

KINETIC AND CONFORMATIONAL
CHARACTERIZATION OF
TRANSCRIPTIONAL ACTIVATOR-
COACTIVATOR INTERACTIONS

by

Ningkun Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2013

Doctoral Committee:

Professor Anna K. Mapp, Chair
Professor Hashim M. Al-Hashimi
Professor Carol A. Fierke
Professor Stephen W. Ragsdale

© Ningkun Wang

All rights reserved, 2013

Acknowledgements

Arriving at this final stage of my PhD career was made possible by so many people. First of all I would like to thank my advisor Prof. Anna K. Mapp. Over the years her patience, encouragement and guidance have helped me grow to be a more mature scientist. Her fearless approach on research greatly inspires me, and the scientific freedom she has given me has taught me to try and be an independent thinker. She has been a great mentor on life in general as well, and I have learned so much from her about things such as how to effectively communicate my science to people and trying to maintain a work-life balance. She has also been very supportive and helpful in the process of my trying to decide my career path. I feel very fortunate to be able to be a student of hers. Her advice for me and outstanding example she has set will follow me for the rest of my career.

I would also like to thank my committee members for spending time with me and giving me ideas on my research. It was always great to talk with them and hear about their ideas regarding my research topics during committee meetings, there were so many ideas bouncing off the walls and I've learned so much from those conversations. I'd like to especially thank Professor Carol Fierke, who has sat down with me multiple times to help me with my kinetic data despite her busy schedule. Similarly, I'd like to thank Professor Hashim Al-Hashimi. We have developed several exciting collaborations and it's always great to talk to him.

I would also like to thank members of the Mapp lab. Especially Amberlyn, who was a senior grad student when I joined and was my mentor and basically introduced me to science and the lab. She taught me so many things I know about science, even now I hear her in my head when I'm doing certain procedures. She has also been a great friend and has looked out for me, and I have looked up to her ever since I joined the lab. I would also like to thank Chinmay, Amy, Jonas and Chris, who were senior grad students in the lab and selflessly helped me in much of my science and career. I feel that along with Amberlyn I owe my thesis work to every one of them. I would like to thank Amanda for being there as a peer and support and coffee buddy. I feel lucky to join the lab along with her, and to have her friendship along the way. I would like to

thank the rest of the current members of the Mapp lab: Jean, Paul, JP, Conor, Aaron, Steve, Vick and James. It's always good to bounce ideas back and forth with them and share a beer as well. They have made the Mapp lab a great place to work, both intellectually and socially. I was fortunate to be a mentor to Jean when she was a rotation student. She is extremely smart and inspiring, and I can't wait to see the great work she will accomplish in the future.

I would like to thank all the people who helped me in my research: Professor Jeanne Stuckey and Jennifer Meagher on teaching me crystallography; Jeet and Katja in the Al-Hashimi lab and Felicia in the Cierpicki lab on NMR HSQC; John Hsieh, Xin Liu, Yu Chen, Noah Wolfson and Mike Howard in the Fierke lab for helpful conversation and guidance with kinetics and stopped-flow technique. I would also like to thank the labs I've rotated in, the Sherman lab and the Ragsdale lab, and Professors Ragsdale and Sherman as well as my mentors there, Tonia Bucholz, Chris Rath, Gunes Bender and Elizabeth Pierce. They have taught me many things that still benefit me, and have been good friends.

I would like to dearly thank fellow students in the Chemical Biology program, especially those from my cohort, Amanda, Rachel, Elin, Hugo, Brian, Chenxi and Eli, also Steffen, Wendy, Felicia and Noah who's basically honorable Chem Bio. I would of course like to thank Katie and Elaina as well who were a year above me and introduced me to Ann Arbor and Chem Bio. We have started grad school together and have bonded like a family ever since. These years in Ann Arbor were made memorable by their friendship and the laughs and tears shared together. I will cherish their friendship and I hope we will keep in touch in the future.

I would also like to thank friends in the Chemistry Department, Xin Liu, Si Yang, Shuwen Sun, Bo Peng and Yu Chen. Thanks for sharing the good times with the hot pots and the picnics and great conversation.

I would like to thank my parents for believing in me from the very beginning and encouraging me to pursue my dreams and never give up. They have served as great examples for me as well as my moral support, and I love them dearly.

Last but not least I would like to thank Tad Ogorzalek. He is my best friend and partner in crime. He has always been supportive and understanding, and has always encouraged me to push myself further. Going through life with him has been a wonderful ride and it reminds me to put things in perspective.

Table of Contents

Acknowledgements.....	ii
List of Figures.....	viii
List of Tables.....	xi
List of Abbreviations.....	xii
Abstract.....	xiv
CHAPTER 1 INTRODUCTION.....	1
A. Introduction.....	1
B. Transcriptional regulation.....	2
<i>B.1. Overview of transcriptional regulation.....</i>	<i>2</i>
<i>B.2. Transcriptional activators.....</i>	<i>3</i>
<i>B.3. Transcriptional coactivators.....</i>	<i>4</i>
C. Protein-protein interactions (PPIs) in transcription.....	4
<i>C.1. Diverse functions and structural architecture of cellular protein-protein interactions.....</i>	<i>4</i>
<i>C.2. Transcription factors as hub proteins--multispecificity and multifunctionality.....</i>	<i>7</i>
<i>C.3. PPIs between transcriptional activators and coactivators.....</i>	<i>9</i>
D. Mechanisms of transcriptional activator-coactivator interactions.....	12
<i>D.1. In vitro techniques employed for studying the mechanism of transcription factor protein-protein interactions.....</i>	<i>12</i>
<i>D.2. Summary of activator-coactivator interaction mechanisms.....</i>	<i>14</i>
E. Small molecules as modulators of the transcriptional PPI network.....	17
<i>E.1. Targeting transcriptional activator-coactivator interactions with small molecules.....</i>	<i>17</i>
<i>E.2. Examples of small molecule modulators of activator-cofactor interactions.....</i>	<i>18</i>
F. Thesis summary.....	20
G. References.....	21
CHAPTER 2 KINETIC CHARACTERIZATION OF A TRANSCRIPTIONAL COACTIVATOR (MED15) IN COMPLEX WITH TRANSCRIPTIONAL ACTIVATORS VP16, GAL4 AND GCN4.....	28
A. Introduction.....	28

B. Background	29
<i>B.1. Transcriptional activation domains (TADs) VP16, Gal4 and Gcn4</i>	30
<i>B.2. Transcriptional coactivator Med15 (Gal11)</i>	33
<i>B.3. Controversy in kinetic models of activator-coactivator interactions</i>	34
<i>B.4. Stopped-flow technique utilized in characterizing transient-state kinetics</i>	35
C. Experimental Design	37
<i>C.1. Gal4 DBD-VP16, Gal4 and Gcn4 TAD fusion proteins as transcriptional activators.</i>	37
<i>C.2. Solubility tags for recombinant protein expression</i>	39
<i>C.3. Experimental setup of fluorescence stopped-flow assay</i>	39
D. Results	40
<i>D.1. Three activators show different levels of potency in S. cerevisiae</i>	40
<i>D.2. Binding affinity of activator•DNA complex to the coactivator Med15</i>	41
<i>D.3. Transient-state analysis of activator-coactivator association reveals biphasic binding kinetics</i>	42
<i>D.4. Microscopic kinetic rate constants calculated for the two-step binding model</i>	42
<i>D.5. Global kinetic simulation studies suggest that Med15 binding precedes the conformational change.</i>	429
E. Discussion	50
F. Conclusions	52
G. Experimental Methods	53
H. References	62

CHAPTER 3 STRUCTURAL CHARACTERIZATION OF TRANSCRIPTIONAL COACTIVATOR KIX IN COMPLEX WITH BINDING PARTNERS	68
A. Introduction	68
B. Background	69
<i>B.1 The KIX domain</i>	69
<i>B.2 Identified small molecule ligands of the CBP/p300 KIX domain and limitations</i>	73
<i>B.3 Tethering and other small molecule ligands stabilizing proteins</i>	75
C. Experimental Design	76
D. Results	77
<i>D. 1. Binding affinity of small molecule-tethered KIX L664C to TAD peptides</i>	77
<i>D.2. Thermal, proteolytic and solvent stability of small molecule-tethered KIX L664C</i>	78
<i>D.3. Crystal structure of 1-10—KIX L664C</i>	81

D.4 ^1H - ^{15}N -HSQC analysis of small molecule-tethered KIX conformational change when binding to ligands at the MLL binding site	83
E. Discussion and Conclusions	88
F. Future Directions	91
H. Experimental Methods	92
G. References	101
CHAPTER 4 KINETIC CHARACTERIZATION OF A TRANSCRIPTIONAL COACTIVATOR (KIX) IN COMPLEX WITH PEPTIDE AND SMALL MOLECULE BINDING PARTNERS	105
A. Introduction	105
B. Background	106
B.1. The KIX domain interacts with the TADs of MLL, c-Myb and pKID	106
B.2. The study of transient-state kinetics of IDP protein-protein interactions	109
C. Experimental Design	111
C.1. Selection of TAD constructs	111
C.2. Complexes of KIX used in binding experiments	113
D. Results	114
D.1. Stopped-flow of MLL and pKID binding KIX	114
D.2. Cooperativity of MLL and pKID binding to KIX	117
D.3. The allosteric effect of 1-10 tethering on the pKID binding affinity of different KIX mutants	118
D.4. Stopped-flow spectroscopy of different states of KIX binding to pKID	120
D.5. Φ values for 1-10 —KIX N627C and 1-10 —KIX L664C suggest 1-10 elicits different allosteric effects when tethered at different positions	124
E. Discussion and Conclusions	125
F. Future Directions	128
F.1. Conformational analysis of small molecule tethered KIX	128
F.2. FRET or fluorescence quenching stopped-flow	128
G. Experimental Methods	129
H. References	134
CHAPTER 5 CONCLUSIONS AND FUTURE DIRECTIONS	137
A. Conclusions	137
A.1. Transient-state kinetic mechanisms of activator-coactivator interactions	137

<i>A.2. Allosteric effects of a small molecule modulator on activator-coactivator interactions</i>	138
B. Future Directions	140
<i>B.1. Application of environmentally sensitive fluorophores</i>	140
<i>B.2. In cell NMR</i>	142
C. References	143
APPENDIX	146
A. Introduction	146
B. Background	147
<i>B.1. Structural information of Gal4</i>	147
<i>B.2 Using NMR to gain structural information of conformationally dynamic proteins</i>	148
<i>B.3. An NMR compatible solubility tag GB1</i>	149
C. Using ^1H-^{15}N HSQC to study allosteric effects between Gal4 domains	150
<i>C.1. Gal4 TAD exhibits increased affinity to Med15 when fused to Gal4 DBD</i>	150
<i>C.2. ^1H-^{15}N HSQC data of GB1-tagged Gal4 constructs</i>	151
D. Future Directions	153
E. Experimental Methods	154
F. References	159

List of Figures

Figure 1.1 Cartoon representation of the transcriptional activation process	2
Figure 1.2 Network diagram showing a map of protein-protein interactions.....	5
Figure 1.3 Example of protein-protein interactions.....	6
Figure 1.4 Schematic representation of protein interaction networks	7
Figure 1.5 Cartoon depiction of protein-protein interaction types	9
Figure 1.6 Interaction networks of transcription factors	10
Figure 1.7 Timeline of the identification of transcriptional activator targets.....	11
Figure 1.8 Schematic of coupled binding and folding.....	16
Figure 1.9 Structures of transcriptional activator-coactivator complexes	16
Figure 2.1 Schematic of transcriptional activation	30
Figure 2.2 Structures of VP16, Gal4 and Gcn4.	30
Figure 2.3 Mediator and Med15	34
Figure 2.4 Schematic of an average stopped-flow setup	36
Figure 2.5 Dissociation constants of activators for DNA	38
Figure 2.6 Analytical gel filtration data for GST-Med15 (1-345).....	39
Figure 2.7 Activities of the Gal4 (1-100)-TAD activators.....	41
Figure 2.8 Dissociation constants of DNA-bound activators for Med15 (Gal11).....	42
Figure 2.9 Biphasic kinetics observed for activator•DNA complex binding to Med15 ..	43
Figure 2.10 Kinetics of activator•DNA complex binding to Med15 at 16 °C.....	44
Figure 2.11 Negative control stopped-flow experiments.....	45
Figure 2.12 Plot of $k_{obs,1}$ and $k_{obs,2}$ of the activators against Med15 concentration	46
Figure 2.13 Schematics of two limiting binding models	47
Figure 2.14 Plot of the ratio of amplitudes.....	47
Figure 2.15 Stopped-flow traces overlaid with simulation traces	49
Figure 2.16 Comparison of equilibrium constants and activity	51
Figure 2.17 Plot of the amplitudes of the slow phase.....	61

Figure 3.1 The KIX domain.....	69
Figure 3.2 Structure of the KIX domain of CBP in complex with two TADs.....	70
Figure 3.3 The allosteric network of KIX	71
Figure 3.4 KIX: c-Myb: MLL ternary complex and KIX: c-Myb binary complex	72
Figure 3.5 Schematic of the Tethering screen	76
Figure 3.6 Direct binding affinity of TAD peptides to KIX.	78
Figure 3.7 Bar graph of KIX melting temperature.....	79
Figure 3.8 Melting temperature curve by CD.	79
Figure 3.9 Proteolytic stability of KIX complexes	81
Figure 3.10 Refined crystal structure of 1-10 --KIX L664C	82
Figure 3.11 ¹ H- ¹⁵ N-HSQC of KIX complexes.	84
Figure 3.12 Chemical shift perturbation experiment of 1-10 -tethered KIX L664C.....	85
Figure 3.13 Chemical shift perturbation experiment of 2xMLL complexed KIX L664C.....	85
Figure 3.14 Chemical shift perturbation experiment of 2-64 -tethered KIX L664C.....	87
Figure 3.15 Molecular dynamics simulations.	89
Figure 4.1 Overlay structure of KIX with MLL, pKID and c-Myb.....	107
Figure 4.2 Coupled folding and binding mechanism for the pKID-KIX interaction.....	108
Figure 4.3 Gal4 constructs binding to DNA.....	111
Figure 4.4 DNA•Gal4 constructs binding to KIX.....	112
Figure 4.5 Sequences of MLL and pKID TAD peptides.....	112
Figure 4.6 Position of FITC relative to the TAD-KIX complex.....	113
Figure 4.7 Structures of fragments 1-10 and 2-64	114
Figure 4.8 Position of mutations L664C and N627C on KIX	114
Figure 4.9 Stopped-flow traces of control experiments	115
Figure 4.10 Representative stopped-flow traces of MLL and pKID binding to KIX.....	116
Figure 4.11 Observed association rates plotted against KIX concentration	117
Figure 4.12 Cooperativity between MLL and pKID binding to KIX	118
Figure 4.13 Equilibrium binding of KIX complexes with pKID	119
Figure 4.14 Stopped-flow traces of pKID association and dissociation with KIX	120
Figure 4.15 k_{obs} of pKID binding KIX complexes plotted against KIX concentration..	121
Figure 4.16 Bar graphs of binding constants between pKID and KIX complexes.....	123

Figure 4.17 Energy diagram of the KIX-pKID interaction	126
Figure 5.1 Schematic of the activator-coactivator interaction mechanism	138
Figure 5.2 Structure of 1-10 —KIX L664C and positions of Cys mutations	139
Figure 5.3 Schematic of utilizing solvatochromic fluorophores	141
Figure 5.4 Artistic representation of a cross-section of an <i>E. coli</i> cell	142
Figure A.1 Structures of Gal4 domains	147
Figure A.2 Schematic of components of the two separate assays	151
Figure A.3 ^1H - ^{15}N -HSQC spectrum of ^{15}N GB1-Gal4 (840-881)	152
Figure A.4 Overlay of the ^1H - ^{15}N -HSQC spectra of ^{15}N labeled Gal4 constructs	153

List of Tables

Table 1.1 Techniques used in probing activator-coactivator interactions.	12
Table 1.2 Small molecule inhibitors of transcriptional activator-coactivator interactions	18
Table 2.1 Experimentally determined thermodynamic and kinetic constants	42
Table 2.2 Calculated rate constants for binding model in Scheme A	48
Table 2.3 Calculated rate and equilibrium constants for binding model in Scheme B.....	48
Table 3.1 Small molecules that target the KIX domain of CBP/p300	74
Table 3.2 K_{DS} for KIX constructs interacting with MLL and pKID.....	78
Table 3.3 Difference in number of amide protons protected from solvent.....	80
Table 3.4 Summary of fragment-protein complexes screened for crystallization.	81
Table 3.5 Data collection, phasing and refinement statistics.....	100
Table 4.1 Summary of K_D values for pKID binding to KIX complexes	119
Table 4.2 On-rates of pKID binding to KIX complexes	121
Table 4.3 Off-rates of pKID dissociating from KIX complexes	121
Table 4.4 Binding Φ values for 1-10 tethered KIX mutants	125
Table 4.5 Peptide sequence and ESI-MS characterization.....	131

List of Abbreviations

4-DMN	4-N,N-DiMethylamino-1,8-Naphthalimide
ABD1	Activator Binding Domain 1 in Med15
ACTR	Activator of Thyroid and Retinoid receptors
AHD	ANC1 Homology Domain
BSA	Bovine Serum Albumin
CBP	CREB Binding Protein
CSL	(CBF1/RBP-J, Su(H), Lag-1)
DBD	DNA Binding Domain
DNA	DeoxyriboNucleic Acid
DR50	Dose Response 50
DTT	DiThioThreitol
E1A	Adenovirus-encoded E1A 13S protein
EG	Ethylene Glycol
ESX	Epithelial-Specific Transcription factor
FOXO3a	Forkhead box class O 3a
FRET	Förster Resonance Energy Transfer
GACKIX	Gal11, Arc105, Creb Binding Protein KIX domain
GB1	G protein B1 domain
GST	Glutathione Transferase
HAT	Histone Acetyl Transferase
HBZ	HTLV-1 Basic leucine Zipper factor
HDAC	Histone DeAcetylase
HIF-1α	Hypoxia-Inducible Factor 1 α subunit
hrs	Hours
HTLV-1	Human T-cell Leukemia Virus type 1
IDP	Intrinsically Disordered Protein
IPTG	IsoPropyl- β -D-1-ThioGalactopyranoside
IRF3	Interferon Regulatory Factor 3
IκBα	Inhibitor of κ B isoform α
KID	Kinase Inducible Domain
KIX	Kinase Inducible Domain Interacting Domain of CBP
LiOAc	Lithium Acetate
MBP	Maltose Binding Protein

MD	Molecular Dynamics
MDM2	Murine Double Minute 2
MEF2	Myocyte Enhancer Factor 2
min	Minutes
MINT	Msx2-Interacting Nuclear Target protein
MLL	Multi-Lineage Leukemia
NCBD	Nuclear Coactivator Binding Domain of CBP
NES	Nuclear Export Sequence
NFκB	Nuclear Factor κB
Ni-NTA	Nickel-Nitrilotriacetic Acid
NLS	Nuclear Localization Sequence
NMR	Nuclear Magnetic Resonance
ONPG	<i>Ortho</i> -NitroPhenyl-β-Galactoside
PC4	Positive Cofactor 4
PDB	Protein DataBase
PEG	Poly-Ethylene Glycol
PMT	PhotoMultiplier Tubes
PPI	Protein-Protein Interactions
RNA	RiboNucleic Acid
RNA Pol II	RNA Polymerase II
rpm	Revolutions Per Minute
s/sec	Seconds
SAGA	Spt-Ada-Gcn5-Acetyltransferase
SDOM	Standard Deviation of the Mean
SPR	Surface Plasmon Resonance
STAT6	Signal Transducer and Activator of Transcription 6
TAD	Transcriptional Activation Domain
TAF	TBP Associated Factor
TAZ2	Transcriptional Adapter Zinc finger 2
TBP	TATA-Binding Protein
TCEP	Tris(2-CarboxyEthyl)Phosphine
TEAD	Transcriptional Enhancer Activator Domain
TFIIB	Transcription Factor II B
VP16	Herpes simplex Virus Protein 16
YAP	Yes-Associated Protein
Zn	Zinc

Abstract

Kinetic and Conformational Characterization of Transcriptional Activator-Coactivator Interactions

by

Ningkun Wang

Chair: Anna K. Mapp

Initiation of transcription is achieved through a series of coupled binding equilibria commenced by interactions between DNA-bound transcriptional activators and coactivators. Abnormalities in either the activator or coactivator can disrupt or enhance the activator-coactivator interaction and result in serious diseases. Hence there is great need to understand the mechanism of these activator-coactivator interactions and design artificial transcriptional regulators as probes or potential therapeutics. However, the key mechanistic features responsible for the differential transcriptional output of these activators are yet to be well-defined. The focus of this dissertation work has been to dissect the kinetic and structural characteristics of transcriptional activator-coactivator interactions and examine the effects of small molecule modulators on these interactions.

Utilizing fluorescence stopped-flow, we measured the transient-state kinetics of the transcriptional activation domains (TADs) of the activators Gal4, Gcn4 and VP16 in their DNA-bound forms binding to the coactivator Med15. We determined that they interact through the

same two-step binding mechanism: an initial rapid bimolecular association step followed by a slower conformational change step. Additional analysis suggests that the tendency for an activator to undergo conformational change correlates with both its overall affinity to the coactivator and its transcriptional activity *in vivo*.

This mechanistic study of activator-coactivator interactions was further applied to the more conformationally defined system of TADs (MLL and pKID) binding cooperatively to the coactivator KIX. The study showed that both TADs bind to KIX through a two-step mechanism similar to that of TADs binding to Med15. A small molecule fragment **1-10** from a Tethering screen covalently tethers to a cysteine mutant of the coactivator KIX domain of CBP at the MLL binding site. The additional stabilizing effect of **1-10** tethering to KIX enabled me to obtain a crystal structure of **1-10**—KIX L664C. This is the first crystal structure of the KIX domain of CBP, and provides a high-resolution snapshot of this domain. Additionally, I found that **1-10** elicits varying allosteric effects on the opposite pKID binding site of KIX, depending on the site at which it tethers. This reveals **1-10** as a powerful small molecule probe to dissect the allosteric mechanisms of the KIX-pKID interaction. I used **1-10** tethered at different cysteine mutations as well as the MLL peptide as probes to study the allosteric effects of KIX's pKID binding through transient-state kinetics. I also observed that the dissociation rate constant k_{off} between pKID and KIX correlated with their overall binding affinity K_{D} . As k_{off} better reflects the slower conformational change step of this interaction, this again suggests that conformational change drives overall affinity between pKID and KIX.

In conclusion, I have elucidated the transient-state kinetic mechanism of two activator-coactivator interaction systems, and have shown the potential of a small molecule as a powerful probe for dissecting interaction mechanisms in ways unique to native peptide ligands.

Chapter 1 Introduction

A. Introduction

In response to certain signals, transcriptional activators bind to their specific DNA recognition sites and recruit proteins known as coactivators followed by components of the transcriptional machinery to the promoter site of the gene, thus initiating transcription (Figure 1.1). Transcriptional activator-coactivator interactions play a key role in regulating the transcription of genes in the cell. Many diseases are known to be caused by abnormalities in transcriptional activators and/or their regulation networks, such as cancer and developmental disorders.¹⁻⁵ This suggests the possibility of developing artificial transcriptional modulators for therapeutic applications.⁶⁻⁸ However, the mechanism of transcriptional activator-coactivator interactions in the transcriptional machinery remains poorly understood and has hampered the progress of targeted development of artificial activator replacements or inhibitors. The work outlined in this dissertation focuses on the elucidation of the mechanism of transcriptional activator-coactivator interactions through the study of their conformation and properties. In this chapter, I will briefly discuss the complex nature of protein-protein interactions, with an emphasis on transcriptional activator-coactivator interactions. In particular, the advancements and limitations in mechanistic studies of these interactions will be discussed, as well as their implications for the discovery of small molecule modulators of transcriptional regulation.

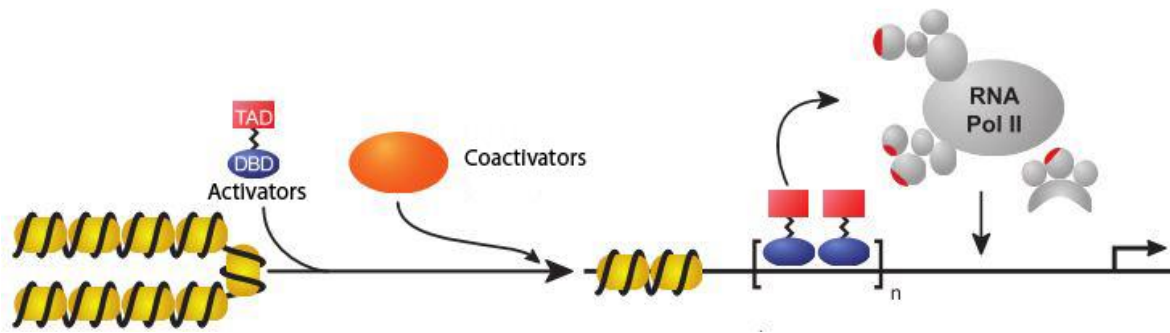


Figure 1.1 Cartoon representation of the transcriptional activation process. A modular transcriptional activator binds to DNA, and in turn recruits coactivators and chromatin remodeling enzymes, which also act as a scaffold to recruit the general preinitiation complex containing RNA Pol II to the promoter region of the gene, thus initiating transcription. Figure adapted from Mapp et al, 2007.⁹

B. Transcriptional regulation

B.1. Overview of transcriptional regulation

Transcription is a highly regulated and complicated process orchestrated by a vast protein network consisting of approximately 300 proteins in yeast and up to 3000 in humans. The ultimate output of this elaborate cascade is the localization of RNA polymerase II (RNA Pol II) to the promoter region of a gene and the catalysis of the template-based synthesis of mRNA. General transcription factors along with RNA Pol II form the preinitiation complex. These are the component that are necessary for basal level transcription.^{10,11} These five general transcription factors, referred to as TFIIB, TFIID, TFIIE, TFIIIF and TFIIH, are responsible for promoter recognition and unwinding the promoter DNA.¹² Apart from these key players, the myriad proteins that regulate the activation and recruitment of the preinitiation complex leading to transcription can be separated into two types. The first type is transcriptional activators that bind directly to DNA. The second type are transcriptional coactivators and co-repressors, protein complexes that do not directly bind to DNA, but instead bind to the transcriptional activators and act as scaffolds to recruit other proteins and/or alter chromatin structure by their own

enzymatic activity.^{13,14} These components along with the preinitiation complex are responsible for activated transcription in cells. The emphasis of this dissertation is the interaction between these transcriptional activators and the coactivators they interact with (Figure 1.1).

B.2. Transcriptional activators

Transcriptional activators are modular proteins with domains that are interchangeable.¹⁵ The major domains are the DNA binding domain (DBD) and the transcriptional activation domain (TAD).^{16,17} Various transcriptional activators also have additional domains such as a nuclear localization sequence (NLS) or a nuclear export sequence (NES), and for certain transcriptional activators such as hormone receptors there is a ligand binding domain.¹⁸ DNA binding domains usually bind to the gene response elements on DNA with high specificity and high affinity (nanomolar K_D)¹⁹ via several different types of motifs. A few well studied motifs include the helix-turn-helix motif,²⁰ the zinc finger motif²¹ and a more simple zinc cluster motif,²² the basic leucine zipper motif,²³ and the basic helix-loop-helix motif.²⁴ Unlike the well-studied DBDs, there is less information known about transcriptional activation domains (TADs), which are responsible for recruiting coactivators and the general transcriptional machinery via protein-protein interactions. These domains tend to be less structured and are often rich in acidic amino acids.²⁵ Many TADs are hub proteins (which will be discussed in the following section) that can bind to multiple binding partners at either the same or different sites with high specificity.²⁶⁻³⁰ However, the mode of action of TADs, such as their specific binding partners and the mechanism of the binding interaction, remains unclear, and there is much to be learned.

B.3. Transcriptional coactivators

Transcriptional coactivators are a class of proteins that modulate transcriptional activity by protein-protein interactions between transcriptional activators and the general transcription factors along with RNA Polymerase II. One important coactivator is the Mediator complex. Mediator is a 25-30 subunit protein complex that is unique to eukaryotes, and is conserved from yeast to humans.³¹⁻³³ It acts as a molecular bridge connecting transcriptional activators with RNA Pol II.³⁴ Although it is not required for basal level transcription *in vitro*, Mediator is an integral part of the preinitiation complex, and is sometimes considered a general transcription factor instead of a coactivator. The other transcriptional coactivators can be organized into two classes: chromatin modification complexes and chromatin remodeling complexes. Chromatin modification complexes such as histone acetyl transferases (HATs)³⁵ act by post translationally modifying histones (acetylation, methylation, phosphorylation, ubiquitination, sumoylation, etc.) to control access of proteins to the DNA. Chromatin remodeling complexes, such as the SWI/SNF complex, slide and disassemble the nucleosome to allow transcription factors to bind to DNA.³⁶

The emphasis of my thesis work is on the mechanism of the protein-protein interactions between transcriptional activators and coactivators, which will be discussed in the following section.

C. Protein-protein interactions (PPIs) in transcription

C.1. Diverse functions and structural architecture of cellular protein-protein interactions

Protein-protein interactions (PPIs) govern most cellular processes. The protein interactome is an intricate and complicated network often depicted as an interwoven web of interaction nodes

and edges (Figure 1.2). It has been estimated that a human cell may contain approximately 130,000 binary interactions between proteins.³⁷ Even in one of the simpler eukaryotic organisms, *Saccharomyces cerevisiae*, there is estimated to be 16,000-25,000 different interaction pairs (excluding homotypic interactions). Enormous effort has been directed towards this field of study in the recent decades and, despite considerable advancements in the understanding of PPIs, there is still a great amount that is unknown about these interactions, such as the underlying kinetic mechanisms and conformational dynamics that enable them to intricately and precisely regulate cellular function.^{14,30,31,38,39}

There is a great range of diversity in the nature of protein-protein interactions, from strong,

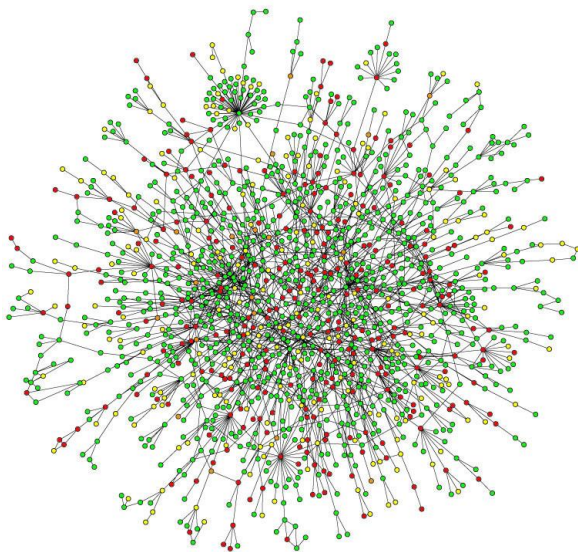


Figure 1.2 Network diagram showing a map of protein-protein interactions in a yeast (*Saccharomyces cerevisiae*) cell. This cluster includes 78 percent of the proteins in the yeast proteome. The color of a node represents the phenotypic effect of removing the corresponding protein (red, lethal; green, nonlethal; orange, slow growth; yellow, unknown). Credit: Hawoong Jeong, KAIST, Korea.

long-lived complex formations to weak, transient interactions.⁴⁰ Of particular interest to us are transient protein-protein interactions, where the separate protein monomers can exist either on their own or in complex *in vivo*, and dynamically interchange between free and complex form.⁴⁰ These PPIs play an important role in cell signaling and homeostasis, and can respond to extra- and intra-cellular changes. To this end, changes in the local environment can cause such actions as protein

oligomerization,⁴¹ changes in local protein concentration (such as in the nucleosome),⁴² and changes in physiological conditions (such as pH, temperature or ionic strength)^{43,44} that would affect the binding affinities of the proteins.^{45,46} Furthermore, these changes could initiate a molecular signal by allosteric or cooperative ligand binding or covalent modification.⁴⁵ Different transient PPIs are commonly compared by examining several properties of the protein-protein interaction interface, including the contact surface area (which can range from ~500 Å² to ~2000 Å²), the planarity of the interface surface, the percent of polar residues at the interface, and the GAP index (which describes the degree of complementarity of the interacting surfaces).⁴⁵ For example, weaker interactions will often have a smaller and flatter interface and less complementary surfaces (Figure 1.3).⁴⁶ It has also been shown that usually a cluster of a few amino acids at the binding interface contribute to most of the binding energy, called the “hot spots” or “hot spot regions”, and designing mutations or small molecules that target these hot spots can be most efficient in modulating the protein-protein interaction.^{47,48}

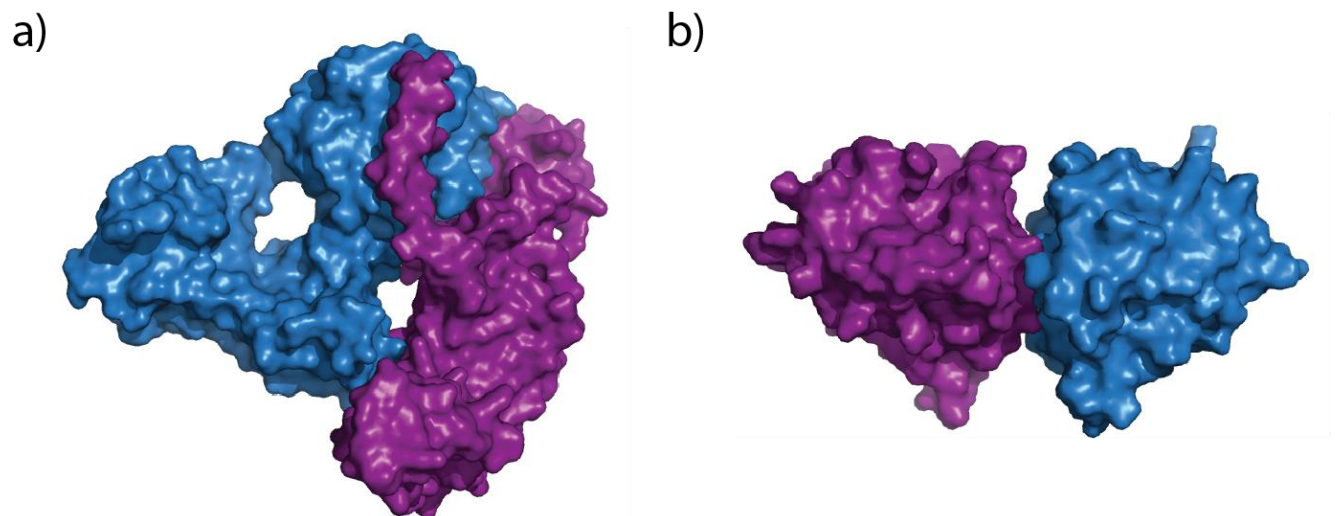


Figure 1.3 Example of protein-protein interactions of high affinity and large contact surface area and low affinity and low contact surface area. **a)** Heterodimer of Ef-Ts (purple) and Ef-Tu (blue) (PDB 1EFU), with binding affinity of 30 nM and contact surface area of 1956 Å².⁴⁹ **b)** Homodimer of β-Lactoglobulin (PDB 1BEB), with binding affinity of 20 μM and contact surface area of 529 Å².⁵⁰

C.2. Transcription factors as hub proteins--multispecificity and multifunctionality

In the PPI network, of specific interest are the proteins that are at the “nodes” of the web, otherwise known as hub proteins.^{51,52} Hub proteins that bind to different proteins at the same binding surface at different times are generally referred to as single-interface hubs, while hub proteins that are able to bind to multiple proteins at the same time are referred to as multi-interface hubs (Figure 1.4).⁵³

Many transcription factors have been shown to be hub proteins, such as subunits of the Mediator complex and the paralogue coactivators CBP/p300.^{51,54–56} Much research has been directed towards the ability of these hub proteins to exhibit multispecificity, as well as multifunctionality, their ability to be involved in multiple non-overlapping pathways.

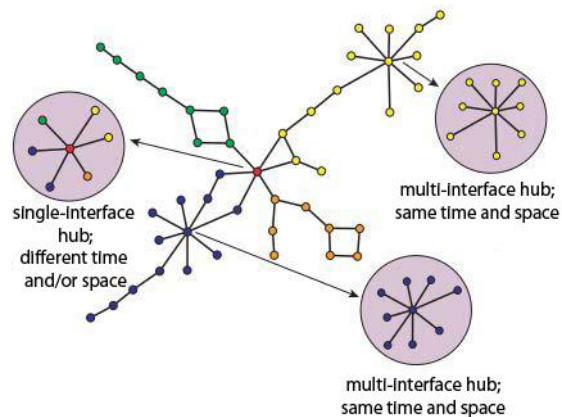


Figure 1.4 Schematic representation of protein interaction networks. Single-interface hubs bind to different partners at different times while multi-interface hubs bind to multiple partners at the same time. Modified from Han et al, 2004.⁵³

A number of studies suggest that structural plasticity plays a key role in the multispecificity of hub proteins.^{57–59} The ability to be structurally flexible enables these proteins to undergo conformational change and adapt to various partner proteins as required. Studies also show that while the single-interface hub proteins tend to be more disordered, both single- and multi-interface hubs have highly structured binding interfaces; however the binding partners of these hubs tend to be highly disordered.⁵⁹ This is exemplified in the example of the hub protein, the KIX domain of CBP: KIX is a three helix bundle, however the TADs binding to KIX at two

different sites (MLL, pKID, c-Myb and p53) are all disordered in isolation and only assume structure when bound.^{30,60-62} A large-scale study on the comparison of non-hub interactions and hub interactions (mainly single-interface hub proteins) summarized several characteristics of common interface motifs of hub proteins: lower conservation of “hot spot” clusters (known as “hot regions”) at the interface; smaller accessible surface area at a more planar interface; a high prevalence of α -helices at multifunctional interfaces; and imperfect packing at the interface with more polar residues. These characteristics could also explain the multispecificity of hub proteins.

How transcription factor hub proteins still retain a level of specificity and are able to function in multiple pathways without creating debilitating overlaps and “short circuiting” is a more complicated question. Several hypotheses regarding this phenomenon have been put forth, but researchers are far from fully understanding this mechanism. Various studies suggest hub proteins can specifically recognize which binding partners to bind to through mechanisms such as “switches” in the form of post translational modifications and interfaces being used in pathways in a sequential matter when a substrate-such as ubiquitin-must be passed from one protein to another.⁶³ In a more general observation, an important factor again appears to be the conformational flexibility of the protein.⁶⁴ While buried interface residues tend to be important for recognizing multiple partners, the exposed interface residues impart specificity for a particular ligand, and diverse side-chain torsion angles of hub proteins are important in accommodating binding to multiple targets.⁶⁵

Additionally, allosteric interactions play a key role in multifunctionality and regulation as well.⁶⁶ This is commonly seen among transcription factor hub proteins such as the KIX domain of CBP/p300.⁶⁷⁻⁶⁹ This is discussed in detail in Chapters 3 and 4.

C.3. PPIs between transcriptional activators and coactivators

As with any complicated signaling network, protein-protein interactions is an integral component in the mechanism by which transcriptional activators and coactivators work together to regulate transcription. These PPIs can be classified into two categories: more stable interactions between several subunits of a complex, commonly found in coactivator complexes such as SWI/SNF, SAGA and Mediator (Figure 1.55a); and weaker, transient interactions between transcriptional activators and coactivators/suppressors (Figure 1.5b). While the stable complexes are relatively well defined now by techniques such as co-purification⁷⁰ and affinity capture⁷¹ and some even directly visualized by x-ray crystallography⁷² and electron microscopy³², the transient activator-coactivator interactions are more elusive and difficult to characterize, both conformationally and kinetically.

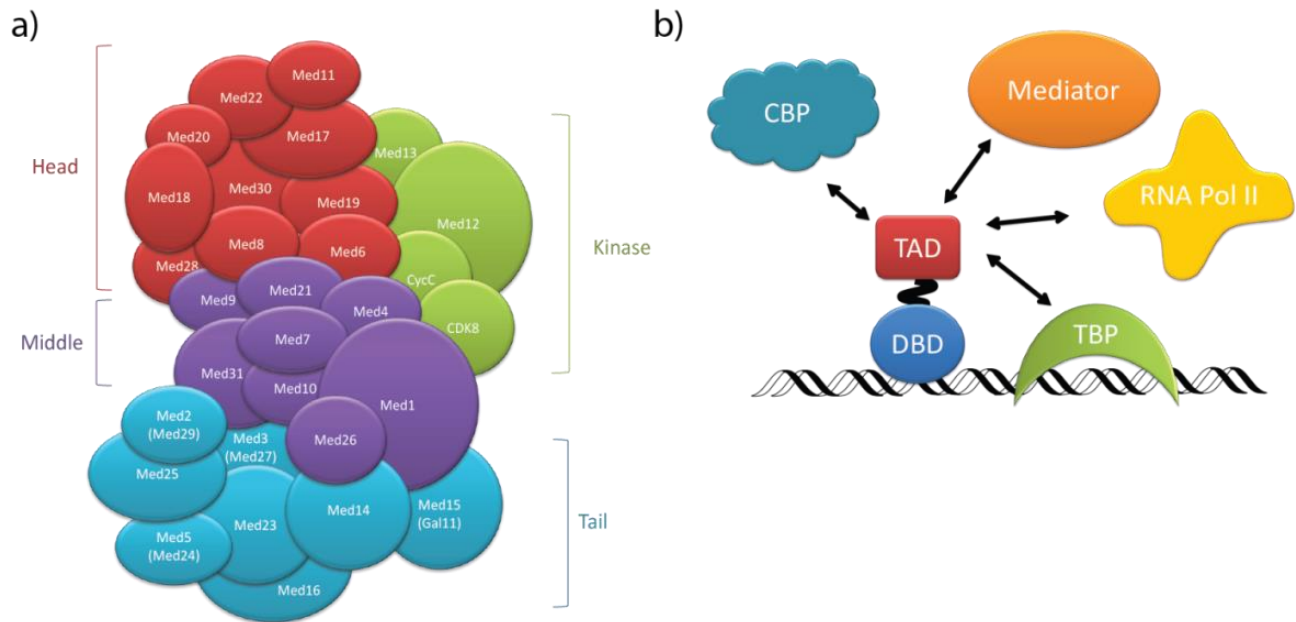


Figure 1.5 Cartoon depiction of **a)** multiple subunits stably interacting with each other, forming the Mediator complex. Figure modified from Malik et al, 2010.⁵⁴ **b)** the same transcriptional activation domain (TAD) interacting with multiple partners such as histone modification coactivators (CBP), coactivators (Mediator) and general transcriptional factors (TBP and RNA Pol II).

A closer look at the nature of activator-coactivator interactions explain to some extent why these are difficult to fully understand: most coactivators and activators would qualify as hub proteins, where one protein will have multiple binding partners involved in several different pathways (Figure 1.6), not surprisingly both transcriptional activators and coactivators tend to be conformationally flexible or disordered until they are bound to their partners.^{25,73}

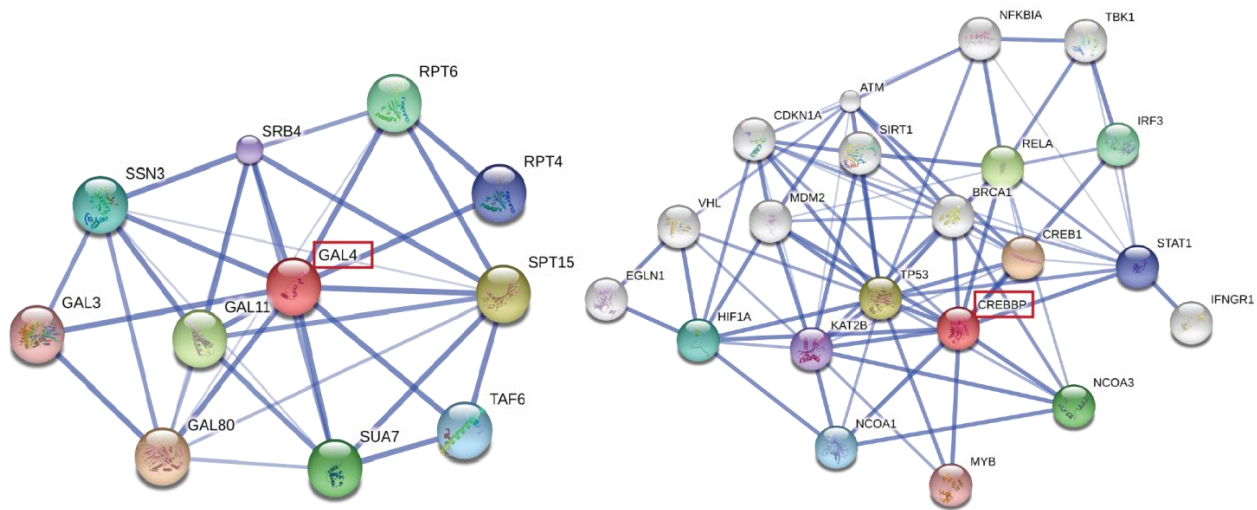


Figure 1.6 Interaction networks (incomplete for clarity) of transcriptional activator Gal4 and coactivator CREB Binding Protein (CREBBP) show that they are hub proteins and demonstrate the complexity of transcriptional activator-coactivator interactions. Networks generated by STRING, a database of PPIs based on experimental data and computational predictions.⁷⁴ Associations with higher confidence are represented by thicker lines.

Extensive effort has been made to reveal this network of interactions. Indeed, over the past few decades, enormous progress has been made in identifying the coactivator binding partners of various transcriptional activators (Figure 1.7) through a variety of *in vivo* techniques such as affinity capture,⁷⁵ phenotypic suppression or enhancement, protein-fragment complementation assays and co-localization experiments.⁷⁶

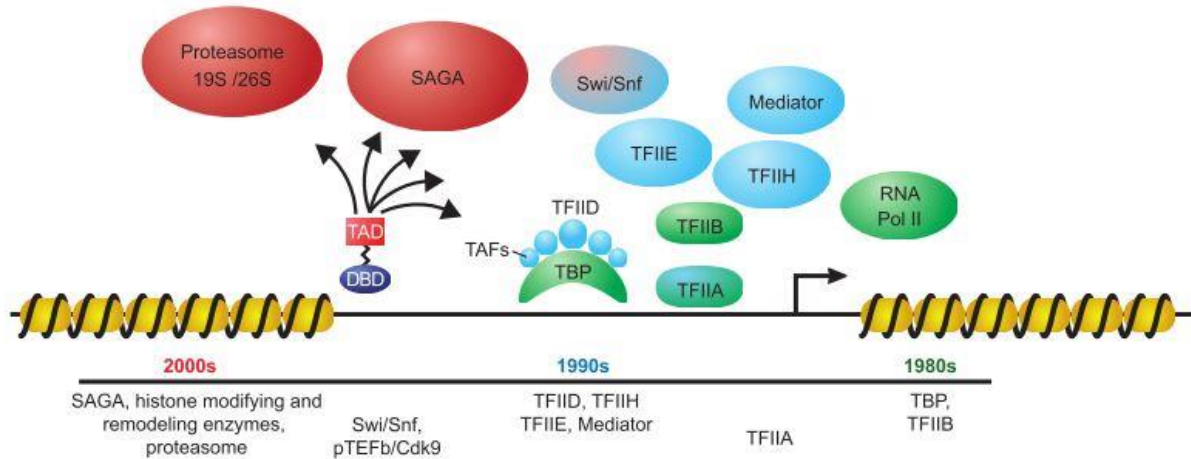


Figure 1.7 Timeline of the identification of transcriptional activator targets since the 1980s. The physiological relevance of many activator targets remains under debate. Figure from Mapp et al, 2007.⁹

While these studies seek to answer the question of “who” the proteins interact with, another equally important question is “how” they interact with each other. Clearly the interaction between activators and coactivators is not as simple as the North and South poles of two magnets. The multispecificity and multifunctionality of these proteins hint at a much more nuanced and complicated mechanism. For one thing, activators interact with coactivators and suppressors with a wide range of binding affinity at the same interaction site. For example, Gal4 TAD binds to its suppressor Gal80 with low nanomolar affinity⁷⁷ whereas the same Gal4 TAD binds to Med15 (Gal11) with a K_D of 0.1 μM .⁷⁸ The KIX domain of the mammalian coactivator CREB Binding Protein (CBP) binds to a FOXO3a TAD with an approximately 100 μM dissociation rate constant⁷⁹ while the TAD of the phosphorylated kinase inducible domain (pKID) of CREB binds to the same KIX domain with a higher affinity (K_D of less than 1 μM).^{80–82} Furthermore, allostery plays an important part in coactivators that are multi-interface hub proteins, where one protein binding to the hub can elicit allosteric effects on a second distal binding site, thus enhancing or abrogating the binding of a different protein at the distal site.⁸³ Evidence suggests that transcriptional activator-coactivator interactions are governed by complicated mechanisms

that likely involve conformational dynamics and interconversion between structured and disordered states.^{61,73,84} Deciphering these mechanisms will contribute to intelligently designing therapeutic strategies to correct disease-related aberrant transcriptional activator-coactivator interactions. The following section will discuss the progress made in this field and the questions that have yet to be addressed.

D. Mechanisms of transcriptional activator-coactivator interactions

D.1. *In vitro* techniques employed for studying the mechanism of transcription factor protein-protein interactions

Over the past few decades, many new techniques have emerged and have been employed in studying transcription factor protein-protein interactions. The following table is a brief overview of commonly used techniques.

Table 1.1 Examples of techniques used in probing activator-coactivator interactions.

Technique	Description	Examples	Notes
Affinity Pull-down Assays	Determines the equilibrium binding affinity of two proteins by quantification of one protein that is bound to another affinity tagged-protein	Gal4-Gal80 ⁷⁷ Gal4-Med15 ⁸⁵ VP16-PC4 ⁸⁶	Used mostly in early studies of transcriptional activator-coactivator interactions. Quantification by band intensity can result in poor accuracy.
Isothermal Titration Calorimetry (ITC)	Determines the equilibrium binding affinity, enthalpy and stoichiometry by monitoring heat change in solution as interaction takes place	p53-KIX(CBP) ³⁰ CSL-MINT ⁸⁷	Can give useful thermodynamic information such as enthalpy. Does not require modification of the protein. Requires large amounts of material.
Fluorescence Anisotropy/Polarization	Determines equilibrium binding affinity by monitoring the increase in polarity (decrease in tumbling rate i.e. increase in complex size)	VP16-TFIIB ⁸⁸ pKID-KIX(CBP) ⁸⁹	Can be performed with small quantities in high throughput fashion. Only suitable when one component is much smaller than its binding partner, also requires a fluorescent probe

			modification to the protein.
Fluorescence Intensity	Determines equilibrium binding affinity by monitoring change in fluorescence intensity of fluorescent probe or intrinsic Tyr/Trp fluorescence	c-Myb-KIX(CBP) ⁶⁰	Use of intrinsic fluorescent can avoid modifying protein. Relies on environmental changes to the fluorophore, which under some circumstances is not large enough to detect a change in fluorescence intensity.
Förster Resonance Energy Transfer (FRET)	Determines equilibrium binding affinity and conformational changes by monitoring acceptor fluorophore fluorescence intensity as excited donor fluorophore comes into close proximity	c-Fos-c-Jun ⁹⁰	High sensitivity to distance, can be used as a “molecular ruler”. Is amenable to single molecule studies, and needs very small amount of material. Strict requirement of proximity of the donor and acceptor fluorophores. Known structural information is almost always required. Protein modification required.
X-ray Crystallography	Structure of protein complex can be compared to structure of free proteins to determine conformational change upon binding	STAT6-NcoA-1 ⁹¹ MEF2-TAZ2(p300) ⁹² TEAD-YAP ⁹³	Can glean detailed structural information including side chain orientation. Proteins need to crystallize, need high concentration of material, the nature of crystal packing might result in artificial and forced interaction interfaces.
Nuclear Magnetic Resonance (NMR)	<u>HSQC and 3D-NMR:</u> Structure of protein complex can be compared to structure of free proteins to determine conformational change upon binding	VP16-Med25 ⁹⁴ Gcn4-Med15 ⁹⁵ MLL/c-Myb/ pKID-KIX(CBP) ^{96,97}	Amenable to proteins in solution, can capture more accurate average pose of PPI in more physiologically relevant conditions. Isotopic labeling of proteins is required, not amenable to large protein complexes (>50 kDa). A high concentration of protein is often needed.
	<u>Residual Dipolar Coupling (RDC):</u> Besides static structural information can also detect domain and local or segmental movements	Oct1-Sox2 ⁹⁸	
	<u>Relaxation Dispersion:</u> By setting molecule in a non-equilibrium state and observing it return to the equilibrium state, can determine transient-state kinetics of interactions	pKID-KIX(CBP) ⁶¹ HIF-1 α -TAZ1(CBP) ⁹⁹	

Surface Plasmon Resonance (SPR)	Interaction between immobilized protein and flow-through protein causes change in light resonance from surface, can determine transient-state kinetics of interaction	Gal4-SRB10/11 ¹⁰⁰ p65 (NFκB)-TBP/TFIIB/E1A 13S ¹⁰¹ Gal4/VP16-Swi1/TBP ¹⁰²	Does not need large amounts of material, can obtain on-rate and off-rate data in same set of experiments. One component must be immobilized to a surface.
Stopped-Flow Spectroscopy (of fluorescence, anisotropy, CD, etc.)	Real-time monitoring of spectroscopy change as two species are rapidly mixed enables determination of transient-state kinetics.	p160-CBP ¹⁰³ NFκB-IκBα ¹⁰⁴ c-Myb-KIX(CBP) ⁶⁰	Can observe interaction in solution, closer to physiological condition. Short dead-time, high time resolution. Often need larger volumes of protein.

While various techniques are being developed and optimized over the years, most techniques have certain shortcomings, and often times a combination of techniques are required to obtain a comprehensive picture of the mechanism of transcriptional activator-coactivator interactions. In addition, computational strategies such as molecular dynamics (MD) are often employed to simulate mechanisms that are difficult to monitor by existing techniques. There are also many techniques utilized to study protein-nucleic acid interactions and protein folding that are extremely useful and have not been employed to study transcriptional activator-coactivator interactions. These techniques include: fluorescent quench assays, single molecule FRET assays and ion-mobility mass spectrometry. New techniques and new combination of techniques are crucial in progressing further in our understanding of the transcriptional activator-coactivator interactions.

D.2. Summary of activator-coactivator interaction mechanisms

Recent studies on transcriptional activator-coactivator interactions have found several common aspects in the interaction mechanism.

One of the key features of transcriptional activator-coactivator interactions is that one or both components of the complex undergo significant conformational change. In many cases, the transcriptional activator TAD is intrinsically disordered and assumes a helical structure upon binding to a coactivator or suppressor, as is the case in pKID binding to KIX of CBP,⁶¹ Hif-1 α binding to Taz1 of CBP,¹⁰⁵ Gcn4 binding to Med15,⁹⁵ Gal4 binding to Gal80,¹⁰⁶ and VP16 binding to hTAFII31.¹⁰⁷ Some studies have found that this interaction follows an induced fit mechanism, as suggested in an SPR study on TADs Gal4 and VP16 binding to coactivators Swi1 and TBP¹⁰² and a fluorescence stopped-flow study on the activation domain from the p160 transcriptional coactivator for thyroid hormone and retinoid receptors (ACTR) binding to the nuclear coactivator binding domain (NCBD) of CBP.¹⁰³ A more comprehensive view suggested by many other studies is that this interaction is a combination of both induced fit and conformational selection. This is often referred to as a coupled binding and folding mechanism where both the activator and the coactivator undergo conformational change upon binding (Figure 1.8).¹⁰⁸⁻¹¹⁰ Examples as followed include NMR relaxation dispersion studies on the phosphorylated kinase inducible domain (pKID) of CREB binding to KIX domain of CBP⁶¹ and the ANC1 homology domain (AHD) of AF9 binding to elongation factor AF4.¹¹¹ In most of the recent studies, the interaction is observed to undergo at least two steps: a fast association step and a slower transition step.^{61,102,112,113} Chapter 2 and Chapter 4 of this dissertation focus on this aspect. All of these previous studies have used isolated peptides as TADs. As the native state of transcriptional activators is a DBD-TAD modular construct that is also bound to DNA, using this simplified TAD peptide in interaction studies might not correctly reflect the kinetics of activator-coactivator interactions. Chapter 2 addresses this issue by studying the interactions between a DNA-bound DBD-TAD fusion transcriptional activator protein and its binding coactivator.⁷⁸

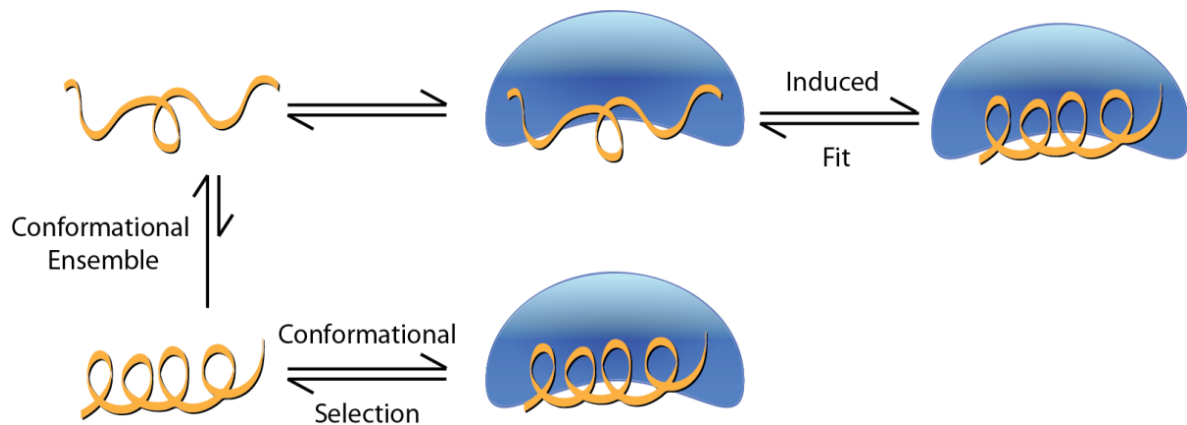


Figure 1.8 Schematic of coupled binding and folding, where the mechanism is a combination of induced folding and conformational selection. Figure modified from Wright et al, 2009.¹¹⁴

Another interesting feature found in several transcriptional activator-coactivator interactions is that there can be more than one conformation or binding mode, which possibly contributes to the multispecificity of these species. For example, NCBP of CBP assumes different conformations when it binds to either ACTR or IRF3 (Figure 1.9a),¹¹⁵ the two TADs of FOXO3a can bind to either one of the two separate binding sites on the KIX domain of CBP with comparable affinity,⁷⁹ and Gcn4 has been shown to bind to Med15 in multiple poses, forming a “fuzzy complex” with Med15 (Figure 1.9b).⁹⁵

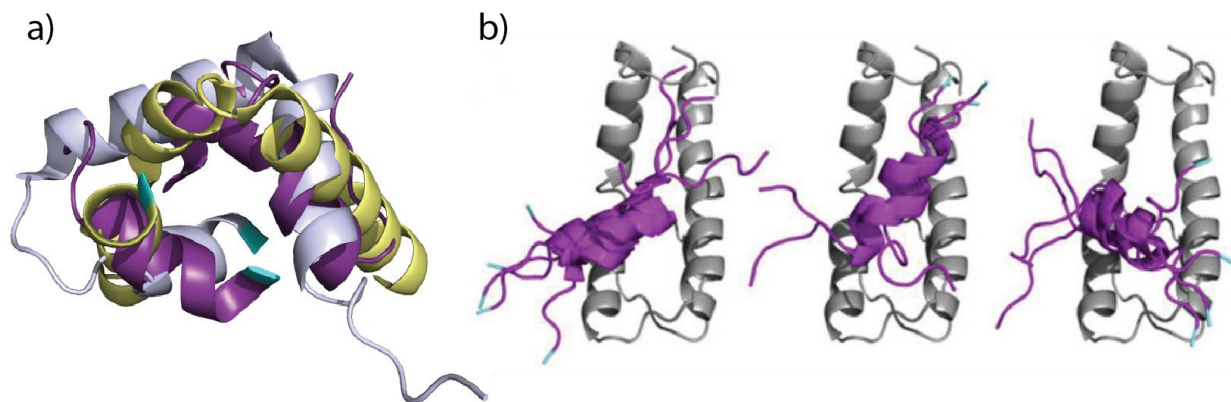


Figure 1.9 a) Overlay of the NCBP of CBP in the free (purple, 1JJS), IRF3-bound (white blue, 1ZOQ) and ACTR-bound (light yellow, 1KBH) conformations show significant variations. **b)** Three orientations of Gcn4-cAD (purple) binding to ABD1 of Med15 (gray) from ensemble of NMR solution structures (1LPB). Adapted from Brzovic et al, 2011.⁹⁵ N-terminals are colored in cyan.

There is also evidence that the entire domains do not contribute equally in binding interactions. Key regions can be more structured than their neighbors and nucleate the coupled binding and folding upon interaction with its binding partners.¹¹⁶ Also the side chains of specific residues can be extremely flexible and undergo a large axis of rotation upon the binding interaction, forming key electrostatic or hydrophobic interactions that nucleate the binding.^{115,117–120} The work in Chapter 3 will exemplify such points.

E. Small molecules as modulators of the transcriptional PPI network

E.1. Targeting transcriptional activator-coactivator interactions with small molecules

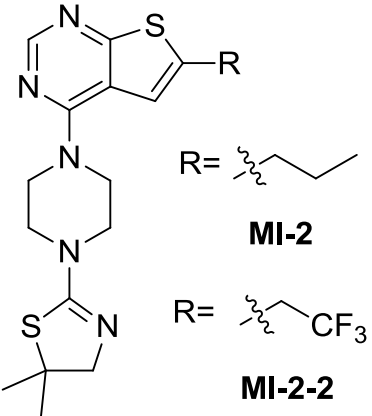
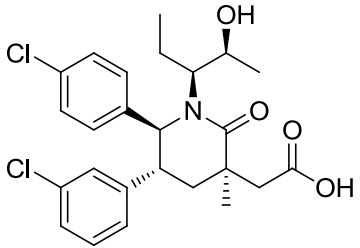
One of the ultimate goals of understanding the mechanism of transcriptional activator-coactivator interactions is to enable guided design of small molecule regulators of transcription. As protein-protein interactions usually consist of shallow, large interfaces, targeting this interface can be challenging.⁴⁶ Understanding the underlying mechanism of these interactions can enable more accurate identification of small molecule modulators. Small molecules can be applied therapeutically for treating diseases that involve malfunction of transcriptional activator-coactivator interactions. These therapeutic modulators can act by various modes of function: they can inhibit an over-regulated interaction, mimic an abrogated interaction or stabilize a weakened interaction.^{121–124} Another extremely useful function of small molecule modulators is their use as a mechanistic probe. There are still many unknown aspects of transcription. The ability to selectively inhibit a certain interaction and observe the resulting consequences will enable researchers to elucidate the related transcriptional pathways and mechanisms. This is especially advantageous when a single transcription factor interacts with multiple binding partners. Instead of genetically knocking down or knocking out the expression of an entire protein, it may be

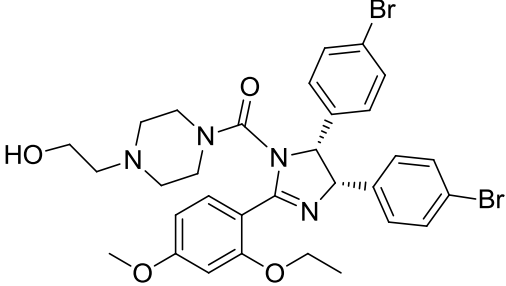
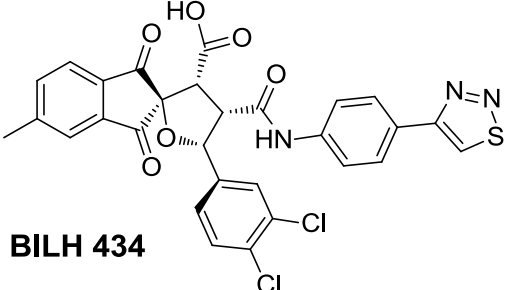
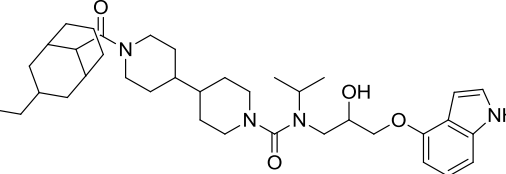
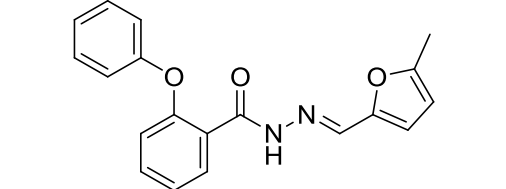
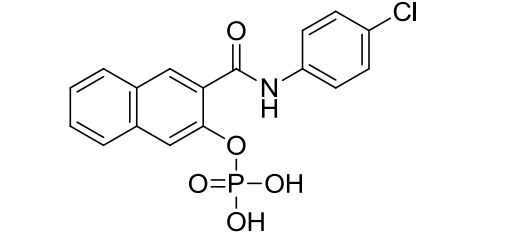
possible for specific small molecule modulators to finely control only one interaction, hence pinpointing the target pathway that is studied. Outlined in Chapter 3 and Chapter 4 are studies on transcription factors in the context of small molecule modulators.

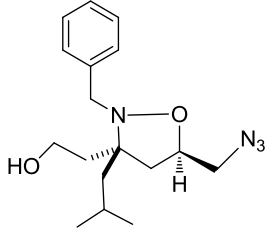
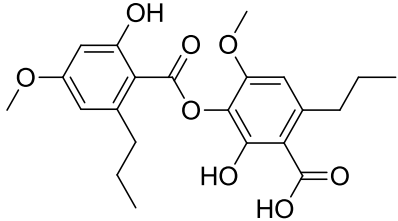
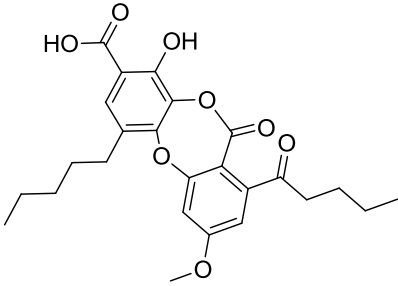
E.2. Examples of small molecule modulators of activator-cofactor interactions

Many small molecule inhibitors of activator-coactivator have emerged in the past decade or so due to high demand of modulators in this crucial pathway. Especially remarkable are the plethora of small molecule inhibitors identified for the p53-MDM2 interaction, which has served as a model system to study with both experimentally and as a data training set to use in computational studies and predictions for small molecule inhibitors of PPIs.^{121,125} Below is a table outlining a few examples of these inhibitors.

Table 1.2 Examples of small molecule modulators of transcriptional activator-coactivator interactions

Structure	Target Interaction	K_D / IC_{50} Values	Citation
 <p>MI-2 MI-2-2</p>	MLL-Menin	<p>MI-2 $K_D = 0.446$ μM</p> <p>MI-2-2 $K_D = 0.046$ μM</p>	Shi et al (2012) ¹²⁶
 <p>AM-8553</p>	p53-Mdm2	$K_D = 0.4$ nM	Bernard et al (2012) ¹²⁷

 <p>Nutlin_16</p>	p53-Mdm2	IC ₅₀ = 0.14 μM	Vassilev et al, (2004) ¹²⁸
 <p>BILH 434</p>	HPV11 TAD E2-HPV18 E1	K _D = 0.040 μM IC ₅₀ = 0.10 μM	Wang et al, (2004) ¹²⁹
 <p>Adamanolol</p>	ESX-Sur2	IC ₅₀ = 8 μM	Asada et al, (2003) ¹³⁰
 <p>PNU-74654</p>	β-catenin-Tcf4	K _D = 0.45 μM	Trosset et al, (2006) ¹³¹
 <p>Naphthol AS-E phosphate (KG-501)</p>	pKID-KIX(CBP)	K _i ≈ 90 μM K _D = 115 μM	Best et al, (2004) ¹³² Pomerantz et al, (2012) ⁸¹

 <p style="text-align: center;">iTAD1</p>	MLL-KIX(CBP)	$K_D = 38 \mu\text{M}$	Buhrlage, S. J. et al (2009) ¹³³
 <p style="text-align: center;">Sekikaic Acid</p>	pKID-KIX(CBP) and MLL-KIX(CBP)	pKID site $IC_{50} = 17 \mu\text{M}$ MLL site $IC_{50} = 34 \mu\text{M}$	Majmudar, C. Y., Højfeldt, J. W. et al (2012) ¹³⁴
 <p style="text-align: center;">Lobaric Acid</p>	pKID-KIX(CBP) and MLL-KIX(CBP)	pKID site $IC_{50} = 25 \mu\text{M}$ MLL site $IC_{50} = 17 \mu\text{M}$	Majmudar, C. Y., Højfeldt, J. W. et al (2012) ¹³⁴

While some small molecules are able to inhibit certain protein-protein interaction with high affinity, the majority of protein-protein interactions between transcriptional activators and coactivators still lack tight, specific small molecule modulators.⁴⁶ Thus the work in this dissertation is focused on elucidating the mechanisms of these interactions to better design small molecule modulators.

F. Thesis summary

In this dissertation, I will describe the studies of the conformational and kinetic mechanism of transcriptional activator-coactivator interactions by focusing on two prototypical *in vitro* systems: the coactivator Med15 and activators Gal4, VP16 and Gcn4 (Chapter 2 and Appendix);

and the more conformationally defined KIX domain of coactivator CBP with activators MLL and pKID (Chapter 3 and Chapter 4). Additionally, the effects of a covalently tethered small molecule inhibitor of one binding site of KIX will be examined in the context of the stability of KIX and allosteric effects of its binding properties to activators at the opposite binding site. In Chapter 2, the transient-state kinetics mechanism of the interactions between Med15 and Gal4, VP16 and Gnc4 will be dissected by fluorescence stopped-flow. In Chapter 3, the conformational effect of a covalently tethered small molecule inhibitor on the KIX domain in CBP/p300 will be examined by biochemical and molecular dynamics methods and structural techniques such as solution protein NMR and X-ray crystallography. In Chapter 4, the kinetics and allosteric effect of a covalently tethered small molecule inhibitor on the KIX domain in CBP/p300 will be examined using fluorescence stopped-flow and other biochemical methods. In summary, this dissertation seeks to shed new light on the mechanism of transcriptional activator-coactivator interactions by both conformational and kinetic studies. The results can provide insight into guided approaches in designing artificial transcription factors as potential probes and therapeutics.

G. References

1. Paul, S. A. M., Simons, J. W. & Mabeesh, N. J. HIF at the crossroads between ischemia and carcinogenesis. *J. Cell. Physiol.* **200**, 20–30 (2004).
2. Karin, M. & Lin, A. NF-kappaB at the crossroads of life and death. *Nat. Immunol.* **3**, 221–7 (2002).
3. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553 (2000).
4. Ernst, P. *et al.* Definitive Hematopoiesis Requires the Mixed-Lineage Leukemia Gene. *Dev. Cell* **6**, 437–443 (2004).
5. Joerger, A. C. & Fersht, A. R. Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582 (2008).
6. Pandolfi, P. P. Transcription therapy for cancer. *Oncogene* **20**, 3116–3127 (2001).
7. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
8. Darnell Jr., J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749 (2002).

9. Mapp, A. K. K. & Ansari, A. Z. Z. A TAD further: exogenous control of gene activation. *ACS Chem. Biol.* **2**, 62–75 (2007).
10. Koleske, A. J. & Young, R. A. The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* **20**, 113–116 (1995).
11. Ptashne, M. Control of gene transcription: An outline. *Nature Medicine* **3**, 1069–1072 (1997).
12. Conaway, R. C. & Conaway, J. W. General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* **62**, 161–90 (1993).
13. Ptashne, M. How eukaryotic transcriptional activators work. *Nature* **335**, 683–689 (1988).
14. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–77 (1997).
15. Brent, R. & Ptashne, M. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729–736 (1985).
16. Ptashne, M. & Gann, A. *Genes and Signals*. (Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 2008).
17. Triezenberg, S. J. Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5**, 190–196 (1995).
18. Kumar, R. & Thompson, E. B. The structure of the nuclear hormone receptors. *Steroids* **64**, 310–319 (1999).
19. Garvie, C. W. & Wolberger, C. Recognition of specific DNA sequences. *Mol. Cell* **8**, 937–946 (2001).
20. Fogh, R. H. *et al.* Solution structure of the LexA repressor DNA binding domain determined by 1H NMR spectroscopy. *EMBO J.* **13**, 3936–44 (1994).
21. Schuetz, A. *et al.* The structure of the Klf4 DNA-binding domain links to self-renewal and macrophage differentiation. *Cell. Mol. Life Sci.* **68**, 3121–31 (2011).
22. Hong, M. *et al.* Structural basis for dimerization in DNA recognition by Gal4. *Structure* **16**, 1019–1026 (2008).
23. Hinnebusch, A. G. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu. Rev. Microbiol.* **59**, 407–450 (2005).
24. Nair, S. K. & Burley, S. K. X-Ray Structures of Myc-Max and Mad-Max Recognizing DNA. *Cell* **112**, 193–205 (2003).
25. Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* **45**, 6873–6888 (2006).
26. Reeves, W. M. & Hahn, S. Targets of the Gal4 transcription activator in functional transcription complexes. *Mol. Cell. Biol.* **25**, 9092–9102 (2005).
27. 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).
28. Jedidi, I. *et al.* Activator Gcn4 employs multiple segments of Med15/Gal11, including the KIX domain, to recruit mediator to target genes in vivo. *J. Biol. Chem.* **285**, 2438–2455 (2010).
29. Rajagopalan, S., Andreeva, A., Rutherford, T. J. & Fersht, A. R. Mapping the physical and functional interactions between the tumor suppressors p53 and BRCA2. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8587–8592 (2010).
30. Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry* **48**, 2115–2124 (2009).
31. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**, 235–239 (2005).
32. Dotson, M. R. *et al.* Structural organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14307–10 (2000).
33. Chadick, J. Z. & Asturias, F. J. Structure of eukaryotic Mediator complexes. *Trends Biochem. Sci.* **30**, 264–271 (2005).
34. Davis, J. A., Takagi, Y., Kornberg, R. D. & Asturias, F. A. Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol. Cell* **10**, 409–415 (2002).

35. Narlikar, G. J., Fan, H. Y. & Kingston, R. E. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* **108**, 475–487 (2002).
36. Burns, L. G. & Peterson, C. L. The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. *Mol. Cell. Biol.* **17**, 4811–4819 (1997).
37. Venkatesan, K. *et al.* An empirical framework for binary interactome mapping. *Nat. Methods* **6**, 83–90 (2009).
38. Parker, D. *et al.* Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol.* **19**, 5601–7 (1999).
39. Korkmaz, E. N., Nussinov, R. & Haliloğlu, T. Conformational Control of the Binding of the Transactivation Domain of the MLL Protein and c-Myb to the KIX Domain of CREB. *PLoS Comput. Biol.* **8**, e1002420 (2012).
40. Nooren, I. M. A. & Thornton, J. M. Diversity of protein-protein interactions. *EMBO J.* **22**, 3486–92 (2003).
41. Eccleston, J. F., Binns, D. D., Davis, C. T., Albanesi, J. P. & Jameson, D. M. Oligomerization and kinetic mechanism of the dynamin GTPase. *Eur. Biophys. J.* **31**, 275–82 (2002).
42. Brasier, A. R. The NF- κ B Regulatory Network. *Cardiovasc. Toxicol.* **6**, 111–130 (2006).
43. Gibbard, R. J., Morley, P. J. & Gay, N. J. Conserved features in the extracellular domain of human toll-like receptor 8 are essential for pH-dependent signaling. *J. Biol. Chem.* **281**, 27503–11 (2006).
44. Groemping, Y. & Reinstein, J. Folding properties of the nucleotide exchange factor GrpE from *Thermus thermophilus*: GrpE is a thermosensor that mediates heat shock response. *J. Mol. Biol.* **314**, 167–78 (2001).
45. Nooren, I. M. . & Thornton, J. M. Structural Characterisation and Functional Significance of Transient Protein–Protein Interactions. *J. Mol. Biol.* **325**, 991–1018 (2003).
46. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chem. Biol.* **7**, 1311–20 (2012).
47. Moreira, I., Fernandes, P. & Ramos, M. Hot spots—A review of the protein–protein interface determinant amino acid residues. *Proteins* **68**, 803–812 (2007).
48. DeLano, W. L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* **12**, 14–20 (2002).
49. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. The structure of the *Escherichia coli* EF-Tu.EF-Ts complex at 2.5 Å resolution. *Nature* **379**, 511–8 (1996).
50. Brownlow, S. *et al.* Bovine β -lactoglobulin at 1.8 Å resolution — still an enigmatic lipocalin. *Structure* **5**, 481–495 (1997).
51. Singh, G. P., Ganapathi, M. & Dash, D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* **66**, 761–5 (2007).
52. Keskin, O. & Nussinov, R. Similar binding sites and different partners: implications to shared proteins in cellular pathways. *Structure* **15**, 341–54 (2007).
53. Han, J.-D. J. *et al.* Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88–93 (2004).
54. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* **11**, 761–72 (2010).
55. Huang, Y. & Liu, Z. Smoothing molecular interactions: The “kinetic buffer” effect of intrinsically disordered proteins. *Proteins* **78**, 3251–9 (2010).
56. Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).
57. Schreiber, G. & Keating, A. E. Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50–61 (2011).
58. Higurashi, M., Ishida, T. & Kinoshita, K. Identification of transient hub proteins and the possible structural basis for their multiple interactions. *Protein Sci.* **17**, 72–8 (2008).
59. Kim, P. M., Sboner, A., Xia, Y. & Gerstein, M. The role of disorder in interaction networks: a structural analysis. *Mol. Syst. Biol.* **4**, 179 (2008).

60. 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).
61. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
62. Arai, M., Dyson, H. J. & Wright, P. E. Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. *FEBS Lett.* 4500–4504 (2010).
63. Tyagi, M., Shoemaker, B. A., Bryant, S. H. & Panchenko, A. R. Exploring functional roles of multibinding protein interfaces. *Protein Sci.* **18**, 1674–83 (2009).
64. Tompa, P., Szász, C. & Buday, L. Structural disorder throws new light on moonlighting. *Trends Biochem. Sci.* **30**, 484–9 (2005).
65. Dasgupta, B., Nakamura, H. & Kinjo, A. R. Distinct roles of overlapping and non-overlapping regions of hub protein interfaces in recognition of multiple partners. *J. Mol. Biol.* **411**, 713–27 (2011).
66. Swapna, L. S., Mahajan, S., De Brevern, A. G. & Srinivasan, N. Comparison of tertiary structures of proteins in protein-protein complexes with unbound forms suggests prevalence of allostery in signalling proteins. *BMC Struct. Biol.* **12**, 6 (2012).
67. 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 (2001).
68. Kerppola, T. K. Transcriptional cooperativity: bending over backwards and doing the flip. *Structure* **6**, 549–554 (1998).
69. Ferreón, J. C. *et al.* Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6591–6596 (2009).
70. Myers, L. C. *et al.* The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**, 45–54 (1998).
71. Cairns, B. R., Kim, Y. J., Sayre, M. H., Laurent, B. C. & Kornberg, R. D. A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1950–4 (1994).
72. Dechassa, M. L. *et al.* Architecture of the SWI/SNF-nucleosome complex. *Mol. Cell. Biol.* **28**, 6010–21 (2008).
73. Motlagh, H. N., Li, J., Thompson, E. B. & Hilser, V. J. Interplay between allostery and intrinsic disorder in an ensemble. *Biochem. Soc. Trans.* **40**, 975–80 (2012).
74. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–15 (2013).
75. Krishnamurthy, M. *et al.* Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo. *ACS Chem. Biol.* **6**, 1321–6 (2011).
76. Bhaumik, S. R., Raha, T., Aiello, D. P. & Green, M. R. In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev.* **18**, 333–343 (2004).
77. Lue, N. F., Chasman, D. I., Buchman, A. R. & Kornberg, R. D. Interaction of GAL4 and GAL80 gene regulatory proteins in vitro. *Mol. Cell. Biol.* **7**, 3446–51 (1987).
78. 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).
79. Wang, F. & Marshall, C. Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6078–83 (2012).
80. Cook, P. R., Polakowski, N. & Lemasson, I. HTLV-1 HBZ protein deregulates interactions between cellular factors and the KIX domain of p300/CBP. *J. Mol. Biol.* **409**, 384–98 (2011).
81. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **7**, 1345–50 (2012).
82. Wang, N. & Majmudar, C. Ordering a dynamic protein via a small molecule stabilizer. *J. Am. Chem. Soc.* (2013).doi:10.1021/ja3122334

83. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *J. Biol. Chem.* **277**, 43168–74 (2002).
84. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
85. Jeong, C. J. *et al.* Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **40**, 9421–9427 (2001).
86. Jonker, H. R. a H. R. A. R. A., Wechselberger, R. W. W. R. W., Boelens, R., Folkers, G. E. G. E. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44**, 827–839 (2005).
87. VanderWielen, B. D., Yuan, Z., Friedmann, D. R. & Kovall, R. A. Transcriptional repression in the Notch pathway: thermodynamic characterization of CSL-MINT (Mx2-interacting nuclear target protein) complexes. *J. Biol. Chem.* **286**, 14892–902 (2011).
88. Shen, F., Triezenberg, S. J., Hensley, P., Porter, D. & Knutson, J. R. Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. *J. Biol. Chem.* **271**, 4827–4837 (1996).
89. Parker, D. *et al.* Analysis of an Activator:Coactivator Complex Reveals an Essential Role for Secondary Structure in Transcriptional Activation. *Mol. Cell* **2**, 353–359 (1998).
90. Vámosi, G. *et al.* Conformation of the c-Fos/c-Jun complex in vivo: a combined FRET, FCCS, and MD-modeling study. *Biophys. J.* **94**, 2859–68 (2008).
91. Razeto, A. *et al.* Structure of the NCoA-1/SRC-1 PAS-B domain bound to the LXXLL motif of the STAT6 transactivation domain. *J. Mol. Biol.* **336**, 319–329 (2004).
92. He, J. *et al.* Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Res.* **39**, 4464–74 (2011).
93. Li, Z. *et al.* Structural insights into the YAP and TEAD complex. *Genes Dev.* **24**, 235–40 (2010).
94. Vojnic, E. *et al.* Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18**, 404–9 (2011).
95. 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).
96. 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**, 1005–13 (2006).
97. 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**, 521–34 (2004).
98. Williams, D. C., Cai, M. & Clore, G. M. Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. *J. Biol. Chem.* **279**, 1449–57 (2004).
99. Sugase, K., Lansing, J. C., Dyson, H. J. & Wright, P. E. Tailoring relaxation dispersion experiments for fast-associating protein complexes. *J. Am. Chem. Soc.* **129**, 13406–7 (2007).
100. Ansari, A. Z. *et al.* Transcriptional activating regions target a cyclin-dependent kinase. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14706–14709 (2002).
101. Paal, K., Baeuerle, P. A. & Schmitz, M. L. Basal transcription factors TBP and TFIIB and the viral coactivator E1A 13S bind with distinct affinities and kinetics to the transactivation domain of NF-kappaB p65. *Nucleic Acids Res.* **25**, 1050–1055 (1997).
102. Ferreira, M. E. *et al.* Mechanism of transcription factor recruitment by acidic activators. *J. Biol. Chem.* **280**, 21779–21784 (2005).
103. 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**, 34316–24 (2012).

104. Sue, S.-C., Alverdi, V., Komives, E. A. & Dyson, H. J. Detection of a ternary complex of NF-kappaB and IkappaBalpha with DNA provides insights into how IkappaBalpha removes NF-kappaB from transcription sites. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1367–72 (2011).
105. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5271–6 (2002).
106. Kumar, P. R., Yu, Y., Sternglanz, R., Johnston, S. A. & Joshua-Tor, L. NADP regulates the yeast GAL induction system. *Science* **319**, 1090 (2008).
107. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277**, 1310–1313 (1997).
108. Zhou, H.-X., Pang, X. & Lu, C. Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *Phys. Chem. Chem. Phys.* **14**, 10466–10476 (2012).
109. Hammes, G. G., Chang, Y.-C. & Oas, T. G. Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13737–41 (2009).
110. Kiefhaber, T., Bachmann, A. & Jensen, K. S. Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr. Opin. Struct. Biol.* **22**, 21–9 (2012).
111. Leach, B. I. *et al.* Leukemia fusion target AF9 is an intrinsically disordered transcriptional regulator that recruits multiple partners via coupled folding and binding. *Structure* **21**, 176–83 (2013).
112. Huang, Y. & Liu, Z. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the “fly-casting” mechanism. *J. Mol. Biol.* **393**, 1143–59 (2009).
113. Dogan, J., Schmidt, T., Mu, X., Engstrom, A. & Jemth, P. Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. *J. Biol. Chem.* M112.399436– (2017).doi:10.1074/jbc.M112.399436
114. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
115. Kjaergaard, M., Teilum, K. & Poulsen, F. M. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12535–12540 (2010).
116. Kim, D. H. *et al.* Multiple hTAF(II)31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16. *BMB Rep.* **42**, 411–417 (2009).
117. Zhang, W., Ganguly, D. & Chen, J. Residual Structures, Conformational Fluctuations, and Electrostatic Interactions in the Synergistic Folding of Two Intrinsically Disordered Proteins. *PLoS Comput. Biol.* **8**, e1002353 (2012).
118. Huang, Y. & Liu, Z. Anchoring Intrinsically Disordered Proteins to Multiple Targets: Lessons from N-Terminus of the p53 Protein. *Int. J. Mol. Sci.* **12**, 1410–30 (2011).
119. Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G. & Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3952–7 (2011).
120. Daily, M. D., Upadhyaya, T. J. & Gray, J. J. Contact rearrangements form coupled networks from local motions in allosteric proteins. *Proteins* **71**, 455–66 (2008).
121. Murray, J. K. & Gellman, S. H. Targeting protein-protein interactions: lessons from p53/MDM2. *Biopolymers* **88**, 657–86 (2007).
122. Berg, T. Inhibition of transcription factors with small organic molecules. *Curr. Opin. Chem. Biol.* **12**, 464–71 (2008).
123. Koehler, A. N. A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.* **14**, 331–40 (2010).
124. Scheuermann, T. H. *et al.* Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat. Chem. Biol.* **9**, 271–276 (2013).

125. Majmudar, C. Y. & Mapp, A. K. Chemical approaches to transcriptional regulation. *Curr. Opin. Chem. Biol.* **9**, 467–474 (2005).
126. Shi, A. *et al.* Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia. *Blood* **120**, 4461–9 (2012).
127. Bernard, D., Zhao, Y. & Wang, S. AM-8553: a novel MDM2 inhibitor with a promising outlook for potential clinical development. *J. Med. Chem.* **55**, 4934–5 (2012).
128. Vassilev, L. T. *et al.* In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–8 (2004).
129. Wang, Y. *et al.* Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor. *J. Biol. Chem.* **279**, 6976–85 (2004).
130. Asada, S., Choi, Y. & Uesugi, M. A gene-expression inhibitor that targets an alpha-helix-mediated protein interaction. *J. Am. Chem. Soc.* **125**, 4992–3 (2003).
131. Trosset, J.-Y. *et al.* Inhibition of protein-protein interactions: the discovery of druglike beta-catenin inhibitors by combining virtual and biophysical screening. *Proteins* **64**, 60–7 (2006).
132. Best, J. L. *et al.* Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17622–7 (2004).
133. Buhrlage, S. J. J. *et al.* Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem. Biol.* **4**, 335–344 (2009).
134. 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).

Chapter 2 Kinetic Characterization of a Transcriptional Coactivator (Med15) in Complex with Transcriptional Activators VP16, Gal4 and Gcn4¹

A. Introduction

Several lines of evidence suggest that the prototypical amphipathic transcriptional activators Gal4, Gcn4 and VP16 interact with the key coactivator Med15 (Gal11) during transcription initiation, despite little sequence homology amongst the activators. Recent crosslinking data further reveal that at least two of the activators utilize the same binding surface within Med15 for transcriptional activation, yet activate transcription to quite different levels.¹ The mechanism of these activator-coactivator interaction are poorly understood. However as they play an important part in the transcription pathway, knowing the mechanism of interaction would provide insight into designing small molecule modulators of activator-coactivator complexes to act as functional probes or even potential therapeutics. Hence this chapter describes a kinetic approach to dissect the activator-coactivator interaction mechanism. To assess if these three activators use a shared binding mechanism for Med15 recruitment, we characterized the thermodynamics and kinetics of Med15•activator•DNA complex formation by fluorescence titration and stopped-flow techniques. Combination of each DNA•activator complex with Med15 produces bi-phasic time-

¹ This was a collaboration with Dr. Amberlyn M. Wands and data obtained by her is explicitly labeled within the chapter. The contents of this chapter are adapted and reproduced from a published article: Wands, A. M., Wang, N., Lum, J. K., Hsieh, J., Fierke, C. A., and Mapp, A. K. (2011), *J. Biol. Chem.* 286, 16238–16245.²

courses.² This is consistent with a minimal, two-step binding mechanism comprised of a bimolecular association step followed by a conformational change step of the Med15•activator•DNA complex. Furthermore, the equilibrium constant for the conformational change (K_2) correlates with the ability of an activator to stimulate transcription. VP16, the most potent of the activators, has the largest value of K_2 while the least potent, Gcn4, has the smallest value. This correlation is consistent with a model in which transcriptional activation is regulated at least in part by the rearrangement of the Med15•activator•DNA ternary complex. These results are the first detailed kinetic characterization of the transcriptional activation machinery and provide a framework for the future design of potent transcriptional modulators

B. Background

The unique transcriptional signatures associated with human disease have spurred enormous efforts towards the discovery of artificial transcriptional regulators.³⁻⁶ However, these efforts are hampered by an incomplete understanding of the mechanism utilized by transcriptional activators to interact with and recruit the transcriptional machinery to a gene promoter (Figure 2.1a).⁷ What is known is that most activators have modular architecture, consisting minimally of a DNA binding domain (DBD) and a transcriptional activation domain (TAD),^{8,9} and that they activate transcription by physically recruiting transcriptional coactivators and the preinitiation complex to the promoter site of a gene.^{10,11} One domain of the activator, the transcriptional activation domain or TAD, carries out this process, interacting with a variety of coactivators that remodel chromatin and mediate assembly of the pre-initiation complex; these interactions regulate both the timing and extent of transcriptional activation.^{7,8}

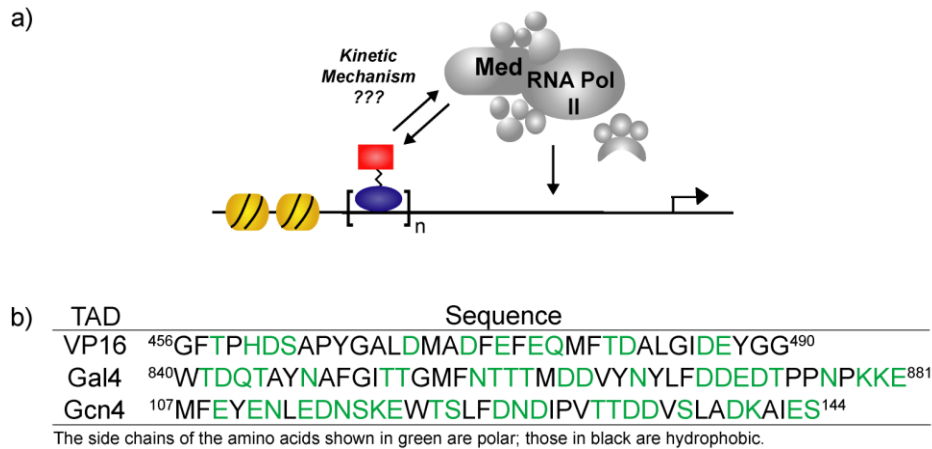


Figure 2.1 a) Schematic of a transcriptional activator localized upstream of a gene promoter through a DNA-binding domain (DBD) (blue circle) which is poised for recruitment of the transcriptional machinery through interaction with the transcriptional activation domain (TAD) (red square). b) Primary amino acid sequences of the amphipathic TADs used in this study.

B.1. Transcriptional activation domains (TADs) VP16, Gal4 and Gcn4

Perhaps the three best-characterized transcriptional activation domains are those of VP16, Gal4 and Gcn4. They are mainly characterized in terms of their sequences, the proteins they target and basic determination of their affinity with their targets. These activators are members of the largest class of activators, the amphipathic class, characterized by interspersed polar and hydrophobic amino acid residues in their TADs (Figure 2.1b).



Figure 2.2 a) Structure of VP16C (red) binding to the Tfb1 subunit of TFIIF (gray) (PDB: 2K2U). b) Structure of Gal4 DBD (1-100) bound to DNA, the two monomers that form a homodimer are depicted in light and dark green, and the Zn molecules in the Zn₂Cys₆ clusters depicted as purple spheres (PDB: 3COQ). c) Structure of Gcn4 (blue) in complex with Gal11/Med15 (gray) (PDB: 2LPB).

VP16: The herpes simplex virus (HSV) virion protein 16 (VP16) is a transcriptional activator that regulates transcription of viral genes within an infected mammalian host cell, thus doing so through the formation of protein-protein interactions with endogenous host transcription factors.¹² The C-terminal residues (residues 412-490) of the 490-residue VP16 protein have been identified as the transcriptional activation domain,¹³ which binds to many proteins involved in the transcriptional machinery, including TBP, TFIIA, TFIIB, TFIIH, hTAF_{II}31, dTAF_{II}40, the human cofactor PC4, CBP, p300¹⁴ and Med15 of the Mediator complex.^{1,15} The VP16 TAD is further dissected into two separate activation subdomains, the amino subdomain (VP16N: residues 412-456) and the carboxyl subdomain (VP16C: residues 456-490, *used in this study*).¹⁶ Similar to many amphipathic transcriptional activators, the VP16 TAD is intrinsically disordered and has little to no secondary structure in its free state, while upon binding to its various partners the TAD assumes an α -helical structure.^{14,17-19} Despite being extensively studied and utilized as a model transcriptional activator, only recently has any high-resolution structural information been obtained on this domain, in which an NMR structure of VP16C in complex with the Tfb1 subunit of TFIIH revealed its structural similarity to the p53 TAD when bound to Tfb1 (Figure 2.2a).²⁰

Gal4: Gal4 is a yeast transcriptional activator from *Saccharomyces cerevisiae* that is involved in the regulation of galactose metabolism. It has also been shown to regulate transcription in other organisms such as *Drosophila*²¹ and mammalian cells.²² From as early as the 1980s, Gal4 has been studied extensively as a model transcriptional activator.²³ As of present day, both the DNA-binding properties of the Gal4 DBD²⁴ and the coactivators/transcription factors recruited by Gal4 TADs are relatively well-defined.²⁵⁻²⁷ Notably, the structure of the Gal4 DBD (residues 1-100, *used in this study*) in complex with DNA is available, depicting a homodimer of

intertwined helical bundles binding to DNA via two Zn_2Cys_6 clusters.²⁴ (Figure 2.2b) Furthermore, truncation studies on the ability of Gal4 to activate transcription have also identified a TAD region at the C-terminus of Gal4 that retains most of its activity,^{28,29} and this region (residues 840-881, *used in this study*) has been used as a minimal TAD for further studies.^{30,31} Studies suggest the Gal4 TAD is an unstructured domain that assumes secondary structure upon binding to coactivators or suppressors.³² To date, the only structural information of the Gal4 TAD is a 20-residue peptide (854-874) bound to its suppressor protein Gal80 in a loop form.³³ This further demonstrates the difficulty of obtaining detailed structures of a disordered TAD.

Gcn4: Gcn4 is a yeast transcriptional activator from *Saccharomyces cerevisiae* that regulates amino acid synthesis and consumption under amino acid starvation circumstances.³⁴ It is a 281-residue transcriptional activator including a C-terminal DBD and two tandem N-terminal TADs.^{35,36} The DNA binding domain binds to DNA via a leucine zipper motif, the structure of which has been studied extensively.³⁷⁻⁴⁰ The two N-terminal TADs are structurally independent, and are deemed the N-terminal TAD (residues 1-100) and the central TAD (residues 101-134, *used in this study*).^{41,42} Site-directed crosslinking studies have shown Gcn4 to interact with the Tra1, Med15/Gal11, and Taf12 subunits of the transcriptional gene regulation complexes SAGA, NuA4, Mediator, and TFIID.^{43,44} Additionally, extensive work has been done focusing on the binding properties of both Gcn4 TADs to the Med15/Gal11 subunit of Mediator.^{1,42,45,46} An NMR study has revealed that the Gcn4 central TAD (101-134) is intrinsically disordered when isolated and assumes an α -helical structure upon binding to coactivators such as Med15 (Figure 2.2c).⁴⁵

There is considerable evidence that, despite little sequence and structural homology, these TADs share key mechanistic features, including a functionally important binding interaction with the coactivator Med15(Gal11).^{27,43,44,47,48} Indeed, Gal4 and Gcn4 were recently shown to interact with the same binding site in the amino-terminus of Med15.¹ Despite these similarities, the activators stimulate transcription to differing levels, with VP16-derived activators being the most potent and Gcn4 the least (see Results section Figure 2.7). The intrinsically disordered nature of these TADs hinder structural studies. Thus, we elected to carry out kinetics studies on the mechanism of these activator-coactivator interactions.

B.2. Transcriptional coactivator Med15 (Gal11)

The transcriptional coactivator Med15 (Gal11) is a subunit of the Mediator protein complex, which functions as a general transcriptional regulator of RNA Pol II transcription to stimulate basal transcription up to ten-fold.^{49,50} Originally identified as playing an essential role in yeast transcription,⁵¹ the Mediator complex has homologs in higher organisms such as murine and human.⁵² It is a large complex (1.2 Mda) consisting of ~ 20 subunits, and is further dissected into three subdomains, termed the “head”, “middle” and “tail” modules (Figure 2.3a).⁵² Biochemical and structural studies have shown that the Mediator complex is in direct contact with RNA polymerase II at all three of its modules.^{50,51} Med15 resides in the “tail module” of the Mediator complex, this module is linked to the regulation of SAGA-dependent and TATA-containing genes in yeast.⁵³ To this end, *in vivo* and *in vitro* experiments have identified several transcriptional activators that bind to Med15, including VP16, Gal4, Gcn4 and Oaf1.^{47,54,55} Med15 is a 1081-residue protein, yet several domains have been dissected as activator binding domains (Figure 2.3b).^{1,45} In particular, the N-terminal residues (~300) have been identified to

be responsible for the majority of binding interactions with the TADs of VP16, Gal4 and Gcn4. Thus, Med15 (1-345) has been used in this study.¹ Limited structural information is known of Med15. This includes an NMR structure of the N-terminal KIX domain (residues 6-90), of which there are homologous domains in coactivators Arc105 and CREB Binding Protein (CBP),^{56,57} and an NMR structure of the ABD1 domain (residues 158-238) in complex with Gcn4 TAD (Figure 2.3b).⁴⁵

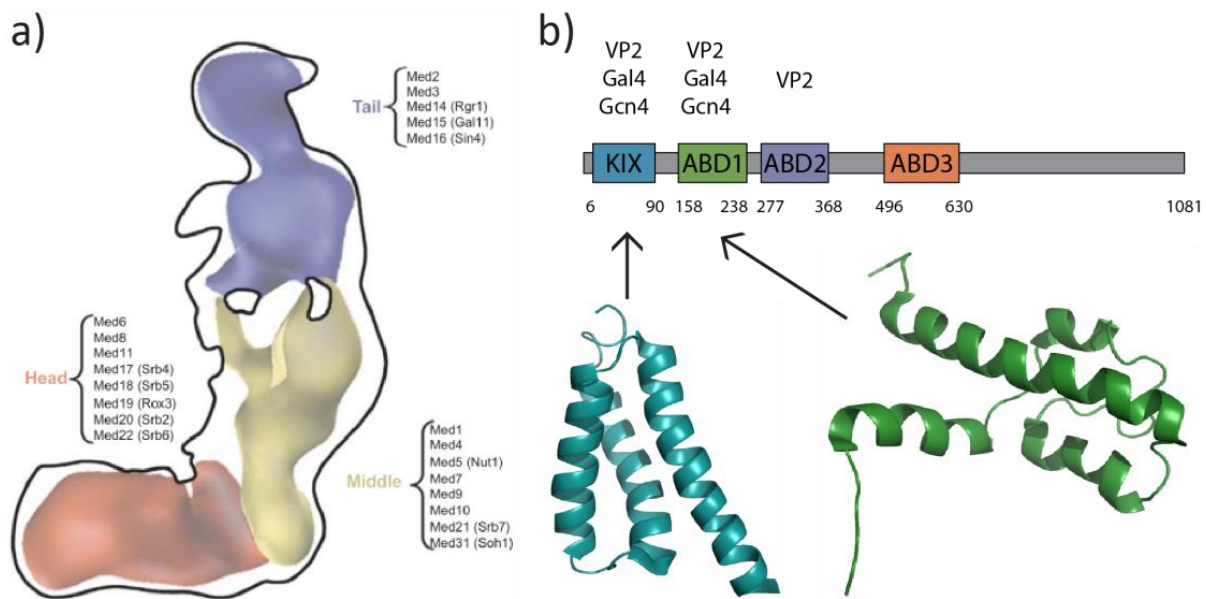


Figure 2.3 a) Schematic of Mediator complex based on the electron microscopy structure and the projected subunits in each module. Figure from Chadick et al, 2005⁵⁸. **b)** Schematic of Med15 and various activator binding domains, the TADs binding to each domain are depicted on top of the graph. Structures of the KIX domain (PDB ID: 2K0N) and the Activator Binding Domain 1 (ABD1, PDB ID: 2LPB) shown.

B.3. Controversy in kinetic models of activator-coactivator interactions

Historically, comparisons of activator-coactivator interactions have centered on equilibrium binding measurements (apparent affinities) but these values are not a uniformly good predictor of function.⁵⁹⁻⁶² Transcription is initiated through a series of coupled binding equilibria and thus a comparison of the kinetic and thermodynamic constants describing the individual steps of activator-coactivator complex formation should be more revealing of activator potency.

However, efforts to kinetically characterize these interactions have produced conflicting models for the mechanism of complex formation.⁶³⁻⁶⁵ In one example,⁶³ the interaction between the TAD of NF- κ B p65 and various coactivators was observed to be a single-step binding event, consistent with a simple co-localization function of the transcriptional activation domain. In another study, however, a two-step binding sequence was observed between the TAD of c-Myc and the coactivator TATA-binding protein (TBP),⁶⁴ suggesting that a conformational change in one or both of the partners may contribute to the function of the activator. In both of these examples, the time resolution of the experiment was within the seconds' range, limiting the ability to detect rapid changes happening in the early stages of the interaction. An additional complication is that the kinetic studies to date have not examined activator-coactivator interactions under conditions where the activator is bound to DNA. This is important because, under physiological conditions, activators are always bound to their DNA recognition site, indicating that DNA-binding must play an important role in the process.

B.4. Stopped-flow technique utilized in characterizing transient-state kinetics

The technique of stopped-flow has been utilized in studying various biochemical interactions, for example: monitoring enzyme activity by absorbance of enzyme cofactors;⁶⁶ monitoring protein-protein interactions by fluorescence;^{67,68} monitoring protein folding by circular dichroism;⁶⁹ and monitoring protein folding by Fourier transform infrared (FT-IR) spectroscopy. The stopped-flow apparatus enables rapid mixing of two species and observation of change in real-time (Figure 2.4). Solutions of two reaction species are first placed into separate syringes. Once the driving motor presses both syringes down at the same time, equal amounts of solution are rapidly mixed in the mixer, which then enters the observation cell and replaces the existing

solution with freshly mixed solution. Spent solution that is pushed from the observation cell by fresh solution entering the stop syringe which causes the plunger to move forward and trigger data collection as well as stop the flow of the solutions. As a means of detection, changes in optical properties of the reaction mixture is monitored in the observation cell over real time (absorbance, fluorescence, light scattering, turbidity, fluorescence anisotropy etc.). The time period from the solutions entering the mixer to arriving at the observation cell is called the “dead time”, which is where one cannot observe changes in the mixture. In modern stopped-flow devices, the dead time is as low as 2 milliseconds; hence one is able to monitor fairly fast reactions.

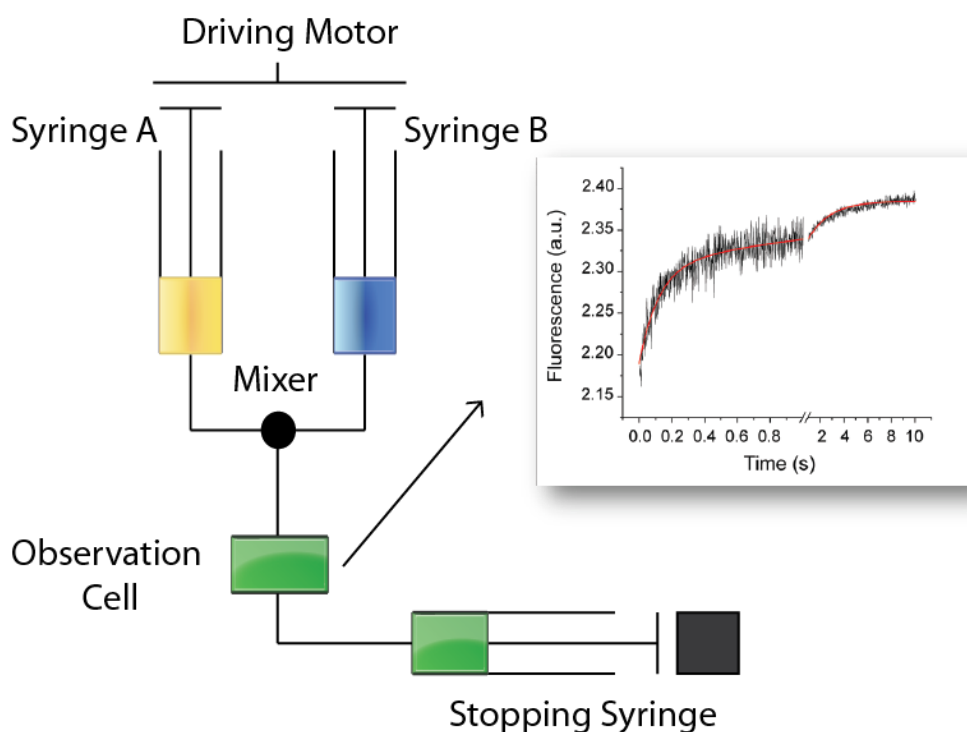


Figure 2.4 Schematic of an average stopped-flow setup. The figure on the right is a typical signal output of a stopped-flow trace.

The common experimental methods used to study the kinetics of protein-protein interactions include NMR relaxation dispersion, stopped-flow and surface plasmon resonance (SPR).⁷⁰ NMR relaxation dispersion requires relatively high concentrations of isotopically labeled protein;

hence this method was not utilized in this project due to the limitations of solubility of the studied proteins. SPR requires immobilization of one of the binding components on a charged surface, and direct comparisons of association rate constants for a number of protein-protein complexes showed that SPR data may provide different transient-state kinetic values from those obtained in solution.^{71,72} Thus we elected to use the stopped-flow technique as no components are immobilized, which is a closer approximation of physiological conditions.

The advantages of using the stopped-flow technique to study transient state kinetics include short dead time, rapid mixing, small sample volume needed for each reading, and the fact that all components of the interaction can be free in solution. For the above reasons, stopped-flow was the technique chosen in this chapter to dissect the mechanism of activator-coactivator interactions.

C. Experimental Design

C.1. Gal4 DBD-VP16, Gal4 and Gcn4 TAD fusion proteins as transcriptional activators

Many transcriptional activators are modular, thus the DBD of one activator can be interchanged onto the TAD of a different activator, and this fused activator will bind to the DNA sequence of the former activator and retain the recruitment and binding characteristics of the latter activator.⁷³

In our experimental design, each TAD is characterized in the context of a DNA-bound activator to mimic its presentation at a gene promoter. For this purpose, the well-studied Gal4 (1-100) DNA-binding domain (DBD) fused to the TADs of Gal4 (residues 840-881), Gcn4 (residues 107-144), and VP16 (residues 456-490) are expressed and purified from bacteria. Each DBD-TAD fusion protein exhibits an identical binding affinity ($K_d = 15 \pm 5$ nM) for a

fluorescently labeled consensus DNA binding site composed of two half sites, as determined through fluorescence polarization experiments (Figure 2.5);² these K_d values are consistent with those previously reported for Gal4(1-100).²⁴ Thus, the DNA binding function is independent of the TAD, indicative of the modular architecture of most transcriptional activators.^{7,8}

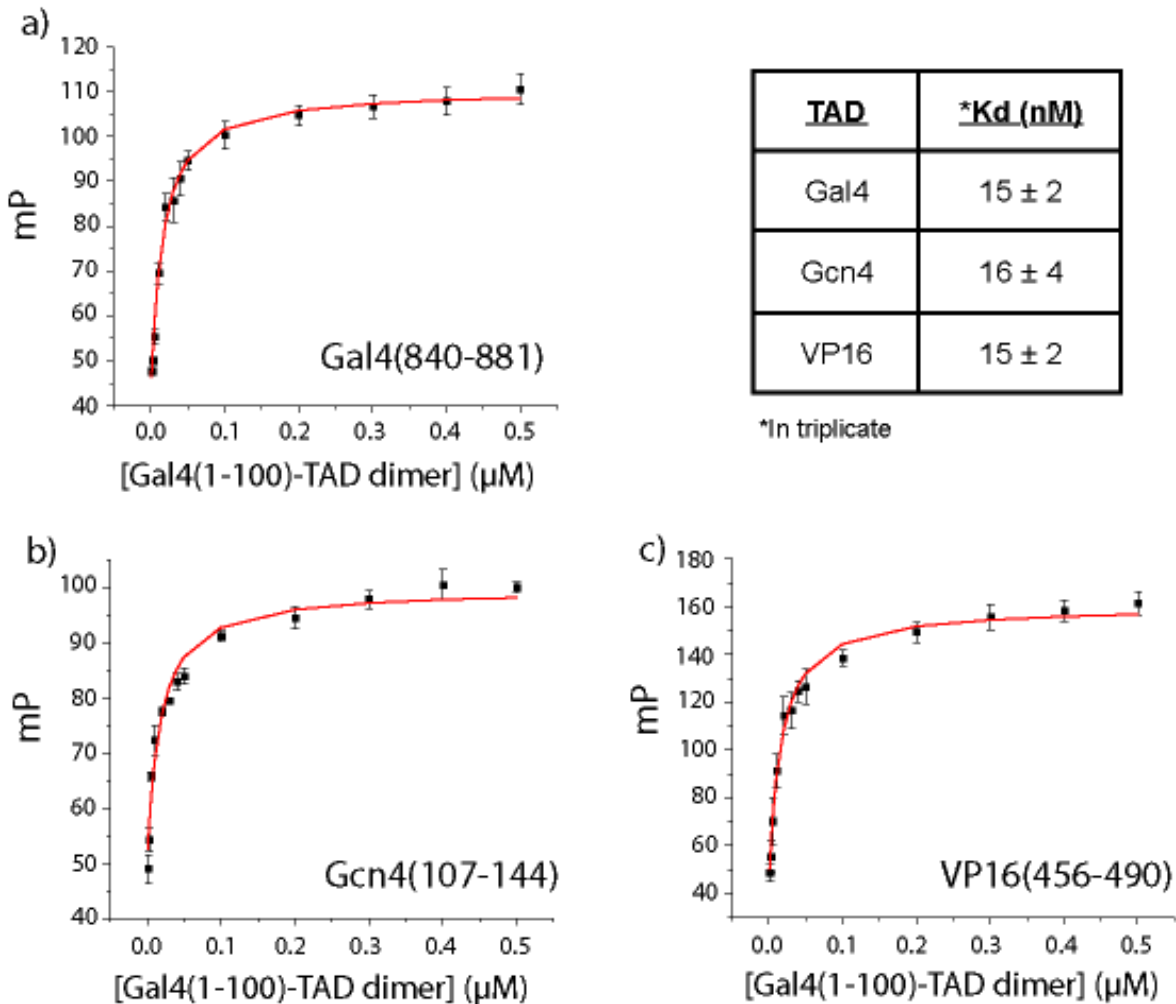


Figure 2.5 Dissociation constants of activators for DNA. A constant 1 nM concentration of the 5'-fluorescein-labeled duplex DNA was incubated with varying concentrations of Gal4(1-100)-TAD for 30 min at room temperature and the resultant polarization values at each protein concentration were obtained on a Beacon 2000 instrument (Pan Vera Corp). Each value is the average of three independent experiments with the indicated error (standard deviation). This data was obtained by Dr. Amberlyn M. Wands.

² This section of work was performed by Dr. Amberlyn M Wands.

C.2. Solubility tags for recombinant protein expression

Although transcriptional activators and Med15 lacking a solubility tag can be expressed and isolated, we observed that the propensity of these constructs to aggregate rendered them unsuitable for stopped-flow experiments, consistent with previous observations.^{27,74} Taking advantage of the ability of transcriptional activation domains to bind and function when fused to a wide variety of proteins, for example,^{29,73,75–77} we incorporated an MBP tag at the amino terminus of each activator to mitigate the aggregation propensity of the Gal4 DNA binding domain. Thus MBP-tagged activators and GST-tagged Med15 were expressed and isolated (See Methods section for complete details), and used in binding and kinetic experiments. Gel filtration experiments (Figure 2.6) indicate that GST-Med15 exists as a homodimer in solution over the range of concentrations used in the stopped-flow kinetic experiments. Thus, the indicated concentrations of Med15 represent that of the dimer.

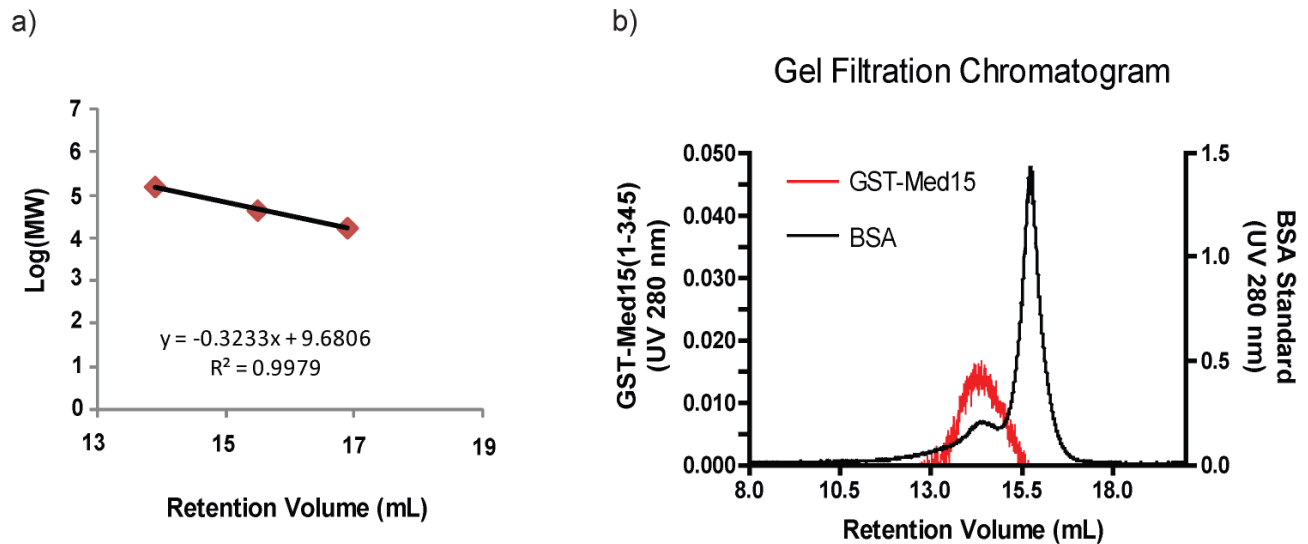


Figure 2.6 Analytical gel filtration data for GST-Med15 (1-345). **a)** Calibration curve for log (molecular weight) plotted against retention volume obtained with standards γ -globulin, ovalbumin and myoglobin. **b)** Chromatogram of 1.5 μ M (based upon a monomer) GST-Med15 (red) compared to chromatogram of BSA (black) (monomer (66.4 kD) and dimer (132.8 kD)).

C.3. Experimental setup of fluorescence stopped-flow assay

In the stopped-flow experiment, syringe A contained constant concentrations of fluorescein-DNA•Gal4 (1-100)-TAD complexes and syringe B contained excess amount of varied concentrations of Med15 to achieve pseudo-first order conditions.

To be consistent across different activators, fluorescein-labeled Gal4 binding site DNA (21 bp oligomers) was utilized to act as the fluorescent probe in complex with Gal4 (1-100)-TAD fusion proteins. When the solutions are mixed in the stopped-flow experiments, there is a 25 nM concentration of DNA and 100 nM concentration of dimeric activator present (a 1:4 ratio). From simulation studies performed in which the concentration of total DNA was set as a fixed parameter of 25 nM, and the $K_{d,app}$ for activator binding to DNA is set as a fixed parameter of 15 nM, and the total concentration of dimeric activator is varied, we estimate that the DNA is 84% bound at the concentrations we are using. This was determined to be the optimal ratio of DNA: activator in order to maximize the amount of DNA in the bound form, minimize the amount of free activator, and still maintain pseudo-first order conditions for the stopped-flow experiments. It is from this preformed DNA•Activator complex that we detect a signal upon Med15 binding.

D. Results

*D.1. Three activators show different levels of potency in *S. cerevisiae**

Med15 resides in the multi-component Mediator complex thought to function as a conduit between DNA-bound transcriptional regulators and RNA polymerase II.^{50,78,79} Both genetic and biochemical studies suggest that Med15 is a key target of amphipathic activators. Recent cross-linking experiments localized functionally important binding interactions within the amino-terminus of this protein.¹ Furthermore, decreases in transcription are observed for Gal4, VP16

and Gcn4-derived activators in Med15 (1-345)-deleted strains compared to wild type yeast, as measured by β -galactosidase assays (Figure 2.7).

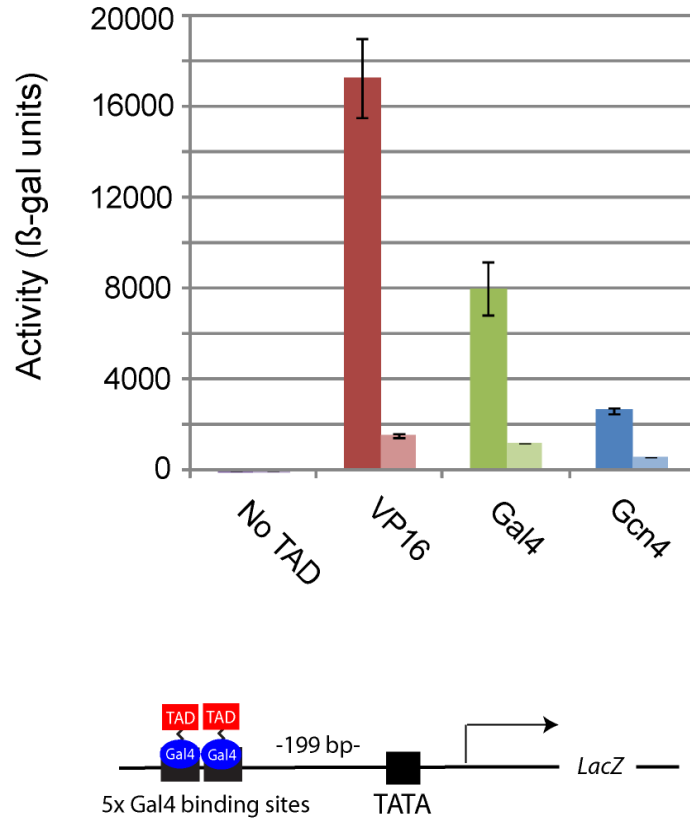


Figure 2.7 Activities of the Gal4 (1-100)-TAD activators (VP16 (456-490) > Gal4 (840-881) > Gcn4 (107-144)) in the presence of full-length Med15 (dark colored bars). Deletion of the N-terminal 345 amino acids of Med15 results in a loss of activity of all three activators (light colored bars). Each β -galactosidase assay was performed at least in triplicate. The errors shown are standard deviation of the mean (SDOM).

D.2. Binding affinity of activator•DNA complex to the coactivator Med15³

We next examined the affinity of the activator•DNA complex for the coactivator Med15 under equilibrium conditions. Titration of micromolar concentrations of GST-tagged Med15(1-345)^{27,74} into a solution of pre-formed BODIPY®FL-DNA•Gal4(1-100)-TAD complex produces a hyperbolic increase in fluorescence (Figure 2.8). The three activators exhibit similar apparent

³ This section of work was performed by Dr. Amberlyn M. Wands and is included here for completeness.

dissociation constants ($K_{d,app}$) with Med15 between 100 to 320 nM (Table 2.1); the affinity of the VP16- and Gal4-derived activators is ~2-fold higher than that of the Gcn4 activator.

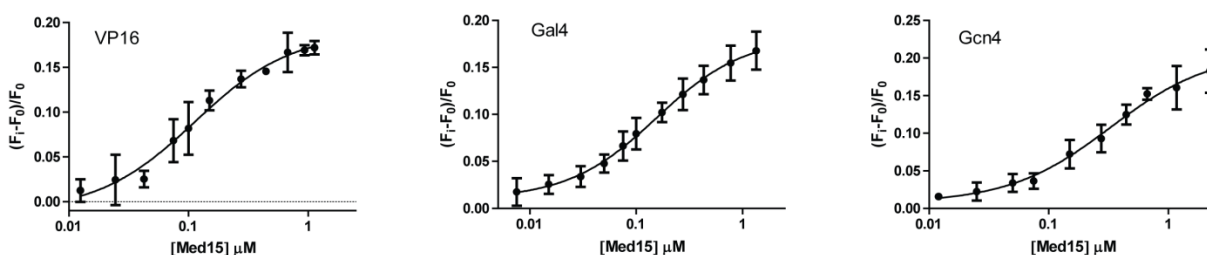


Figure 2.8 Dissociation constants of DNA-bound activators for Med15 (Gal11). A solution of 25 nM 5'-BODIPY@FL-DNA•Gal4(1-100)-TAD complex was titrated with increasing amounts of GST-Med15(1-345) in DNA-binding buffer (final concentrations: 20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β ME, 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) at 25 °C and the resultant fluorescence enhancement ($\lambda_{ex} = 502$ nm; $\lambda_{em} > 510$ nm) at each coactivator concentration (dimer concentration) was monitored on an Eclipse spectrofluorometer (Varian Corp). Each value is the average of three independent measurements and the errors associated with these values are the standard deviation. This data was obtained by Dr. Amberlyn M. Wands.

Table 2.1 Experimentally determined thermodynamic and kinetic constants^a

TAD	$K_{d,app}$ ^b (nM)	k_{on} ($\mu\text{M}^{-1}\text{s}^{-1}$)	$k_{obs,1}$ (y-intercept) (s^{-1})	$k_{obs,2}^{\max}$ ^c (s^{-1})
VP16	100 ± 20	23 ± 2	12.1 ± 0.7	4.6
Gal4	143 ± 7	9.1 ± 0.8	3.6 ± 0.4	0.67
Gcn4	320 ± 50	66 ± 8	26 ± 4	2.6 ± 0.4

- Assays were carried out in DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β -mercaptoethanol, 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) at 25°C.
- Data fit according to a 1:1 binding interaction of DNA-bound activator to dimeric Med15. Each value is the average of three independent experiments with the indicated error (standard deviation).
- The value of $k_{obs,2}^{\max}$ is extrapolated as described in Experimental methods.

D.3. Transient-state analysis of activator-coactivator association reveals biphasic binding kinetics⁴

The association kinetics for formation of the BODIPY@FL-DNA•activator•Med15 complex

⁴ This section was performed collaboratively with Dr. Amberlyn M. Wands.

were assessed by mixing Med15 with the BODIPY-DNA•activator complex in a stopped-flow fluorescence spectrometer under pseudo-first order conditions and measuring the time-dependent changes in fluorescence. Gal4, Gcn4, and VP16 all exhibit an increase in fluorescence over time at each Med15 concentration tested, producing time courses that are biphasic and best fit by a double exponential at both 25 °C (Figure 2.9) and 16 °C (Figure 2.10).

Analogous experiments performed by mixing DNA-bound Gal4 (1-100) lacking a TAD with Med15 or mixing the DNA•activator complexes with GST produced no increase in fluorescence over background (Figure 2.11), indicating that the fluorescence change reports only the Med15-TAD interaction.

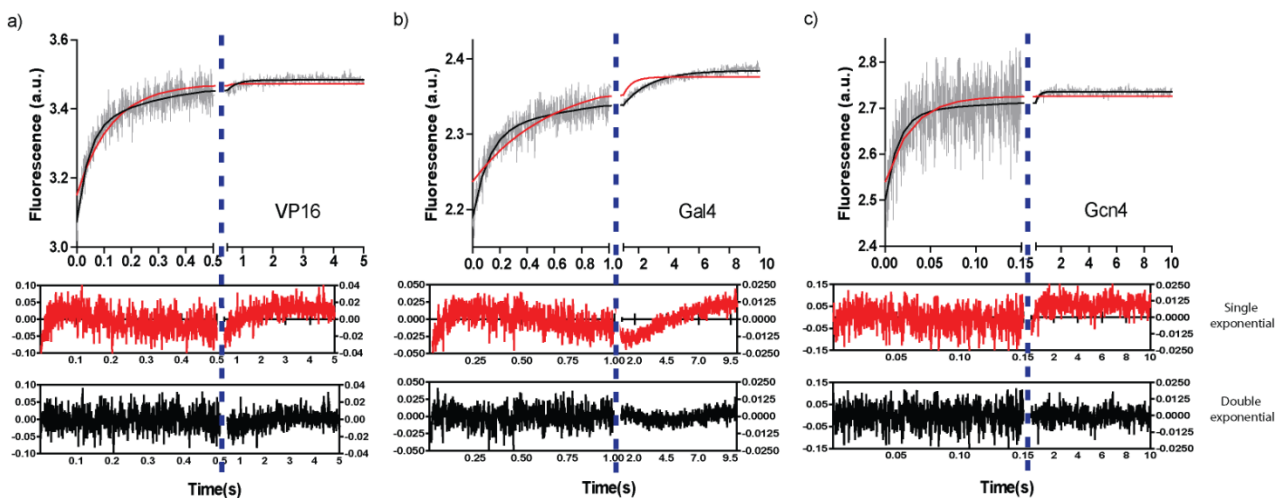


Figure 2.9 Biphasic kinetics observed for activator•DNA complex binding to Med15. Time-dependent changes in fluorescence ($\lambda_{\text{ex}} = 502 \text{ nm}$; $\lambda_{\text{em}} > 510 \text{ nm}$) after mixing either **a**) 5'-BODIPY@FL-DNA•Gal4(1-100)- VP16(456-490), **b**) 5'-BODIPY@FL-DNA•Gal4(1-100)-Gal4(840-881), or **c**) 5'-BODIPY@FL-DNA•Gal4(1-100)- Gcn4(107-144) activator complex (25 nM DNA and 100 nM DBD-TAD fusion protein after mixing) with an equal volume of dimeric GST-Med15 (0.5 μM with VP16; 0.625 μM with Gal4 and 0.75 μM with Gcn4 after mixing) in a stopped-flow apparatus in DNA-binding buffer at 25 °C (grey line). The red line superimposed on the time course is the best single exponential fit to the data, while the black line is the best fit of the sum of two exponentials to the data. Residuals for the single exponential fit (red line) or double exponential fit (black line) to the time courses above reveal that a double exponential is the best fit to the data. Values for the first and second time domains are plotted with respect to the left and right axes, respectively. This data was performed collaboratively with Dr. Amberlyn M. Wands.

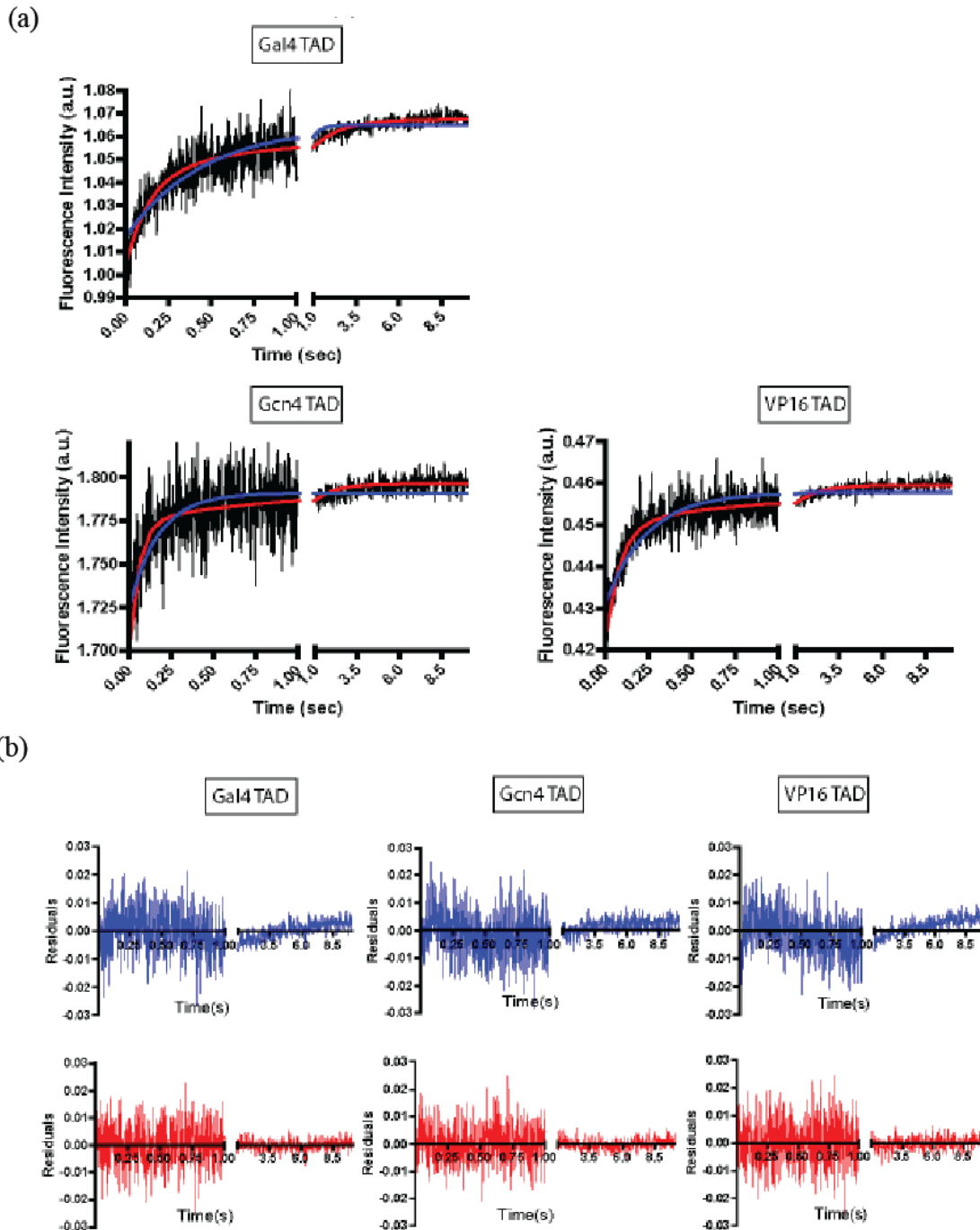


Figure 2.10 **a**) Binding time course of 50 nM 5'-BODIPY®FL-labeled, double-stranded oligonucleotides pre-complexed to 200 nM dimeric MBP-tagged Gal4 (1-100)-TAD mixed with Med15 (0.375 μ M after mixing) in a stopped-flow apparatus in DNA-binding buffer at 16 °C. The blue line superimposed on the time course is the best fit of the sum of one exponential to the data, while the red line is the best fit of the sum of two exponentials to the data. **b**) Residuals for the single exponential fit (blue line) or double exponential fit (red line) to the time course in **a**) reveals that a double exponential is the best fit to the data.

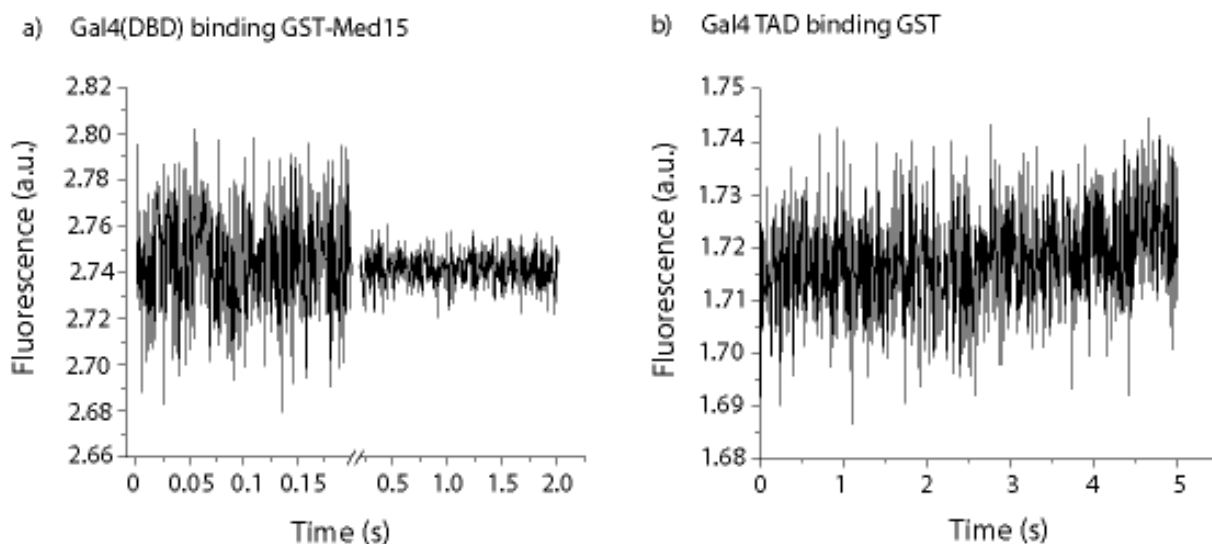


Figure 2.11 Negative control stopped-flow experiments. **a)** No fluorescence enhancement is observed in the time-course when 50 nM 5'-BODIPY@FL-labeled, double-stranded oligonucleotides pre-complexed to 200 nM dimeric MBP-tagged Gal4 (1-100) is mixed with Med15 (0.375 μ M after mixing) at 25 $^{\circ}$ C. **b)** No fluorescence enhancement is observed in the time-course when 5'-BODIPY@FL-DNA•Gal4 (1-100)-Gal4 (840-881) activator complex (5 nM after mixing) is mixed with GST (0.25 μ M after mixing) at 25 $^{\circ}$ C. This data was obtained by Dr. Amberlyn M. Wands.

The observed rate for the faster phase ($k_{\text{obs},1}$) is linearly dependent on the Med15 concentration for all three activator constructs with a positive slope reflecting an apparent bimolecular association rate constant that varies from $9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ for Gal4 to $66 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ for Gcn4 (Table 2.1, Figure 2.12a). These values are in the range of diffusion-controlled processes ($10^6 - 10^7 \text{ M}^{-1}\text{s}^{-1}$),⁸⁰ suggesting that this step is the binding of Med15 to the activator•DNA complex.⁵ The observed rates for the slower phase ($k_{\text{obs},2}$) (Figure 2.12b) are dependent on the identity of the TAD (Figure 2.12b and Table 2.1) and have a modest dependence on the concentration of Med15. The higher errors in the rates are due to the smaller amplitude of this slower phase. Lowering the temperature to 16 $^{\circ}$ C results in a more prominent slower phase with rates that are decreased 2-4 fold (Figure 2.10).

⁵ To a first approximation, these values are identical. Additionally, the occurrence of a rapid conformational change within the time-scale of this initial binding phase could contribute to the small differences in the apparent bimolecular association rate constants of Gal4 and VP16 over that of Gcn4.

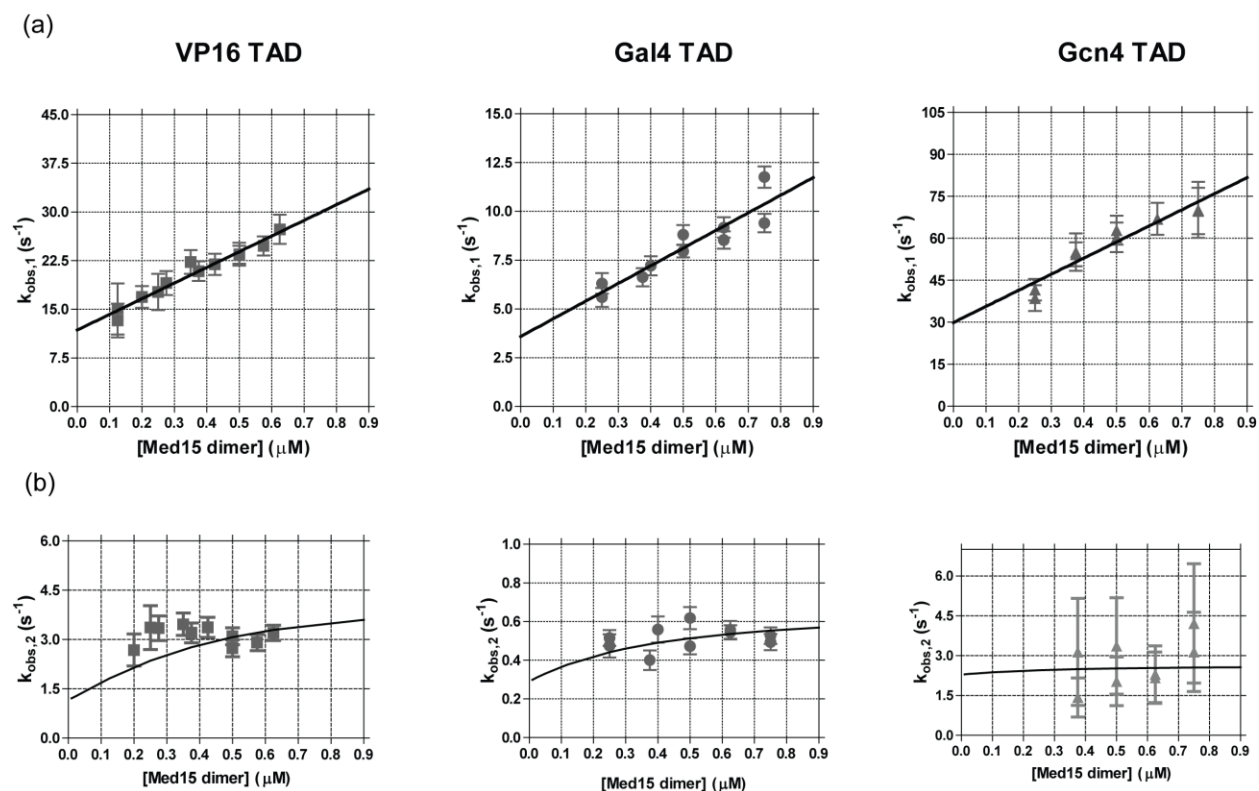


Figure 2.12 Plot of $k_{obs,1}$ **a)** and $k_{obs,2}$ **b)** of the activators VP16 (■), Gal4 (●) and Gcn4 (▲) against the concentration of Med15 (0.125 - 0.75 μM). The standard errors for both observed rates are indicated. The solid line in a) is a curve-fit to these data. The solid line in b) is simulated $k_{obs,2}$ according to Scheme A using the extrapolated $k_{obs,2}$ max values of 4.6 s^{-1} (VP16), 0.67 s^{-1} (Gal4) and 2.6 s^{-1} (Gcn4). Simulated $k_{obs,2}$ values at each concentration of Med15 were obtained by fitting Eq. 2.4 to the simulated traces of the time-dependent formation of $[\text{DNA}\cdot\text{Activator}\cdot\text{Med15}] + [(\text{DNA}\cdot\text{Activator})\cdot\text{Med15}]$. In these simulations, the microscopic rate constants k_1 and k_{-1} (Table 2.2), as well as calculated values for k_2 and k_{-2} , were constants (VP16: $k_2=3.2 \text{ s}^{-1}$, $k_{-2}=1.4 \text{ s}^{-1}$; Gal4: $k_2=0.37 \text{ s}^{-1}$, $k_{-2}=0.30 \text{ s}^{-1}$; Gcn4: $k_2=0.28 \text{ s}^{-1}$, $k_{-2}=2.3 \text{ s}^{-1}$, all values are calculated within 20% propagated error).

These biphasic binding kinetics suggest a minimal kinetic mechanism that requires a two-step binding process. The fact that the rate of the slow phase is sensitive to temperature and does not exhibit a linear dependence on Med15 concentration (Figure 2.12 and Figure 2.10) is consistent with the occurrence of a unimolecular conformational change step that occurs either after or before the bimolecular binding step (Schemes A and B, respectively, in Figure 2.13). Additionally, the relative amplitudes of the two phases depend on the TAD of each activator (Figure 2.14), consistent with this kinetic mechanism.

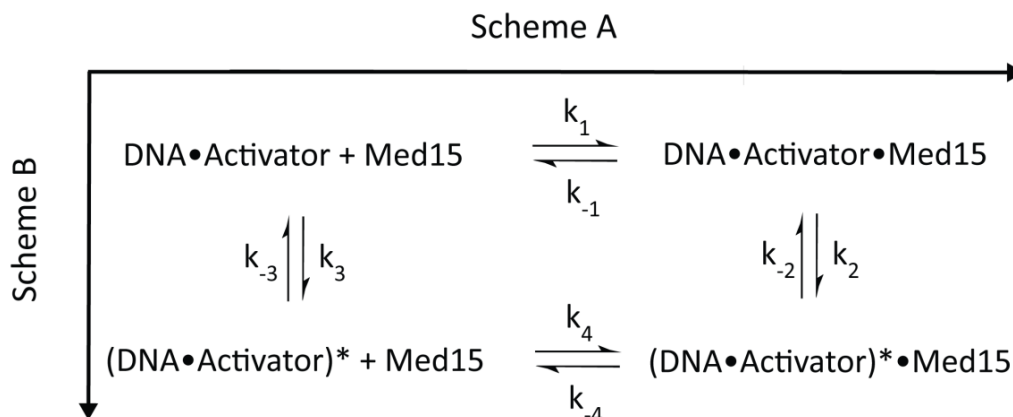


Figure 2.13 Schematics of the two limiting binding models for DNA-bound activators interacting with Med15. In Scheme A, the conformational change occurs after an initial binding event, whereas in Scheme B, the DNA-bound activator undergoes a conformational change prior to associating with Med15. The pathway used by an activator may change depending upon the concentration(s) of the individual binding partners.

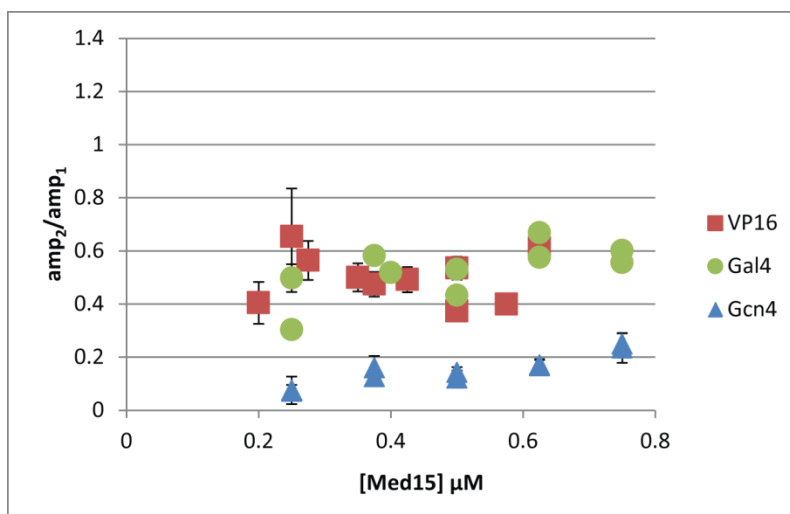


Figure 2.14 Plot of the ratio of amplitude of the slow phase (amp_2) relative to the amplitude of the fast phase (amp_1) for all three activators VP16 (■), Gal4 (●) and Gcn4 (▲) against the concentration of Med15 (0.125 - 0.75 μM).

D.4. Microscopic kinetic rate constants calculated for the two-step binding model.

The full solutions to the rate equations describing the two limiting binding mechanisms shown in Figure 2.13 have been published elsewhere.⁸¹ The microscopic rate constants included in $k_{\text{obs},1}$ and $k_{\text{obs},2}$ can be approximated for each of the limiting pathways (Figure 2.13, Schemes

A and B) according to Equations 2.3-2.8 and Equations 2.9-2.14, respectively (see Experimental methods).⁸² Using these equations, the microscopic rate (k_1 , k_{-1}) and equilibrium (K_1) constants for Scheme A, the binding mechanism in which a conformational change occurs after the bimolecular collision step, were calculated from the experimentally obtained values presented in Table 2.1 for k_{on} , $k_{obs,1}^{(y-intercept)}$ and the values of $k_{obs,2}^{max}$ extrapolated to saturating Med15 (see Experimental methods). Specifically, k_1 is determined from the slope of $k_{obs,1}$ versus Med15 concentration, k_{-1} equals the value of $k_{obs,1}^{(y-intercept)}$ minus $k_{obs,2}^{max}$, and K_1 is the ratio k_1/k_{-1} . The value of K_2 was then calculated from $K_{d,app}$ (Table 2.1) and K_1 . The results are summarized in Table 2.2. The values for k_2 and k_{-2} were also estimated (from $k_{obs,2}^{max}$ and K_2) for use in simulation studies presented later.⁶

Table 2.2 Calculated rate constants for binding model in Scheme A
(Propagation of error for calculations resulted in ~ 20% error for all values.)

TAD	k_1 ($\mu\text{M}^{-1}\text{s}^{-1}$)	k_{-1} (s^{-1})	K_1 (μM^{-1})	K_2
VP16	23	7.6	3.1	2.3
Gal4	9.1	2.9	3.1	1.3
Gcn4	66	24	2.8	0.12

Similarly the microscopic rate and equilibrium constants for the binding mechanism presented in Scheme B, in which a conformational change occurs before the bimolecular collision step, are presented in the Table 2.3.

Table 2.3 Calculated rate and equilibrium constants for the binding mechanism presented in Scheme B

TAD	k_3 (s^{-1})	k_{-3} (s^{-1})	k_4 ($\mu\text{M}^{-1}\text{s}^{-1}$)	k_{-4} (s^{-1})	K_4 (μM^{-1})	K_3
VP16	4.6	6.9 ± 0.8	23 ± 2	0.95 ± 0.3	24 ± 14	0.70 ± 0.05
Gal4	0.67	2.7 ± 0.5	9 ± 1	0.26 ± 0.5	34 ± 40	0.25 ± 0.04
Gcn4	2.6 ± 0.4	21 ± 5	66 ± 8	2.3 ± 2.9	29 ± 36	0.12 ± 0.03

⁶ While k_1 and k_{-1} are calculated directly from the kinetic data, to obtain k_2 and k_{-2} one must use the K_d determined through equilibrium binding measurements (Figure 2) that were carried out under conditions that differ slightly. Experiments to directly measure the off rate (k_{off}) in order to better define k_2 and k_{-2} were unsuccessful due to aggregation propensities of the TADs at the high concentrations needed for these experiments.

D.5. Global kinetic simulation studies suggest that Med15 binding precedes the conformational change.

The data from individual stopped flow experiments were globally analyzed using the KinTek Global Kinetic Explorer Program with either the Scheme A or Scheme B models. The values for k_1 (Scheme A) or k_4 (Scheme B) were used as a constraint in these fits (Table 2.2 and Table 2.3). Scheme A provided a modest to significantly better fit to the experimental data relative to Scheme B.⁷ This suggests that a model in which the conformational change step occurs after the formation of the Med15•activator•DNA complex (Scheme A) is the more appropriate framework for describing this interaction.

To examine whether the limiting kinetic models are sufficient to describe the data, the transient kinetic curves were simulated using the rate constants for Scheme A (calculated by Equations 2.3-2.8: k_1 and k_{-1} (Table 2.2), as well as k_2 and k_{-2}) as fixed parameters, and allowing the change in fluorescence for each step to vary. The simulated curves agree well with the

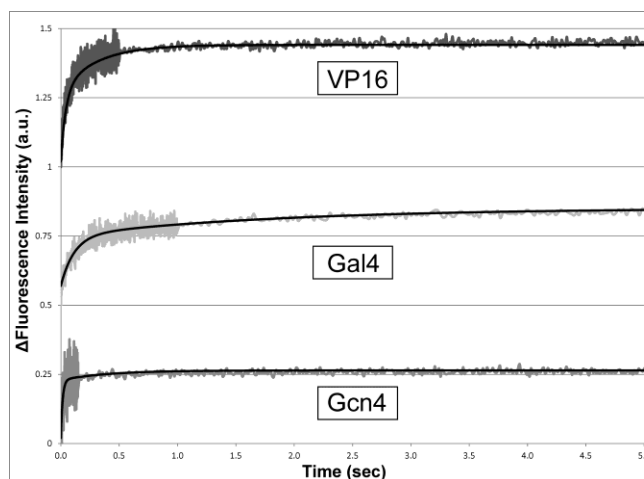


Figure 2.15 Stopped-flow traces of 25 nM 5'-BODIPY®FL-DNA•Gal4(1-100)-VP16(456-490), 5'-BODIPY®FL-DNA•Gal4(1-100)-Gal4(840-881), and 5'-BODIPY®FL-DNA•Gal4(1-100)-Gcn4(107-144) mixed with 0.625 μ M dimeric GST-Med15 are overlaid with simulated traces (black line) based on Scheme A using Kintek Global kinetic explorer. The values for the microscopic rate constants k_1 and k_{-1} (Table 2.2), as well as calculated values for k_2 and k_{-2} , were used as constraints in the simulations. (VP16: $k_2=3.2$ s⁻¹, $k_{-2}=1.4$ s⁻¹; Gal4: $k_2=0.37$ s⁻¹, $k_{-2}=0.30$ s⁻¹; Gcn4: $k_2=0.28$ s⁻¹, $k_{-2}=2.3$ s⁻¹, all values are calculated within 20% propagated error.) Experimental data and simulated traces reflect a change in fluorescence intensity. The three TADs are offset by 1.0 (VP16), 0.5 (Gal4) and 0 (Gcn4) a.u. for clarity purposes.

⁷ Goodness of fit was determined by Chi²/Degrees of Freedom (DoF) values. Chi²/DoF values for Scheme A: VP16:1.07; Gal4:1.06; Gcn4:1.07. Chi²/DoF values for Scheme B: VP16:1.14; Gal4:1.52; Gcn4:1.14.

experimental stopped-flow traces (Figure 2.15). In comparing the three activators, the traces show a decrease in amplitude of the slow phase relative to the fast phase (Figure 2.14) consistent with the decreasing value of K_2 .

In summary, these simulations demonstrate that the data are best-described by a scheme in which the observed conformational change occurs after the DNA-bound activator associates with Med15 (Figure 2.13, Scheme A).⁸³⁻⁸⁵ Consistent with this mechanism, structural studies performed on the isolated TADs of Gal4 and VP16 have shown them to be mostly unstructured in the absence of their binding partners, but helical in the bound form^{18,86-88} and this behavior is characteristic of other isolated TADs of this class as well.^{83,89-91}

E. Discussion

Here, for the first time, we report the transient kinetics of prototypical activator•DNA complexes interacting with the key coactivator Med15. Importantly, these data demonstrate that three distinct amphipathic activators interact with Med15 via a two-step binding mechanism that includes a conformational change. The limiting kinetic pathways for this two-step mechanism differ only in the order of the two steps (Scheme A and Scheme B, Figure 2.13). While our kinetic data do not absolutely distinguish these mechanisms, Scheme A, in which a conformational change step occurs after the association step, is most consistent with the high value of the bimolecular rate constant, the simulation data, and previous data demonstrating the formation of helical structure in the isolated TADs of Gal4 and VP16 upon interaction with their binding partners.^{18,86-88} We therefore further analyze these data using this kinetic mechanism.

A comparison of the measured rate and equilibrium constants determined for Scheme A reveals trends that provide insight into the differential activity of the three activators Gal4, Gcn4, and VP16. In general, the TAD activity correlates with a more favorable $K_{d,app}$ for Med15 (VP16

> Gal4 > Gcn4) (Figure 2.16, Table 2.1).

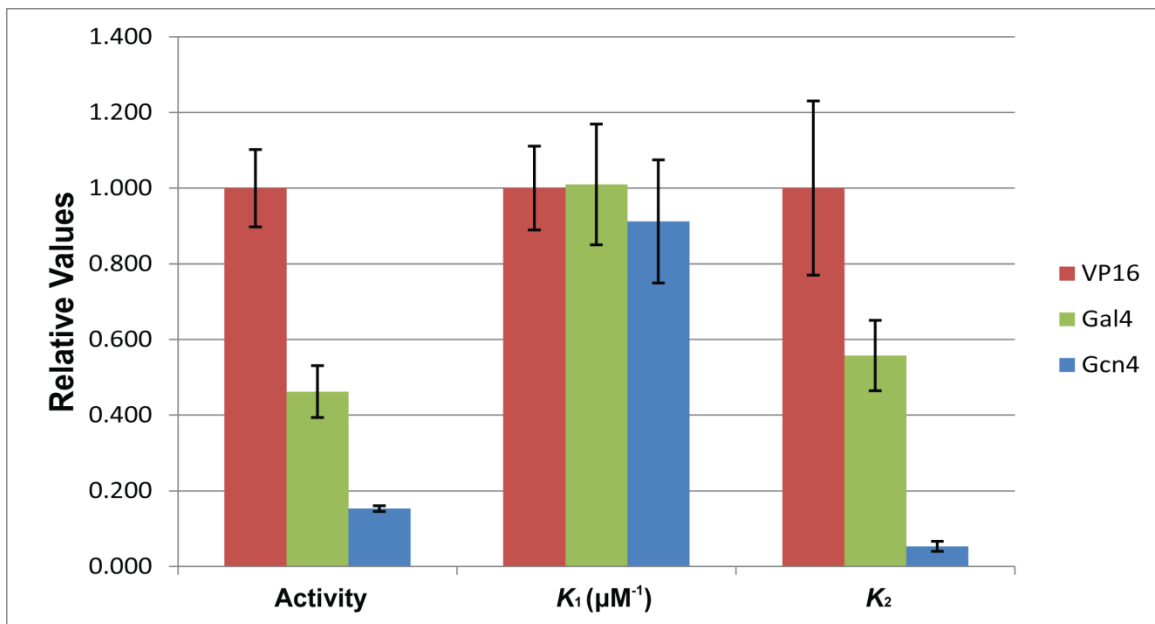


Figure 2.16 Comparison of the equilibrium constants (K_1 and K_2) for the binding mechanism presented in Scheme A. Values are normalized to the most potent activator, VP16. Activity values are obtained by *in vivo* β -gal assays. Error bars depict standard deviation (SD) of each averaged value.

Analysis of the thermodynamics of the individual steps within the kinetic mechanism reveals that this correlation is governed not by the equilibrium constant for the initial bimolecular association step (K_1), which is invariant among all three TADs, but by the value of the subsequent conformational change step (K_2) (Figure 2.16). In particular, a greater value for K_2 (more favorable conformational change step) correlates with a more active TAD. To play an important role in regulating activation by varying the concentration of the active (DNA•activator)*•Med15 complex, the conformational change must be unfavorable for some or all of the TADs. Consistent with this model, for VP16, the equilibrium constant for the conformational change is moderately favorable ($K_2 = 2.3$) indicating that at saturating Med15, more than 70% of the DNA•VP16•Med15 complex forms the active conformation. In contrast, the conformational change is unfavorable for Gcn4 ($K_2 = 0.12$) indicating that only ~10% of the DNA•Gcn4•Med15 complex is in the active conformation. Although less well-defined by the

data, a similar trend is observed for kinetic partitioning between the conformational change and dissociation of Med15 (k_2/k_{-1}); the most active TAD, VP16, has the largest value of k_2 and the largest partition ratio, indicating that the intermediate DNA•activator•Med15 complex undergoes the conformational change the fastest of the three activators and has the highest partition ratio to form the new conformer relative to dissociation.

F. Conclusions

In conjunction with structural evidence that isolated TADs of this class are intrinsically disordered and undergo structural enhancement upon interaction with a target protein^{8,92,93} we hypothesize that the different degree of transient structure inherent to each TAD may play a role in the differences observed in the kinetics during the conformational change step to form the final DNA•activator•Med15 complexes (k_2). For example, the faster conformational change observed for VP16 compared to that of Gal4 and Gcn4 may be attributed to the fact that residues 472-479 within the isolated VP16 TAD have a propensity to form a relatively well-defined helix,⁹⁴ although additional structural studies will be required to refine this model. Furthermore, it is important to note that although the current method of detection through an enhancement in fluorescence of a probe on the DNA limits our abilities to determine which component(s) within the complex is undergoing the conformational change, future studies in which the probes are placed at the TAD-target interface will provide additional details in this regard.

Finally, the favorability of the conformational change (K_2), which is encoded in the primary amino acid sequence of each TAD by the particular arrangement of the acidic, polar, and hydrophobic amino acid residues,⁹⁵ may explain why three TADs that target the same protein via a shared binding mechanism can lead to differences in the assembly of the pre-initiation complex

and transcriptional output. Thus, the identification of artificial TADs with additional structural stability may lead to artificial activators with enhanced transcriptional activity.

G. Experimental Methods

Plasmid Construction

Plasmids encoding MBP-tagged Gal4 (1-100) fused to different TADs and GST-tagged Med15 (1-345) were constructed and prepared by standard molecular cloning methods.

Table 2.4: Plasmids used in this study

Plasmid name	Reference	Function
pGal4(1-100)-Gal4(840-881)	⁹⁶	ARS/CEN yeast expression plasmid under the control of a β -actin promoter with a HIS ⁺ selection marker
pGal4(1-100)-Gcn4(107-144)	⁹⁶	
pGal4(1-100)-VP16(456-490)	²	
ycplac111-Med15	¹	ARS/CEN yeast expression plasmid under the control of a Med15 native promoter with a LEU ⁺ selection marker
ycplac111-Med15 Δ (1-345)	¹	
pMCSG9-Gal4(1-100)-Gal4(840-881)	²	Expresses activators fused to the His ₆ -MBP solubility tag in <i>E. coli</i>
pMCSG9-Gal4(1-100)-Gcn4(107-144)	²	
pMCSG9-Gal4(1-100)-VP16(456-490)	²	
pGEX-Med15(1-345)	²	Expresses Med15 fused to the GST solubility tag in <i>E. coli</i>

β -galactosidase Assays

LS41 Δ Med15 [JPY9::ZZ41, *Mat α his3 Δ 200 leu2 Δ 1 trp1 Δ 63 ura3-52 lys2 Δ 385 gal4 URA::pZZ41 Med15::TRP*] yeast was co-transformed with plasmids encoding each Gal4 (1-100)-TAD fusion and a fragment of Med15. The activity of each activator construct was monitored using β -galactosidase assays as previously described.⁹⁷

Protein Expression and Purification

His₆-MBP-tagged activators: Expression of the activators fused to a His₆-MBP tag was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen) as previously described.⁹⁸ Maltose-binding protein (MBP) has been demonstrated to effectively enhance the solubility of aggregation-prone proteins;⁹⁹ in addition, fusion of solubility tags to the N-terminus of the Gal4 DBD has been reported previously not to impact activator function.^{77,100} Briefly, cultures (50 mL) inoculated with single colonies were grown overnight at 37 °C (250 rpm) in Lennox L Broth (Research Products International) supplemented with ampicillin (100µg/mL) and chloramphenicol (34 µg/mL) before dilution (50-fold) into 8 x 50 mL cultures of Lennox L Broth supplemented with ampicillin (100 µg/mL). After an OD₆₀₀ of 0.4 was reached, protein over-expression was induced with IPTG (final concentration 1 mM) in the presence of 20 µM ZnSO₄ for 5 hours. Cells in 50 mL culture were pelleted by centrifugation, resuspended in 10 mL lysis buffer A (10 mM Tris, pH 8.0 at 4 °C, 500 mM NaCl, 10% glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), and Roche Complete Protease Inhibitor Cocktail), then lysed using sonication. His-tagged protein was isolated by incubating cell lysate with 200 µL of Ni-NTA beads (Qiagen) for 1 hour at 4 °C, followed by washing 8 times with 1 mL wash buffer A (20 mM Tris, pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 1 mM β-ME, 0.1% Tween* 20 (v/v), 30 mM imidazole). The protein was eluted from the beads by incubation at 4 °C overnight with 1 mL elution buffer A (20 mM Tris pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 250 mM imidazole). The protein solution was buffer exchanged into storage buffer A (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM β-ME, 1 mM EDTA, 20 µM ZnSO₄) using a PD-10 column (GE Healthcare), and the protein concentration was measured using absorbance at 280 nm. The

identity and purity (>90%) of the protein was verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

GST-tagged Med15 (1-345): Expression of GST-Med15 (1-345) was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen). Glutathione S-transferase (GST) has been demonstrated to effectively enhance the solubility and stability of aggregation-prone proteins (8); in addition, fusion of GST to the N-terminus of Med15(Gal11) has been reported previously not to impact activator binding to this target protein.^{1,27,74} Briefly, cultures (50 mL) from single colonies were grown overnight at 37 °C (250 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100µg/mL) and chloramphenicol (34 µg/mL) before dilution (100-fold) into 4 x 1 L of Select APS Super Broth supplemented with ampicillin (100 µg/mL). After an OD₆₀₀ of 0.3 was reached, the cultures were cooled for 45 min at 16 °C (150 rpm), and expression was induced with IPTG (final concentration 0.1 mM) for 5-6 hours at 250 rpm. Each cell pellet was resuspended in 25 mL lysis buffer B (100 mM PBS pH 7.4 (Pierce), 0.2% NP-40 Substitute (Fluka), 10% glycerol (v/v), 1 mM DTT and Roche Complete Protease Inhibitor Cocktail), lysed using sonication, and the GST-tagged protein was isolated using Glutathione Sepharose 4B (GE Healthcare). The cell lysate was incubated with 2 x 1 mL of glutathione beads for 1 hour at 4 °C. The beads were washed 6 times with 10 mL wash buffer B (100 mM PBS pH 7.4 (Pierce), 0.2% NP-40 Substitute (Fluka), 10% glycerol (v/v), 1 mM DTT), and the protein was eluted from the beads by incubation at 4 °C overnight with 1 mL elution buffer B (50 mM Tris pH 8.0 at 4 °C, 0.015 M reduced glutathione, 0.1% NP-40 Substitute). Additional protein was eluted from the column by twice incubating the beads with elution buffer for 1 hour at 4 °C. The protein samples were combined and concentrated using a Centriprep 10K centrifugal filter device before buffer exchange into storage buffer B (10 mM PBS pH 7.4 (Pierce), 10% glycerol (v/v), 0.01%

NP-40 Substitute, 1 mM DTT) using a PD-10 column (GE Healthcare). The protein was then concentrated using a Vivaspin 30K centrifugal filter device, and the protein concentration was measured using absorbance at 280 nm. The identity and purity (>85%) of the protein was verified by reducing SDS-PAGE with appropriate molecular weight standards.

Analytical Gel Filtration

Analytical gel filtration was performed to determine the oligomeric state of GST-Med15 (1-345) over the concentration range used in the fluorescence stopped-flow kinetic experiments described below. A final concentration of 1.5 μM and 0.5 μM (based on monomer concentrations) of GST-Med15 (1-345) was run through a Superose 6 gel filtration column equilibrated with DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β -mercaptoethanol, 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA).²⁷ At the highest concentration of GST-Med15 (1-345) tested (1.5 μM), only a single peak eluted as detected by UV (280 nm), with a projected molecular weight of 114 kDa as determined from molecular weight standards (Figure 2.6). These data are consistent with a dimeric state of GST-Med15 (1-345) (monomer = 62.5kD, dimer = 125 kDa). This same species was the only species observed at the lowest concentration tested (0.5 μM) as determined by a western blot probing for GST on eluted fractions (Figure 2.6).

Fluorescence Polarization Assays to Measure DNA Affinity to Transcriptional Activators

The 20 bp oligonucleotide 5'-TCC GGA GGA CTG TCC TCC GG-3' (26) labeled at the

5'end with fluorescein or BODIPY®FL was purchased from Invitrogen. The fluorescently labeled oligonucleotide was then annealed with an unlabeled complementary oligonucleotide (5' - GCC GGA GGA CAG TCC TCC GG-3') in annealing buffer (10 mM HEPES pH 7.5, 150 mM NaCl) by heat denaturation for 7 min at 95 °C, followed by cooling at room temperature for 30 min and 4 °C for 30 min.

Annealed fluorescein-labeled oligonucleotide (45 μM) was diluted in DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β-mercaptoethanol (βME), 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) to a concentration of 1.25 nM. Then 200 μL of the DNA solution was added to a series of 50 μL solutions of varying activator concentrations in storage buffer A (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM βME, 1 mM EDTA, 20 μM ZnSO₄) to obtain the final concentrations of up to 0.5 μM. The samples were incubated for 30 min at room temperature before the degree of fluorescence polarization was measured (Beacon 2000, Pan Vera Corp). A binding isotherm that accounts for ligand depletion⁶² (assuming a 1:1 binding model of dimeric activator to duplex DNA) was fit to the observed mP values as a function of activator to obtain the apparent equilibrium dissociation constant, K_d :

$$y = c + (b - c) \times \frac{[(Kd+a+x) - \sqrt{(Kd+a+x)^2 - 4ax}]}{2a} \quad (\text{Eq. 2.1})$$

where “a” and “x” are the total concentrations of duplex DNA and dimeric activator, respectively, “y” is the observed polarization at any activator concentration, “b” is the maximum observed polarization value, and “c” is the minimum observed polarization value. Each data point in Figure 2.5 is an average of three independent experiments with the indicated error (standard deviation). Data analysis was performed using Origin software (Originlab Corp).

Fluorescence Titration Assays to Measure Med15 Affinity to Transcriptional Activators

GST-Med15(1-345) in storage buffer B (10 mM phosphate-buffered saline pH 7.4 (Pierce), 10% glycerol (v/v), 0.01% NP-40 substitute, 1 mM dithiothreitol (DTT)) was titrated (without exceeding a 5% volume increase) into the following solution to obtain the final Med15 concentrations indicated in Figure 2: 25 nM 5'-BODIPY®FL-labeled, double-stranded oligonucleotide pre-incubated with 100 nM dimeric activator in DNA-binding buffer. Under these conditions 84% of the DNA is estimated to be complexed with activator, as calculated using Equation 2.1. The fluorescence intensity of BODIPY®FL was monitored on an Eclipse spectrofluorometer (Varian Corp) ($\lambda_{\text{ex}} = 500 \text{ nm}$, $\lambda_{\text{em}} = 512 \text{ nm}$; 5 nm band pass). The fluorescence intensity (F_i) was corrected for dilution effects and background fluorescence from the DNA•activator complex (F_0), such that $\Delta F = F_i - F_0$. The observed fluorescence fraction increase, $\Delta F/F_0$, was plotted as a function of Med15 concentration, and a binding isotherm that accounts for ligand depletion (Equation 2.1) (assuming a 1:1 binding model of DNA•activator complex to dimeric GST-Med15) was fit to the data using Origin 7.0 software to obtain the apparent equilibrium dissociation constant, K_d . The parameters “a” and “x” represent the total concentrations of DNA•activator complex and dimeric Med15, respectively, “y” is the observed fluorescence enhancement at any Med15 concentration, “b” is the maximum observed relative fluorescence enhancement value, and “c” is the minimum observed relative fluorescence enhancement value. Each data point in Figure 2.8 is an average of three independent experiments with the indicated error (standard deviation).

Fluorescence Stopped-flow Kinetic Experiments

Stopped-flow experiments were performed on a KinTek model SF-2001 stopped-flow

equipped with a 75W Xe arc lamp in two-syringe mode. 50 nM 5'-BODIPY®FL-labeled, double-stranded oligonucleotide pre-complexed to 200 nM dimeric MBP-tagged Gal4(DBD) fused to different TADs in DNA-binding buffer (25 nM and 100 nM after mixing) was mixed with an equal volume of GST-tagged Med15(1-345) (final concentration after mixing 0.125-0.75 µM for homodimer) in DNA binding buffer at 25 °C. BODIPY®FL was excited at 502 nm and its emission was monitored at wavelengths > 510 nm using a long-pass filter (Corion). All kinetic traces reported are an average of four to six independent determinations. Sum of exponentials was fit to the transient kinetic time courses, $F(t)$ as in Equation 2.2, to obtain the fluorescence amplitude (A) and the observed rate, k_{obs} , for each exponential phase where $F(0)$ is the initial fluorescence intensity, and t , time:

$$F(t) = \sum A_n [1 - \exp(-k_{obs,n} t)] + F(0) \quad (\text{Eq. 2.2})$$

Two control experiments were performed to ensure that the fluorescence changes are from a Med15-TAD interaction: 50 nM 5'-BODIPY®FL-labeled, double-stranded oligonucleotides pre-complexed to 200 nM dimeric MBP-tagged Gal4(1-100) was mixed with Med15 (0.375 µM after mixing) at 25 °C, and 5'-BODIPY®FL-DNA•Gal4(1-100)-Gal4(840-881) activator complex (5 nM after mixing) was mixed with dimeric GST (0.25 µM after mixing) at 25 °C. No time-dependent fluorescence enhancement was observed in either experiment (Figure 2.11).

Analysis of the time courses was performed using Kintek software, and the reported errors are the asymptotic standard errors. The dependence of the observed rates on Med15 concentration was analyzed using GraphPad Prism 4.0 software. Each kinetic constant reported in Table 1 is an average of two independent experiments with propagation of the experimental error.

Calculation of the Microscopic Kinetic Rate and Equilibrium Constants for the Two-step

Binding Models

Approximate solutions for $k_{obs,1}$ and $k_{obs,2}^{max}$ in terms of the microscopic rate and equilibrium constants can be obtained from Equations 2.3 & 2.4 for Scheme A and Equations 2.9 & 2.10 for Scheme B, respectively.⁸² These equations are approximations from the full solutions derived for these mechanisms and previously published.⁸¹

$$k_{obs,1} \approx k_1[Med15] + k_{-1} + k_2 + k_{-2} \quad (\text{Eq. 2.3})$$

$$k_{obs,2} \approx \frac{k_1[Med15](k_{-2}+k_2)+k_{-1}k_{-2}}{k_1[Med15]+k_{-1}+k_2+k_{-2}} \quad (\text{Eq. 2.4})$$

$$k_{obs,2}^{max} \approx k_2 + k_{-2} \quad (\text{Eq. 2.5})$$

$$K_{d,app} = 1/[K_1(1 + K_2)] \quad (\text{Eq. 2.6})$$

$$K_1 = k_1/k_{-1} \quad (\text{Eq. 2.7})$$

Equations used to calculate the microscopic rate and equilibrium constants, according to the binding model presented in Scheme B, are as follows:

$$k_{obs,3} \approx k_3 + k_{-3} + k_4[Med15] + k_{-4} \quad (\text{Eq. 2.9})$$

$$k_{obs,4} \approx \frac{k_3(k_{-4}+k_4[Med15])+k_{-3}k_{-4}}{k_3+k_{-3}+k_4[Med15]+k_{-4}} \quad (\text{Eq. 2.10})$$

$$K_{d,app} = (1 + K_3)/K_3K_4 \quad (\text{Eq. 2.11})$$

$$k_{obs,4}^{max} \approx k_3 \quad (\text{Eq. 2.12})$$

$$K_3 = k_3/k_{-3} \quad (\text{Eq. 2.13})$$

$$K_4 = k_4/k_{-4} \quad (\text{Eq. 2.14})$$

As described in the text, the experimentally obtained values for $K_{d,app}$, k_{on} , $k_{obs,1}^{(y\text{-intercept})}$ (summarized in Table 2.1) were used to calculate the microscopic rate constants for the fast phase and equilibrium constants for both phases, according to Equations 2.3-2.7 for Scheme A

and Equations 2.9-2.14 for Scheme B. The value of $k_{\text{obs},2}^{\text{max}}$ used in these calculations was extrapolated to saturating Med15 concentration, as outlined in the subsequent section. The calculated values for the microscopic constants for Scheme A (k_1 , k_{-1} , K_1 and K_2) are summarized in Table 2.2. The calculated values for the microscopic constants for Scheme B (k_4 , k_{-4} , K_3 and K_4) are summarized in Table 2.3.

KinTek Global Kinetic Explorer Analysis

Experimental fluorescence traces were fit with KinTek Global Kinetic Explorer¹⁰¹ based on the Scheme A and Scheme B models (Figure 2.13) including parameters describing the fluorescence change in each step. The value of k_1 (k_4 in the case of Scheme B) was calculated from experimental data using Equation 2.3 (Equation 2.9 in the case of Scheme B) and served as a fixed constraint in the simulated fitting process. Values for Med15 concentrations and BODIPY-DNA•activator complex concentrations were set as the experimental values.

Due to the aggregation propensity of Med15 at higher concentrations, stopped-flow experiments were not carried out at saturating Med15 concentrations.

To estimate the value of $k_{\text{obs},2}^{\text{max}}$, kinetic traces of the concentrations of the intermediates were simulated (KinTek Global Explorer) at

varying Med15 concentrations and $k_{\text{obs},2}^{\text{max}}$ values. The value of $k_{\text{obs},2}^{\text{max}}$ that gave the best

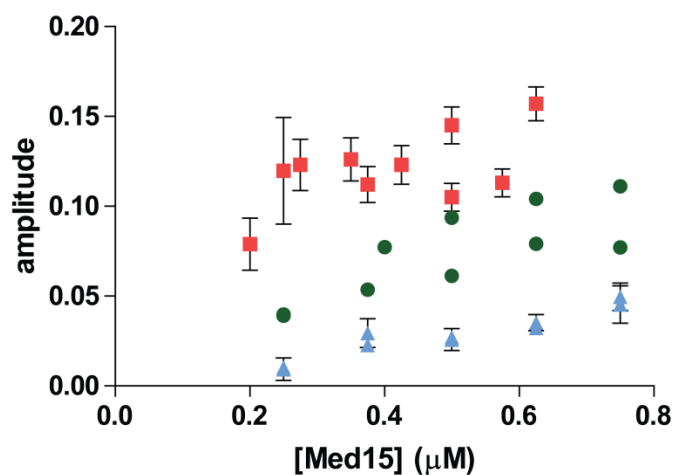


Figure 2.17 Plot of the amplitudes of the slow phase of the activators VP16 (■), Gal4 (●) and Gcn4 (▲) against the concentration of Med15 (0.125 - 0.75 μM).

description of the measured data is listed in Table 1 and used in subsequent calculations. In each case, the value of $k_{\text{obs},2}^{\text{max}}$ increases $\leq 50\%$ from the measured value at the highest [Med15]. The Med15 concentration dependence of $k_{\text{obs},2}$ is also consistent with the $K_{1/2}$ value for the [Med15]-dependence of the second phase amplitude (Figure 2.17).

Values for k_2 and k_{-2} were also calculated using Equations 2.5 and 2.8:

$$K_2 = k_2/k_{-2} \quad (\text{Eq. 2.8})$$

The calculated rate constants k_1 and k_{-1} in Table 2.2, along with calculated values for k_2 and k_{-2} were used to simulate experimental traces based on the Scheme A model (Figure 2.13), allowing the change in fluorescence to vary for each step.

H. References

1. Majmudar, C. Y. *et al.* A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew. Chem., Int. Ed.* **48**, 7021–4 (2009).
2. 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).
3. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
4. Chen, X. *et al.* Gene expression patterns in human liver cancers. *Mol. Cell. Biol.* **13**, 1929–1939 (2002).
5. Darnell Jr., J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749 (2002).
6. Lee, L. W. & Mapp, A. K. Transcriptional switches: chemical approaches to gene regulation. *J. Biol. Chem.* **285**, 11033–11038 (2010).
7. Mapp, A. K. K. & Ansari, A. Z. Z. A TAD further: exogenous control of gene activation. *ACS Chem. Biol.* **2**, 62–75 (2007).
8. Ptashne, M. & Gann, A. *Genes & Signals*. (Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., 2002).
9. Triezenberg, S. J. Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5**, 190–196 (1995).
10. Ptashne, M. How eukaryotic transcriptional activators work. *Nature* **335**, 683–689 (1988).
11. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–77 (1997).

12. Triezenberg, S. J., LaMarco, K. L. & McKnight, S. L. Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes Dev.* **2**, 730–742 (1988).
13. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2**, 718–729 (1988).
14. Jonker, H. R. a H. R. A. R. A., Wechselberger, R. W. W. R. W., Boelens, R., Folkers, G. E. G. E. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44**, 827–839 (2005).
15. Krishnamurthy, M. *et al.* Caught in the act: covalent cross-linking captures activator-coactivator interactions in vivo. *ACS Chem. Biol.* **6**, 1321–6 (2011).
16. Sullivan, S. Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus [published erratum appears in *Nucleic Acids Res* 1998 Dec 1;26(23):537-8]. *Nucleic Acids Res.* **26**, 4487–4496 (1998).
17. Shen, F., Triezenberg, S. J., Hensley, P., Porter, D. & Knutson, J. R. Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. *J. Biol. Chem.* **271**, 4827–4837 (1996).
18. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277**, 1310–1313 (1997).
19. Grossmann, J. G., Sharff, A. J., O’Hare, P. & Luisi, B. Molecular shapes of transcription factors TFIIB and VP16 in solution: implications for recognition. *Biochemistry* **40**, 6267–6274 (2001).
20. Langlois, C. *et al.* NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: structural similarities between VP16 and p53. *J. Am. Chem. Soc.* **130**, 10596–10604 (2008).
21. Fischer, J. A., Giniger, E., Maniatis, T. & Ptashne, M. GAL4 activates transcription in *Drosophila*. *Nature* **332**, 853–856 (1988).
22. Kakidani, H. & Ptashne, M. GAL4 activates gene expression in mammalian cells. *Cell* **52**, 161–167 (1988).
23. Giniger, E., Varnum, S. M. & Ptashne, M. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* **40**, 767–774 (1985).
24. Hong, M. *et al.* Structural basis for dimerization in DNA recognition by Gal4. *Structure* **16**, 1019–1026 (2008).
25. Traven, A., Jelacic, B. & Sopta, M. Yeast Gal4: a transcriptional paradigm revisited. *EMBO Rep.* **7**, 496–9 (2006).
26. Chang, C. *et al.* The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. *J. Biol. Chem.* **276**, 30956–30963 (2001).
27. Jeong, C. J. *et al.* Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **40**, 9421–9427 (2001).
28. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847–853 (1987).
29. Ma, J. & Ptashne, M. The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. *Cell* **50**, 137–142 (1987).
30. Ansari, A. Z. *et al.* Transcriptional activating regions target a cyclin-dependent kinase. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14706–14709 (2002).

31. Mondal, K. *et al.* Design and isolation of temperature-sensitive mutants of Gal4 in yeast and Drosophila. *Journal of molecular biology* **370**, 939–50 (2007).
32. Van Hoy, M., Leuther, K. K., Kodadek, T. & Johnston, S. A. The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets. *Cell* **72**, 587–594 (1993).
33. Kumar, P. R., Yu, Y., Sternglanz, R., Johnston, S. A. & Joshua-Tor, L. NADP regulates the yeast GAL induction system. *Science* **319**, 1090 (2008).
34. Hinnebusch, A. G. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu. Rev. Microbiol.* **59**, 407–450 (2005).
35. Hope, I. A. & Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885–894 (1986).
36. 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).
37. Ciani, B. *et al.* Molecular basis of coiled-coil oligomerization-state specificity. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19850–5 (2010).
38. Keller, W., König, P. & Richmond, T. J. Crystal structure of a bZIP/DNA complex at 2.2 Å: determinants of DNA specific recognition. *J. Mol. Biol.* **254**, 657–67 (1995).
39. König, P. & Richmond, T. J. The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. *J. Mol. Biol.* **233**, 139–54 (1993).
40. Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α Helices: Crystal structure of the protein-DNA complex. *Cell* **71**, 1223–1237 (1992).
41. Drysdale, C. *et al.* The transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids. *Mol. Cell. Biol.* **15**, 1220–1233 (1995).
42. Herbig, E. *et al.* Mechanism of Mediator recruitment by tandem Gcn4 activation domains and three Gal11 activator-binding domains. *Mol. Cell. Biol.* **30**, 2376–90 (2010).
43. Fishburn, J., Mohibullah, N. & Hahn, S. Function of a eukaryotic transcription activator during the transcription cycle. *Mol. Cell* **18**, 369–378 (2005).
44. Reeves, W. M. & Hahn, S. Targets of the Gal4 transcription activator in functional transcription complexes. *Mol. Cell. Biol.* **25**, 9092–9102 (2005).
45. 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).
46. Jedidi, I. *et al.* Activator Gcn4 employs multiple segments of Med15/Gal11, including the KIX domain, to recruit mediator to target genes in vivo. *J. Biol. Chem.* **285**, 2438–2455 (2010).
47. Park, J. M. *et al.* In vivo requirement of activator-specific binding targets of mediator. *Mol. Cell. Biol.* **20**, 8709–8719 (2000).
48. Zhang, F., Sumibcay, L., Hinnebusch, A. G. & Swanson, M. J. A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol. Cell. Biol.* **24**, 6871–6886 (2004).
49. Myers, L. C. & Kornberg, R. D. Mediator of transcriptional regulation. *Annu. Rev. Biochem.* **69**, 729–49 (2000).

50. Davis, J. A., Takagi, Y., Kornberg, R. D. & Asturias, F. A. Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol. Cell* **10**, 409–415 (2002).
51. Myers, L. C. *et al.* The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**, 45–54 (1998).
52. Dotson, M. R. *et al.* Structural organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14307–10 (2000).
53. Ansari, S. A. *et al.* Distinct role of Mediator tail module in regulation of SAGA-dependent, TATA-containing genes in yeast. *EMBO J.* **31**, 44–57 (2012).
54. Lee, Y. C., Park, J. M., Min, S., Han, S. J. & Kim, Y. J. An activator binding module of yeast RNA polymerase II holoenzyme. *Mol. Cell. Biol.* **19**, 2967–76 (1999).
55. Thakur, J. K. *et al.* Mediator subunit Gal11p/MED15 is required for fatty acid-dependent gene activation by yeast transcription factor Oaf1p. *J. Biol. Chem.* **284**, 4422–8 (2009).
56. Thakur, J. K. *et al.* A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **452**, 604–609 (2008).
57. Novatchkova, M. & Eisenhaber, F. Linking transcriptional mediators via the GACKIX domain super family. *Curr. Biol.* **14**, R54–R55 (2004).
58. Chadick, J. Z. & Asturias, F. J. Structure of eukaryotic Mediator complexes. *Trends Biochem. Sci.* **30**, 264–271 (2005).
59. Wu, Y., Reece, R. J. & Ptashne, M. Quantitation of putative activator-target affinities predicts transcriptional activating potentials. *EMBO J* **15**, 3951–3963 (1996).
60. 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).
61. Wu, Z. *et al.* Targeting the transcriptional machinery with unique artificial transcriptional activators. *J. Am. Chem. Soc.* **125**, 12390–12391 (2003).
62. Volkman, H. M., Rutledge, S. E. & Schepartz, A. Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain. *J. Am. Chem. Soc.* **127**, 4649–58 (2005).
63. Paal, K., Baeuerle, P. A. & Schmitz, M. L. Basal transcription factors TBP and TFIIB and the viral coactivator E1A 13S bind with distinct affinities and kinetics to the transactivation domain of NF-kappaB p65. *Nucleic Acids Res.* **25**, 1050–1055 (1997).
64. Hermann, S., Berndt, K. D. & Wright, A. P. How transcriptional activators bind target proteins. *J. Biol. Chem.* **276**, 40127–40132 (2001).
65. Ferreira, M. E. *et al.* Mechanism of transcription factor recruitment by acidic activators. *J. Biol. Chem.* **280**, 21779–21784 (2005).
66. Fagan, R. L., Jensen, K. F., Björnberg, O. & Palfey, B. A. Mechanism of flavin reduction in the class 1A dihydroorotate dehydrogenase from *Lactococcus lactis*. *Biochemistry* **46**, 4028–36 (2007).
67. 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**, 34316–24 (2012).
68. Rogers, J. M., Steward, A. & Clarke, J. Folding and Binding of an Intrinsically Disordered Protein: Fast, but Not “Diffusion-Limited”. *J. Am. Chem. Soc.* **135**, 1415–1422 (2013).
69. Chen, C. *et al.* Study of pH-induced folding and unfolding kinetics of the DNA i-motif by stopped-flow circular dichroism. *Langmuir* **28**, 17743–8 (2012).

70. Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev. (Washington, DC, U. S.)* **109**, 839–60 (2009).
71. Nieba, L., Krebber, A. & Plückthun, A. Competition BIAcore for Measuring True Affinities: Large Differences from Values Determined from Binding Kinetics. *Anal. Biochem.* **234**, 155–165 (1996).
72. Piehler, J. & Schreiber, G. Biophysical analysis of the interaction of human ifnar2 expressed in E. coli with IFN α 2. *J. Mol. Biol.* **289**, 57–67 (1999).
73. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563–564 (1988).
74. Hidalgo, P. *et al.* Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes Dev.* **15**, 1007–1020 (2001).
75. 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).
76. Archer, C. T. & Kodadek, T. The hydrophobic patch of ubiquitin is required to protect transactivator-promoter complexes from destabilization by the proteasomal ATPases. *Nucleic Acids Res* **38**, 789–796 (2010).
77. Ferdous, A., Gonzalez, F., Sun, L., Kodadek, T. & Johnston, S. A. The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. *Mol. Cell* **7**, 981–991 (2001).
78. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**, 235–239 (2005).
79. Kim, Y. J. *et al.* A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599–608 (1994).
80. Northrup, S. H. & Erickson, H. P. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 3338–3342 (1992).
81. Johnson, K. A. Rapid kinetic analysis of mechanochemical adenosinetriphosphatases. *Methods Enzymol.* **134**, 677–705 (1986).
82. Johnson, K. A. Transient-state kinetic analysis of enzyme reaction pathways. *The Enzymes* 1–61 (1992).
83. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
84. Hsieh, J. & Fierke, C. A. Conformational change in the Bacillus subtilis RNase P holoenzyme--pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA (New York, N.Y.)* **15**, 1565–77 (2009).
85. Rea, A. M., Thurston, V. & Searle, M. S. Mechanism of ligand-induced folding of a natively unfolded helixless variant of rabbit I-BABP. *Biochemistry* **48**, 7556–64 (2009).
86. Ansari, A. Z., Reece, R. J. & Ptashne, M. A transcriptional activating region with two contrasting modes of protein interaction. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13543–13548 (1998).
87. Donaldson, L. & Capone, J. P. Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1. *J. Biol. Chem.* **267**, 1411–1414 (1992).

88. Thoden, J. B., Ryan, L. A., Reece, R. J. & Holden, H. M. The Interaction between an Acidic Transcriptional Activator and Its Inhibitor. *J. Biol. Chem.* **283**, 30266–30272 (2008).
89. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5271–6 (2002).
90. Wells, M. *et al.* Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5762–5767 (2008).
91. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–953 (1996).
92. Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).
93. Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
94. Kim, D. H. *et al.* Multiple hTAF(II)31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16. *BMB Rep.* **42**, 411–417 (2009).
95. Demchenko, A. P. Recognition between flexible protein molecules: induced and assisted folding. *J. Mol. Recognit.* **14**, 42–61 (2001).
96. Lum, J. K., Majmudar, C. Y., Ansari, A. Z. & Mapp, A. K. Converting inactive peptides into potent transcriptional activators. *ACS Chem. Biol.* **1**, 639–643 (2006).
97. Majmudar, C. Y., Lum, J. K., Prasov, L. & Mapp, A. K. Functional specificity of artificial transcriptional activators. *Chem. Biol.* **12**, 313–321 (2005).
98. Reece, R. J., Rickles, R. J. & Ptashne, M. Overproduction and single-step purification of GAL4 fusion proteins from *Escherichia coli*. *Gene* **126**, 105–107 (1993).
99. Donnelly, M. I. *et al.* An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expression Purif.* **47**, 446–454 (2006).
100. Archer, C. T., Delahodde, A., Gonzalez, F., Johnston, S. A. & Kodadek, T. Activation Domain-dependent Monoubiquitylation of Gal4 Protein Is Essential for Promoter Binding in Vivo. *J. Biol. Chem.* **283**, 12614–12623 (2008).
101. 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).

Chapter 3 Structural Characterization of Transcriptional Coactivator KIX in Complex with Binding Partners⁸

A. Introduction

Transcriptional coactivators are among the most conformationally malleable of proteins and contain binding surfaces that undergo rapid remodeling as complexes are formed with their cognate ligands.^{1,2} This plasticity is essential to their function, enabling recognition of an often diverse array of transcriptional activator sequences.^{3,4} Perhaps the best-studied example of this is the KIX domain of the coactivator CBP/p300, a small (90 amino acid) domain that is known to interact with >10 distinct amphipathic sequences at two distinct binding sites (Figure 3.1) in order to stimulate transcription at hundreds of genes,⁵⁻⁹ including those regulating hematopoiesis, memory formation and the inflammatory response.¹⁰⁻¹² Not surprisingly, the malleability of this class of proteins renders them especially intractable to crystallographic characterization, either alone or in complex with their binding partners. In the case of the KIX domain, there are no crystal structures of either free protein or any complexed form. Here we demonstrate that a covalently linked small-molecule ligand of this conformationally dynamic protein enables, for the first time, a high resolution snapshot of the coactivator interacting with a ligand. This first crystal structure of KIX provides important insight to the side chain orientations of this domain

⁸ The contents of this chapter are adapted and reproduced from a published article: Wang, N., Majmudar, C. Y., Pomerantz, W. C., Gagnon, J. K., Sadowsky, J. D., Meagher, J. L., Johnson, T. K., Stuckey, J. A., Brooks, C. L., Wells, J. A., and Mapp, A. K. (2013) Ordering a dynamic protein via a small-molecule stabilizer., *J. Am. Chem. Soc.*, 135, 3363-6.⁷⁷

in the context of ligand recognition, particularly with regard to small molecules. Furthermore, these results show that the ligand discovery strategy of Tethering¹³⁻¹⁶ can be expanded to targeting conformationally dynamic proteins and enable their structural characterization.



Figure 3.1 The KIX domain is in the N-terminal region of CBP/p300. KIX interacts with >10 amphipathic transcriptional activators using two distinct sites.⁵⁻⁹ MLL, HBZ and c-Jun target a smaller, deeper site while the activation domains of c-Myb and CREB (pKID) utilize a second, broader site.

B. Background

B.1 The KIX domain

The coactivators CBP and p300 are homologous and exist in a wide range of organisms from plants¹⁷ and *C. elegans*¹⁸ to rats and humans.¹⁹ The CBP/p300 proteins are transcription hubs that interact with numerous activators using multiple distinct well folded domains.²⁰ The GACKIX domain, also known simply as the KIX domain, is one of these domains. This is a 90 residue domain that consists of a three helix bundle along with two 3_{10} helices.⁵ The KIX domain is known to interact with more than ten distinct transcriptional activators²¹ via two different binding sites as shown in Figure 3.2. The complexes formed when KIX is bound to transcriptional activation domains (TADs) at one or both of its binding sites have been extensively studied by solution NMR:^{1,5,22-24} the helices α_3 and α_2 form a deep, narrow groove that interacts with transcriptional activators such as MLL, c-Jun and HBZ whereas on the

opposite side of the protein the helices α_3 and α_1 form a shallower, broader groove that binds to transcriptional activators such as pKID of CREB and c-Myb. There are also transcriptional activators that have been found to bind to either one of the two sites. For example, there are two TADs in FOXO3a, CR2C and CR3. Two different NMR structures of these tandem TADs binding to KIX exist: one with CR2C binding to the MLL site and CR3 binding to the pKID site and the other vice versa.²²

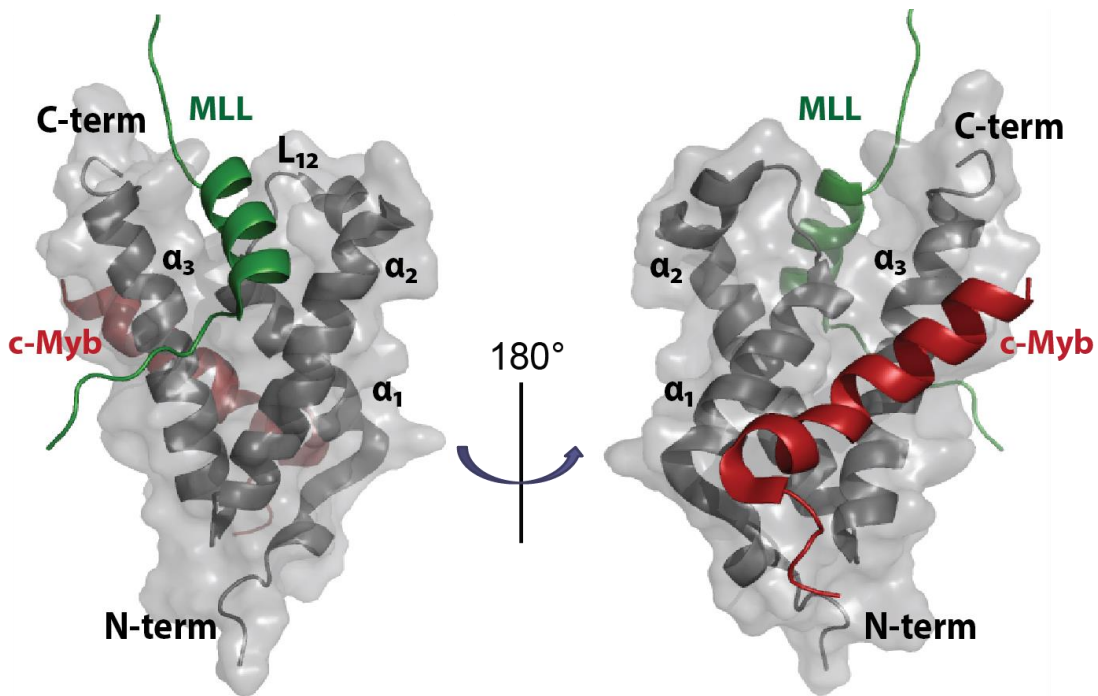


Figure 3.2 Structure of the KIX domain of CBP in complex with two transcriptional activation domains, MLL and c-Myb at two distinct sites, α_1 , α_2 and α_3 label the three helices in KIX and L_{12} depicts the loop region between α_1 and α_2 . Based on PDB structure 2AGH

Binding cooperativity is observed between the two sites when certain activators are involved.^{5,6,8} For example when either pKID or c-Myb is pre-bound to KIX, the binary complex shows a 2-fold increase in binding affinity to MLL. Similarly when either MLL or HBZ is pre-bound to KIX, the binary complex shows a 2 and 6 fold increase in binding affinity to c-Myb, pKID will also bind to the KIX: MLL binary complex with a 2-fold increase affinity. Various

NMR and computational studies suggest that this cooperativity is due to an allosteric network interconnecting these two sites.^{25,26} Ile 611, Ile 657 and Ile 660 serve as a bridge connecting Phe 612 at the MLL binding site with a series of residues on the pKID/c-Myb binding site, including Tyr 650, His 651, Ala 654 and Tyr 658 (Figure 3.3). This domain has also been identified in other proteins, such as transcriptional coactivators Gal11 (Med15), Arc105²⁷ and a Malarial erythrocyte-binding-antigen (EBA-175 region VI).²⁸

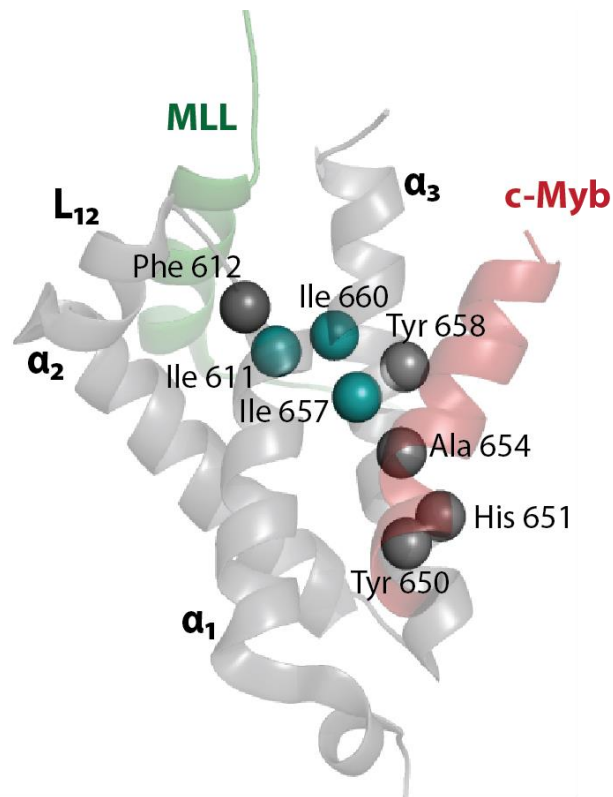


Figure 3.3 The allosteric network: Phe 612 at the MLL site connects with Tyr 658, Ala 654, His 651 and Tyr 650 at the pKID/c-Myb via Ile 611, Ile 660 and Ile 657. Adapted from PDB 2AGH.

There are several structures of the KIX domain in the Protein Data Base, and this information has proved invaluable in understanding the interaction between KIX and its activator binding partners as well as designing artificial ligands to target KIX. However, aside from the KIX-like-domain of EBA-175 region VI, all the other structures are obtained by NMR in solution. We and others have attempted to crystallize the KIX domain (Dr. Malathy Krishnamurthy, data unpublished) under various screening conditions but with no success. Moreover, while both the NMR structures of Gal11(Med15)²⁹ and Arc105 KIX³⁰ domains are that of the free, unbound protein, all the structures of the KIX domain from CBP/p300 are solution structures of the

protein in complex with a peptide sequence of a transcriptional activator binding partner. Not only is there no structure of the free KIX domain of CBP, there is also no structure of the KIX domain in complex with a ligand solely at the MLL binding site. An atomic resolution structure of the KIX domain of CBP with no ligand at the pKID binding site would increase our understanding of this coactivator and its interactions with TADs.

The lack of crystal structures and ligand-free structures of the CBP KIX domain might be

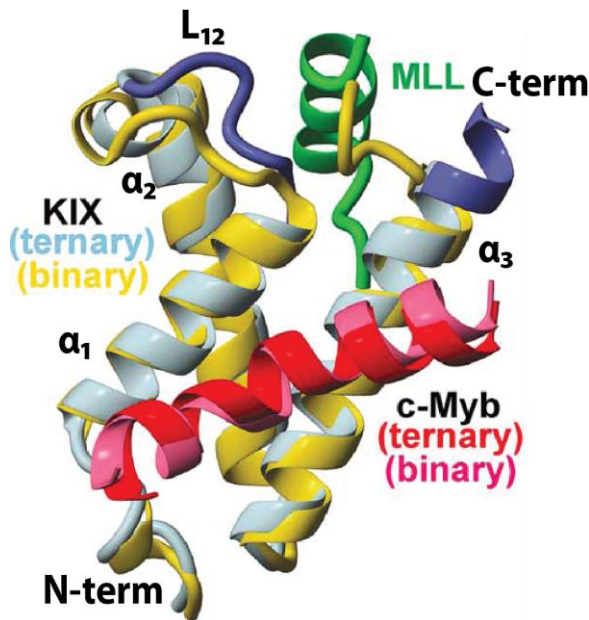


Figure 3.4 Superposition of the KIX: c-Myb: MLL ternary complex (KIX: light blue, c-Myb: red, MLL: green) with the KIX: c-Myb binary complex (KIX: yellow, c-Myb: pink), showing the movement of loop L₁₂. Modified from Guzman, et al. 2006.⁵

allosterically communicate between the two binding sites.

As mentioned earlier, abnormalities in the interaction of KIX with its various TADs have been linked to a multitude of diseases. For example, atherosclerosis has been linked to transcriptional activators such as cAMP response element-binding protein (CREB) and c-Myb;^{31,32} the generation of hematopoietic stem cells is dependent on the transcriptional activator mixed

attributed to the dynamic nature of this protein. Previous studies have shown both by protein solution NMR and molecular dynamics simulation methods that there are movements in KIX when transitioning between the binary (KIX:MLL or KIX:c-Myb) and ternary (MLL:KIX:c-Myb) complex, as shown in Figure 3.4.^{5,26}

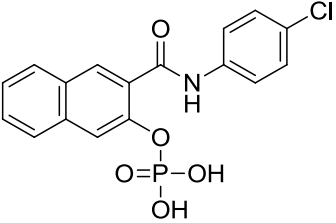
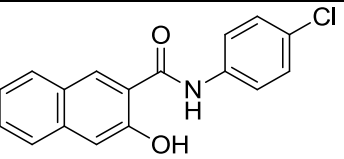
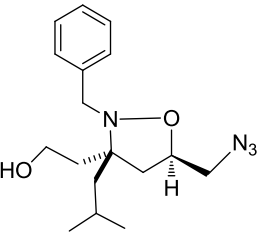
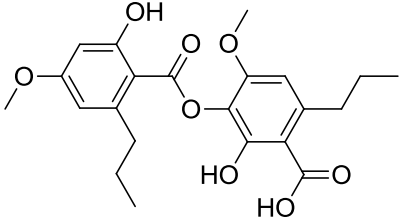
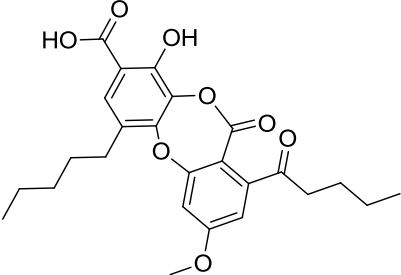
Notably, the L₁₂ loop region and the C-term region of helix α_3 undergo a significant conformational shift. It is suggested that this movement enables KIX to

lineage leukemia (MLL);³³ cardiac hypertrophy has been linked to the master transcriptional coactivator CREB binding protein (CBP);³⁴ and adult T-cell leukemia is linked to HTLV-1 basic leucine zipper factor (HBZ).⁸ Thus, it is very important to develop small molecule modulators that target this interaction to act as therapeutics or mechanistic probes. However, there are very few small molecules that have been identified as KIX binders (see section B.2), and of the limited small molecules that have been found, low affinity and poor specificity remains a major road block for further developing them as useful modulators.³⁵⁻³⁷ Among the various obstacles in discovering small molecule modulators of protein-protein interactions, the lack of a high resolution and well-defined crystal structure of KIX stands out as an impediment in further progress on this front as well as for understanding the underlying mechanism of KIX-TAD interactions.

B.2 Identified small molecule ligands of the CBP/p300 KIX domain and limitations

A brief summary of the small molecules found to target the CBP/p300 KIX domain is listed below in Table 3.1. In short, despite the important role the KIX domain plays in various pathways across organisms, discoveries of small molecules that target this domain are few and far between. It is crucial to develop more small molecules that inhibit KIX in order to probe the mechanism of action of this allosterically regulated transcriptional coactivator to both further develop therapeutic strategies and to elucidate the general mechanism of transcriptional protein-protein interactions from this prototypical system. New screening approaches must be used to find small molecule binding partners of KIX. Also appropriate small molecule ligands would possibly serve as a stabilizer for KIX conformation, hence enable the crystallization of KIX and provide us with more details of the mechanism.

Table 3.1 Summary of small molecules known to target the KIX domain of CBP/p300

Compound	Method of Discovery	Target Site (K_i or K_D)	Citations
 <p>Naphthol AS-E phosphate (KG-501)</p>	2D ^1H - ^{15}N HSQC NMR based screen	pKID/c-Myb site ($K_i \approx 90 \mu\text{M}$, $K_D = 115 \mu\text{M}$)	Best, J. L. et al (2004) ³⁶ Pomerantz, W. C et al (2012) ³⁸
 <p>Naphthol AS-E</p>	Modification of naphthol AS-E phosphate	pKID/c-Myb site ($\text{IC}_{50} = 2.9 \mu\text{M}$, $K_D = 8.6 \mu\text{M}$)	Li, B. X. et al (2009) ³⁹ Li, B. X. et al (2012) ³⁷
 <p>iTAD1</p>	In-cell assays 2D ^1H - ^{15}N HSQC NMR perturbation experiments	MLL/c-Jun site ($K_D = 38 \mu\text{M}$)	Buhrlage, S. J. et al (2009) ⁴⁰
 <p>Sekikaic Acid</p>	Fluorescence polarization-based high throughput screen of natural product extracts	pKID/c-Myb site ($\text{IC}_{50} = 17 \mu\text{M}$) MLL/c-Jun site ($\text{IC}_{50} = 34 \mu\text{M}$)	Majmudar, C. Y., Højfeldt, J. W. et al (2012) ³⁵
 <p>Lobaric Acid</p>	Similar structure to sekikaic acid	pKID/c-Myb site ($\text{IC}_{50} = 25 \mu\text{M}$) MLL/c-Jun site ($\text{IC}_{50} = 17 \mu\text{M}$)	Majmudar, C. Y., Højfeldt, J. W. et al (2012) ³⁵

B.3 Tethering and other small molecule ligands stabilizing proteins

In recent years, the concept of using small molecules to target large protein surfaces at protein-protein interaction (PPI) interfaces, including transcriptional complexes, has become increasingly visited,⁴¹⁻⁴³ despite the many difficulties such as lack of defined deep binding pockets and in many cases even lack of defined structure and conformation.^{2,44} While the search for a small molecule inhibitor is the more common approach in these cases,^{36,45-50} there are also studies emerging where small molecules are demonstrated to stabilize the protein-protein interaction⁵¹ or even trapping an otherwise conformationally dynamic protein in a single more desirable conformation. Some recent examples of the latter include small molecules locally stabilizing two 10-20 residue long sequences of c-Myc forcing it into a conformation unable to form heterodimers with binding partner Max;⁵² an allosteric inhibitor of the hepatitis C virus NS3 protein that locks the helicase/protease protein into a “closed” conformation hence inhibiting its proteolytic activity;⁵³ and small molecules with a Michael acceptor moiety that covalently bind to cysteine sites in p53 and enhance the thermal stability of several oncogenic p53 mutants.⁵⁴

Identifying small molecule ligands targeting PPI surfaces pose many challenges including large binding surface and lack of structure; initial hits often bind to the target protein with comparatively low affinity, and are easy to be presented as false negative results in a screen. One way to avoid missing these low affinity ligands would be to use the Tethering technique designed by James Wells’ laboratory.¹⁶ This strategy utilizes a small library of disulfide-containing fragments that are able to form covalent bonds with native or engineered cysteines in the target protein. The covalently bound small molecules are then amenable to detection by techniques such as mass spectrometry. Varying the concentration of competing disulfide-forming

molecule such as beta-mercaptoethanol controls the stringency of the screen. Once these weak-binding fragments are identified as leads one can improve upon the binding affinity by synthesizing a small set of analogs.¹⁶ Also these can be converted into irreversible small molecule binders without extensive modifications (Mapp lab, data unpublished). Additionally, from a more interesting perspective, this technique can also serve as a probe for studying protein mechanisms of action. The Wells lab has been able to uncover a new common allosteric site on caspases using Tethering to trap disulfide inhibitory small molecules,^{14,15} and have also used this method to characterize an allosteric site on protein kinase PDK1 as well as demonstrate that targeting a single site on the protein kinase can result in both inhibitory and activating small molecule hits.¹³

C. Experimental Design

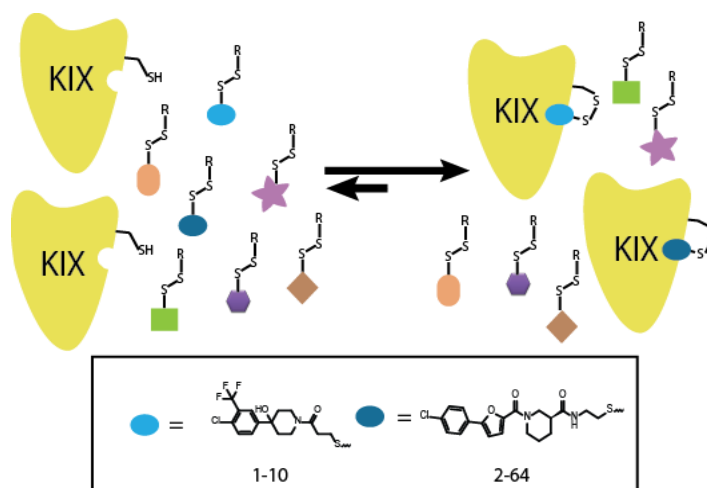


Figure 3.5 Schematic of the Tethering screen used to identify small molecule fragments (**1-10** and **2-64**) that form a disulfide bond with a cysteine introduced at position 664 (L664C) within KIX.

We screened for small molecules that interact with the KIX domain using the Tethering approach,¹⁶ a strategy that provides a mechanism for the rapid discovery of covalent ligands (Figure 3.5). Attention was focused on the binding site that is targeted by the transcriptional

activation domains of proteins such as the Mixed Lineage Leukemia (MLL) activator and c-Jun; the Tethering approach is a fragment discovery method and the smaller, deeper MLL/c-Jun binding site appeared the more targetable by low molecular weight compounds.^{35,38} Towards this end, a residue at the rim of the binding surface, L664, was mutated to a cysteine and the resulting KIX L664C mutant fully characterized. Small molecule fragments containing a disulfide motif were then screened for the ability to form a disulfide bond with KIX L664C in the presence of a competitor, β -mercaptoethanol. Two fragment ligands emerged from the screen with high Tethering efficiency to KIX L664C as quantified by DR (Dose Response)₅₀ values (2-8 μ M), fragments **1-10** and **2-64** (Figure 3.5).

D. Results

D. 1. Binding affinity of small molecule-tethered KIX L664C to TAD peptides.

To assess the effect of tethered **1-10** or **2-64** on the binding properties of KIX, fluorescent anisotropy binding assays were used to measure the binding affinity of wild type KIX, KIX L664C and fragment-tethered KIX L664C complexes to native transcriptional activator ligands that target the two different binding sites (Table 3.1 and Figure 3.6). Consistent with the screen design, the presence of **1-10** or **2-64** decreased MLL binding to KIX L664C by ~22 to 33-fold (Table 3.2). Also, while tethered **2-64** does not affect KIX's binding affinity for pKID, the transcriptional activation domain of CREB that interacts with the distal binding site,²³ KIX tethered to fragment **1-10** does exhibit attenuated binding to pKID (~2-fold). This suggests that **1-10** engages the amino acid side chains comprising the allosteric network connecting the two binding sites.^{6,25,26}

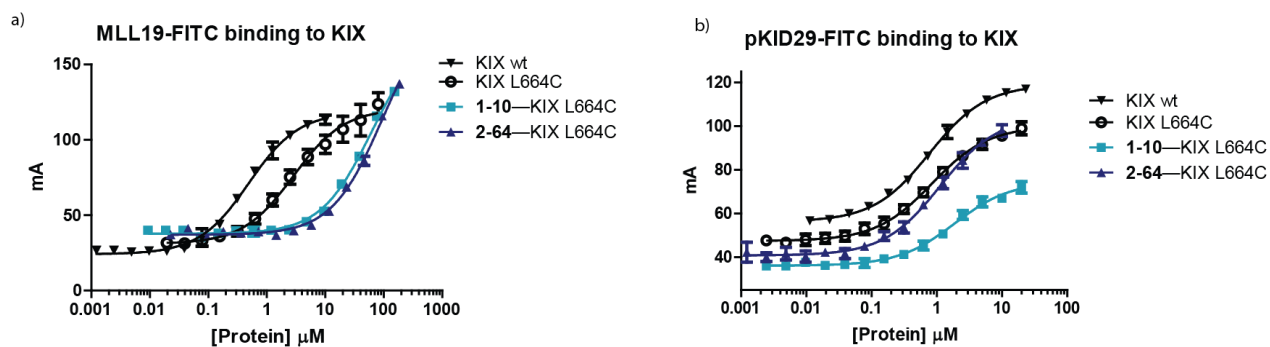


Figure 3.6 Direct binding affinity of fluorescently labeled peptides **a)** MLL19-FITC or **b)** pKID29-FITC to varying concentrations of KIX wt (black triangles), KIX L664C (black circles), **1-10**—KIX L664C (teal squares) and **2-64**—KIX L664C (blue triangles) was determined by fluorescence anisotropy. Each data point was taken in triplicate, error bars depict standard deviation (SD). See methods for more information.

Table 3.2 K_D s for KIX constructs interacting with fluorescein-labeled MLL and pKID peptides were determined by fluorescent anisotropy. Each K_D is a fitted result of experiments performed in triplicate with the indicated error (SD).

KIX construct	K_D (μM)	
	MLL	pKID
KIX wt	0.5 ± 0.1	0.9 ± 0.1
KIX L664C	2.8 ± 0.3	0.8 ± 0.1
1-10 —KIX L664C	62 ± 4	1.9 ± 0.2
2-64 —KIX L664C	94 ± 7	1.1 ± 0.1

D.2. Thermal, proteolytic and solvent stability of small molecule-tethered KIX L664C

The tethered fragments significantly altered the stability of the KIX domain. This was assessed for each of the fragment-protein pairs by measuring changes in CD-monitored thermal melting temperature, amide hydrogen-deuterium (H-D) exchange and thermolysin-mediated proteolysis (Figure 3.7).

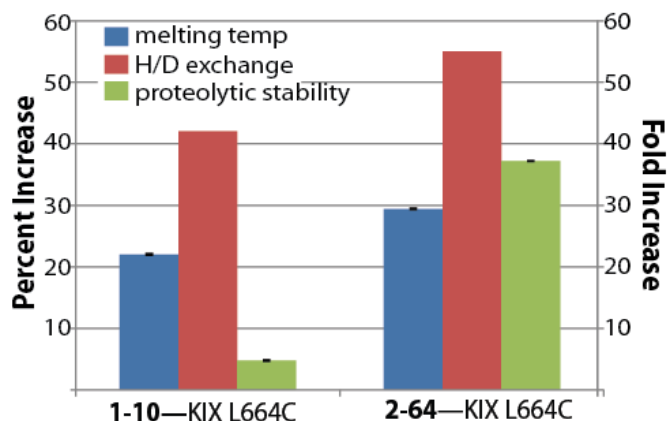


Figure 3.7 Bar graph depicts the percent increase in melting temperature (T_M) upon tethering to either **1-10** or **2-64** as monitored by circular dichroism (blue bars) and the percent of backbone amides protected from H-D exchange upon attachment of the small molecules (red bars). The green bars represent the fold-increase in resistance to thermolysin degradation of the KIX mutants when tethered to **1-10** and **2-64**. Data is normalized to KIX L664C.

To assess the effect of tethered small molecules to the global protein stability of KIX, we obtained the melting temperature of **1-10**—KIX L664C and **2-64**—KIX L664C by fitting the decrease of helical content (reflected by circular dichroism) over a temperature scan. As shown in Figure 3.8, the thermal stability was greatly increased in both complexes: the **1-10**—KIX L664C and the **2-64**—KIX L664C complexes exhibit a 15-18 °C ($\geq 20\%$) increase in melting temperature.

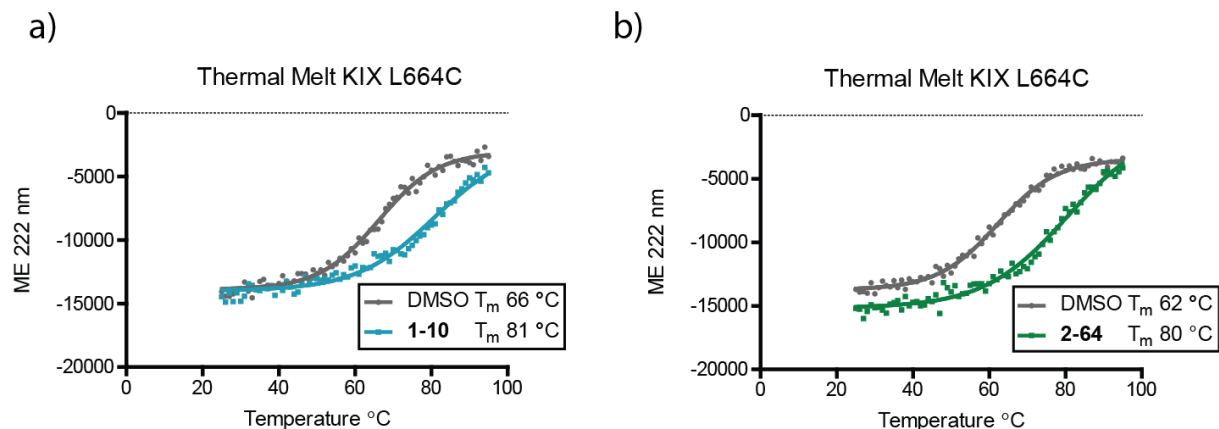


Figure 3.8 Melting temperature curve obtained by plotting molar ellipticity (obtained from circular dichroism measurements) at 222 nm during heating of the protein solutions at a rate of 1 degree/minute from 20-95°C.

To probe for the protection of backbone amide protons by small molecules **1-10** and **2-64**, we carried out hydrogen-deuterium (HD) exchange experiments with tethered and free KIX L664C.¹³ The mass of free KIX L664C shifted 29 Da upon exposure to D₂O for 1 min as monitored by mass spectrometry,¹³ whereas the mass shift was 17 and 13 Da when **1-10** and **2-64** were tethered to KIX L664C, respectively, showing that 40%-55% of the exchangeable amides were protected from H-D exchange compared to the free protein. (Table 3.3)

Table 3.3 Difference in number of amide protons protected from solvent.

	Δ m.w. (Da)	% amides protected
KIX L664C	29	-
1-10 —KIX L664C	16.7	42
2-64 —KIX L664C	13.1	55

To probe the energetic stability of KIX when tethered to small molecules, we carried out pulse proteolysis with a nonspecific protease thermolysin.^{13,55} The protease reaction was quenched at several time points and run on SDS-PAGE. The amount of protein remaining in each lane was quantified on the gel using ImageJ. The proteolytic stability (half-life) of the tethered complex increased 5-37 fold compared to the untethered protein, (for **1-10** for example: $T_{1/2}$ of 10 minutes versus 2.1 minutes) as shown in Figure 3.9.⁵⁵ This demonstrated that 1-10 and 2-64 stabilize KIX from degradation when covalently tethered to the protein. These findings encouraged pursuit of crystallization of fragment—KIX L664C complexes.

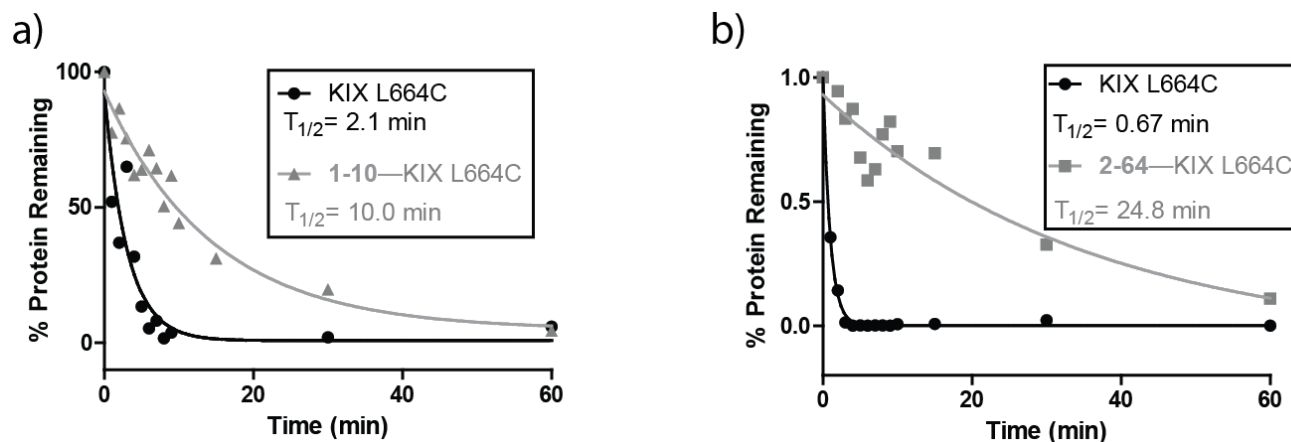


Figure 3.9 Proteolytic stability assays for free KIX L664C and KIX L664C covalently tethered to **a) 1-10** or **b) 2-64**. Half-life of the protein-small molecule complex was compared to free KIX L664C half-life in the same experiment using the same batch of thermolysin, avoiding systemic errors such as inconsistent enzyme activity. See methods for more information.

D.3. Crystal structure of **1-10**—KIX L664C

Of the various fragment—protein complexes and conditions that were screened (see Table 3.4), the best results were obtained with **1-10**—KIX L664C under the crystallizing condition of 1.8 M ammonium sulfate and 0.1 M Tris, pH 7.0 at 25° C, leading to crystals amenable for diffraction. However, only microcrystals of **2-64** tethered to KIX L664C were obtained and were of too poor quality to solve. Initially, molecular replacement strategies using the NMR structures of KIX bound to native transcriptional activation domains were used but did not lead to the **1-10**—KIX structure.^{5,23} Therefore, a selenomethionine-incorporated KIX L664C tethered to **1-10** were prepared and the X-ray structure was solved. Using these data, the structure of **1-10**—KIX L664C was determined to 2.0 Å resolution.

Table 3.4 Summary of fragment-protein complexes screened for crystallization.

Complex	1-10 --KIX L664C	2-64 --KIX L664C	1-10 --KIX L627C	2-64 --KIX L627C	1-10 --KIX L664C and pKID	1-10 --KIX L664C and c-Myb
Results	Crystal	None	None	None	None	None

As illustrated in Figure 3.10a, the small molecule **1-10** sits within the MLL/c-Jun binding site of KIX, and is oriented toward the core of the protein between helices α_3 (residues 646-664) and α_2 (residues 623-638). Notably, the aromatic ring of **1-10** is positioned relatively deep in a hydrophobic pocket lined by the side chains of Ile611, Leu628, Leu607, Val635, and Tyr631 (Figure 3.10c); Leu628 and Tyr631 have previously been shown to be key residues involved in KIX interacting with MLL.^{5,56} Tyr631 in particular, closely contacts the aromatic ring of **1-10** ($\sim 4\text{\AA}$), illustrated by the above 2σ deviation of the Tyr631 ϕ and ψ angles. This is consistent with data from a solution binding study of untethered 1-10 interacting with KIX containing ^{19}F -labeled Tyr631 that showed a dose-dependent change in the ^{19}F chemical shift.³⁸ Consistent with these data, chemical shift perturbation experiments with ^{15}N -labeled KIX L664C free and covalently tethered to **1-10** (Figure 3.10a) revealed significant changes in the backbone amide shifts of the residues lining the hydrophobic binding surface for **1-10** (Ile611, Leu628, Leu607, Val635 and Tyr631).

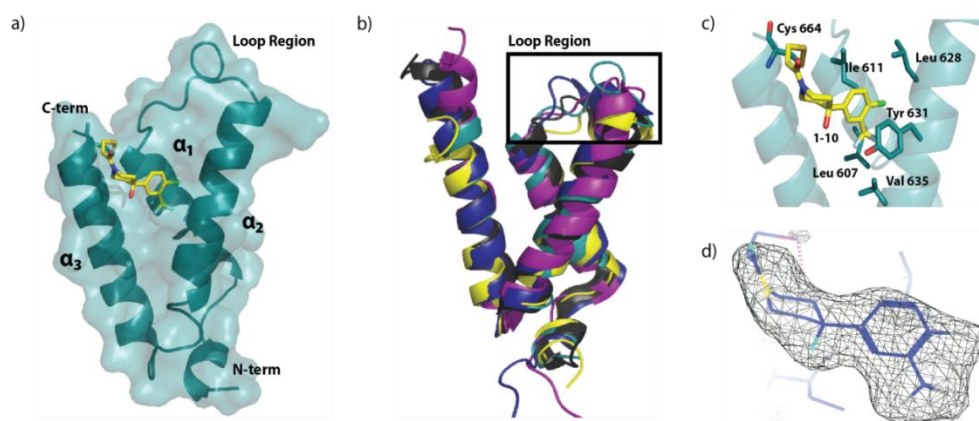


Figure 3.10 **a)** Refined crystal structure of KIX L664C covalently tethered to fragment 1-10. Refined resolution = 2.0 Å, $R_{\text{work}}/R_{\text{free}} = 0.2064/0.2329$. **b)** Crystal structure of KIX L664C tethered to 1-10 (teal) superimposed using Coot on the NMR solution structures of KIX in complex with cognate transcriptional activation domains: pKID (yellow, PDB ID 1KDX, R.M.S.D. =1.40 Å); with MLL and c-Myb (deep blue, PDB ID 2AGH, R.M.S.D. =1.80 Å); with PCET (purple, PDB ID 2KWF, R.M.S.D. =1.81 Å); and with FOXO3A (black, PDB ID 2LQH, R.M.S.D. =1.07 Å). **c)** Interactions between 1-10 (yellow) and residue side chains of KIX L664C (blue) at the binding surface. **d)** 3σ electron density map ($F_o - F_c$) of 1-10 illustrates the fit of the small molecule.

The prevailing structural model of the amphipathic class of activator-coactivator complexes is that the activator forms an amphipathic helix upon binding to the surface of the coactivator.^{1,57,58} Although only a limited suite of surfaces have been characterized, the available data suggest that the binding surfaces are often broad,^{2,59} making them particularly challenging to target with small molecules that have far less volume and surface area than the typical helix of a transcriptional activator.⁶⁰ Overlay of the **1-10**—KIX L664C structure with the averages of the previously reported NMR structures of KIX-ligand complexes^{5,22,24} yields R.M.S.D. values between 1.07-1.81 Å, demonstrating the overall similarities in the backbone structure. The exception to this similarity is in the loop region (residues 612-622) between helices α_1 and α_2 , which deviates significantly with R.M.S.D. values between 2.73-3.11 Å (Figure 3.10b). This difference is not surprising, as conformational changes in the loop regions are thought to be integral to the ability of KIX to accommodate diverse native ligands.^{5,22,23,26}

D.4 ¹H-¹⁵N-HSQC analysis of small molecule-tethered KIX conformational change when binding to ligands at the MLL site

Chemical shift perturbation experiments with ¹⁵N-labeled KIX L664C free and covalently tethered to **1-10** (Figure 3.11 and Figure 3.12) revealed significant changes in the backbone amide shifts of the residues lining the hydrophobic binding surface for **1-10** (Ile611, Leu628, Leu607, Val635 and Tyr631). In addition, the significant perturbations of residues Ile611 and Ile660 upon **1-10** tethering are consistent with the attenuated affinity of pKID for the distal binding site, as these residues are a critical part of the allosteric network that comprises the communication between the two binding sites within KIX. Upon comparison to chemical shift perturbation experiments performed on KIX L664C free and incubated with 2 molar equivalents

of MLL peptide, there are many similar regions of large chemical shift perturbations between that elicited by MLL and **1-10** (Figure 3.12 and Figure 3.13). An identical chemical shift perturbation experiment with 2-64 tethering to ^{15}N -labeled KIX L664C showed similar trends where residues not in direct contact with the small molecule were perturbed as well (Figure 3.14).

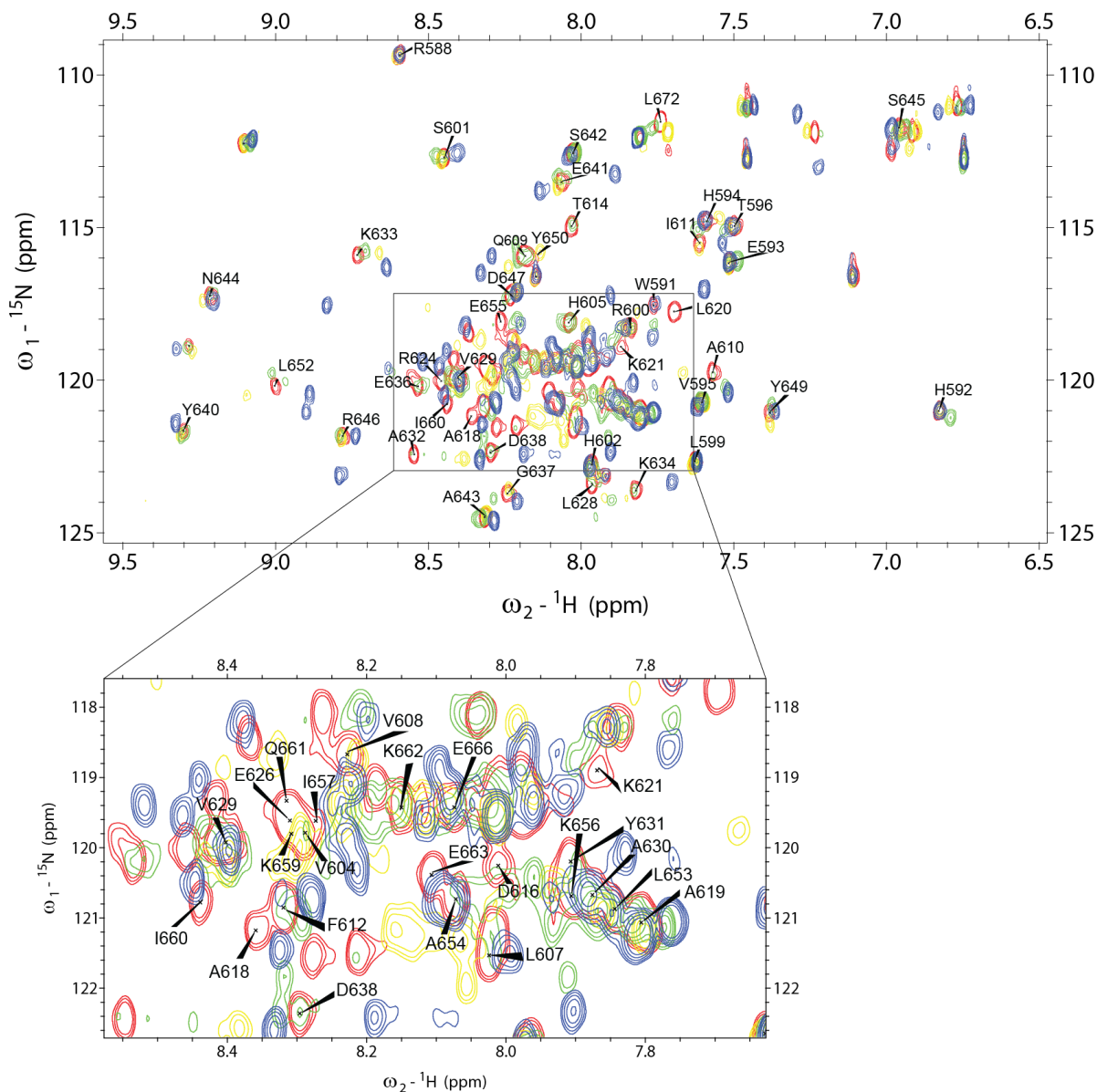
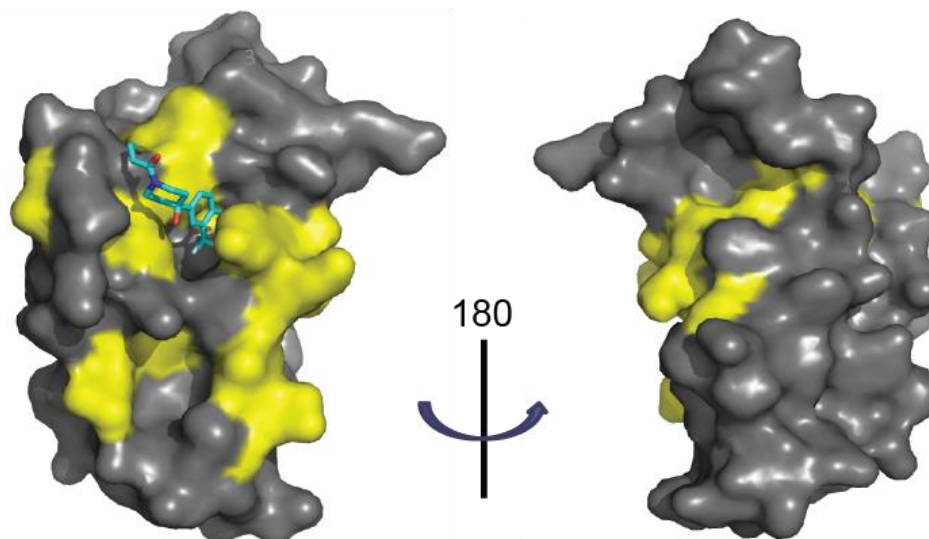
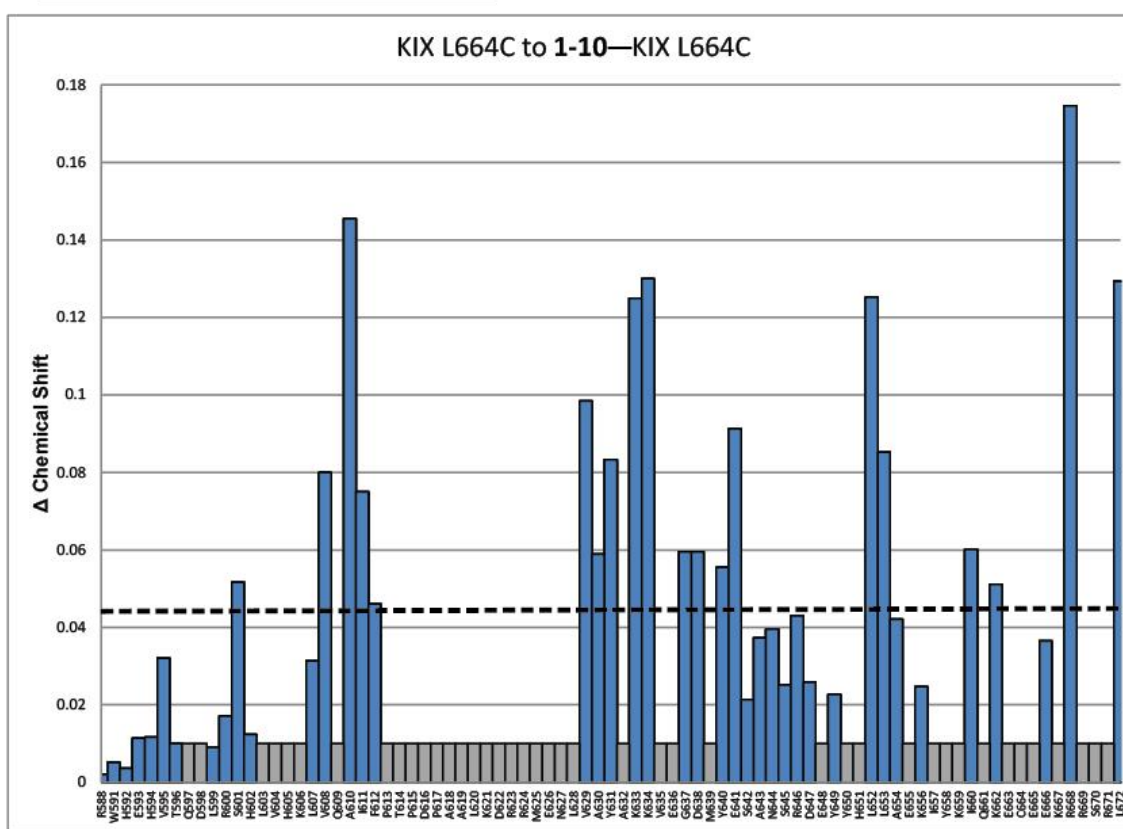


Figure 3.11 ^1H - ^{15}N -HSQC of ^{15}N KIX L664C (red) overlaid with ^{15}N **1-10**—KIX L664C (blue) and ^{15}N **2-64**—KIX L664C (green) and 2xMLL incubated with ^{15}N KIX L664C (yellow). Small molecules were at least 95% tethered to KIX L664C as confirmed by Q-TOF LC-MS (Agilent).

a)



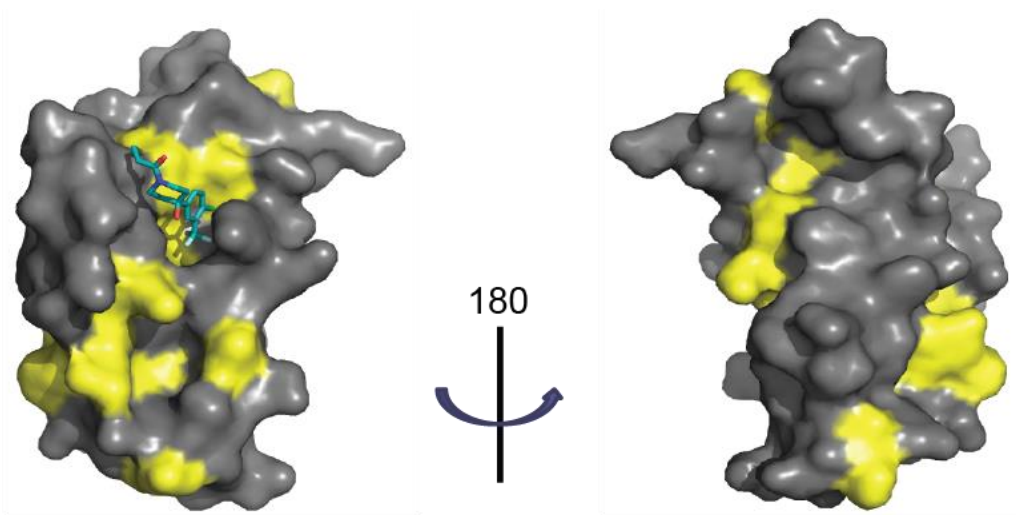
b)



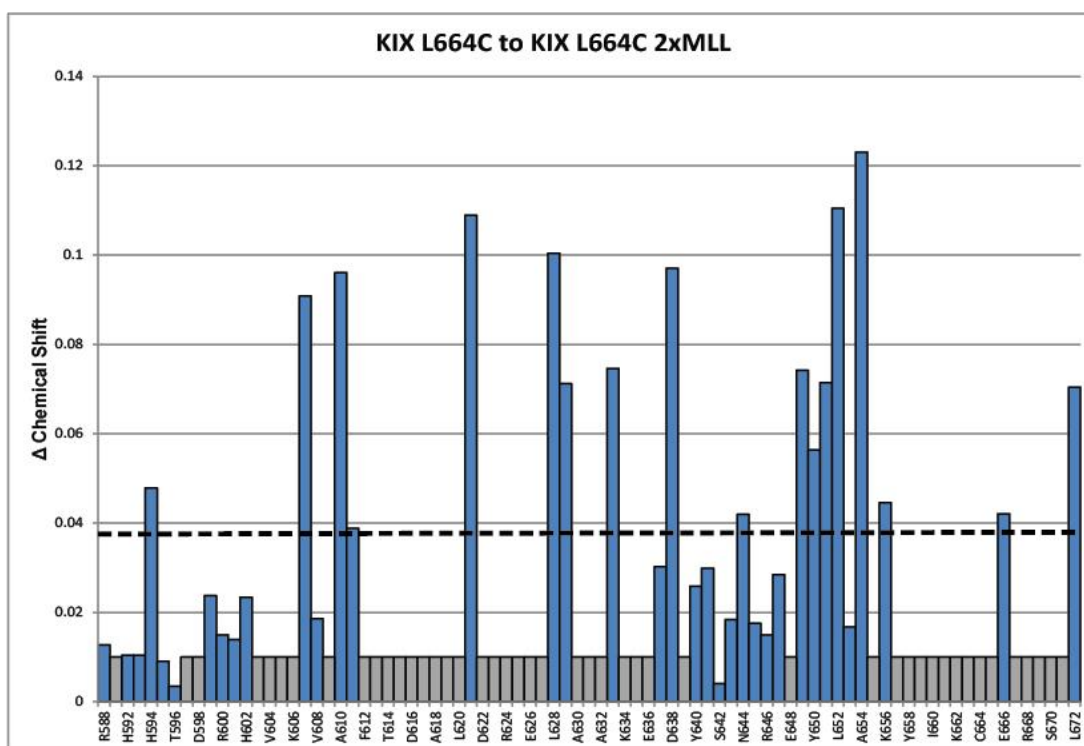
Not detected

Figure 3.12 a) Results from chemical shift perturbation experiment (^1H - ^{15}N -HSQC) with **1-10**-tethered KIX L664C. Residues that shifted more than 1 SD upon **1-10** tethering are in yellow and include Ile611, Leu628, Leu607, Val635 and Tyr631, and Ile660; **b)** Chemical shift perturbation mapping of KIX L664C residues upon tethering to **1-10**. Residues that shift >1 SD are assumed to be significant (dotted line), see methods for details. Residues that are unable to be detected/assigned are shown in gray.

a)



b)



Not detected

Figure 3.13 a) Results from chemical shift perturbation experiment (^1H - ^{15}N -HSQC) with 2x MLL pre-incubated with KIX L664C (using structure of **1-10**—KIX L664C for comparison). Residues that shifted more than 1 SD upon addition of MLL are in yellow; **b)** Chemical shift perturbation mapping of KIX L664C residues upon pre-incubating with 2 molar equivalents of MLL. Residues that shift >1 SD are assumed to be significant (dotted line), see methods for details. Residues that are unable to be detected/assigned are shown in gray.

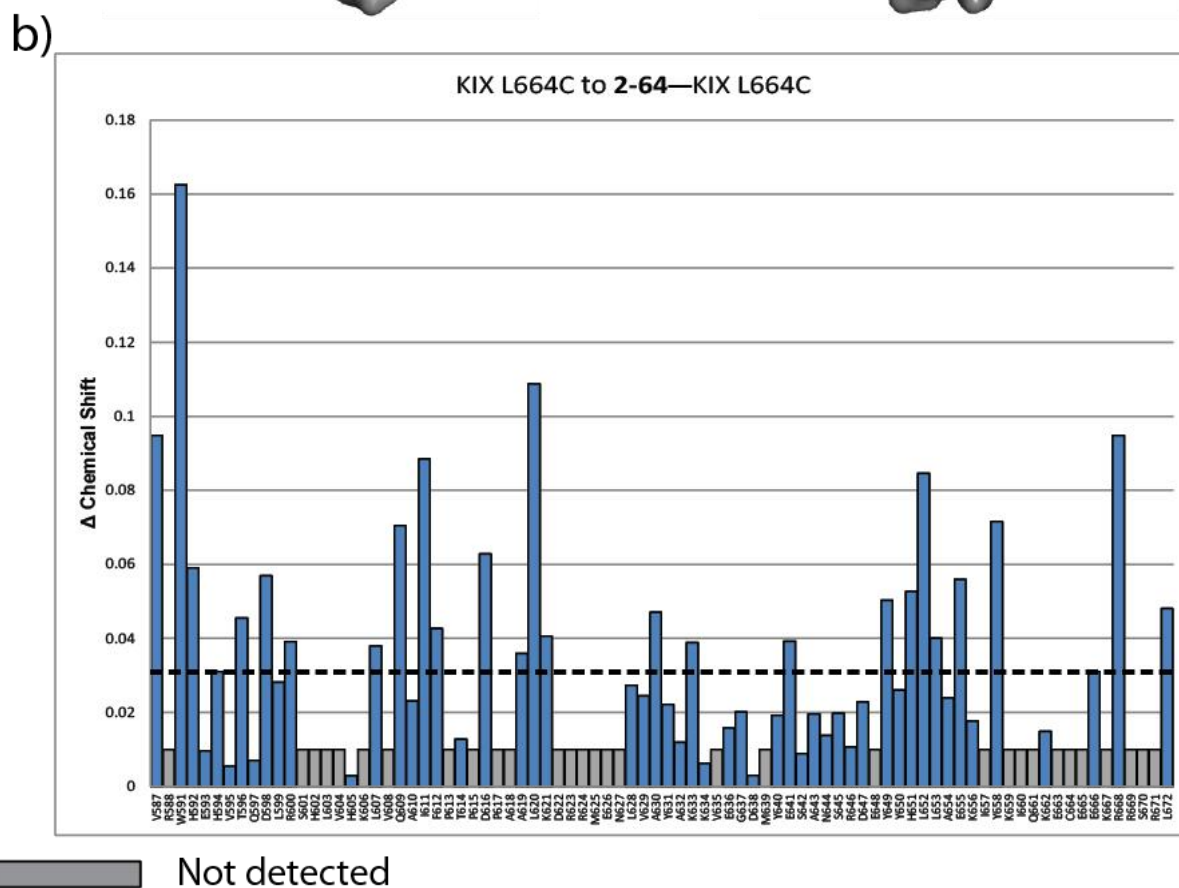
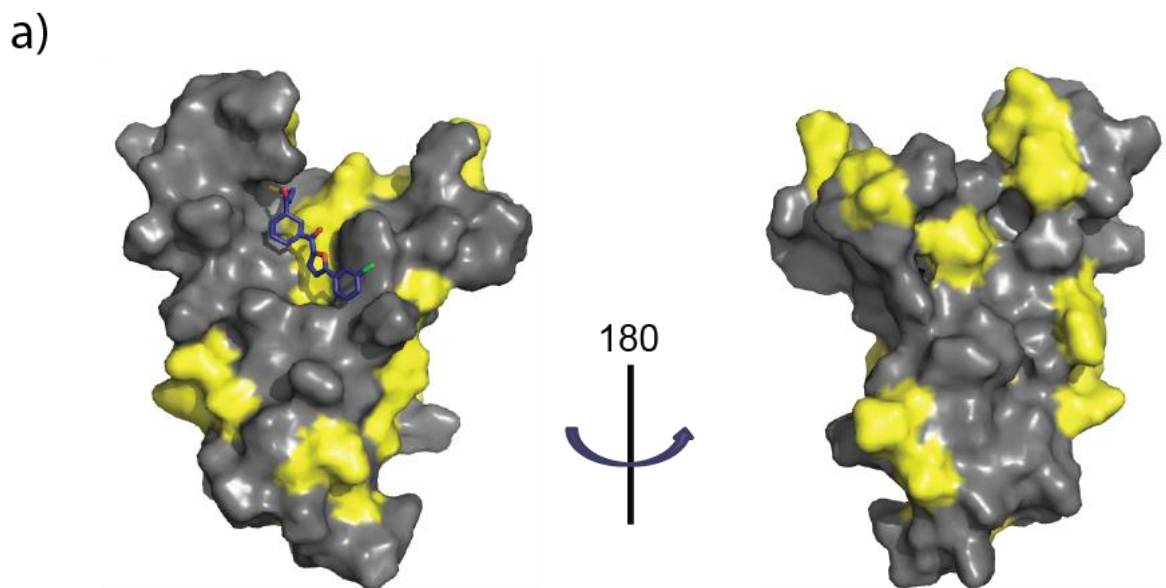


Figure 3.14 a) Results from chemical shift perturbation experiment (^1H - ^{15}N -HSQC) with **2-64**-tethered KIX L664C. Structure of **2-64**---KIX L664C is a simulated structure closest to the calculated average of the 40 ns molecular dynamics simulation started from the top docking result. Residues that shifted more than 1 SD upon **2-64** tethering are in yellow; b) Chemical shift perturbation mapping of KIX L664C residues upon tethering to **2-64**. Residues that shift >1 SD are assumed to be significant (dotted line), see methods for details. Residues that are unable to be detected/assigned are shown in gray.

E. Discussion and Conclusions

To dissect in more detail how the KIX surface remodels itself to recognize fragment **1-10** Jessica Gagnon from the Brookes lab carried out 40 ns molecular dynamics simulations of the KIX crystal structure with or without ligand **1-10**. A gross comparison of the backbone reveals that a change in the loop conformation is the most significant. These changes are often difficult to visualize by solution methods because the loop region contains several proline residues, but mutagenesis and NMR methods have suggested that conformational plasticity in this region underlies the ability of KIX to recognize diverse amphipathic sequences.^{5,25,26} It is this movement of the loop and a rotation of helix α_1 that enable the formation of a narrower binding surface to accommodate a molecule that is considerably smaller than a peptidic helix (~77% smaller volume).

The binding surface that is targeted by **1-10** is also significantly different, both as a result of loop conformational changes and because of side chain motions as demonstrated by the change in solvent accessible surface area of the residues when the fragment is tethered (Figure 3.15a). For example, the liganded KIX shows a population shift in the Tyr 631 side chain χ angles relative to the simulated untethered protein, leading to a hydrophobic binding surface for deeper interactions (Figure 3.15b). Simulations of **2-64** tethered to KIX L664C suggest that the binding mode of this ligand is similar to that of **1-10** and further demonstrates the ability of this protein to adapt to different binding partners (Figure 3.15c).

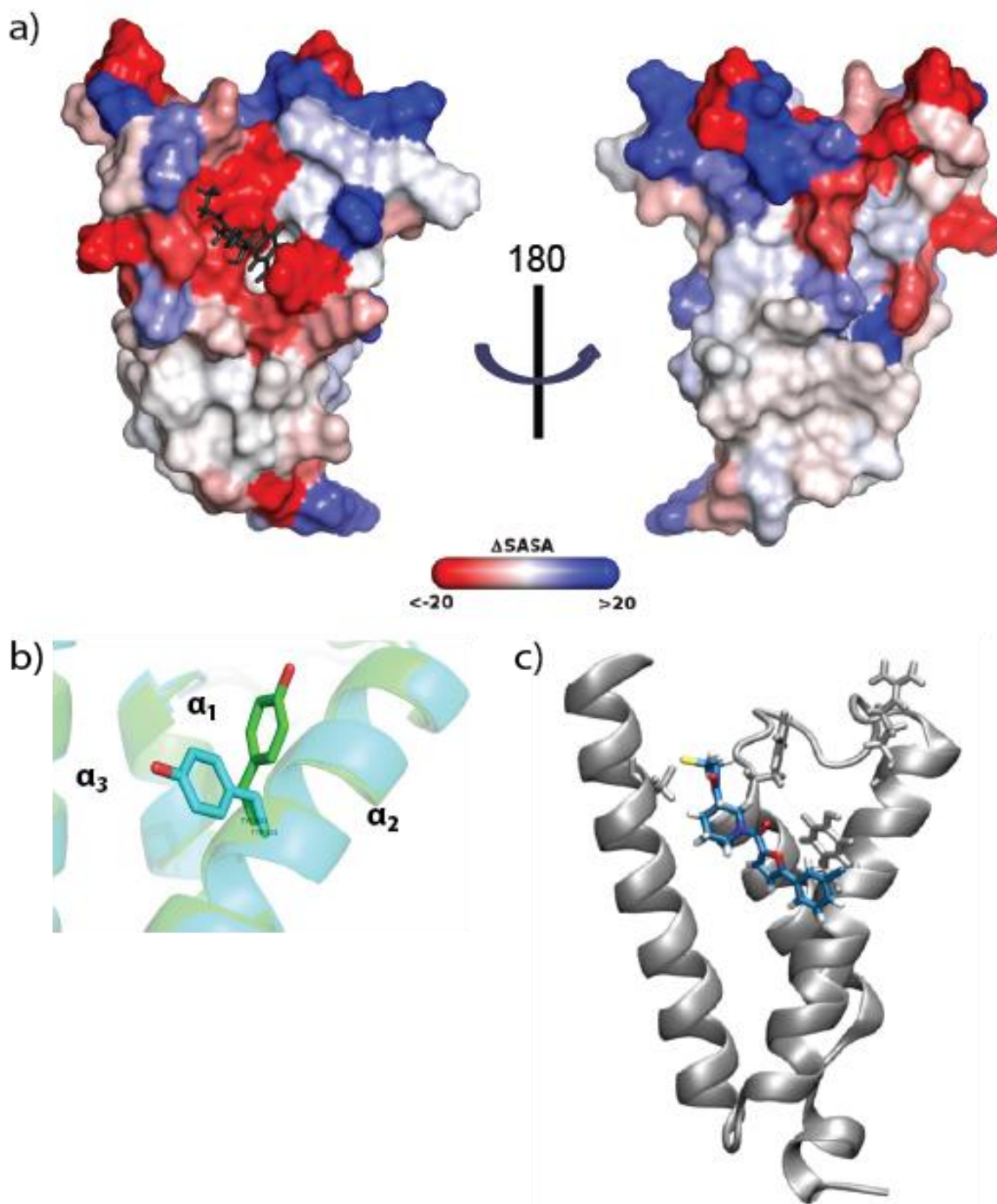


Figure 3.15 a) The difference in the average SASA (Solvent Accessible Surface Area) calculated by residue between simulations of KIX L664C untethered and tethered to **1-10** in units of \AA^2 . A residue colored red is less solvent-exposed in the **1-10**-tethered structure, with color intensity indicating the extent of the change; blue residues are more solvent-exposed in the **1-10**-tethered structure. b) Comparison of the orientation of Tyr631 populations in the **1-10**-tethered KIX L664C (blue) and the simulated free KIX L664C (green). c) Fragment **2-64** tethered to KIX L664C structure closest to the calculated average of the 40 ns molecular dynamics simulation started from the top docking result, with the fragment in blue and protein in gray. Simulations performed by Jessica Gagnon.

While there are several crystal structures of proteins with small molecules covalently tethered, the proteins involved all had crystal structures without tethering to a small molecule as well. This is the first time a small molecule has been shown to stabilize a formerly uncrystallizable protein to form diffraction-quality homogenous crystals. This suggests a great potential for using small molecule tethering as a tool to stabilize and enable conformationally dynamic proteins to form crystals in addition to established methods such as pre-incubating protein with small molecule inhibitors and “stapling” certain residues on a protein.

While this structure of **1-10**--KIX L664C does not fully reflect the conformation of the free KIX domain, it is highly possible that this conformation is one of many conformations that the KIX domain samples in solution, trapped and “frozen” by a small molecule stabilizer. This presents the potential of further crystallizing small molecule-tethered-KIX in complex with different small molecules and TAD peptides. As crystal structures enable us to examine the exact side chain orientations at atomic levels, comparing these structures will enable us to identify specific side chain changes and helix twist changes upon different ligands binding to KIX.

In conclusion, we have obtained a 2Å-resolution snapshot of the conformationally dynamic coactivator KIX domain complexed with a small molecule. This will significantly facilitate using rational structure-based approaches to design more potent analogs; for example, current efforts include extending the molecule **1-10** at the C4 position of the aromatic ring in order to more effectively engage with the hydrophobic space within the KIX site. From a broader perspective, these results in combination with recent studies showing noncovalent small molecules stabilizing conformationally dynamic proteins^{53,61} suggest that Tethering may be an exceptionally enabling approach to obtain long-sought x-ray crystallography data of conformationally dynamic proteins. This includes transcriptional coactivators such as CBP/p300 targeted here, but also members of

other cellular machines that rely upon conformationally dynamic interfaces to recognize binding partners.^{62,63}

F. Future Directions

Efforts to move this project further include plans to modify the reversible disulfide moiety of the tethered fragment hits into a series of irreversible alkylating moiety-containing molecules. These modifications will be examined for their ability to modulate MLL and pKID or c-Myb binding to KIX. There will also be modifications on the fragments such as a biotin handle to enable in-cell pull down studies based on the streptavidin-biotin interaction. Through this we will be able to globally analyze the proteins that the small molecules target within cellular environments and assess the specificity of these small molecules. There has also been simulation work done by Jessica Gagnon and Sarah Graham from the Brookes lab evaluating the theoretical binding affinities of **1-10** fragments with several side chains modified. They have identified a few key modifications that prove to be promising as a KIX binder *in silico*. Future plans include synthesizing these small molecules and testing them against various KIX mutations. Both of these endeavors will be carried out by a graduate student Jean Lodge and a postdoctoral fellow Dr. James Clayton in the lab.

Additional KIX cysteine mutants have also been cloned and expressed where the cysteine is positioned at the pKID/c-Myb binding site of KIX. These KIX constructs will undergo a modified Tethering screen where initial screening will be through an FP based assay (designed by former postdoctoral fellow Dr. William Pomerantz in the lab) and then the initial hits confirmed by mass spectrometry as the previous study. The screens will be carried out by the Wells lab at UCSF in collaboration with Jean Lodge in the Mapp lab.

Crystallization efforts will be continued with additional fragment hits and modified fragments identified and evaluated in the aforementioned projects. We will also focus on crystallizing tethered KIX mutants in complex with transcriptional activator peptides that bind to the other unoccupied site of KIX, in hope of elucidating the allosteric effect of small molecule tethering to the conformation of both the activator peptide and KIX itself.

H. Experimental Methods

Protein Expression and Purification

As previously described,³⁸ the DNA sequence encoding the KIX domain of mouse CBP (586-672) was cloned into a pRSETB vector incorporating a hexahistidine tag with a short linker fused to the N-terminus of KIX for protein expression resulting in a protein with sequence (tag and linker residues are shown in lower case):

mrgshhhhhgmasGVRKGWHEHVTQDLRSHLVHKL VQAI FPTDPAALKDRRMENLV
AYAKKVEGDMYESANSRDEYYHLLAEKIYKIQKELEEKRRSRL

The cysteine mutant at KIX L664C was created using site-directed mutagenesis as previously described.³⁸ For protein expression, the plasmid was transformed into Rosetta2 (DE3) pLysS *E. coli* (Novagen) and grown in LB media. After an O.D.₆₀₀ of 0.8 was reached (37 °C, 250 rpm), the cultures were cooled to 25 °C and expression was induced with 0.5 mM IPTG for 4 h (250 rpm). The His-tagged protein was isolated using Ni-NTA beads (Qiagen) according to the manufacturer's instruction and eluted with 400 mM imidazole. Final purification was carried out by ion-exchange column chromatography on a Source S column, (GE Healthcare) in 50 mM phosphate buffer, 1 mM DTT, pH 7.2 eluting with increasing concentrations of NaCl (0-1 M NaCl). Purified protein solutions were buffer exchanged into 10 mM sodium phosphate, 100

mM NaCl, 10 % glycerol 0.01 % NP40, pH 6.8 using a PD-10 column (GE-Healthcare) and stored at -80 °C.

Selenomethionine incorporated KIX L664C was expressed as previously described using Rosetta2 (DE3) pLysS *E. coli*.⁶⁴ KIX L664C was expressed in minimal media M9 supplemented by amino acid mixture containing Selenomethionine as previously described⁶⁴ and purified as described above. Selenomethionine incorporation was confirmed by Q-TOF LC-MS (Agilent).

Peptide Synthesis and Purification

All peptides (Fl-MLL19, Fl-pKID29, Fl-Myb25) were synthesized as previously described.³⁸

Fragment Screening and DR₅₀ Determination

To identify molecules that interacted with KIX a 480-member disulfide fragment library (SMDC, UCSF) was screened in a high-throughput format using a mass spectrometry based assay.¹⁶ Immediately prior to screening, the bacterially expressed KIX L664C was buffer exchanged to 10 mM sodium phosphate pH 6.8, 100 mM NaCl, 5 mM β-ME using a PD-10 column (GE Healthcare). Next, 2 μM of the buffer exchanged KIX cysteine mutant was incubated with 500 μM of each fragment with shaking for 1 h at room temperature (25 μL total volume in a 96-well plate). Subsequently, the plate was moved to 4 °C and the mass of each well measured using a LCT-Premier LC/electrospray ionization-MS instrument (Waters). Protein masses were deconvoluted using the Max-Ent algorithm within the MassLynx software. Tethering efficiency was measured by comparing peak areas for the tethered versus untethered protein.

Dose Response (DR)₅₀ values were determined as previously described.¹³ Varying concentrations (500-0.2 μM) of the fragment molecules were incubated with 2 μM KIX mutant protein (at 1mM β-ME) for 1 hour at RT. The percent of protein tethered to fragment molecules were determined by Q-TOF LC-MS (Agilent). The concentration of fragment molecule required for 50% maximum tethering (DR₅₀) was determined by data analysis in GraphPad Prism software, fitting to Equation 3.1, where x is the log of fragment molecule concentration and y is the normalized response from 1 to 100 (percent of protein tethered to fragment molecule).

$$y = \frac{100}{1+10^{(\text{Log}DR_{50}-x)}} \quad (\text{Eq 3.1})$$

Fragment Tethering

KIX L664C was incubated with 10 equivalents of **1-10** or **2-64** in 50 mM phosphate buffer at pH 6.8 overnight, excess small molecule was removed by NAP-5 desalting column and buffer exchanged into 20 mM HEPES, 100 mM NaCl, pH 8.0. Small molecule tethering efficiency was confirmed by Q-TOF LC-MS (Agilent) and tethering reached at least 95% complete. Protein-small molecule complexes were flash frozen in liquid N₂ and stored at -80°C.

Circular Dichroism

CD spectra were acquired on a J-715 spectropolarimeter equipped with a temperature control unit (Jasco, Inc., Easton, MD). Briefly, samples of 10 μM KIX L664C were treated with DMSO or small molecule fragments (10 eq.) and tethered >95% (overnight, room temperature and confirmed by quantitative LC-MS) in 10 mM sodium phosphate pH 6.8, 100 mM NaCl, < 2% DMSO, 0.1 mM β-ME. CD spectra were recorded using a 1 mm path length quartz cuvette. The CD signal resulting from the buffer alone was subtracted from the spectrum of each protein

solution. Variable temperature CD was acquired by monitoring the molar ellipticity at 222 nm during heating of the protein solutions at a rate of 1 degree/minute from 20-95 °C. Data were converted to mean residue ellipticity, $[\Theta]$ (deg cm² dmol⁻¹) according to the Equation 3.2:

$$[\Theta] = \Psi / (1000 * n * l * c) \quad (\text{Eq. 3.2})$$

Where Ψ is the CD signal in degrees, n is the number of amides, l is the path length in centimeters, and c is the concentration in decimoles per cm³. Data analysis was performed using GraphPad Prism software to calculate the melting temperature.

Hydrogen-Deuterium Exchange

Hydrogen-deuterium (HD) exchange experiments were carried out with tethered and free KIX L664C to probe for the protection of backbone amide protons by small molecules **1-10** and **2-64** following procedures similar to previous work by Wells et al.¹³ Frozen aliquots (100 μM) of KIX L664C, **1-10**—KIX L664C and **2-64**—KIX L664C were thawed on ice, these solutions were then diluted 4-fold into 12.5 mM Tris 50 mM NaCl in 97.5 % D₂O and 2.5 % H₂O, pH 7.0 directly in a mass spectrometry sample vial. The vial was immediately placed in the automatic sampling plate (4°C) of the Q-TOF HPLC-MS (Agilent). The sample was injected onto the mass spectrometer 1 minute after exposure to D₂O. Subsequently for comparison, another vial with the protein/ protein-small molecule complexes diluted 4-fold into 12.5 mM Tris 50 mM NaCl in 100 % H₂O, pH 7.0 was injected on the mass spectrometer as well. Data was analyzed by Agilent Qualitative Analysis Program. The difference in molecular weight with and without deuterium exchange was calculated, and percent amides protected was calculated as 100 % - (Δ (**1-10** or **2-64**) — KIX L664C / Δ KIX L664C) as shown in Table 3.3.

Proteolysis

25 μL of KIX L664C, **1-10**—KIX L664C and **2-64**—KIX L664C were diluted in 225 μL of proteolysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM CaCl_2 , 4.5 M Urea) to yield a final protein concentration of 15 μM and incubated at RT for 1 hr. Thermolysin from a 10 mg/mL stock solution was added to the protein at a 1:30 molar ratio. 15 μL of the proteolysis reaction was added to 5 μL of 50 mM EDTA to quench proteolysis at various time points and stored at $-20\text{ }^\circ\text{C}$. The quenched samples were analyzed by SDS-PAGE (12 % Bis-Tris gel in MES running buffer, staining with comassie blue). Band intensities were analyzed by ImageJ imaging software. Percent protein remaining was plotted against time and fit to an exponential decay equation to obtain the half-life of the protein using GraphPad Prizm software.

Fluorescent Anisotropy Assays

The fluorescent anisotropy assays were done in triplicate with a final sample volume of 10 μL in a low volume, non-binding, black, 384-well plate (Corning), and read using a Tecan Genios Pro plate reader with polarized excitation at 485 nm and emission intensity measured through

a parallel and perpendicularly polarized 535 nm filter. FITC (Fluorescein isothiocyanate) labeled peptides were diluted in storage buffer (10 mM Phosphate, 100 mM NaCl, 0.01 % NP-40, 10 % Glycerol, pH 6.8) to a concentration of 25 nM. Then 10 μL of the peptide solution was added to a series of 50 μL solutions of varying KIX concentrations in storage buffer to obtain the final concentrations of up to 20 μM . The samples were incubated for 30 min at room temperature

before the degree of fluorescence anisotropy was measured (Tecan Genios Pro). A binding isotherm that accounts for ligand depletion (assuming a 1:1 binding model of peptide to KIX) was fit to the observed anisotropy values as a function of KIX to obtain the apparent equilibrium dissociation constant, K_D :

$$y = c + (b - c) \times \frac{[(K_D + a + x) - \sqrt{(K_D + a + x)^2 - 4ax}]}{2a} \quad (\text{Eq 3.3})$$

where “a” and “x” are the total concentrations of fluorescent peptide and KIX, respectively, “y” is the observed anisotropy at any KIX concentration, “b” is the maximum observed anisotropy value, and “c” is the minimum observed anisotropy value. Each data point is an average of three independent experiments with the indicated error (standard deviation). Data analysis was performed using GraphPad Prizm 5 software.

Protein Crystallization

Initial conditions found for crystallization of 1-10—KIX L664C was from Index-HT™ condition A5 (2.0 M Ammonium sulfate, 0.1 M HEPES pH 7.5).

Purified Selenomethionine incorporated 1-10—KIX L664C was concentrated to 4.5 mg/mL in 25 mM HEPES, 100 mM NaCl buffer (pH 6.5) prior to crystallization. Crystals were grown by hanging drop vapor diffusion at 20 °C with drops containing 2 μL protein and 2 μL precipitant (1.8 M (NH₄)₂SO₄, 0.1 M Tris-HCl pH 7.0). Crystals were soaked in well solution containing 10 % ethylene glycol, then transferred to a solution of 20 % ethylene glycol for cryoprotecting prior to flash freezing in liquid nitrogen for data collection.

Data Collection and Refinement

Data were collected at the Advanced Photon Source (LS-CAT Beamline 21-ID-F) at the Argonne National Laboratory equipped with a Mar225 detector at wavelength of 0.97852 Å and -180 °C. Data were processed and scaled with HKL2000.⁶⁵ **1-10**—KIX L664C crystallized in space group P4₃2₁2, with unit cell parameters of a = b = 48.330 Å, c = 85.464 Å, and $\alpha = \beta = \gamma = 90^\circ$. Phases were initially determined by single wavelength anomalous X-ray scattering of the selenium atoms using AutoSol in Phenix.⁶⁶ The program identified three Se atom sites, pertaining to the two ordered Se-methionines in the KIX structure. One Se-Met residue had two equivalent positions; the other two Se-met residues were located in the disordered N-terminal region. The resulting structure was fit in Coot⁶⁷ using the $2F_o - F_c$ and $F_o - F_c$ electron density maps from Phenix, followed by rigid body and restrained refinement in Buster.⁶⁸ After refinement clear $F_o - F_c$ electron density was apparent for compound **1-10**. Three-dimensional coordinates and buster style restraint files for the compound were generated in GRADE⁶⁹ using a smile string created by the PRODRG web server.⁷⁰ Iterative rounds of fitting in Coot and refinement in Buster of the **1-10**—KIX L664C structure to 2.0 Å resulted in $R_{\text{work}} = 20.6\%$ and $R_{\text{free}} = 23.3\%$.

All residues from the three structures are in the allowed regions of the Ramachandran plot. Structures were validated with Molprobit, ⁷¹ Parvati, ⁷² and whatcheck.⁷³ Ligand statistics were obtained from the Uppsala Electron-Density Server.⁷⁴ Areas of poor electron density were not modeled. These areas include side chains of residues Arg 588, Lys 621, Asp 622, Asp 638, Lys 662 and Glu 665. The N-terminal His tag along with N-terminal residues 584-587 and C-terminal residues 666-672 were disordered in the structure. Data refinement and statistics are given in Table 3.5.

¹H-¹⁵N-HSQC NMR Experiments

Uniformly ¹⁵N labeled KIX L664C protein was expressed and purified as previously described.⁴⁰ Samples of the purified ¹⁵N labeled KIX L664C were tethered with small molecules **1-10** or **2-64** as described above. A 45 μM solution of ¹⁵N- labeled KIX L664C with or without small molecule was prepared in a 9:1 H₂O:D₂O 10 mM sodium phosphate buffer containing 100 mM NaCl at pH 7.2. ¹H-¹⁵N HSQC experiments were recorded at 27 °C on an Avance Bruker 600 MHz NMR spectrometer equipped with a 5 mm cryogenic probe. HSQCs were collected with the protein itself, the protein covalently tethered to **1-10** or **2-64** and the chemical shifts were compared. Data was processed using NMRpipe⁷⁵ and analyzed in Sparky (UCSF).⁷⁶ Chemical shifts of residues were identified based on previous assignments.³⁵ Chemical shift changes for individual peaks were quantified as $((0.2 * \Delta\delta^{15\text{N}})^2 + (\Delta\delta^1\text{H})^2)^{0.5}$, which is a weighted length of the vector from free KIX L664C to protein complexed with small molecule.

Table 3.5 Data collection, phasing and refinement statistics**Data Collection and Phasing Statistics**

SpaceGroup	P4 ₃ 2 ₁ 2
Unit Cell	a =48.330 b =48.330 c =85.464 $\alpha = 90^\circ \quad \beta =90 \quad \gamma =90$
Wavelength (Å)	0.97852
Resolution (Å)	50-2.0 (2.03)
R _{svm} (%)	0.061 (0.228)
<I/sI>	20 (10)
Completeness (%)	100 (100)
Redundancy	20.8 (21.2)
Phase Resolution (Å)	2.5
Figure of Merit	0.44
BAYES-CC	48.1 ± 18.6
R Factor	0.33/0.37
Skew	0.16
Map Correlation	0.07
Selenomethionine Sites	3 (M625, M639A, M639B)
Final Refined Occupancy	M625 (1.0), M639A (0.70), M639B (0.30)
Final Refined B-Factor	M625 (59.79), M639A (57.80), M639B (23.22)

Refinement Statistics

Resolution (Å)	42.07-2.00
R _{work} (%)	0.2064
R _{free} (%)	0.2329
Protein atoms	636
Ligand Atoms	79
Water Molecules	27
Unique Reflections	7312
Correlation Coefficient	0.9233
R.m.s.d.	
Bonds	0.01
Angles	0.87
MolProbity Score	1.34
Clash Score	3.99

Ligand Statistics

Real-space R-value	0.28
Real-Space Correlation Coefficient	0.81

G. References

1. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
2. Thompson, A. D., Dugan, A., Gestwicki, J. E. & Mapp, A. K. Fine-tuning multiprotein complexes using small molecules. *ACS Chem. Biol.* **7**, 1311–20 (2012).
3. Singh, G. P., Ganapathi, M. & Dash, D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* **66**, 761–5 (2007).
4. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
5. 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**, 1005–13 (2006).
6. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *J. Biol. Chem.* **277**, 43168–74 (2002).
7. 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 (2001).
8. Cook, P. R., Polakowski, N. & Lemasson, I. HTLV-1 HBZ protein deregulates interactions between cellular factors and the KIX domain of p300/CBP. *J. Mol. Biol.* **409**, 384–98 (2011).
9. Vendel, A. C., McBryant, S. J. & Lumb, K. J. KIX-mediated assembly of the CBP-CREB-HTLV-1 tax coactivator-activator complex. *Biochemistry* **42**, 12481–7 (2003).
10. Kasper, L. H. *et al.* A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. *Nature* **419**, 738–43 (2002).
11. Wood, M. A., Attner, M. A., Oliveira, A. M. M., Brindle, P. K. & Abel, T. A transcription factor-binding domain of the coactivator CBP is essential for long-term memory and the expression of specific target genes. *Learn. Mem.* **13**, 609–17 (2006).
12. Liu, G.-H., Qu, J. & Shen, X. NF-kappaB/p65 antagonizes Nrf2-ARE pathway by depriving CBP from Nrf2 and facilitating recruitment of HDAC3 to MafK. *Biochim. Biophys. Acta* **1783**, 713–27 (2008).
13. Sadowsky, J. D. *et al.* Turning a protein kinase on or off from a single allosteric site via disulfide trapping. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6056–61 (2011).
14. Hardy, J. A., Lam, J., Nguyen, J. T., O'Brien, T. & Wells, J. A. Discovery of an allosteric site in the caspases. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12461–6 (2004).
15. Scheer, J. M., Romanowski, M. J. & Wells, J. A. A common allosteric site and mechanism in caspases. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7595–600 (2006).
16. Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9367–9372 (2000).
17. Bordoli, L. Plant orthologs of p300/CBP: conservation of a core domain in metazoan p300/CBP acetyltransferase-related proteins. *Nucleic Acids Res.* **29**, 589–597 (2001).
18. Shi, Y. & Mello, C. A CBP/p300 homolog specifies multiple differentiation pathways in *Caenorhabditis elegans*. *Genes Dev.* **12**, 943–955 (1998).
19. Chan, H. M. & La Thangue, N. B. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.* **114**, 2363–73 (2001).
20. Giordano, A. & Avantaggiati, M. L. p300 and CBP: partners for life and death. *J. Cell. Physiol.* **181**, 218–30 (1999).
21. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553 (2000).
22. Wang, F. & Marshall, C. Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6078–83 (2012).

23. 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).
24. 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**, 521–34 (2004).
25. Brüschweiler, S. *et al.* Direct observation of the dynamic process underlying allosteric signal transmission. *J. Am. Chem. Soc.* **131**, 3063–8 (2009).
26. Korkmaz, E. N., Nussinov, R. & Haliloğlu, T. Conformational Control of the Binding of the Transactivation Domain of the MLL Protein and c-Myb to the KIX Domain of CREB. *PLoS Comput. Biol.* **8**, e1002420 (2012).
27. Novatchkova, M. & Eisenhaber, F. Linking transcriptional mediators via the GACKIX domain super family. *Curr. Biol.* **14**, R54–R55 (2004).
28. Withers-Martinez, C. *et al.* Malarial EBA-175 region VI crystallographic structure reveals a KIX-like binding interface. *J. Mol. Biol.* **375**, 773–81 (2008).
29. Thakur, J. K. *et al.* A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **452**, 604–609 (2008).
30. Yang, F. *et al.* An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442**, 700–704 (2006).
31. Ichiki, T. Role of cAMP response element binding protein in cardiovascular remodeling: good, bad, or both? *Arterioscler., Thromb., Vasc. Biol.* **26**, 449–55 (2006).
32. Chen, Y. *et al.* The c-Myb functions as a downstream target of PDGF-mediated survival signal in vascular smooth muscle cells. *Biochem. Biophys. Res. Commun.* **360**, 433–6 (2007).
33. Ernst, P. *et al.* Definitive Hematopoiesis Requires the Mixed-Lineage Leukemia Gene. *Dev. Cell* **6**, 437–443 (2004).
34. Gusterson, R. J., Jazrawi, E., Adcock, I. M. & Latchman, D. S. The transcriptional co-activators CREB-binding protein (CBP) and p300 play a critical role in cardiac hypertrophy that is dependent on their histone acetyltransferase activity. *J. Biol. Chem.* **278**, 6838–47 (2003).
35. 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).
36. Best, J. L. *et al.* Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17622–7 (2004).
37. Li, B. X., Yamanaka, K. & Xiao, X. Structure-activity relationship studies of naphthol AS-E and its derivatives as anticancer agents by inhibiting CREB-mediated gene transcription. *Bioorg. Med. Chem.* **20**, 6811 (2012).
38. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **7**, 1345–50 (2012).
39. Li, B. X. & Xiao, X. Discovery of a small-molecule inhibitor of the KIX-KID interaction. *ChemBioChem* **10**, 2721–4 (2009).
40. Buhrlage, S. J. J. *et al.* Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem. Biol.* **4**, 335–344 (2009).
41. Metallo, S. J. Intrinsically disordered proteins are potential drug targets. *Curr. Opin. Chem. Biol.* **14**, 481–8 (2010).
42. Meireles, L. M. C. & Mustata, G. Discovery of modulators of protein-protein interactions: current approaches and limitations. *Curr. Top. Med. Chem.* **11**, 248–57 (2011).
43. Berg, T. Inhibition of transcription factors with small organic molecules. *Curr. Opin. Chem. Biol.* **12**, 464–71 (2008).
44. Koehler, A. N. A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.* **14**, 331–40 (2010).
45. Chung, C.-W., Dean, A. W., Woolven, J. M. & Bamborough, P. Fragment-based discovery of bromodomain inhibitors part 1: inhibitor binding modes and implications for lead discovery. *J. Med. Chem.* **55**, 576–86 (2012).

46. Bamborough, P. *et al.* Fragment-based discovery of bromodomain inhibitors part 2: optimization of phenylisoxazole sulfonamides. *J. Med. Chem.* **55**, 587–96 (2012).
47. Shi, A. *et al.* Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia. *Blood* **120**, 4461–9 (2012).
48. Bernard, D., Zhao, Y. & Wang, S. AM-8553: a novel MDM2 inhibitor with a promising outlook for potential clinical development. *J. Med. Chem.* **55**, 4934–5 (2012).
49. Kung, A. L. *et al.* Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell* **6**, 33–43 (2004).
50. Scheuermann, T. H. *et al.* Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat. Chem. Biol.* **9**, 271–276 (2013).
51. Thiel, P., Kaiser, M. & Ottmann, C. Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery? *Angew. Chem., Int. Ed.* **51**, 2012–8 (2012).
52. Follis, A. V., Hammoudeh, D. I., Wang, H., Prochownik, E. V & Metallo, S. J. Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. *Chem. Biol.* **15**, 1149–55 (2008).
53. Saalau-Bethell, S. M. *et al.* Discovery of an allosteric mechanism for the regulation of HCV NS3 protein function. *Nat. Chem. Biol.* **8**, 920–5 (2012).
54. Kaar, J. L. *et al.* Stabilization of mutant p53 via alkylation of cysteines and effects on DNA binding. *Protein Sci.* **19**, 2267–78 (2010).
55. Park, C. & Marqusee, S. Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nat. Methods* **2**, 207–12 (2005).
56. Arai, M., Dyson, H. J. & Wright, P. E. Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. *FEBS Lett.* 4500–4504 (2010).
57. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J. & Verdine, G. L. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277**, 1310–1313 (1997).
58. Thoden, J. B., Ryan, L. A., Reece, R. J. & Holden, H. M. The interaction between an acidic transcriptional activator and its inhibitor. The molecular basis of Gal4p recognition by Gal80p. *J. Biol. Chem.* **283**, 30266–30272 (2008).
59. Fuller, J. C., Burgoyne, N. J. & Jackson, R. M. Predicting druggable binding sites at the protein-protein interface. *Drug Discovery Today* **14**, 155–61 (2009).
60. Fry, D. Protein–protein interactions as targets for small molecule drug discovery. *Pept. Sci.* **84**, 535–552 (2006).
61. Basse, N. *et al.* Toward the rational design of p53-stabilizing drugs: probing the surface of the oncogenic Y220C mutant. *Chem. Biol.* **17**, 46–56 (2010).
62. Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18**, 756–64 (2008).
63. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
64. Begley, M. J. *et al.* Crystal Structure of a Phosphoinositide Phosphatase, MTMR2. *Mol. Cell* **12**, 1391–1402 (2003).
65. Otwinowski, Z. & Minor, W. Processing of X-Ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307–326 (1997).
66. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–21 (2010).
67. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 486–501 (2010).
68. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., P. W. & Roversi P, Sharff A., Smart O.S., Vonnrhein C., W. T. O. Buster version 2.10.0. (2011).
69. O. S. Smart, T. O. Womack, A. Sharff, C. F. & P. Keller, W. Paciorek, C. Vonnrhein, and G. B. GRADE version 1.1.1. (2011).at <<http://www.globalphasing.com>>

70. Schüttelkopf, A. W. & Van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 1355–63 (2004).
71. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 12–21 (2010).
72. Zucker, F., Champ, P. C. & Merritt, E. A. Validation of crystallographic models containing TLS or other descriptions of anisotropy. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 889–900 (2010).
73. Hoof, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).
74. Kleywegt, G. J. *et al.* The Uppsala Electron-Density Server. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2240–9 (2004).
75. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, (1995).
76. T. D. Goddard and D. G. Kneller Sparky 3.
77. Wang, N. *et al.* Ordering a dynamic protein via a small-molecule stabilizer. *J. Am. Chem. Soc.* **135**, 3363–6 (2013).

Chapter 4 Kinetic Characterization of a Transcriptional Coactivator (KIX) in Complex with Peptide and Small Molecule Binding Partners

A. Introduction

Chapter 2 outlined our study on the kinetic mechanism of TAD-Med15 interactions. However there is little structural information on either the TADs involved or the coactivator Med15. This deters further study on the conformational change aspect of the interaction, which we have found to play a key role in TAD potency.¹ Hence a model system with more detailed structural information is needed. One such system is the TAD-KIX complex. The interaction between disordered transcriptional activators and the conformationally dynamic KIX domain of the coactivator CBP has long served as a prototypical case study for the kinetics and conformational dynamics of protein-protein interactions involving highly flexible components.²⁻⁶ The conformational changes that the TADs and KIX undergo upon binding along with the allosteric interaction between different TADs at distal binding sites of KIX have been examined by various NMR studies.^{2,7,8} However, aspects of these interactions remain poorly understood. For example, the underlying mechanism of the allostery between the two binding sites on the KIX domain has yet to be well defined, as well as the ability of different protein ligands to elicit different degrees of allosteric effects.^{9,10} In addition, there is limited study on small molecule

modulators that target this interaction.^{11,12} To address this problem, Mapp and coworkers have recently performed a disulfide screen using Tethering¹³ technology to identify several small molecule fragments that can selectively tether to cysteine mutations introduced within the KIX domain.^{14,15} Two molecule fragments that stand out as having high binding affinity to KIX mutants are **1-10** and **2-64**. Interestingly, these tethered small molecules are able to allosterically affect pKID binding at the distal site of KIX to varying extents. In this chapter the underlying transient-state kinetics mechanism of this small-molecule elicited allosteric regulation is studied in detail using stopped-flow spectroscopy. While pKID association constants are similar across different protein-protein and protein-small molecule complexes, we find that the dissociation constant best reflect the change in KIX's affinity to pKID. We further noticed that **1-10** can elicit distinct allosteric effects on the pKID binding of KIX depending on the site at which it is tethered. Thus, **1-10** is a potentially powerful probe to dissect the allosteric mechanism that governs this prototypical coactivator motif.

B. Background

B.1. The KIX domain interacts with the TADs of MLL, c-Myb and pKID

Three well-studied TADs that bind to KIX with low micromolar affinity are those of MLL, pKID (phosphorylated kinase inducible domain of CREB), and c-Myb. The binding site targeted by MLL is located on the opposite binding surface as that targeted by both pKID and c-Myb, albeit in different orientations (Figure 4.1).^{7,8} As mentioned in Chapter 3, c-Myb and pKID can each bind cooperatively to KIX with MLL,^{7,16} NMR studies have suggested that this is enabled by an allosteric network between the two binding sites.¹⁷ All three of these TADs are considered to be intrinsically disordered proteins (IDPs). While there are detailed studies on each of the

three activators binding to KIX separately, research in the field has mainly focused on elucidating the mechanisms of the c-Myb and pKID binding interactions.^{2,3,18-23} As is the case with many studies on TAD-coactivator interactions, all the TADs used in the following studies as well as structure determination were isolated peptide fragments.

KID is the kinase inducible domain of the cAMP response element binding protein

(CREB), and when isolated *in vitro*, this domain is unstructured.²¹ However, upon phosphorylation of Ser133 by protein kinase A, the domain (now named pKID) binds to the KIX domain and assumes a helical structure.^{8,21} Computational and experimental data suggest the phosphorylated serine contributes to pKID's affinity to KIX by both electrostatic interactions as well as hydrogen bonding with Lys 662 or Tyr 658 on KIX.^{8,24} Interestingly, phosphorylation of a different serine residue on KID, Ser142, attenuates its ability to bind KIX, thus suggesting this is a fine-tuned interaction that requires the interplay between a well-defined collection of side chains.²⁰

NMR relaxation studies and simulations using a Go-type model have suggested a three-step mechanism for the pKID-KIX interaction (Figure 4.2).^{2,25} in which pKID is initially unstructured when it encounters KIX,²⁶ and only assumes the helical structure in the final bound complex. Furthermore, the overall on-rate (k_{on}) for this binding process from the “encounter”

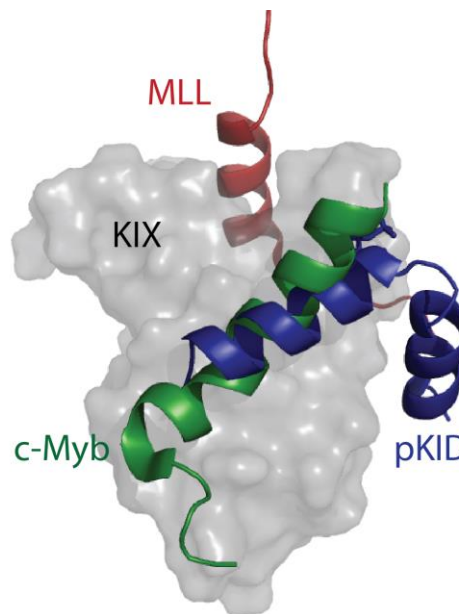


Figure 4.1 Overlay of structures 1KDX and 2AGH, showing both c-Myb (green) and pKID (blue) binding to the KIX domain (grey). MLL (red) is shown at the distal site.

complex to the “bound” complex has been determined to be an average of $6.3 \mu\text{M}^{-1}\text{s}^{-1}$ by NMR relaxation studies.

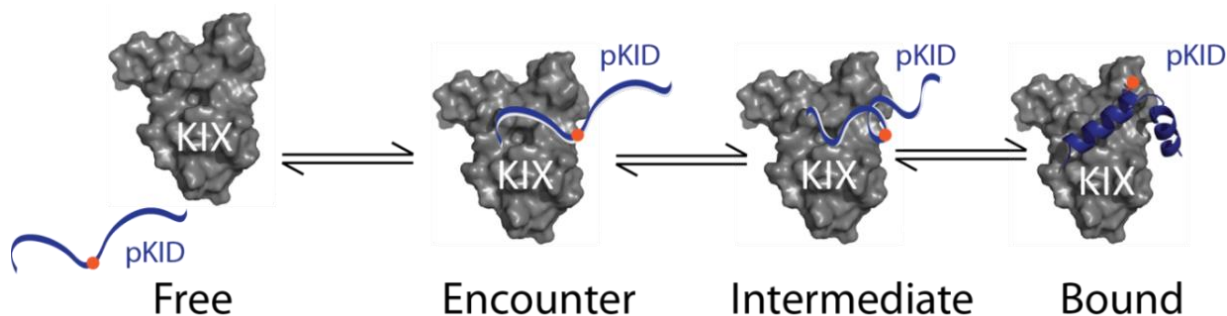


Figure 4.2 Schematic of the coupled folding and binding mechanism proposed for the pKID-KIX interaction. Orange sphere refers to the phosphorylated Ser133 on pKID.

In particular, it must be noted that the first step from “free” to “encounter” was proposed as a theory, but the time scale on which this step would occur is too fast to observe by the techniques employed. This step was supported, however by computational simulations.²⁵ The same computational study also observed that increasing the helical propensity of pKID actually led to a decrease in its association rate with KIX. This finding is consistent with the “fly casting” model,²⁷ where a more flexible, less structured protein has larger available surface area to encounter a binding partner, resulting in a faster association rate.

Studies have shown that although c-Myb and pKID bind to KIX at the same site, the underlying mechanisms bear distinctive features. c-Myb binds to KIX at the same site as pKID, but with 4-25 fold lower binding affinity (values vary between studies based on the length of c-Myb peptide used and techniques employed).^{14,23,28} Similar to pKID, c-Myb becomes more helical upon binding KIX and is proposed to follow a similar folding after binding mechanism.²² However CD studies have shown that there is significantly higher content of helicity in free c-Myb than in the free pKID peptide.²⁸ The driving force of c-Myb binding to KIX is mainly the hydrophobic interactions between the side chains and structural re-organization is less

dramatic.²⁸ This further illustrates that elements such as conformational dynamics and interaction interface can differentiate two separate ligands binding to the same site of a protein.

Interestingly, a recent study showed that a potent transcriptional activator HBZ (which binds to KIX at the MLL/cJun site) potentiates c-Myb binding by 6-fold, but does not have any potentiation effect on pKID.¹⁰ This supports the hypothesis that c-Myb and pKID bind KIX through different modes and are connected by different allosteric networks. This study also reveals the fact that different ligands binding at the same site of KIX can have different allosteric effects on ligands at the opposite site. However there have not been detailed transient-state kinetics studies comparing different allosteric effects on KIX and its binding partners. Our studies seek to elucidate the mechanism of a similar situation where small molecule-ligands elicit varying allosteric effects on pKID binding to KIX.

B.2. The study of transient-state kinetics of IDP protein-protein interactions

Protein-protein interactions between intrinsically disordered proteins (IDPs) and their binding partners are complicated processes. In addition to electrostatic interactions and hydrogen bonding, various other elements must be considered, such as hydrophobic effects and the folding of the disordered protein upon binding. This requires not only studying the thermodynamic binding affinity of the interaction (K_{DS}), but additionally examining the kinetic properties such as transient-state association (k_{on}) and dissociation (k_{off}) rate constants.

An important characteristic of IDP interactions exemplifies the importance of understanding the transient-state kinetics involved: it is widely suggested that IDPs are able to fine-tune multiple branching signaling pathways by a combination of high specificity and low affinity.^{29–31} A closer look at the transient-state kinetics reveals that IDPs are able to achieve this low binding

affinity by having extremely high association constants^{2,32,33} (likely due to a larger capture radius caused by being structurally flexible)^{4,27} but at the same time also have very high dissociation constants.³⁴ This short-lived, fast association contributes to the high level of specificity.

Controversial opinions exist as to whether it is the association or the dissociation rate constant that governs the equilibrium binding affinity in IDP interactions. A review by Prakash (2011)³⁵ found a positive correlation between k_{on} and K_{D} by examining the mutation studies of several protein-protein interactions involving IDPs (these rate constants were obtained by either SPR or fluorescence stopped-flow experiments).^{33,36,37} However a study by Raza Haq et al (2011)³⁸ determined that the dissociation rate constant k_{off} governs K_{D} . This was based on mutation studies of intrinsically disordered peptides binding to different PDZ domains and the rate constants were obtained by FRET stopped-flow. The author argued that in the Prakash paper many of the mutations were involved in electrostatic interactions and no mutations that affected hydrophobicity were included, hence skewing the results. Clearly further studies on a wider array of systems need to be carried out for a more concrete verdict.

The work in this chapter takes into consideration the advancements and controversies in the field of PPI interaction kinetics, and seeks to establish a model for the kinetic mechanism of transcriptional activator-KIX interactions with the tool of a small molecule probe to elucidate the role of allostery in this interaction.

C. Experimental Design

C.1. Selection of TAD constructs

Nonspecific Binding of Gal4 (1-100) to KIX

Initial plans involved using a similar setup as outlined in Chapter 2, where Gal4 (1-100) would be fused to the TADs of MLL, pKID and c-Myb, and fluorescence intensity of the Gal4-bound DNA would be monitored upon interaction with KIX. However our initial studies have determined that Gal4 (1-100) alone binds to KIX with an affinity of approximately 3-14 μM (Figure 4.4). As it is challenging to express Gal4 (1-100) in high quantity without solubility tags, both MBP-Gal4(1-100) and GB1-Gal4(1-100) were expressed and both exhibited affinity for KIX by fluorescent anisotropy binding assays, ruling out the possibility that the solubility tag contributed to KIX affinity. Gal4 (1-147) was also tested for KIX affinity and showed similar affinity to KIX as Gal4 (1-100). Hence we were not able to use this fusion construct to study the binding kinetics of various TADs to KIX.

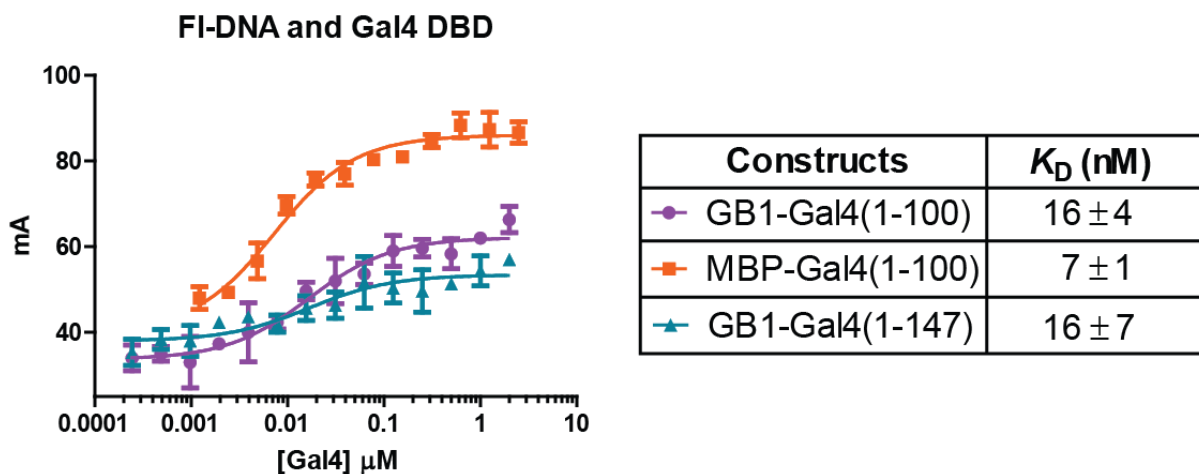


Figure 4.3 Graph of a binding experiment of three Gal4 constructs binding to fluorescein-DNA as a control of the integrity and structure of Gal4 DBD. K_D values were similar to each other and similar to previous experiments as shown in Chapter 2. All measurements were performed in triplicate and the error bars indicate standard deviation (SD). K_D was fitted using GraphPad Prism software.

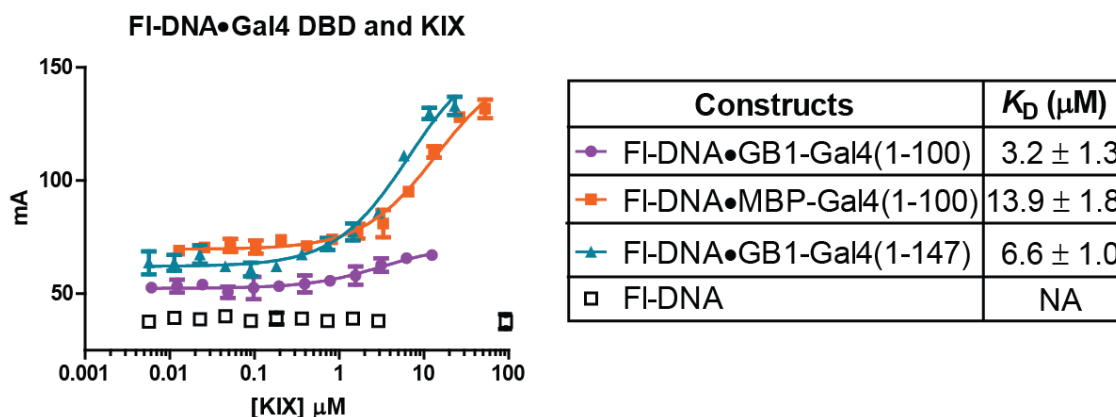


Figure 4.4 Graph of a binding experiment of three Gal4 constructs in complex with fluorescein-DNA binding to KIX. All measurements were performed in triplicate and the error bars indicate standard deviation (SD). K_D was fitted using GraphPad Prism software.

Use of fluorescently labeled and unlabeled TAD peptides in binding experiments.

As the Gal4 (1-100)-TAD fusion construct was not feasible to study TAD-KIX interactions, peptides derived from the TADs of MLL and pKID were used to monitor the activator-coactivator interaction,^{7,8} and fluorescein isothiocyanate (FITC) was coupled to the N-terminus of the peptides via a β -Alanine linker (Figure 4.5).



Figure 4.5 a) Sequences of MLL and pKID TADs used in peptide. (* denotes a phosphorylated Serine). **b)** Structure of fluorescein isothiocyanate, where the isothiocyanate group attached to fluorescein enables the coupling of FITC to amino acids via an amide bond. FITC was placed at the N-terminal of both peptides.

The FITC molecule was placed at the N-terminal of both peptides. A closer look of the structure of both MLL and pKID bound to KIX (Figure 4.6) shows that the position of the fluorophore will be far away from the binding interface and it will point outward towards

solution. This suggests that the fusion of a hydrophobic fluorophore at this position might minimally perturb the binding properties of the peptide TADs. The binding affinity of these FITC-TADs to KIX show a similar trend as values published by other groups (Figure 4.12). Moreover, these results are also within 10-fold difference of the values obtained by isothermal titration calorimetry (ITC) without any tags on the components.^{7,9,18} Nonetheless, to avoid any inconsistencies, the absolute kinetics values obtained in this work are only compared with other values that are obtained with FITC-TADs.

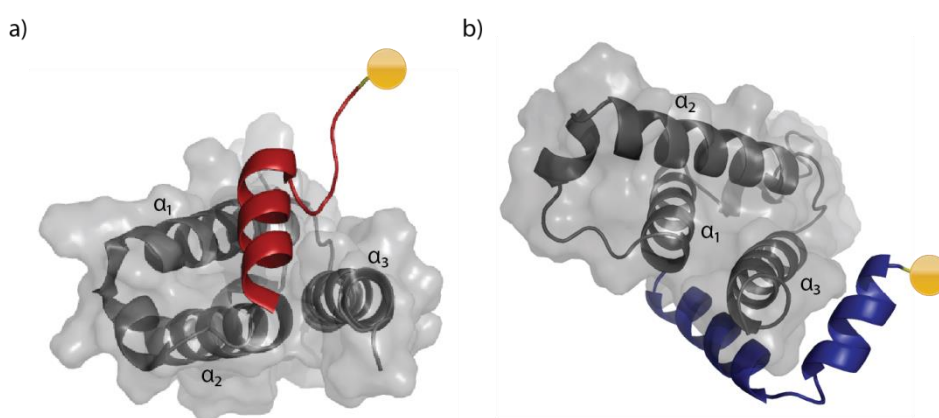


Figure 4.6 Position of FITC relative to the TAD-KIX complex. **a)** FITC (yellow circle) at N-terminus (labeled yellow) of MLL (red) in the MLL-KIX (grey)-c-Myb complex (PDB 2AGH). **b)** FITC (yellow circle) at N-terminus (labeled yellow) of pKID (blue) in the pKID-KIX (grey) complex (PDB 1KDX). Helices α_1 , α_2 and α_3 are labeled for clarity.

C.2. Complexes of KIX used in binding experiments.

Apart from free KIX and KIX in complex with excess amounts of MLL peptide, additional complexes of KIX mutants tethered to small molecule fragments were also used to study the kinetics of allosterically perturbed KIX binding to pKID.

The cysteine mutants KIX N627C and KIX L664C (Figure 4.8) were used in this study based on their high affinity to fragment **1-10** (as quantified by Dose Response DR_{50} , see Chapter 3). In addition, KIX L664C also has affinity to a larger fragment **2-64** (Figure 4.7). The integrity

of these two mutants compared to KIX wt was determined by circular dichroism scans (data not shown) and binding affinity to MLL and pKID as shown in Figure 4.12 .

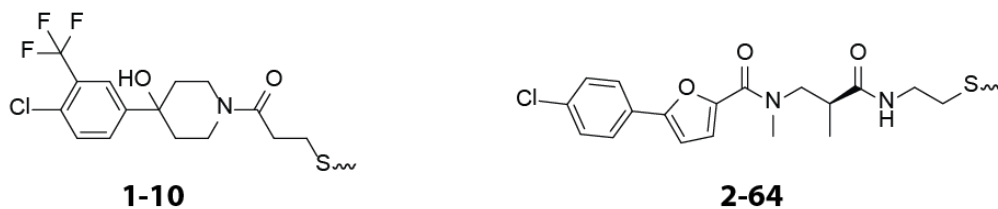


Figure 4.7 Structures of fragments **1-10** and **2-64** that emerged from a disulfide screening library as tethering fragments to KIX at the MLL binding site.

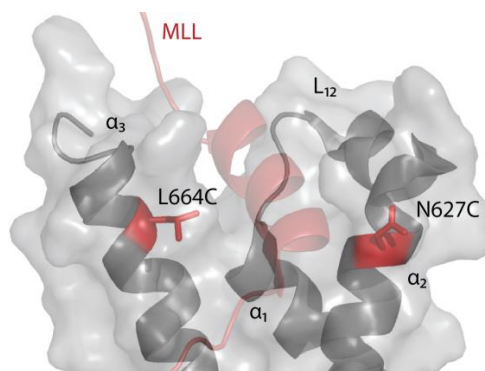


Figure 4.8 Position of mutations L664C and N627C on KIX (grey) and their relative orientation with MLL (red). Modified from PDB 2AGH. Helices α_1 , α_2 and α_3 and loop L₁₂ are labeled for clarity.

D. Results

D.1. Stopped-flow of MLL and pKID binding KIX

The transient-state binding kinetics between KIX and MLL or pKID was examined by fluorescence stopped-flow.

25 nM of pKID-FITC was rapidly mixed with either KIX storage buffer (10 mM Sodium Phosphate, 100 mM NaCl, 0.01 % NP-40, 10 % Glycerol, pH 6.8) (Figure 4.9a) to control for potential photo bleaching of the fluorophore or 0.1mg/ml of BSA in KIX storage buffer (Figure 4.9b) to control for effects of nonspecific binding and crowding of protein. Both control

experiments revealed no change in fluorescence intensity over a 120 second time period. Similar controls were carried out for MLL-FITC with similar results (data not shown).

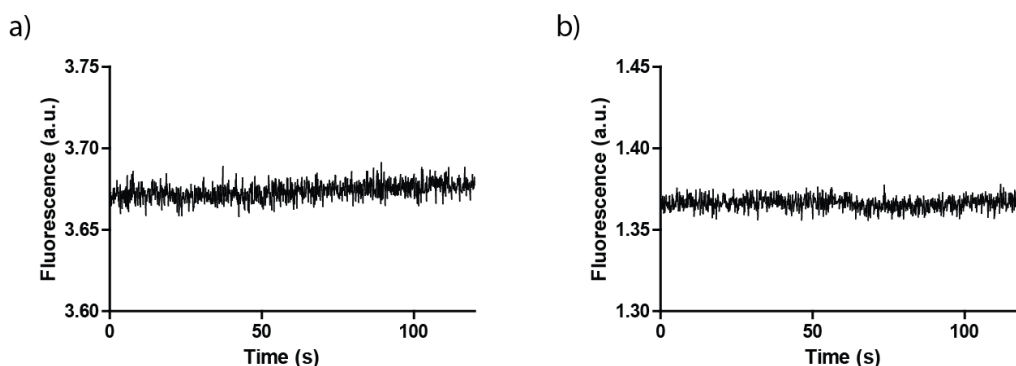


Figure 4.9 Stopped-flow traces of control experiments **a)** 25 nM pKID-FITC rapidly mixed with KIX storage buffer and **b)** 25 nM pKID-FITC rapidly mixed with 0.1 mg/ml BSA in KIX storage buffer. Both spectra obtained by excitation at 493 nm and emission collected with a 510 (LP) filter. Both spectra are an average of 5-8 separate traces.

25 nM of MLL-FITC or pKID-FITC in KIX storage buffer was rapidly mixed with excess amount of KIX complexes at varying concentrations of 0.25-5 μM in KIX storage buffer. The resulting change in fluorescence intensity (Figure 4.10) was monitored over several time periods varying from 0.05 seconds to 120 seconds. Two phases were observed for both MLL and pKID binding to KIX. However the faster phase of the MLL-KIX interaction took much less time to complete (0.05 seconds) as compared to the slower phase (120 seconds), so two different time domains were used to monitor the interaction, and each trace was fit to a monophasic equation. It was possible to record both time domains for the pKID-KIX interaction in one single trace, which were then fit to a biphasic equation to obtain k_{obs} (See Experimental methods section for detailed description).

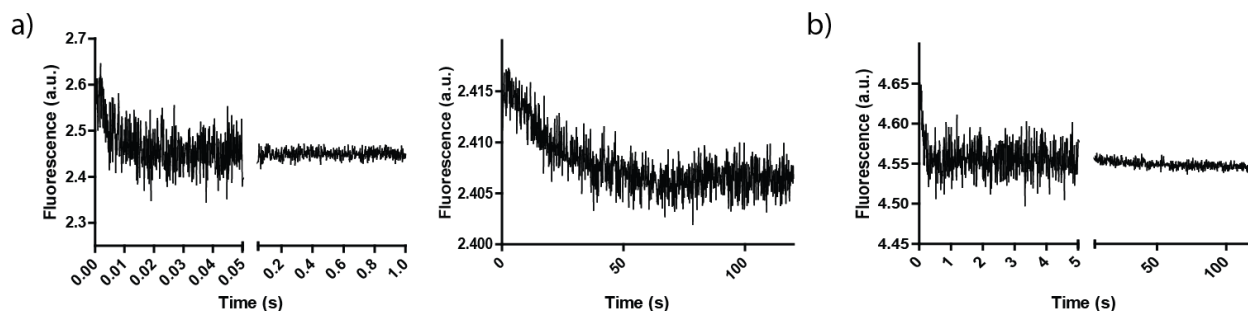


Figure 4.10 Representative stopped-flow traces of **a)** association of MLL with KIX, a shorter time domain to observe the fast phase (0.05 sec) and a longer time domain to monitor the slow phase (120 sec) and **b)** association of pKID with KIX, using a split time domain to capture both the fast phase (5 sec) and the slow phase (120 sec). All spectra were obtained by excitation at 493 nm and emission collected with a 510 (LP) filter. All spectra are an average of 3-5 separate traces.

Both peptides bind to KIX via a two-step mechanism, similar to observations made in previous studies (chapter 2) of Gal4, Gcn4 and VP16 TADs binding to Med15.¹ The observed rate of the fast phase linearly correlates with KIX concentration, suggesting a bi-molecular association step, while the observed rate of the slower phase does not exhibit positive linear correlation to KIX concentration (Figure 4.11). The slope value from linear regression fits of observed rate constant plotted over KIX concentration are the on-rates (k_{on}) of the association step. The on-rate of the association step of pKID binding to KIX ($k_{on} = 14.8 \pm 1.3 \mu\text{M}^{-1}\text{s}^{-1}$) agrees with values obtained by NMR relaxation dispersion in literature ($6.3 \mu\text{M}^{-1}\text{s}^{-1}$).² The on-rate of MLL was ten-fold faster than that of pKID ($k_{on} = 126 \pm 8 \mu\text{M}^{-1}\text{s}^{-1}$). Further studies will focus on the fast, bi-molecular association phase of the pKID-KIX interaction.

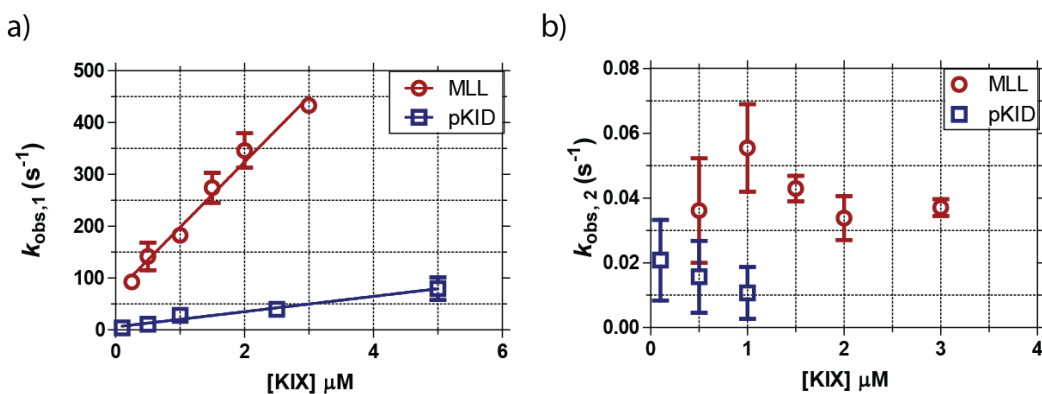


Figure 4.11 Observed association rate of **a)** the fast phase ($k_{obs,1}$) and **b)** the slow phase ($k_{obs,2}$) of the MLL/pKID-KIX interaction plotted against KIX concentration. The slow phase was not observable at high KIX concentrations in the pKID-KIX interaction. The error bars represent the error of the fitted k_{obs} values for the average (of 3-5) traces per KIX concentration point.

D.2. Cooperativity of MLL and pKID binding to KIX

In agreement with literature,^{7,9} fluorescent anisotropy experiments have shown that KIX will bind to the pKID peptide with higher affinity if it is pre-complexed with excess amounts of MLL, and vice versa (Figure 4.12a). By incubating KIX with various molar equivalents of either MLL or pKID unlabeled peptide and monitoring the dissociation constant (K_D) of this complex with a FITC (fluorescein isothiocyanate)-labeled pKID or MLL peptide, we can conclude that 4 molar equivalents of MLL or pKID peptide pre-complexed with KIX is enough to elicit the approximate maximum response of binding cooperativity for the complementary TAD (Figure 4.12b).

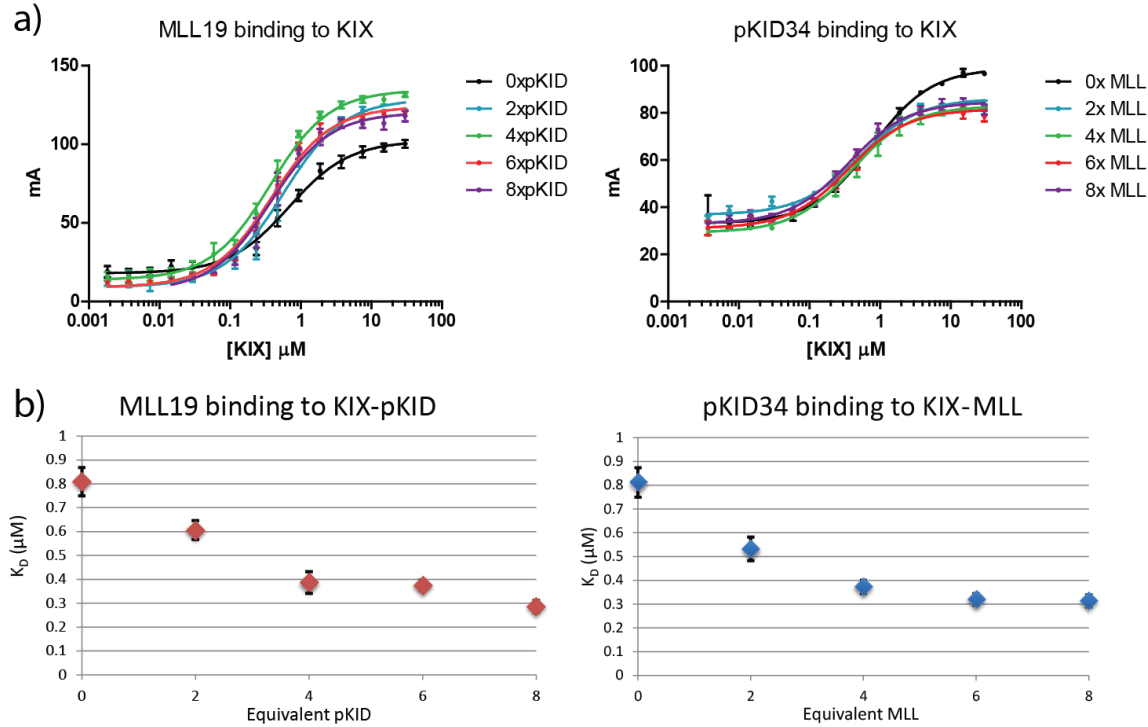


Figure 4.12 Cooperativity between MLL and pKID binding to KIX **a)** Anisotropy binding curves of KIX binding to either MLL or pKID when pre-complexed with indicated molar equivalents of the other TAD peptide. All measurements were performed in triplicate and the error bars indicate standard deviation (SD). **b)** Graph of K_D s from GraphPad Prism fits of curves in a) plotted over molar equivalents of the other TAD peptide shows the extent of cooperativity is “saturated” when approximately 4 molar equivalents of both MLL and pKID is pre-complexed with KIX, error bars represent the standard error of non-linear fit in Prism.

D.3. The allosteric effect of **1-10** tethering on the pKID binding affinity of different KIX mutants

To assess the allosteric effect of **1-10** tethering to either KIX N627C or KIX L664C, the equilibrium dissociation constants of **1-10**—KIX N627C and **1-10**—KIX L664C and the pKID-FITC peptide were recorded. pKID binds to the opposite allosteric site of KIX that **1-10** is not in direct contact with. Equilibrium constants were obtained by fluorescent anisotropy binding assays. These dissociation constants were compared to that of two other KIX complexes: free KIX mutants with pKID-FITC, and when 4 molar equivalents of MLL were pre-complexed with KIX mutants (Figure 4.13, Table 4.1).

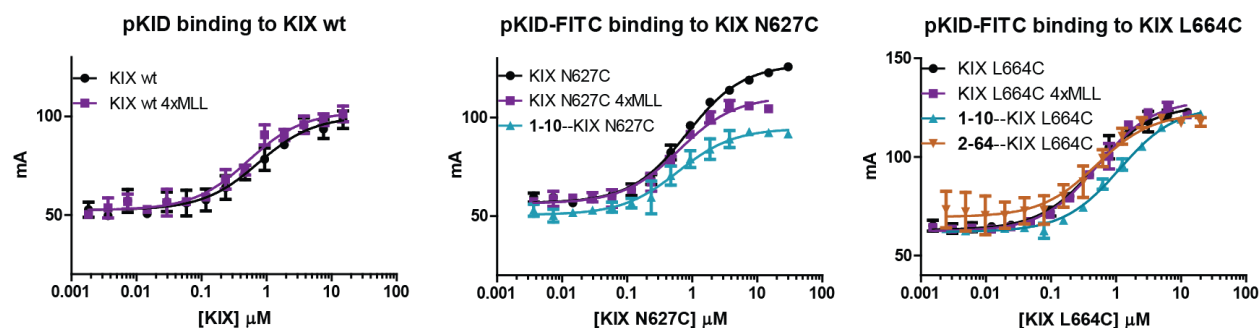


Figure 4.13 Anisotropy binding curves of complexes of KIX wt, KIX N627C and KIX L664C binding to pKID. All measurements were performed in triplicate and the error bars indicate standard deviation (SD) of the three measurements.

Table 4.1 Summary of K_D values from GraphPad Prism fits of binding curves in Figure 4.13. Errors reflect the standard error of non-linear fits in Prism.

KIX construct	Free	with 4xMLL	tethered to 1-10	tethered to 2-64
KIX wt	$0.72 \pm 0.14 \mu\text{M}$	$0.50 \pm 0.08 \mu\text{M}$	NA	NA
KIX N627C	$0.82 \pm 0.05 \mu\text{M}$	$0.62 \pm 0.08 \mu\text{M}$	$0.60 \pm 0.09 \mu\text{M}$	NA
KIX L664C	$0.46 \pm 0.04 \mu\text{M}$	$0.50 \pm 0.04 \mu\text{M}$	$1.12 \pm 0.08 \mu\text{M}$	$0.49 \pm 0.11 \mu\text{M}$

In summary, the tethering of **1-10** to KIX N627C seems to have similar allosteric effects on KIX as that of 4 molar equivalents of MLL in terms of the increase of pKID affinity. This is highly intriguing that a small molecule can cause similar allosteric effects as a 19-residue peptide. **1-10** tethered to KIX L664C on the other hand abrogates pKID binding to KIX. Interestingly, while KIX N627C shows similar potentiation of pKID affinity by MLL as that of KIX wt, no such potentiation is observed when MLL is pre-complexed with the KIX L664C mutant. This suggests that certain components of the allosteric network might be disrupted upon the mutation at Leu 664. This observation is consistent with that of previous studies (dissertation work from Sven Brüsweiler) where the mutation of Ile 660 to Val, a residue close to Leu 664, resulted in no cooperative binding of pKID when pre-complexed with MLL.

D.4. Stopped-flow spectroscopy of different states of KIX binding to pKID

Stopped-flow fluorescence spectroscopy was applied to compare the transient-state kinetics effects of KIX complexed with either MLL or small molecule tethering fragments. Association and dissociation stopped-flow fluorescence assays were carried out by monitoring the fluorescence intensity of fluorescein isothiocyanate (FITC) at the N-terminus of the pKID peptide when pKID is either a) rapidly mixed with KIX (association) or b) competed off KIX with excess amounts of unlabeled pKID peptide (dissociation).

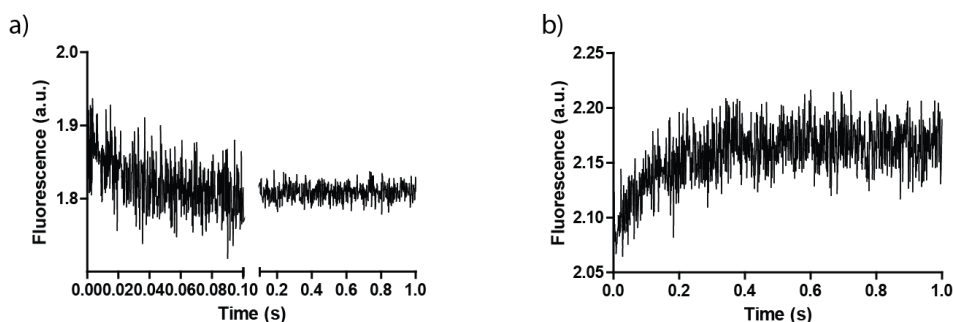


Figure 4.14 Representative stopped-flow traces of **a)** association of pKID with various KIX complexes and **b)** dissociation of pKID from KIX complexes by competition of excess amounts of unlabeled pKID peptide. All spectra were obtained by excitation at 493 nm and emission collected with a 510 (LP) filter. All spectra are an average of 5-8 separate traces.

For direct association binding assays, 25 nM of pKID-FITC (after mixing) was rapidly mixed with excess amount of KIX complexes (free protein, pre-complexed with 4 equivalents of MLL peptide, pre-tethered with **1-10** or **2-64**) at varying concentrations of 0.1-5 μM (after mixing). The resulting change in fluorescence intensity (Figure 4.14a) was monitored over several time periods varying from 0.05 seconds to 10 seconds and fit best to a single exponential equation to obtain k_{obs} (See Experimental methods section for detailed description).

For dissociation assays, 25 nM of pKID-FITC in KIX storage buffer was pre-equilibrated with 500 nM KIX complex and rapidly mixed with 12.5 μM (500 molar equivalents) of unlabeled pKID peptide in KIX storage buffer. The resulting change in fluorescence intensity

(Figure 4.14b) was monitored over several time periods varying from 0.1 second to 10 seconds and fit best to a single exponential equation to obtain k_{obs} (See Methods section for detailed description). The time domain of 1 second was selected for data analysis as it was closest to the predicted best time-frame by the fits in the Kintek software.

The observed rates obtained from fitting the association spectra were plotted against KIX concentration and displayed a linear relationship to KIX concentration (Figure 4.15). The slopes of the linear regression fits are the overall on-rate of pKID interacting with KIX as summarized in Table 4.2. The observed rates obtained from fitting the dissociation spectra are the off-rates of pKID being competed off KIX complexes, summarized in Table 4.3.

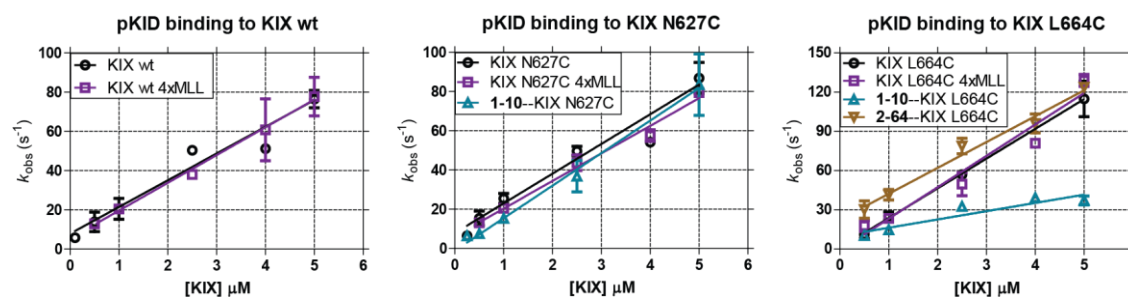


Figure 4.15 k_{obs} of pKID binding to KIX complexes plotted against KIX concentration for KIX wt, KIX N627C and KIX L664C. Each data point is an average of the k_{obs} of two separate experiments; each experiment is an average of five to eight traces. The error bars represent the standard deviation of the two separate k_{obs} values.

Table 4.2 On-rates of pKID binding to KIX complexes from the slopes of linear regression fits of Figure 4.15, errors are from the GraphPad Prism fits.

k_{on} ($\mu\text{M}^{-1}\text{s}^{-1}$)	Free	with 4xMLL	tethered to 1-10	tethered to 2-64
KIX	14 ± 1	14 ± 1	NA	NA
KIX N627C	15 ± 1	14 ± 1	17 ± 2	NA
KIX L664C	23 ± 2	24 ± 2	6 ± 1	20 ± 2

Table 4.3 Off-rates of pKID dissociating from KIX complexes. Errors are standard deviations of the average of two separate values.

k_{off} (s^{-1})	Free	with 4xMLL	tethered to 1-10	tethered to 2-64
KIX	7.0 ± 0.4	5.0 ± 1.3	NA	NA
KIX N627C	7.3 ± 1.3	4.1 ± 2.5	4.0 ± 0.8	NA
KIX L664C	8.7 ± 0.4	7.7 ± 0.3	13.8 ± 0.6	10.2 ± 3.5

Comparison of the on and off rates between pKID and various KIX complexes similar trends among KIX complexes between transient-state off-rates and equilibrium binding constants (Figure 4.16). Also of interest is while most of the on-rates are similar for all KIX complexes binding to pKID, **1-10**—KIX L664C seems to have an exceptionally slow on-rate. This could be related to the fact that **1-10** tethering is hindering the flexibility of the KIX protein and thus limiting the extent of conformational plasticity that is required for KIX to associate with the pKID peptide.

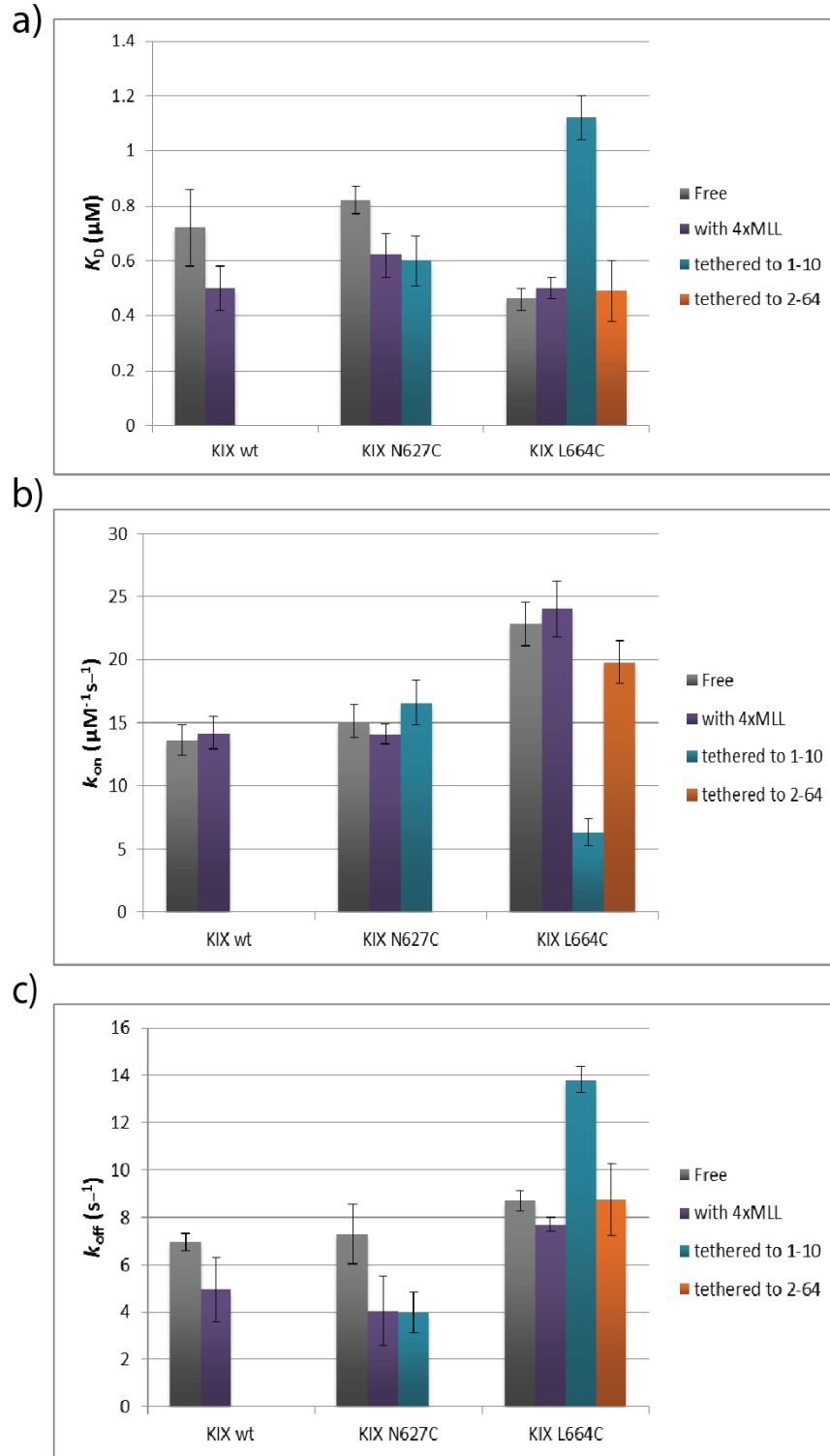


Figure 4.16 a) Bar graph of equilibrium binding constants of pKID binding to KIX complexes. b) Bar graph of on-rate of pKID binding to KIX complexes. c) Bar graph of off-rate of pKID dissociating from KIX complexes. The error bars represent standard deviation of the average of two values from separate experiments.

*D.5. Φ values for **1-10**—KIX N627C and **1-10**—KIX L664C suggest **1-10** elicits different allosteric effects when tethered at different positions*

Φ values were calculated for **1-10**—KIX N627C and **1-10**—KIX L664C, the two small molecule-KIX complexes that notably altered KIX affinity to pKID, to further investigate the observation that **1-10** is having different effects on different KIX mutants.

Φ value analysis was applied to examine the roles of specific amino acids during the transition state of protein folding.³⁹ It has been shown that this analysis can also be more broadly applied to study the transition state of protein-protein association.⁴⁰⁻⁴⁴ In Φ value analysis, the change in free energy for the binding association kinetics upon mutation ($\Delta\Delta G^\ddagger$) is related to the change in free energy for the overall binding reaction ($\Delta\Delta G^{eq}$), calculated by the following equations:

$$\Phi = \Delta\Delta G^\ddagger / \Delta\Delta G^{eq} \quad (\text{Eq. 4.1})$$

where

$$\Delta\Delta G^\ddagger = RT \ln\left(\frac{k_{on}^{WT}}{k_{on}^{mut}}\right) \quad (\text{Eq. 4.2}) \quad \text{and} \quad \Delta\Delta G^{eq} = RT \ln\left(\frac{K_D^{WT}}{K_D^{mut}}\right) \quad (\text{Eq. 4.3})$$

A Φ value of 1 means that the mutations/changes on the protein have an effect on the transition state of the interaction, indicating that the change in K_D is governed by the change in k_{on} . A Φ value of 0, on the other hand, means the mutations/changes have an effect after the transition state of the interaction, indicating that the change in K_D is governed by the change in k_{off} . While Φ value analysis is usually used to study the effect of single-residue mutations on protein-protein interactions,^{38,45,46} in this study we employ it to examine the effect of the tethered small molecule **1-10** on KIX-pKID interaction. The results are shown in Table 4.4. It shows **1-10**—KIX N627C with a low Φ value, suggesting the **1-10** tethering is having an effect on the

interaction after the initial bimolecular association step. **1-10**—KIX L664C however displays a much larger Φ value, indicative of effects occurring during the initial association step.

Table 4.4 Binding Φ values for 1-10 tethered KIX mutants. The “ $k_{\text{on}}^{\text{WT}}$ ” and “ K_{D}^{WT} ” used in the calculations for $\Delta\Delta G^{\text{eq}}$ and $\Delta\Delta G^{\ddagger}$ are the values of the corresponding mutants without 1-10 tethered.

Complex	$k_{\text{on}} (\text{s}^{-1}\mu\text{M}^{-1})$	$\Delta\Delta G^{\ddagger}$ (kcal mol ⁻¹)	$K_{\text{D}} (\mu\text{M})$	$\Delta\Delta G^{\text{eq}}$ (kcal mol ⁻¹)	Φ
1-10 —KIX N627C	16.6 ± 1.8	-0.056	0.60 ± 0.09	-0.19	0.3
1-10 —KIX L664C	6.3 ± 1.1	0.766	1.12 ± 0.08	0.53	1.4

E. Discussion and Conclusions

Overall, these transient-state kinetic analyses suggest the KIX-pKID interaction undergoes at least two steps: a fast bi-molecular association step and a slow conformational change step.

Using the available data it is not possible to distinguish if the conformational change step occurs

before or after the bi-molecular association step. However based on literature results^{1,2,22,25} and

for ease of discussion, the mechanism $Free \leftarrow \rightarrow Intermediate \leftarrow \rightarrow Bound$ will be employed to

demonstrate key conclusions. As shown earlier in Figure 4.2 and in literature², there might be an

even faster phase between the Free and Intermediate states, but similar to the NMR relaxation

dispersion data used in literature, this phase is not observable under stopped-flow conditions.

Hence the following discussion will address the $Free \leftarrow \rightarrow Intermediate$ overall interaction as one step.

The bar graphs of K_{D} and k_{off} values in Figure 4.16 (a) and (c) show a similar trend across different KIX complexes. This suggests that there is a positive correlation between the change in KIX’s affinity to pKID (K_{D}) and the dissociation rate of pKID from KIX (k_{off}). This phenomenon has been observed before in studies on both protein-protein interactions involving disordered proteins^{38,45} and protein-protein interactions involving structured proteins.³⁵ The way our stopped-flow experiment is set up means that k_{on} only reflects the on-rate of the bimolecular

association step, but k_{off} reflects the overall off-rate of pKID from KIX including both the conformational change step and the biomolecular dissociation step. The fact that k_{off} is governing K_D suggests that the changes in KIX (complex with MLL or small molecule) are affecting the slow conformational change step of the interaction by eliciting allosteric effects on the pKID binding site.

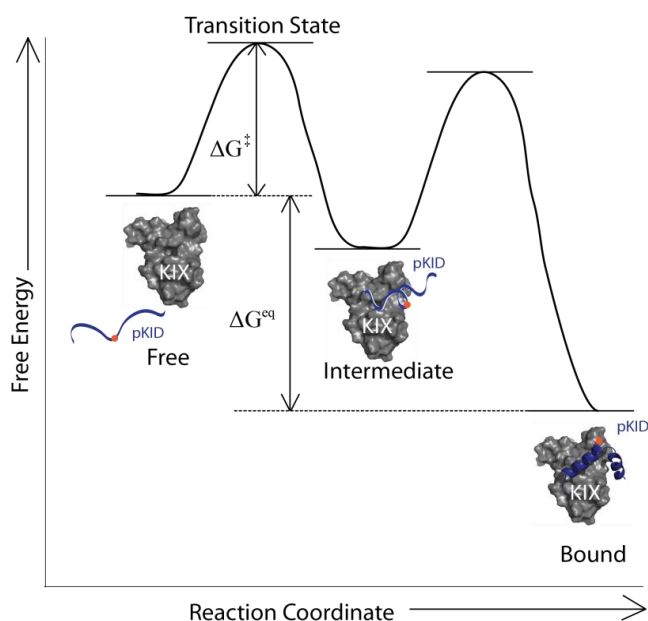


Figure 4.17 Energy diagram of the KIX-pKID interaction, depicting the free energy of association kinetics (ΔG^\ddagger) and the free energy of the overall binding interaction (ΔG^{eq}).

Equilibrium association values and Φ value analysis of **1-10** tethered to KIX N627C and KIX L664C binding to pKID show that the same small molecule is eliciting different allosteric effects on the distal pKID binding site. As mentioned in the results section and demonstrated in Figure 4.17, when Φ is close to 0, as in $\Delta\Delta G^\ddagger$ is much smaller than $\Delta\Delta G^{\text{eq}}$, it means that **1-10** tethering is affecting side chain involved in steps occurring after the transition state, and

vice versa when Φ is close to 1. The Φ value for **1-10**—KIX N627C is 0.3, which is in the low to mid range⁴⁵, whereas Φ for **1-10**—KIX L664C is 1.4, which is a much higher value. This is suggesting that when tethered to KIX N627C, **1-10** allosterically affects KIX amino acid side chains that are involved in the conformational change step in KIX-pKID binding, whereas it is affecting KIX side chains that are involved in the bimolecular association step of the KIX-pKID interaction when tethered to KIX L664C. Projecting further in the future, the fact that a same

small molecule can elicit entirely different and even opposite allosteric effects on the same binding site of a protein when tethered at different positions suggests the potential of using small molecules as a switch in cells. One could foresee changing the cellular environment, by pH, temperature or excitation wavelength, and being able to change to property of a small molecule to either up regulate or down regulate a certain pathway.

In conclusion, we have performed fluorescence stopped-flow assays to study the transient state kinetics of various KIX constructs binding to TADs pKID and MLL. We observed that MLL and pKID bind to KIX through a similar two-step mechanism as that of TADs binding to Med15: a rapid bimolecular association step and a slower conformational change step. We find that small molecule fragment **1-10** can elicit allosteric effects on KIX at the distal pKID binding site with the same order of magnitude as a known peptide ligand MLL. In addition we show that the off-rate (k_{off}) of these interactions is the driving force of the change in equilibrium binding affinity (K_D) between KIX and pKID. More importantly, we show that **1-10** can allosterically affect different sets of amino acids in KIX depending on the position of the tethering. One set of these residues are more involved in the association of KIX and pKID, while the other set is more involved in the conformational change step of KIX-pKID interaction. As a result the same molecule can either enhance or abrogate the affinity of KIX to pKID when at different positions, some of which are different effects from that elicited by the native peptide ligand MLL. These results provide insight into the exciting potential of using small molecules as powerful probes in dissecting the allosteric mechanisms of protein-protein interactions in ways native peptide ligands are incapable of.

F. Future Directions

F.1. Conformational analysis of small molecule tethered KIX

^1H - ^{15}N HSQC or even ^1H - ^{13}C HSQC NMR of both **1-10**—KIX L664C and **1-10**—KIX N627C as well as **2-64**—KIX L664C would provide more direct structural information on the allosteric effects of small molecules on KIX. Comparing the chemical shift of the tethered KIX constructs to free KIX would define which specific residues are perturbed by the small molecules in different positions. A complementary method would be to try and obtain crystal structures of **1-10**—KIX N627C and **2-64**—KIX L664C. While one would not be able to compare these structures to that of free KIX, one could however see the differences in side chain interactions between the different small molecules at different positions on KIX. This can identify in detail what residues are important for KIX-pKID association and what residues are important for the coupled folding conformational change occurring with binding.

F.2. FRET or fluorescence quenching stopped-flow

The fluorescence probe FITC used in this study displayed a change in fluorescence intensity upon TAD-KIX binding, likely due to changes in the local environment of the fluorophore. While we were able to glean kinetic constants from this change, we were not able to directly observe more details such as the difference in distance between TAD and KIX over time. Installing a small non-invasive FRET pair on the pKID or c-Myb TAD and KIX will enable higher resolution read-outs of the interaction as well as conformational change. Strategically placed small molecule FRET pairs such as Cy3 and Cy5 will directly reflect changes in distance by change in FRET efficiency and act as a molecular ruler.⁴⁷ Similarly, a fluorophore and a quenching molecule can achieve the same purpose. The real-time change in FRET or quench

efficiency will aid in accurately identifying which steps are association (large change in distance and efficiency) and which steps are conformational changes (smaller changes in distance and efficiency).⁴⁸⁻⁵¹ This will enable us to propose a more detailed mechanism scheme for TAD-KIX association and the effects of small molecule tethering to KIX.

G. Experimental Methods

Protein Expression and Purification

As previously described,¹⁴ the DNA sequence encoding the KIX domain of mouse CBP (586-672) was cloned into a pRSETB vector incorporating a hexahistidine tag with a short linker fused to the N-terminus of KIX for protein expression resulting in a protein with sequence (tag and linker residues are shown in lower case):

mrgshhhhhgmasGVRKGWHEHVTQDLRSHLVHKL VQAIPTPDPAALKDRRMENLV
AYAKKVEGDMYESANSRDEYYHLLAEKIYKIQKELEEKRRSRL

The cysteine mutant at KIX L664C was created using site-directed mutagenesis as previously described.¹⁴ For protein expression, the plasmid was transformed into Rosetta2 (DE3) pLysS *E. coli* (Novagen) and grown in LB media. After an O.D.₆₀₀ of 0.8 was reached (37 °C, 250 rpm), the cultures were cooled to 25 °C and expression was induced with 0.5 mM IPTG for 4 h (250 rpm). The His-tagged protein was isolated using Ni-NTA beads (Qiagen) according to the manufacturer's instruction and eluted with 400 mM imidazole. Final purification was carried out by ion-exchange column chromatography on a Source S column, (GE Healthcare) in 50 mM phosphate buffer, 1 mM DTT, pH 7.2 eluting with increasing concentrations of NaCl (0-1 M NaCl). Purified protein solutions were buffer exchanged into 10 mM sodium phosphate, 100

mM NaCl, 10 % glycerol 0.01 % NP40, pH 6.8 using a PD-10 column (GE-Healthcare) and stored at -80 °C.

Peptide Synthesis and Purification

All peptides were synthesized by standard N-9-Fluorenylmethoxycarbonyl (Fmoc) solid phase synthesis methods¹ on CLEAR amide resin (Peptide International, 0.48 mmol/g). In the case of MLL15 and Myb25, a c-terminal tyrosine or tryptophan were added respectively to facilitate concentration determination. All peptides were cleaved from the solid support in a mixture of 95/2.5/2.5 trifluoroacetic acid (TFA)/triisopropylsilane/water for 3-4 hrs followed by evaporation of solvent under a nitrogen stream except for Myb25 and FI-Myb25 which were cleaved in 92.5% TFA 2.5% thioanisole and 2.5% ethanedithiol followed by addition of TMS-bromide during the last 15 minutes of cleavage to prevent methionine oxidation. The crude peptides were precipitated into cold ether and purified by reverse phase HPLC on a Waters C18 column using water with 0.1% TFA as the A solvent and CH₃CN as the B solvent. Product molecular weight was confirmed by electrospray ionization mass spectrometry (ESI-MS) using a Micromass LCT Time-of-Flight mass spectrometer.

Peptide	Sequence	ESI-MS
Mll19	AcWADAGILPSDIMDFVLKNTYCONH ₂	Calculated M-1: 2310 Experimental ES- m/1 = 2309.4
FI-Mll19	FITC-beta-ADAGNILPSDIMDFVLKNTPCONH ₂ ,	Calculated M-1: 2517 Experimental ES-: m/2 = 1285.7 (m-1; 2516)
pKID29	AcTDSQKRREILSRPS(Phos)YRKILNDLSSDAPGCONH ₂	Calculated M+1:3480 Experimental ES+: m/3 = 1161.1 (m+1; 3481)

Fl-pKID29	FITC-AEEA-TDSQKRREILSRP(S)(Phos)YRKILNDLSSDAPGCONH ₂	Calculated M+1: 3972 Experimental ES+: m/4 = 994.0 (M+1; 3973)
Myb25	Ac-KEKRIKELELLLMSTENELKGQALWCONH ₂	Calculated M-1: 3169 Experimental ES-: m/2 = 1584.3 (M-1; 3168)
Fl-Myb25	FITC-beta-AKEKRIKELELLLMSTENELKGQALWCONH ₂	Calculated M-1: 3587 Experimental ES-: m/2 = 1792.5 (M-1;3586); ES-m/3 = 1195.0 (M-1; 3587)
All single letter amino acid abbreviations are use unless indicated: beta-A (beta-alanine), AEEA (2-(2-amino-ethoxy)-ethoxy acetic acid, S (phos) (phosphoserine).		

Table 4.5 Peptide sequence and ESI-MS characterization

Fragment Tethering

KIX L664C or KIX N627C was incubated with 10 equivalents of **1-10** or **2-64** in 50 mM phosphate buffer at pH 6.8 overnight, excess small molecule was removed by NAP-5 desalting column and buffer exchanged into 20 mM HEPES, 100 mM NaCl, pH 8.0. Small molecule tethering efficiency was confirmed by Q-TOF LC-MS (Agilent) and tethering reached at least 95% complete. Protein-small molecule complexes were flash frozen in liquid N₂ and stored at -80°C.

Fluorescent Anisotropy Assays

The fluorescent anisotropy assays were done in triplicate with a final sample volume of 10 µL in a low volume, non-binding, black, 384-well plate (Corning), and read using a Tecan Genios Pro plate reader with polarized excitation at 485 nm and emission intensity measured through a parallel and perpendicularly polarized 535 nm filter. FITC (Fluorescein isothiocyanate) labeled peptides were diluted in storage buffer (10 mM Phosphate, 100 mM NaCl, 0.01 % NP-40, 10 % Glycerol, pH 6.8) to a concentration of 25 nM. Then 10 µL of the peptide solution was added to a series of 50 µL solutions of varying KIX concentrations in storage buffer to obtain the

final concentrations of up to 20 μM . The samples were incubated for 30 min at room temperature before the degree of fluorescence anisotropy was measured (Tecan Genios Pro). Anisotropy data was corrected for change in fluorescence intensity using Equation 4.4.

$$f_B = \frac{r - r_F}{R(r_B - r) + r - r_F} \quad (\text{Eq. 4.4})$$

A binding isotherm that accounts for ligand depletion (assuming a 1:1 binding model of peptide to GACKIX) was fit to the observed anisotropy values as a function of KIX to obtain the apparent equilibrium dissociation constant, K_D :

$$y = c + (b - c) \times \frac{[(K_D + a + x) - \sqrt{(K_D + a + x)^2 - 4ax}]}{2a} \quad (\text{Eq. 4.5})$$

where “a” and “x” are the total concentrations of fluorescent peptide and KIX, respectively, “y” is the observed anisotropy at any KIX concentration, “b” is the maximum observed anisotropy value, and “c” is the minimum observed anisotropy value. Each data point in figures in Results section is an average of three independent experiments with the indicated error (standard deviation). Data analysis was performed using GraphPad Prism 5 software.

Fluorescence Stopped-Flow Kinetic Assays

Stopped-flow experiments were performed on a KinTek model SF-2001 stopped-flow equipped with a 75W Xe arc lamp in two-syringe mode. FITC was excited at 493 nm and its emission was monitored at wavelengths > 510 nm using a long-pass filter (Corion).

Association Experiments: 25 nM (final concentration after mixing) of pKID-FITC in KIX storage buffer was rapidly mixed with excess amount of KIX complexes (free protein, pre-complexed with 4 equivalents of MLL peptide, pre-tethered with **1-10** or **2-64**) at varying

concentrations of 0.1-5 μM (final concentration after mixing) in KIX storage buffer at 25°C. The time domains (0.1 and 1 second or 0.05 and 0.5 second for fast phase and 120 sec for slow phase) were selected for data analysis as they were closest to the predicted best time-frame by the fits in the Kintek software.

Dissociation Experiments: 25 nM (final concentration after mixing) of pKID-FITC in KIX storage buffer was pre-equilibrated with 500 nM KIX complex (final concentration after mixing) and rapidly mixed with 12.5 μM (500 molar equivalents, final concentration after mixing) of unlabeled pKID peptide in KIX storage buffer at 25°C.

All kinetic traces reported are an average of five to eight independent determinations. Sum of exponentials was fit to the transient kinetic time courses, $F(t)$ as in Equation 4.6, to obtain the fluorescence amplitude (A) and the observed rate, k_{obs} , for each exponential phase where $F(0)$ is the initial fluorescence intensity, and t , time:

$$F(t) = \sum A_n [1 - \exp(-k_{\text{obs},n}t)] + F(0) \quad (\text{Eq. 4.6})$$

Two control experiments were performed to ensure that the fluorescence changes are from a KIX-pKID interaction: 25 nM of pKID-FITC was rapidly mixed with either KIX storage buffer (10 mM Sodium Phosphate, 100 mM NaCl, 0.01 % NP-40, 10 % Glycerol, pH 6.8) or 0.1mg/ml of BSA in KIX storage buffer at 25°C. Both control experiments exhibited no time-dependent change in fluorescence intensity over a 10 second time period.

Analysis of the time courses was performed using Kintek software, and the reported errors are the asymptotic standard errors. The dependence of the observed rates on KIX concentration was analyzed using GraphPad Prism 4.0 software.

H. References

1. 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).
2. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
3. Huang, Y. & Liu, Z. Smoothing molecular interactions: The “kinetic buffer” effect of intrinsically disordered proteins. *Proteins* **78**, 3251–9 (2010).
4. Huang, Y. & Liu, Z. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the “fly-casting” mechanism. *J. Mol. Biol.* **393**, 1143–59 (2009).
5. Zhou, H.-X., Pang, X. & Lu, C. Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *Phys. Chem. Chem. Phys.* **14**, 10466–10476 (2012).
6. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
7. 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**, 1005–13 (2006).
8. 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).
9. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *J. Biol. Chem.* **277**, 43168–74 (2002).
10. Cook, P. R., Polakowski, N. & Lemasson, I. HTLV-1 HBZ protein deregulates interactions between cellular factors and the KIX domain of p300/CBP. *J. Mol. Biol.* **409**, 384–98 (2011).
11. Best, J. L. *et al.* Identification of small-molecule antagonists that inhibit an activator:coactivator interaction. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17622–7 (2004).
12. 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).
13. Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9367–9372 (2000).
14. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **7**, 1345–50 (2012).
15. Wang, N. & Majmudar, C. Ordering a dynamic protein via a small molecule stabilizer. *J. Am. Chem. Soc.* (2022).doi:10.1021/ja3122334
16. 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 (2001).
17. Brüschweiler, S. *et al.* Direct observation of the dynamic process underlying allosteric signal transmission. *J. Am. Chem. Soc.* **131**, 3063–8 (2009).
18. Arai, M., Dyson, H. J. & Wright, P. E. Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. *FEBS Lett.* 4500–4504 (2010).

19. Matsuno, H., Furusawa, H. & Okahata, Y. Kinetic study of phosphorylation-dependent complex formation between the kinase-inducible domain (KID) of CREB and the KIX domain of CBP on a quartz crystal microbalance. *Chem.--Eur. J.* **10**, 6172–8 (2004).
20. Parker, D. *et al.* Analysis of an Activator:Coactivator Complex Reveals an Essential Role for Secondary Structure in Transcriptional Activation. *Mol. Cell* **2**, 353–359 (1998).
21. Parker, D. *et al.* Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism. *Mol. Cell. Biol.* **16**, 694–703 (1996).
22. 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).
23. Zor, T., Mayr, B. M., Dyson, H. J., Montminy, M. R. & Wright, P. E. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J. Biol. Chem.* **277**, 42241–8 (2002).
24. Dal Peraro, M. Ser133 phosphate-KIX interactions in the CREB-CBP complex: *Eur. Biophys. J.* **75** (2001).
25. Turjanski, A. G., Gutkind, J. S., Best, R. B. & Hummer, G. Binding-induced folding of a natively unstructured transcription factor. *PLoS Comput. Biol.* **4**, e1000060 (2008).
26. Radhakrishnan, I., Pérez-Alvarado, G. C., Dyson, H. J. & Wright, P. E. Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett.* **430**, 317–322 (1998).
27. Shoemaker, B. a, Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8868–73 (2000).
28. Parker, D. *et al.* Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol.* **19**, 5601–7 (1999).
29. Rezaei-Ghaleh, N., Blackledge, M. & Zweckstetter, M. Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *Chembiochem* **13**, 930–50 (2012).
30. Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).
31. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–64 (2010).
32. Lacy, E. R. *et al.* p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol.* **11**, 358–64 (2004).
33. Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M.-F. & Ahmadian, M. R. An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Mol. Cell* **20**, 313–24 (2005).
34. Zhou, H.-X. Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem. Sci.* **37**, 43–8 (2012).
35. Prakash, M. K. Insights on the role of (dis)order from protein-protein interaction linear free-energy relationships. *J. Am. Chem. Soc.* **133**, 9976–9 (2011).
36. Panayotou, G. *et al.* Interactions between SH2 Domains and Tyrosine-Phosphorylated Platelet-Derived Growth Factor , B-Receptor Sequences : Analysis of Kinetic Parameters by a Novel Biosensor-Based Approach. *Mol. Cell. Biol.* **13**, 3567–3576 (1993).

37. Umulis, D., O'Connor, M. B. & Blair, S. S. The extracellular regulation of bone morphogenetic protein signaling. *Development* **136**, 3715–28 (2009).
38. Haq, S. R. *et al.* Side-chain interactions form late and cooperatively in the binding reaction between disordered peptides and PDZ domains. *J. Am. Chem. Soc.* **134**, 599–605 (2012).
39. Fersht, A. R. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3–9 (1997).
40. Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev. (Washington, DC, U. S.)* **109**, 839–60 (2009).
41. Taylor, M. G., Rajpal, A. & Kirsch, J. F. Kinetic epitope mapping of the chicken lysozyme.HyHEL-10 Fab complex: delineation of docking trajectories. *Protein Sci.* **7**, 1857–67 (1998).
42. Mateu, M. G., Sánchez Del Pino, M. M. & Fersht, A. R. Mechanism of folding and assembly of a small tetrameric protein domain from tumor suppressor p53. *Nat. Struct. Mol. Biol.* **6**, 191–8 (1999).
43. Wu, L. C., Tuot, D. S., Lyons, D. S., Garcia, K. C. & Davis, M. M. Two-step binding mechanism for T-cell receptor recognition of peptide MHC. *Nature* **418**, 552–6 (2002).
44. Kiel, C., Selzer, T., Shaul, Y., Schreiber, G. & Herrmann, C. Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9223–8 (2004).
45. 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**, 34316–24 (2012).
46. Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G. & Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3952–7 (2011).
47. Abelson, J. *et al.* Conformational dynamics of single pre-mRNA molecules during in vitro splicing. *Nat. Struct. Mol. Biol.* **17**, 504–12 (2010).
48. Gakamsky, D. M. *et al.* Kinetic evidence for a ligand-binding-induced conformational transition in the T cell receptor. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16639–16644 (2007).
49. Burton, R. L., Chen, S., Xu, X. L. & Grant, G. A. Transient kinetic analysis of the interaction of L-serine with Escherichia coli D-3-phosphoglycerate dehydrogenase reveals the mechanism of V-type regulation and the order of effector binding. *Biochemistry* **48**, 12242–51 (2009).
50. Chen, Y. & Erickson, H. P. Rapid in vitro assembly dynamics and subunit turnover of FtsZ demonstrated by fluorescence resonance energy transfer. *J. Biol. Chem.* **280**, 22549–54 (2005).
51. Liu, Y. & Schepartz, a Kinetic preference for oriented DNA binding by the yeast TATA-binding protein TBP. *Biochemistry* **40**, 6257–66 (2001).

Chapter 5 Conclusions and Future Directions

A. Conclusions

In this dissertation, the experiments have resulted in furthering our knowledge of the mechanism of transcriptional activator-coactivator interactions, a process that has been studied for several decades but is still not completely understood. Particularly, various biochemical and biophysical methods were applied to the study of two different activator-coactivator systems: the TADs VP16, Gal4 and Gcn4 and their shared coactivator Med15, and a more conformationally defined system which is the TADs MLL and pKID and their shared coactivator the KIX domain of CBP. We have detailed the transient-state kinetics characteristics of these activator-coactivator interactions, and have determined the effects on the activator-coactivator interaction and coactivator conformation when covalently tethering a small molecule to the coactivator.

A.1. Transient-state kinetic mechanisms of activator-coactivator interactions

In Chapters 2 and 4, transient state kinetic analysis has established that the mechanisms for both the TAD (VP16, Gal4, Gcn4)-Med15 interaction and the TAD (MLL, pKID)-KIX interaction are at least biphasic, involving a fast, bimolecular association step and a slow, conformational change step (Figure 5.1).

We have compared the kinetic properties of different TADs (VP16, Gal4 and Gcn4) binding to the same coactivator (Med15)¹ in Chapter 2 and different complexes of a coactivator (KIX with MLL or small molecules) binding to the same TAD (pKID) in Chapter 4.

In both cases, there is a strong correlation between the overall binding affinity and the slow conformational change step. This suggests that conformational flexibility plays a key role in the ability of both activators and coactivators to distinguish between different binding partners by varying the overall binding affinity. There has been similar observations in literature regarding the multispecificity of hub transcription factors by examining their structural characteristics.²⁻⁷ In this study, a unique kinetic approach has been taken to address and confirm such a correlation. The further understanding of this interaction mechanism will enable better guided designs of small molecule modulators of activator-coactivator interactions in the future. By tuning the extent of flexibility of a small molecule we might be able to control the binding affinity of that molecule to transcriptional activators or coactivators, hence control the potency of the small molecule.

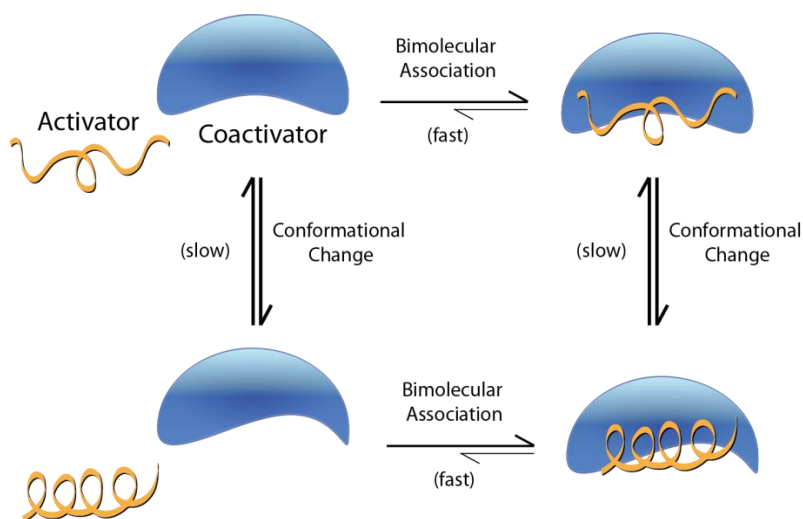


Figure 5.1 Schematic of the activator-coactivator interaction mechanism. A fast bimolecular association step is coupled with a slower conformational change step.

A.2. Allosteric effects of a small molecule modulator on activator-coactivator interactions

The work outlined in this dissertation has also demonstrated the potential of using a small molecule modulator to probe transcriptional activator-coactivator interactions. A small molecule

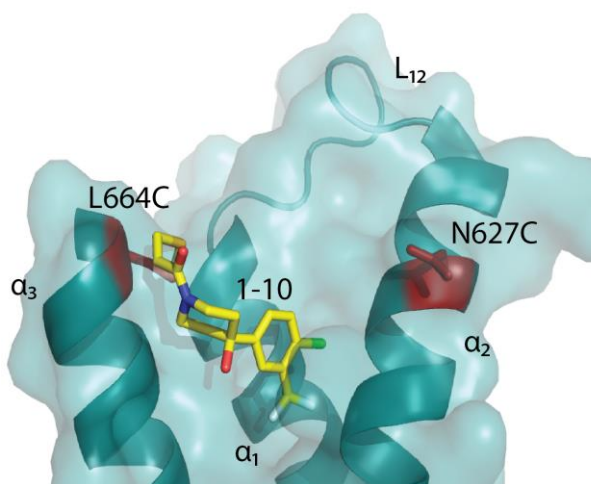


Figure 5.2 Structure of **1-10**—KIX L664C (PDB 4I9O). N627C is labeled as well. This snapshot shows that the two mutations, L664C and N627C are at opposite sides of the MLL binding site and on different helices. This possibly contributes to the different effects **1-10** has when tethered to the different mutants.

fragment **1-10** emerged from a Tethering screen against the MLL binding site of KIX. In Chapter 3, we were able to obtain a homogenous crystal of **1-10**—KIX L664C, and solved the first crystal structure of the KIX domain of CBP.⁸

This enabled us to have a closer look at various side chain orientations and examine the structural shifts in KIX, specifically the loop region, in its different complexes. In Chapter 4, we have also

been able to use **1-10** to allosterically

modulate the affinity of KIX to pKID at a distal binding site, and used this as a tool to study the correlation between transient state kinetics and equilibrium binding affinity as outlined above.

More importantly, we discovered that the same small molecule **1-10** can have different allosteric effects depending on what position it is tethered to on KIX (Figure 5.2). While **1-10** tethering to KIX N627C seemed to affect the residues governing conformational change after pKID binding (similar to the effect of MLL), **1-10** tethering to KIX L664C affected residues governing the initial bimolecular association of pKID to KIX. As a result the same molecule can either enhance or abrogate the affinity of KIX to pKID when at different positions. This enabled us to use **1-10** to examine allosteric effects on KIX that MLL was not able to elicit. Furthermore, this provides important insight into the possible consequences of the application of small molecule modulators. It reveals the potential of designing a small molecule as a transcriptional “switch”, where a

certain extracellular signal, such as change in pH or oxidation state can convert a transcriptional repressor to a transcriptional activator and vice versa.

B. Future Directions

B.1. Application of environmentally sensitive fluorophores

In Chapters 2 and 4, fluorescein was utilized as the fluorescent probe to detect changes in the local environment upon binding. While we were able to glean useful data from this, the amplitude of intensity change was quite low, ranging from 5-20% of the initial fluorescence intensity (see Figures 2.9, 2.10 and 4.10, 4.14). This results in lower signal to noise resolution and we might miss or misinterpret certain steps. A simple solution to this is to use solvatochromic fluorophores instead. Solvatochromic fluorophores can be extremely useful in monitoring protein-protein interactions.⁹⁻¹¹ These are environmentally sensitive fluorophores that have low quantum yield when exposed to polar solvents such as water, and have up to 100-fold higher quantum yield when buried in hydrophobic environments.¹²⁻¹⁴

Labeling either the more structured coactivator or the unstructured activator with such a fluorophore could provide higher resolution data reflecting the binding interaction (Figure 5.3). By labeling the binding site of a coactivator such as KIX with a solvatochromic fluorophore, one can directly monitor the conformational changes on KIX instead of the activator TAD. This will be more applicable to testing the hypothesis that small molecules are potentiating pKID binding to KIX by changing the conformation of KIX so less conformational change is required upon pKID binding (Figure 5.3a). By labeling an activator that transitions from unstructured to structured upon coactivator binding, such as c-Myb, pKID, HBZ or Gal4, one can obtain information regarding the formation of a helix upon binding.¹¹ This is especially advantageous to

monitoring TADs such as c-Myb that do not display N-terminal FITC fluorescent intensity change upon KIX-binding (Mapp lab, data unpublished) as well TADs that do not have a defined structure such as HBZ (Figure 5.3b).¹⁵

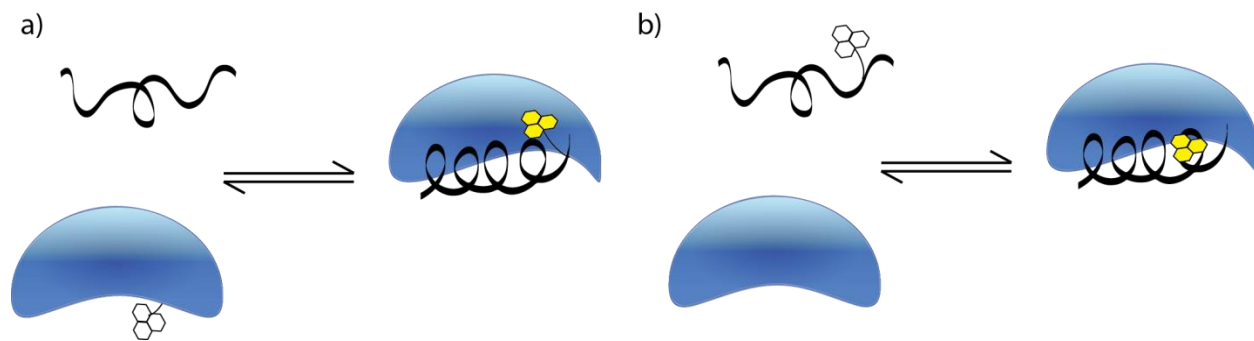


Figure 5.3 Schematic of utilizing solvatochromic fluorophores to monitor activator-coactivator interactions in two ways. **a)** The fluorophore is tagged on the more structured coactivator at the binding site, upon binding of an activator, the coactivator undergoes conformational change, resulting in the fluorophore being buried in a more hydrophobic environment and emission increases. **b)** The fluorophore is tagged on the unstructured TAD, upon the TAD assuming a helical structure after binding to a coactivator, the fluorophore is buried in the hydrophobic core of the helix and emission increases.

While there are several commercially available solvatochromic fluorophores, such as MDCC, IANBD, BADAN, IAEDANS and PyMPO, a fluorophore developed by the Imperiali group is especially attractive for this purpose. 4-N,N-dimethylamino-1,8-naphthalimide (4-DMN) has been shown to be superior to several commercially available solvatochromic fluorophores due to its low background signal, large increase in quantum yield upon protein binding/folding, long excitation wavelength (408 nm) deeming it more bio-orthogonal, and high stability in aqueous solution.^{16,17} In addition, the protocols are available for synthesizing versions of this fluorophore for both solid phase peptide synthesis incorporation and cysteine-labeling,^{13,18} so one can incorporate it in both synthesized peptides and recombinantly expressed proteins.

B.2. In cell NMR

As multiple studies have shown, the inner cellular environment is extremely crowded (Figure 5.4). Estimations suggest that the protein concentration in a living cell is 200-300 mg/ml, while the RNA concentration is 75-150 mg/ml, in total these proteins and nucleic acids may occupy up to 40% of the intracellular environment.¹⁹⁻²¹ This crowding effect can mean that the structural characterizations of intrinsically disordered proteins are different in cells than that in a dilute homogenous buffer.^{22,23} This also holds true for the protein-protein interactions. The high viscosity can limit rotational diffusion of proteins and alter the binding kinetics.²⁴

To address this shortcoming of studying protein interactions in buffer, in cell NMR has been developed to examine the conformation of proteins in a cellular environment in the context of *Xenopus* oocytes, *Escherichia coli* and mammalian cells.²⁵⁻³⁰ A method has also been described to monitor protein-protein interactions in cells (STINT-NMR),³¹ and has been applied to study the interaction between Ubiquitin and its various binding partners.³²

In cell NMR will be a powerful tool for studying the conformational dynamics of disordered TADs undergoing the binding interaction with coactivators, such as Gal4 and Med15 or its suppressor Gal80. We are already able to obtain a clear ¹H-¹⁵N HSQC spectra of GB1 tagged Gal4 TAD (residues 840-881, see

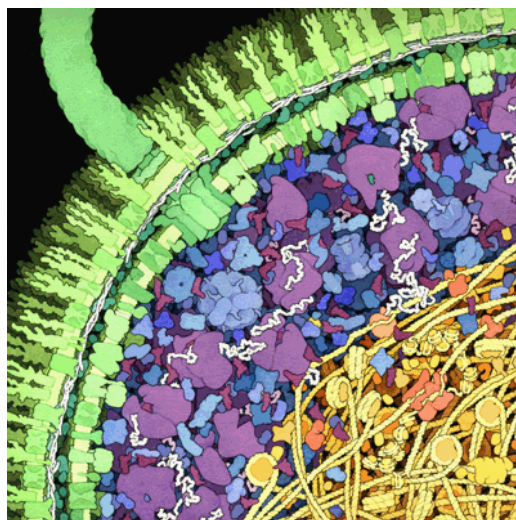


Figure 5.4 Artistic representation of a cross-section of a small portion of an *E. coli* cell. The cell wall and a flagellar motor are in green, the cytoplasmic area includes enzymes (blue) and ribosomes (purple) with RNA depicted in white. The nucleoid region is shown in yellow and orange. Illustration reproduced with permission by David S. Goodsell, the Scripps Research Institute.

Appendix). This suggests there is potential in pursuing in cell NMR on this construct and monitoring its protein-protein interactions by collaborating with highly established NMR laboratories on campus such as the Al-Hashimi lab. Also, a simpler and faster ^{19}F -NMR technique to study the basic conformational dynamics of KIX has been developed in the Mapp lab,³³ this can be adapted to in cell NMR methods as well, enabling us to monitor the interactions of KIX with its various ligands in a cellular environment.

C. References

1. 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).
2. Tompa, P., Szász, C. & Buday, L. Structural disorder throws new light on moonlighting. *Trends Biochem. Sci.* **30**, 484–9 (2005).
3. Motlagh, H. N., Li, J., Thompson, E. B. & Hilser, V. J. Interplay between allostery and intrinsic disorder in an ensemble. *Biochem. Soc. Trans.* **40**, 975–80 (2012).
4. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
5. Schreiber, G. & Keating, A. E. Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50–61 (2011).
6. Higurashi, M., Ishida, T. & Kinoshita, K. Identification of transient hub proteins and the possible structural basis for their multiple interactions. *Protein Sci.* **17**, 72–8 (2008).
7. Kim, P. M., Sboner, A., Xia, Y. & Gerstein, M. The role of disorder in interaction networks: a structural analysis. *Mol. Syst. Biol.* **4**, 179 (2008).
8. Wang, N. *et al.* Ordering a dynamic protein via a small-molecule stabilizer. *J. Am. Chem. Soc.* **135**, 3363–6 (2013).
9. Wang, Q. & Lawrence, D. S. Phosphorylation-driven protein-protein interactions: a protein kinase sensing system. *J. Am. Chem. Soc.* **127**, 7684–5 (2005).
10. Venkatraman, P. *et al.* Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells. *Nat. Chem. Biol.* **3**, 222–8 (2007).
11. Fuller, A. a., Seidl, F. J., Bruno, P. a., Plescia, M. a. & Palla, K. S. Use of the environmentally sensitive fluorophore 4-N,N-dimethylamino-1,8-naphthalimide to study peptoid helix structures. *Biopolymers* **96**, 627–638 (2011).
12. Loving, G. S., Sainlos, M. & Imperiali, B. Monitoring protein interactions and dynamics with solvatochromic fluorophores. *Trends Biotechnol.* **28**, 73–83 (2010).
13. Sainlos, M. & Imperiali, B. Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores through SPPS-based “building block” approach. *Nature Protocols* **2**, 3210–8 (2007).

14. Sainlos, M. & Imperiali, B. Tools for investigating peptide-protein interactions: peptide incorporation of environment-sensitive fluorophores via on-resin derivatization. *Nature Protocols* **2**, 3201–9 (2007).
15. Cook, P. R., Polakowski, N. & Lemasson, I. HTLV-1 HBZ protein deregulates interactions between cellular factors and the KIX domain of p300/CBP. *J. Mol. Biol.* **409**, 384–98 (2011).
16. Eugenio Vazquez, M., Rothman, D. M. & Imperiali, B. A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis. *Org. Biomol. Chem.* **2**, 1965–6 (2004).
17. 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**, 13630–8 (2008).
18. Loving, G. & Imperiali, B. Thiol-reactive derivatives of the solvatochromic 4-N,N-dimethylamino-1,8-naphthalimide fluorophore: a highly sensitive toolset for the detection of biomolecular interactions. *Bioconjugate Chem.* **20**, 2133–41 (2009).
19. Ellis, R. J. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* **11**, 114–119 (2001).
20. Zimmerman, S. B. & Trach, S. O. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J. Mol. Biol.* **222**, 599–620 (1991).
21. Hall, D. & Minton, A. P. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1649**, 127–139 (2003).
22. Cino, E. A., Karttunen, M. & Choy, W.-Y. Effects of molecular crowding on the dynamics of intrinsically disordered proteins. *PLoS One* **7**, e49876 (2012).
23. Miklos, A. C., Sarkar, M., Wang, Y. & Pielak, G. J. Protein crowding tunes protein stability. *J. Am. Chem. Soc.* **133**, 7116–20 (2011).
24. Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev. (Washington, DC, U. S.)* **109**, 839–60 (2009).
25. Maldonado, A. Y., Burz, D. S. & Shekhtman, A. In-cell NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 197–212 (2011).
26. Serber, Z., Corsini, L., Durst, F. & Dötsch, V. In-cell NMR spectroscopy. *Methods Enzymol.* **394**, 17–41 (2005).
27. Pielak, G. J. *et al.* Protein nuclear magnetic resonance under physiological conditions. *Biochemistry* **48**, 226–34 (2009).
28. Selenko, P., Serber, Z., Gadea, B., Ruderman, J. & Wagner, G. Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11904–9 (2006).
29. Robinson, K. E., Reardon, P. N. & Spicer, L. D. In-cell NMR spectroscopy in *Escherichia coli*. *Methods Mol. Biol. (N. Y.)* **831**, 261–77 (2012).
30. Bekei, B. *et al.* In-cell NMR in mammalian cells: part 1. *Methods Mol. Biol. (N. Y.)* **895**, 43–54 (2012).
31. Burz, D. S., Dutta, K., Cowburn, D. & Shekhtman, A. In-cell NMR for protein-protein interactions (STINT-NMR). *Nature Protocols* **1**, 146–52 (2006).
32. Burz, D. S. & Shekhtman, A. In-cell biochemistry using NMR spectroscopy. *PLoS One* **3**, e2571 (2008).

33. Pomerantz, W. C. *et al.* Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **7**, 1345–50 (2012).

Appendix⁹

A. Introduction

Many transcriptional activators are intrinsically disordered proteins or have intrinsically disordered domains.¹⁻⁶ This enables allosteric interactions between domains and between binding partners, giving these proteins the advantage of multispecificity.⁷⁻⁹ However this unique characteristic does impede the understanding of these proteins and their interactions on a structural level, as it is more difficult to obtain a structure or ensemble of conformations using traditional structural techniques such as protein crystallography. Nuclear Magnetic Resonance (NMR) has proved to be a powerful tool in detecting the conformation of more structurally flexible proteins.¹⁰⁻¹³ Gal4 has been studied as a prototypical eukaryotic transcriptional activator for several decades,¹⁴⁻¹⁷ and is involved now in routine systems used to probe transcriptional pathways in cells.^{18,19} However the structural details of Gal4 remain vague, and additional information is crucial in fully understanding the mechanism of this activator. Here, using structural information gained by NMR spectroscopy, we are able to investigate the allosteric effects between two Gal4 domains, the DNA binding domain (DBD) and the transcriptional activation domain (TAD). This can aid in further understanding of Gal4 when used as a probing tool, as well as provide insight into the structural information on transcriptional activators in general.

⁹ This work was performed in collaboration with Amanda Dugan and Felicia Gray, and also under guidance of Dr. Jeetender Chugh and Dr. Katja Petzold in Prof. Hashim M. Al-Hashimi's laboratory.

B. Background

B.1. Structural information of Gal4

Gal4 is an 881-residue protein, deletion analysis²⁰ has shown that the N-terminal 147 residues and the last 100 residues at the C-terminal are the key domains that govern DNA-binding and transcriptional activation respectively.²¹ Further studies of Gal4 have used the simplified modular construct of these two domains with slight variations on sequence length.^{21–26} Structural studies of Gal4 have mainly focused on the DNA binding domain (sometimes further categorized as a DNA binding domain and a dimerization domain)²⁷, as this is a structurally stable domain. NMR and X-ray crystallography data have shown that the DBD binds to DNA via Zn_2Cys_6 clusters, and exists as a homodimer of intertwined helices (Figure 0.1a).^{27–29}

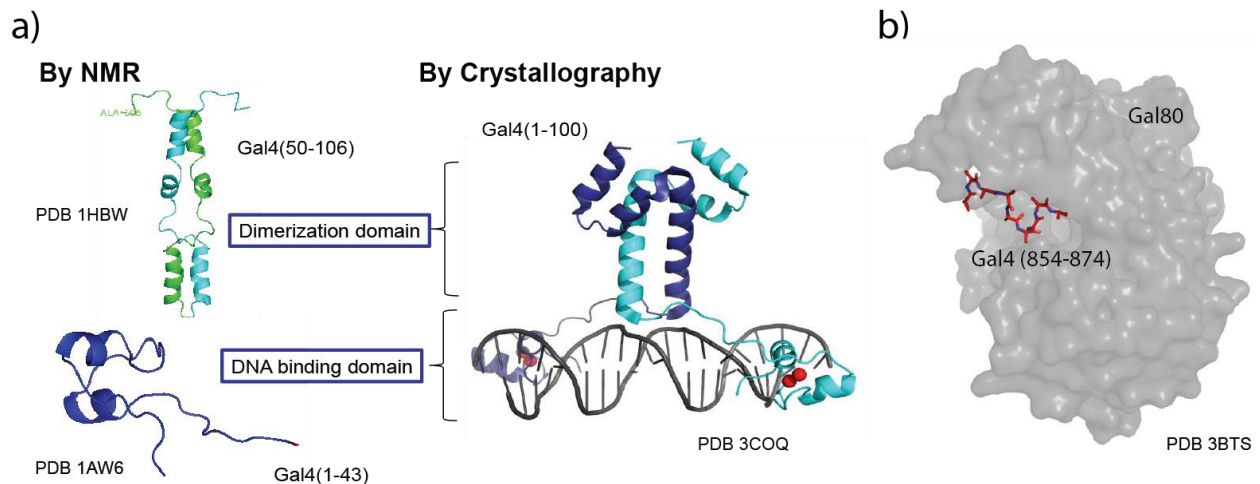


Figure 0.1 Structures of Gal4 domains. **a)** Various structures of the Gal4 DBD and the methods used to solve them. The structure of Gal4 (50-106) clearly shows different structural organization than that of Gal4 (1-100). This could be due to allosteric effects by residues 1-49 or by the DNA oligomer, or due to crystal packing that could occur in x-ray crystallography. **b)** Structure of a Gal4 TAD peptide (red) bound to Gal80 (gray), its suppressor.

Interestingly the conformation of the published DBD structures are quite different, likely caused by certain variables, such as domain length, structural method employed, and with or without DNA. The conformation of the transcriptional activation domain of Gal4, however, is

poorly understood, and the only structure of the TAD is a 21-residue peptide bound to its suppressor, Gal80 (Figure 0.1b).³⁰ There is much more that needs to be understood about the conformational characteristics of the Gal4 TAD, as this is the domain of the protein that governs interaction with and recruitment of various different transcription factors.

B.2 Using NMR to gain structural information of conformationally dynamic proteins

The utilization of NMR techniques to obtain structural information of proteins has been steadily progressing over the past few decades.^{10,31} Today, NMR has been proven complementary to crystallography in producing protein structures.³² While crystallography can produce very specific atomic level structures, it relies heavily on a protein forming a homogenous crystal. Alternatively, solution NMR provides access to the solution structure and dynamics of protein. This method is even more desirable in studying the conformations of intrinsically disordered or conformationally dynamic proteins.^{12,13} It is harder for flexible proteins to crystallize, and even though various methods have been developed to aid their crystallization^{33,34} only one conformation will be captured, and might not comprehensively reflect the ensemble nature of these flexible proteins. NMR, on the other hand, does not require a homogenous solution, and can record an ensemble of conformations that could be flexible and “fuzzy” in certain regions.³⁵ This gives us a more accurate view of the state of the protein in solution.

Even more enticing, different NMR techniques can also provide insights into protein structure or dynamics without necessarily requiring complete structural data. In particular, valuable information is available through monitoring changes in protein chemical shifts using simple 1D or 2D experiments.^{12,13} ¹H, ¹⁵N or ¹³C chemical shifts report on the

chemical and electric environment of a particular nuclei, which can be correlated to the local protein secondary structure.³⁶⁻⁴⁰ Common 2D NMR experiments for studying proteins in solution are 2D NMR ^1H - ^{15}N or ^1H - ^{13}C HSQC (heteronuclear single quantum correlation) experiments. These experiments are advantageous as protein ^{15}N and ^{13}C chemical shifts typically have greater dispersion than ^1H chemical shifts, and it is possible to assign the chemical shifts of individual protein residues.⁴¹ This allows for higher resolution characterization of changes in protein structure or environment. The conformational information of many transcriptional activation domains have been probed by this technique. For example, ^{15}N -labeled VP16 TAD was titrated with a binding partner PC4, and the chemical shift perturbation shown by the overlaid ^1H - ^{15}N HSQC spectra identified two regions of VP16 TAD that were most effected by PC4 binding.⁴² Another example is the studies of HIF-1 α binding to the TAZ1 domain of CBP/p300. The ^1H - ^{15}N HSQC spectra for the complex showed that ^1H signals for the unstructured HIF-1 α had greater dispersion beyond the random coil region (7.8-8.4 ppm) indicating that upon TAZ1 binding the unstructured HIF-1 α adopts an α -helical structure.^{13,43} Similarly, observing the chemical shift perturbation of Gal4 ^1H - ^{15}N HSQC spectra in different constructs could help us gain insight into the conformations adopted by different constructs of the protein.

B.3. An NMR compatible solubility tag GBI

A common characteristic of many flexible proteins, especially transcriptional activators is the low solubility of the isolated protein in vitro. To this end, solubility enhancement tags are often fused to the N-terminal of the target protein. Unfortunately many of the common solubility tags, such as glutathione S-transferase (GST) and maltose binding protein (MBP),

are fairly large in size. This poses as a problem in NMR studies, as conventionally it is much harder to obtain high resolution spectra from larger proteins (>30 kDa).⁴⁴ Cleaving the solubility tag after expression is not an option either, as most 2D NMR techniques require a sample concentration of approximately 100 μ M. To address this problem, a much smaller solubility tag is often used in NMR studies, the 56 amino acid protein G B1 domain (GB1).⁴⁵ This domain has been shown to be highly soluble, highly stable, and has minimal nonspecific interactions with target proteins.^{46,47} For these reasons, we have chosen to use GB1 as the solubility tag for the Gal4 proteins.

C. Using ^1H - ^{15}N HSQC to study allosteric effects between Gal4 domains

C.1. Gal4 TAD exhibits increased affinity to Med15 when fused to Gal4 DBD

We noticed that the affinity of the Gal4 DBD-TAD fusion protein (Gal4 (1-100)-(840-881)) to Med 15 (1-345) was much higher than that of the Gal4 TAD peptide (854-874) (Figure 0.2). While there were different solubility tags on Med15 (GB1 was used in the first experiment while GST used in the latter), the drastic difference of 20-fold cannot be simply explained by nonspecific interactions between solubility tags (GB1 does not dimerize at the pH conditions used in this assay)⁴⁸ or the small difference in the TAD sequences (residues 855-870 were determined to play the most important role in TAD function)⁴⁹. We then hypothesized that the fusion of the Gal4 DBD to Gal4 TAD allosterically affects the conformational dynamics of the TAD region. To test this hypothesis, NMR HSQC techniques were employed to gain more structural information.

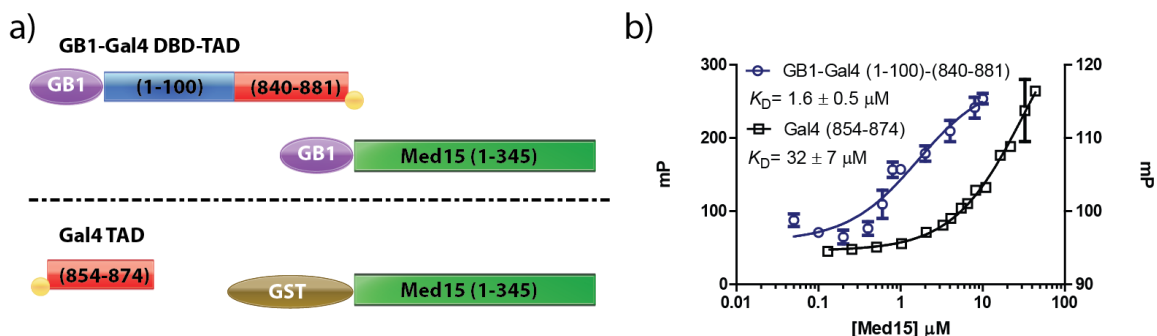


Figure 0.2 a) Schematic of components of the two separate assays. In the first assay, the C-terminus cysteine of Gal4 is labeled with a fluorescein, in the second assay, the N-terminus of Gal4 peptide is coupled with FITC. GB1 and GST are solubility tags fused to either Gal4 or Med15. b) Fluorescent polarization binding curves of Gal4 DBD-TAD (blue, right Y-axis) and Gal4 TAD (black, left Y-axis) binding to Med15 (1-345). The Gal4 DBD-TAD binding assay shows smaller dynamic range due to the smaller difference in size after binding to Med15 compared to the Gal4 TAD peptide. Data for the Gal4 DBD-TAD assay were performed in triplicate; standard deviation is depicted as error bars. Data for the Gal4 TAD assay was performed in duplicate, standard deviation is depicted as error bars. K_D values were obtained by fitting data in GraphPad Prism; standard error of fitting is included. See Experimental methods section for further details.

C.2. ^1H - ^{15}N HSQC data of GB1-tagged Gal4 constructs

Two dimensional ^1H - ^{15}N HSQC (heteronuclear single quantum correlation) NMR experiments were carried out first with a ^{15}N GB1-Gal4 (840-881) construct. The resulting spectrum had clear resolution (Figure 0.3). GB1's NMR assignments are already available⁵⁰ and have been assigned on the spectrum. The unassigned peaks are those of the Gal4 TAD, which are mostly in the random coil region between 7.8-8.4 ppm on the ^1H axis.

Similar ^1H - ^{15}N HSQC experiments were carried out for ^{15}N GB1-Gal4 (1-100) and ^{15}N GB1-Gal4 (1-100)-(840-881). While these two constructs did not yield spectrum with optimal resolution, overlay of the three spectra (Figure 0.4) clearly show peak shifts in the GB1-Gal4 (1-100)-(840-881) spectra that correspond to the Gal4 (840-881) region display perturbation compared to that of isolated GB1-Gal4 (840-881). This suggests that there is indeed a certain degree of conformational change happening to Gal4 TAD when it is fused to the Gal4 DBD. (See Experimental methods section for details.) Also of note is that many

peaks have “disappeared” from the random coil region in the Gal4 (840-881) spectra when Gal4 (1-100) is fused to it. While more vigorous controls and higher resolution data is required to be certain, this suggests some residues in the TAD might have assumed a more structured conformation and shifted out of the random coil region.

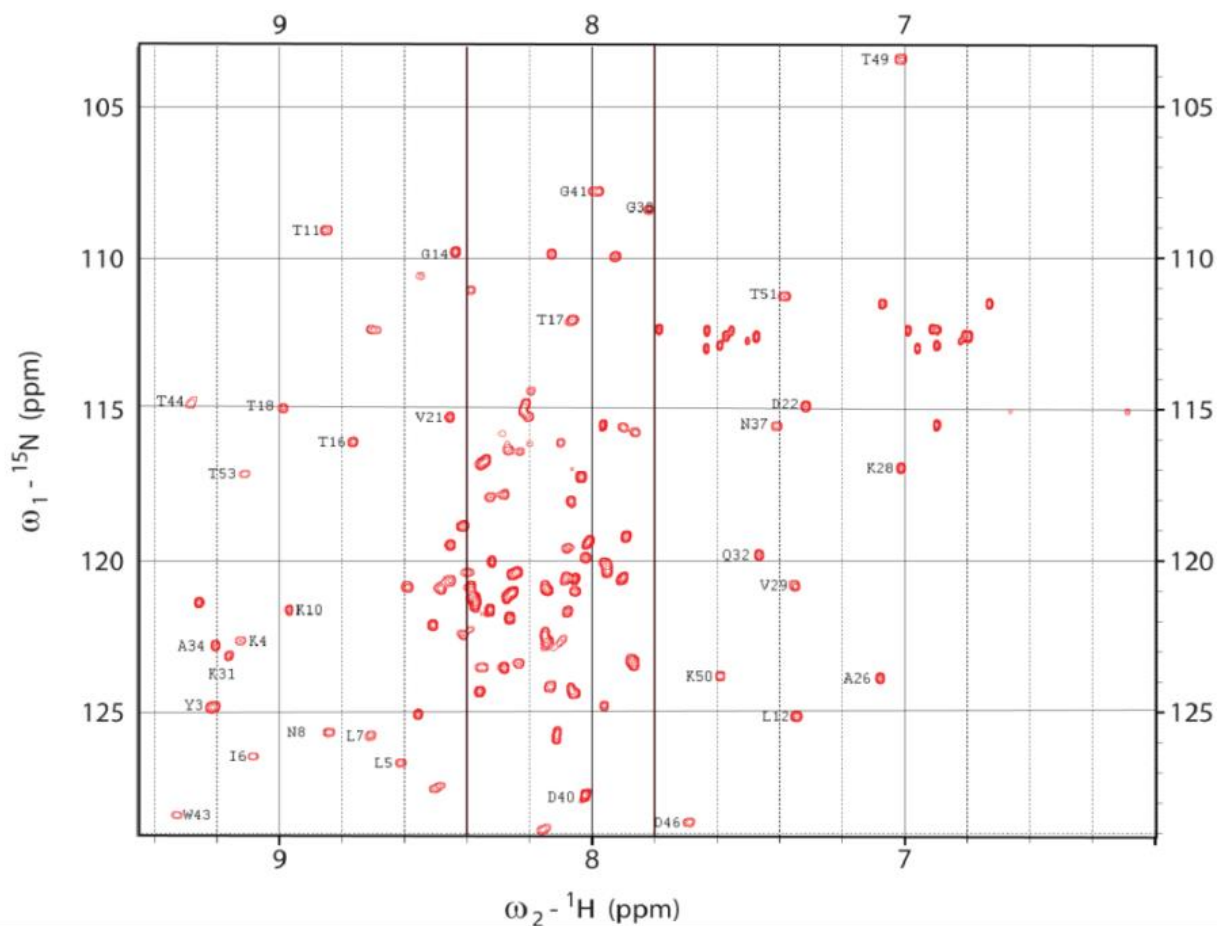


Figure 0.3 ^1H - ^{15}N HSQC spectrum of ^{15}N GB1-Gal4 (840-881). The annotated peaks are residues from GB1. The random coil region between 7.8-8.4 ppm is indicated by vertical lines.

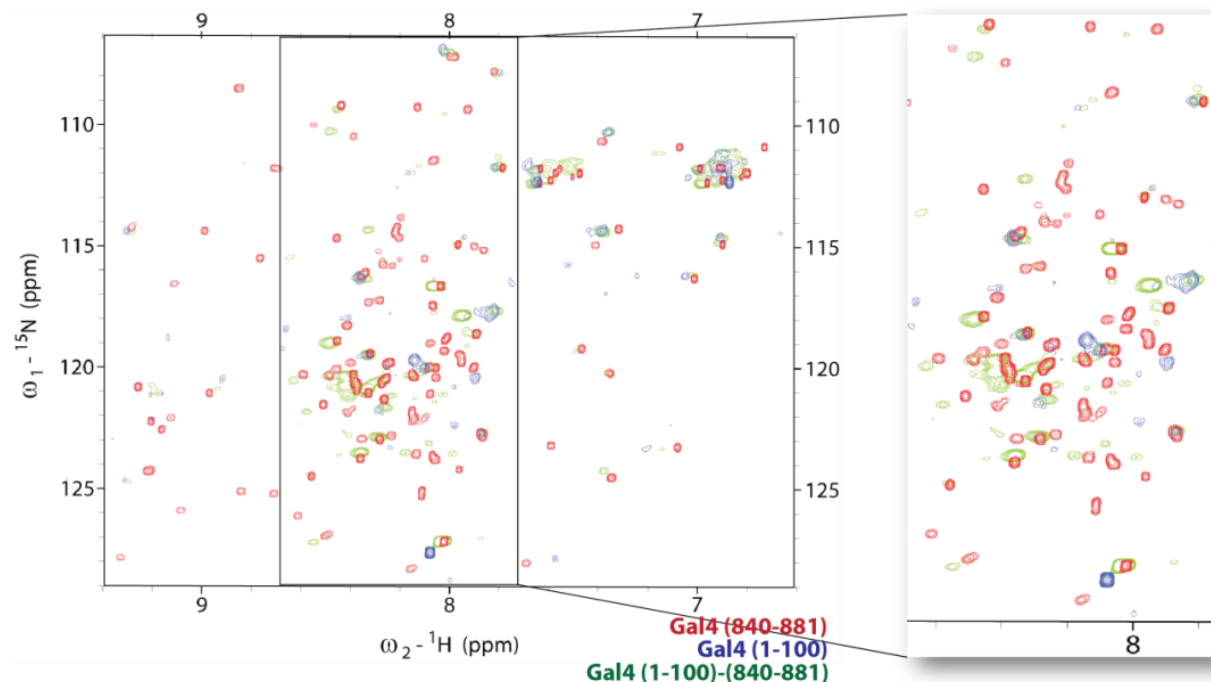


Figure 0.4 Overlay of the ^1H - ^{15}N HSQC spectra of ^{15}N labeled GB1-Gal4 (840-881) (red), GB1-Gal4 (1-100) (blue) and GB1-Gal4 (1-100)-(840-881) (green). The random coil region between 7.8-8.4 ppm is indicated by vertical lines and zoomed in for better resolution.

D. Future Directions

There is still work to be done before definitive answers can be drawn on this study. First of all the residues in the Gal4 (840-881) HSQC spectra should be assigned. Many studies have been focused on predicting, identifying and calibrating chemical shifts in the random coil region,^{12,51,52} which will be instrumental in this process. In addition, alternative detection methods such as ^{13}C -NMR can also be employed. Heteronuclear detection is advantageous for studying intrinsically disordered proteins as ^{13}C and ^{15}N show much better random coil dispersion than protons.⁵³⁻⁵⁶

Also, as mentioned in the Results section, poor resolution spectra were collected for the Gal4 (1-100) and Gal4 (1-100)-(840-881) constructs. Identifying conditions to increase protein solubility are essential to recording high quality spectra. While the GB1 tag is necessary for protein solubility, it is possible that the increased molecular weight from the tag could decrease

spectral quality. One way to avoid this would be by introducing an NMR silent (non ^{15}N labeled) tag: labeled GB1-Gal4 would be expressed and purified as usual, then an unlabeled GB1 tag can be ligated to the other terminus of Gal4 by a pre-installed protein ligation sequence, and the ^{15}N GB1 tag is cleaved off at an introduced cleavage site.⁴⁶

Once the chemical shifts of Gal4 (840-881) are correctly assigned and the spectra of Gal4 (1-100) and Gal4 (1-100)-(840-881) are obtained in high resolution, overlay of Gal4 (1-100) with Gal4 (1-100)-(840-881) will reveal the resonance peaks corresponding to the TAD region in the Gal4 (1-100)-(840-881) construct. These resonance peaks will then be compared to those of the assigned peaks in the Gal4 (840-881) spectrum to identify the residues that have shifted in conformation upon fusion to Gal4 DBD.

E. Experimental Methods

Protein Expression and Purification

His₆-GB1-tagged Gal4 proteins: The DNA sequences of Gal4 (1-100), Gal4 (1-100)-(840-881) and Gal4 (840-881) were cloned into the plasmid pMCSG9 containing a His₆-GB1 tag at the N-terminus. Expression of proteins were carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen) as previously described.⁵⁷ Briefly, cultures (50 mL) inoculated with single colonies were grown overnight at 37 °C (250 rpm) in Lennox L Broth (Research Products International) supplemented with ampicillin (100µg/mL) and chloramphenicol (34 µg/mL) before dilution (50-fold) into 4x 1 L cultures of either Lennox L Broth (Research Products International) or minimal media M9 (including $^{15}\text{NH}_4\text{Cl}$ as previously described)⁵⁸ supplemented with ampicillin (100 µg/mL). After an OD₆₀₀ of 0.4 was reached, protein over-expression was induced with IPTG (final concentration 1 mM) in the presence of 20 µM

ZnSO₄ for 5 hours. Cells in 50 mL culture were pelleted by centrifugation, resuspended in 10 mL lysis buffer A (10 mM Tris, pH 8.0 at 4 °C, 500 mM NaCl, 10% glycerol (v/v), 10 mM β-ME, 0.1% Tween* 20 (v/v), and Roche Complete Protease Inhibitor Cocktail), then lysed using sonication. His-tagged protein was isolated by incubating cell lysate with 200 μL of Ni-NTA beads (Qiagen) for 1 hour at 4 °C, followed by washing with 8 times with 1 mL wash buffer A (20 mM Tris, pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 1 mM β-ME, 0.1% Tween* 20 (v/v), 30 mM imidazole). The protein was eluted from the beads by incubation at 4 °C overnight with 1 mL elution buffer A (20 mM Tris pH 8.0 at 4 °C, 100 mM NaCl, 20% glycerol (v/v), 250 mM imidazole). The protein solution was buffer exchanged into storage buffer A (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM β-ME, 1 mM EDTA, 20 μM ZnSO₄) using a PD-10 column (GE Healthcare), and the protein concentration was measured using absorbance at 280 nm. The identity and purity (>90%) of the protein was verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

GST-tagged Med15 (1-345): The DNA sequence of Med15 (1-345) **was** cloned into the plasmid pGEX containing a GST at the N terminus. Expression of GST-Med15 (1-345) was carried out in Rosetta2 (DE3) pLysS *E. coli* cells (Novagen). Briefly, cultures (50 mL) from single colonies were grown overnight at 37 °C (250 rpm) in Select APS Super Broth (Difco) supplemented with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL) before dilution (100-fold) into 4 x 1 L of Select APS Super Broth supplemented with ampicillin (100 μg/mL). After an OD₆₀₀ of 0.3 was reached, the cultures were cooled for 45 min at 16 °C (150 rpm), and expression was induced with IPTG (final concentration 0.1 mM) for 5-6 hours at 250 rpm. Each cell pellet was resuspended in 25 mL lysis buffer B (100 mM PBS pH 7.4 (Pierce), 0.2% NP-40 Substitute (Fluka), 10% glycerol (v/v), 1 mM DTT and Roche Complete Protease Inhibitor

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

GB1-tagged Med15 (1-345): The DNA sequence of Med15 (1-345) was cloned into the plasmid pMCSG9 containing a His₆-GB1 tag at the N-terminus. The protein was expressed as described for GST-Med15 (1-345), and purified and verified for purity and concentration as described for the Gal4 proteins.

Expression and Fluorescein Labeling of Gal4 Cysteine Mutant

A cysteine residue was added to the C terminus of the His₆-GB1-Gal4 (1-100)-(840-881) by site directed mutagenesis to create His₆-GB1-Gal4 (1-100)-(840-881)-Cys. The protein was expressed and purified as described for His₆-GB1-tagged Gal4 proteins. Fluorescein-5-

Maleimide (Pierce) was dissolved in small amounts of DMSO, then added to His₆-GB1-Gal4(1-100)-(840-881)-Cys at 25-molar excess in labeling buffer (20 mM HEPES, pH 7.5 at 4 °C, 200 mM NaCl, 10% glycerol (v/v), 1 mM TCEP, 1 mM EDTA, 20 μM ZnSO₄). The reaction was incubated overnight at 4°C protected from light. Excess Fluorescein-5-maleimide was removed by desalting with a PD-10 column (GE Healthcare). The degree of labeling was calculated by the following equation:

$$\text{Moles fluorophore per moles protein} = \frac{A_{max}}{\epsilon' \times [\text{protein}](M)} \times \text{dilution factor}$$

(Eq. A.1)

In this equation A_{max} is the absorbance of fluorophore (at 495 nm), and ϵ' is the molar extinction coefficient of the fluorophore (68,000 M⁻¹cm⁻¹). The protein concentration is obtained by equation A.2, in which CF is the correction factor of $A_{280}/A_{max}=0.3$ and ϵ is the molar extinction coefficient of the protein in question at A_{280} .

$$[\text{protein}](M) = \frac{A_{280} - (A_{max} \times CF)}{\epsilon} \times \text{dilution factor}$$

(Eq. A.2)

The Gal4 protein was labeled with 50-60% efficiency, and concentration of the protein used in calculations for binding assays was corrected for the concentration of fluorescein-labeled protein.

Peptide Synthesis and Purification

The Gal4 TAD (854-874) peptide was synthesized and purified as previously described.⁵⁹

The sequence is as follows:

N term-FITC-CGMFNTTMDDVYNYLFDDEDT-C term

Fluorescence Polarization Assays to Measure Gal4 affinity to Med15 (1-345)

FITC-labeled Gal4 (854-874) peptide or Fluorescein labeled His₆-GB1-Gal4(1-100)-(840-881)-Cys were diluted in DNA-binding buffer (20 mM HEPES pH 7.5, 75 mM potassium acetate, 0.02 mM zinc sulfate, 4 mM magnesium acetate, 1 mM β-mercaptoethanol (βME), 0.05 mM EDTA, 10% glycerol, and 0.1 mg/mL BSA) to a concentration of 100 nM. Then 200 μL of the DNA solution was added to a series of 50 μL solutions of varying Med15(1-345) concentrations in storage buffer B (10 mM PBS pH 7.4 (Pierce), 10% glycerol (v/v), 0.01% NP-40 Substitute, 1 mM DTT) to obtain the final concentrations of up to 80 μM. The samples were incubated for 30 min at room temperature before the degree of fluorescence polarization was measured (Beacon 2000, Pan Vera Corp). A binding isotherm that accounts for ligand depletion⁶⁰ (assuming a 1:1 binding model of dimeric activator to duplex DNA) was fit to the observed mP values as a function of activator to obtain the apparent equilibrium dissociation constant, K_d :

$$y = c + (b - c) \times \frac{[(Kd+a+x) - \sqrt{(Kd+a+x)^2 - 4ax}]}{2a} \quad (\text{Eq. A.3})$$

where “a” and “x” are the total concentrations of duplex DNA and dimeric activator, respectively, “y” is the observed polarization at any activator concentration, “b” is the maximum observed polarization value, and “c” is the minimum observed polarization value. Each data point is an average of either two or three independent experiments with the indicated error (standard deviation). Data analysis was performed using GraphPad Prism.

¹H-¹⁵N-HSQC NMR Experiments

Uniformly ¹⁵N labeled GB1-Gal4 protein was expressed and purified as previously described. A 50-150 μM solution of ¹⁵N- labeled GB1-Gal4 constructs was prepared in a 9:1 H₂O:D₂O 10

mM sodium phosphate buffer containing 100 mM NaCl at pH 7.2. ^1H - ^{15}N HSQC experiments were recorded at 27 °C on an Avance Bruker 600 MHz NMR spectrometer equipped with a triple-resonance 5 mm cryogenic probe. Data was processed using NMRpipe⁶¹ and analyzed in Sparky (UCSF).⁶² Chemical shifts of residues were identified based on previous assignments.⁶³

F. References

1. Liu, J. *et al.* Intrinsic disorder in transcription factors. *Biochemistry* **45**, 6873–6888 (2006).
2. Leach, B. I. *et al.* Leukemia fusion target AF9 is an intrinsically disordered transcriptional regulator that recruits multiple partners via coupled folding and binding. *Structure* **21**, 176–83 (2013).
3. Wells, M. *et al.* Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5762–5767 (2008).
4. Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).
5. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–64 (2010).
6. Uversky, V. N., Florida, S., Academy, R. & Region, M. *3 . 9 Intrinsically Disordered Proteins*. **3**, 170–211 (2012).
7. Singh, G. P., Ganapathi, M. & Dash, D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* **66**, 761–5 (2007).
8. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
9. Tsai, C.-J., Ma, B. & Nussinov, R. Protein-protein interaction networks: how can a hub protein bind so many different partners? *Trends Biochem. Sci.* **34**, 594–600 (2009).
10. Wuthrich, K. The Development of Nuclear Magnetic Resonance Spectroscopy as a Technique for Protein Structure Determination. *Structure* 36–44 (1989).
11. Felli, I. C. & Pierattelli, R. Recent progress in NMR spectroscopy: Toward the study of intrinsically disordered proteins of increasing size and complexity. *IUBMB Life* **64**, 473–481 (2012).
12. Jensen, M. R., Ruigrok, R. W. & Blackledge, M. Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr. Opin. Struct. Biol.* **23**, 1–10 (2013).
13. Dyson, H. J. & Wright, P. E. Unfolded proteins and protein folding studied by NMR. *Chem. Rev. (Washington, DC, U. S.)* **104**, 3607–3622 (2004).
14. Giniger, E., Varnum, S. M. & Ptashne, M. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* **40**, 767–774 (1985).
15. Kakidani, H. & Ptashne, M. GAL4 activates gene expression in mammalian cells. *Cell* **52**, 161–167 (1988).
16. Fischer, J. A., Giniger, E., Maniatis, T. & Ptashne, M. GAL4 activates transcription in *Drosophila*. *Nature* **332**, 853–856 (1988).

17. Collins, G. a, Lipford, J. R., Deshaies, R. J. & Tansey, W. P. Gal4 turnover and transcription activation. *Nature* **461**, E7; discussion E8 (2009).
18. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–15 (1993).
19. Joung, J. K., Ramm, E. I. & Pabo, C. O. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7382–7 (2000).
20. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847–853 (1987).
21. Johnston, S. A. The DNA Binding and Activation Domains of Gal4p Are Sufficient for Conveying Its Regulatory Signals. *Microbiology* 2538–2549 (2008).
22. Liang, S. D., Marmorstein, R., Harrison, S. C. & Ptashne, M. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn₂ Cys₆ binuclear cluster proteins recognizes DNA. *Mol. Cell. Biol.* **16**, 3773–3780 (1996).
23. Ding, W. V & Johnston, S. A. The DNA binding and activation domains of Gal4p are sufficient for conveying its regulatory signals. *Mol. Cell. Biol.* **17**, 2538–2549 (1997).
24. Jeong, C. J. *et al.* Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* **40**, 9421–9427 (2001).
25. Fukuda, A. *et al.* Transcriptional coactivator PC4 stimulates promoter escape and facilitates transcriptional synergy by GAL4-VP16. *Mol. Cell. Biol.* **24**, 6525 (2004).
26. Reeves, W. M. & Hahn, S. Targets of the Gal4 transcription activator in functional transcription complexes. *Mol. Cell. Biol.* **25**, 9092–9102 (2005).
27. Hidalgo, P. *et al.* Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes Dev.* **15**, 1007–1020 (2001).
28. Baleja, J. D., Thanabal, V. & Wagner, G. Refined solution structure of the DNA-binding domain of GAL4 and use of 3 J (113 Cd , 1 H) in structure determination. *Biol. Chem.* **10**, 397–401 (1997).
29. Hong, M. *et al.* Structural basis for dimerization in DNA recognition by Gal4. *Structure* **16**, 1019–1026 (2008).
30. Kumar, P. R., Yu, Y., Sternglanz, R., Johnston, S. A. & Joshua-Tor, L. NADP regulates the yeast GAL induction system. *Science* **319**, 1090 (2008).
31. Billeter, M., Wagner, G. & Wuthrich, K. Solution NMR structure determination of proteins revisited. *J. Biomol. NMR* **42**, 155–158 (2008).
32. Yee, A. A. *et al.* NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *J. Am. Chem. Soc.* **127**, 16512–7 (2005).
33. Zhan, Y., Song, X. & Zhou, G. W. Structural analysis of regulatory protein domains using GST-fusion proteins. *Gene* **281**, 1–9 (2001).
34. Buck, M. Crystallography: Embracing Conformational Flexibility in Proteins. *Structure* **11**, 735–736 (2003).
35. 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).
36. Wishart, D. S. & Sykes, B. D. [12] Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**, 363–392 (1994).
37. Braun, D., Wider, G. & Wuthrich, K. Sequence-Corrected 15N “Random Coil” Chemical Shifts. *J. Am. Chem. Soc.* **116**, 8466–8469 (1994).

38. Wishart, D. S., Sykes, B. D. & Richards, F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–333 (1991).
39. Kjaergaard, M. & Poulsen, F. M. Disordered proteins studied by chemical shifts. *Prog. Nucl. Magn. Reson. Spectrosc.* **60**, 42–51 (2012).
40. Mulder, F. A. A. & Filatov, M. NMR chemical shift data and ab initio shielding calculations: emerging tools for protein structure determination. *Chem. Soc. Rev.* **39**, 578–90 (2010).
41. Yao, J., Dyson, H. J. & Wright, P. E. Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Lett.* **419**, 285–289 (1997).
42. Jonker, H. R. a H. R. A. R. A., Wechselberger, R. W. W. R. W., Boelens, R., Folkers, G. E. G. E. E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44**, 827–839 (2005).
43. Freedman, S. J. *et al.* Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5367–72 (2002).
44. Yu, H. Extending the size limit of protein nuclear magnetic resonance. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 332–334 (1999).
45. Huth, J. R. *et al.* Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. *Protein Sci* **6**, 2359–2364 (1997).
46. Zhou, P. & Wagner, G. Overcoming the solubility limit with solubility-enhancement tags : successful applications in biomolecular NMR studies. *J. Biomol. NMR* **46**, 23–31 (2010).
47. Zhou, P., Lugovskoy, A. A. & Wagner, G. A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. *J. Biomol. NMR* **20**, 11–14 (2001).
48. Tomlinson, J. H., Craven, C. J., Williamson, M. P. & Pandya, M. J. Dimerization of protein G B1 domain at low pH: a conformational switch caused by loss of a single hydrogen bond. *Proteins* **78**, 1652–61 (2010).
49. Ansari, A. Z., Reece, R. J. & Ptashne, M. A transcriptional activating region with two contrasting modes of protein interaction. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13543–13548 (1998).
50. Wilton, D. J., Tunnicliffe, R. B., Kamatari, Y. O., Akasaka, K. & Williamson, M. P. Pressure-induced changes in the solution structure of the GB1 domain of protein G. *Proteins* **71**, 1432–40 (2008).
51. Tamiola, K., Acar, B. & Mulder, F. A. A. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–3 (2010).
52. Schwarzinger, S. *et al.* Sequence-Dependent Correction of Random Coil NMR Chemical Shifts. *J. Am. Chem. Soc.* **123**, 2970–2978 (2001).
53. Showalter, S. A. NMR assignment of the intrinsically disordered C-terminal region of Homo sapiens FCP1 in the unbound state. *Biomol. NMR Assignments* **3**, 179–81 (2009).
54. Hsu, S.-T. D., Bertocini, C. W. & Dobson, C. M. Use of protonless NMR spectroscopy to alleviate the loss of information resulting from exchange-broadening. *J. Am. Chem. Soc.* **131**, 7222–3 (2009).
55. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391**, 136–48 (2009).

56. Csizmok, V., Felli, I. C., Tompa, P., Banci, L. & Bertini, I. Structural and dynamic characterization of intrinsically disordered human securin by NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 16873–9 (2008).
57. Reece, R. J., Rickles, R. J. & Ptashne, M. Overproduction and single-step purification of GAL4 fusion proteins from *Escherichia coli*. *Gene* **126**, 105–107 (1993).
58. Buhrlage, S. J. J. *et al.* Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem. Biol.* **4**, 335–344 (2009).
59. Majmudar, C. Y. *et al.* A high-resolution interaction map of three transcriptional activation domains with a key coactivator from photo-cross-linking and multiplexed mass spectrometry. *Angew. Chem., Int. Ed.* **48**, 7021–4 (2009).
60. Volkman, H. M., Rutledge, S. E. & Schepartz, A. Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain. *J. Am. Chem. Soc.* **127**, 4649–58 (2005).
61. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, (1995).
62. T. D. Goddard and D. G. Kneller Sparky 3.
63. 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).