the 1 L growth (all labeled components were purchased from Cambridge Isotopes).

To purify Med25, cell pellets were resuspended in 25 mL lysis buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.2, 1.4 µL/mL β-mercaptoethanol, 1 Roche complete mini protease inhibitor tablet) and lysed by sonication. Insoluble material was then pelleted by centrifugation (9500 rpm, 20 min), the supernatant was removed and re-sonicated, and then filtered using a 0.45 syringe filter (CellTreat) and loaded onto an AKTA Pure FPLC equipped with a Ni HisTrap HP column (GE Healthcare) pre-equilibrated with wash buffer (50 mM phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.2). Med25 was then purified using a gradient of 10–300 mM imidazole (other buffer components were constant), and fractions containing Med25 were pooled and subjected

to secondary purification using a HiTrap SP HP cation exchange column (GE Healthcare) using a gradient of 0–1 M NaCl (50 mM sodium phosphate, 1 mM DTT, pH 6.8). Pooled fractions were dialyzed into stopped-flow buffer (10 mM sodium phosphate, 100 mM NaCl, 1% glycerol, 0.001% NP-40, pH 6.8) or NMR buffer (20 mM sodium phosphate, 150 mM NaCl, pH 6.5). Concentration was determined by a NanoDrop instrument using an extinction coefficient at 280 nm of 22,460 $M^{-1}cm^{-1}$. Aliquots were flash frozen in liquid N_2 and stored at –80 ºC until use. Protein identity was confirmed by mass spectrometry (Agilent Q-TOF).

Peptide Synthesis

The peptides used in this study were prepared using standard Fmoc solid-phase peptide synthesis on a Liberty Blue Microwave Peptide Synthesizer (CEM). Deprotection was accomplished by 20% piperidine (ChemImpex) in DMF supplemented with 0.2 M Oxyma Pure (CEM), with irradiation at 90 ºC for 1 min. Coupling reactions were completed with 5 equivalents of Fmoc-amino acid (CEM), 7 equivalents of diisopropylcarbodiimide (ChemImpex), and 5 equivalents of Oxyma Pure in DMF, with irradiation at 90 ºC for 4 min. Between steps, the resin was rinsed four times with an excess of DMF. Selectively ^{15}N-labeled peptides were synthesized using Fmoc-^{15}N-Leu in place of Fmoc-Leu in the specified positions.

Unlabeled peptides were acetylated at the *N*-terminus through reaction of a mixture of acetic anhydride (Sigma), triethylamine (Sigma), and dichloromethane in a 1:1:8 ratio for 30 min after the conclusion of the synthesis. 4-DMN labeled peptides were coupled with ~1.5 equivalents of 4-DMN-β-Alanine, as described previously.[29] All peptides were cleaved from the resin using a cocktail of 95:2.5:2.5 trifluoroacetic acid (Sigma), ethanedithiol (Sigma), and water for 3 hr, followed by filtration. The peptide solution was concentrated under a stream of N_2, then precipitated with cold diethyl ether and pelleted by centrifugation (4,500 rpm, 5 min). The ether was then discarded and the pellet was taken up in 7:3 100 mM Ammonium acetate-acetonitrile. The peptide was purified by an Agilent 1260 preparatory HPLC using a 40 minute gradient of 10-50% acetonitrile, with 100 mM ammonium acetate as the stationary phase. The flow rate was 40 mL/min. Fractions containing the correct peptide were pooled and lyophilized, and the resulting

powders were dissolved in minimal DMSO and stored at –20 ºC. Concentrations were taken with a 1:100 dilution of the DMSO stock into 6 M Guanidinium Chloride on a NanoDrop instrument, using extinction coefficients of 5,690 $M^{-1}cm^{-1}$ (280 nm, unlabeled) or 10,800 $M^{-1}cm^{-1}$ (450 nm, 4-DMN-labeled). Identity of the peptides were confirmed by mass spectrometry.

Table 3.1 Sequences of peptides used for this study

Peptide	Sequence
ETV1	Ac-DLAHDSEELFQDLSQLQETWLAEAQVPDNDEQ
ETV4	Ac-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
ETV5	Ac-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
4-DMN-ETV1	4-DMN-βAla-DLAHDSEELFQDLSQLQETWLAEAQVPDNDEQ
4-DMN-ETV4	4-DMN-βAla-LPPLDSEDLFQDLSHFQETWLAEAQVPDSDEQ
4-DMN-ETV5	4-DMN-βAla-DLAHDSEELFQDLSQLQEAWLAEAQVPDDEQ
4-DMN-ETV4$^{LPPL/HL}$	4-DMN-βAla-LPPLDSEDLFQDLSHLQETWLAEAQVPDSDEQ
4-DMN-ETV4$^{LPPL/QL}$	4-DMN-βAla-LPPLDSEDLFQDLSQLQETWLAEAQVPDSDEQ
4-DMN-ETV4$^{DLAH/HL}$	4-DMN-βAla-DLAHDSEDLFQDLSHLQETWLAEAQVPDSDEQ
4-DMN-ETV4$^{DLAH/HF}$	4-DMN-βAla-DLAHDSEDLFQDLSHFQETWLAEAQVPDSDEQ
4-DMN-ETV4$^{DLAH/QF}$	4-DMN-βAla-DLAHDSEDLFQDLSQFQETWLAEAQVPDSDEQ
4-DMN-ETV1$^{\Delta Nt}$	4-DMN-βAla-DSEELFQDLSQLQETWLAEAQVPDNDEQ
4-DMN-ETV4$^{\Delta Nt}$	4-DMN-βAla-DSEDLFQDLSHFQETWLAEAQVPDSDEQ
ETV4^{F60L}	Ac-LPPLDSEDLFQDLSHLQETWLAEAQVPDSDEQ
ETV1^{HL}	Ac-DLAHDSEELFQDLSHLQETWLAEAQVPDNDEQ
ETV4^{QF}	Ac-LPPLDSEDLFQDLSQFQETWLAEAQVPDSDEQ
ETV1^{DVAH}	Ac-DVAHDSEELFQDLSQLQETWLAEAQVPDNDEQ
ETV4^{LPPV}	Ac-LPPVDSEDLFQDLSHFQETWLAEAQVPDSDEQ

Stopped-flow kinetics

Stopped-flow kinetic assays were performed using a Kintek SF-2001 stopped flow instrument equipped with a 100-W Xe arc lamp in two-syringe mode. All experiments were completed at 10 ºC in stopped-flow buffer (10 mM sodium phosphate, 100 mM NaCl, 2% DMSO, 1% glycerol, 0.001% NP-40, pH 6.8). All concentrations reported are after mixing. The 4-DMN fluorophore was excited at 440 nm, and fluorescence intensity was measured at wavelengths >510 nm using a long-pass filter (Corion). Association experiments were completed by 1:1 mixing of a constant concentration of 0.25 µM 4-DMN-labeled peptide with variable concentrations of Med25. Dissociation experiments were performed by mixing 50 µM unlabeled peptide with a preformed complex of 0.5-1 µM Med25 and 0.25 µM labeled peptide. Unlabeled peptides for dissociation experiments mutants were

typically the parent peptide, but no unique effects were observed from using different competitors. Typically, 30-40 traces were averaged before fitting.

Traces were fitted using a series of exponential equations (first equation below), where $F(t)$ is the fluorescence at time t, F_∞ is the endpoint fluorescence, ΔF_n are the fluorescence amplitudes, and $k_{obs,n}$ are the observed rate constants. The individual $k_{obs,n}$ values were plotted as a function of concentration and fit to square hyperbola (second equation below) 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 fitting purposes, but the value itself is defined by the corresponding $k_{obs,n,off}$ value from dissociation experiments and thus the value from fitting was not used for calculations. The microscopic rate constants were calculated using a combined rapid equilibrium and steady-state approximation, detailed in the next section. This approach was enabled by optimized conditions for dissociation experiments from Chapter 2, as the $k_{obs,n,off}$ phases were much more clearly defined under the conditions used. To reduce erroneous error inflation from several error propagation steps, values of all microscopic rate and equilibrium constants were calculated for single datasets and averaged across 2-3 results from independent datasets. All errors reported are the standard deviation between the results from separate datasets.

$$F(t) = F_\infty + \sum \Delta F_n \exp(-k_{obs,n} \times t)$$

$$k_{obs,n} = \frac{k_{obs,n,max} * [Med25]}{[Med25] + K_{1/2,n}} + k_{obs,n,min}$$

Calculation of Microscopic Rate Constants

Here, we chose to use a combined rapid equilibrium and steady state approach to determine rate parameters as a straightforward way to handle the large amount of complex kinetic data collected. This method for calculating the microscopic rate constants is split into two "sections". First, the mechanism is considered as only the first two steps: the initial binding step to form conformer C1 and its transition to C2. After calculation of all first order microscopic rate constants from this "section" of the mechanism, the transition from C2 to C3 is considered. This is similar to treating the first two steps of the

binding mechanism as a rapid equilibrium before the final conformational transition to C3. Beginning from the first two steps, the maximal observed rate constant of the first conformational change phase ($k_{obs,2,max}$) is:

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

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

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

The value of k_{off} can be retrieved from the observed rate constants from dissociation experiments ($k_{obs,n,off}$) in an analogous way to Chapter 2. That is, by the exact expression[29] k_{off} is equivalent to:

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

In all cases except for ETV4, the fast dissociation phase ($k_{obs,1,off}$) is well defined in dissociation experiments, enabling the use of this method. For ETV4, where it is not well defined, this value was set to the minimal value we observed for other variants tested, 300 s^{-1}. To substitute a directly measurable value for $k_{F,1}$, the equation defined above for $k_{obs,2,max}$ can be used. Thus, by substitution and simplification:

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

And by rearrangement:

$$k_{R,1} = \frac{k_{obs,2,off} \times (k_{obs,1,off} + k_{obs,2,off})}{k_{obs,1,off} + 2 \times k_{obs,2,off} - k_{obs,2,max}}$$

Calculation of $k_{F,1}$ is then obtained by subtracting the $k_{R,1}$ value from $k_{obs,2,max}$. Next, the transition from C2 to C3 was considered. In dissociation experiments, the observed rate constant of the kinetic phase corresponding to this step ($k_{obs,3,off}$) is given below by the steady-state approximation:

$$k_{obs,3,off} = k_{R,2} \times \frac{k_{R,1}'}{k_{R,1}' + k_{F,2}}$$

Here, $k_{R,1}'$ is the net rate constant for successful dissociation from the C2 conformer and is identical to the expression for $k_{obs,2,off}$ above. Thus, the value of $k_{obs,2,off}$ that is obtained from fitting is used in this calculation. Similarly, $k_{obs,3,max}$ is the sum of $k_{F,2}$ and $k_{R,2}$. Thus, substitution in a analogous manner as above gives:

$$k_{obs,3,off} = k_{R,2} \times \frac{k_{obs,2,off}}{k_{obs,2,off} + k_{obs,3,max} - k_{R,2}}$$

And by rearrangement:

$$k_{R,2} = \frac{k_{obs,3,off} \times (k_{obs,2,off} + k_{obs,3,max})}{k_{obs,2,off} + k_{obs,3,off}}$$

Again, $k_{F,2}$ is obtained by subtracting the calculated value of $k_{R,2}$ from $k_{obs,3,max}$. The calculated values for all tested variants in this study are shown in Table 3.2. Relative populations of C1, C2, and C3 at equilibrium were then determined by the conformational equilibrium constants ($K_{C,n} = k_{F,n}/k_{R,n}$), which by definition are ratios between the conformational states. Below are the equations used to calculate the population of each conformational state, using C1 as the reference state. Only the relative conformational populations of the bound state were considered, thus the values are concentration-independent.

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

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

$$Population\ C3 = \frac{K_{C,1} \times K_{C,2}}{1 + K_{C,1} + K_{C,1} \times K_{C,2}} \times 100\%$$

Table 3.2 Calculated rate and equilibrium constants for all 4-DMN-labeled PEA3 TADs

ETV Variant	$K_{C,1}$	$k_{F,1}$ (s⁻¹)	$k_{R,1}$ (s⁻¹)	$K_{C,2}$	$k_{F,2}$ (s⁻¹)	$k_{R,2}$ (s⁻¹)	K_d (µM)
ETV1	3.6 ± 0.7	52 ± 5	15 ± 3	0.5 ± 0.1	2.4 ± 0.3	4.9 ± 0.8	1.2 ± 0.2
ETV4	3.2 ± 0.5	56 ± 8	18 ± 1	2.3 ± 0.5	3.3 ± 0.7	1.4 ± 0.1	0.7 ± 0.2
ETV5	3.4 ± 0.6	46 ± 4	14 ± 2	0.6 ± 0.3	2.5 ± 1.0	4.1 ± 0.3	0.9 ± 0.1
ETV4$^{DLAH/HL}$	2.7 ± 0.4	101 ± 7	38 ± 3	0.6 ± 0.1	4.6 ± 0.5	7.8 ± 0.1	1.2 ± 0.2
ETV4$^{DLAH/QF}$	5.4 ± 0.5	54 ± 6	10 ± 0.3	–	–	–	0.9 ± 0.1
ETV4$^{DLAH/HF}$	4.2 ± 0.2	63 ± 8	15 ± 2	–	–	–	0.8 ± 0.3
ETV4$^{LPPL/QL}$	0.7 ± 0.1	27 ± 4	37 ± 1	–	–	–	2.0 ± 0.1
ETV4$^{LPPL/HL}$	0.9 ± 0.1	34 + 2	38 ± 1	–	–	–	1.7 ± 0.1
ETV4$^{LPPL/QF}$	2.8 ± 0.4	52 ± 1	13 ± 1	1.6 ± 0.2	13 ± 0.7	8.1 ± 0.8	0.6 ± 0.1
ETV1$^{\Delta Nt}$	1.0 ± 0.3	21 ± 3	21 ± 3	–	–	–	0.7 ± 0.1
ETV4$^{\Delta Nt}$	3.9 ± 0.7	62 ± 1	16 ± 3	4.9 ± 1.5	4.9 ± 0.9	1.0 ± 0.1	0.4 ± 0.1

Table 3.3 Calculated conformational populations for all 4-DMN-labeled PEA3 TADs

ETV Variant	C1 Population (%)	C2 Population (%)	C3 Population (%)
ETV1	16 ± 2	56 ± 6	28 ± 5
ETV4	9 ± 1	28 ± 4	63 ± 5
ETV5	16 ± 2	53 ± 8	31 ± 9
ETV4$^{DLAH/HL}$	19 ± 2	51 ± 3	30 ± 2
ETV4$^{DLAH/QF}$	16 ± 1	84 ± 1	–
ETV4$^{DLAH/HF}$	19 ± 1	81 ± 1	–
ETV4$^{LPPL/QL}$	58 ± 4	42 ± 4	–
ETV4$^{LPPL/HL}$	53 ± 1	47 ± 1	–
ETV4$^{LPPL/QF}$	12 ± 1	34 ± 3	54 ± 3
ETV1$^{\Delta Nt}$	50 ± 7	50 ± 7	–
ETV4$^{\Delta Nt}$	4 ± 2	17 ± 4	79 ± 6

Discussion of Kinetic Data Analysis

The major strength of the calculation method outlined above is that it is operationally simple, and performs very well when tested against simulated data in the relative rate ranges we observed experimentally. Specifically, calculated values of microscopic rate constants obtained from simulated data were well within 20% of the input values, which is on the level of normal experimental error. However, we must note that this methodology performs poorly under conditions where $k_{R,2}$ approaches the value of $k_{R,1}$. In kinetic simulations, the range where this becomes problematic is when $k_{R,2} \geq 0.5^*k_{R,1}$; above this value of $k_{R,2}$ the equilibrium constant $K_{C,2}$ can be significantly underestimated. In the current study, this scenario only occurred for the variant ETV4$^{LPPL/QF}$ (see Tables 3.2 and 3.3), and this likely led to an underestimated $K_{C,2}$ value. Nonetheless, this variant was still detected by kinetic analysis as conformationally similar to wild-type ETV4, which is further supported by NMR analysis (See Figs. 3.6A and 3.7). As the value of $k_{R,2}$ increases beyond $k_{R,1}$ it becomes especially challenging to fit all three phases accurately, in this case because this step is no longer the amplitude of the corresponding kinetic phase shrinks considerably. This specific issue is not necessarily a drawback of our analysis approach, as other calculation methods and global fitting strategies can not typically detect a fast step that follows a slow step. This kinetic behavior also may be why the Group 2 ETV4 (Fig. 3.3B) and the ETV1$^{\Delta Nt}$ (Fig. 3.6C) variants only have one observable conformational change. In the case where the second conformational change phase becomes undetectable due to an increase in $k_{R,2}$, the apparent population of C2 (as presented in the figures) would be the sum of the true populations of C2 and C3 conformers. We note that this possibility does not affect any of the conclusions presented in this work.

NMR Spectroscopy

Constant time ^{1}H,^{13}C-HSQC experiments were performed with 50-75 µM uniformly ^{13}C,^{15}N-labeled Med25 in NMR buffer (20 mM sodium phosphate pH 6.5, 150 mM NaCl, 3 mM DTT, 10% D$_2$O, and 2% DMSO) on a Bruker 600 MHz instrument equipped with a

cryoprobe. HSQC experiments were processed in NMRPipe[44] and visualized with NMRFAM-Sparky.[45] All chemical shift perturbation analyses were performed on samples with 1.1 equivalents of unlabeled binding partner, which results in ≥96% bound Med25 based on measured K_d values. Peak assignments of PEA3-bound complexes were achieved by titration experiments with the unlabeled TAD with titration points of 0.2, 0.5, 0.8, and 1.1 equivalents of TAD, and assignments of mutations were made based on the parent TAD complexes. Native ETV1- and ETV4-bound spectra were representative of at least three biological replicates using different protein and peptide stocks. Chemical shift perturbations (Δδ) were calculated from the proton ($\Delta\delta_H$) and carbon ($\Delta\delta_C$) chemical shifts by:

$$\Delta\delta = \sqrt{(\Delta\delta_H)^2 + (0.25 \times \Delta\delta_C)^2}$$

Assignments of side-chain methyl resonances of free Med25 were achieved through 3D H(CCCO)NH and (H)CC(CO)NH TOCSY experiments (23 ms TOCSY mixing time) performed with a sample of 600 μM ^{13}C,^{15}N Med25 on a Bruker 800 MHz instrument equipped with a cryoprobe. An additional non-uniformly sampled 4D HCC(CO)NH TOCSY experiment[46] (12 ms TOCSY mixing time and 25% sampling) with a sample of 400 μM ^{13}C,^{15}N Med25 was performed to supplement the 3D experiments. ^{1}H,^{13}C resonance assignment from these experiments was enabled by a previous assignment of the Med25 ^{1}H,^{15}N-HSQC spectrum from Chapter 2,[29] however of 141 possible assignable non-proline amide resonances only 117 had been assigned, which precluded full assignment of the methyl spectrum. Thus, a supplemental set of TROSY-based HNCACB and CBCA(CO)NH experiments were performed to enable more complete assignment of ^{1}H,^{15}N resonances using a 600 μM sample of ^{13}C,^{15}N Med25 on a Bruker 800 MHz instrument. From these experiments, 132 of 141 amide resonances along with 83 of 89 methyl resonances were assigned. Unassigned methyl resonances include one diastereotopic methyl of Leu399 and both methyls of Leu544, the latter of which is a cloning artifact and not part of the native Med25 sequence. Three residues, Leu427, Ile453, and Ile526 are not possible to assign from the TOCSY experiments as they are directly before Pro residues in the primary sequence. However, at least one methyl from

each of these residues were previously assigned by another group[47] and were well separated from other resonances. The full ^{1}H,^{13}C-HSQC spectral assignment is shown below (Fig. 3x) along with a structural representation of the assignment (Fig. 3x).

Figure 3.10. Assigned methyl ^{1}H,^{13}C-HSQC spectrum of Med25. Experiment was conducted at 600 μM Med25 and 800 MHz field strength.

Figure 3.11. Assigned residues with methyl groups plotted on structure of Med25 (PDB ID 2XNF).[31] Green: all possible methyl groups for the residue were assigned using either the two 3D HCC(CO)NH TOCSY experiments, or the HNCACB experiment (for remaining Ala residues before Pro in the sequence). Yellow: only one of two diastereotopic methyls were assigned from the TOCSY experiments. Blue: well-dispersed methyls for the residue were assigned using a previous methyl assignment.[47] In all cases, these residues were before Pro residues in the sequence and thus were not observed in the HCC(CO)NH TOCSY experiments. Black: unassigned.

3.6 References

1. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113–119 (1987).

2. Hope, I. A., Mahadevan, S. & Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333**, 635–640 (1988).

3. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).

4. Mapp, A. K. & Ansari, A. Z. A TAD Further: Exogenous Control of Gene Activation. *ACS Chem. Biol.* **2**, 62–75 (2007).

5. Piskacek, S. *et al.* Nine-amino-acid transactivation domain: Establishment and prediction utilities. *Genomics* **89**, 756–768 (2007).

6. Krishnamurthy, M. *et al.* Caught in the Act: Covalent Cross-Linking Captures Activator–Coactivator Interactions *in Vivo*. *ACS Chem. Biol.* **6**, 1321–1326 (2011).

7. Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).

8. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

9. 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).

10. 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.* **111**, E3506–E3513 (2014).

11. 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).

12. Ravarani, C. N. *et al.* High-throughput discovery of functional disordered regions: investigation of transactivation domains. *Mol. Syst. Biol.* **14**, (2018).

13. Arnold, C. D. *et al.* A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J.* **37**, (2018).

14. Staller, M. V. *et al.* A High-Throughput Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. *Cell Syst.* **6**, 444-455.e6 (2018).

15. Minter, A. R., Brennan, B. B. & Mapp, A. K. A Small Molecule Transcriptional Activation Domain. *J. Am. Chem. Soc.* **126**, 10504–10505 (2004).

16. Rowe, S. P. & Mapp, A. K. Assessing the permissiveness of transcriptional activator binding sites. *Biopolymers* **89**, 578–581 (2008).

17. 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).

18. Majmudar, C. Y. *et al.* Sekikaic Acid and Lobaric Acid Target a Dynamic Interface of the Coactivator CBP/p300. *Angew. Chem. Int. Ed.* **51**, 11258–11262 (2012).

19. Chrivia, J. C. *et al.* Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855–859 (1993).

20. Ruas, J. L., Poellinger, L. & Pereira, T. Functional Analysis of Hypoxia-inducible Factor-1α-mediated Transactivation: IDENTIFICATION OF AMINO ACID RESIDUES CRITICAL FOR TRANSCRIPTIONAL ACTIVATION AND/OR INTERACTION WITH CREB-BINDING PROTEIN. *J. Biol. Chem.* **277**, 38723–38730 (2002).

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

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. 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).

24. Darnell, J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749 (2002).

25. Bushweller, J. H. Targeting transcription factors in cancer — from undruggable to reality. *Nat. Rev. Cancer* **19**, 611–624 (2019).

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

27. 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).

28. Currie, S. L. *et al.* ETV4 and AP1 Transcription Factors Form Multivalent Interactions with three Sites on the MED25 Activator-Interacting Domain. *J. Mol. Biol.* **429**, 2975–2995 (2017).

29. Henderson, A. R. *et al.* Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *Proc. Natl. Acad. Sci.* **115**, 8960–8965 (2018).

30. Loving, G. & Imperiali, B. A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4- *N,N* -Dimethylamino-1,8-naphthalimide: A Powerful Tool for the Study of Dynamic Protein Interactions. *J. Am. Chem. Soc.* **130**, 13630–13638 (2008).

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

32. Zhuravleva, A. & Gierasch, L. M. Substrate-binding domain conformational dynamics mediate Hsp70 allostery. *Proc. Natl. Acad. Sci.* **112**, E2865–E2873 (2015).

33. Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).

34. Fuxreiter, M. Fuzziness in Protein Interactions—A Historical Perspective. *J. Mol. Biol.* **430**, 2278–2287 (2018).

35. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

36. Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. Binding Rate Constants Reveal Distinct Features of Disordered Protein Domains. *Biochemistry* **54**, 4741–4750 (2015).

37. Yang, K. *et al.* Structural basis for cooperative regulation of KIX-mediated transcription pathways by the HTLV-1 HBZ activation domain. *Proc. Natl. Acad. Sci.* **115**, 10040–10045 (2018).

38. 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).

39. Brüschweiler, S., Konrat, R. & Tollinger, M. Allosteric Communication in the KIX Domain Proceeds through Dynamic Repacking of the Hydrophobic Core. *ACS Chem. Biol.* **8**, 1600–1610 (2013).

40. Wang, N. *et al.* Ordering a Dynamic Protein Via a Small-Molecule Stabilizer. *J. Am. Chem. Soc.* **135**, 3363–3366 (2013).

41. Nishikawa, J. L. *et al.* Inhibiting fungal multidrug resistance by disrupting an activator–Mediator interaction. *Nature* **530**, 485–489 (2016).

42. Wu, Z. *et al.* Targeting the Transcriptional Machinery with Unique Artificial Transcriptional Activators. *J. Am. Chem. Soc.* **125**, 12390–12391 (2003).

43. Cesa, L. C. *et al.* X-linked inhibitor of apoptosis protein (XIAP) is a client of heat shock protein 70 (Hsp70) and a biomarker of its inhibition. *J. Biol. Chem.* **293**, 2370–2380 (2018).

44. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, (1995).

45. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

46. Mobli, M., Stern, A. S., Bermel, W., King, G. F. & Hoch, J. C. A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *J. Magn. Reson.* **204**, 160–164 (2010).

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

CHAPTER 4

Conclusions and Future Directions

4.1 Conclusions

The work presented in this dissertation contributes to a deeper understanding of the mechanisms by which transcriptional activator•coactivator protein-protein interactions (PPIs) form. Particularly, this work sheds significant light on the mechanisms that enable a single activator binding domain (ABD) in a coactivator to interact with several distinct activator sequences that often share little sequence homology.

Several proposals have been put forward to account for this longstanding observation. Initial functional mutagenesis studies of transcriptional activation domains (TADs) indicated that acidic and hydrophobic amino acids were critical to function, although the specific amino acid sequence was not as important.[1–5] This result was interpreted to mean that the intrinsically disordered TAD uses nonspecific hydrophobic and electrostatic interactions to mediate PPIs with the ABDs of the transcriptional machinery.[3] Thus, according to this model, an ABD can interact with several unique TADs by virtue of the promiscuous and nonspecific character of the TAD sequence. Recent biophysical work from the Hahn group appears to support this notion.[6–8] Opposing this nonspecific model is the proposition that the TADs use their intrinsic flexibility to adapt to the topology of an ABD surface. [9] In this model, different TAD sequences that bind to the same ABD can form specific complexes because the TADs utilize the binding interface(s) in distinct manners. This latter view has considerable support from structural studies of ABDs in complex with unique TADs.[9–14]

A common theme between these prominent molecular recognition models is that they emphasize the properties of the intrinsically disordered TAD as a means to understand how TAD•ABD complexes form. The role of the ABD in molecular recognition, on the

94

other hand, tends to be ignored or diminished in structural and mechanistic studies of TAD•ABD complex formation.[6–8,15,16] Nonetheless, both recognition models have implications for the mechanism by which ABDs function. The modern nonspecific models, for example, posit that the ABD serves as little more than a shallow and featureless hydrophobic surface to interact with the "hydrophobic cloud" formed by hydrophobic sidechains of the TAD.[8] The negatively charged side-chains of the TAD are then complemented by positively charged residues surrounding the hydrophobic binding site. Conversely, specific models stress that recognition occurs through the flexible TAD adapting to the ABD interface, which implies that the mechanism by which ABDs recognize TADs is through the general topology of hydrophobic grooves and charged patches in the ABD surface.[9] Thus, the ABD serves as a scaffold to bind to TADs that can complement the characteristics of its binding surface. As discussed in chapter one, a less-studied aspect of specific recognition additionally includes the ability of the ABD to adapt its surfaces to better complement the TAD.

In this dissertation, we examined these molecular recognition mechanisms using the structurally distinct ABD from Mediator subunit Med25 as a model system.[17–19] Because many of the model ABDs used to investigate activator molecular recognition are structurally similar, the unique structure of the Med25 ABD enabled us to dissect the underlying principles of activator molecular recognition in a manner that is less biased to specific structural motifs. Below, we discuss the general themes that emerged from these studies and their implications to activator molecular recognition.

The Central Role of Dynamic Substructures in Molecular Recognition

In chapter two, we examined how Med25 forms complexes with distinct activator sequences. We found that, analogous to well-studied ABDs, Med25 uses two topologically unique binding sites that each interact with distinct sets of binding partners. Recognition of activators is further facilitated by conformational changes in dynamic substructures of the Med25 ABD that occur after binding. Conformational shifts in these substructures also mediates allosteric communication between the two Med25 binding surfaces, and directly targeting a dynamic loop with a small molecule recapitulated the allosteric effects of a native activator.

Together, these results provide significant evidence that a key element of activator molecular recognition is conformationally plastic ABD binding interfaces. In addition to assisting the formation of binary activator•coactivator complexes, highly dynamic structural elements can facilitate allosteric coupling[20,21] and thus are likely a critical general mechanism for the formation of ternary activator•coactivator complexes. Functionally, cooperative formation of these ternary complexes are thought to increase selectivity of specific ABDs for certain enhancers. For example, enhancers containing PEA3 binding sites also often contain binding sites for AP-1 transcription factors, and there is some evidence that this enhances association of Mediator through ternary complex formation of AP-1 and PEA3 TADs with the Med25 ABD.[22]

Dynamic substructures of ABDs also provide handles to modulate ABD function with small molecule "cochaperones" that stabilize specific conformations of the ABD. This strategy proved successful here with Med25 and previously with the KIX ABD.[23–25] Studies of KIX–cochaperone complexes indicate that stabilization of dynamic regions in the ABD can differentially modulate activator binding in a manner dependent on the molecule and the region it targets, suggesting that screening hits can be optimized into either agonists or antagonists of the ABD.[23–25] As a general targeting mechanism, conformational stabilization will likely be a more successful tactic than direct targeting of activator binding surfaces, which are typically too broad and shallow for small molecules to bind with high affinity.[26] A critical next step in the development of these molecules will therefore to be to demonstrate that they can be optimized for both affinity and allosteric activity.

"Sequence-Independent" Function of Activators is Tied to the Plasticity of ABDs

In chapter three, we examined how the Med25 ABD forms complexes with a set of highly related activators from the PEA3 family of Ets transcription factors. PEA3 family activators form highly dynamic complexes with Med25, with multiple conformations populated at equilibrium. These interactions are thus analogous to the dynamic interactions formed by the yeast activator Gcn4 with Med15, the key model system for nonspecific activator•coactivator molecular recognition.[6–8] However, in contrast to the expectations from nonspecific recognition models, our detailed mechanistic analysis of

PEA3•Med25 interactions demonstrated that there is a significant degree of conformational sensitivity to slight changes in the TAD sequence. This result strongly suggests that the PEA3 TADs make specific contacts in the TAD•Med25 complex, and indeed structural analysis demonstrated that elements of the PEA3 TADs localize to specific regions on the Med25 interface. Conformational changes in the Med25 ABD also play a role in enabling specific complex formation, further demonstrating the critical role ABD plasticity plays in molecular recognition.

A key observation that underlies nonspecific models of TAD recognition is that no specific sequence motifs predict TAD function except for the presence of hydrophobic and acidic amino acids and the absence of basic amino acids.[1,5,27–29] In the face of this surprising result, a reasonable interpretation appears to be that TADs are functionally nonspecific. That is, they bind indiscriminately to other proteins in the transcriptional apparatus, which is sufficient to fulfil their function as recruiters. However, an alternative explanation—supported by the data presented in this dissertation—is that the ABDs in the transcriptional machinery are highly plastic and can individually adapt to a variety of TAD sequences by remodeling their binding surfaces. Further, there are several structurally distinct ABDs in the transcriptional machinery, so a sequence that meets the general criteria of a TAD can likely find at least one binding partner that can accommodate its specific display of acidic and hydrophobic sidechains. This scenario does not require all complexes to form through nonspecific interactions. Critically, this indicates that the key to understanding TAD molecular recognition in a systems-based manner is to assay how the interactome of a TAD changes when it is mutated, in addition to whether or not it functions. Such an experiment could also assist target identification for small molecule intervention by identifying nodes in TAD PPI networks that are critical to function.

4.2 Future Directions

The conclusions from this dissertation have implicated ABD plasticity as a critical player in TAD recognition. Future efforts will build on this model, focusing on testing several key hypotheses that emerged from this work.

A key hypothesis that emerges from the observation that members of the PEA3 family of activators bind to Med25 with unique conformational ensembles is that individual

PEA3•Med25 PPIs could be differentially affected by conformational shifts in the Med25 ABD. In an initial test of this hypothesis, we generated the point mutation Med25^{M523E} to affect the conformation of a dynamic allosteric loop and tested the effects on the binding of PEA3 activators (Fig. 4.1). Strikingly, this mutation uniquely affected the binding of PEA3 activators in a conformationally-specific manner. The binding affinity of ETV1 and ETV5, which have similar conformational ensembles, were weakened by ~2 and ~1.4-fold, respectively. In contrast, the binding affinity of ETV4 was enhanced by ~3-fold. The functional result of this mutation was therefore to cause a shift in Med25 to prefer binding to ETV4 over ETV1 and ETV5 by ~9 and ~4-fold, respectively. This indicates that allosteric small molecules—targeting dynamic substructures—may be able to induce selective changes of an ABD interactomes between even highly related TAD sequences. Such a scenario opens up the possibility of being able to directly assay the function of individual TAD•ABD complexes *in vivo*.
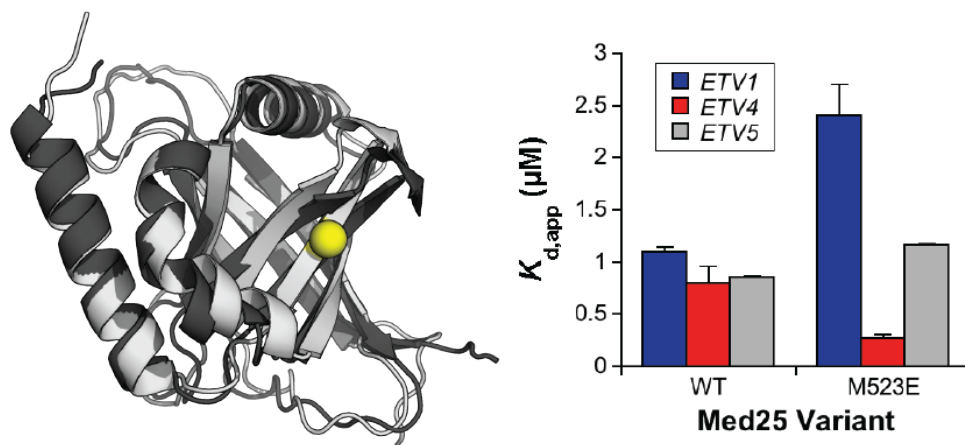
Figure 4.1. Modulation of an allosteric loop induces selectivity changes between PEA3•Med25 complexes. Left, molecular dynamics simulations (yellow sphere is the site of the Med25^{M523E} mutation) demonstrate a conformational shift in the allosteric loop (above mutation). Right, changes in affinity induced by the Med25M523E mutation. Molecular dynamics simulations were performed by Amanda Peiffer. Binding experiments were performed as described in chapter three.

Further, while no other known coactivators contain an ABD with a similar sequence or structure to Med25, the PTOV1 protein contains two domains that are highly homologous to the Med25 ABD (73 and 81% sequence identity).[30–33] The function of

PTOV1 is currently unknown, but it has been proposed that it functions to squelch Med25 PPIs, based on the expectation that the PTOV1 domains bind to the same set of partners as Med25. However, based on the conclusions of this dissertation, the sequence differences between PTOV1 and Med25 may be equally as likely to result in distinct interactomes of each individual domain. Thus, a next step is to examine how the PTOV1 domains bind to native Med25 partners, including investigating changes in binding mechanisms. At the same time, it will be critical to identify any unique PPIs made by the PTOV1 domains. Taken together, this will provide critical insight into the relationships between ABD sequence, conformation, and function.

A final undertaking moving forward is to consider additional technologies to enable development of ABDs cochaperones. Certainly, the disulfide Tethering technology[34] used in this dissertation is a useful approach, however recent developments in protein and peptide engineering have opened new avenues for developing conformationally selective nanobodies[35] and peptide macrocycles.[36] The former is especially intriguing because nanobodies can be encoded onto inducible plasmids, avoiding issues with cell permeability that typically plague biologic approaches. Critically, these approaches provide molecules with high selectivity and affinity in a rapid manner compared to traditional medicinal chemistry approaches, which enables the development of high-quality cellular probes on a larger scale.

4.3 References

1. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113–119 (1987).
2. Hope, I. A., Mahadevan, S. & Struhl, K. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333**, 635–640 (1988).
3. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).
4. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Amp Dev.* **2**, 718–729 (1988).
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.* **90**, 883–887 (1993).
6. 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).

7. 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.* **111**, E3506–E3513 (2014).

8. 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).

9. 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).

10. Radhakrishnan, I. *et al.* Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell* **91**, 741–752 (1997).

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

12. Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).

13. Langlois, C. *et al.* NMR Structure of the Complex between the Tfb1 Subunit of TFIIH and the Activation Domain of VP16: Structural Similarities between VP16 and p53. *J. Am. Chem. Soc.* **130**, 10596–10604 (2008).

14. Krois, A. S., Ferreon, J. C., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc. Natl. Acad. Sci.* **113**, E1853–E1862 (2016).

15. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

16. Dogan, J., Jonasson, J., Andersson, E. & Jemth, P. Binding Rate Constants Reveal Distinct Features of Disordered Protein Domains. *Biochemistry* **54**, 4741–4750 (2015).

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

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

19. Bontems, F. *et al.* NMR structure of the human Mediator MED25 ACID domain. *J. Struct. Biol.* **174**, 245–251 (2011).

20. Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl. Acad. Sci.* **104**, 8311–8315 (2007).

21. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).

22. Currie, S. L. *et al.* ETV4 and AP1 Transcription Factors Form Multivalent Interactions with three Sites on the MED25 Activator-Interacting Domain. *J. Mol. Biol.* **429**, 2975–2995 (2017).

23. Wang, N. *et al.* Ordering a Dynamic Protein Via a Small-Molecule Stabilizer. *J. Am. Chem. Soc.* **135**, 3363–3366 (2013).

24. 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.* **111**, 12061–12066 (2014).

25. Lodge, J. M., Justin Rettenmaier, T., Wells, J. A., Pomerantz, W. C. & Mapp, A. K. FP tethering: a screening technique to rapidly identify compounds that disrupt protein–protein interactions. *Med Chem Commun* **5**, 370–375 (2014).
26. Mapp, A. K., Pricer, R. & Sturlis, S. Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **11**, 891–894 (2015).
27. Ravarani, C. N. *et al.* High-throughput discovery of functional disordered regions: investigation of transactivation domains. *Mol. Syst. Biol.* **14**, (2018).
28. Staller, M. V. *et al.* A High-Throughput Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. *Cell Syst.* **6**, 444-455.e6 (2018).
29. Arnold, C. D. *et al.* A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J.* **37**, (2018).
30. Benedit, P. *et al.* PTOV1, a novel protein overexpressed in prostate cancer containing a new class of protein homology blocks. *Oncogene* **20**, 1455–1464 (2001).
31. Mittler, G. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* **22**, 6494–6504 (2003).
32. Youn, H.-S., Park, U.-H., Kim, E.-J. & Um, S.-J. PTOV1 antagonizes MED25 in RAR transcriptional activation. *Biochem. Biophys. Res. Commun.* **404**, 239–244 (2011).
33. Cánovas, V., Lleonart, M., Morote, J. & Paciucci, R. The role of prostate tumor overexpressed 1 in cancer progression. *Oncotarget* **8**, (2017).
34. Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proc. Natl. Acad. Sci.* **97**, 9367–9372 (2000).
35. McMahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).
36. Huang, Y., Wiedmann, M. M. & Suga, H. RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chem. Rev.* **119**, 10360–10391 (2019).

Appendix A

Characterization of Synthesized Peptides

This appendix contains analytical HPLC chromatograms of all synthesized peptides used in this dissertation.

DAD1 A, Sig=280,4 Ref=off (C:\HPLC\MJH\DATA\2017_03_09_VP16_L2-L3_NTAC_F2_REPUR2 1\001-0101.D)
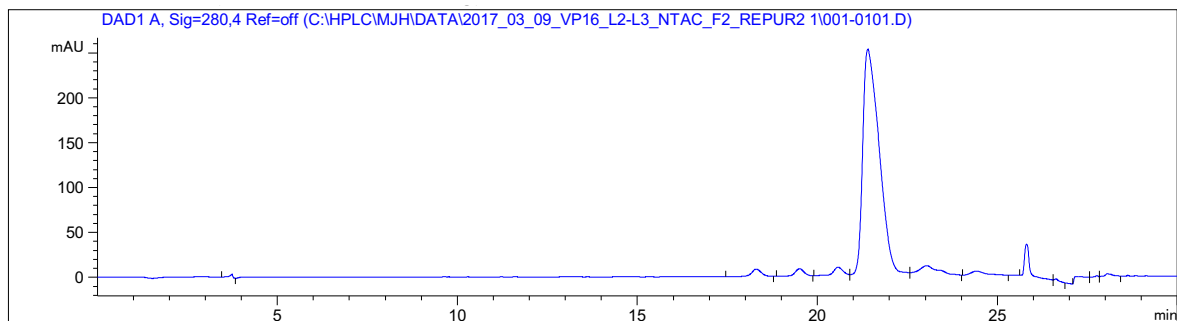
Analytical HPLC trace of **VP16 (438-490)**, monitored at 280 nm. Analytical sample was run in a 100 mM ammonium acetate(pH 7)/acetonitrile binary solvent system. The sample was injected with an isocratic flow of 82% 100 mM ammonium acetate and 18% acetonitrile. After 2 mins, the solvent gradient was increased from 18-28% ACN over 20 mins.


DAD1 A, Sig=280,4 Ref=off (C:\HPLC\MJH\DATA\2017_01_20_VP16_H1_NTAC_ANALYT\001-0101.D)
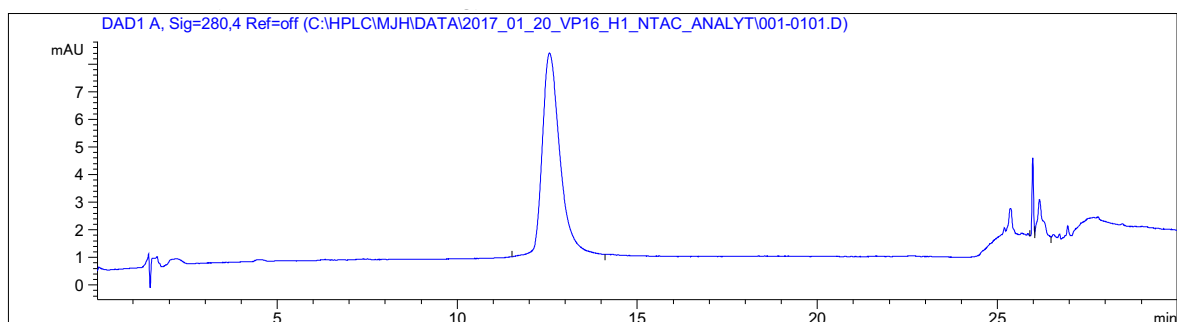
Analytical HPLC trace of **VP16 (413-454)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 84% water (with 100 mM ammonium acetate) and 16% acetonitrile. After 2 mins, the solvent gradient was increased from 16-26% ACN over 20 mins.

DAD1 A, Sig=280,4 Ref=off (C:\HPLC\MJH\DATA\2017_01_20_VP16_H2_NTAC_ANALYT\002-0101.D)
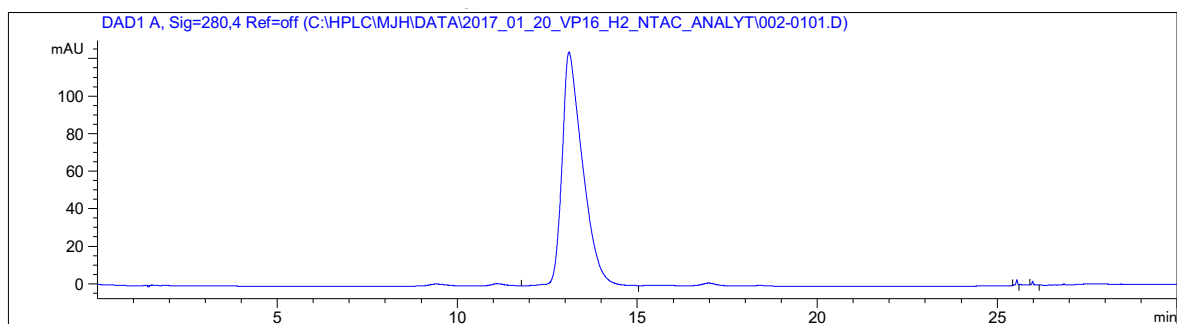
Analytical HPLC trace of **VP16 (455-490)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 78% water (with 100 mM ammonium acetate) and 18% acetonitrile. After 2 mins, the solvent gradient was increased from 22-32% acetonitrile over 20 mins.
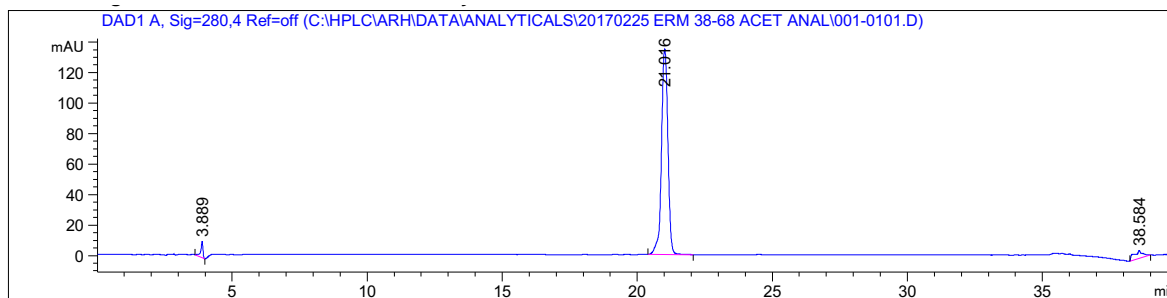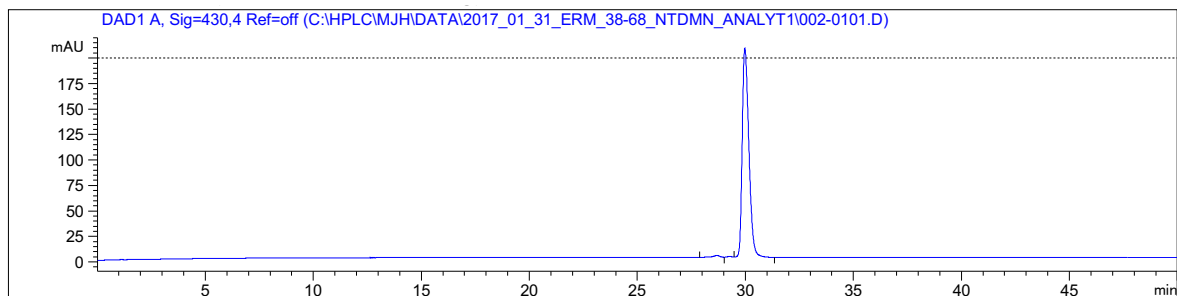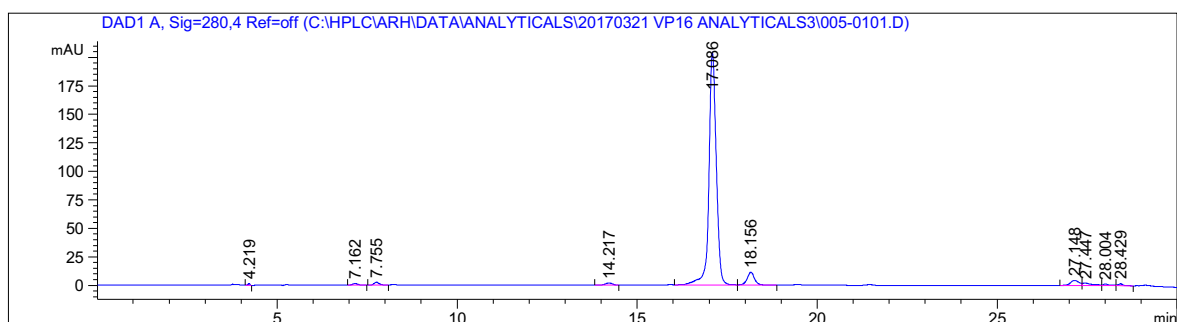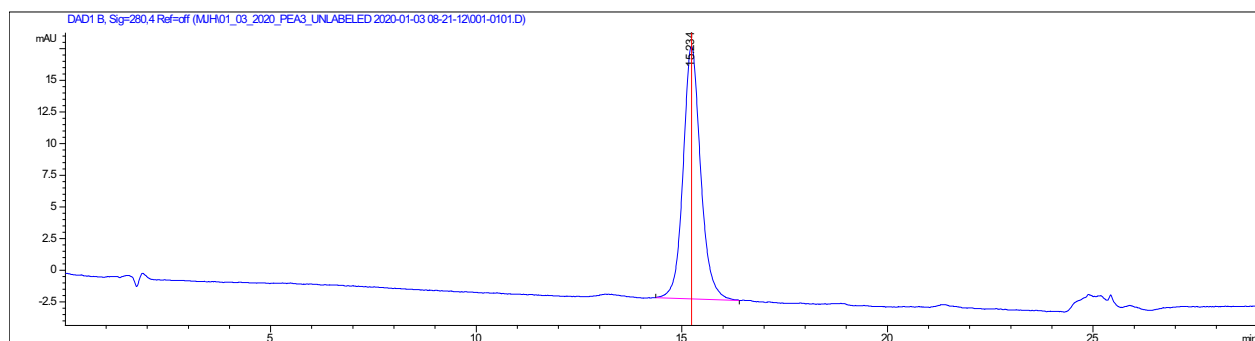
Analytical HPLC trace of **ERM (38-68)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 85% water (with 100 mM ammonium acetate) and 15% acetonitriile. After 2 mins, the solvent gradient was increased from 15-30% acetonitrile over 20 mins.
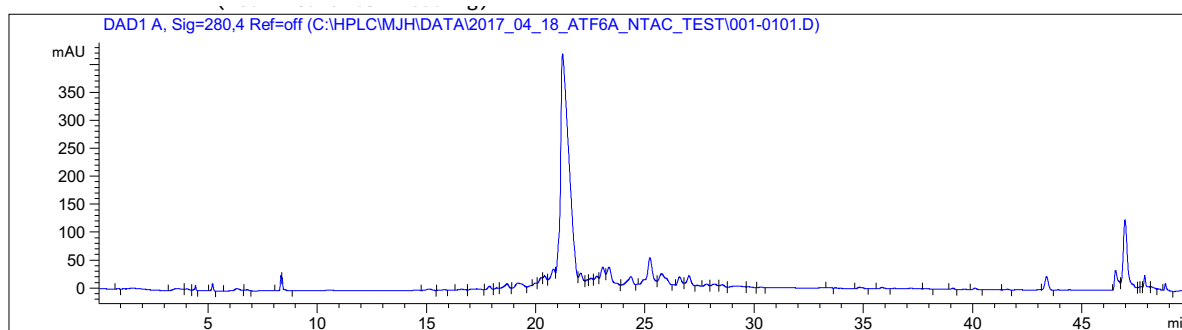
Analytical HPLC trace of **ATF6α (38-64)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 85% water (with 100 mM ammonium acetate) and 15% acetonitrile. After 2 mins, the solvent gradient was increased from 15-30% acetonitrile over 20 mins.
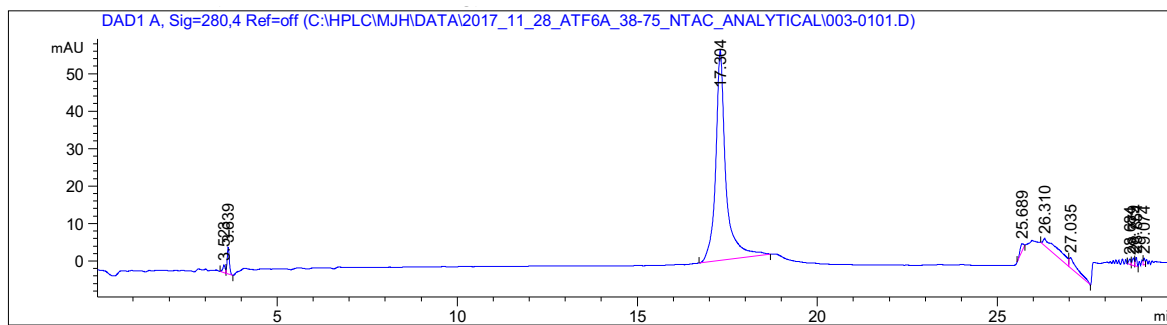
Analytical HPLC trace of **ATF6α (38-75)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 80% water (with 100 mM ammonium acetate) and 20% acetonitrile. After 2 mins, the solvent gradient was increased from 20-45% acetonitrile over 20 mins.
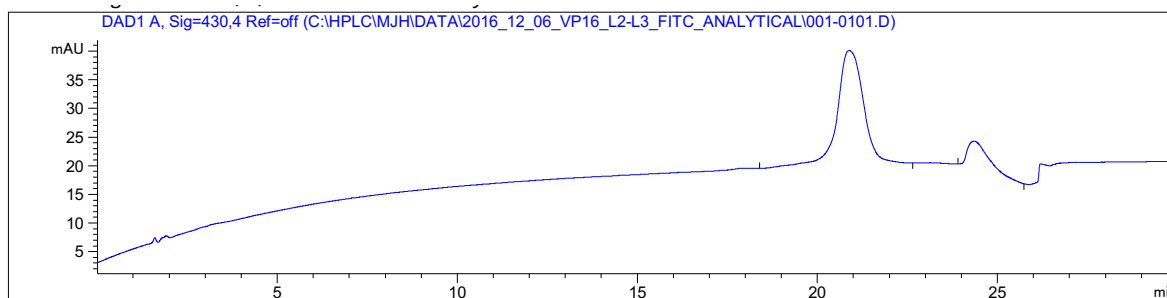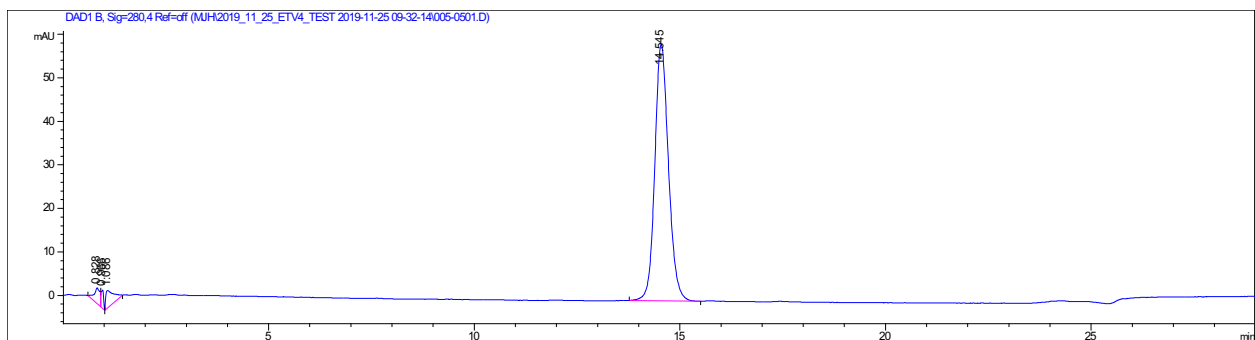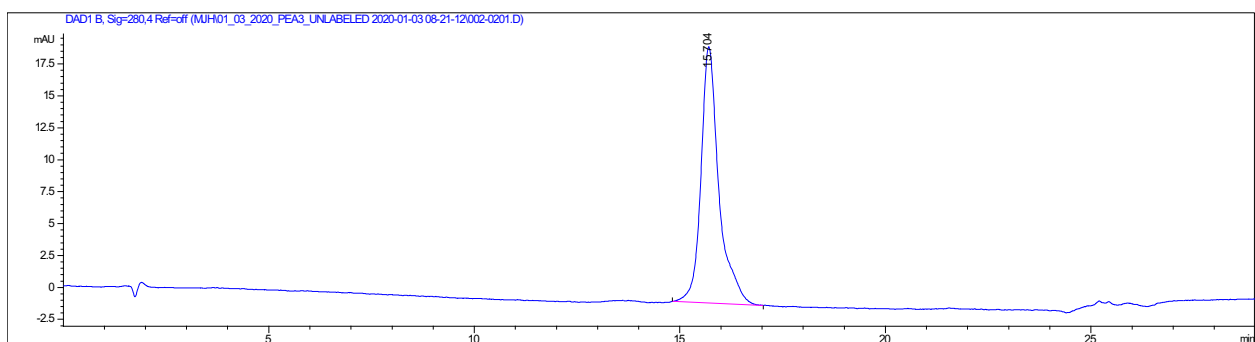
Analytical HPLC trace of **FITC-VP16 (438-490)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 82% water (with 100 mM ammonium acetate) and 18% acetonitrile. After 2 mins, the solvent gradient was increased from 18-28% acetonitrile over 20 mins.
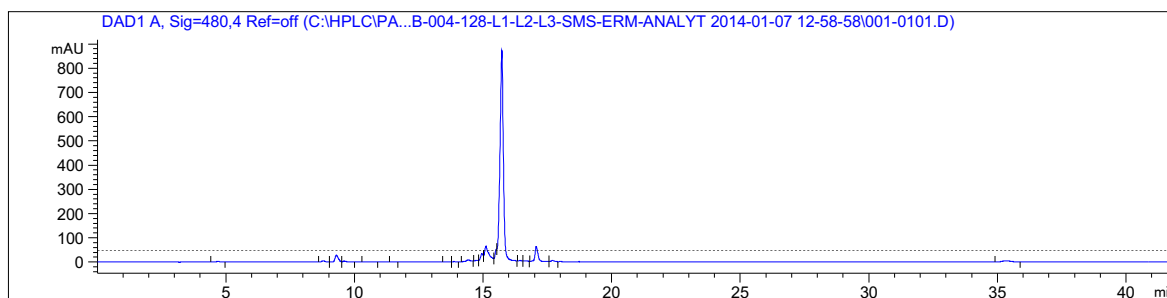
Analytical HPLC trace of **FITC-VP16 (413-437)**, monitored at 480 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 90% water (with 100 mM ammonium acetate) and 10% acetonitrile. After 2 mins, the solvent gradient was increased from 10-40% acetonitrile over 30 mins.
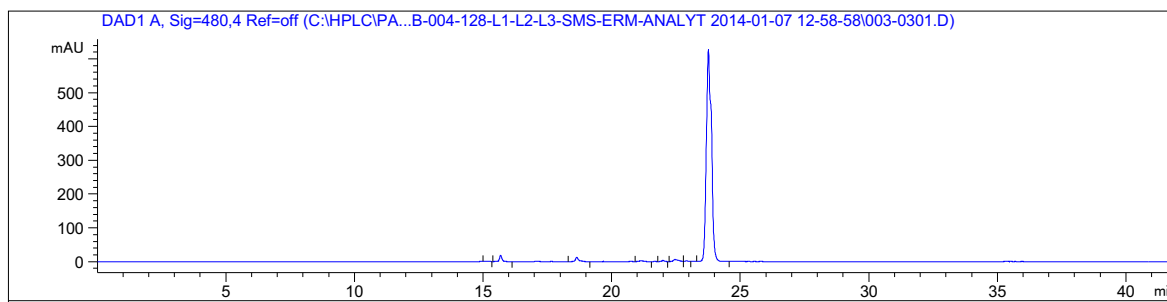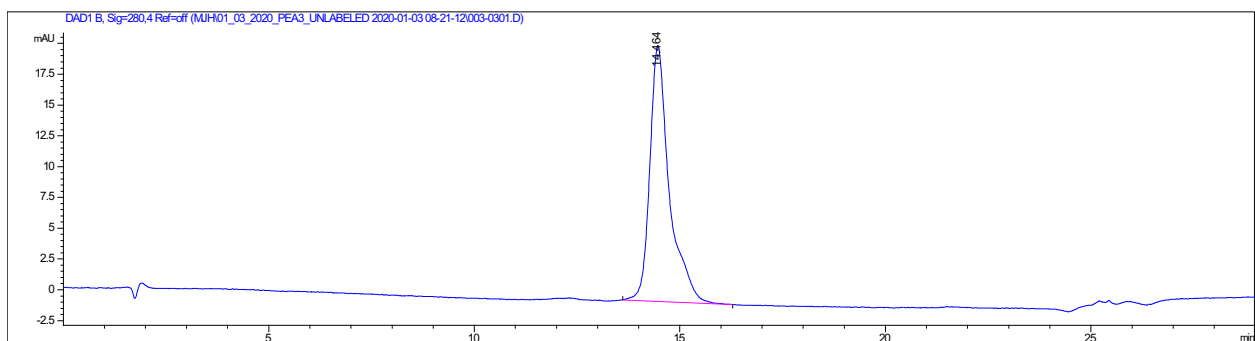

Analytical HPLC trace of **FITC-VP16 (465-490)**, monitored at 480 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 90% water (with 100 mM ammonium acetate) and 10% acetonitrile. After 2 mins, the solvent gradient was increased from 10-40% acetonitrile over 30 mins.
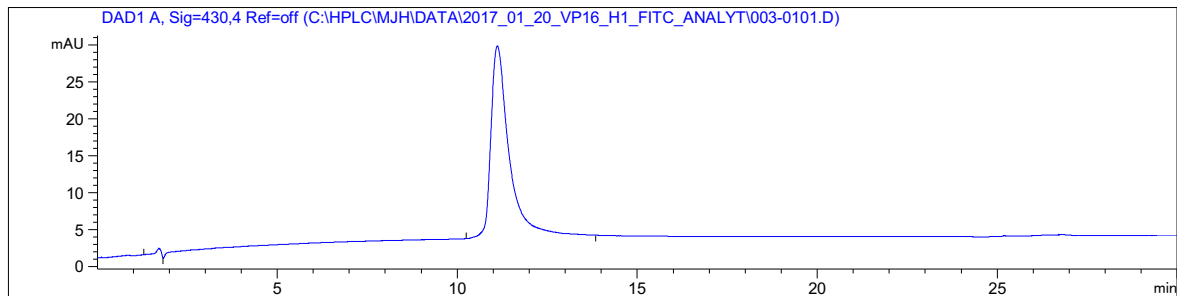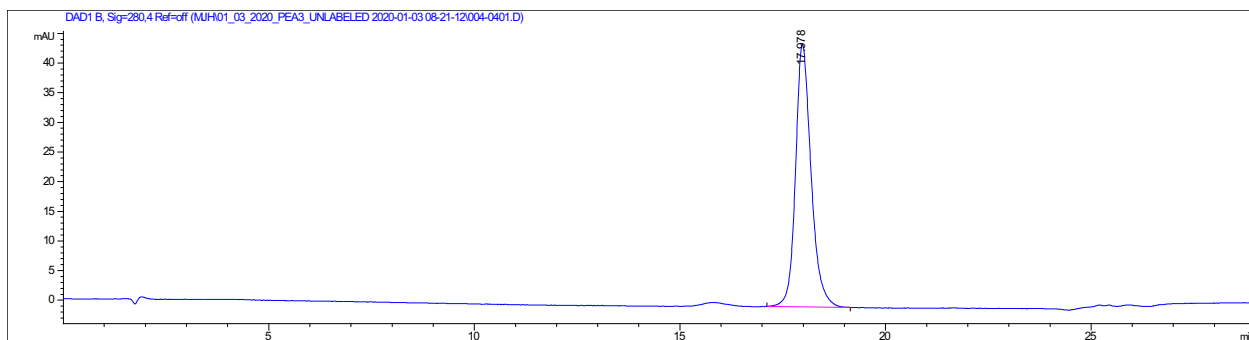
Analytical HPLC trace of **FITC-VP16 (413-454)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/acetonitrile system. The sample was injected with an isocratic flow of 90% water (with 100 mM ammonium acetate) and 10% acetonitrile. After 2 mins, the solvent gradient was increased from 10-40% acetonitrileover 30 mins.
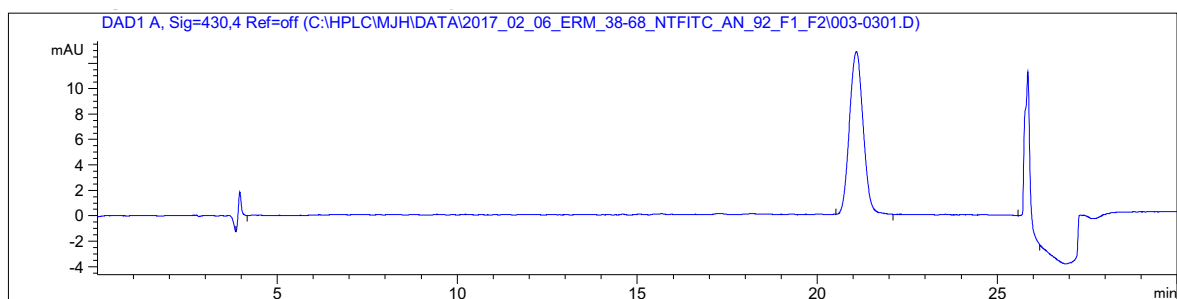
Analytical HPLC trace of **FITC-ERM (38-68)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 82% water (with 100 mM ammonium acetate) and 18% acetonitrile. After 2 mins, the solvent gradient was increased from 18-28% acetonitrile over 20 mins.
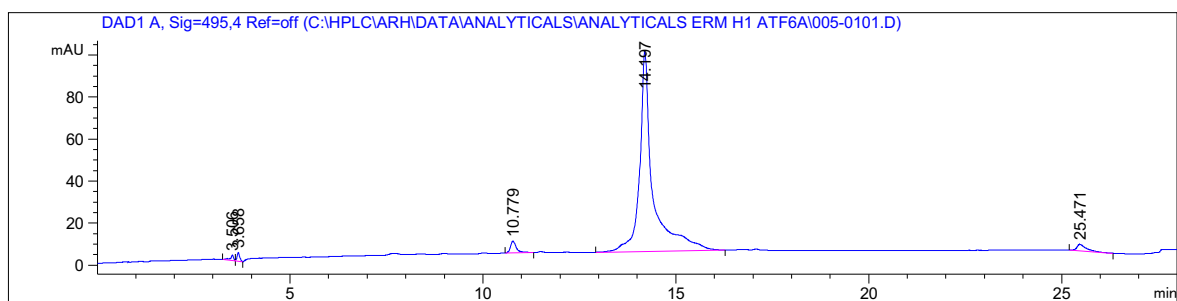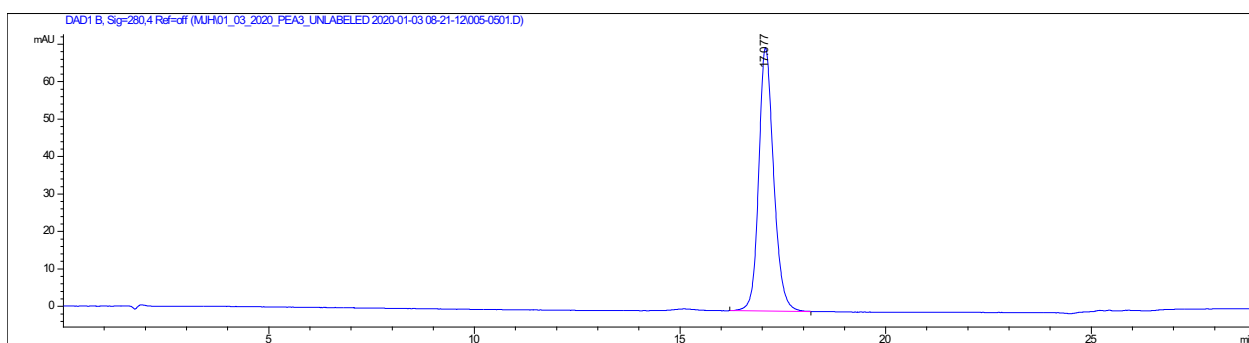

Analytical HPLC trace of **FITC-ATF6α (38-64)**, monitored at 495 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 90% water (with 100 mM ammonium acetate) and 10% acetonitrile. After 2 mins, the solvent gradient was increased from 10-40% acetonitrile over 20 mins.
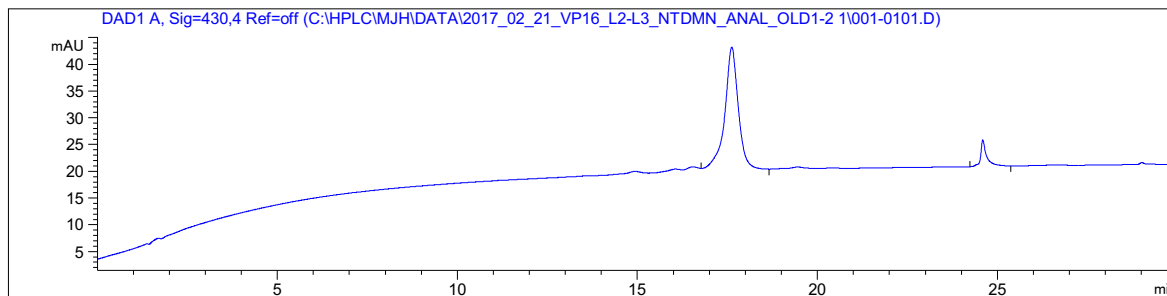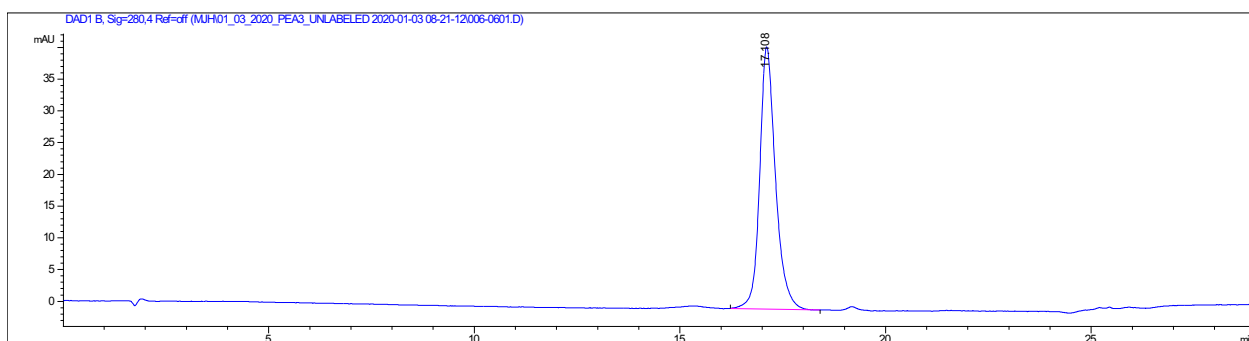
Analytical HPLC trace of **4-DMN-VP16 (438-490)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 78% water (with 100 mM ammonium acetate) and 22% acetonitrile. After 2 mins, the solvent gradient was increased from 22-32% acetonitrile over 20 mins.
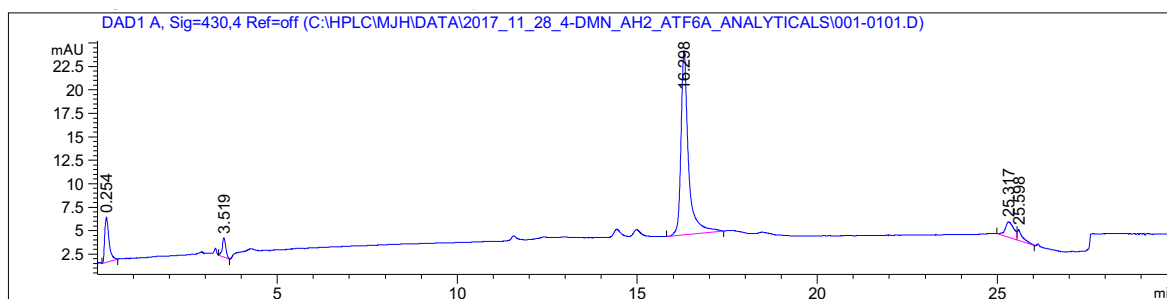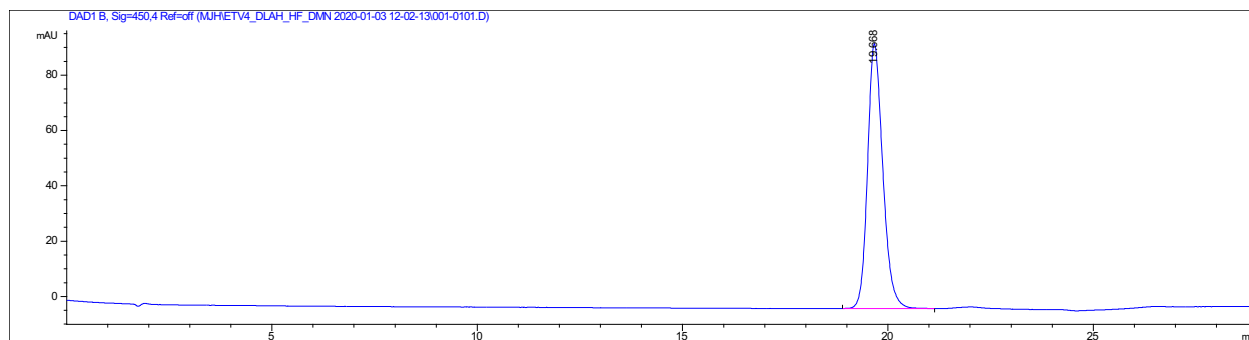

Analytical HPLC trace of **4-DMN-VP16 (467-488)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 80% water (with 100 mM ammonium acetate) and 20% acetonitrile. After 2 mins, the solvent gradient was increased from 20-45% acetonitrile over 20 mins.
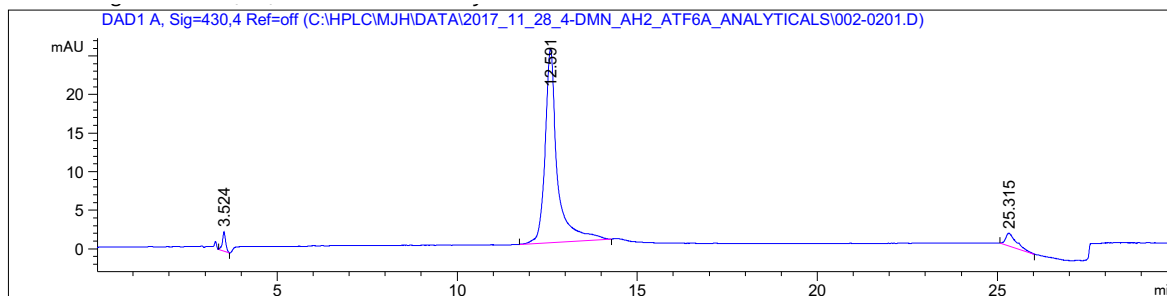
Analytical HPLC trace of **4-DMN-ATF6α (38-75)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 80% water (with 100 mM ammonium acetate) and 20% acetonitrile. After 2 mins, the solvent gradient was increased from 20-45% acetonitrile over 20 mins.
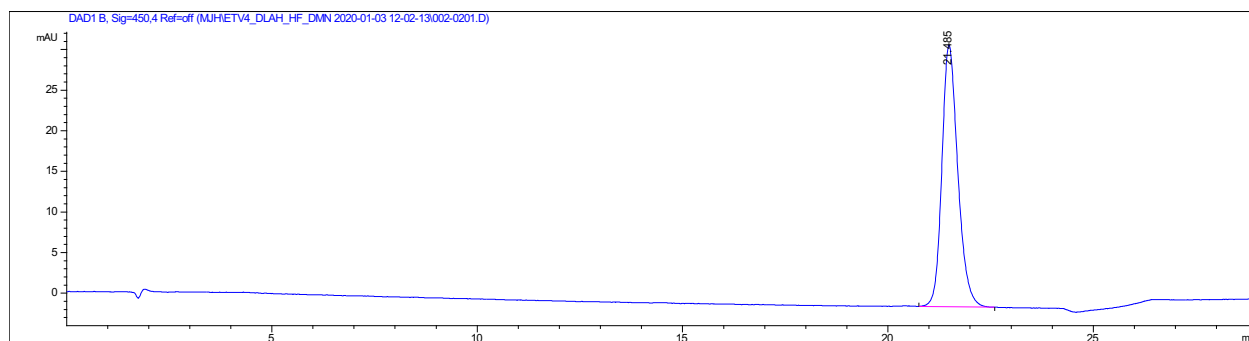
107

Analytical HPLC trace of **4-DMN-ERM (38-68)**, monitored at 430 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 90% water (with 100 mM ammonium acetate) and 10% acetonitrile. After 2 mins, the solvent gradient was increased from 10-40% acetonitrile over 40 mins.
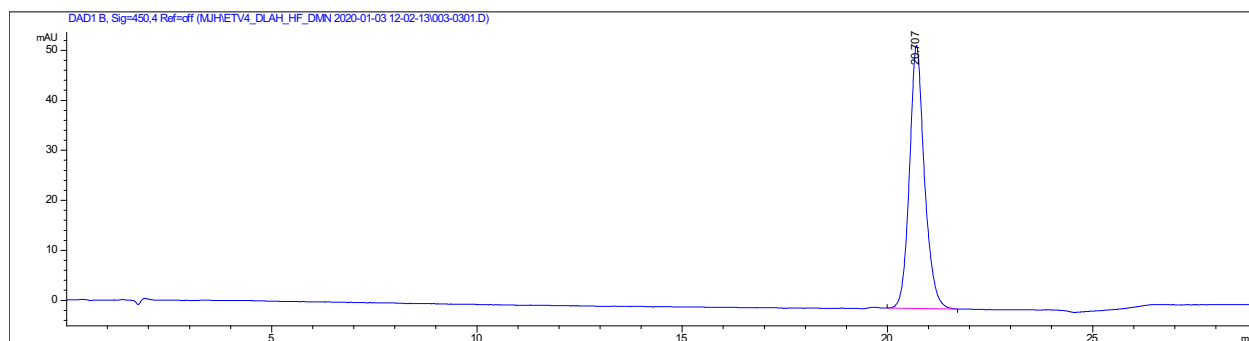

Analytical HPLC trace of **VP16 (438-454) G450C**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 85% water (with 100 mM ammonium acetate) and 15% acetonitrile. After 2 mins, the solvent gradient was increased from 15-30% acetonitrile over 20 mins.
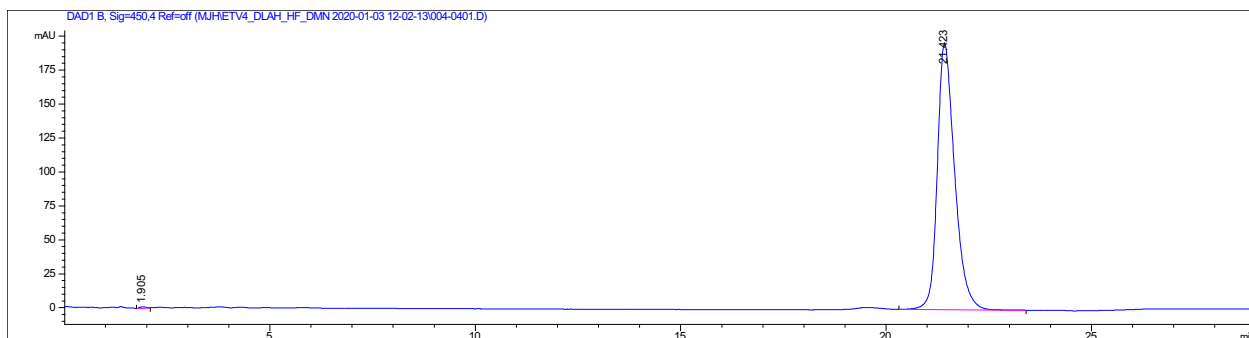

Analytical HPLC trace of **ETV1(38-69)**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

108

Analytical HPLC trace of **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 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.
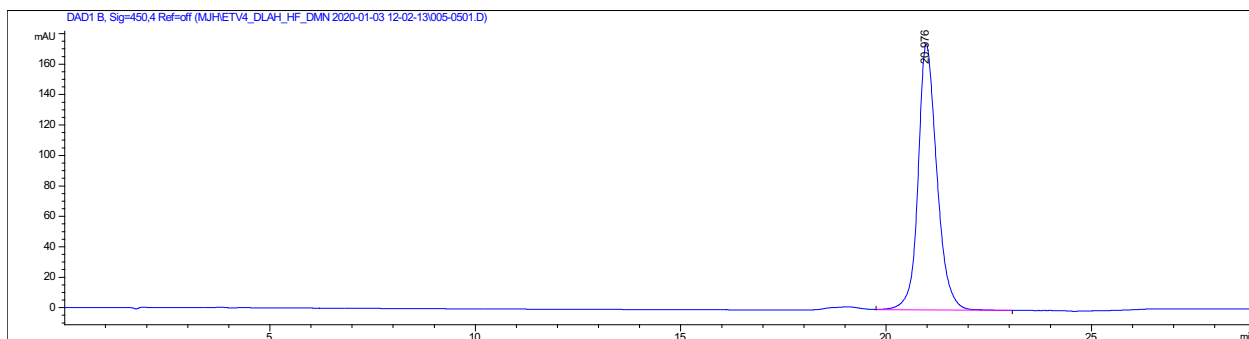

Analytical HPLC trace of **ETV1(38-69)HL**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.
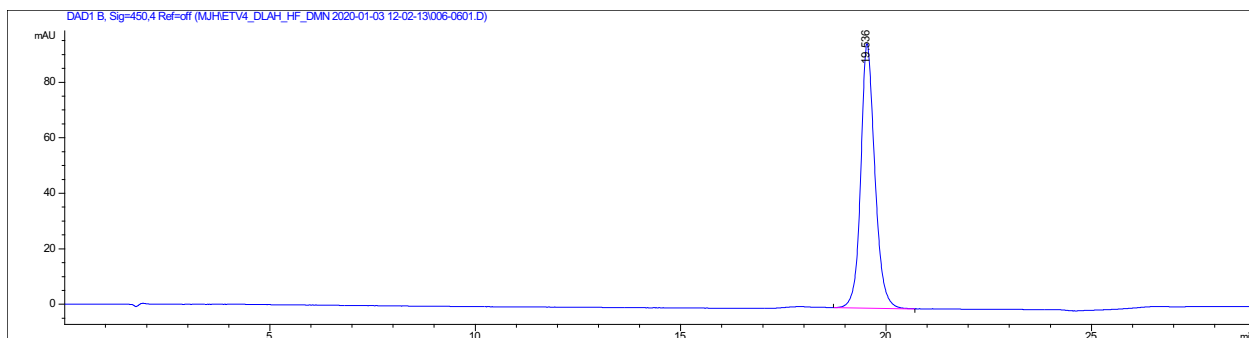

Analytical HPLC trace of **ETV1(38-69)DVAH**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **ETV4(45-76)^F60L**, 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 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.
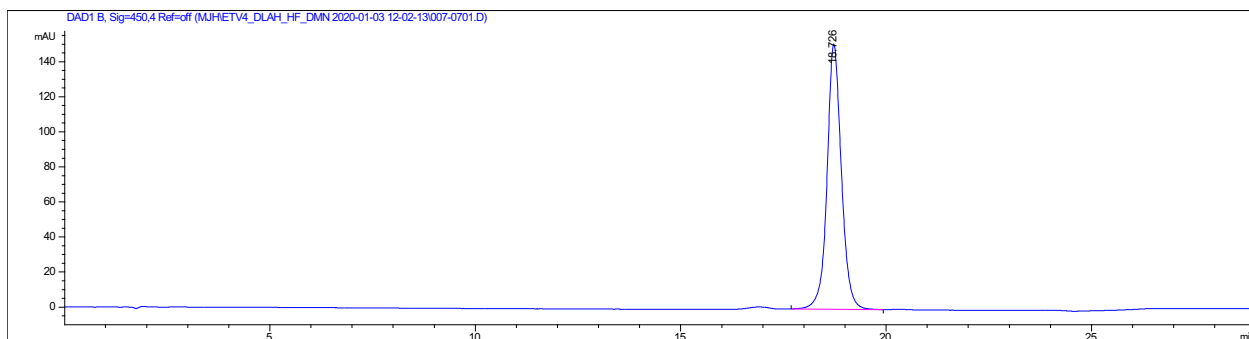

Analytical HPLC trace of **ETV4(45-76)^QF**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient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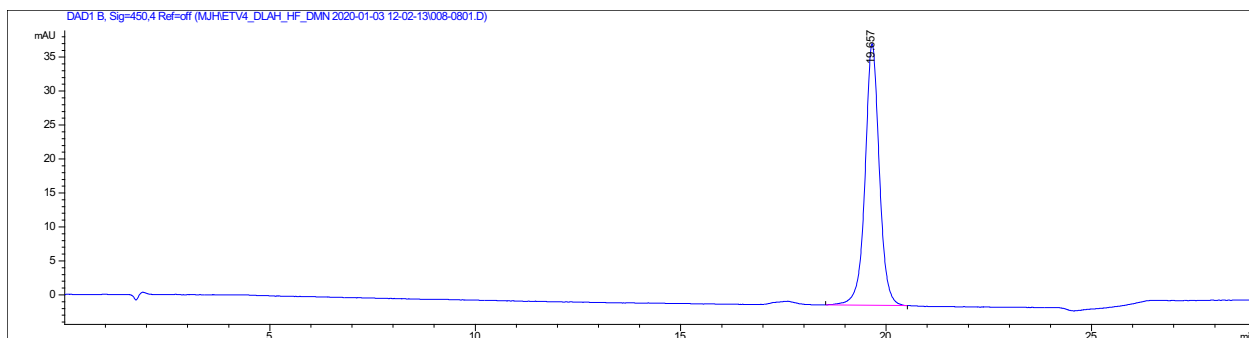

Analytical HPLC trace of **ETV4(45-76)^LPPV**, monitored at 280 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.
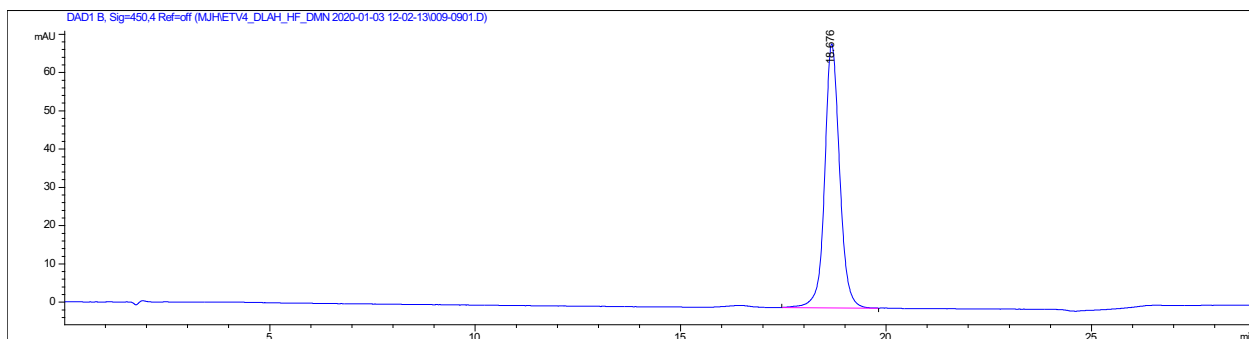
Analytical HPLC trace of **4-DMN-ETV1(38-69)**, monitored at 450 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **4-DMN-ETV4(45-76)**, monitored at 450 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 10-35% acetonitrile over 20 mins.
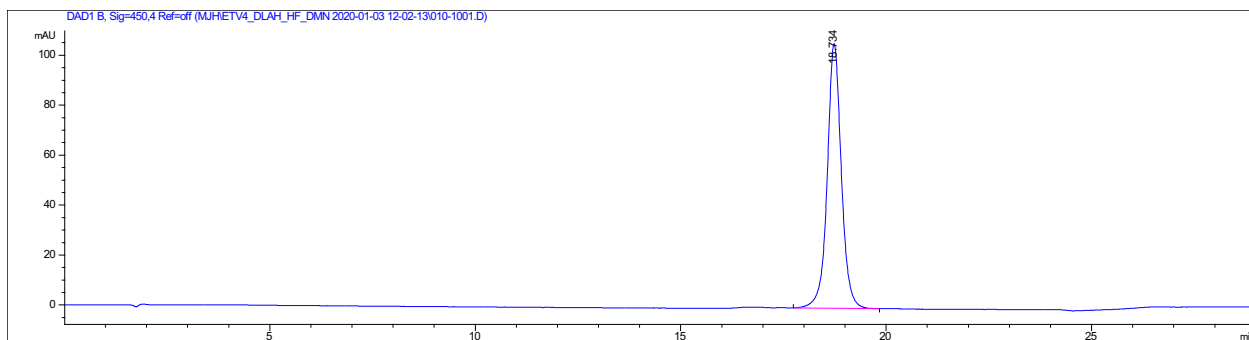

Analytical HPLC trace of **4-DMN-ETV4(45-76)^{QF}**, monitored at 450 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.
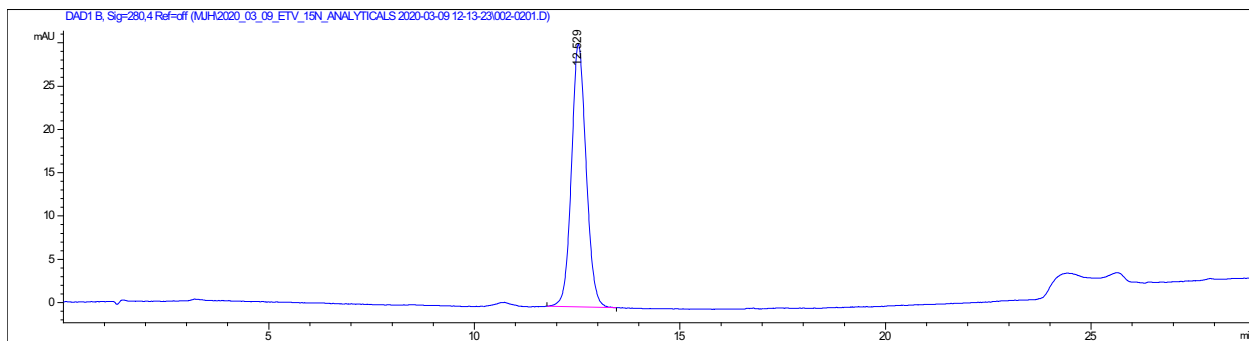
Analytical HPLC trace of **4-DMN-ETV4(45-76)^HL**, monitored at 450 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **4-DMN-ETV4(45-76)^QL**, monitored at 450 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 10-35% acetonitrile over 20 mins.
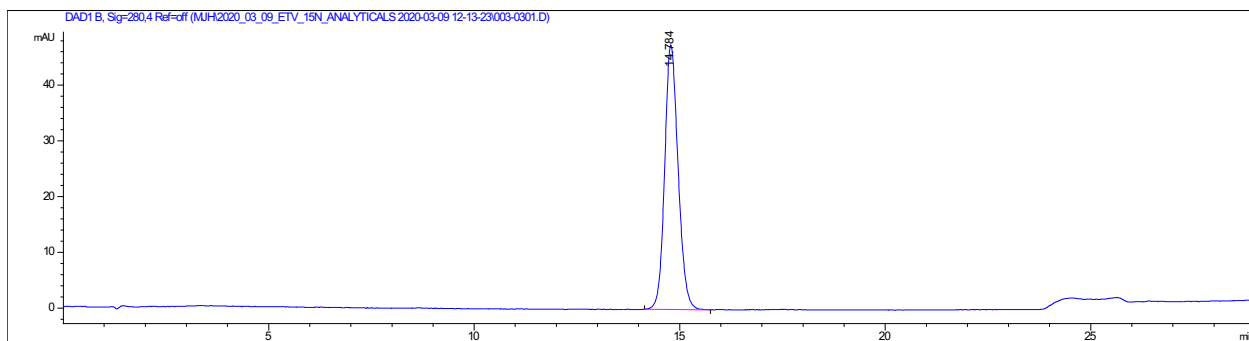
Analytical HPLC trace of **4-DMN-ETV4(45-76)^{DLAH/HF}**, monitored at 450 nm. Analytical sample was run in a water (with 100 mM ammonium acetate)/ acetonitrile system. The sample was injected with an isocratic flow of 70% water (with 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **4-DMN-ETV4(45-76)^{DLAH/QF}**, monitored at 450 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 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **4-DMN-ETV4(45-76)**^{DLAH/HL}, monitored at 450 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 10-35% acetonitrile over 20 mins.


Analytical HPLC trace of **4-DMN-ETV1(42-69)**, monitored at 450 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 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **4-DMN-ETV4(49-76)**, monitored at 450 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 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **ETV1(38-69) Leu39 ^{15}N**, 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 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Analytical HPLC trace of **ETV4(45-76) Leu48 ^{15}N**, 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 100 mM ammonium acetate) and 30% acetonitrile. After 2 mins, the solvent gradient was increased from 10-35% acetonitrile over 20 mins.

Appendix B

Kinetic Data for ETV/PEA3 Variants

This appendix contains raw kinetic data for all ETV/PEA3 variants tested in Chapter 3.
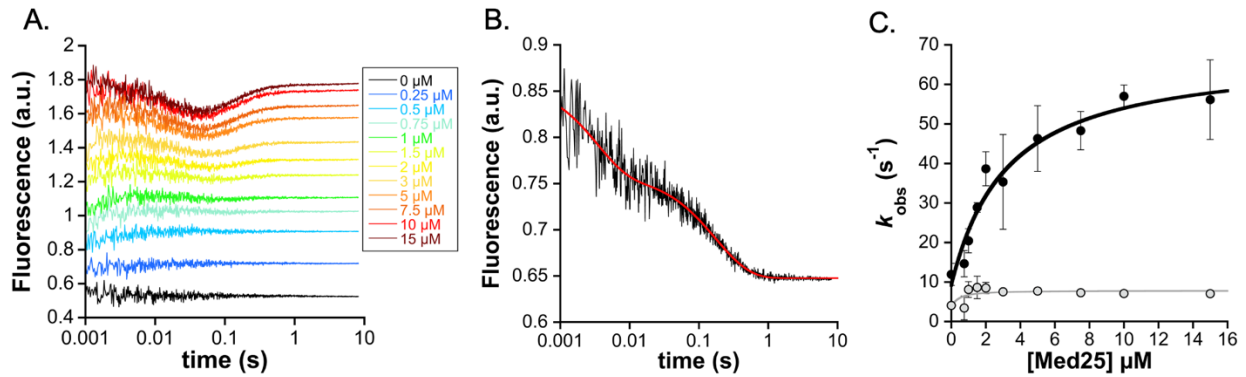
A. Association traces for binding of 0.25 µM 4-DMN-ETV1(38-69) to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV1(38-69) prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV1(38-69). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=3 biological replicates, and error represents the standard deviation.
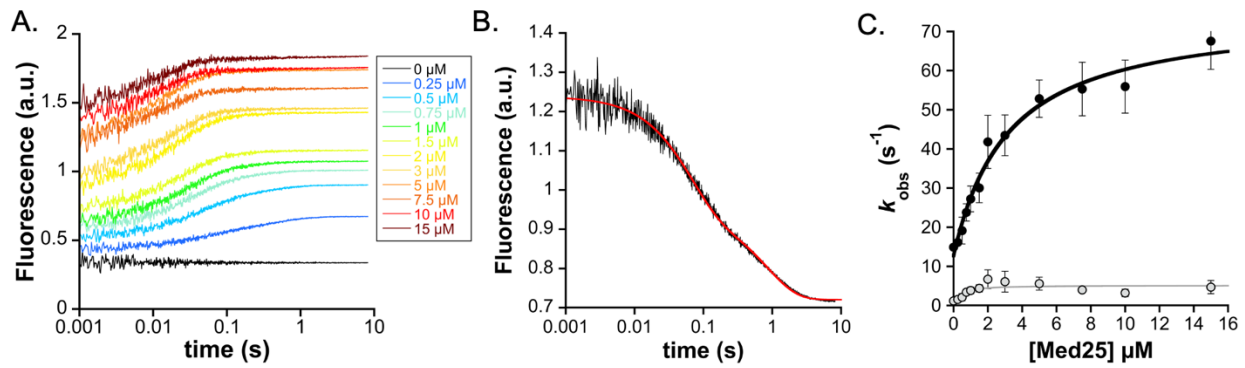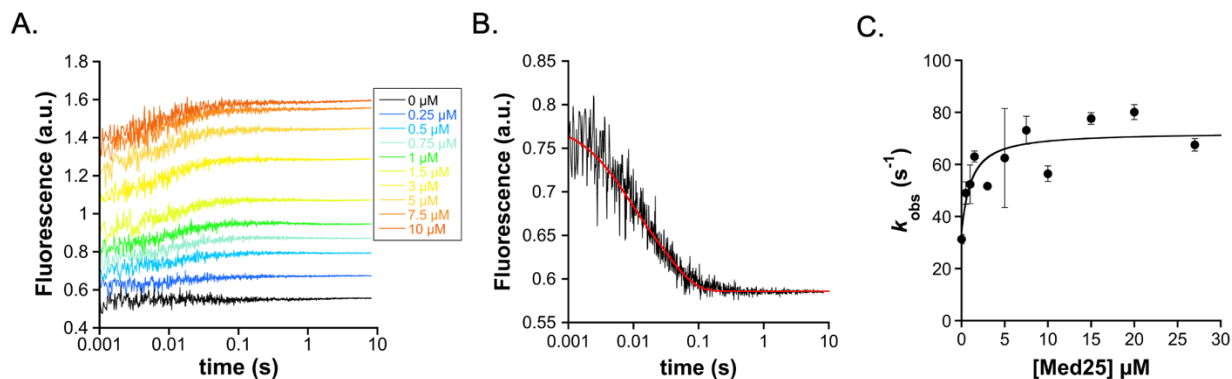

A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76) to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76) prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=4 biological replicates, and error represents the standard deviation.
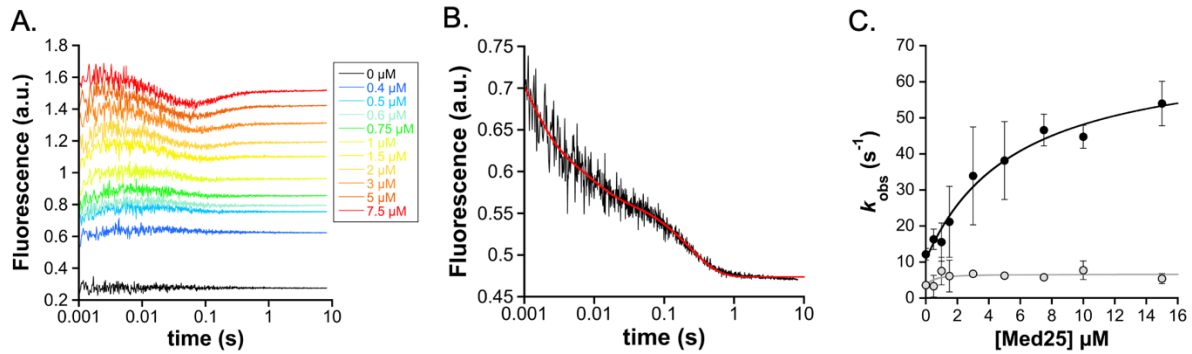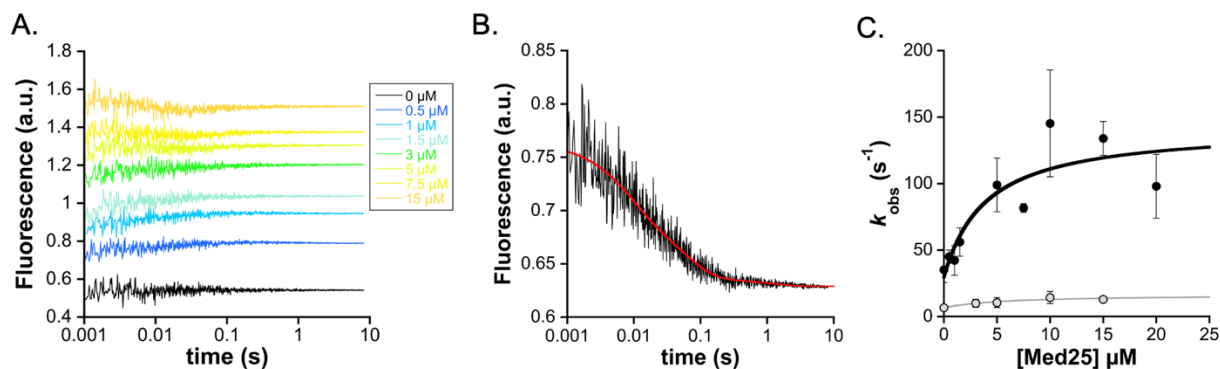
A. Association traces for binding of 0.25 µM 4-DMN-ETV5(38-68) to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV5(38-68) prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV5(38-68). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=3 biological replicates, and error represents the standard deviation.
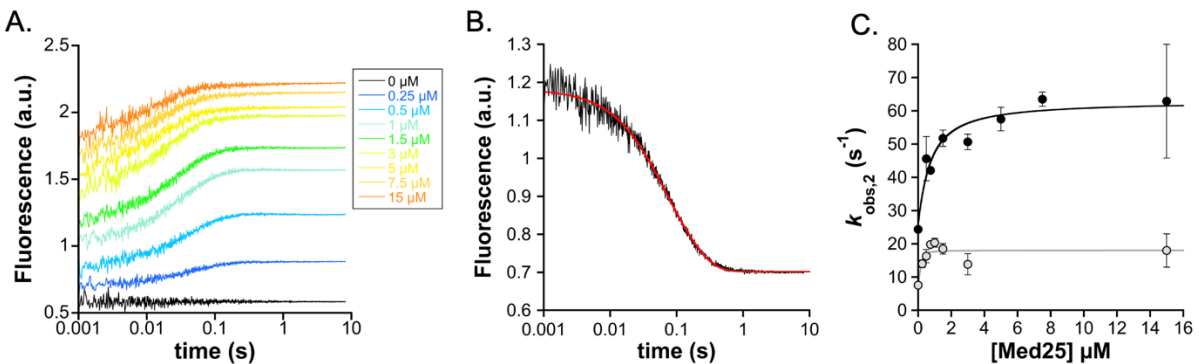
A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76)^{QF} to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76)^{QF} prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.
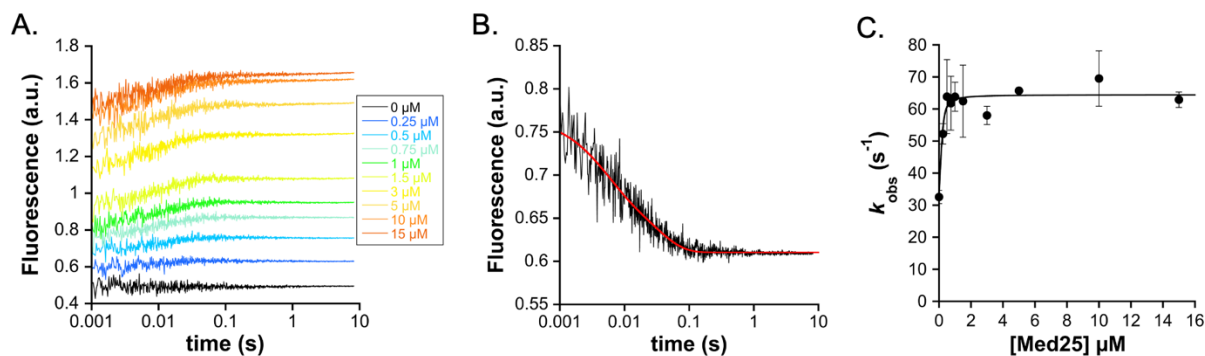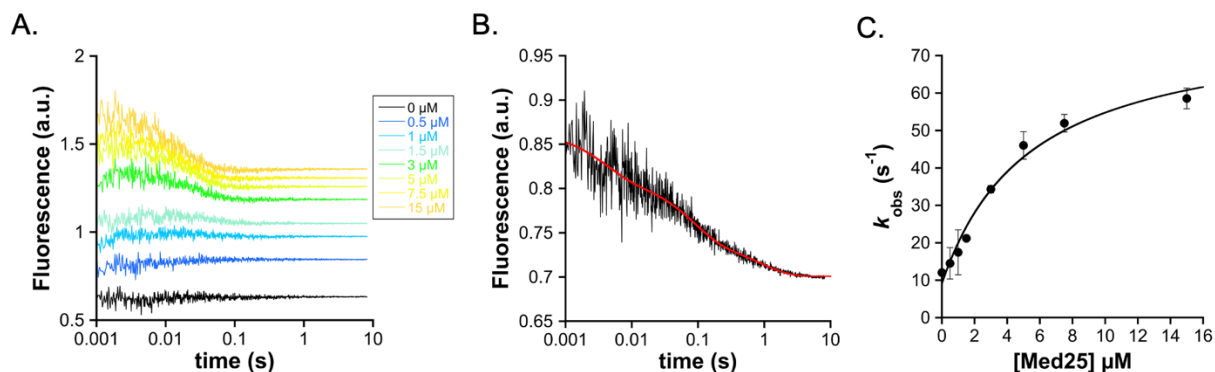
A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76)^{QL} to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76)^{QL} prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.
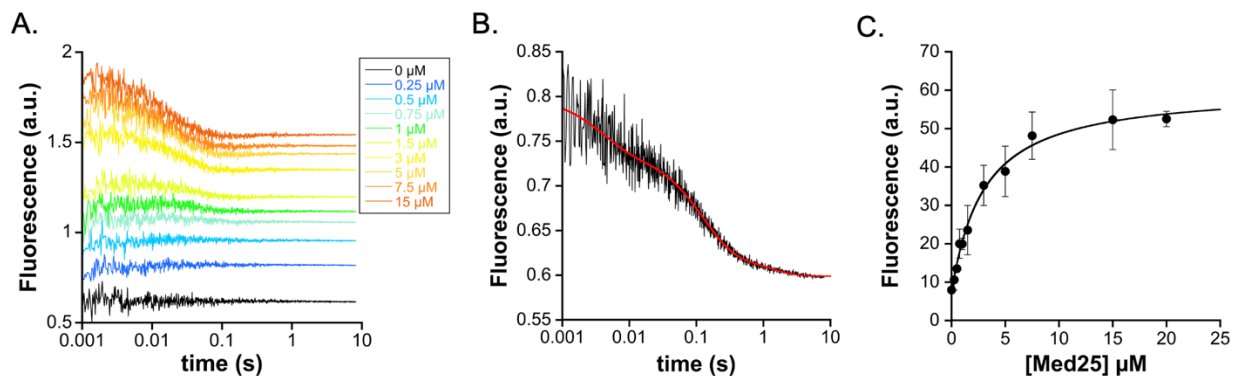

A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76)^{HL} to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76)^{HL} prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.
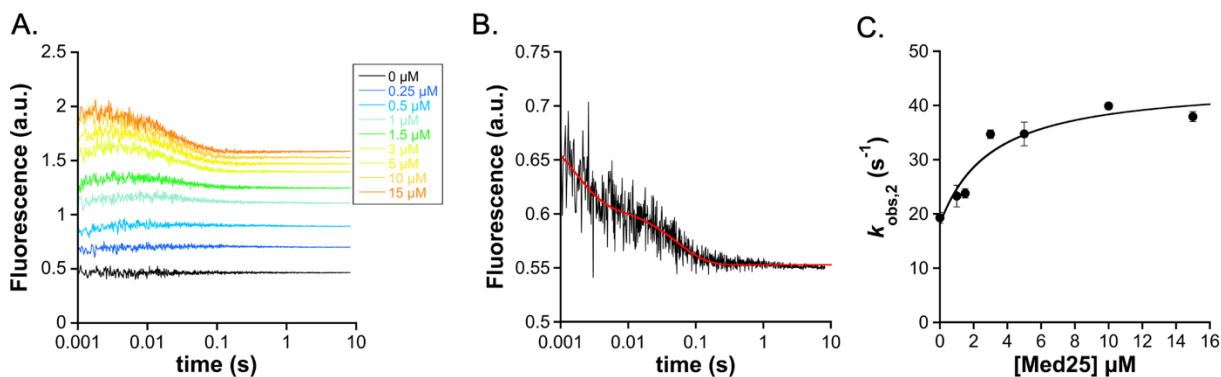
A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76)^{DLAH/HL} to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76)^{DLAH/HL} prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.
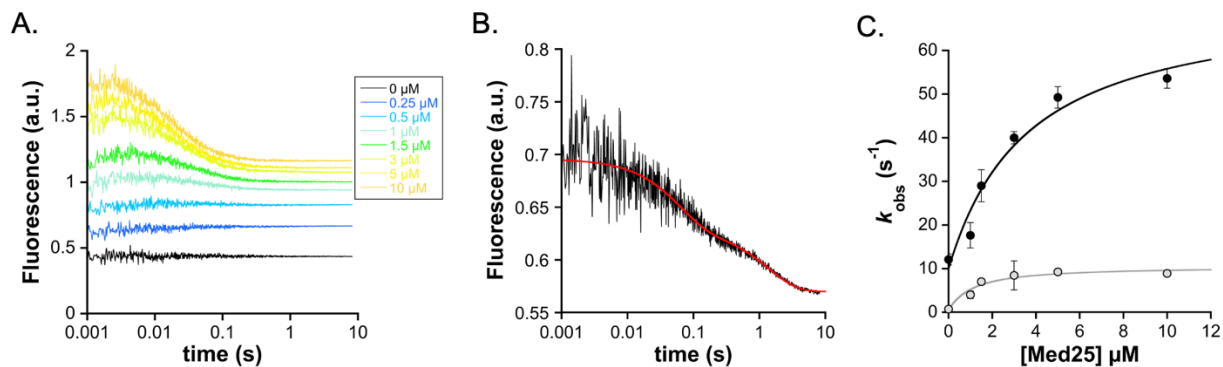

A. Association traces for binding of 0.25 µM 4-DMN-ETV4(45-76)^{DLAH/HF} to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(45-76)^{DLAH/HF} prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.

A. Association traces for binding of 0.25 μM 4-DMN-ETV4(45-76)$^{DLAH/QF}$ to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 μM 4-DMN-ETV4(45-76)$^{DLAH/QF}$ prebound to 0.5 μM Med25, mixed with 50 μM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.

A. Association traces for binding of 0.25 μM 4-DMN-ETV1(42-69) to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 μM 4-DMN-ETV1(42-69) prebound to 0.5 μM Med25, mixed with 50 μM unlabeled ETV1(38-69). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.

A. Association traces for binding of 0.25 µM 4-DMN-ETV4(49-76) to Med25. Concentrations of Med25 are as noted. B. Dissociation trace from 0.25 µM 4-DMN-ETV4(49-76) prebound to 0.5 µM Med25, mixed with 50 µM unlabeled ETV4(45-76). C. Observed rate constants for conformational change phases from curve fitting (black=fast phase, grey=slow phase). Values are average of n=2 biological replicates, and error represents the standard deviation.