

**Exogenous Control of the Assembly of Transcriptional
Complexes at Gene Promoters**

by

Jonas Westergaard Hojfeldt

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in The University of Michigan
2011

Doctoral Committee:

Professor Anna K. Mapp, Chair
Professor Carol A. Fierke
Associate Professor Jason E. Gestwicki
Associate Professor Jorge A. Iniguez-Lluhi

© Copyright by Jonas Westergaard Hoffeldt 2011
All Rights Reserved

Acknowledgements

At the final stage of my formal education I wish to thank teachers and advisors that have inspired my development as a student and a scientist. Peder Rosted, Frank Nielsen and Kurt Gothelf were mentors in my early education that I particular enjoyed learning from, and I am grateful for their teachings and inspiration.

I would like to thank Anna Mapp for the 5 years that I have worked in her lab. I was very excited to come to Michigan with the plan to join her lab, because it was the most inspiring research I had seen in my search for graduate studies. I have learned that this inspiration comes from Anna fully embracing what I consider the spirit of chemical biology, and I could not have had a better mentor to appreciate this exciting field. I have great admiration for her venturous approach to science. I have felt an enormous freedom, faith and respect to develop as a creative and critical scientist, and this I have appreciated the most. I am also thankful of a gracious and genuine personal relationship.

I would like to thank Carol Fierke, Jason Gestwicki and Jorge Iñiguez-Lluhí for serving as my committee members and giving helpful advice. I feel fortunate to have met each of them and they are in my opinion great role models for younger scientists. Carol also deserves credit for an idea that was relayed to me in my first year and that I have transformed into the ideas presented in Chapter 4. Jorge has been a very resourceful scientist for me to collaborate with during the past year, and I am very appreciative of that experience.

I would also like to acknowledge people that I have worked with in my research efforts. Much of the research that is presented in this dissertation has been a collaborative effort and contributions are specified in the chapters. Special thanks to Chinmay Majmudar, Steve Rowe, Carl Arevang, Pam Schultz, David Sherman, Will Pomerantz, Tomek Cierpicki, Amy Danowitz, Jolanta Grembecka, Anna Mapp for contributions to work that involved a high-throughput screen and follow-up studies, and to Yasuhiro Imaeda, Aaron Van Dyke, Osvaldo Cruz, James Carolan, Steven Sturlis, Jorge Iñiguez-Lluhí, Anna Mapp for work on collaborative efforts of targeting nuclear receptors.

I would like to thank all fellow members of the Mapp lab that have overlapped with my time in the lab. I appreciate all help that I have received experimentally and through discussions, and I am thankful for the solidary nature of many lab members. Chinmay Majmudar and Amberlyn Wands introduced me to molecular biology techniques, which I ended up using extensively. Chinmay has been a fantastic capacity, and was a driving force to push me through rough patches of a long follow-up effort to a high-throughput screen. Ryan Casey helped shape my ideas in early years in lab by engaging in many enjoyable discussions and worked with me on several smaller projects. I would like to thank Aaron Van Dyke for a really incredible effort and attitude in a joint project.

Finally I would like to thank my family and friends for support. Our friends in Ann Arbor have given Martina and I many joyful moments here and they have become part of a Michigan family that we will miss and look forward to visit in the future. My parents and sisters have always offered unwavering support and encouragement for my aspirations, and they can never be thanked enough for all their love. Martina is an incredible wife, who is a joy to come home to everyday and she is a fantastic mom to Lukas and Annika. She is also one of the smartest individuals I have met, and many topics in biology were made accessible to me through her. Annika and Lukas have been true miracles in our lives, and we love them very much.

Table of Contents

Acknowledgements	ii
List of Figures	vii
List of Abbreviations	x
Abstract	xiii
Chapter	
1. Artificial transcriptional regulators - ligands and bifunctional recruiters	1
1.1 Abstract	1
1.2 Introduction	1
1.3 Basic mechanisms of transcription initiation	2
1.4 Principles of inhibiting recruitment of coregulatory proteins to DNA	4
1.5 Tactics and limitations to positively modulate recruitment	7
1.6 Discovery of ligands of coactivator proteins	9
1.6.1 Synthetic transcriptional activation domain mimics	9
1.6.2 Screening for coactivator ligands	12
1.7 High affinity, drug-like ligands of transcriptional coregulatory proteins	18
1.8 Conclusions	20
1.9 Research goals	21
1.10 References	21
2. Discovery of coactivator ligands	28
2.1 Abstract	28
2.2 Introduction to coactivators as targets of high-throughput screening	29

2.3	Cell-based assays for transcriptional inhibitors	32
2.4	Conclusion and discussion of efforts with cell-based HTS assays	44
2.5	Identification of a lichen natural product as ligand of the coactivator CBP from a biochemical high-throughput screen	45
2.6	Conclusion and discussion of sekikaic acid discovery	70
2.7	Materials and methods	73
	2.7.1 Materials and methods used with cell-based assays	73
	2.7.2 Materials and methods used in discovery and characterization of sekikaic acid	77
2.8	Characterization of compounds made or purified by Jonas W. Hojfeldt	87
2.9	References	92
3.	Extrinsic recruitment of transcriptional coregulators by bifunctional nuclear receptor ligands	101
3.1	Abstract	101
3.2	Introduction to nuclear receptors as ligand-regulated factors	102
3.3	The glucocorticoid receptor	104
3.4	Nuclear receptor extrinsic recruitment	105
3.5	Recruitment of non-native proteins to the glucocorticoid receptor	106
3.6	Conclusion and discussion	127
3.7	Materials and methods	129
3.8	References	140
4.	Towards extrinsic recruitment of HDAC-associated complexes to the glucocorticoid receptor	146
4.1	Abstract	146
4.2	Introduction	147
4.3	Synthesis of GR-ligand - HDAC inhibitor conjugates	150
4.4	HDAC inhibition by conjugates	152
4.5	SDex and RU486 conjugated to SAHA can induce GR localization at GREs	154

4.6	Conclusion and discussion	156
4.7	Materials and methods	157
4.8	References	160
5.	Conclusions and Future Directions	163
5.1	Conclusions	163
5.2	Future directions	164
5.3	Concluding remarks	170
5.4	References	171

List of Figures

Figure 1.1 Ligands of coregulatory proteins can control their recruitment	3
Figure 1.2 Transcriptional activators orchestrate PIC formation	3
Figure 1.3 Inhibitors of DNA-TF and TF-coregulator interactions	6
Figure 1.4 Amphipathic activation domain mimics	11
Figure 1.5 Coactivator ligand discovery	14
Figure 2.1 Structures of activator-coactivator complexes	31
Figure 2.2 Schematic of a cell-based luciferase reporter inhibition screen	33
Figure 2.3 Activity of three fusion activators	35
Figure 2.4 Test for inhibition of Gal4-Elf3	37
Figure 2.5 Performance of a high-throughput Gal4-VP16 inhibition assay	38
Figure 2.6 Pilot screen of 2000 compounds	39
Figure 2.7 A gain-of-signal assay for activator-coactivator inhibitors	41
Figure 2.8 Knock-down of firefly luciferase gene by miR-luc	42
Figure 2.9 Dose response of transcriptional inhibition and gain-of-signal	43
Figure 2.10 The CBP/p300 coactivator and its ligands	47
Figure 2.11 Fluorescence polarization assay of the MLL-KIX interaction	49
Figure 2.12 Theoretical detection limit of KIX-MLL FP assay	50
Figure 2.13 Variability of KIX-MLL FP assay	51
Figure 2.14 High-throughput screen of 66882 compounds for inhibitors of KIX-MLL	52
Figure 2.15 Structure of the lichen depside 2'-O-methylanziaic acid	53
Figure 2.16 Activity of lichen extracts in MLL-KIX and other FP assays	54
Figure 2.17 Structures of sekikaic acid and microphyllinic acids	55
Figure 2.18 Inhibition of FI-MLL19 – KIX interaction by sekikaic acid	57
Figure 2.19 Inhibition of FI-MLL15dm – KIX interaction by sekikaic acid	58

Figure 2.20 Inhibition of FI-KID29 – KIX interaction by sekikaic acid	59
Figure 2.21 1D- ¹ H-NMR of sekikaic acid binding to KIX	60
Figure 2.22 ¹ H-NMR of sekikaic acid to monitor indications of aggregation at high concentration	62
Figure 2.23 1H-15N HSQC of KIX with sekikaic acid	64
Figure 2.24 Map of chemical shift changes in sekikaic acid-bound KIX	65
Figure 2.26 Cytotoxicity of sekikaic acid	66
Figure 2.27 Effect of sekikaic acid on MLL and CREB activated transcription	66
Figure 2.28 KIX-MLL two-hybrid	67
Figure 2.29 Competitive inhibition of KIX tracers with lecanoric acid and lobaric acid	69
Figure 2.30 Carboxyl-derivatives of sekikaic acid	70
Figure 3.1 Nuclear receptor extrinsic recruitment	105
Figure 3.2 Ligands for constructing bifunctional molecules	107
Figure 3.3 Functionalization of ligands	109
Figure 3.4 Conjugates of GR ligands with FK506	111
Figure 3.5 Activity of unconjugated ligands in the three-hybrid assay	112
Figure 3.6 Activity of SDex-FK506 conjugates	113
Figure 3.7 Activity of RU486-FK506 conjugates	114
Figure 3.8 Overview of compounds chosen for studies with full-length GR	115
Figure 3.9 Binding affinity of conjugates for GR	116
Figure 3.10 Functional recruitment of VP16-FKBP to GR by SDex-conjugates	118
Figure 3.11 Functional recruitment of VP16-FKBP to GR by RU486-conjugates	120
Figure 3.12 Effect of competitive binding of RU486 conjugates in the presence of Dex	121
Figure 3.13 Expression levels of VP16-FKBP and endogenous FKBP1A	124
Figure 3.14 Effect of RU468-O ₃ -FK506 on expression of endogenous genes	125

Figure 3.15 Chromatin immunoprecipitation of GR and VP16-FKBP-myc ₆	127
Figure 4.1 HDAC inhibitors	148
Figure 4.2 Recruitment of an HDAC complex with HDAC inhibitors	150
Figure 4.3 The HDAC inhibitor SAHA	151
Figure 4.4 Synthesis of GR-ligand - HDAC inhibitor conjugates	152
Figure 4.5 Test of HDAC1 inhibition	153
Figure 4.6 Conjugate-induced GR binding to the FKBP5 promoter	155
Figure 4.7 Effect on acetylation levels	156
Figure 5.1 Multiplex screening of coactivator-activator interactions	165
Figure 5.2 Identifying binding site-proximal amines to target with an electrophile	167
Figure 5.3 Synthesis of a “bumped” ligand of F37V-FKBP1A	168
Figure 5.4 Synthesis of an intermediate for SDex conjugation with p300 HAT inhibitor	170

List of Abbreviations

ATP	adenosine triphosphate
CID	chemical inducer of dimerization
CBP	CREB binding protein
ChIP	chromatin immunoprecipitation
CMV	cytomegalovirus
DBD	DNA binding domain
DSC	disuccinimidyl carbonate
DMEM	Dulbecco's modified eagle medium
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
FBS	fetal bovine serum
FDA	US Food and Drug Administration
FKBP	FK506 binding protein
FP	fluorescence polarization
EC50	half maximal effective concentration
GR	glucocorticoid receptor
GRE	glucocorticoid response element
HAT	histone acetyltransferase
HDAC	histone deacetylase
HGNC	HUGO gene nomenclature committee
HPLC	high-performance liquid chromatography
HTS	high-throughput screening
HSQC	heteronuclear single-quantum correlation
IC50	half maximal inhibitory concentration
iTAD	isoxazolidine-based TAD

IKK	inhibitor of kabbeB kinase
LBD	ligand-binding domain
Med	mediator complex subunit
miRNA	microRNA
MLL	mixed lineage leukemia protein
NMO	N-methylmorpholine-N-oxide
NMR	nuclear magnetic resonance
mRNA	messenger RNA
MsCl	methanesulfonyl-chloride
NR	nuclear receptor
LNA	locked nucleic acids
p300	E1A binding protein p300
PCR	polymerase chain reaction
PIC	pre-initiation complex
Pol2	RNA polymerase II
PR	progesterone receptor
PTM	post-translational modification
qPCR	quantitative PCR
REST	RE1-silencing transcription factor
RNA	ribonucleic acid
RT-qPCR	reverse transcription qPCR
SAHA	suberoylanilide hydroxamic acid
shRNA	short hairpin RNA
STAT3	signal transducer and activator of transcription 3
SUMO	small ubiquitin-like modifier
TAD	transcriptional activation domain
TBS	tert-butyldimethylsilyl
TPAP	tetrapropylammonium perruthenate
tBu	tert-butyl
TF	transcription factor
TFIID	transcription factor II D

TRE	tetracycline responsive enhancer
tTA	tetracycline-responsive transcriptional activator
UV	ultraviolet
VIS	visible
VP16	herpex simplex virion protein 16
Å	angstrom

Abstract

Exogenous Control of the Assembly of Transcriptional Complexes at Gene Promoters

by

Jonas Westergaard Hojfeldt

Chair: Anna K. Mapp

Establishment of gene expression patterns defines cellular function and the first critical step in this process is transcription. Exogenous agents that can alter transcription hold tremendous utility as chemical genetic probes and as therapeutics. In my dissertation research I have pursued strategies for controlling the transcription of genes with small drug-like molecules.

The transcriptional status of genes is defined by coregulatory complexes recruited to gene promoters. To achieve control of this recruitment, small molecules that bind to these proteins must be discovered. By use of a high-throughput biochemical screen, sekikaic acid was discovered as a potent ligand

of the transcriptional coactivator protein CBP. Sekikaic acid binds to the KIX domain of CBP and its binding precludes the *in vitro* interaction of transcriptional activation domains from the human transcription factors MLL and CREB to two distinct sites on KIX.

In order for a small molecule to recruit target proteins to individual genes, it must have the ability to localize to DNA. Nuclear receptors were identified as potential targets that could tether bifunctional molecules to their target genes. Conjugates of glucocorticoid receptor (GR) ligands and a ligand of the FK506 binding protein (FKBP) were prepared, and it was demonstrated that the conjugates bind the receptor and recruit an FKBP fusion protein that influence the transcriptional control of target genes in concert with intrinsic GR activity. The ability to recruit nuclear receptor extrinsic functionality to target genes with receptor targeting bifunctional molecules has potential to greatly increase our level of control over the therapeutically relevant genes that can be targeted via nuclear receptors.

The traditional paradigm of artificial transcription factor design is an attempt to mimic natural transcription factors by binding their direct interaction partners. The arsenal of ligands that bind these proteins is limiting, and instead untraditional recruitment targets have been identified: the enzymatic subunits of chromatin modifying complexes. High affinity ligands of these subunits have been discovered and it is demonstrated in this work that a histone deacetylase inhibitor can be incorporated into NR targeting bifunctional recruiters that retain the ability of each conjugate partner to bind its respective protein target.

Chapter 1

Artificial transcriptional regulators - ligands and bifunctional recruiters*

1.1 Abstract

The human body is comprised of several hundred distinct cell types that develop from a single cell through regulated expression of individual genes in the common genome.¹ The first critical step in this process is transcription and is governed by a large number of transcription factors. Small molecules that can alter transcription hold tremendous utility as chemical genetic probes and as therapeutics. This introductory chapter defines the principles of direct transcriptional regulation by small molecules and surveys previous discoveries of small ligands (drug-like molecules and short peptides) that bind transcriptional coregulatory proteins and have the potential or demonstrated capability to control assembly of coregulatory proteins at gene promoters.

1.2 Introduction

Transcription, the initial step in gene expression, brings to life an organism's heritable code. The master regulators of this process, proteins termed transcription factors (TFs), generate immense cellular diversity from a common genetic template by activating or repressing specific gene targets.² Given this critical role, it is not surprising that misregulated transcription strongly correlates with disease.³ Thus, one strategy for combating disease is to modulate transcription in a way that restores the cell's healthy phenotype.⁴ The highly

* This chapter is adapted from a published review article: Hojfeldt, J. W., Van Dyke, A. R. & Mapp, A. K. *Chem. Soc. Rev.*, 2011, DOI: 10.1039/C1CS15050B - Reproduced by permission of The Royal Society of Chemistry (RSC).

regulated action of transcription factors can be modulated indirectly by small drug-like molecules targeting cellular signaling events.⁵ Predictable gene selective effects through these targets can be realized with a complete understanding of signaling networks and the transcription factors they effect, and are reliant on non-aberrant signal cascades in disease states. Neither of these criteria is a reality (e.g. non-canonical transcription factor targets of nuclear β -catenin^{6,7} and constitutive activation of NF- κ B in Hodgkin lymphoma^{8,9}), thus direct targeting of transcription processes is an attractive goal. Transcriptional coregulatory complexes recruited to gene promoters define the transcriptional status of genes, and small molecules capable of recruiting these complexes to promoters or block their recruitment by endogenous transcription factors will directly affect the transcription of genes. To achieve control of recruitment of target proteins to gene promoters, small molecules that bind to these proteins (ligands of transcriptional coregulatory proteins) must be discovered and linked to functionality that can tether the recruiter at a target gene,¹⁰ or these same ligands may be used as negative recruiters by blocking interactions between target proteins and endogenous recruiters (transcription factors) (Figure 1.1).¹¹

1.3 Basic mechanisms of transcription initiation

All protein-coding genes are transcribed by RNA polymerase II (Pol2), which is a multi-subunit enzyme. Pol2 must bind at a gene promoter, located near the beginning of a gene, together with a set of general transcription factors (TFIIA-TFIIF) forming what is called the pre-initiation complex (PIC). At many genes the formation of the PIC is the limiting step in transcriptional activation, while certain steps such as promoter clearance (in which the bound polymerase proceeds from the promoter along the gene) can be limiting in other contexts.^{12,13} Thus, in most cases transcriptional activation faces the challenge of facilitating PIC formation at a promoter (Figure 1.2), while transcriptional repression results from the blockage of this assembly.

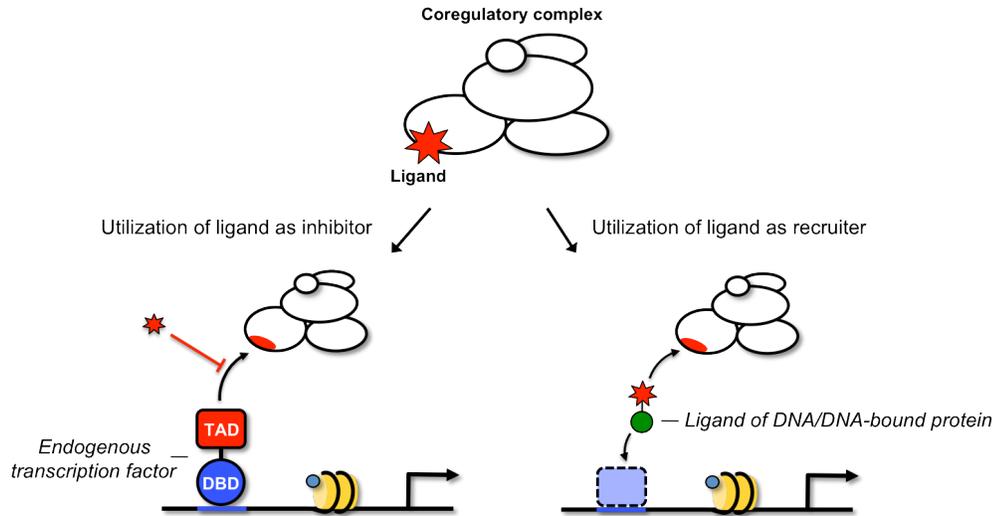


Figure 1.1 Ligands of coregulatory proteins can control their recruitment Ligands that bind transcriptional coregulatory proteins can be incorporated into bifunctional molecules that can recruit target complex to promoters, or, for ligands that perturb interaction with endogenous transcription factors, they can be used to inhibit recruitment.

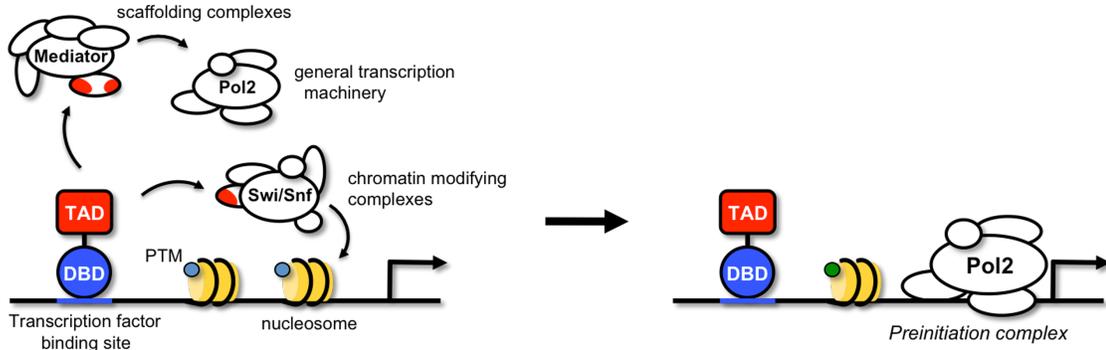


Figure 1.2 Transcriptional activators orchestrate PIC formation One class of eukaryotic transcription factors are transcriptional activators. These are bifunctional molecules, minimally composed of a DNA-binding domain (DBD) that provides gene-targeting specificity and a transcriptional activation domain (TAD) that contacts multiple surfaces within transcriptional coregulatory complexes. Through these contacts the TAD orchestrates recruitment of factors that remodel chromatin (catalyze nucleosome repositioning and changes in post-translation modification (PTM) patterns of histones) and factors that scaffold recruitment of the general transcription factors to form the preinitiation complex (PIC).

An intrinsic barrier to PIC formation is chromatin, the protein scaffold of genomic DNA. In chromatin, 146 DNA base pairs wrap around an octamer of histone proteins to form the smallest chromatin entity called a nucleosome.¹⁴ Nucleosomes placed along a strand of DNA form a structure resembling beads on a string. Chromatin facilitates packaging of approximately two meters of DNA in each human cell and maintains a dynamically accessible genome by forming different higher-order structures in which nucleosomes are highly or loosely compacted. Condensed chromatin is associated with inactive genes; this is logical as the DNA in these structures is buried, hindering its interaction with protein factors. Even in open chromatin, also called euchromatin, single nucleosomes can pose a barrier.¹⁵ For example, a nucleosome located at a TATA box (a common sequence element in promoters) prevents binding of the TFIID subunit, TBP, and this alone prevents formation of the PIC.¹⁶ Therefore, one class of transcriptional cofactors that often needs to be recruited to genes to regulate PIC formation are comprised of ATP-dependent chromatin remodeling enzymes, which catalyze the repositioning of nucleosomes.¹⁷ Chromatin also serves a role beyond a simple packaging scaffold. Residues in the amino and carboxyl tails of the histone proteins are subject to an extensive combination of posttranslational modifications. This includes acetylation, methylation, phosphorylation, ubiquitylation, and SUMOylation.¹⁸ These modifications serve as marks allowing protein factors to distinguish a location in the genome independent of the underlying DNA sequence. Some of these marks are inherited during subsequent cell divisions and are therefore classified as epigenetic factors. The enzymes that place or remove these marks constitute an additional set of transcriptional regulators that are recruited to genes, and the study of these proteins and their associated regulation mechanisms are currently an intense research focus.

1.4 Principles of inhibiting recruitment of coregulatory proteins to DNA

It is the role of transcription factors to orchestrate the previously discussed processes at gene promoters: chromatin remodeling, cofactor recruitment, and

PIC assembly (Figure 1.2). TFs are minimally composed of a sequence selective DNA-binding domain (DBD) that specifies the locality of action and a transcriptional regulatory domain that recruits necessary coregulatory complexes through a series of protein-protein interactions. One obvious strategy to modulate transcription is through inhibition of these interactions.

Small molecules binding directly to DNA are necessarily required to bind to a large surface area to gain sequence specificity and must in addition bind with high affinity to compete with sub-nanomolar affinities typical of DBD-DNA interactions.¹⁹ This criterion makes even designed DNA-binding molecules large structures that lack the attractive advantages of small molecules. Progress is being made in the field, as relatively short synthetic oligomers have been demonstrated to bind genomic DNA in living cells and block binding of transcription factors: pyrrole-imidazole polyamides²⁰ and locked nucleic acids (LNAs)²¹ (Figure 1.3). However, the LNAs need to be transfected into cells, and the polyamides show general cytotoxicity at concentrations very close to efficacies doses, which may be an indication that their effect is indirect or is non-specific. If strategies could be devised to give these structures improved properties in biological systems,^{22,23} programmable sequence-selective DNA ligands may emerge with many attractive applications. Inhibition of DNA binding can also be achieved through ligand binding to the transcription factor rather than DNA; for example, an inhibitor of homodimerization of the human transcription factor, STAT3, prevents STAT3 binding to DNA²⁴ (Figure 1.3).

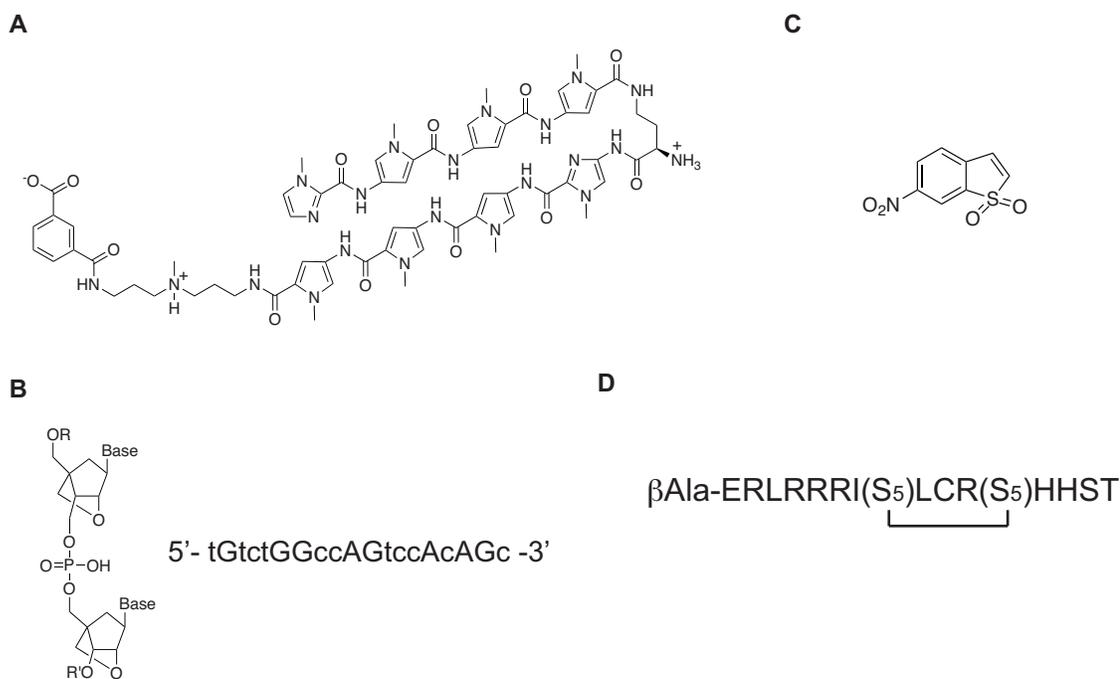


Figure 1.3 Inhibitors of DNA-TF and TF-coregulator interactions A. The shown pyrrole-imidazole polyamide inhibits glucocorticoid receptor binding to DNA; B. The structure of two connected locked nucleic acids is shown on left. The sequence of a mixed LNA-DNA oligomer that blocks progesterone receptor binding to DNA is shown to the right. LNA-monomers are in capital letters; C. Inhibitor of STAT3 homodimerization; D. Sequence of an α -helix stabilized peptide. The side chains of unnatural (S)-C α -methyl-C α -alkenyl amino acid, S₅, have been connected with ring-closing metathesis.

Inhibiting the recruitment of coregulatory proteins may also be achieved through inhibition of their interactions with transcription factors. In these interactions one partner will typically have a relatively concave surface that the other partner docks into. The class of amphipathic transcriptional activators, for example, contains short domains that comprise the majority of the activation potential.²⁵⁻²⁷ These short peptides are furthermore thought to bind their coactivator targets as α -helices,^{28,29} and bind to shallow grooves present in the coactivators.^{30,31} Thus, for the interaction between amphipathic transcriptional activation domains and their target coactivators, small molecule ligands that competitively inhibit this interaction are far more likely to bind the invaginated surface (the coactivator), where hydrophobic contacts can be buried. One consequence of this is that the

gene specificity will not be derived from the activator, but instead be defined by the binding site on the target coactivator, which may or may not be shared with additional activators. With a fewer number of coactivators relative to transcription factors,³² the coactivator likely has lower selectivity, although individual coactivators can have several distinct interaction domains and binding surfaces.³³ Coactivator ligands have been discovered in recent years (discussed in section 1.6), and recently, an inhibitor that binds the DNA-bound factor and precludes coactivator recruitment was also reported³⁴ (Figure 1.3D).

1.5 Tactics and limitations to positively modulate recruitment

No therapeutics are yet developed that inhibit recruitment of coregulatory components by competitive blockage of binding sites, but in contrast a vast number of drugs lead to modulation of transcription factor function directly or indirectly. These mechanisms can stimulate recruitment or change recruitment patterns, rather than merely eliminating a subset of interactions. One class of drugs binds directly to a unique family of transcription factors: the nuclear receptors (NRs). NRs bind endogenous lipophilic ligands,³⁵ and upon binding, these small molecules cause a conformational change in their cognate receptor and modulate an interaction surface that recruits coregulatory proteins to elicit a transcriptional response. It has proven possible to design synthetic ligands of nuclear receptors that stabilize conformations leading to both transcriptionally activating and repressing states.³⁶ A fundamental limitation of this approach, though, is that the drug acts entirely through the nuclear receptor surfaces to intervene in the transcription process; it is unable to orchestrate protein recruitment on its own. Thus, mutations in the nuclear receptor can render these drugs ineffective, as is observed in androgen-ablated prostate cancers that evolve into a lethal, ligand-independent state.³⁷

The nuclear receptors have evolved to respond to binding of ligands, but are unique in this ability among transcription factors. In contrast, all transcription factors are regulated by intracellular signaling events. Drugs targeting cell

surface receptors and enzymes in these signaling cascades constitute the largest class of transcription-targeting drugs.³⁸ While these approaches have yielded successful drugs, they do not represent a general solution for gene-targeted therapy. One disadvantage of targeting the upstream signaling events is that there can be significant overlap between signaling networks, resulting in the modulation of undesired transcription factors and, consequently, undesired gene regulation.³⁹ Additionally, targeting an upstream cascade member is ineffective when a downstream member is autonomous or non-functional due to an alteration such as a constitutively activating mutation or a gene deletion.⁹

A case study that highlights the limitation of transcription modulation through TF function is evident from the role of the neural repressor REST in medulloblastoma, a childhood-associated brain cancer. REST is a transcription factor that represses neuronal differentiation genes and the expression of REST must be turned off during proper neuronal development.⁴⁰ It was discovered that the cancerous tissue of medulloblastoma results from improperly developed neuronal lineage cells that still express REST.⁴¹ Inhibition of REST is not sufficient to reactivate its target genes and REST does not possess latent activation function that can be stimulated through signaling events. In addition, no other transcription factor that can bind and reactivate REST target genes is known. Thus, none of the drug strategies outlined above could be applied to control transcription of REST target genes. However, it has been shown that a synthetic chimera of REST's DNA-binding domain and a strong activation domain from the viral transcription factor VP16 does produce an artificial transcription factor that can reactivate the neuronal differentiation genes and kill the tumor cells.⁴¹ Therefore, drug-like molecules that could directly address the transcriptional status of REST-regulated genes could be the ultimate goal for treating medulloblastoma.

Thus, it is clear that the agents in need of development are artificial transcription factors that on their own can localize at target genes and recruit coregulatory

complexes. Both of these events - localization and recruitment - simply involve binding interactions: binding to DNA or to a DNA-bound protein will localize a molecule at the DNA, and binding a transcriptional complex member will recruit this complex to the locale of the ligand. Thus, artificial transcription factors are bifunctional molecules that can be constructed by linking together DNA-localizing moieties with molecules that bind transcription complexes,¹⁰ and they differ from the majority of drugs: their mode of action is a gain of function; they do not inhibit an endogenous process, but aim to orchestrate one. Such bifunctional molecules mimic the design of endogenous transcription factors, with distinct DNA-localizing and regulatory domains. They are also conceptually related to the chemical inducer of dimerization (CID) concept originally developed in the labs of Stuart Schreiber and Gerald Crabtree.⁴² CIDs are made by linking together two ligands that each can bind a target protein; the conjugate consequently brings together (or dimerizes) the two target proteins. CIDs are excellent general tools as many cellular processes occur simply as a consequence of co-localization⁴³, and this is also true for events leading to transcription.⁴⁴ In the case of artificial transcription factors, they should reproduce the basic mechanism of transcription: bringing together the gene promoter and relevant transcriptional complexes.

The ligands previously described as potential recruitment inhibitors (as well as NR ligands) conveniently meet the requirements of each half of an artificial transcription factor: these are either ligands of DNA or DNA-bound proteins, or they are ligands of coregulatory proteins. With advances being made on both fronts, realization of artificial transcription factors is increasingly plausible.⁴⁵ The direct DNA-binding molecules are fundamentally different structures and labs specialize in their development.^{46,47} The following section surveys discoveries of ligands of coactivator proteins.

1.6 Discovery of ligands of coactivator proteins

1.6.1 Synthetic transcriptional activation domain mimics

Transcriptional activators recruit coactivator proteins and the Pol2 holoenzyme to genes that are selected by its DNA-binding domain. With the goal of developing bifunctional molecules that can orchestrate a transcriptional response, much effort has been devoted towards understanding which coactivators participate in direct binding interactions with endogenous transcriptional activation domains. Such knowledge would aid in the identification of ligands of these surfaces, ligands that themselves could then be used as artificial transcriptional activation domains. The certain identification of direct binding targets, however, has remained elusive, and instead nearly all coactivators have been proposed as a direct binding target of activators. It appears likely that activators make contact with more than one coactivator in an ordered fashion and that these interactions are dependent on the promoter context.^{48,49} With this presumptive scenario of multiple elusive binding partners, we and others have pursued strategies for discovering artificial activation domains in which essential features of natural activation domains are incorporated into synthetic structures, rather than screening for ligands of putative activation domain targets.^{29,50-52}

Amphipathic peptides

Early studies of natural transcriptional activators identified short sequences of interspersed acidic and hydrophobic residues that could reproduce most of the activation potential when fused to a DNA-binding domain. These short sequences can be fused to heterologous DNA-binding domains, which on their own lack activation function, transforming them into transcriptional activators.²⁵⁻²⁷ Further studies have identified other classes of minimal activation domains, but the acidic activators remain the largest and most potent class in all eukaryotes. Comparison of such minimal acidic activation domains from different activators revealed no sequence homology. Instead, a pattern emerged that these sequences could fold as alpha helices in which acidic or polar residues would align on one face of the helix and the hydrophobic residues on the other side, forming what is termed an amphipathic helix. Indeed a 15-mer peptide was designed as a test of this hypothesis that did not share sequence homology with

any identified activators, but was predicted to form a similar amphipathic helix (Figure 1.4A). This 15-mer peptide, termed amphipathic helix (AH), activates reporter genes to 20% the level of full length Gal4 activator when fused the Gal4 DNA-binding domain.²⁹ This was the first rationally designed artificial activation domain and it gave credence to the idea that the amphipathic helix and not a particular sequence of amino acids was the determining factor in the activation potential of this class of activators. Taken together with the multi-partner binding profiles of activators, it could be proposed that multiple coactivator proteins contain permissive binding surfaces for such amphipathic helices.^{29,53,54}

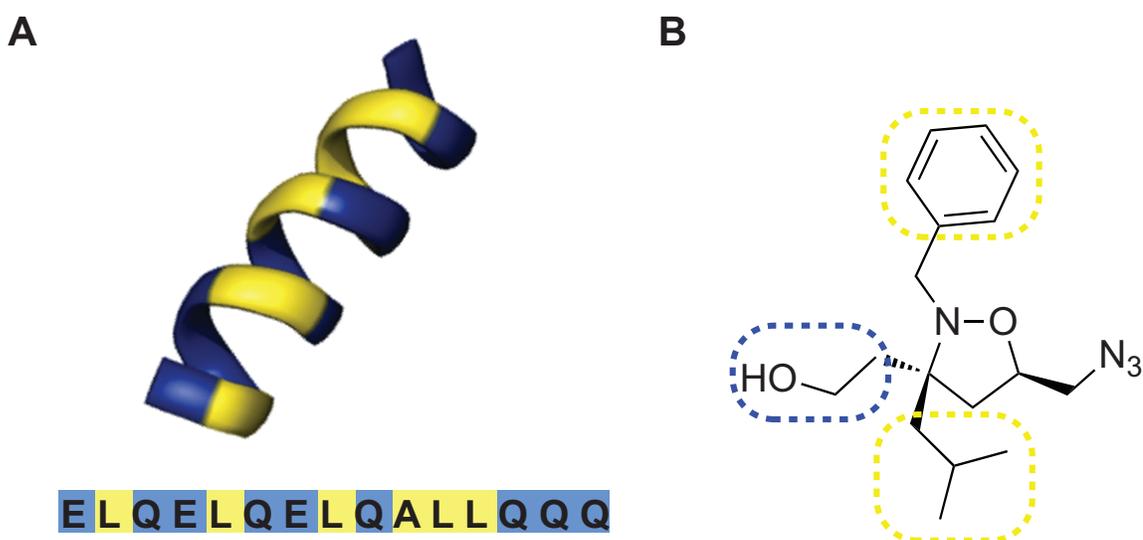


Figure 1.4 Amphipathic activation domain mimics A. An amphipathic helix with polar residues (blue) on one face of the helix and hydrophobic residues (yellow) on the other. The sequence of a peptide is shown below; B. Structure of an isoxazolidine based transcriptional activation domain mimic (iTAD). Hydrophobic and polar moieties are highlighted.

A small molecule transcriptional activation domain mimic

In order to create a small synthetic molecule that would mimic the physicochemical nature of amphipathic helices, our lab decorated a 5-membered heterocyclic scaffold with combinations of polar and hydrophobic amino acid side chains commonly featured in minimal activation domains (Figure 1.4B). Members of this small series of isoxazolidines demonstrated activation potential in an *in*

vitro assay comparable to a minimal activation domain of a natural acidic activator, and they represent the first examples of small-molecule activation domain mimics.^{50,51} Since an organic molecule cannot easily be attached to a protein DNA-binding domain; the isoxazolidine was attached to a second small molecule, methotrexate, that can bind a protein fusion of dihydrofolate reductase and a DNA-binding domain. Thus in this iteration of a chemical inducer of dimerization, one end of the molecule binds an engineered protein, while the other binds native transcriptional components. The isoxazolidine-based transcriptional activation domain mimic (iTAD) also demonstrated activity in cell-based assays in which it activates a luciferase reporter 80-fold above a control lacking the activation domain.⁵⁵ In further cell-based studies of iTAD analogs, an amphipathic balance proved to be a critical determinant for function.⁵² While hydrophobic contacts have been proposed to be the strongest contributors of binding affinity to coactivator targets,⁵⁶ it appears that in a small molecule they must be balanced by polar groups for activation potential. This property is similar to natural activation domains and it is plausible that increased hydrophobicity causes the structures to become sequestered by non-specific interactions in the cell. A photo-crosslinker has been incorporated in the linker between an iTAD and a DNA localizing functionality to capture its direct binding targets. Out of several yet unidentified proteins that are cross-linked to the iTAD, the mammalian coactivator CBP was demonstrated to be one of its targets. By an HSQC-NMR experiment with a domain from CBP, the iTAD was shown to bind a site shared by several natural activators.⁵⁷ Thus, a small molecule designed to mimic the amphipathic nature of natural transcriptional activation domains shares at least one binding surface with these, and further demonstrates the permissive nature of these. This increases the likelihood that additional coactivators with permissive binding surfaces are amongst the additional uncharacterized binding partners, and the multi-binding partner profile may be a functionally determining feature of iTADs.

1.6.2 Screening for coactivator ligands

An alternative strategy to rationally design mimics of natural activation domains is to screen for structures with innate activation potential. Given the power of modern molecular biology, large libraries of peptide fragments attached to DNA-binding domains can be generated. For example, a screen of peptides encoded by random DNA-fragments from *E. coli*, yielded potent activators that appear by sequence to be prototypical amphipathic activation domains.²⁷ In principle, it is possible to conduct a similar screen to identify small molecules that activate a given reporter gene (screening without a defined binding target). However, in order to be successful, each member of the screening library would need to possess a DNA-localizing functionality that targets the small molecule to the reporter gene. In essence, each library member would need to be a bifunctional molecule. To circumvent this issue, current screening efforts first involve screening for binding to putative coactivator proteins, a challenging task in itself given our limited knowledge of transcriptionally relevant binding interactions. The hits obtained from these screens must then be incorporated into bifunctional molecules to test their activation potential (Figure 1.5C). The following cases are representative examples illustrating high-throughput screening (HTS) efforts in which ligands of individual transcriptional coactivator proteins were successfully identified.

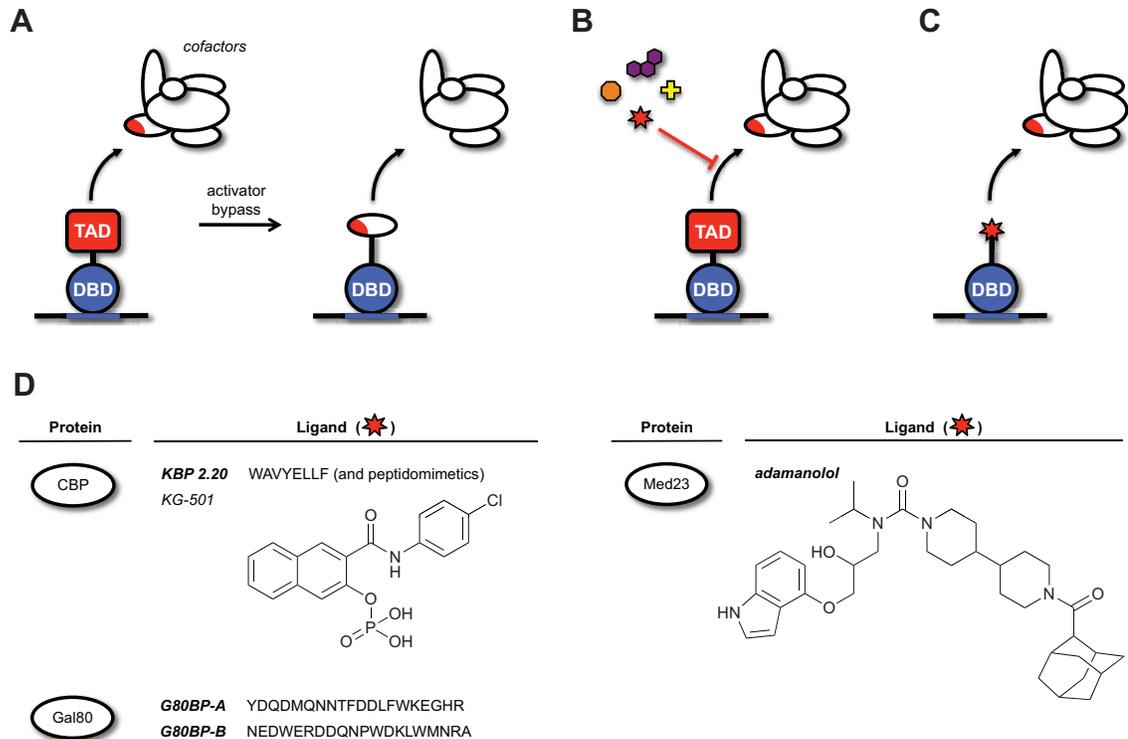


Figure 1.5 Coactivator ligand discovery A. An activator bypass experiment can be used to identify coactivator targets whose direct recruitment suffices to activate transcription; B. Ligands of coactivators can be identified from screening for inhibitors of interaction between TAD and coactivators; C. A ligand of a coactivator target can be tethered to a DBD to test its ability to recruit coactivators to the DNA; D. Examples of cofactor proteins for which ligands (small molecules or peptides) have been identified (bold: ligand has been demonstrated to function as an artificial TAD).

Ligands of the CBP KIX domain

The coactivator CREB binding protein (CBP) is one of the known direct binding targets of mammalian transcriptional activators. It is a direct binding target of CREB, which binds to the KIX domain in CBP.⁵⁸ Biochemical evidence suggests that CBP is a direct binding target for numerous other activators, and there is surmounting evidence that CBP acts as a transcriptional hub interacting with hundreds of transcription factors and coactivators.³³ CBP functions as both a scaffolding component and as a chromatin modifier via its intrinsic histone acetyltransferase (HAT) domain. This hub-like role of CBP indicates that its direct recruitment to promoters may suffice for recruitment of other necessary cofactors and nucleation of PIC assembly. To test this hypothesis, ligands of CBP were

sought through screening efforts. A peptide phage display selection was used to discover peptides that bind the KIX domain of CBP.⁵⁹ An eight amino acid peptide was discovered that binds to the KIX domain with a 16 μ M affinity (Figure 1.5D). This is comparable to the affinity reported for natural activation domains that, in general, have weak affinity for their coactivator targets. The 8-mer peptide activates transcription in cells to levels comparable to natural TADs. The sequence of this peptide is amphipathic like most natural activation domains and shares a binding surface on KIX with several of these. Because of this resemblance, it is likely to have a multi-partner binding profile. Our lab has studied peptidomimetic versions (d-peptide, beta-peptide, peptoid) of the same sequence.⁶⁰ Unexpectedly, the binding affinity for KIX is largely unchanged in these analogs, indicating that the binding surface in the KIX domain is permissive of structural differences in the ligands. Thus, it appears that the amphipathic character is the important feature. The CBP KIX domain was also used in an NMR-based screen for small molecule ligands and the molecule KG-501 (naphthol AS-E phosphate) (Figure 1.5D) was discovered as a high micromolar binder that competes with CREB for binding.⁶¹ While not yet reported, it would be interesting if this ligand binds other coactivators and if it can function as an activation domain when used as part of a bifunctional molecule that localizes it to DNA. Other notable synthetic biopolymer and peptoid ligands of the KIX domain have been reported that potentially activate transcription.^{62,63}

Ligands of human mediator subunit Med23

One coactivator complex is believed to play an essential role in all transcriptional activation by bridging transcription factors to the general transcriptional machinery. This complex is named Mediator. A screen for small molecule inhibitors of the interaction between the human transcription factor Elf3 and its binding target within the mediator complex, Med23, was performed. Evidence suggests that Elf3 binds Med23 through an 8-mer alpha-helix containing a critical tryptophan residue and a small molecule, adamanolol (Figure 1.5D), was discovered that blocks this interaction.⁶⁴ Similar to the discovery of KG-501, the

purpose of this screen was to discover inhibitors of the activator:coactivator interaction. As found with several other inhibitors of activator-coactivator interactions, adamanolol binds the coactivator. The researchers subsequently used a derivative of this molecule as part of a bifunctional DNA-localizing conjugate to show that it can activate transcription *in vitro* and in cells.^{65,66} Through a series of synthetic derivatives, the authors found a molecule with enhanced transcriptional activation potency as well as enhanced binding to the original target coactivator Med23. As the authors point out, it will be valuable to learn if the derivative is more potent due to a more selective binding of Med23 over other non-productive cellular components or whether the enhanced potency stems from increased binding to other coactivators (the original molecule, adamanolol, was shown to bind multiple proteins in cell extracts).⁶⁴

Ligands of activator masking proteins

A different example of the multi-partner binding profile of a transcriptional activation domain is seen with masking proteins. Some transcriptional activators are kept inactive by binding of the activation domain to a masking protein. Unmasking of the activation domain, then, is under the control of signaling events. One of the best examples of this interaction is between the yeast activator Gal4 and its masking protein Gal80.⁶⁷ Gal4 can bind its target genes in the absence of galactose, but Gal80 binds to and masks the activation domain of Gal4. In the presence of galactose, Gal80 binds Gal3 and its repression of Gal4 is relieved. A phage display library was screened to select peptides that bind to Gal80.⁶⁸ Two peptide binders identified from this screen were then attached to the DNA binding domain of Gal4 and found to activate transcription to approximately 35% the level of the minimal activation domain of Gal4. Although the peptides do not share significant homology with the Gal4 activation domain, they share at least one coactivator target, Gal11, and Gal80 masks their activity in yeast.

In mammalian cells, another TAD-masking protein has received significant attention with regards to discovery of small molecule ligands. Hdm2 is a masking

protein of the human tumor suppressor and transcriptional activator p53. The masking of p53 by Hdm2 helps cancer cells evade apoptosis. Intense efforts towards finding small molecule inhibitors of this interaction have therefore been undertaken and yielded a number of drug candidates that have progressed to clinical trials.⁶⁹ Analogous to the Gal80 ligands, it is conceivable that these Hdm2 inhibitors could be used as transcriptional activation domains. However, the inhibitors facilitate p53 transcription, indicating that they are selective for Hdm2 over essential p53 coactivators. Nonetheless, it may be a fruitful effort to take a further look at some of the many structures related to the current lead drugs, as some of these may be less selective for the masking protein compared to coactivator targets. Hypothetically, the discovery of ligands that bind to both coactivators and masking proteins could yield artificial transcription factors that additionally are responsive to physiological signals.

Coactivator ligands that bind non-competitively with activators

For all of the above discovered artificial activation domains, it is unclear if they function by binding one coactivator (the one they were screened against) or by binding multiple coactivators. While the multi-partner binding profile of activators is well supported, there is evidence that direct recruitment of a single coactivator can suffice to activate transcription. Evidence for this comes from so-called activator bypass studies (Figure 1.5A). In these studies a DNA-binding domain is fused directly to a coactivator and the chimera is tested for activation potential. In yeast, such a study identified several coactivator fusions that activate transcription.⁷⁰ This is not to say that other coactivators are not recruited to the promoter, but it would have to be indirectly. The important point is that a ligand of those coactivators that function in the bypass experiment should, on its own, function as an activation domain.

In a yeast activator bypass study, the coactivator with the strongest activation potential was the mediator subunit Med15. A peptide library was screened in our lab for binding against Med15 and two motifs were discovered that bind Med15

and possess activation potential.⁷¹ These peptides bind to separate surfaces on Med15, compared to a naturally-derived activator, and it would therefore not necessarily be expected that they, like natural activation domains, would bind other coactivators. These novel Med15 ligands had lower levels of activation potential than a minimal natural activation domain, despite having similar binding affinity. It may be that the binding orientation of Med15 is important, but it may also indicate that binding a larger set of coactivators is needed for higher activation. To discover a ligand for a coactivator complex that synergizes with Med15 in bypass experiments, a SAGA-complex member (Tra1) was screened for peptide ligands.⁷² In this case, however, the discovered peptides bind the same site as a natural activator and thus may be considered a typical amphipatic TAD. Indeed the activity of the Tra1 ligands is dependent on other coactivators, such as Med15, in addition to Tra1. Nevertheless, at least one of the Tra1 ligands synergizes with the previously mentioned Med15 ligands.

1.7 High affinity, drug-like ligands of transcriptional coregulatory proteins

Inhibitors of chromatin modifying enzymes as potential recruiters

One significant challenge in the discovery of molecules that can bind and recruit coregulatory transcription complexes is the nature of the shallow protein surfaces that comprise the target binding surfaces of transcription factors. The binding affinities in the native interactions are often weak (high nanomolar to low micromolar), and these structures fall in the regime of traditionally undruggable targets.⁷³ Outside of transcription factor interaction surfaces, however, there are numerous proteins involved in transcription with known high affinity ligands. Chromatin modifying enzymes are emerging drug targets with many active inhibitor discovery programs.⁷⁴ Most numerous among these targets are the enzymes responsible for controlling methylation and acetylation patterns of residues in the histone tails. These enzymes lack sequence specific DNA binding capability and are recruited by transcription factors or by proteins recognizing distinct histone modifications. The recruited activity of these enzymes establishes patterns of histone tail modifications, and these epigenetic changes are critical in

many cancers and establish a therapeutic window despite their ubiquitous function in cells. Inhibitors of histone deacetylases were the first members of this class and include suberoylanilide hydroxamic acid (SAHA), which is an FDA approved drug (Vorinostat).⁷⁵ The use of these inhibitors as bifunctional recruiters has to my knowledge never been proposed and this would perhaps be considered unintuitive since their activity would be inhibited by the recruiting moiety. I hypothesize that they can nonetheless be useful when recruited, because of the prospective corecruitment of their native multi-protein corepressor complexes, and this new strategy is explored in Chapter 4 of this thesis. Conjugates of SAHA to a DNA-binding pyrrole-imidazole polyamides were recently reported, although the rationale for these structures was not to achieve recruitment of co-repressor complexes to target genes.⁷⁶

Other putative druggable domains

In addition to TF interaction domains and catalytic domains, transcriptional coregulatory proteins have been found to contain other conserved domains, found primarily in nuclear proteins, or reside in complexes that include these domains. One such domain is the bromodomain which recognizes and binds acetylated histones.⁷⁷ Histones are acetylated by histone acetyltransferases (HATs) leading to active transcription, opposing the action of HDACs. Although these domains are not enzymatic, their binding pockets are not shallow grooves like those found in activator-coactivator interfaces. Thus, it is likely that high affinity ligands can be discovered for these domains. The use of such ligands to recruit a bromodomain containing complex would, unlike the recruitment strategy for HDAC complexes, not inhibit any enzymatic function in the recruited complex and would even be predicted to recruit complexes in a natural orientation. Recently the first bromodomain inhibitors were discovered.⁷⁸ Although these inhibitors do not bind the bromodomains of two of the most common coactivator complexes, GCN5 and p300/CBP, or to the general transcription factor, TAFII250 (subunit of TFIID), I predict that ligands of these bromodomains will soon follow,

and it will be exciting to test their potential in bifunctional DNA-localizing conjugates.

1.8 Conclusions

Small molecule ligands of DNA, transcription factors and coregulatory factors all have the potential to directly modulate transcription. As single agents they can inhibit endogenous interactions, and when combined into bifunctional molecules they can themselves control recruitment of coregulators to gene promoters. Currently, NR ligands are the only molecules in this category that are effective drugs, and few of the remaining structures have drug characteristics. NR are also the only DNA-bound proteins with small molecule ligands (peptidic ligands can be used in cell culture²² and stabilized peptides may share drug-like attributes with small molecules³⁴). The DNA-binding ligands are continuously improved, but suffer particularly from cell permeability issues preventing their effective studies in live cells. Thus NR ligands are the only drug-like molecules in our arsenal that can be included in bifunctional transcriptional regulators to contribute DNA-localizing function.

The rational design of small molecule transcriptional activation domains have proven that this biological event indeed can be controlled by drug-like molecules, and has strengthened a model of transcriptional activation occurring through multiple coactivators with similar binding surfaces. Furthermore, screening for ligands of one of these coactivator surfaces typically yields ligands with similar physiochemical characteristics when compared to the amphipathic class of activators, and which therefore may also bind other coactivator surfaces and function in agreement with this model. As illustrated by the Med15 ligands that bind non-overlapping surfaces relative to endogenous activators, it is however also possible to recruit specific coregulatory proteins in a non-native orientation and achieve transcriptional control.

I furthermore believe that there is a large set of available ligands of coregulatory proteins, which includes inhibitors of chromatin modifying enzymes, that do not bind interfaces typically used for recruitment, but that will be worth testing as part of bifunctional molecules. Like the NR ligands these molecules are approved drugs or have drug-like properties, and they bind with high affinity to their targets.

1.9 Research goals

One goal of this research is to expand the class of coactivator ligands. The discovery of several small molecule ligands has been encouraging, but none of these have yet been successfully applied as a chemical genetic probe and all have shown recruiting functionality in only limited scenarios. Different discovery strategies may yield ligands with higher efficacy as inhibitors and greater versatility as recruiters, and this is explored in Chapter 2 of this thesis.

A second goal is to explore the use of high affinity ligands in bifunctional molecules. Direct recruitment of activity to nuclear receptors may greatly increase our control of these important drug targets, and this is the topic of Chapter 3 and Chapter 4. Exploration of recruitment of coregulatory complexes through inhibitors of chromatin modifiers may identify a large number of new targets of artificial transcription factors to consider. This class of proteins has known drugs, and many more will surely be discovered in the near future. The first efforts to incorporate these inhibitors into bifunctional molecules with potential to orchestrate their recruitment is described in Chapter 4.

1.10 References

1. Vickaryous, M.K. & Hall, B.K. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol Rev Camb Philos Soc* **81**, 425-55 (2006).
2. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev* **19**, 1129-55 (2005).

3. Perou, C.M. et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A* **96**, 9212-7 (1999).
4. Pandolfi, P.P. Transcription therapy for cancer. *Oncogene* **20**, 3116-27 (2001).
5. Chang, F. et al. Signal transduction mediated by the Ras/Raf/MEK/ERK pathway from cytokine receptors to transcription factors: potential targeting for therapeutic intervention. *Leukemia* **17**, 1263-93 (2003).
6. Olson, L.E. et al. Homeodomain-mediated beta-catenin-dependent switching events dictate cell-lineage determination. *Cell* **125**, 593-605 (2006).
7. Gordon, M.D. & Nusse, R. Wnt signaling: multiple pathways, multiple receptors, and multiple transcription factors. *J Biol Chem* **281**, 22429-33 (2006).
8. Bargou, R.C. et al. Constitutive nuclear factor-kappaB-RelA activation is required for proliferation and survival of Hodgkin's disease tumor cells. *J Clin Invest* **100**, 2961-9 (1997).
9. Jost, P.J. & Ruland, J. Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. *Blood* **109**, 2700-7 (2007).
10. Ansari, A.Z. & Mapp, A.K. Modular design of artificial transcription factors. *Curr Opin Chem Biol* **6**, 765-72 (2002).
11. Majmudar, C.Y. & Mapp, A.K. Chemical approaches to transcriptional regulation. *Curr Opin Chem Biol* **9**, 467-74 (2005).
12. Roeder, R.G. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* **579**, 909-15 (2005).
13. Orphanides, G. & Reinberg, D. A unified theory of gene expression. *Cell* **108**, 439-51 (2002).
14. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251-60 (1997).
15. Wyrick, J.J. et al. Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418-21 (1999).

16. Martinez-Campa, C. et al. Precise nucleosome positioning and the TATA box dictate requirements for the histone H4 tail and the bromodomain factor Bdf1. *Mol Cell* **15**, 69-81 (2004).
17. Muchardt, C. & Yaniv, M. ATP-dependent chromatin remodelling: SWI/SNF and Co. are on the job. *J Mol Biol* **293**, 187-98 (1999).
18. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
19. Letovsky, J. & Dynan, W.S. Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. *Nucleic Acids Res* **17**, 2639-53 (1989).
20. Muzikar, K.A., Nickols, N.G. & Dervan, P.B. Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci U S A* **106**, 16598-603 (2009).
21. Beane, R.L. et al. Inhibiting gene expression with locked nucleic acids (LNAs) that target chromosomal DNA. *Biochemistry* **46**, 7572-80 (2007).
22. Stafford, R.L., Arndt, H.D., Brezinski, M.L., Ansari, A.Z. & Dervan, P.B. Minimization of a protein-DNA dimerizer. *J Am Chem Soc* **129**, 2660-8 (2007).
23. Gagnon, K.T. et al. Antisense and antigene inhibition of gene expression by cell-permeable oligonucleotide-oligospermine conjugates. *J Am Chem Soc* **133**, 8404-7 (2011).
24. Schust, J., Sperl, B., Hollis, A., Mayer, T.U. & Berg, T. Stattic: a small-molecule inhibitor of STAT3 activation and dimerization. *Chem Biol* **13**, 1235-42 (2006).
25. Hope, I.A. & Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885-94 (1986).
26. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847-53 (1987).
27. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113-9 (1987).
28. 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**, 1310-3 (1997).

29. Giniger, E. & Ptashne, M. Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. *Nature* **330**, 670-2 (1987).
30. 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**, 741-52 (1997).
31. Parker, D. et al. Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Mol Cell* **2**, 353-9 (1998).
32. Messina, D.N., Glasscock, J., Gish, W. & Lovett, M. An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Res* **14**, 2041-7 (2004).
33. Bedford, D.C., Kasper, L.H., Fukuyama, T. & Brindle, P.K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9-15 (2010).
34. Moellering, R.E. et al. Direct inhibition of the NOTCH transcription factor complex. *Nature* **462**, 182-8 (2009).
35. Gronemeyer, H., Gustafsson, J.A. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3**, 950-64 (2004).
36. Bruning, J.B. et al. Coupling of receptor conformation and ligand orientation determine graded activity. *Nat Chem Biol* **6**, 837-43 (2010).
37. Devlin, H.L. & Mudryj, M. Progression of prostate cancer: multiple pathways to androgen independence. *Cancer Lett* **274**, 177-86 (2009).
38. Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat Rev Drug Discov* **5**, 993-6 (2006).
39. Kerkela, R. et al. Sunitinib-induced cardiotoxicity is mediated by off-target inhibition of AMP-activated protein kinase. *Clin Transl Sci* **2**, 15-25 (2009).
40. Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C. & Mandel, G. REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell* **121**, 645-57 (2005).
41. Fuller, G.N. et al. Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST)/neuron-restrictive silencer factor, which can be functionally countered by REST-VP16. *Mol Cancer Ther* **4**, 343-9 (2005).

42. Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019-24 (1993).
43. Ullrich, A. & Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203-12 (1990).
44. Ptashne, M. & Gann, A. *Genes and Signals*, (Cold Spring Harbor Laboratory Press, 2002).
45. Rodriguez-Martinez, J.A., Peterson-Kaufman, K.J. & Ansari, A.Z. Small-molecule regulators that mimic transcription factors. *Biochim Biophys Acta* **1799**, 768-74 (2010).
46. Dervan, P.B. Molecular recognition of DNA by small molecules. *Bioorg Med Chem* **9**, 2215-35 (2001).
47. Nielsen, P.E. Peptide nucleic acid targeting of double-stranded DNA. *Methods Enzymol* **340**, 329-40 (2001).
48. Cosma, M.P., Tanaka, T. & Nasmyth, K. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **97**, 299-311 (1999).
49. Metivier, R. et al. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751-63 (2003).
50. Minter, A.R., Brennan, B.B. & Mapp, A.K. A small molecule transcriptional activation domain. *J Am Chem Soc* **126**, 10504-5 (2004).
51. Buhrlage, S.J., Brennan, B.B., Minter, A.R. & Mapp, A.K. Stereochemical promiscuity in artificial transcriptional activators. *J Am Chem Soc* **127**, 12456-7 (2005).
52. Casey, R.J., Desaulniers, J.P., Hojfeldt, J.W. & Mapp, A.K. Expanding the repertoire of small molecule transcriptional activation domains. *Bioorg Med Chem* **17**, 1034-43 (2009).
53. Jonker, H.R., Wechselberger, R.W., Boelens, R., Folkers, G.E. & Kaptein, R. Structural properties of the promiscuous VP16 activation domain. *Biochemistry* **44**, 827-39 (2005).
54. 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).

55. Rowe, S.P., Casey, R.J., Brennan, B.B., Buhrlage, S.J. & Mapp, A.K. Transcriptional up-regulation in cells mediated by a small molecule. *J Am Chem Soc* **129**, 10654-5 (2007).
56. Cress, W.D. & Triezenberg, S.J. Critical structural elements of the VP16 transcriptional activation domain. *Science* **251**, 87-90 (1991).
57. Buhrlage, S.J. et al. Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem Biol* **4**, 335-44 (2009).
58. Xu, W., Kasper, L.H., Lerach, S., Jeevan, T. & Brindle, P.K. Individual CREB-target genes dictate usage of distinct cAMP-responsive coactivation mechanisms. *EMBO J* **26**, 2890-903 (2007).
59. Frangioni, J.V., LaRicca, L.M., Cantley, L.C. & Montminy, M.R. Minimal activators that bind to the KIX domain of p300/CBP identified by phage display screening. *Nat Biotechnol* **18**, 1080-5 (2000).
60. Rowe, S.P. & Mapp, A.K. Assessing the permissiveness of transcriptional activator binding sites. *Biopolymers* **89**, 578-81 (2008).
61. 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).
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. Xiao, X., Yu, P., Lim, H.S., Sikder, D. & Kodadek, T. A cell-permeable synthetic transcription factor mimic. *Angew Chem Int Ed Engl* **46**, 2865-8 (2007).
64. 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).
65. Kwon, Y. et al. Small molecule transcription factor mimic. *J Am Chem Soc* **126**, 15940-1 (2004).
66. Jung, D. et al. Wrenchnolol derivative optimized for gene activation in cells. *J Am Chem Soc* **131**, 4774-82 (2009).
67. Johnston, S.A., Salmeron, J.M., Jr. & Dincher, S.S. Interaction of positive and negative regulatory proteins in the galactose regulon of yeast. *Cell* **50**, 143-6 (1987).

68. Han, Y. & Kodadek, T. Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. *J Biol Chem* **275**, 14979-84 (2000).
69. Shangary, S. & Wang, S. Small-molecule inhibitors of the MDM2-p53 protein-protein interaction to reactivate p53 function: a novel approach for cancer therapy. *Annu Rev Pharmacol Toxicol* **49**, 223-41 (2009).
70. Cheng, J.X., Gandolfi, M. & Ptashne, M. Activation of the Gal1 gene of yeast by pairs of 'non-classical' activators. *Curr Biol* **14**, 1675-9 (2004).
71. Wu, Z. et al. Targeting the transcriptional machinery with unique artificial transcriptional activators. *J Am Chem Soc* **125**, 12390-1 (2003).
72. Majmudar, C.Y., Labut, A.E. & Mapp, A.K. Tra1 as a screening target for transcriptional activation domain discovery. *Bioorg Med Chem Lett* **19**, 3733-5 (2009).
73. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nat Rev Drug Discov* **1**, 727-30 (2002).
74. Copeland, R.A., Olhava, E.J. & Scott, M.P. Targeting epigenetic enzymes for drug discovery. *Curr Opin Chem Biol* **14**, 505-10 (2010).
75. Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R. & Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12**, 1247-52 (2007).
76. Ohtsuki, A. et al. Synthesis and properties of PI polyamide-SAHA conjugate. *Tetrahedron Letters* **50**, 7288-7292 (2009).
77. Marmorstein, R. & Berger, S.L. Structure and function of bromodomains in chromatin-regulating complexes. *Gene* **272**, 1-9 (2001).
78. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* (2010).

Chapter 2

Discovery of coactivator ligands*

2.1 Abstract

Efforts to modulate transcription directly with small molecules rely on the availability of ligands, which ideally are drug-like molecules that bind to a protein surface of interest. One of our main objectives is to identify new ligands of transcriptional coactivators; the proteins recruited as facilitators by transcriptional activators to upregulate gene expression. Such ligands can be tethered to DNA, to facilitate recruitment of target coactivators to the tethered location, or they can be used to inhibit recruitment of a coactivator by endogenous transcriptional activators through direct blockage or allosteric perturbation of the activator-coactivator contact surface. Discovery of new ligands was pursued through high-throughput screening. Cell-based assays were explored, because these would ensure that molecules identified would be functional in downstream biological applications. From empirical and theoretical analysis it was determined that effective ligand discovery in cell-based assays would only be facilitated by a gain-of-signal assay. An assay was designed in which inhibitors down-regulate transcription of a miRNA, and consequent loss of RNAi knockdown provides a gain-of-signal for the inhibitors. The assay is functional, but demonstrated poor sensitivity and was judged difficult to optimize to satisfactory performance. A

* The research described in section 2.5 is a collaborative effort. J.W. Hojfeldt, C.Y. Majmudar, S.P. Rowe, W.C. Pomerantz, and A. Lipinski contributed to preparation of proteins and peptides as detailed in section 2.7. P.J. Schultz and D. Sherman obtained extracts from Costa Rica for study presented in Figure 2.16. C.J. Arevang purified and identified structures shown in Figure 2.17. All figures are from experiments done by J.W. Hojfeldt. C.Y. Majmudar contributed to work presented in Figure 2.11, 2.13, 2.16, 2.18, 2.19, 2.20, and 2.29. W.C. Pomerantz and T. Cierpicki assisted with work presented in Figure 2.21 and 2.23.

robust biochemical assay was chosen instead to screen for inhibitors of the interaction between the activator MLL and the coactivator CBP. A high-throughput screen with this assay identified sekikaic acid and other lichen-derived depsides as inhibitors of MLL-CBP. Sekikaic acid was found to bind cooperatively to the KIX domain of CBP in competition with activators binding to two separate sites on KIX. Sekikaic acid is the most potent molecule ligand of the KIX domain identified to date and its binding mode is unique amongst coactivator ligands.

2.2 Introduction to coactivators as targets of high-throughput screening

The protein-protein interaction surface between transcriptional activators and coactivators is an attractive target for small molecule ligand discovery. Ligands of the activator binding surface on coactivators can be utilized as chemical genetic probes¹ and, possibly, therapeutic drugs² that inhibit the native interactions, or they can be incorporated into DNA-localizing bifunctional molecules that can recruit the target coactivator to a DNA-locus and act as an artificial transcription factor.^{3,4} Ligands that are intended for incorporation into bifunctional recruiters do not necessarily need to bind the activator binding surface, but it may be functionally beneficial to recruit through this surface to maintain native orientation of recruited complexes.⁵ It may therefore be advantageous to screen for ligands of coactivators in assays that report on activator-coactivator inhibition to bias for hits against this particular surface.

The majority of protein-protein interactions are traditionally considered undruggable⁶; this classification implies that it is not cost-effective for pharmaceutical companies to screen for drugs against this type of surface, and although review articles report successful pursuits of such ligands and assert progress in the tractability of these targets,⁷⁻¹⁰ they remain state-of-the-art discoveries. Transcriptional coactivators share several challenges general to discovery of protein-protein inhibitors. Unlike enzymes and receptors, the coactivators have no natural small-molecule ligand or substrate that can be used

as starting point for development of new ligands, and the majority of characterized protein-protein interactions are large shallow surfaces with typically 1200-2000 Å² of buried surface area on each protein.¹¹ Small molecules may however not need to bind this entire surface to function as competitive inhibitors, as smaller areas in the interface, known as hot spots, contributes the majority of binding affinity.¹¹⁻¹³

Only a few activator-coactivator interactions have been structurally characterized, posing an additional challenge for ligand discovery. Most numerous of these are structures of various nuclear receptors (NRs) binding an LxxLL motif in the p160 family of coactivators: e.g., GRIP1^{14,15}, NCOA1,¹⁶ TIF2¹⁷. In the NR-p160 interactions a short LxxLL motif-containing peptide binds in a groove in the activator, and it would be difficult to find a small molecule that binds the interaction partner with the non-invaginated surface, i.e. the coactivator in this case. This is, however, opposite of the relationship assumed for most typical amphipathic transcriptional activation domains; since the function of these are comprised within a short peptide,¹⁸⁻²⁰ they likely bind grooves on their coactivator targets. Indeed, the remaining activator-coactivator interactions that are structurally characterized show this relationship (Figure 2.1), including solution NMR structures of short α -helical transcriptional activation domains from several different activators (and p160 coactivators) bound to domains of one coactivator, the CREB binding protein (CBP),²¹⁻²⁶ as well as a crystal structure of the STAT6 transcriptional activation domain (TAD) bound to a p160 coactivator (NCOA1).²⁷ These structures reveal that the buried surface between activator and coactivator are smaller (e.g. 600 Å²)²¹ than typical protein-protein interfaces. Such small shallow binding surfaces are predicted to involve weak and transient binding interactions,¹¹ which is also in agreement with affinities measured by us and others for minimal transcriptional activation domains binding domains of CBP.²⁸ The low binding affinities measured in vitro may not be accurate estimates of native interactions that involve full-length interacting proteins within a context of

additional protein complexes and DNA, and this must be considered as a possible challenge for discovered ligands to function as inhibitors in cells.

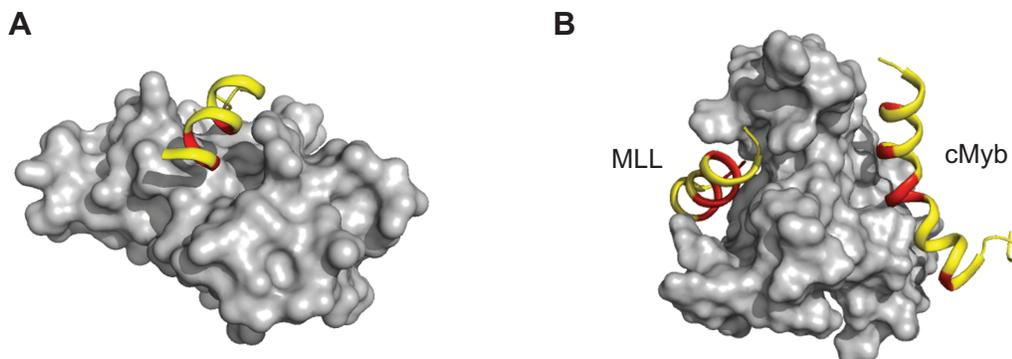


Figure 2.1 Structures of activator-coactivator complexes Only a few structures of activator-coactivator complexes are available. Two representative structures of the amphipathic class of transcriptional activation domains binding to a target coactivator are shown: A. Crystal structure of STAT6 activation domain (helix) binding a groove on PAS-B domain of NCOA1 (PDB:1OJ5)²⁷; B. The KIX domain of the coactivator CBP has two activator binding spots. Solution NMR structure of MLL and cMyb activation domains (helices) binding to their respective grooves (PDB:2AGH)²⁹. Hydrophobic residues in the helices that are critical for binding affinity are colored red.

A handful of small molecule ligands of coactivators have been identified in the past decade. From our lab, the isoxazolidine-based transcriptional activation domain mimics (iTADs) have emerged as the only designed ligands,³⁰⁻³² while other ligands have been discovered in high-throughput screens.³³⁻³⁸ The iTADs are thought to bind several coactivators at binding surfaces shared with natural TADs, and indeed they have been cross-linked to several intracellular targets. The coactivator CBP is the only identified target, and the affinity of one iTAD for this surface is in the low millimolar range (W. C. Pomerantz, unpublished results). The low affinity and promiscuous binding profile makes the iTADs ambiguous probes of individual activator-coactivator interactions. In experiments performed in our lab, in which I have participated, the ability of iTADs and iTAD-derivatives to inhibit natural TADs have been evaluated, but direct effect on targets can only be inferred.^{39,40} As a transcriptional activation domain, the activity of iTADs has only been demonstrated in cells when tethered to the ligand binding domain of

the glucocorticoid receptor.^{41,42} The surface of the receptor may contribute to this activity, which is further evidenced by the inability of iTAD to activate transcription when tethered via a GR antagonist (J. W. Hojfeldt, A. R. Van Dyke, Y. Imaeda, unpublished results). We have limited experience with ligands discovered in other labs, but few reports of studies with these ligands, subsequent to their initial discovery, exist. Chetomin, which was discovered as inhibitor of the interaction between hypoxia-inducible factor 1 α (Hif1 α) and the coactivator CBP,³⁴ may act through zinc ejection rather than binding the coactivator⁴³ and chetomin additionally has other notable intracellular interactions.⁴⁴ Thus, available ligands do not currently represent a satisfactory arsenal in terms of numbers, potency and specificity, and we have therefore explored high-throughput screening options to discover new classes of coactivator ligands.

2.3 Cell-based assays for transcriptional inhibitors

Design of a luciferase-based reporter assay

Cell-based assays, when they can be designed with satisfactory sensitivity and precision for a high-throughput screening format, have the advantage of identifying only compounds that are cell permeable and that function within the complex environment of the cell. Cell-based assays on the other hand face additional challenges; since many intracellular targets may influence the assay, it is important to design assays with minimal response to off-target effects. Transcriptional activators are the downstream effectors of signaling cascades, which not only include multiple proteins, but the majority of these signaling components have enzymatic function and are considerably more druggable.⁶ Signaling cascades must therefore be uncoupled from the transcriptional output of the chosen assay, so that off-target hits do not predominate screening hits (Figure 2.2).

We decided to use a minimal transcriptional activation domain, defined as a polypeptide sequence from an endogenous activator that comprises the majority of its activation function, and fused it a heterologous DNA-binding domain (DBD)

taken from the yeast activator, Gal4 (residues 1-147). By only including a small portion of the native mammalian activator, many signal responsive domains are excluded. This fusion activator is introduced into a human cell line together with a luciferase reporter plasmid with five Gal4-binding sites in the promoter. In the assay, inhibition of transcription by the fusion activator would be seen as a drop in luciferase signal. While many potential off-target proteins are eliminated by the use of a signal-independent activator, inhibitors of the general transcription or translation machinery as well as cytotoxic compounds, would also produce a drop in signal (Figure 2.2). We reasoned that a comparison of hits to data from other screening campaigns would facilitate filtering of cytotoxic compounds. In fact, many screening centers have tested their libraries in cell viability assays to help filter such compounds. An alternative option is to simultaneously use an orthogonal assay for cell viability.

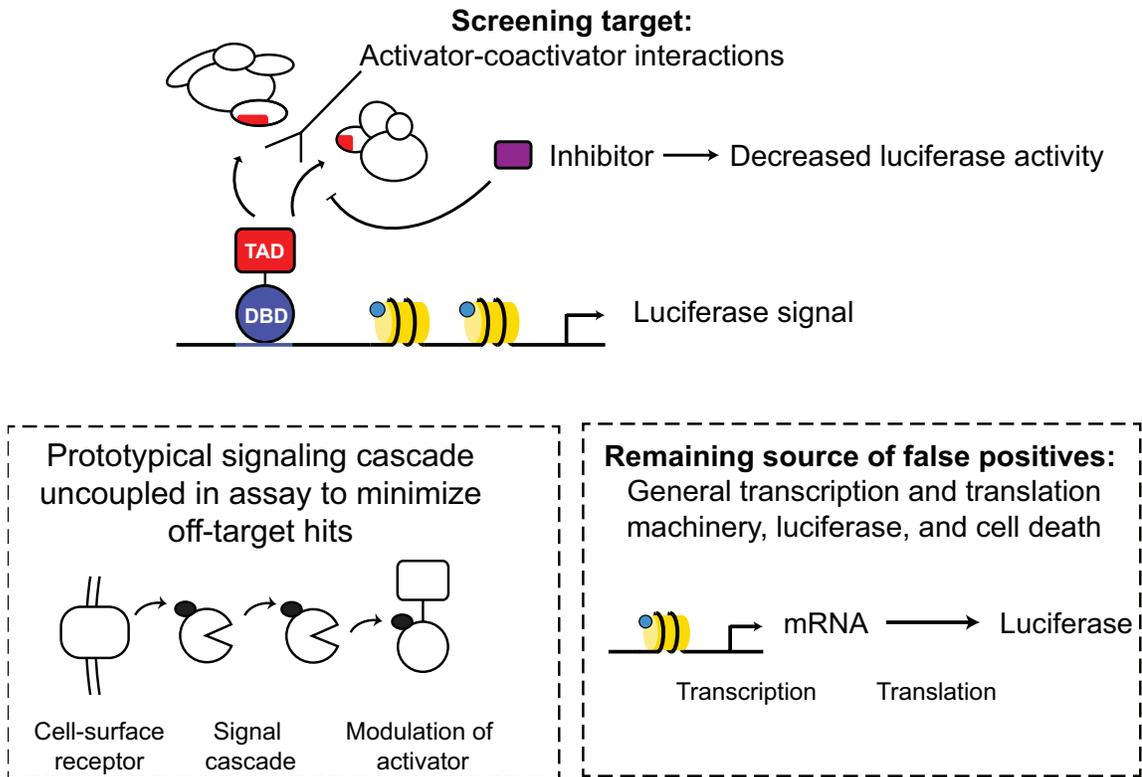


Figure 2.2 Schematic of a cell-based luciferase reporter inhibition screen A minimal transcriptional activation domain, TAD, is fused to heterologous DNA-binding domain (DBD). Inhibitors of the activator-coactivator interaction, which is

necessary for transcription of the luciferase reporter gene, will cause a decrease in luciferase activity. Use of a minimal TAD and a heterologous DBD largely uncouples the fusion protein from regulation by cellular signaling.

The activation domain (residues 413-490) from the herpes simplex virion protein 16 (VP16) is one of the strongest known activation domains in mammalian cells.^{45,46} The strength of VP16 is beneficial in a HTS, because it can produce a strong signal. The direct coactivator binding partners of VP16 and other activators have been difficult to affirm^{47,48}, although many have been proposed.⁴⁹ Identification of ligands that inhibit VP16 would help elucidate these coactivator targets, and discovery of ligands to coactivators that mediate such strong activation potential could lead to structures that can recruit these same proteins and make very potent artificial transcriptional activators when incorporated into DNA-targeting bifunctional conjugates. The VP16 activation domain has advantages in a HTS assay over two other activation domains that were considered: the activation domains from the human activators RelA (NF- κ B p65) and Elf3, both of which are studied in our lab.

RelA is a well-studied transcriptional activator with several proposed coactivator targets that could facilitate follow-up studies,⁵⁰⁻⁵² and it is a therapeutically relevant target for inhibitors.⁵³ The C-terminal activation domain, residues 508-550,^{54,55} however, is the direct target of NF- κ B signaling events.^{56,57} It is therefore possible that a Gal4 fusion protein would still respond to regulation by endogenous kinases, and indeed a GST-RelA(TAD) fusion protein is phosphorylated *in vitro* by the inhibitor of kappaB kinase (IKK).⁵⁶ The Elf3 activation domain⁵⁸ was considered because two small molecules already have been discovered that inhibit the transcription of the breast-cancer associated ErbB2 oncogene^{33,39}, whose transcription is proposed to be driven by Elf3.⁵⁹⁻⁶¹ If these small molecules could inhibit a Gal4-Elf3 activation assay, they would be useful controls in a screen, and screening hits may add to the list of ErbB2 transcription inhibitors. An additional attraction to Elf3 was the discovery of

Med23 as an essential coactivator target⁶², which would present an immediate candidate target protein in secondary assays.

Preparation of assay for a high-throughput pilot screen

The primary advantage of VP16 over the two mammalian activators was its potency. The future goal with the assay was to make a stable cell line with an integrated reporter and fusion-activator expression construct. We therefore tested the activators in a commercially available HeLa derived cell line, HLR (Stratagene); HLR cells have an integrated reporter with 5xGal4 binding sites and a minimal TATA-box containing promoter upstream of the *Photinus pyralis* (American firefly) luciferase gene. The number of integration sites and location are not characterized. Gal4(DBD)-RelA(522-551) and Gal4(DBD)-Elf3(129-159) activated transcription 8- and 20-fold respectively, measured as luciferase activity relative to activity in untransfected HLR cells. In comparison Gal4-VP16(413-490) activated transcription 70-fold above background (Figure 2.3). This difference is significant since the assay format was not yet reduced to a 384-well format to be used in the high-throughput screen, and generation of stable cell lines would most likely result in lower expression than seen with the transiently transfected cells. An early attempt to make stable cell lines with Gal4-Elf3 and Gal4-RelA lead to only two clones of Gal4-RelA expressing cells. Because the activity in these clones was only 3-4 fold above untransfected HLR cells (Figure 2.3), this

effort was not pursued further.

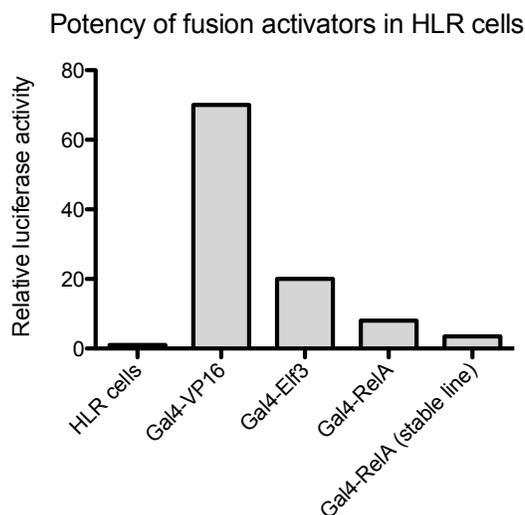


Figure 2.3 Activity of three fusion activators Expression plasmids for three fusion activators was transfected into a cell line with an integrated luciferase reporter gene and the activity determined.

The two reported inhibitors of ErbB2-gene transcription (N2-biphenyl iTAD³⁹ and adamanolol³³) were synthesized and tested for their ability to inhibit the activity of Gal4-Elf3, but showed no inhibition in this assay. Any reduction in relative luciferase expression was a consequence of the compounds being cytotoxic as determined from the activity of a coexpressed *Renilla reniformis* luciferase gene and a simultaneous WST-1 viability test (Figure 2.4). Complicating this analysis is the fact that they may be cytotoxic through inhibition of various activator-coactivator interactions. A small library (~ 50 compounds) of small molecules made by various members of the Mapp lab, which included isoxazolidines and spirooxindoles functionalized with both hydrophobic and hydrophilic groups to mimic the amphipathic nature of natural TADs, was tested for the ability to inhibit fusion activators of Gal4 with Elf3/RelA/VP16 as well as 5 other natural TADs. However, none of these compounds functioned as inhibitors for any of the activators. The consequence of these tests was that no real positive control was available for the screen. Instead, the translation inhibitor, cycloheximide, was used as a surrogate control.

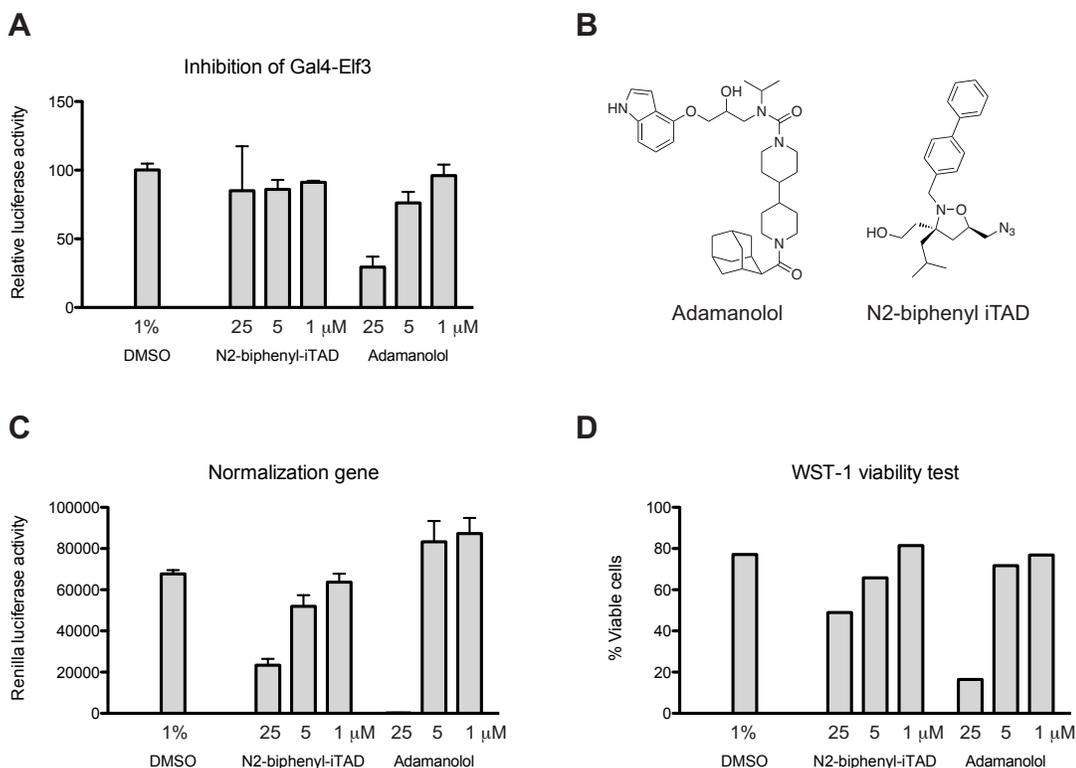


Figure 2.4 Test for inhibition of Gal4-Elf3 Adamanolol and N2-biphenyl iTAD were added to HeLa cells transfected with a firefly luciferase reporter, a Gal4-Elf3 activator and a *Renilla reniformis* luciferase expression plasmid. A; The activity of the Gal4-Elf3 induced firefly luciferase normalized to *Renilla reniformis* luciferase was determined. The activity is reported relative to the DMSO control; B. The two ErbB2 transcription inhibitors; C. The *Renilla reniformis* luciferase activity is affected by the small molecules; D. The drop in normalization gene is likely due to cell death.

Because VP16 is the most potent activator with no known signal-responsiveness, Gal4-VP16 was chosen for a pilot screen. HLR cell lines were transiently transfected with a Gal4-VP16 expression plasmid. In the 384-well format with the translation inhibitor, cycloheximide, as a surrogate control, an 8-fold dynamic range was obtained with a Z'-factor of 0.55. The Z-factor (calculated from screen data) and Z'-factor (controls only) are measures of the variability of samples in an assay relative to the dynamic range of the assay, and a Z-factor > 0.5 is considered an excellent assay for HTS.⁶³

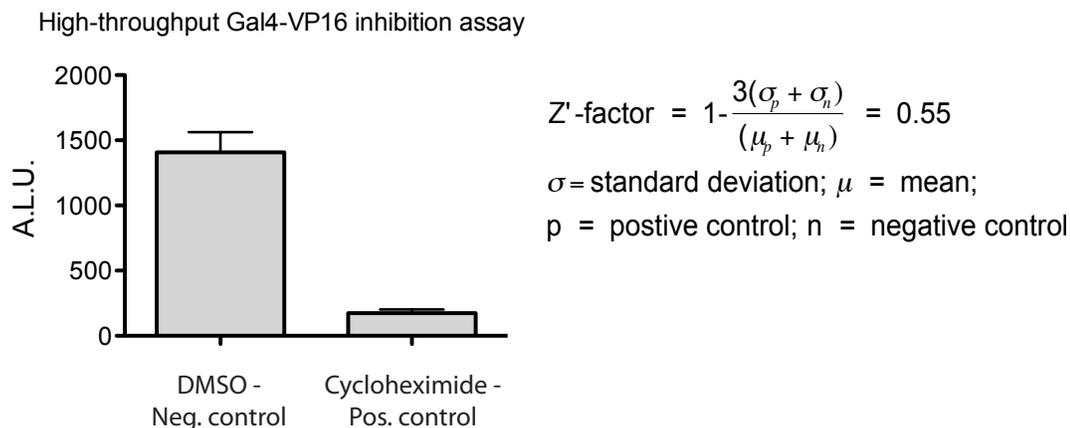
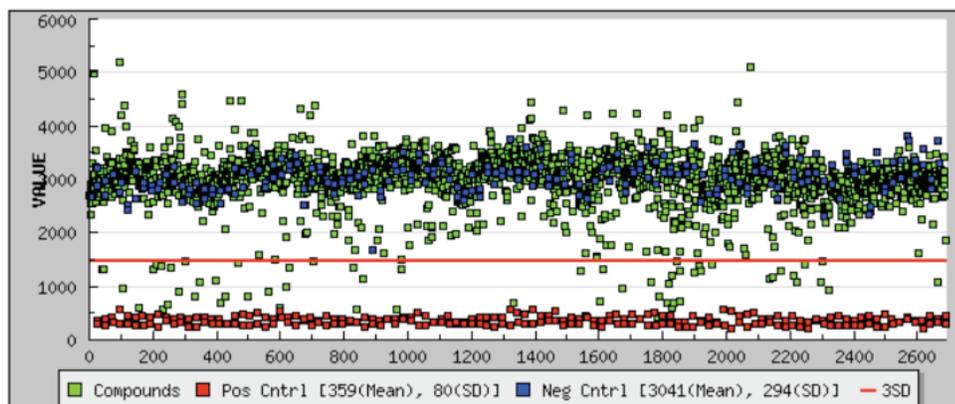
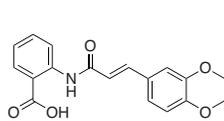


Figure 2.5 Performance of a high-throughput Gal4-VP16 inhibition assay

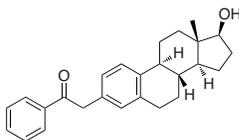
Bar graph shows average luciferase activity measured for 320 samples treated with 0.5% DMSO (negative control) and average activity for 64 samples treated with 10 μ g/mL cycloheximide (mock positive control). Error bars represent one standard deviation. Z'-factor is calculated from the mean and standard deviation of luciferase activities of positive and negative controls.

Evaluation of the assay in a pilot screen

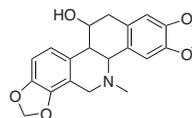
A library of 2000 compounds (Spectrum library, MicroSource), which includes 50% approved drugs, 30 % diverse collection of natural products and 20 % other bioactive small molecules, was screened at 7.5 μ M to characterize assay performance. This screen had a Z-factor of 0.58, with 71 initial hits based on a 2.5 standard deviation cut-off. The activity of the hits was compared to the activity against two other mammalian cell-based loss-of-function assays used in the same screening facility. One of these assays also uses a luciferase reporter gene, and the other assay uses a luciferase-based viability reporter. 14 compounds remained as hits after filtering, with 9 of the hits inhibiting more than 50% and more than 3 standard deviations. Of the 9 top hits, 5 are steroids. To more closely monitor the effects of some of the hit compounds on the cells and the assay output, three of the top 9 hits were purchased (including two estradiol derivatives) as well as one cytotoxic hit that was filtered out (Figure 2.6B).

A $Z\text{-factor} = 0.58$ **B**

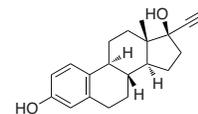
Tranilast



Estradiol benzoate



Chelidonium



Ethinyl estradiol

Figure 2.6 Pilot screen of 2000 compounds A. Luciferase activity measured for 2000 compounds; B. Structures of 4 hits tested in follow-up studies. The four compounds are all false-positives in the screen.

The dose-response of these four compounds was tested in the assay. As previously determined the WST-1 viability reagent can be used prior to addition of the luciferase reagents without interfering with the later, and this was included to simultaneously monitor if the compounds were cytotoxic. The filtered hit compound, chelidonium, was indeed very toxic, while the other non-steroid hit compound tested, tranilast, was neither toxic nor inhibited the assay. Thus, tranilast is a false positive that would likely have been eliminated in a larger screen during confirmatory tests of initial hits. The two estradiol derivatives showed reduction of luciferase activity, but this was associated with cell death.

This pilot screen demonstrated the difficulty of accurately eliminating cytotoxic compounds in a loss-of function cell-based assay. It was not surprising that the library contained many cytotoxic compounds (especially since these are tested at

a single dose), but it is problematic that these cannot be easily eliminated by comparison to other cell based assays. Further complicating this is the possibility that a true positive hit is also cytotoxic and should not be eliminated, and since many cell-based assays are reporter gene-based assays, non-selective inhibition relative to these assays may also eliminate true positives. It appears that the best option for this assay format would be to use an internal control for cell viability. Cotransfection of a second orthogonal luciferase reporter is commonly used in our lab to control for variation in transfection efficiency and cell numbers, and could in a high-throughput assay report on viability or non-selective inhibition of transcription against the second luciferase construct. However, my experience with dual-luciferase assays that have been used to test a small number of potential inhibitors is, that the activities of the two luciferases aren't always closely matched to each other in response to a significant drop in viability. This can make samples appear as either inhibitors or enhancers of transcription.

Evaluating a novel gain-of-signal transcription assay

Cytotoxicity poses a problem in the cell-based loss-of-signal assay, because it can't be distinguished from inhibition. In addition, the inhibitors that we seek may indeed be cytotoxic as a consequence of their specific function. An assay, in which the inhibition event leads to a rise in signal, would solve the issue of cytotoxic false-positives and could still detect cytotoxic true-positives. A novel gain-of-function assay was designed (Figure 2.7): a luciferase gene is expressed from a constitutively active promoter; this signal is however suppressed by a microRNA transcribed by the transcription factor of interest.

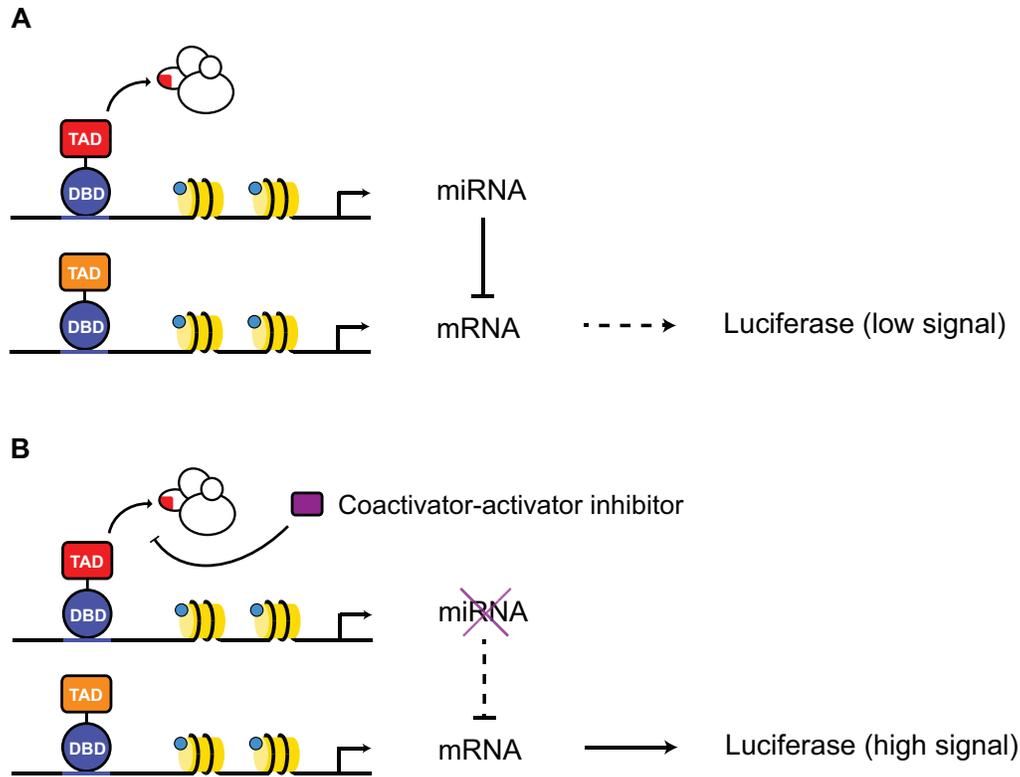


Figure 2.7 A gain-of-signal assay for activator-coactivator inhibitors A. The transcribed luciferase gene is knocked down by a microRNA, whose expression is controlled by the activator of interest (red TAD); B. An inhibitor blocks transcription of the miRNA but not the luciferase gene, and the luciferase activity increases.

MicroRNAs are endogenously expressed small single stranded RNA sequences that mediate gene silencing through components shared with the RNAi pathway⁶⁴⁻⁶⁶ In contrast to short hairpin RNA (shRNA), which must be transcribed from a plasmid by RNA polymerase III, endogenous miRNA are processed from Pol2 transcripts.⁶⁷ Recently, Pol2-driven vectors were developed to express engineered miRNAs.⁶⁸ In these expression vectors, an engineered miRNA sequence (designed against target transcript)⁶⁹ is cloned in between flanking regions based on murine miR-155⁷⁰, which facilitate processing of mature miRNA. The engineered miRNA and miR-155-based flanking regions are located in a gene intron, and the exons can simultaneously code any gene of interest. This allows fluorescent proteins to be simultaneously expressed in the same transcript as the miRNA to serve as a control for the transcription of this gene.

A miRNA was designed to target firefly luciferase (miR-luc), and when this was expressed from a highly active CMV promoter in cells cotransfected with a luciferase expressing plasmid, the luciferase signal was reduced to 9% (Figure 2.8).

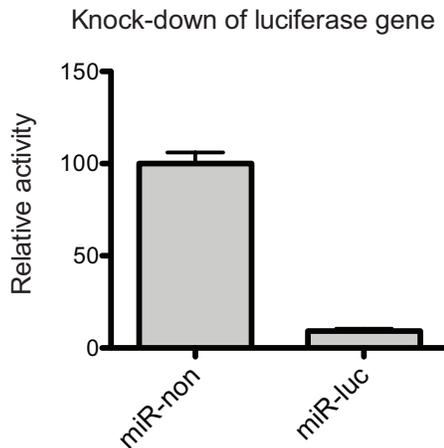


Figure 2.8 Knock-down of firefly luciferase gene by miR-luc Luciferase activity in cells transfected with luciferase expression plasmid and a plasmid encoding either a miRNA against luciferase (miR-luc) or a non-targeting miRNA (miR-non).

The CMV enhancer of the miRNA expression vector was replaced with a tetracycline responsive enhancer (TRE).^{71,72} The TRE has high activity when bound by tetracycline-responsive transcriptional activator, tTA, which contains a VP16-derived transcriptional activation domain. The promoter becomes nearly silent when tetracycline (or tetracyclin analog, doxycyclin) binds tTA and disrupts its ability to bind DNA. Thus, miRNA expression in this promoter can be blocked by the addition of doxycyclin. To study the transcriptional dose-response of the TRE to doxycyclin, a TRE-regulated luciferase reporter was used (Figure 2.9A). Addition of doxycyclin reduced the luciferase activity 100-fold in a dose-responsive fashion with an IC₅₀ of 2.5 ng/mL and over 92% inhibition at 10 ng/mL. When the tetracycline-regulated miR-luc expression vector was used to knockdown the signal from a cotransfected luciferase (with a constitutive SV40 enhancer), a doxycyclin dose-response showed a 3-fold gain in signal with an EC₅₀ of 33 ng/mL (Figure 2.9B). The level of luciferase activity at high concentration of doxycycline was equivalent to the activity when pTRE-miR-non

was used (TRE-regulated expression of non-targeting miRNA), and this signal was independent of doxycyclin concentration.

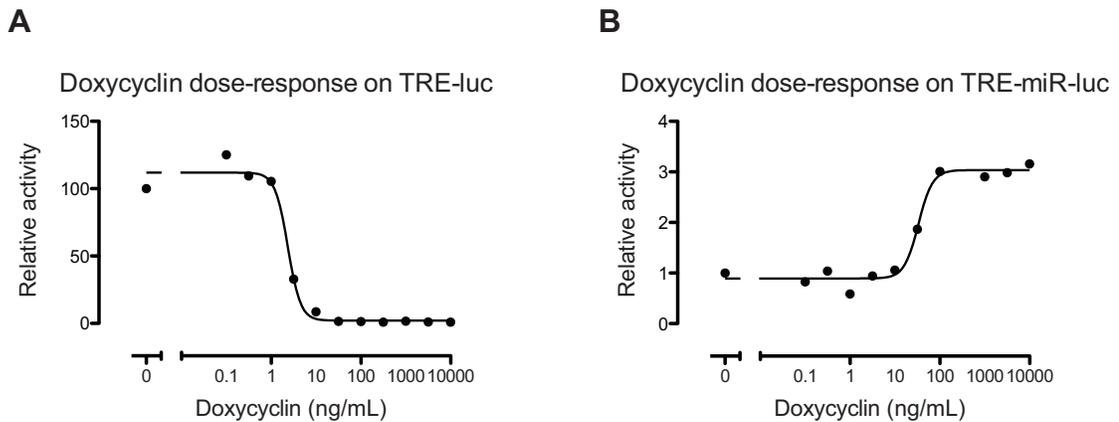


Figure 2.9 Dose response of transcriptional inhibition and gain-of-signal A. Luciferase activity in cells transfected with luciferase expression plasmid and a plasmid encoding either a miRNA against luciferase (miR-luc) or a non-targeting miRNA; B. Transcriptional dose-response of doxycyclin determined in cells cotransfected with a doxycyclin-responsive activator tTA (inhibited by doxycyclin) and a luciferase reporter with binding sites for tTA; C. The gain of luciferase activity in response to different doses of doxycyclin, when miR-luc is expressed from a tTA regulated gene.

The >10-fold dynamic range demonstrated by the CMV-driven miRNA suppression is likely sufficient for a HTS assay. It is also in agreement with typical RNAi results, which provide knockdowns from 90-98%. The 3-fold induction of signal obtained with doxycyclin regulated inhibition of miRNA production is too low, but can likely be improved. In these assays, three plasmids are introduced into cells in a co-transfection, and if a small number of cells receive the luciferase expression plasmid but little or none of either pTRE-miR-luc or tTA expression plasmid, a high baseline is obtained from this population. The signal-to-background ratio of the gain-of-function experiment was 1000, and thus the signal can be greatly reduced in attempts to optimize the dynamic range of knockdown. Ideally stable cell lines should be made, which must receive three plasmids: expression of tTA activator, luciferase and miR-luc. If the assay is modified to report an endogenously expressed transcriptional activator, only the luciferase and a miR-luc plasmids would need to be introduced, where the

enhancer of the miR-luc plasmid would be changed to include binding sites for the activator of interest. The more concerning result from the dose-response study is the shift in IC₅₀/EC₅₀ (the dose at which half of the maximal response is seen). At 10 ng/mL doxycyclin, transcription from the TRE promoter is reduced to 10% (based on luciferase activity), while there is no rise in signal at this concentration in the gain-of-function assay. The EC₅₀ for doxycyclin is raised >10-fold in the gain-of-function assay. Thus, in a screen, only compounds that inhibit transcription by at least 90% would be seen as hits. This is particularly concerning because full inhibition of a single activator-coactivator interaction may only partially reduce the activity of the activator.⁷³

Optimization of the assay was judged too difficult to pursue further. The levels of the expressed miRNA likely need to be controlled to a level of strong knockdown, yet still limiting to show a response to low levels of inhibition. Such balance is difficult to obtain with transient transfections and with the low dynamic range found initially. Creation of numerous stable cell lines is likely necessary to achieve such balance, but may be practically unfeasible in the system with three separate plasmids.

2.4 Conclusion and discussion of efforts with cell-based HTS assays

For identification of inhibitors of individual transcriptional activation domains in cells, a loss-of-function assay will yield unsatisfactorily high false positive hit rates. If these are eliminated by an internal viability control or comparison to screening data, a significant fraction of true hits may be lost. These problems are, at least in theory, overcome by a gain-of-signal assay. A gain-of-signal assay for inhibitors can be constructed by making the target component suppress a constitutive signal, which was attempted for an activator-coactivator interaction via an RNAi knockdown mechanism. The gain-of-signal mechanism is functional, but the added complexity makes the assay difficult to optimize. Most cell-based HTS assays use a signal produced from a luciferase or fluorescent protein, but a viability assay could also be used and would simplify the assay since no added

reporter gene is needed. The assayed component could in this case induce a toxic gene, and inhibitors would increase the cell number. Regardless of the chosen signal or suppression mechanism, it will be critical to adjust the level of suppression to a starting point that is sensitive to inhibitors. The dose response of the suppression mechanisms will additionally determine if the assay output is a graded response or an active/inactive type of response.

We have additionally realized that data for inhibitors of transcriptional activation domains may already exist in screening databases. Many assays use transcription as an output coupled to the intended target function. Transcription inhibitors would likely be eliminated from such screens in secondary assays, and labeled as false positives in those assays. Thus data mining of publicly available HTS results may be a fruitful effort in place of or in combination with a primary screen for inhibitors.

2.5 Identification of a lichen natural product as ligand of the coactivator CBP from a biochemical high-throughput screen

Introduction to the multi-domain CBP/p300 coactivators

The cell-based assays explored in the previous sections were designed to discover inhibitors of a specific transcription factor, but with no specific coactivator as target. This was done because the direct binding targets of most activators are unknown. There are, however, coactivator proteins that are known to be direct targets of some activators, and these activator-coactivator interactions can be assayed for inhibitors in biochemical assays. The highly homologous human transcriptional coactivators CREB binding protein (CBP or HGNC: CREBPB) and E1A binding protein p300 (p300 or HGNC: EP300) are considered to be transcriptional hubs.⁷⁴ They consist of 9 conserved domains: a histone acetyltransferase domain gives the protein intrinsic chromatin modifying ability, while the remaining domains are protein interaction domains that are proposed to interact with several hundred transcription factors and coactivators (Figure 2.10).^{74,75} The role of CBP/p300 in animal physiology has been difficult to

study using standard genetic approaches, as knockouts of either protein cause embryonic lethality.^{76,77} Cells lacking both proteins can be generated but do not divide.⁷⁸ These observations are expected as a consequence of the proteins' ubiquitous role in transcription of numerous genes in all cell types.

Disruption of individual interaction domains of CBP or p300, in contrast, show more restricted phenotypes, indicating that they are each utilized by only a subset of transcriptional activators. Knock-in mice have been generated in which CBP or p300 is replaced by a protein with a mutation that disrupts an activator interaction surface in the KIX or the CH1 domain.^{79,80} Mice with mutations in p300's KIX domain that disrupt binding to transcriptional activators cMyb and CREB, are viable but show mild defects in hematopoiesis⁷⁹ and impairment of long-term memory.⁸¹ Mutations in the CH1 domain of p300 and CBP block activation of hypoxia-responsive genes by Hif1alpha.⁸⁰ The restricted use of individual activator interaction sites within p300/CBP is encouraging from a perspective of probing activator function with small molecule inhibitors of these domains. It also increases the likelihood that ligands to these domains can have therapeutic value, since a limited set of genes will be affected by their disruption.

Complicated genetic strategies were needed to demonstrate the domain specific interactions of p300/CBP. The two proteins can often complement each other, and to study the role of the KIX domain of either p300 or CBP in CREB mediated transcription, mice had to be generated with three mutant alleles and one loxP-flanked allele, which is eliminated in isolated cells by transfection of Cre recombinase.⁷³ The influence of these CBP/p300 KIX-mutants can only be studied through this genetic strategy in the generated mouse cell lines, and the employment of such strategy in any human cell line would be impossible. Thus, to complement genetic strategies, small molecules that bind to domains of CBP/p300 and block interactions with transcription factors would have great value to help elucidate the complex network that these proteins function within. Indeed, several groups have reported discoveries of ligands/inhibitors of

CBP/p300 domains (Figure 2.10), including the nuclear receptor interaction domain (NRID) (ICG-001: inhibitor of β -catenin-CBP)³⁶, TAZ1 domain (chetomin: inhibitor of Hif1 α -CBP, not shown in figure)³⁴, KIX domain (Naphthol AS-E phosphate: inhibitor of CREB-CBP)³⁵, HAT domain (C646, histone acetyltransferase inhibitor).⁸² In addition, our lab determined that iTAD1 interacts with the MLL binding site in the KIX domain of CBP.⁴²

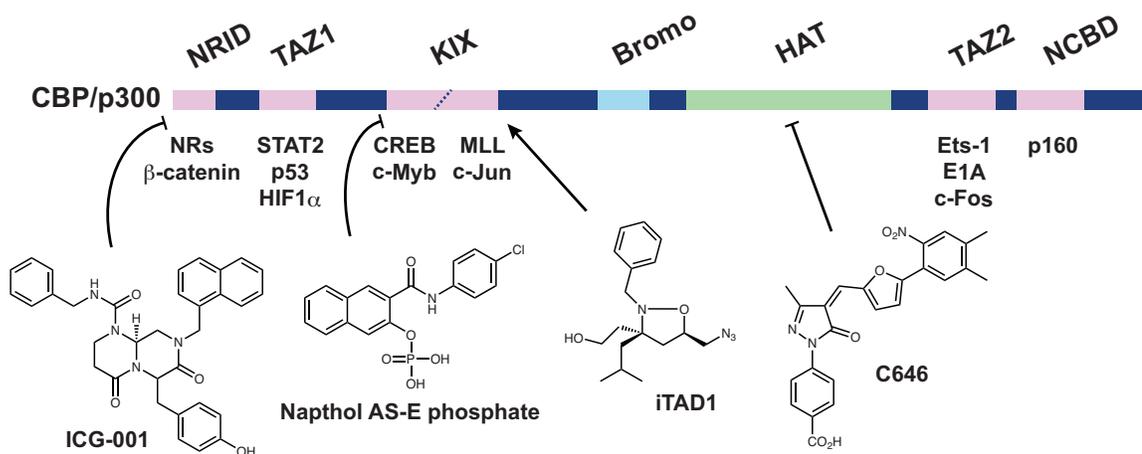


Figure 2.10 The CBP/p300 coactivator and its ligands The CBP/p300 coactivator contains multiple interaction domains. Domain labels are above the schematic representation of the protein (NRID: nuclear receptor interaction domain; HAT: histone acetyltransferase domain; NCBD: NR-coactivator interaction domain) Domains that interact with numerous transcriptional activators and coactivators are shown in pink, and representative examples are listed below (NRs: nuclear receptors; p160: p160 class of NR coactivators). Structures of four molecules that bind to individual domains are shown (references in text).

The KIX domain of CBP has two interaction surfaces.⁸³ Several activators bind to a surface shared with CREB, and other activators bind a surface shared with MLL (Figure 2.1B).^{21,84} Two activation domains can bind simultaneously to their respective target sites in a cooperative fashion.^{28,83} The p53 transcriptional activator binds to both sites of KIX with two separate activation domains,⁸⁵ and the FOXO3A TAD can bind both sites with similar affinity.⁸⁶ Of the two sites, the MLL site presents a deeper cleft, which likely makes it easier to target with small molecules. The binding affinity of iTAD for this site is weak (1-4 mM, W.C.

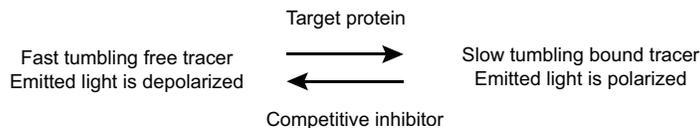
Pomerantz, unpublished data) making it unlikely to function as an inhibitor, and since no other ligands of the site is known, we wished to screen for inhibitors of the KIX-MLL interaction.

A fluorescence polarization assay of the MLL-KIX interaction

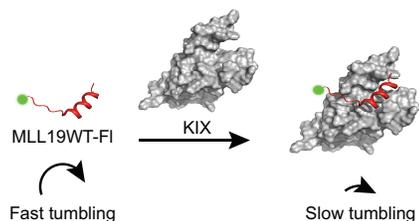
To screen for ligands that interact with the MLL binding site in the KIX domain of CBP, we used a fluorescence polarization-based assay^{87,88} with bacterially expressed KIX-domain and a fluorescein-labeled version of the MLL transcriptional activation domain. The principle of the fluorescence polarization assay is illustrated in Figure 2.11. The assay is attractive in a high-throughput format because the readout is a ratio of emitted light measured through two perpendicular filters and this ratio is stable to variations in concentrations, total volumes, and compounds that absorb light, thus producing a signal with robust tolerance to interference. In a solution NMR structure of the KIX-MLL complex, MLL makes contact to KIX via a short peptide (residues 2840-2858)²⁸. Eleven amino acids of the 19-mer peptide fold as an alpha helix when bound to KIX and the majority of residues in the helical portion in addition to 3 N-terminal residues are critical for the binding interaction.⁸³ The 19-mer MLL peptide (2840-2858) was synthesized with fluorescein attached to the N-terminus (FI-MLL19). A direct binding experiment was used to determine that FI-MLL19 binds to KIX with a dissociation constant (Kd) of 1.2 μM (Figure 2.11B), which is in agreement with published values.^{28,89} A 15-mer unlabeled peptide, corresponding to the 14 critical residues in MLL (2844-2857) and a C-terminal tyrosine for accurate determination of concentration, was used in a competition assay to inhibit FI-MLL19 (Figure 2.11C). MLL15 peptide inhibits with an IC₅₀ of 10.9 μM , and the inhibition curve can be fitted to a theoretically derived equation⁸⁸ from which a K_i of 1.2 μM is determined (the fitted values of anisotropy of free and bound tracer closely match the anisotropies from the maximum and minimum values in the direct binding experiment, and the goodness of fit is: $R^2 = 0.9841$). Thus, with similar binding affinity of FI-MLL19 and MLL15, it appears that fluorescein does

not contribute to the binding affinity, and false-positives in the HTS will not be found due to inhibition of fluorescein binding.

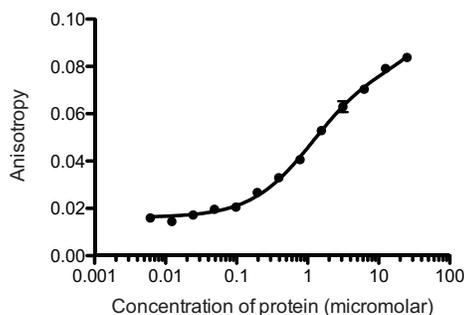
A FP assay principle



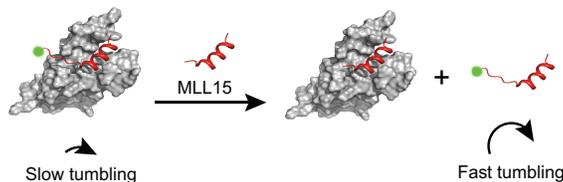
B



Binding curve for protein (KIX) to tracer (FI-MLL19)



C



Competitive inhibition of FI-MLL19 - KIX

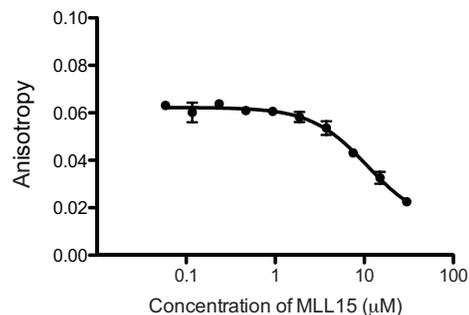


Figure 2.11 Fluorescence polarization assay of the MLL-KIX interaction A. In the FP assay, the MLL peptide is labeled with a fluorophore. When the fluorophore is excited with polarized light, the light emitted will be polarized to a degree that depends on the rotational frequency of the molecule in solution. When the molecule is bound to a large protein, rotation (or tumbling) is slowed down significantly and light remains polarized; B. Dose-response curve for KIX binding to FI-MLL19. With increasing concentration of KIX, more tracer is tumbling slowly and the emitted light is polarized (measured in anisotropy units); C. Competitive inhibition with MLL15.

In the inhibition assay, the concentration of KIX (5 - 10 μM) is chosen so that 80 - 90% of ligand is bound. This ensures that the assay is sensitive to detect inhibition while maintaining a large dynamic range. An equation that defines the

fraction of bound tracer in presence of inhibitors as a function of total input concentrations and binding affinities of assay components is derived in a theoretical treatment of FP assays.⁸⁸ A curve was generated from equation 17 in this paper to show the theoretical fraction of bound FI-MLL19 as a function of the binding constant, K_i , of a competitive inhibitor present at 22 μM , which will be the concentration of small molecules in the high-throughput screen (Figure 2.12). Notice that the starting fraction (at 5 μM KIX) is 80% bound tracer. From this curve the detection threshold of the assay is determined, when it is assumed that 15% change in anisotropy (equivalent to a change in bound tracer from 80% to 68%) can be detected: the assay is sensitive to detect inhibitors with a binding constant of 18 μM .

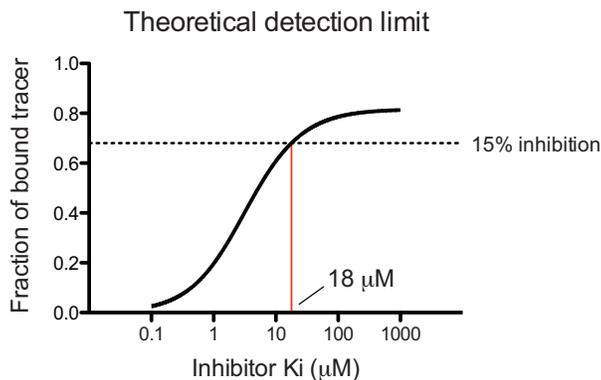


Figure 2.12 Theoretical detection limit of KIX-MLL FP assay In the HTS, compounds will be tested at 22 μM . A competitive inhibitor with a $K_i < 18 \mu\text{M}$ will be detected in the assay.

The FP-assay was tested in a 384-well, 10 μL format with DMSO as a negative control and no added protein as a surrogate positive control. With exclusion of KIX, the tracer is unbound and gives same low anisotropy value as complete inhibition by MLL15. The no KIX control makes assay setup easier, saves on peptide reagent, and defines the dynamic range of the assay. The assay has a Z'-factor of 0.71 (Z-factor > 0.5 is an excellent HTS assay)⁶³ (Figure 2.13).

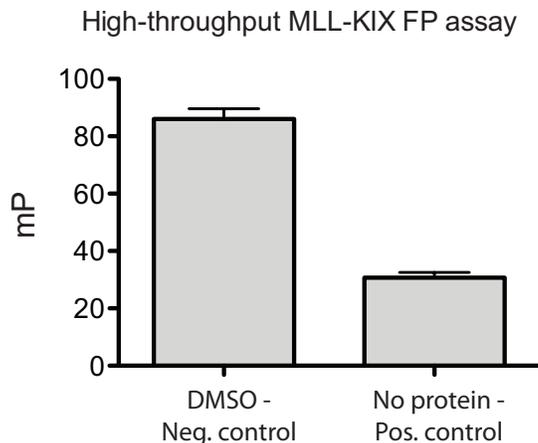


Figure 2.13 Variability of KIX-MLL FP assay The mean mP values for positive and negative controls in a 384-well test plate.

A high-throughput screen with the KIX-MLL FP assay

66882 samples were screened at the University of Michigan Center for Chemical Genomics, which includes 50562 commercially available compounds with known identity, and 16320 natural product extracts with high biodiversity deposited by the laboratory of Dr. Sherman at our institution. The Z-factor during the screen, which was done in multiple sessions ranged from 0.69 to 0.88 (Figure 2.14A). Threshold values for initial hits and the number of hits that met the criteria are shown in Figure 2.14B. Fluorescent molecules interfere with the assay, and samples with a total fluorescence signal higher than 3 standard deviations from the mean of the total fluorescence of all samples were eliminated from hits after the primary screen. The filtered hits were retested in triplicate to confirm their activity. Of the confirmed hits, only the commercial compounds were tested for a dose-response, as the natural product extracts are too sparse. Three potent compounds were purchased, but their activity was determined to be an artifact of their fluorescence, meaning that they had escaped the early filtering of fluorescent compounds. The 64 confirmed hits from the natural product extract collection were tested for inhibition of an unrelated protein-peptide interaction used by another group at the screening facility. This assay also uses a short fluorescein-labeled peptide, but unlike MLL19 the peptide is not believed to bind

as an α -helix. 22 of 64 natural product extracts were selective for the MLL-KIX assay. These extracts all originate from lichens collected in Costa Rica.

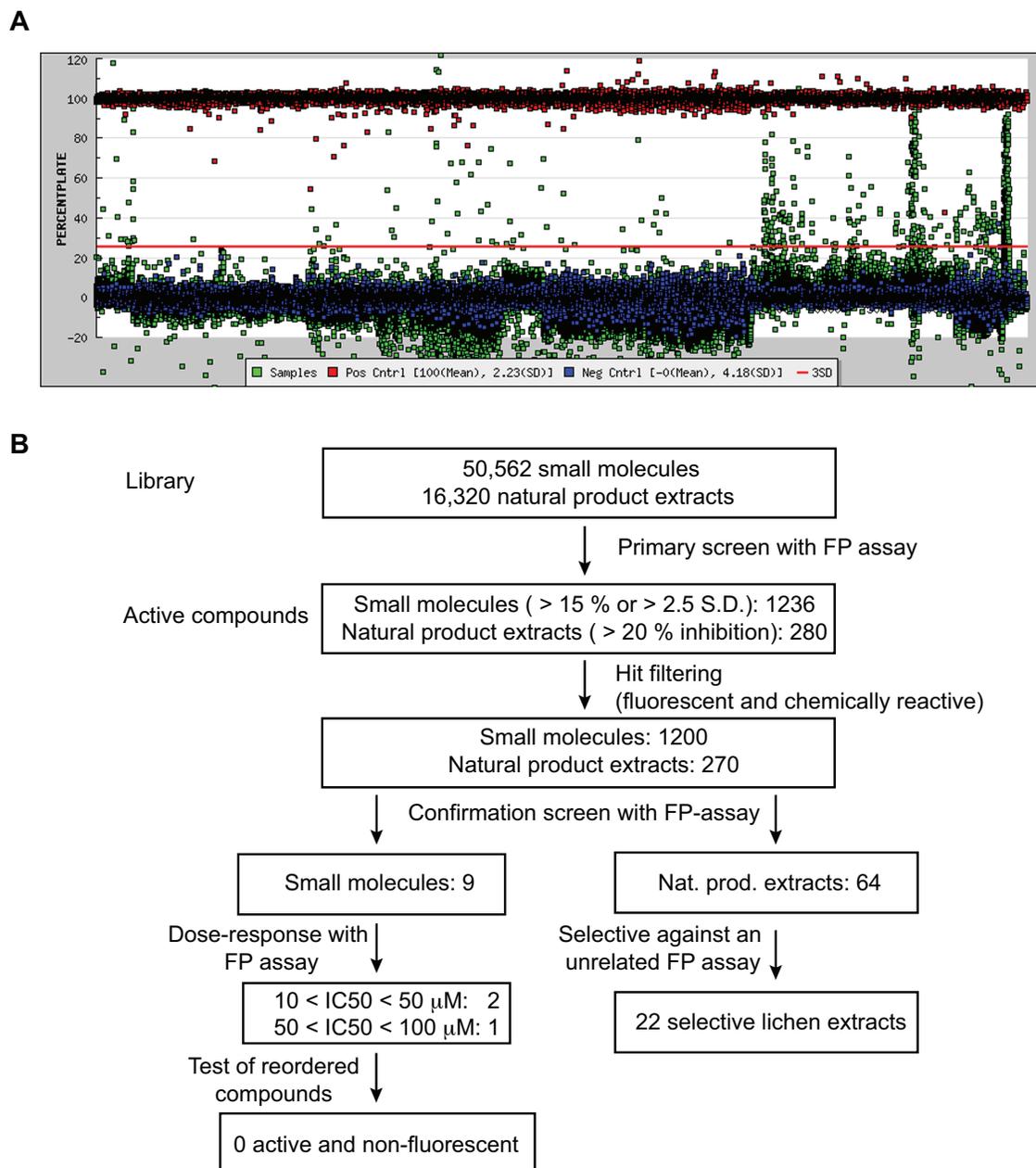


Figure 2.14 High-throughput screen of 66882 compounds for inhibitors of KIX-MLL A. Graphical representation of the % inhibition of all samples from initial screen; B. Flow-chart of hit selection.

Identification of two depsides as selective ligands of the KIX domain

To facilitate follow up studies, additional aliquots of the 22 lichen extracts were requested from the National Institute for Biodiversity in Costa Rica (INBio), who were able to provide 13 of these, of which 7 had shown more than 50% inhibition in the confirmatory screens. For one of the 13 lichen samples, the major component in the extract had previously been identified as 2'-O-methylaniziaic acid and this was supplied in pure form (Figure 2.15). This compound belongs to a common class of secondary lichen metabolites called depsides (named by Emil Fisher in 1910⁹⁰); the defining structure of depsides is two condensed phenol carboxylic acids.

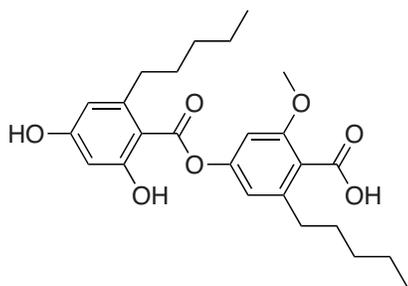


Figure 2.15 Structure of the lichen depside 2'-O-methylaniziaic acid

Of the 12 remaining samples received, 7 were received as both crude extracts and HPLC separated fractions. All crude extracts and fractions were tested in the MLL-KIX FP assay at 3 concentrations to confirm the activity and identify the most active fractions. In addition to 2'-O-methylaniziaic acid, 6 of the extracts shows activity (Figure 2.16), and it is possible to identify individual fractions with highest activity that likely contain the active compound(s) of the extract (Figure 2.16, 3rd column). Although the hits are selective for KIX-MLL inhibition over the counter screen used in the assay, they were tested against two additional FP-assays of activator-coactivator interactions: cMyb-KIX (cMyb binds to the other site on KIX) and VP2-Med15 (VP2 is an artificial dimer derived from the VP16 TAD sequence, Med15 is a yeast coactivator). They were also tested against a protein-DNA binding interaction (Gal4-DNA). Based on these tests, extract 2 appears to contain a compound in fraction 4, which selectively inhibits the two

KIX interactions (with MLL and cMyb), and the same selectivity is observed for extract 6 fraction 5.

Natural product extracts and fractions received March 2010 for reconfirmation of activity against MLL-KIX interaction		% Inhibition data in original screen		% Inhibition of MLL-KIX interaction in FP assay			% Inhibition against other FP assays			
		MLL-KIX	Counter screen	NI. = No inhibition			KIX-MLL	KIX-cMyb	Med15-VP2	Gal4-DNA
Extract #	Type	~ 0.3 mg/mL		0.8 mg/mL	0.08 mg/mL	0.008 mg/mL	0.3 mg/mL			
1	2'-O-methylanziaic acid	88	-9	58	70	N.I.		57	20	61
2	Fraction 1	85	15	N.I.	N.I.	N.I.				
	Fraction 2			6	N.I.	N.I.				
	Fraction 3			50	N.I.	N.I.				
	Fraction 4			65	57	N.I.	45	68	13	-20
	Fraction 5			77	8	N.I.				
	Crude			73	4	N.I.				
3	Fraction 1	78	3	N.I.	N.I.	N.I.				
	Fraction 2			31	N.I.	N.I.				
	Fraction 3			53	11	N.I.				
	Fraction 4			65	39	N.I.	38	63	29	42
	Fraction 5			65	28	N.I.	77	72	29	27
	Crude			74	1	N.I.				
4	Fraction 1	54	6	N.I.	N.I.	N.I.				
	Fraction 2			52	N.I.	N.I.				
	Fraction 3			37	6	N.I.				
	Fraction 4			53	84	N.I.	68	47	55	2
	Fraction 5			54	29	N.I.	71	44	24	-17
	Crude			62	56	2				
5	Fraction 1	46	3	N.I.	N.I.	N.I.				
	Fraction 2			67	41	N.I.				
	Fraction 3			78	25	N.I.				
	Fraction 4			67	22	N.I.				
	Fraction 5			80	64	N.I.	40	7	29	31
	Crude			73	22	N.I.				
6	Fraction 1	65	18	N.I.	N.I.	N.I.				
	Fraction 2			71	N.I.	N.I.				
	Fraction 3			65	62	N.I.				
	Fraction 4			65	68	N.I.				
	Fraction 5			77	77	N.I.	42	73	-4	-3
	Crude			68	13	N.I.				
7	Fraction 1	25	8	N.I.	N.I.	N.I.				
	Fraction 2			41	N.I.	N.I.				
	Fraction 3			67	70	12	50	26	39	38
	Fraction 4			53	36	N.I.				
	Fraction 5			49	N.I.	N.I.				
	Crude			59	N.I.	N.I.				

Figure 2.16 Activity of lichen extracts in MLL-KIX and other FP assays
 Orange color indicates high level of inhibition; yellow indicates mild inhibition.

(Figure 2.18A). This showed that sekikaic acid inhibits KIX-MLL with an IC_{50} of 58 μ M. The Hill-slope is significantly higher for sekikaic acid than the MLL-peptide. High Hill-slopes are indicators of cooperativity, but can alternatively be caused by irreversible inhibitors or by small-molecule aggregators, found to be common promiscuous inhibitors in high-throughput screening libraries.^{91,92} As the depsides lack reactive groups that would irreversibly react with amino acid side chains in a protein, it was assumed not to be an irreversible inhibitor (It is also evident in NMR studies presented below that the binding is not kinetically irreversible). With the large Hill-slope and the activity against both binding surfaces of the KIX domain, it is important to consider the possibility that the active species could be a promiscuous aggregate species. Early evidence against this, is the fact that the HTS was done in an assay buffer containing 0.01% NP-40 detergent to help prevent such aggregates, and the identified depsides had shown selectivity for KIX in the tests summarized in Figure 2.16. To further study the effects of detergents on the inhibition curve by sekikaic acid, buffers with varying amounts of Tween-20 were used (Figure 2.18B). The most shallow dose-response curve was seen with 0% Tween-20. Addition of 0.01-0.02% Tween-20 did not affect the IC_{50} significantly, but the Hill-slope increased dramatically. This observation does not point to activity from an aggregate species; detergents have the potential to disrupt aggregates making them less potent (increasing the IC_{50}).⁹² Furthermore, the high Hill-slope of aggregators is due to cooperativity of aggregate formation as well as high binding affinity of the aggregate for proteins, and detergents would be expected to lower this cooperativity. Thus, the steeper dose-response curves with added detergent are not consistent with an active aggregate species. The inhibition of FI-MLL19 – KIX interaction by both MLL15 and sekikaic acid was studied in a buffer that contained only 0.001% NP40 (which comes from the protein stock). MLL15 inhibits with an IC_{50} of 13.3 μ M and a Hill-slope of 1.18. It is worth noting that with a protein concentration of 5 μ M and a K_i of MLL15 at \sim 1 μ M ([KIX]: K_i 5:1) the Hill-slope is expected to be slightly higher than 1.^{93,94} The IC_{50} for sekikaic acid in this buffer is 52.1 μ M and the Hill-slope is 2.8. The Hill-slope was even

lower (1.6) in the detergent experiment (Figure 2.18B), but this may be due to the fewer data points used for the fit. Because of the high Hill-slope, a K_i for sekikaic acid cannot be determined by fitting the data to theoretical equations. For comparison, a hypothetical inhibitor that binds to only one binding site and has an IC_{50} of 52.1 μM would have a K_i of 3.1 μM (determined with equation: $K_i = (L_b)(IC_{50})(K_d)/((L_o)(R_o) + L_b(R_o - L_o + L_b - K_d))$; L_b = concentration of bound tracer, K_d = K_d of tracer, L_o = total concentration of tracer, R_o = total concentration of protein).⁹⁵

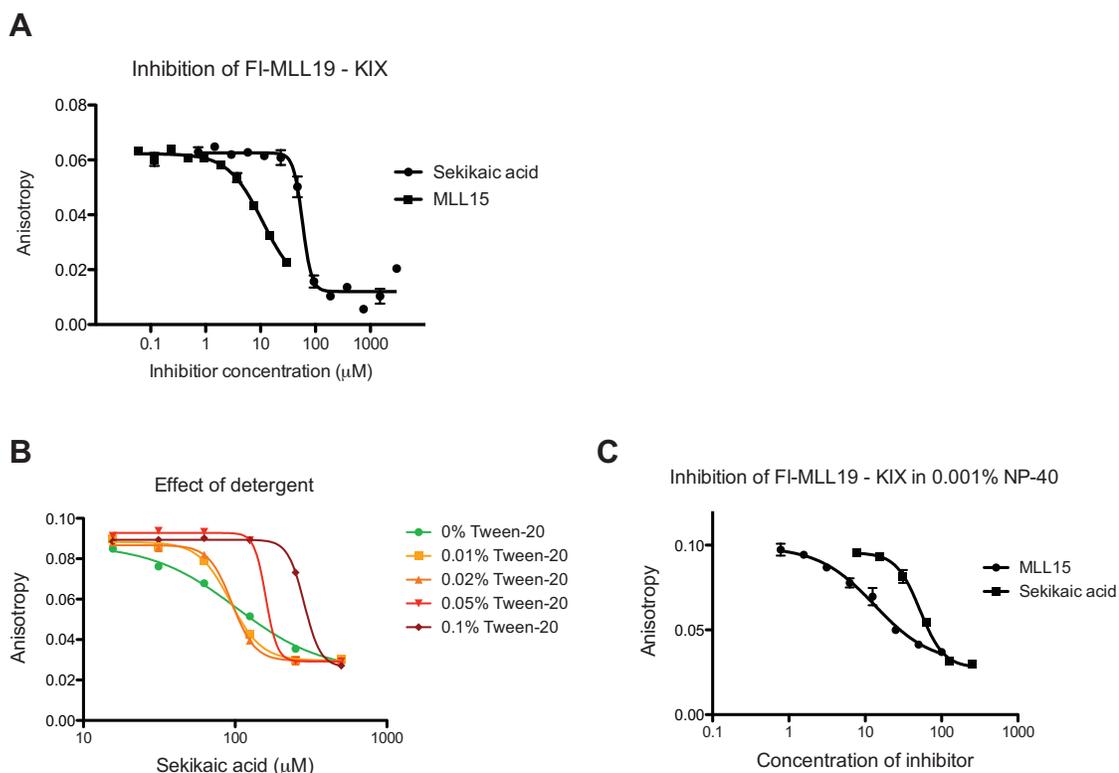


Figure 2.18 Inhibition of FI-MLL19 – KIX interaction by sekikaic acid A. Dose response of purified sekikaic acid in assay used in the screen; B. Dose-response of sekikaic acid with different concentrations of detergent; C. Dose-response of MLL15 and sekikaic acid in low detergent buffer.

A mutant version of the MLL15 peptide had been generated in the lab (by W.C. Pomerantz), through a systematic mutation of amino acids to yield a peptide with higher affinity for the MLL site on KIX. This peptide, named MLL15dm, contains M2850Nle (Nle: norleucine) and V2853I mutations, and binds with 3-fold higher affinity. A fluorescein-labeled version was made, FI-MLL15dm, and used as

tracer in an inhibition assay with sekikaic acid (Figure 2.19). The higher affinity gives this tracer the advantage of allowing a 5-fold dilution of KIX in the assay while maintaining a good dynamic range. Sekikaic acid inhibits FI-MLL15dm – KIX with an IC₅₀ of 37 μ M, which is compared to its IC₅₀ of 52.1 μ M against FI-MLL19 – KIX, where the concentration of KIX is 5-fold higher. If the active species is an aggregate with very low K_i ($K_i \ll [KIX]$), the IC₅₀ is expected to change nearly linearly with change in protein concentration (in this case it should change approximately 5-fold).⁹¹ This is further evidence that the active species is not an aggregate.

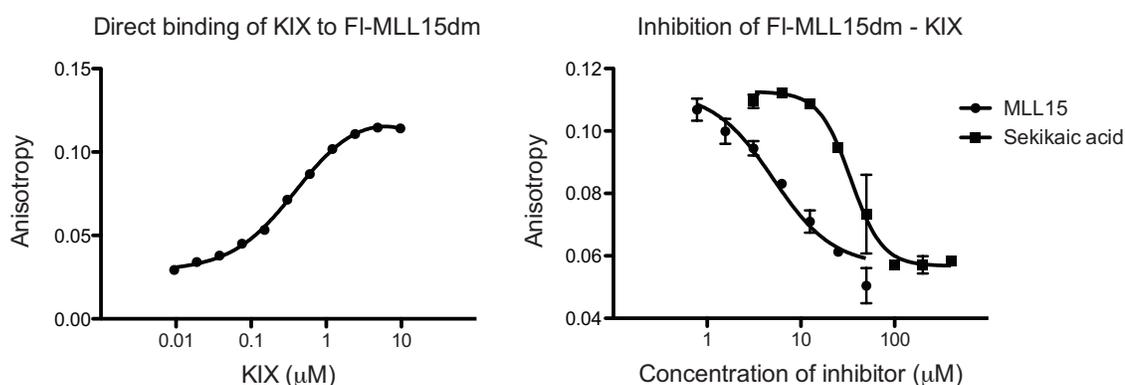


Figure 2.19 Inhibition of FI-MLL15dm – KIX interaction by sekikaic acid A. Direct binding of KIX to the FI-MLL15dm tracer; B. Dose-response of sekikaic acid and MLL15 on inhibition of FI-MLL15dm – KIX.

These binding studies all suggest that the binding mode of sekikaic acid is cooperative, and this can only be achieved if sekikaic acid has more than one binding site on KIX, since KIX is believed to be a monomer in solution. Because natural TADs can bind cooperatively to two separate sites on KIX it is appealing to suggest that sekikaic acid may also bind these same two sites. To determine if sekikaic acid can inhibit binding of a TAD to the other site on KIX, a 29-mer phosphorylated peptide derived from CREB, KID29, was used as tracer (FI-KID19, $K_d = 1.37 \mu$ M) in an FP assay (Figure 2.20). Sekikaic acid inhibits FI-KID29 – KIX with an IC₅₀ of 63.7 μ M and a Hill-slope of 2.7 (KIX at 2.5 μ M). A previously reported inhibitor of KIX – KID interaction, 2-naphthol-AS-E-phosphate (Figure 2.10), was also tested in the assay. From a fitted curve, this compound

has an IC₅₀ of 387 μ M and a Hill-slope of 1.7 (only a partial inhibition curve was observed in the tested concentration range).

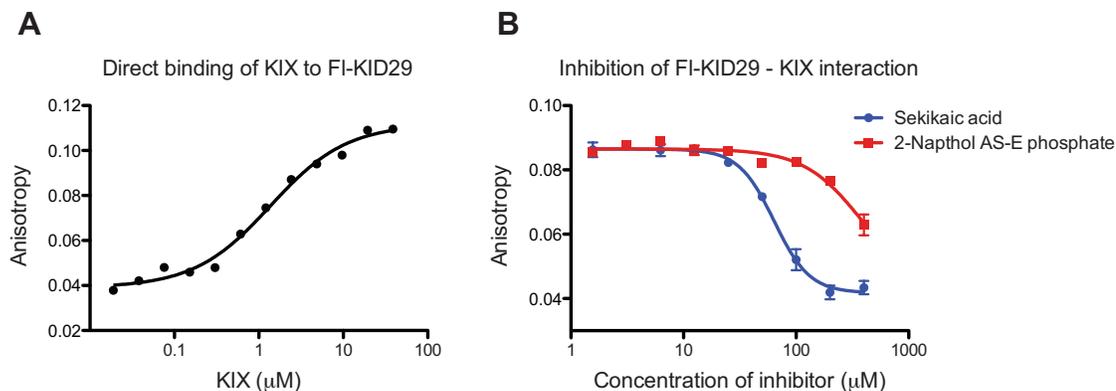


Figure 2.20 Inhibition of FI-KID29 – KIX interaction by sekikaic acid A. Direct binding of KIX to the FI-KID29 tracer; B. Dose-response of sekikaic acid and 2-naphthol AS-E phosphate on inhibition of FI-KID29 – KIX.

To determine if the binding of sekikaic acid to KIX is reversible, and can be inhibited by MLL15dm and KID29 peptides, a ligand-observed 1D-¹H-NMR experiment was used. Upon addition of KIX to sekikaic acid, the chemical shifts of protons in sekikaic acid are perturbed (Figure 2.21). Most diagnostic are the chemical shifts of the aromatic protons, which are isolated in the spectrum from KIX and peptide chemical shifts. The chemical shifts of these three residues shift and broaden slightly in presence of KIX. The change in chemical shift and mild broadening are both evidence of a bound species of sekikaic acid. A single peak is observed as an average of bound and unbound sekikaic acid, because the exchange of bound and unbound species is fast relative to time scale of the NMR acquisition parameters, consistent with a $K_D > 0.1 \mu$ M (approximated by the IC₅₀).⁹⁶ In above analysis of data presented in Figure 2.18 it was determined that the IC₅₀ of sekikaic acid would be identical to a single site competitive inhibitor with a K_i of 3.1 μ M. If it is assumed that sekikaic acid binds to only one binding site on KIX with this affinity the fraction of bound sekikaic acid in the 1D-NMR experiment (+KIX trace) can be determined to be 26%, by using equation 6 from Roehrl et al.⁸⁸ Addition of either MLL15dm or KID29 at equimolar amounts

relative to sekikaic acid, does not cause a significant change in the fraction of bound sekikaic acid as determined by the lack of perturbation of chemical shifts towards the chemical shifts of unbound sekikaic acid. When both peptides are added, however, the chemical shifts change significantly towards the spectrum of free sekikaic acid. Thus, this experiment shows that the binding of sekikaic acid to KIX is reversible and can be inhibited when both MLL15 and KID29 are added, but not as readily when each of the peptides are added alone. Competition by single peptides at twice the concentration gives the same result. The cooperative binding mode of sekikaic acid is stronger than the binding of either peptide alone. The two peptides also bind KIX cooperatively,⁸³ and this may explain how they can compete with sekikaic acid binding when added together.

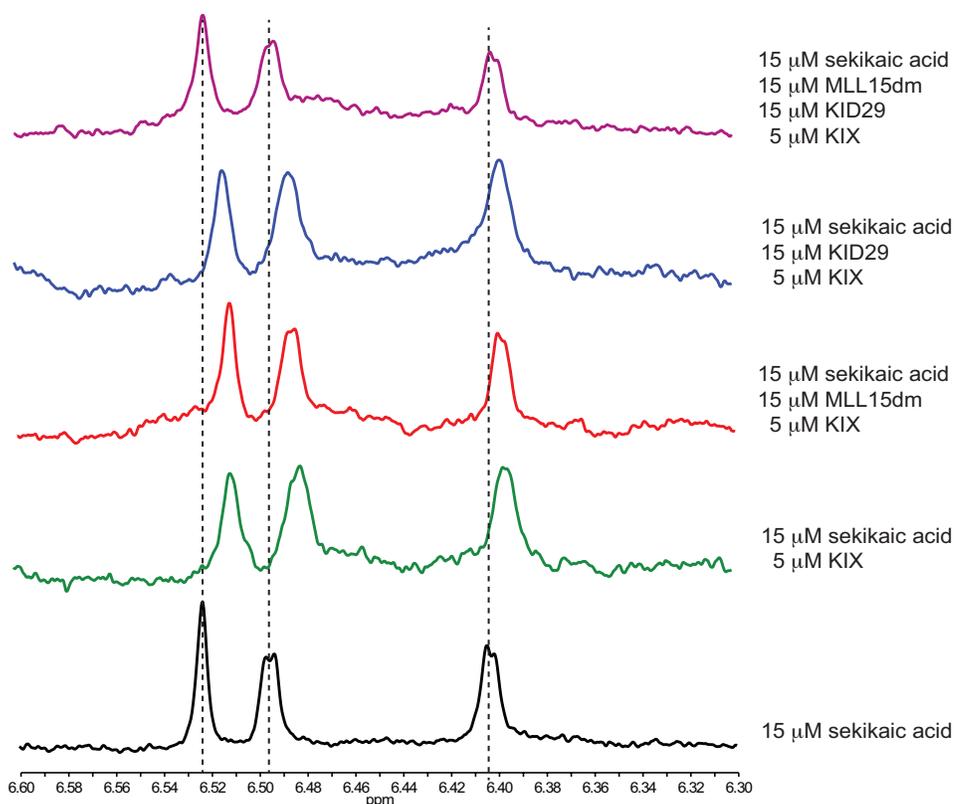


Figure 2.21 1D-¹H-NMR of sekikaic acid binding to KIX The chemical shifts of aromatic proton resonances of sekikaic acid are isolated from other proton resonances in the presence of KIX and TAD peptides. Binding to KIX perturbs the resonances of sekikaic acid as shown in green trace. Addition of MLL15dm and KID29 peptides (purple) but not either of peptide alone (red and blue), inhibits binding of sekikaic acid to KIX.

To map the binding sites of sekikaic acid on the KIX domain, an HSQC-NMR spectrum of ^{15}N -labeled KIX in the absence and presence of increasing relative equivalents of sekikaic acid was recorded. This type of protein observed NMR experiment is preferably done with high concentration of protein to achieve good signal-to-noise spectra. At high concentrations of both KIX and sekikaic acid, however, the protein precipitated out of solution (visibly observed), this was confirmed by loss of detectable protein resonances by NMR. The exact limits in concentrations of KIX and sekikaic acid were difficult to ascertain, but in general sekikaic acid concentrations higher than 150-175 μM tended to cause precipitation of protein, and the precipitation was more predominant with KIX concentrations above 35 μM . The cause of the precipitation is not known. It is possible that precipitation is caused by aggregates of sekikaic acid, which may form at high concentrations of sekikaic acid. It is also possible that the conformation of KIX associated with fully bound sekikaic acid is prone to self-aggregation, which would explain the dependence on protein concentration. In fact the KIX protein is difficult to concentrate beyond 100 μM on its own and is perhaps only a semi-stable fold.⁹⁷ A ^1H -NMR spectrum of sekikaic acid at 500, 250 and 125 μM in D_2O was recorded (Figure 2.22). These spectra show very mild line broadening of sekikaic acid chemical shifts, and no change in the chemical shifts of sekikaic acid except for an emergence of weak, broad peaks in the aromatic proton region. The later peaks could be evidence of aggregates. The 500 μM solution was slightly turbid, and this turbidity was also observed as light scattering in a UV/VIS absorption spectrum (not shown). Light scattering was not seen for samples at 125 or 250 μM . All protein observed NMR experiments are therefore done at a low concentration of KIX (10-25 μM) to allow ≥ 5 -fold equivalents of sekikaic acid to be added.

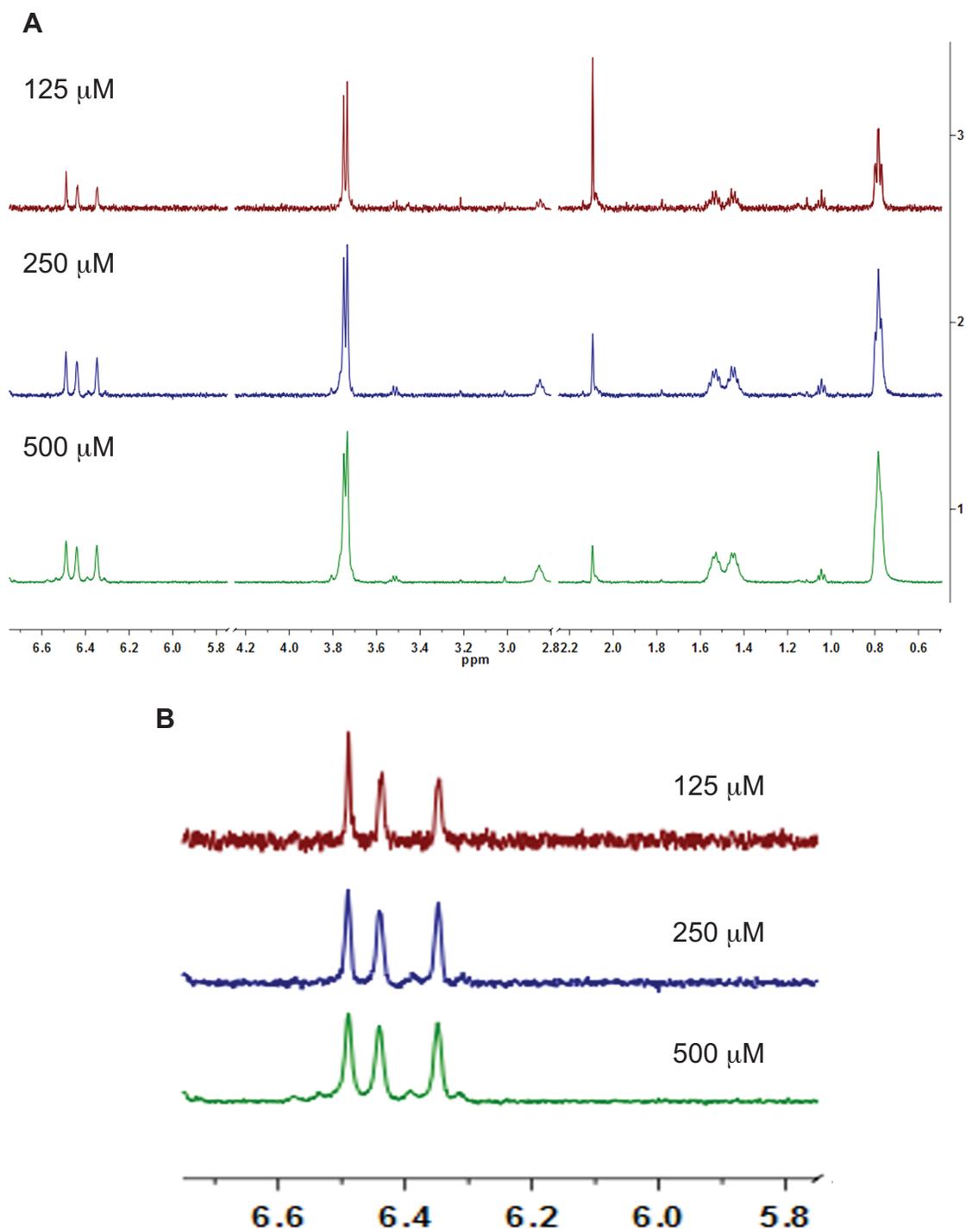


Figure 2.22 $^1\text{H-NMR}$ of sekikaic acid to monitor indications of aggregation at high concentration A. All peaks; B. Aromatic resonances.

A ^1H - ^{15}N HSQC spectrum was recorded of KIX using 10 μM protein with 0, 1, 2.5, and 5 equivalents of sekikaic acid added. Binding of sekikaic acid is evident from perturbation of a chemical shift observed of protein backbone amides for a few residues. The majority of shifts are unchanged indicating that any conformational change associated with binding to the protein is subtle (Figure 2.23). The assignment of the spectrum is done by comparison to reported spectra, and the majority of peaks can be assigned in this way, although with less certainty in assignments.^{89,97} Chemical shift changes for individual peaks are quantified as $((0.2 \cdot \Delta\delta^{15}\text{N})^2 + (\Delta\delta^1\text{H})^2)^{0.5}$, which is a weighted length of the vector from free KIX to 1, 2.5, or 5 equivalents of sekikaic acid. The 0.2 fractional weight of $\Delta\delta^{15}\text{N}$ in this vector is also used by others,⁸⁹ and correspond in this experiment to the relative differences in standard deviations of $\Delta\delta^{15}\text{N}$ and $\Delta\delta^1\text{H}$. The chemical shift changes for KIX with 5 equivalents of sekikaic acid are graphed in Figure 2.23C. The threshold on the graph is based on the average chemical shift change + 3 standard deviations from the sample with 1 equivalent of sekikaic acid (this sample had no large changes in shifts, and the variation is a better indicator of random variation). Significant chemical shift changes are observed for 5 residues. These are identified in spectra and mapped onto a structure of KIX (Figure 2.24). Of the peaks with shifts close to the threshold, several do not observe systematic changes when comparing to sample with 2.5 equivalents of sekikaic acid, and some are caused by low peak height and imprecise peak placement. These may reveal true but mild perturbations, but none of them map closely together on KIX to indicate a binding site. Although not all peaks are assigned and some peaks are not identified because they overlap, no other peaks show as significant shifts as the 5 identified residues. Mapping of the backbone amides of these 5 residues indicate that sekikaic acid may bind in a pocket formed by three helices, to which it can bind on the inside (2.24A) or outside (2.24B) face relative to MLL. It is also possible that a molecule binds on both faces, but this is speculative. In a solution NMR structure, this region was found to be the most flexible.²¹

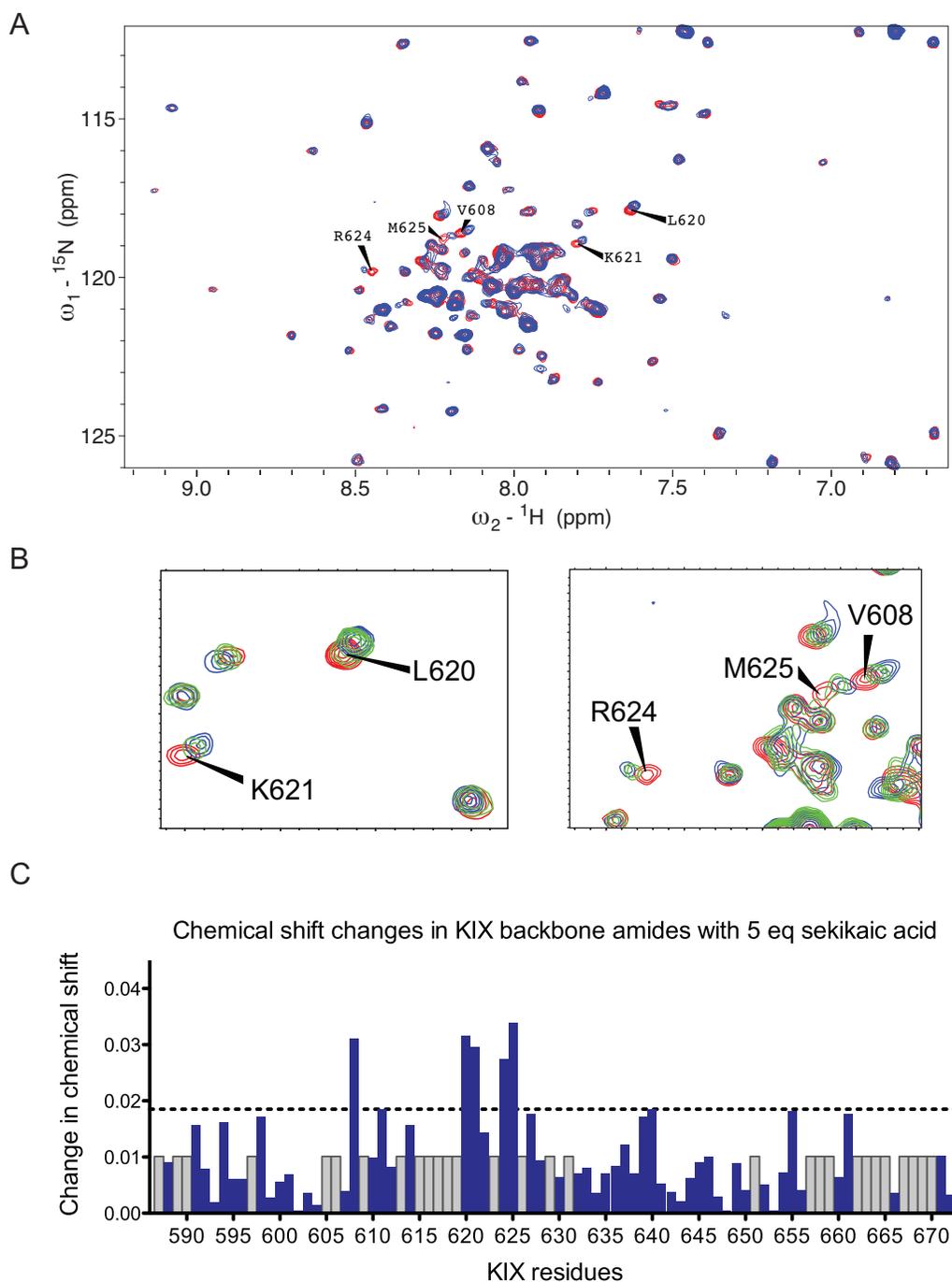


Figure 2.23 ${}^1\text{H}$ - ${}^{15}\text{N}$ HSQC of KIX with sekikaic acid A. Overlaid spectra of free KIX (red) and KIX + 5 eq. sekikaic acid (blue); B. An enlarged view of regions with residues that observe significant chemical shift changes. Spectrum from 2.5 eq. sekikaic acid (green) is added; graph of chemical shift changes for each residue in KIX. Unassigned residues are represented with a grey bar with an average shift change as height.

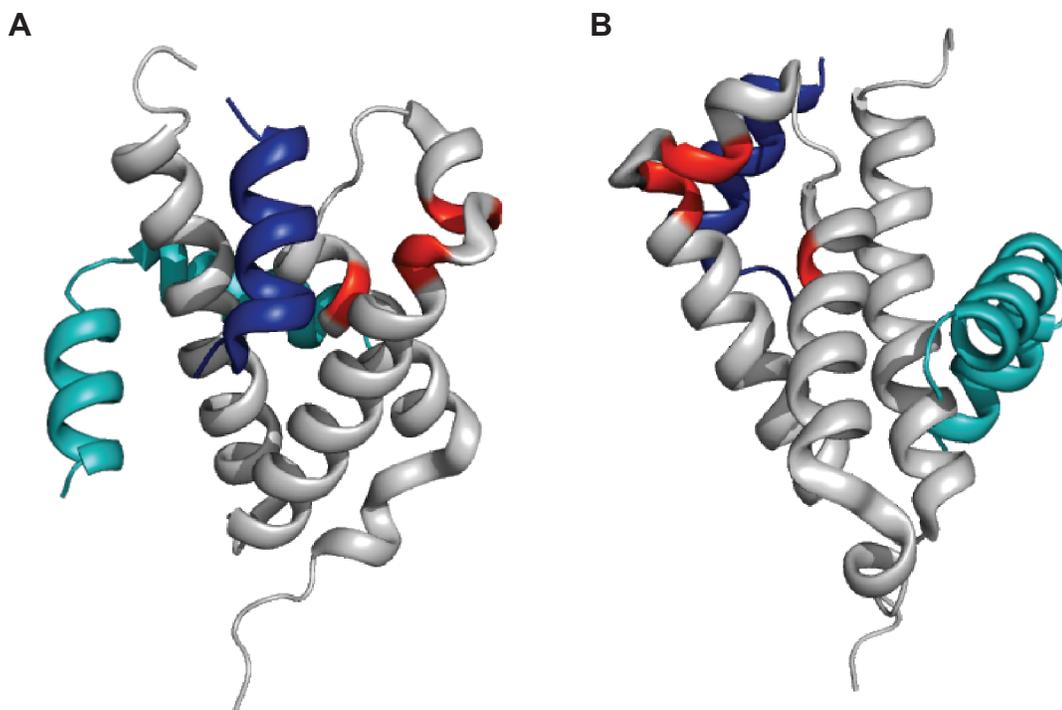


Figure 2.24 Map of chemical shift changes in sekikaic acid-bound KIX KIX (grey) with perturbed backbone amides (red) (PDB:2AGH).²⁹ MLL is blue and KID (Superimposed from PDB:1KDX)²¹ is turquoise.

Cell-based studies with sekikaic acid

Sekikaic acid was identified as a new ligand of the CBP KIX domain in a biochemical screen, but it remains to be seen if cellular activity can be demonstrated. Sekikaic acid's effect on viability of human cell lines, HeLa and HEK293T, was studied to define concentrations at which functional studies can be carried out (Figure 2.26). The concentration of sekikaic acid that causes 50% cell death (LC50) is 37 μM and 32 μM in A549 cells at 24 and 48 hours treatment respectively, and it is 25 μM and 18 μM in HEK293T cells at 24 and 48 hours respectively.

Sekikaic acid was tested in HEK293T transfected with a luciferase reporter plasmid and a Gal4-fusion activator with either the MLL or the CREB activation domain. Cells transfected with Gal4-CREB also received an expression plasmid encoding protein kinase A (PKA), which phosphorylates the CREB TAD and is necessary for CREB binding to the KIX domain of CBP/p300.⁹⁸ At concentrations

of up to 20 μM sekikaic acid, no inhibition of Gal4-CREB was observed. The activity of Gal4-MLL looks to be inhibited to a small degree, and further experiments will assess the significance of this result (Figure 2.27).

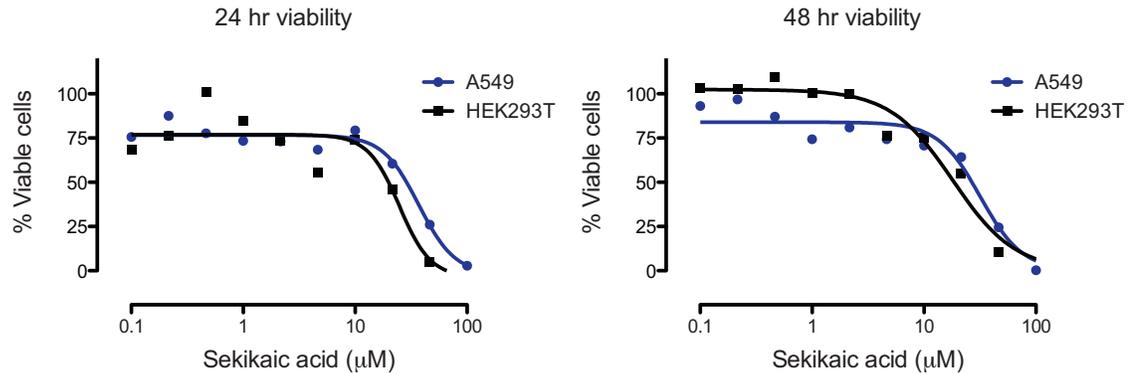


Figure 2.26 Cytotoxicity of sekikaic acid The viability of two human cell lines at varying concentrations of sekikaic acid is determined with WST-1 viability dye added 24 or 48 hours after addition of compound. Values are relative to DMSO.

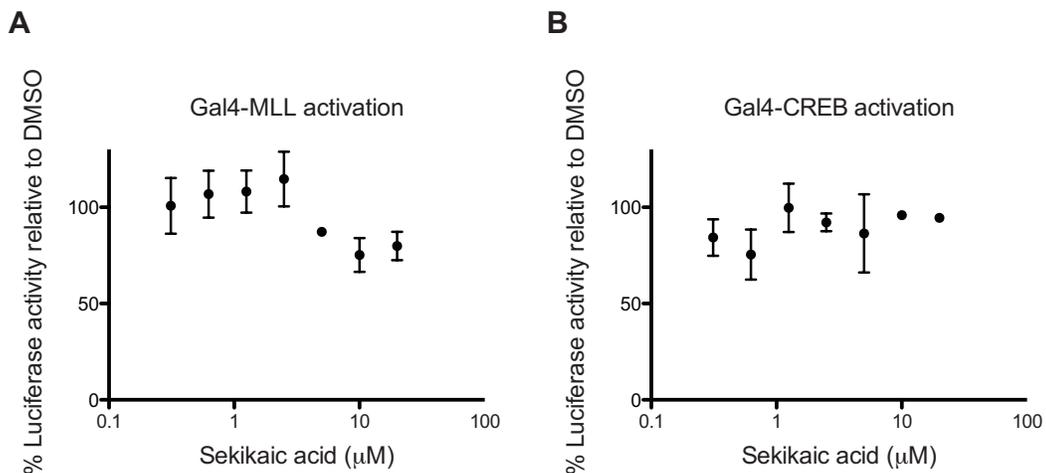


Figure 2.27 Effect of sekikaic acid on MLL and CREB activated transcription A. Gal4-MLL fusion activator introduced together with luciferase reporter with Gal4-binding sites into HEK293T cells. The activity of sekikaic acid treated samples relative to DMSO treatment is shown. B. Same as in A, except fusion activator is Gal4-CREB and it is cotransfected with PKA. Data is from one experiment individual samples done in triplicate.

It is not known to what extent the activators are depended on CBP, and in fact CBP dependent CREB activity is context dependent.⁷³ To explore an assay that

would be entirely depended on MLL-KIX interaction, a mammalian-two-hybrid assay was tested (Figure 2.28). Fusions of MLL and KIX with Gal4-DBD and with the strong VP16 activation domain were made, and these were cotransfected with a luciferase reporter into HeLa cells. Gal4-MLL activates transcription, but with no increase in transcription when cotransfected with VP16-KIX, despite VP16 being a stronger activation domain. Gal4-KIX did not activate transcription, either on its own or when cotransfected with VP16-MLL. Thus, the KIX-MLL interaction is not detected by this two-hybrid assay. It is possible that the interaction is too weak and a two-hybrid assay may give better result in presence of a peptide that binds cooperatively with MLL to KIX.

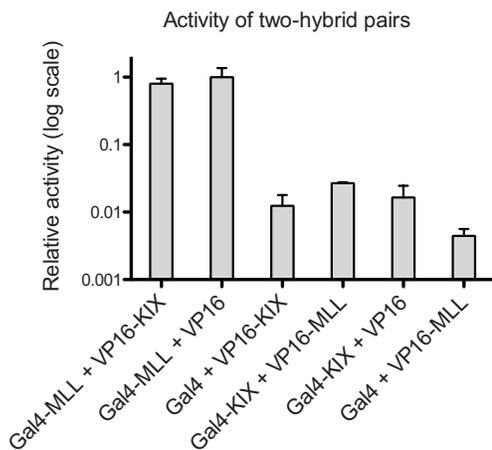


Figure 2.28 KIX-MLL two-hybrid An interaction between KIX and MLL is probed with a two-hybrid assay.

Test of a few sekikaic acid related structures

The only hits from the high-throughput screen were 22 lichen extracts. In the two most KIX selective extracts, the depsides sekikaic acid and microphyllinic acid were identified. Depsides are common secondary metabolites of lichens, and it wouldn't be surprising if all the active lichen extracts contain a depside or related structure as its active component. In fact, one of the extracts was already known to contain the depside, 2-O-methylaniziaic acid, which is active in the MLL-KIX competition FP assay and not in the counter screen, and thus displays some level of target specificity, despite inhibiting VP2-Med15 and Gal4-DNA interactions. These interactions are not completely unrelated to KIX-MLL; VP2-

Med15 is an interaction between an amphipathic helix and a coactivator that binds multiple other amphipathic activators, and amphipathic helices are also known to bind the DNA-binding domain of Gal4. Commercially available depsides or depside-related compounds could be a useful source for further studies with this class of compounds as KIX ligands. Two compounds were purchased (Figure 2.29): Lecanoric acid is a depside whose most obvious difference from the three depsides identified in the screen, is the lack of an aliphatic chain (it has a methyl group) on each aromatic ring; Lobaric acid is a depsidone, which is another common secondary metabolite of lichens and is made from a depside precursor. The two aromatic rings in depsidones are connected by both an ester and an ether linkage and are consequently more conformationally restrained. Lobaric acid has aliphatic chain substituents. The two structures were tested for inhibition potential in the FP assays (Figure 2.29). Lobaric acid (blue curve), but not lecanoric acid (red curve), displayed similar inhibition potential as sekikaic acid and inhibited MLL binding with an IC₅₀ 16.9 μM and inhibited KID binding with an IC₅₀ of 24.7 μM. The Hill-slope is ~ 2.5.

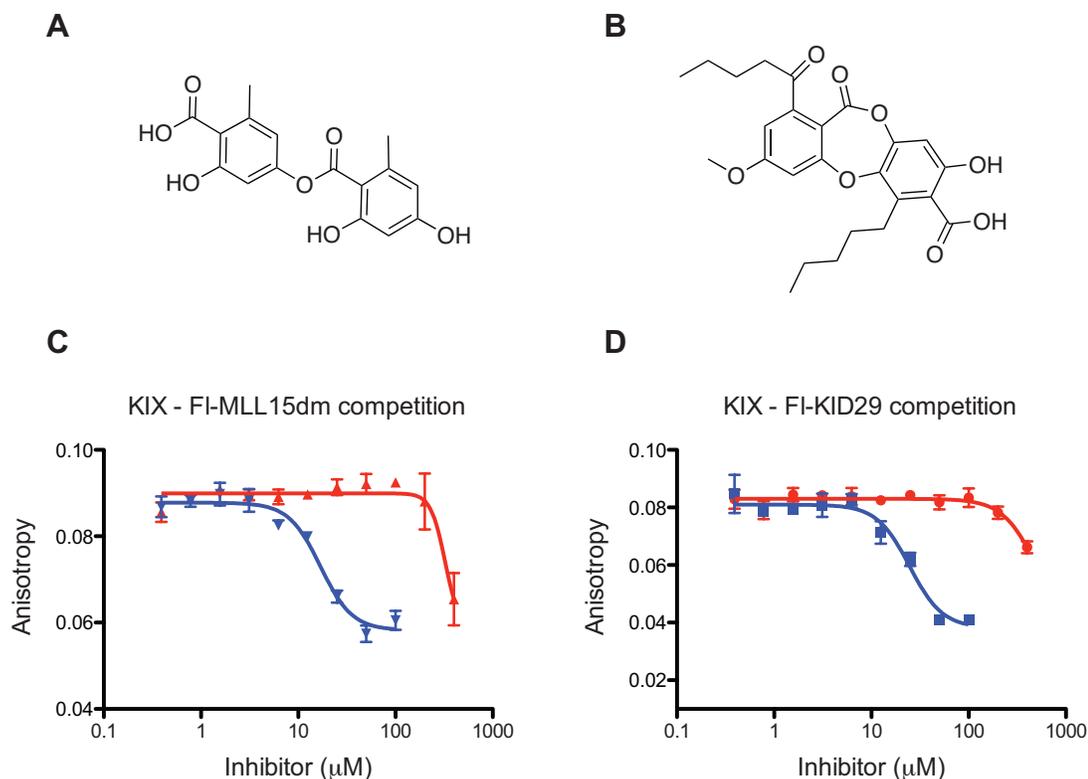


Figure 2.29 Competitive inhibition of KIX tracers with lecanoric acid and lobaric acid A. Lecanoric acid; B. Lobaric acid; C. and D. Competitive inhibition curves of lobaric acid (blue) and lecanoric acid (red).

The carboxylic acid of sekikaic acid is important for binding

One ultimate goal with ligands of coactivator proteins is to link them to DNA-localizing molecules. The depsides contain a carboxylic acid, which is an easy chemical handle to use for conjugations, but may be important for binding to KIX. As a first test to see if the acidic group is needed, the carboxylate was methylated with trimethylsilyldiazomethane.⁹⁹ The methylated sekikaic acid (Figure 2.30), however, had poor solubility in water with visible precipitation above 75 μM . In an FP competition assay, no inhibition was seen below 75 μM . Thus, the polar group helps with solubility, but it cannot be determined from this modification if the acid group is important for binding to KIX. The carboxylic acid group was instead reacted with an amine of cystamine to link cystamine to sekikaic acid with an amide. The goal with this molecule was to functionalize

sekikaic acid with a disulfide, which can undergo reversible disulfide exchange with other thiols, including cysteines in proteins. If a thiol-functionalized ligand binds a spot on a protein surface that is within sufficient proximity of a cysteine for a disulfide bond to form, the disulfide can tether the molecule to the protein and greatly enhance its apparent affinity. This has been utilized to facilitate fragment-based screening¹⁰⁰ and covalent tethering has also been used to enhance inhibitor potencies.¹⁰¹ To my knowledge, the same principle has not been used to map binding sites of discovered ligands. The proposed strategy was to use the disulfide-modified sekikaic acid (Sek-SSEA) in binding studies with multiple different KIX proteins bearing different cysteine point mutations. Identification of individual mutants to which Sek-SSEA has enhanced inhibition potency would effectively map the binding site to a region in proximity of the Cys-mutation. Sek-SSEA is water soluble and can be tethered to KIX with Cys-mutations (by mass spec), but Sek-SSEA does not inhibit the MLL – KIX interaction in an FP competition experiment. Thus, it appears that the derivative has lost its ability to interact with KIX.

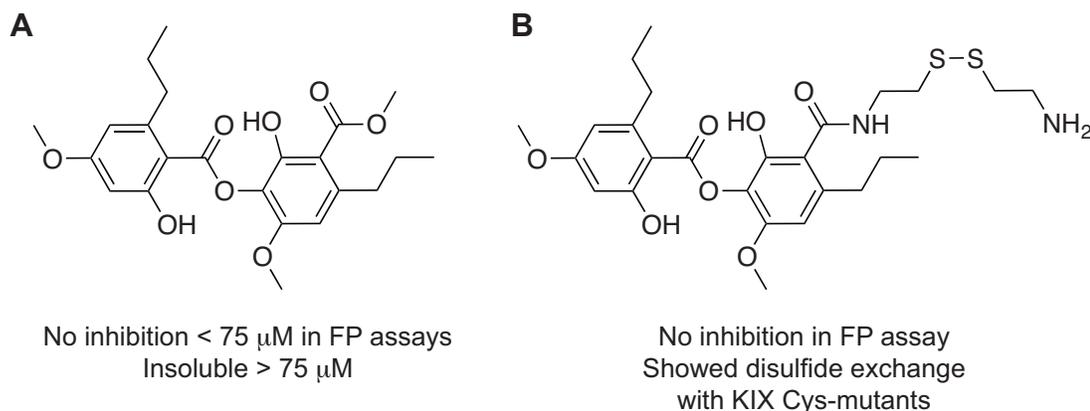


Figure 2.30 Carboxyl-derivatives of sekikaic acid A. Methylated sekikaic acid (Sek-OMe); B. Cystamine-modified sekikaic acid (Sek-SSEA).

2.6 Conclusion and discussion of sekikaic acid discovery

A simple biochemical assay facilitated the discovery of a new class of ligands of the CBP/p300 KIX domain. The lichen depsides were discovered as active

components in extracts from lichen, which are part of a unique natural product library available at the University of Michigan Center for Chemical Genomics. One of these depsides is sekikaic acid, which showed selectivity for inhibiting interactions between the CBP KIX domain and activators binding to two distinct sites over other protein-peptide interactions. The binding mode of sekikaic acid was studied, and a steep dose-response curve appears to be the product of cooperative binding of sekikaic acid to the KIX domain.

From ligand observed 1D-¹H-NMR spectra, reversible binding of sekikaic acid to KIX is evident. Furthermore, the cooperative binding of sekikaic acid is likely the reason that peptides to each of the two binding sites of KIX, added at equimolar (or double) concentration of sekikaic acid, do not effectively compete with sekikaic acid binding, despite showing lower IC₅₀ values in competition FP experiments. Simultaneous addition of peptide ligands to both binding sites, however, effectively competes off sekikaic acid. This result rejects a model, where a sekikaic acid molecule binds to each to the two TAD-KIX interaction sites on KIX and achieves cooperative binding through the same allosteric conformational change that accompanies cooperative binding of MLL and KID. If this was the case, MLL would bind cooperatively with sekikaic acid bound in the KID site, and vice versa for KID, and MLL and KID should be able to compete effectively with sekikaic acid bound at their respective binding sites to produce partial competition in the 1D-NMR experiment. Thus, the cooperative binding of sekikaic acid must be a consequence of either: A) two (or more) molecules of sekikaic acid binding close together on KIX with favorable interactions with each other and KIX (allostery not necessary); B) two (or more) molecules of sekikaic acid binding distant sites that allosterically influence each other, but with an associated conformational change in KIX that is different from conformational change associated with MLL and KID cooperativity. In scenario A, protein allostery is not necessary to explain the cooperative binding behavior, but two molecules of sekikaic acid binding close to each other can not overlap with both

binding sites on KIX. There must therefore still be an allosteric effect on at least one of the two binding sites.

In an HSQC experiment, residues within a small region of KIX observed chemical shift perturbations when sekikaic acid was bound. This is a likely binding spot of sekikaic acid. It is not clear if two molecules of sekikaic acid bind this region, but in no other region of KIX can sekikaic acid binding be detected. It is puzzling that the KID-binding site appears unperturbed by sekikaic acid, but perhaps changes occur without changing the chemical environment of the backbone.

Some structure-activity relationships have emerged from the studies. Since 22 lichen extracts were active, several depside or depside-like structures likely make up a class of compounds with ability to bind KIX. Differential activity of the extracts suggests that the active components are not identical. The structure of only a few of these was determined. In the active depsides the ester-linked aromatic rings have both phenolic groups and extended aliphatic chains as substituents. Lecanoric acid, which lacks the aliphatic substituents, is not active. The depsidone, lobaric acid, was active, despite its more conformationally restricted structure. The cyclic ester-ether connection of the two aromatic rings may also make this compound more stable in biological studies, although depsides are thought to be stable structures resistant to non-enzymatic hydrolysis.¹⁰² The depsides (and depsidone) all contain a carboxylic acid, which for sekikaic acid was not tolerant of substitutions. The substitutions may not be tolerated because of steric clashes in binding site, or perhaps the acidic group participates in ionic interactions. In the later case, acyl sulfonamide derivatives may be tolerated.¹⁰³

We were not able to show effects in cell culture indicative of inhibition of KIX recruitment. Sekikaic acid is toxic above 10 μ M, but the extent to which it is cell permeable can't be deduced. The CBP dependence of activators at individual genes or reporters is difficult to determine, which is one reason that we are

pursuing inhibitors. Parallel efforts in our lab are aimed at exploring such dependence through knockdowns and competition with expressed single domains of CBP. Activators identified through these efforts to be dependent on the KIX domain, will be valuable model systems to test the activity of sekikaic acid and related structures in cells.

Since sekikaic acid does not appear to bind in the binding sites for MLL or KID, although proximal to MLL, there is a better chance that it will not bind promiscuously to other coactivators, unlike TAD mimics. This selectivity was even evident from experiments (Figure 2.16). The cooperative binding is most likely unique to the KIX domain. This may be the result of an inherent property of the KIX domain, as cooperativity is also found for TADs binding the two canonical binding sites.^{28,29,83} The cooperativity could provide sekikaic acid with an additional level of specificity. The human mediator protein, Med15, also contains a KIX domain, and could be a second target of sekikaic acid.

2.7 Materials and methods

2.7.1 Materials and methods used with cell-based assays

Materials for loss-of-signal luciferase reporter assays

HeLa cells were purchased from ATCC. HEK293T cells were a gift from Gary Glick (University of Michigan). HLR cells (Pathdetect HLR cell line) were purchased from Stratagene. All cells are grown in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen, 11965) supplemented with 10% fetal bovine serum (Invitrogen, 16000) in 37°C incubator with 5% CO₂. HLR cells are grown in media additionally supplemented with 100 µg/mL hygromycin B (Calbiochem). Hygromycin B is not included in media at any step of the assays. OptiMEM media (Invitrogen) and Lipofectamine 2000 (Invitrogen) was used in transfections. G418 (Research Product International Corp.) was used for selection of stable HLR-derived cell lines at 800 µg/mL.

The Gal4-fusion vectors are from lab stocks. pGal4-VP16 is identical to pM3VP16¹⁰⁴. pGal4-RelA and pGal4-Elf3 were made by subcloning DNA oligomers (Integrated DNA Technologies, IDT) encoding RelA(522-551) and Elf3(129-159) into BglIII and NotI digested pGal4(1-147)-hGR(499-777)¹⁰⁵ (Gift from Thomas Kodadek, The Scripps Research Institute).

N2-biphenyl-iTAD and adamanolol were synthesized according to published procedures.^{39,106} Cycloheximide was purchased from Sigma-Aldrich. Luciferase activities are measured using Dual-Luciferase Assay Reporter System or One-Glo substrate (Promega) on single cuvette luminometer (Lumat LB 9507, Berthold), or plate luminometers (Genios Pro, Tecan or PheraStar Plus, BMG Labtech). Cell viability is measured with WST-1 reagent (Roche) at 440 nm on multi-well spectrophotometer (Genios Pro, Tecan).

Activation potency of Gal4-fusion activators

15,000 HLR cells were plated in 96-well plate with 100 μ L DMEM + 10% FBS. After overnight incubation, media is replaced with transfection mix: 100 μ L OptiMEM media with 1 μ g pGal4-VP16, pGal4-Elf3 or pGal4-RelA and 3 μ L Lipofectamine 2000. After 3 hours media is changed back to DMEM + 10% FBS. 24 hours after transfection the luciferase activity is measure with Firefly substrate from Dual-Luciferase Assay Reporter System on Berthold luminometer.

Generation of stable cell lines

HLR cells in a 10cm dish were incubated for 3 hours with transfection mix: 10 mL OptiMEM with 5ug pGal4-ESX or pGal4-RelA and 15 μ L Lipofectamine 2000. Following transfection, cells were kept in DMEM + 10% FBS for 24 hours. Cells were split onto plates at 1:20 – 1:100 dilutions. After 24 hours, media is replaced with DMEM + 10% FBS + 100 μ g/mL hygromycin + 800 μ g/mL G418. G418 concentration was previously determined with a kill-curve experiment: 800 μ g/mL G418 is lowest dose that kills untransfected HLR cells in 7 days. After 9 days, massive cell death has occurred and 1:20 split plates have 10-20 resistant

colonies. All cells are trypsinized and replated in 96-well plate at 1 cell/well density (confirmed by visual inspection after small colonies grew in wells). A sample from the suspension of cells for replating showed luciferase activity. Only 3 clones from pGal4-Elf3 transfected cells and 2 from pGal4-RelA transfected cells were isolated from 96-well plates. These clones were tested with firefly substrate from Dual-Luciferase Assay Reporter System.

Assay for Gal4-Elf3 inhibition by N2-biphenyl-isoxazolidine and adamanolol

15,000 HeLa cells were plated in 96-well plate with 100 μ L DMEM + 10% FBS. After overnight incubation, media is replaced with transfection mix: 100 μ L OptiMEM media with 0.5 μ g pGal4-Elf3, 0.5 μ g pGa5luc, 10ng pRLSV40, and 3 μ L Lipofectamine 2000. After 3 hours media is changed back to DMEM + 10% FBS. Compounds are added at 25, 5 and 1 μ M final concentration in 1% final concentration of DMSO. 24 hours after transfection, WST-1 reagent was added to media of one of three replicate samples and the absorption measured on Tecan Genios Pro. Following the WST-1 read, the luciferase activity is measured with Firefly substrate from Dual-Luciferase Assay Reporter System on Berthold luminometer. The WST-1 treated samples did not appear to affect the luciferase readout.

Gal4-VP16 inhibition assay used in the pilot screen

The pilot screen is done at the Center for Chemical Genomics, University of Michigan. 1.5×10^6 HLR cells plated in 10 cm dish with DMEM + 10% FBS. After 18 hours, the media is replaced by transfection mixture: 5 μ g pGal4-VP16, 15 μ L Lipofectamine 2000, 10 mL OptiMEM. After 5 hours transfection mixture is removed, and cells are trypsinized and resuspended in phenol red-free DMEM + 10% FBS at 300,000 cells/mL. At HTS facility, 20 μ L media is added to wells of white, sterile, TC-treated 384-well plate (cells were also plated in a clear 384-well plate to monitor cells with negative and positive control). Compounds are then added to wells with pin-tool, and 20 μ L of resuspended transfected cells are added to the wells (6000 cells/well). Positive control is added at 10 μ g/mL (lowest

concentration determined to completely inhibit translation of luciferase mRNA in a dose-response test). After 18 hours, the assay is developed by: removal of 30 μ L media, addition of 10 μ L One-Glo substrate, and measurement of luminescence on PheraStar luminometer.

Materials used in gain-of-signal assay

HeLa cells, DMEM and OptiMEM are the same as above, but tetracyclin-free fetal bovine serum is used (Clontech, 631105). Transfections are done with Eugene HD (Roche). Luciferase assay reagents and instruments are the same as above. Restriction enzymes and T4 DNA ligase purchased from New England Biolabs.

Plasmids

miRNA expression vectors were purchased from Invitrogen (pcDNA 6.2-GW/EmGFP-miR, BLOCK-iT Pol II miR RNAi Expression vector Kits). This includes a non-targeting miRNA (pcDNA 6.2-GW/EmGFP-miR-neg), which is renamed pCMV-miR-non. pCMV-miR-luc was created by cloning purchased DNA oligo into the parent vector. The antisense target sequence of pCMV-miR-luc is TATTCAGCCCATATCGTTTCA, which targets the luciferase gene (luc+) found in Promega's pGL3 series of luciferase reporter vectors. pTRE-miR-luc is made by sub-cloning the gene from pCMV-miR-luc into pTREtight (Clontech) with SacI and XbaI restriction enzyme cut sites. pTRE-luc was purchased from Clontech. pSV40-luc, was made by sub-cloning the SV40 enhancer from pRLSV40 (Promega) into pGL3basic (Promega) with BglII and HindIII restriction enzyme cut sites. pCMV-tTA is from Clontech (pTet-Off-Advanced).

All plasmid changes were confirmed by sequencing.

Knockdown of firefly luciferase

12,000 HeLa cells plated in 96-well plate with 100 μ L DMEM + 10% FBS (tetracycline free) are transfected by addition of transfection mixture: 180 ng

pCMV-miRluc or pCMV-miR-non, 20 ng pSV40-luc, 0.8 μ L Fugene HD in 10 μ L OptiMEM. After 4 hours, media is replaced with 100 μ L DMEM + 10% FBS (tetracycline free). After 24 hours, luciferase activity is measured on Berthold luminometer with firefly substrate from Dual-Luciferase Assay Reporter System.

Doxycyclin dose-response on TRE promoter

12,000 HeLa cells plated in 96-well plate with 100 μ L DMEM + 10% FBS (tetracycline free) are transfected by addition of transfection mixture: 50 ng pTRE-luc, 50 ng pCMV-tTA, 1 ng pRLSV40, 0.4 μ L Fugene HD in 10 μ L OptiMEM. After 4 hours, media is replaced with 100 μ L DMEM + 10% FBS (tetracycline free) and doxycyclin is added. After 24 hours, luciferase activity is measured on Berthold luminometer with firefly substrate from Dual-Luciferase Assay Reporter System.

Doxycyclin dose-response on luciferase knockdown by TRE regulated miR-luc

12,000 HeLa cells plated in 96-well plate with 100 μ L DMEM + 10% FBS (tetracycline free) are transfected by addition of transfection mixture: 100 ng pSV40-luc, 50 ng pCMV-tTA, 50 ng pTRE-miR-luc or pTRE-miR-non, 1 ng pRLSV40, 0.8 μ L Fugene HD in 10 μ L OptiMEM. After 4 hours, media is replaced with 100 μ L DMEM + 10% FBS (tetracycline free) and doxycyclin is added. After 24 hours, luciferase activity is measured on Berthold luminometer with firefly substrate from Dual-Luciferase Assay Reporter System.

2.7.2 Materials and methods used in discovery and characterization of sekikaic acid

Protein expression

Non-labeled KIX protein was expressed and purified by Chinmay Y. Majmudar and Steven P. Rowe: An expression plasmid encoding His₆-tagged KIX domain from mouse CBP (residues 586-672) was described previously (pHis₆-PL-KIX).⁴² Protein was expressed with IPTG induction from pHis₆-PL-KIX transformed Rosetta pLysS cells (Novagen), purified using Ni-NTA beads (Qiagen) and buffer

exchanged on PD-10 columns (GE Healthcare). The protein is > 90% pure as determined by silver stained polyacrylamide gel. The protein is stored as frozen aliquots at 51 μM in 10 mM PBS pH 7.2, 10% glycerol, 0.01% NP-40. Protein concentration is determined by UV ($\epsilon=14440 \text{ M}^{-1}\text{cm}^{-1}$)

^{15}N -labeled KIX protein was expressed together with William C. Pomerantz: Rosetta pLysS transformed with pHis₆-PL-KIX are used to inoculate 3x50 mL LB starter cultures with 0.1mg/mL ampicillin and 0.034 mg/mL chloramphenicol. Next morning 4x25 mL starter culture is added to 4x1L LB with ampicillin, and bacteria are grown at 37 °C until an OD₆₀₀ of 0.67 is reached. Cells are pelleted and washed with M9 minimal media. Cells from 4x1L cultures are resuspended in 1L M9 minimal media with ampicillin and BioXpress (Cambridge Isotope Laboratories) added. Cells grown in shaker at 37 °C for 1 hr before temperature is reduced to 25 °C and 0.1M IPTG is added. After 8 hours cells are pelleted and frozen at -80 °C. Pellets from two such expressions are thawed on ice and resuspended in 70 mL 50 mM phosphate pH 7.1, 300 mM sodium chloride, 10 mM imidazole. Cells are lysed by sonication on ice. Lysates cleared by centrifugation at 9000xg, 20 min at 4 °C, and added to Ni-NTA beads. After two hours incubation at 4 °C, resin is pelleted with centrifugation at 2500xg, 2 min., 4 °C and washed with 30 mM imidazole containing buffer. Wash is repeated 5 times. Protein is eluted with 5 mL, 400 mM imidazole buffer three times, and the combined eluates buffer exchanged on PD-10 columns to 100 mM NaCl, 50 mM phosphate, 10 % glycerol. The buffer exchanged protein is further purified by cation exchange FPLC (Source 15S column, GE Healthcare; 0-1 M NaCl in 50 mM PBS buffer). FPLC purified protein is buffer exchanged on PD-10 columns to 10 mM phosphate pH 7.2, 100 mM NaCl and concentrated to 90-100 μM . Protein concentration is determined by UV ($\epsilon=14440 \text{ M}^{-1}\text{cm}^{-1}$).

KIX with ^{19}F -labeled tyrosine incorporated is expressed, purified and characterized by William C. Pomerantz.

Lab stocks of proteins Med15(1-347) and Gal4(1-100) were used. These have been previously reported.^{107,108}

Peptide synthesis

Peptides are synthesized on CLEAR amide resin (Peptides international) using standard HBTU/HOBT/DIEA coupling conditions.

FI-MLL19 was synthesized and purified by Steven P. Rowe: Fluorescein isothiocyanate and beta-alanine was coupled to a sequence containing MLL residues 2840-2858 to make FITC- β Ala-DCGNILPSDIMDPVLKNTP. The concentration of FI-MLL19 is determined by UV in 10mM Tris pH 9.0, using $\epsilon = 77,000 \text{ M}^{-1}\text{cm}^{-1}$ at 494 nM (The Handbook, A Guide to Fluorescent Probes and Labeling Technologies, Invitrogen).

MLL15 was synthesized and purified by William C. Pomerantz. The sequence contains MLL residues 2844-2857 in addition to a C-terminal tyrosine; the sequence is ILPSDIMDPVLKNTY. The concentration of MLL15 is determined by UV in 8M Urea, Tris pH 8.0 using $\epsilon = 1420 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nM.¹⁰⁹

MLL15dm was synthesized and purified by William C. Pomerantz. The sequence is ILPSDI(nLe)DPILKNTY.

FI-MLL15dm was synthesized by Ashley Lipinski. It was purified by Jonas W. Hojfeldt by reverse phase HPLC on C18 column with 0.01% trifluoroacetic acid in water and acetonitrile as eluents. A gradient of 30-55% acetonitrile over 25 min. was used. HPLC traces and mass spec. analysis are placed in section 2.8.

Phosphorylated KID29 was synthesized by William C. Pomerantz. It is the same construct reported by others.²³ It was purified by Jonas W. Hojfeldt on reverse phase HPLC with a 10-35 % CH₃CN over 23 min gradient. Spectra attached in section 2.8.

FI-cMyb, which encodes cMyb residues 291-315 was made and purified by Steven P. Rowe. This is the same construct reported by others.²³

FI-VP2 and FI-DNA are previously reported.^{107,108}

Fluorescent polarization assay – direct binding experiments

General: All samples are done in triplicate. The assays are done with a final sample volume of 10 μ L in a low-volume, non-binding, black, 384-well plate (Corning), and read on Tecan Genios Pro with polarized excitation at 485 nm and emission intensity measured through a parallel, I(par), and perpendicular, I(per) polarized 535 nm filter. The relative intensities of these filters have been calibrated with a G-factor¹¹⁰ and the anisotropy is then defined as $A = [I(\text{par}) - I(\text{per})] / [I(\text{par}) + 2 * I(\text{per})]$. The term $I(\text{par}) + 2 * I(\text{per})$ is directly proportional to the total fluorescence intensity of the sample. The data is plotted in Graphpad Prism 5 and fitted with nonlinear regression using built-in equation “One site – Total and nonspecific binding”, from which the Kd is calculated in Prism.

Direct binding of KIX to FI-MLL19: His₆-KIX(586-672) protein is added at 25 μ M to a solution with 25 nM FI-MLL19 and assay buffer (10mM PBS pH 7.4, 10% glycerol, 0.01% NonidetP-40, 1mM DTT). This solution is serially diluted into a solution of 25nM FI-MLL19 in assay buffer.

Direct binding of KIX to FI-MLL15dm and FI-KID29 is done as described for FI-MLL19.

Fluorescent polarization assay – competitive binding experiments

General: Samples are done in triplicate, and the assay is read as in direct binding experiments. The data is plotted in Graphpad Prism 5 (Prism) and fitted with nonlinear regression using built-in equation “log(inhibitor) vs. response – Variable slope (four parameters)” from which the IC50 is calculated in Prism. Inhibition

curves with Hillslope near 1 could be fitted with theoretical equation⁸⁸ to determine a Ki. This equation was entered into Prism as:

$$\begin{aligned}
 d &= Kd + Ki + 10^X + T - P \\
 e &= (10^X - P) * Kd + (T - P) * Ki + Kd * Ki \\
 f &= -Kd * Ki * P \\
 \text{Root} &= (d^2 - 3 * e)^{0.5} \\
 \text{theta} &= 81 \text{rcos}((-2 * d^3 + 9 * d * e - 27 * f) / (2 * \text{Root}^3)) \\
 C &= \cos(\text{theta} / 3) \\
 \text{Top} &= (2 * \text{Root} * C - d) \\
 \text{Fsb} &= \text{Top} / (3 * Kd + \text{Top}) \\
 Y &= \text{Fsb} * (Ab - Af) + Af
 \end{aligned}$$

Kd: Binding affinity of tracer; Ki: Binding affinity of inhibitor; X: Log(concentration of inhibitor); T: Concentration of tracer; P: Concentration of protein; Ab: Anisotropy of bound tracer; Af: Anisotropy of free tracer.

Ki, Af and Ab are the fitted parameters.

Theoretical detection limit: A theoretical curve is plotted in Prism, in which the fraction of bound tracer is defined as a function of Ki (X).⁸⁸ Equation is entered in Prism as:

$$\begin{aligned}
 d &= Kd + 10^X + I + T - P \\
 e &= (I - P) * Kd + (T - P) * 10^X + Kd * 10^X \\
 f &= -Kd * 10^X * P \\
 \text{Root} &= (d^2 - 3 * e)^{0.5} \\
 \text{theta} &= 81 \text{rcos}((-2 * d^3 + 9 * d * e - 27 * f) / (2 * \text{Root}^3)) \\
 C &= \cos(\text{theta} / 3) \\
 \text{Top} &= (2 * \text{Root} * C - d) \\
 \text{Fsb} &= \text{Top} / (3 * Kd + \text{Top})
 \end{aligned}$$

Competitive inhibition of FI-MLL19 – KIX with MLL15: MLL15 at 30 μ M in a solution of 25 nM FI-MLL19 and 5 μ M KIX in assay buffer (10mM PBS pH 7.4, 10% glycerol, 0.01% NonidetP-40, 1mM DTT) is serially diluted into wells with same components except MLL15.

Determination of Z'-factor: This was done on equipment at the Center for Chemical Genomics, University of Michigan. To columns 1-22 of a low volume, NBS, black, 284-well Corning plates (#3676) is added 5 μ L of KIX at 20 μ M in assay buffer (10mM PBS pH 7.4, 10% glycerol, 0.01% NonidetP-40, 1mM DTT), and to columns 23-24 is added 5 μ L assay buffer with multi-channel, automated liquid handlers. DMSO (200 nL, 2% final) is added with pin-tool, followed by addition of 5 μ L, 50 nM FI-MLL19 in assay buffer with liquid handlers. Samples are incubated for 30 min before reading on PheraStar Plus (BMG Labtech) with FI-FP module (485nm/520nm). The performance was improved by addressing two sources of error: (1) The tubing of the liquid handlers used to fill plates with assay reagents, bound one or more assay components and had to be equilibrated prior to plating by passing assay reagent through. (2) The largest source of variability turned out to be a mis-alignment in the plate-reader, which was apparent from a systematic drift pattern of data from control plates, and this was corrected by repair of the instrument.

High-throughput screen: The FI-MLL19 – KIX inhibition assay was set up as during determination of Z'factor. Compounds are added in the pin tool step. Initial hits were analyzed with consultation of Paul D. Kirchhoff (Vahlteich Medicinal Chemistry Core, University of Michigan) and Martha Larsen (Center for Chemical Genomics, University of Michigan). The confirmation tests were done in an assay with identical final components, but with a different order of addition. For this, the tracers are added first, followed by compounds, and the plates are read once before addition of protein. This read is used to determine if any of the compounds quenches the fluorescence of the tracer. The counter screen used to determine selectivity of natural product extracts is an assay developed by Jolanta

Grembecka's lab (University of Michigan). Details of their assay are not presented here.

Specificity tests with lichen extracts from Costa Rica: 13 extracts and fractions received from Costa Rica were tested in the FL-MLL19 – KIX FP assay as described above and three other FP-assays. For these, FI-cMyb and FI-VP2 were used at 25 nM and FI-DNA was used at 10 nM. KIX and Med15 were used at 10 μ M and Gal4 was used at 50 nM.

Competition studies with sekikaic acid (and derivatives): Inhibition of FL-MLL19 – KIX FP assay studied as above. In experiments with varying detergents the detergents were added to detergent free assay buffer at defined concentrations. The NonidetP-40 (NP-40) present in stocks of KIX was not removed. FI-MLL15 – KIX inhibition curves were done with KIX at 1 μ M, and no added detergent. FI-KID29 – KIX inhibition curves were done with KIX at 2.5 μ M.

Compounds purchased for comparative competitive binding studies: 2-naphthol-AS-E-phosphate (Sigma-Aldrich), lecanoric acid(ChromaDex), lobaric acid(ChromaDex).

Identification of depsides

HPLC purification and compound identification was done by Carl J. Arevang (Sherman lab). HPLC fractions were tested in FP assays for activity and only active fractions were characterized. HPLC traces and NMR spectra are attached in section 2.7. A UV spectrum of sekikaic acid was recorded, and $\epsilon_{260\text{nm}} = 12196 \text{ M}^{-1}\text{cm}^{-1}$ was measured and used for determination of concentrations.

1D-¹H-NMR studies of sekikaic acid in presence of KIX and inhibitors

Spectra were recorded on a Bruker Avance III 600-MHz spectrometer equipped with cryogenic probe with advice from Tomek Cierpicki (University of Michigan). The spectra are properly referenced as determined by identical shift of d5-DMSO

peak. Stocks of sekikaic acid and KID29 used, were made in d6-DMSO, and MLL15dm was used as a non-deuterated DMSO stock.

¹⁵N-¹H-HSQC NMR studies of KIX in presence of sekikaic acid

Spectra were recorded on a Bruker Avance III 600-MHz spectrometer equipped with cryogenic probe with help from Tomek Cierpicki (University of Michigan).

¹D-¹⁵F-NMR studies of ¹⁵F-labeled KIX in presence of sekikaic acid

Collection and assignment of spectra was done by William C. Pomerantz.

Viability studies

3000 A549 cells or 4000 HEK293T cells were plated in 96-well plate with phenolred-free DMEM. Compounds are added from a DMSO stock (final conc. of DMSO is 1%). After 24 or 48 hours, 10 μ L WST-1 (Roche) is added to media. After 45 min. incubation, the absorbance at 440nm is read with Tecan Genios Pro. Background (absorbance from wells with no cells, but with media and WST-1 added) is subtracted from all values. % viability is determined as absorbance(sample)/absorbance(DMSO).

Plasmids for luciferase reporter studies

pGal4(1-147)-MLL(2829-2883) was made by William C. Pomerantz by same method used for pGal4-Elf3 described above. pGal4-CREB and pPKA were purchase from Stratagene (Product: Pathdetect HLR cell line). pG5luc (Gal4-luciferase reporter) and pRLSV40 (constitutive Renilla luciferase control) are purchased from Promega.

For two-hybrid constructs, commercial vectors pAct (VP16(411-456)-fusion vector) and pBind (Gal4(1-147)-fusion vector) (Promega) were modified to allow ligation-independent cloning.¹¹¹ The following oligos were annealed and ligated to EcoRI and BamH1 digested pAct and pBind: 5'- GAT CTT GGG AAG CAC CGG TTC TGG TGA -3' and 5'-CTA GTC ACC AGA ACC GGT GCT TCC CAA -

3'. The resulting vectors are termed pVP16-LIC and pGal4-LIC. KIX (CBP(586-672) and MLL(411-456) were cloned out of vectors previously described with PCR, using primers with the following sequences appended to the 5' end of each primer to facilitate LIC cloning: (sense strand primer – LIC appendage) 5' - GGG AAG CAC CGG T ; (antisense strand primer – LIC appendage) 5' CAC CAG AAC CGG T. The PCR products and linearized LIC vectors are processed with T4 DNA polymerase (Novagen), combined, and used to transform *E. coli* cells. The resulting vectors are called pGal4-KIX, pVP16-KIX, pGal4-MLL, pVP16-MLL.

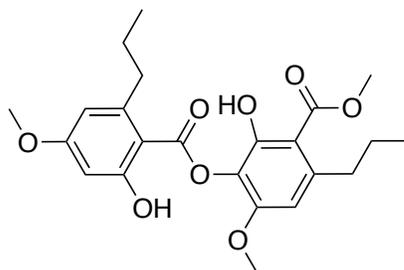
Gal4-MLL and Gal4-CREB luciferase reporter studies

20,000 HEK293T cells are plated in 96 well plate and transfected the following day by replacing media with 50 ng pG5luc, 50 ng pGal4-MLL, 1.5 ng pRLSV40, 0.3 μ L Lipofectamine 2000, 100 μ L OptiMEM (Gal4-MLL) or 50 ng pG5luc, 25 ng pGal4-CREB, 25 ng pPKA, 1.5 ng pRLSV40, 0.3 μ L Lipofectamine 2000, 100 μ L OptiMEM (Gal4-MLL). After 4 hours transfection mix is replaced with normal media, and compounds are added. After 24 hours, assay is read with Dual Luciferase Assay Reporter-System (Promega).

KIX:MLL two-hybrid assay

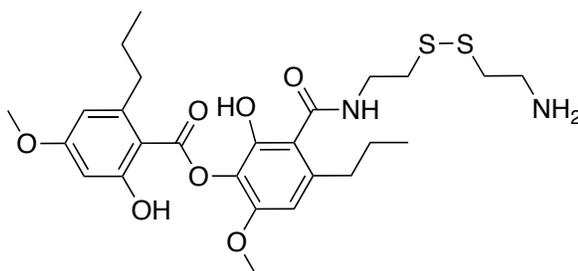
12,000 HeLa cells are plated in wells of 96-well plate. Transfection was carried out by addition of 20 μ L of transfection mix with (in OptiMEM): 30 ng pG5luc, 10 ng pGal4-KIX/MLL/LIC, 10 ng pVP16-KIX/MLL/LIC, 0.2 μ L Lipectamine LTX. 24 hours post transfection, media was replaced with regular media and compounds added. After 24 hours, luciferase activities were determined with the Dual Luciferase Assay Reporter-System (Promega).

Synthesis of sekikaic acid derivatives



Sek-OMe (methyl 2-hydroxy-3-((2-hydroxy-4-methoxy-6-propylbenzoyl)oxy)-4-methoxy-6-propylbenzoate):

To sekikaic acid (1mg, 2.4 μmol) dissolved in 300 μL toluene and 200 μL methanol, is added a 2M ethereal solution of trimethylsilyldiazomethane (3 μL , 6 μmol). Reaction monitored by persistence of yellow color. After 30 minutes of stirring at room temperature, reaction is concentrated, dissolved in acetonitrile and 0.1% aqueous TFA, filtered, and purified by reverse phase HPLC with elution gradient 5-75 % acetonitrile over 30 min. The product is confirmed by ESI-MS (observed m/z ($M+\text{Na}^+$): 455.1), and by $^1\text{H-NMR}$. HPLC chromatogram and $^1\text{H-NMR}$ spectrum are attached in section 2.8.



Sek-SSEA (3-((2-((2-aminoethyl)disulfanyl)ethyl)carbamoyl)-2-hydroxy-6-methoxy-4-propylphenyl 2-hydroxy-4-methoxy-6-propylbenzoate):

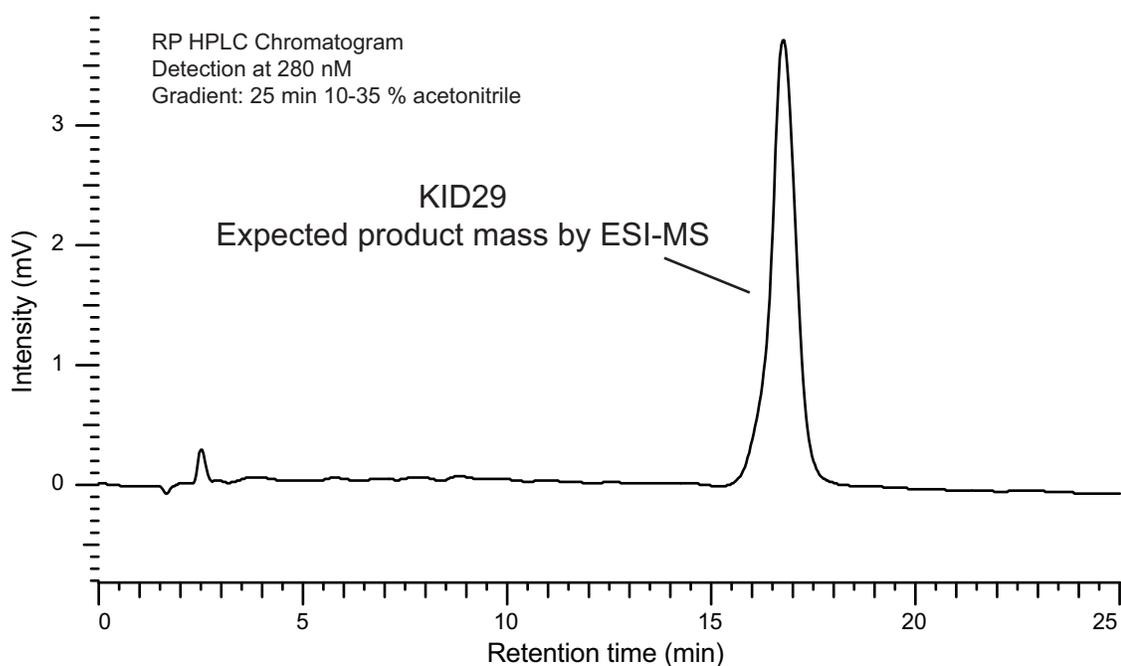
Cystamine dihydrochloride is dissolved in 1M NaOH and extracted with dichloromethane. Dichloromethane solution is passed through Pasteur pipette packed with Na_2SO_4 , concentrated and weighed. Cystamine (4.5 mg, 30 μmol) dissolved in 150 μL dichloromethane is added to a solution of sekikaic acid (2mg, 4.8 μmol) in 150 μL dichloromethane. Distilled trimethylamine (4.2 μL , 30 μmol) and PyBroP (2.5 mg, 5.2 μmol) are added. After 3 hours starting material is consumed and crude product is purified on reverse phase HPLC with elution

gradient 20-50 % acetonitrile over 25 min. The purified product is a mix of disulfide and reduced species, with the major component being the disulfide. These two species can both participate in reversible disulfide exchange reactions, and were not separated for initial studies. The product is confirmed by ESI-MS (observed m/z ($M+H^+$): 553.3), and by 1H -NMR. HPLC chromatogram and 1H -NMR spectrum are attached in section 2.8.

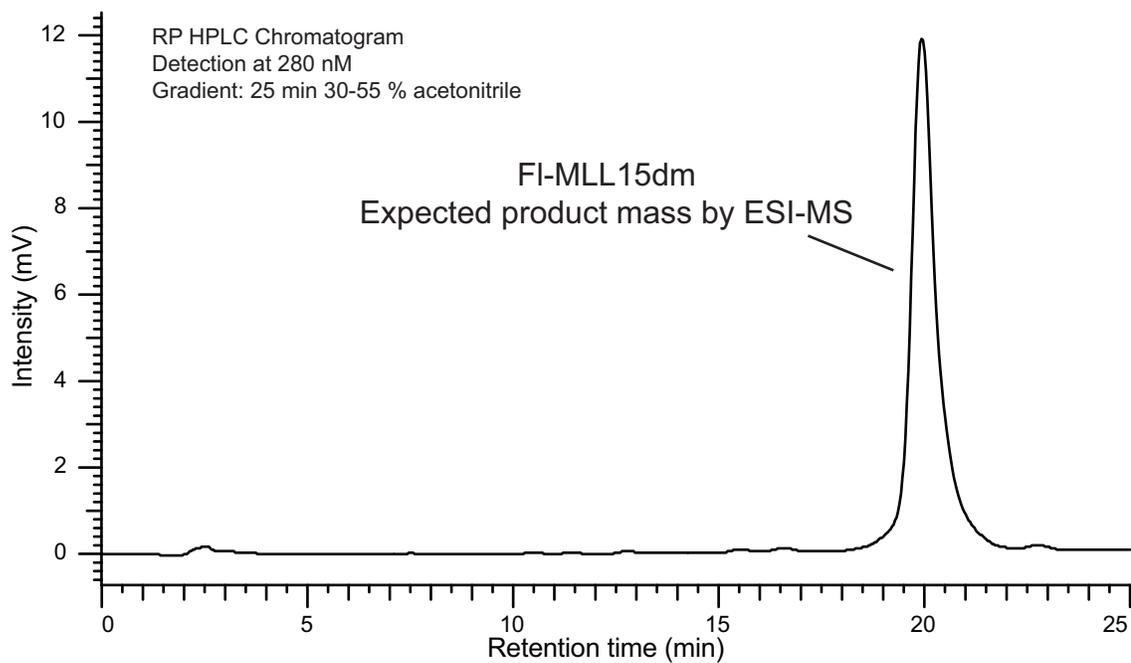
2.8 Characterization of compounds made or purified by Jonas W. Hojfeldt

HPLC chromatograms

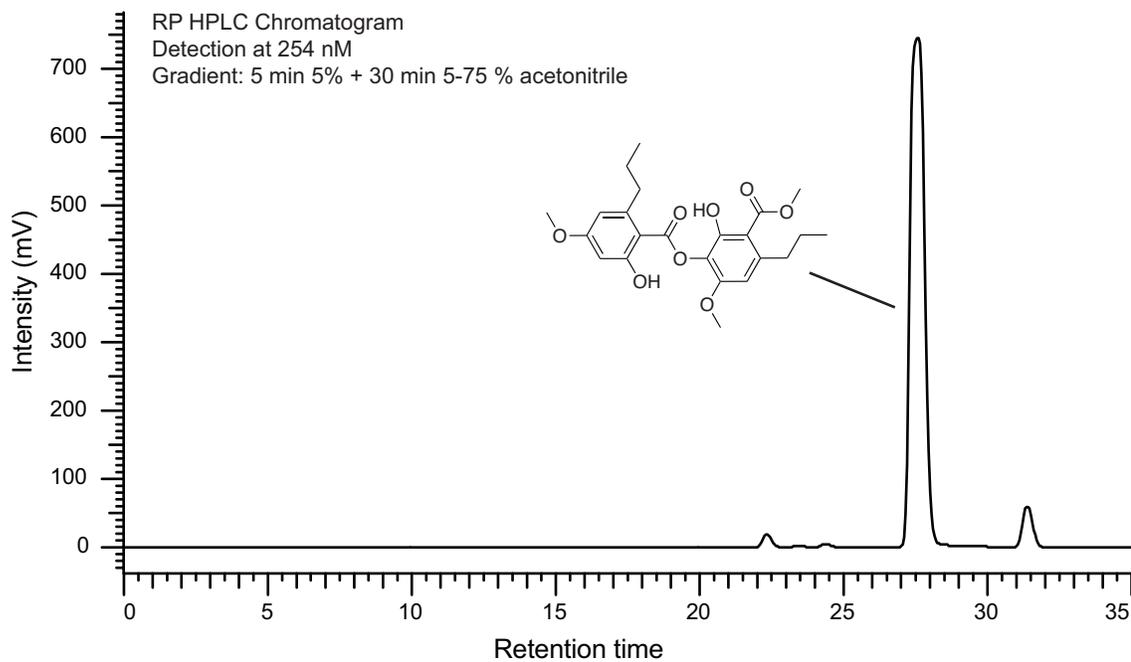
Analytical HPLC chromatogram of purified KID29

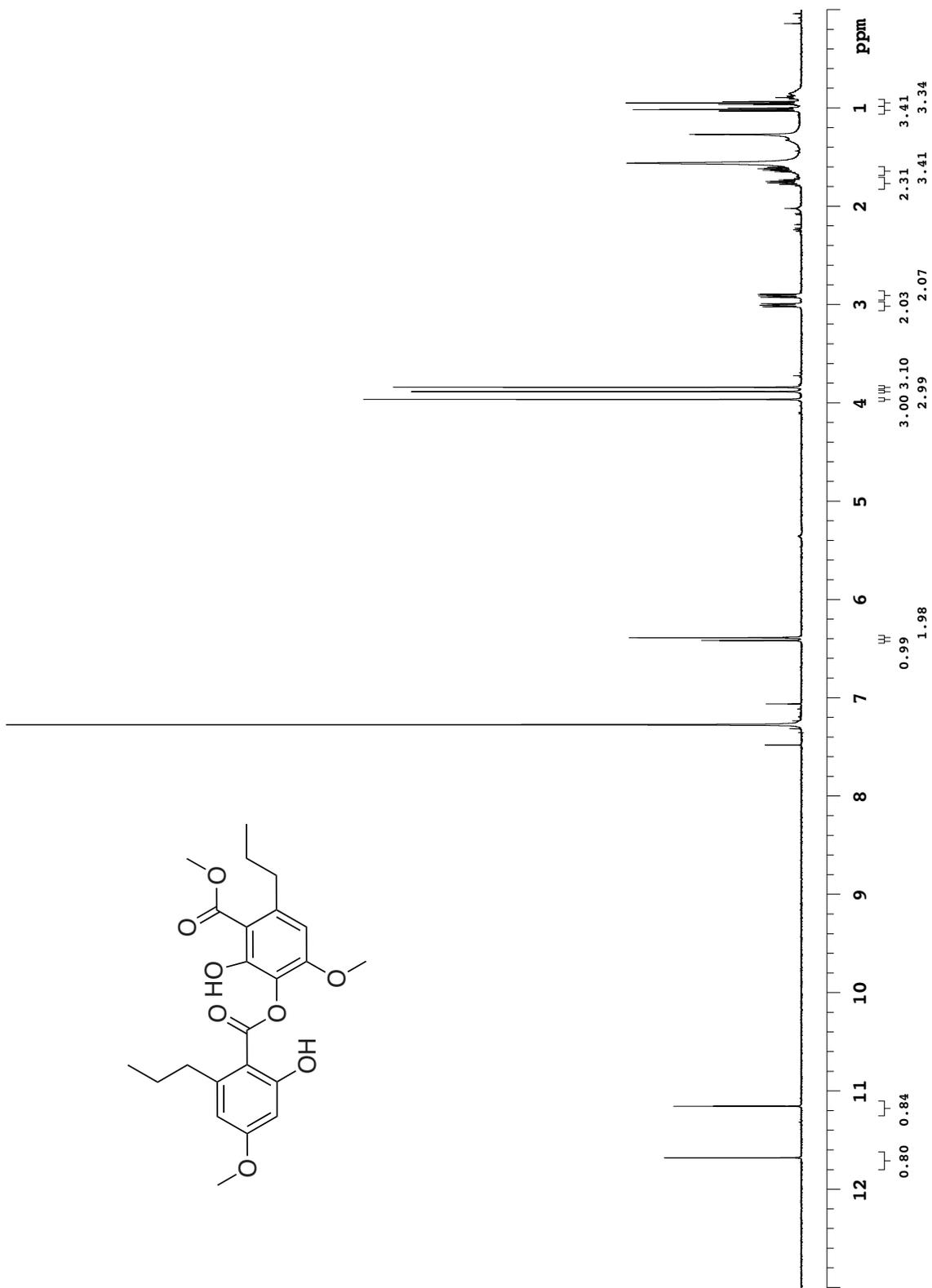
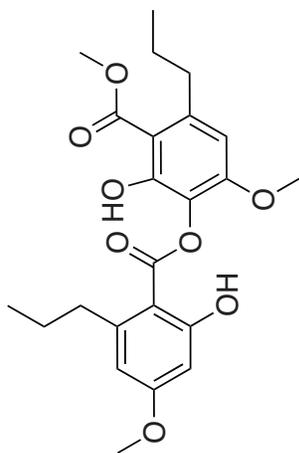


Analytical HPLC chromatogram of purified FI-MLL15dm

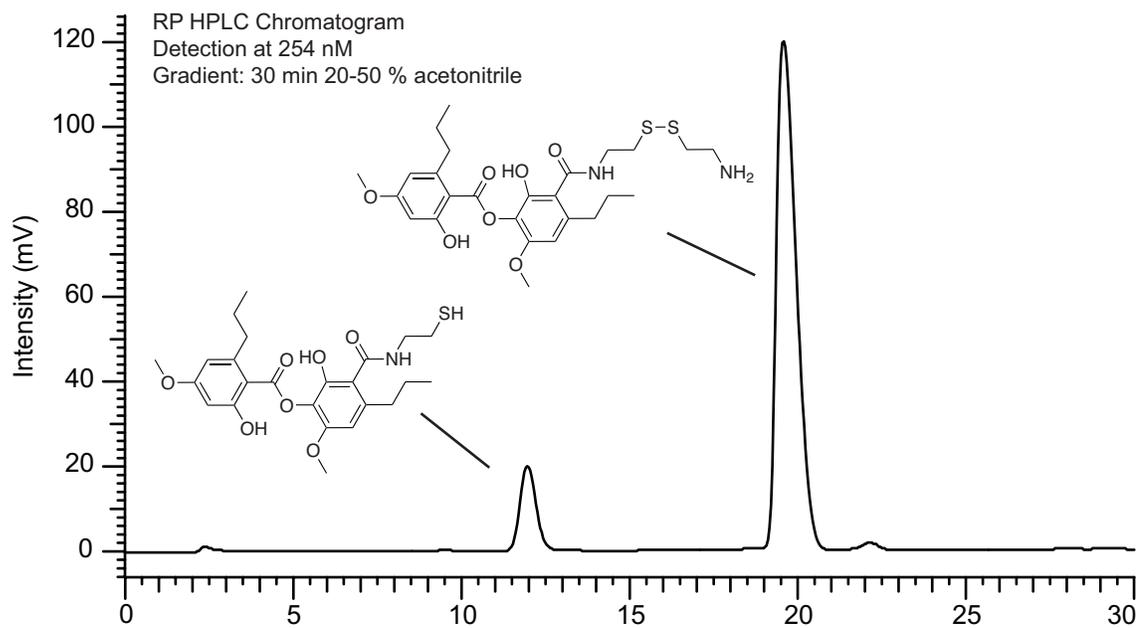


Analytical HPLC chromatogram and ¹H-NMR of purified Sek-OME





Analytical HPLC chromatogram and ¹H-NMR of purified Sek-SSEA



2.9 References

1. Stockwell, B.R. Chemical genetics: ligand-based discovery of gene function. *Nat Rev Genet* **1**, 116-25 (2000).
2. Lee, L.W. & Mapp, A.K. Transcriptional switches: chemical approaches to gene regulation. *J Biol Chem* **285**, 11033-8 (2010).
3. Ansari, A.Z. & Mapp, A.K. Modular design of artificial transcription factors. *Curr Opin Chem Biol* **6**, 765-72 (2002).
4. Rodriguez-Martinez, J.A., Peterson-Kaufman, K.J. & Ansari, A.Z. Small-molecule regulators that mimic transcription factors. *Biochim Biophys Acta* **1799**, 768-74 (2010).
5. Wu, Z. et al. Targeting the transcriptional machinery with unique artificial transcriptional activators. *J Am Chem Soc* **125**, 12390-1 (2003).
6. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nat Rev Drug Discov* **1**, 727-30 (2002).
7. Cochran, A.G. Antagonists of protein-protein interactions. *Chem Biol* **7**, R85-94 (2000).
8. Toogood, P.L. Inhibition of protein-protein association by small molecules: approaches and progress. *J Med Chem* **45**, 1543-58 (2002).
9. Arkin, M.R. & Wells, J.A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* **3**, 301-17 (2004).
10. Zinzalla, G. & Thurston, D.E. Targeting protein-protein interactions for therapeutic intervention: a challenge for the future. *Future Med Chem* **1**, 65-93 (2009).
11. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *J Mol Biol* **285**, 2177-98 (1999).
12. DeLano, W.L., Ultsch, M.H., de Vos, A.M. & Wells, J.A. Convergent solutions to binding at a protein-protein interface. *Science* **287**, 1279-83 (2000).
13. Moreira, I.S., Fernandes, P.A. & Ramos, M.J. Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **68**, 803-12 (2007).

14. Darimont, B.D. et al. Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev* **12**, 3343-56 (1998).
15. Shiau, A.K. et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927-37 (1998).
16. Nolte, R.T. et al. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* **395**, 137-43 (1998).
17. Warnmark, A. et al. Interaction of transcriptional intermediary factor 2 nuclear receptor box peptides with the coactivator binding site of estrogen receptor alpha. *J Biol Chem* **277**, 21862-8 (2002).
18. Hope, I.A. & Struhl, K. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885-94 (1986).
19. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847-53 (1987).
20. Ma, J. & Ptashne, M. A new class of yeast transcriptional activators. *Cell* **51**, 113-9 (1987).
21. 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**, 741-52 (1997).
22. Demarest, S.J. et al. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-53 (2002).
23. 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).
24. Mujtaba, S. et al. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell* **13**, 251-63 (2004).
25. Waters, L. et al. Structural diversity in p160/CREB-binding protein coactivator complexes. *J Biol Chem* **281**, 14787-95 (2006).
26. Lee, C.W., Martinez-Yamout, M.A., Dyson, H.J. & Wright, P.E. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* **49**, 9964-71 (2010).

27. 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**, 319-29 (2004).
28. 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).
29. 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).
30. Minter, A.R., Brennan, B.B. & Mapp, A.K. A small molecule transcriptional activation domain. *J Am Chem Soc* **126**, 10504-5 (2004).
31. Buhrlage, S.J., Brennan, B.B., Minter, A.R. & Mapp, A.K. Stereochemical promiscuity in artificial transcriptional activators. *J Am Chem Soc* **127**, 12456-7 (2005).
32. Casey, R.J., Desaulniers, J.P., Højfeldt, J.W. & Mapp, A.K. Expanding the repertoire of small molecule transcriptional activation domains. *Bioorg Med Chem* **17**, 1034-43 (2009).
33. 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).
34. Kung, A.L. et al. Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell* **6**, 33-43 (2004).
35. 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).
36. Emami, K.H. et al. A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. *Proc Natl Acad Sci U S A* **101**, 12682-7 (2004).
37. Li, B.X. & Xiao, X. Discovery of a small-molecule inhibitor of the KIX-KID interaction. *Chembiochem* **10**, 2721-4 (2009).
38. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067-73 (2010).
39. Lee, L.W. et al. Inhibition of ErbB2(Her2) expression with small molecule transcription factor mimics. *Bioorg Med Chem Lett* **19**, 6233-6 (2009).

40. Bates, C.A., Pomerantz, W.C. & Mapp, A.K. Transcriptional tools: Small molecules for modulating CBP KIX-dependent transcriptional activators. *Biopolymers* **95**, 17-23 (2011).
41. Rowe, S.P., Casey, R.J., Brennan, B.B., Buhrlage, S.J. & Mapp, A.K. Transcriptional up-regulation in cells mediated by a small molecule. *J Am Chem Soc* **129**, 10654-5 (2007).
42. Buhrlage, S.J. et al. Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem Biol* **4**, 335-44 (2009).
43. Cook, K.M. et al. Epidithiodiketopiperazines block the interaction between hypoxia-inducible factor-1alpha (HIF-1alpha) and p300 by a zinc ejection mechanism. *J Biol Chem* **284**, 26831-8 (2009).
44. Yano, K. et al. Chetomin induces degradation of XIAP and enhances TRAIL sensitivity in urogenital cancer cells. *Int J Oncol* **38**, 365-74 (2011).
45. Cousens, D.J., Greaves, R., Goding, C.R. & O'Hare, P. The C-terminal 79 amino acids of the herpes simplex virus regulatory protein, Vmw65, efficiently activate transcription in yeast and mammalian cells in chimeric DNA-binding proteins. *EMBO J* **8**, 2337-42 (1989).
46. 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**, 718-29 (1988).
47. Reeves, W.M. & Hahn, S. Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* **25**, 9092-102 (2005).
48. Majmudar, C.Y. et al. Impact of nonnatural amino acid mutagenesis on the in vivo function and binding modes of a transcriptional activator. *J Am Chem Soc* **131**, 14240-2 (2009).
49. Herrera, F.J. & Triezenberg, S.J. VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection. *J Virol* **78**, 9689-96 (2004).
50. Owen, H.R., Quadroni, M., Bienvenut, W., Buerki, C. & Hottiger, M.O. Identification of novel and cell type enriched cofactors of the transcription activation domain of RelA (p65 NF-kappaB). *J Proteome Res* **4**, 1381-90 (2005).
51. Gerritsen, M.E. et al. CREB-binding protein/p300 are transcriptional coactivators of p65. *Proc Natl Acad Sci U S A* **94**, 2927-32 (1997).

52. Perkins, N.D. et al. Regulation of NF-kappaB by cyclin-dependent kinases associated with the p300 coactivator. *Science* **275**, 523-7 (1997).
53. Karin, M., Cao, Y., Greten, F.R. & Li, Z.W. NF-kappaB in cancer: from innocent bystander to major culprit. *Nat Rev Cancer* **2**, 301-10 (2002).
54. Moore, P.A., Ruben, S.M. & Rosen, C.A. Conservation of transcriptional activation functions of the NF-kappa B p50 and p65 subunits in mammalian cells and *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**, 1666-74 (1993).
55. Blair, W.S., Bogerd, H.P., Madore, S.J. & Cullen, B.R. Mutational analysis of the transcription activation domain of RelA: identification of a highly synergistic minimal acidic activation module. *Mol Cell Biol* **14**, 7226-34 (1994).
56. Sakurai, H., Chiba, H., Miyoshi, H., Sugita, T. & Toriumi, W. IkappaB kinases phosphorylate NF-kappaB p65 subunit on serine 536 in the transactivation domain. *J Biol Chem* **274**, 30353-6 (1999).
57. Wang, D. & Baldwin, A.S., Jr. Activation of nuclear factor-kappaB-dependent transcription by tumor necrosis factor-alpha is mediated through phosphorylation of RelA/p65 on serine 529. *J Biol Chem* **273**, 29411-6 (1998).
58. Chang, C.H., Scott, G.K., Baldwin, M.A. & Benz, C.C. Exon 4-encoded acidic domain in the epithelium-restricted Ets factor, ESX, confers potent transactivating capacity and binds to TATA-binding protein (TBP). *Oncogene* **18**, 3682-95 (1999).
59. Chang, C.H. et al. ESX: a structurally unique Ets overexpressed early during human breast tumorigenesis. *Oncogene* **14**, 1617-22 (1997).
60. Tymms, M.J. et al. A novel epithelial-expressed ETS gene, ELF3: human and murine cDNA sequences, murine genomic organization, human mapping to 1q32.2 and expression in tissues and cancer. *Oncogene* **15**, 2449-62 (1997).
61. Andreoli, J.M. et al. The expression of a novel, epithelium-specific ets transcription factor is restricted to the most differentiated layers in the epidermis. *Nucleic Acids Res* **25**, 4287-95 (1997).
62. Asada, S. et al. External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc Natl Acad Sci U S A* **99**, 12747-52 (2002).

63. Zhang, J.H., Chung, T.D. & Oldenburg, K.R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* **4**, 67-73 (1999).
64. Ambros, V. The functions of animal microRNAs. *Nature* **431**, 350-5 (2004).
65. Zeng, Y., Yi, R. & Cullen, B.R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J* **24**, 138-48 (2005).
66. McManus, M.T. & Sharp, P.A. Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* **3**, 737-47 (2002).
67. Cullen, B.R. Transcription and processing of human microRNA precursors. *Mol Cell* **16**, 861-5 (2004).
68. Chung, K.H. et al. Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Res* **34**, e53 (2006).
69. Zeng, Y., Wagner, E.J. & Cullen, B.R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* **9**, 1327-33 (2002).
70. Lagos-Quintana, M. et al. Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12**, 735-9 (2002).
71. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-51 (1992).
72. Resnitzky, D., Gossen, M., Bujard, H. & Reed, S.I. Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. *Mol Cell Biol* **14**, 1669-79 (1994).
73. Xu, W., Kasper, L.H., Lerach, S., Jeevan, T. & Brindle, P.K. Individual CREB-target genes dictate usage of distinct cAMP-responsive coactivation mechanisms. *EMBO J* **26**, 2890-903 (2007).
74. Bedford, D.C., Kasper, L.H., Fukuyama, T. & Brindle, P.K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9-15 (2010).
75. Dyson, H.J. & Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208 (2005).

76. Yao, T.P. et al. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-72 (1998).
77. Tanaka, Y. et al. Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. *Mech Dev* **95**, 133-45 (2000).
78. Kasper, L.H. et al. CBP/p300 double null cells reveal effect of coactivator level and diversity on CREB transactivation. *EMBO J* **29**, 3660-72 (2010).
79. Kasper, L.H. et al. A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. *Nature* **419**, 738-43 (2002).
80. Kasper, L.H. et al. Two transactivation mechanisms cooperate for the bulk of HIF-1-responsive gene expression. *EMBO J* **24**, 3846-58 (2005).
81. Wood, M.A., Attner, M.A., Oliveira, A.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).
82. Bowers, E.M. et al. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol* **17**, 471-82 (2010).
83. 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**, 2249-58 (2001).
84. Campbell, K.M. & Lumb, K.J. Structurally distinct modes of recognition of the KIX domain of CBP by Jun and CREB. *Biochemistry* **41**, 13956-64 (2002).
85. 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**, 2115-24 (2009).
86. Wang, F. et al. Synergistic interplay between promoter recognition and CBP/p300 coactivator recruitment by FOXO3a. *ACS Chem Biol* **4**, 1017-27 (2009).
87. Jameson, D.M. & Seifried, S.E. Quantification of protein-protein interactions using fluorescence polarization. *Methods* **19**, 222-33 (1999).
88. Roehrl, M.H., Wang, J.Y. & Wagner, G. A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* **43**, 16056-66 (2004).

89. 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* **584**, 4500-4 (2010).
90. Fischer, E. & Freudenberg, K. The carbomethoxy derivative of the phenol carboxylic acids and their application for synthesis IV. *Liebigs Ann* **372**, 32-68 (1910).
91. McGovern, S.L., Caselli, E., Grigorieff, N. & Shoichet, B.K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem* **45**, 1712-22 (2002).
92. McGovern, S.L., Helfand, B.T., Feng, B. & Shoichet, B.K. A specific mechanism of nonspecific inhibition. *J Med Chem* **46**, 4265-72 (2003).
93. Straus, O.H. & Goldstein, A. Zone Behavior of Enzymes : Illustrated by the Effect of Dissociation Constant and Dilution on the System Cholinesterase-Physostigmine. *J Gen Physiol* **26**, 559-85 (1943).
94. Shoichet, B.K. Interpreting steep dose-response curves in early inhibitor discovery. *J Med Chem* **49**, 7274-7 (2006).
95. Kenakin, T.P. *Pharmacologic analysis of drug-receptor interaction*, xii, 483 p. (Raven, New York, 1993).
96. Lepre, C.A., Moore, J.M. & Peng, J.W. Theory and applications of NMR-based screening in pharmaceutical research. *Chem Rev* **104**, 3641-76 (2004).
97. Radhakrishnan, I. et al. Structural analyses of CREB-CBP transcriptional activator-coactivator complexes by NMR spectroscopy: implications for mapping the boundaries of structural domains. *J Mol Biol* **287**, 859-65 (1999).
98. Chrivia, J.C. et al. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855-9 (1993).
99. Narui, T. et al. NMR assignments of depsides and tridepsides of the lichen family umbilicariaceae. *Phytochemistry* **48**, 815-822 (1998).
100. Erlanson, D.A., Wells, J.A. & Braisted, A.C. Tethering: fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* **33**, 199-223 (2004).
101. Blair, J.A. et al. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat Chem Biol* **3**, 229-38 (2007).

102. Armaleo, D., Zhang, Y. & Cheung, S. Light might regulate divergently deposite and deposite accumulation in the lichen *Parmotrema hypotropum* by affecting thallus temperature and water potential. *Mycologia* **100**, 565-76 (2008).
103. Ronn, R. et al. Exploration of acyl sulfonamides as carboxylic acid replacements in protease inhibitors of the hepatitis C virus full-length NS3. *Bioorg Med Chem* **14**, 544-59 (2006).
104. Sadowski, I., Bell, B., Broad, P. & Hollis, M. GAL4 fusion vectors for expression in yeast or mammalian cells. *Gene* **118**, 137-41 (1992).
105. Liu, B., Alluri, P.G., Yu, P. & Kodadek, T. A potent transactivation domain mimic with activity in living cells. *J Am Chem Soc* **127**, 8254-5 (2005).
106. Shimogawa, H. et al. A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction. *J Am Chem Soc* **126**, 3461-71 (2004).
107. Majmudar, C.Y., Labut, A.E. & Mapp, A.K. Tra1 as a screening target for transcriptional activation domain discovery. *Bioorg Med Chem Lett* **19**, 3733-5 (2009).
108. Wands, A.M. et al. Transient-state kinetic analysis of transcriptional activator-DNA complexes interacting with a key coactivator. *J Biol Chem* **286**, 16238-45 (2011).
109. Creighton, T.E. *Proteins: Structures and Molecular Properties*, (W.H. Freeman and Company, New York, 1993).
110. Tetin, S.Y. & Hazlett, T.L. Optical spectroscopy in studies of antibody-hapten interactions. *Methods* **20**, 341-61 (2000).
111. Aslanidis, C. & de Jong, P.J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069-74 (1990).

Chapter 3

Extrinsic recruitment of transcriptional coregulators by bifunctional nuclear receptor ligands*

3.1 Abstract

Nuclear receptors (NRs) are ligand-regulated transcription factors. They offer a means to directly affect the transcription of their target genes with small molecules without simultaneous involvement of signaling cascades, and this property underlies their success as drug targets. Transcriptional control is achieved through recruitment of transcriptional coregulators. The conformation of the receptor, which is modulated by ligand binding, determines which cofactors are recruited. Synthetic ligands have been discovered that can stabilize either active or repressive conformations and bind selectively to receptor subtypes, and these ligands consequently induce gene expression patterns that differ from the natural ligands and can have improved therapeutic effects and tissue specificities. The current strategies for ligand discovery, however, limit the transcriptional control to that defined by patterns of coregulator recruitment intrinsic to the available receptor conformations. This is particularly problematic in diseases where receptor function becomes insensitive to the ligand through mutation or dominant aberrant post-translational modifications. We have shown that bifunctional molecules, made as conjugates of NR ligands to ligands of

* This chapter presents research from a collaborative effort. J.W. Hojfeldt, A.K. Mapp and J.A. Iñiguez-Lluhí planned experiments. J.W. Hojfeldt, A.R. Van Dyke, and Y. Imaeda synthesized compounds. Data presented in Figures 3.9, 3.10, 3.11, and 3.12 was obtained by O. Cruz and J.A. Iñiguez-Lluhí. All other data figures are from experiments done by J.W. Hojfeldt. Y. Imaeda contributed with experiments presented in Figures 3.5, 3.6, and 3.7. A.R. Van Dyke contributed to experiment presented in Figure 3.15. J.P. Carolan assisted with experiments presented in Figure 3.13 and 3.14.

transcriptional components that are non-native to NRs, are able to recruit these non-native targets to the NR and affect transcription. This strategy holds promise to greatly expand the functional control that can be achieved by the important class of NR-targeting drugs.

3.2 Introduction to nuclear receptors as ligand-regulated factors

Nuclear receptors (NRs) are members of a large family of multi-domain¹ transcription factors, characterized by a conserved DNA-binding domain (DBD) and a moderately conserved ligand-binding domain (LBD).² The DNA-binding domain defines the genomic sites that the receptors can bind to, while the receptor's influence on transcription of target genes is modulated by integration of multiple signals by the receptor, including variations in the target DNA-sequence,³ intracellular signaling cascades,⁴ and, uniquely, small cell-permeable ligands.² The canonical mode of action of endogenous ligands involves their binding to the NR LBD and consequent reorientation of one of this domain's 12 α -helices (helix 12).⁵ This conformational change alters what was an interaction surface for corepressors to an interaction surface for coactivators, and it is these coregulatory proteins that are responsible for controlling chromatin remodeling and transcriptional activation.⁶

NRs are critical regulators of a variety of physiological processes, including embryonic development, growth, homeostasis of differentiated phenotypes, metabolism and reproduction, and consequently many pathological processes, such as obesity, inflammation, cardiovascular disease and cancer.⁷ In these processes, NRs are key determinants of gene expression patterns and consequently cellular phenotype. This central role, coupled to their ability to be influenced directly by exogenous ligands with attractive pharmacological properties, has made them a very successful class of drug targets. Roughly 13% of FDA approved drugs target NRs.⁸ Besides the evolved property of NRs to respond to ligands that can be exogenously supplied as drugs, the success of targeting the receptors has been enhanced by the realization that synthetic

ligands can be developed to minimize undesired side effects. Synthetic ligands have been discovered that share the full or partial agonistic properties of cognate ligands or antagonize (partially or fully) the function of natural ligands. Furthermore, some of these synthetic ligands selectively bind to specific receptor paralogs or isoforms, thus providing them with a more narrow range of effects.⁹ These ligands modulate their target receptor according to the canonical model: they stabilize the active (full agonist) or inactive (full antagonist) conformation of helix 12. Partial (ant)agonistic effects can be achieved by ligands that establish mixed populations of active and inactive helix 12 conformations, or ligands that stabilize one conformation, but with a slight distortion compared to the native conformation.¹⁰

Although the nuclear receptor targeting drugs are very successful they still face difficulties of undesired side effects¹¹ as well as resistance in cancers.¹²⁻¹⁴ One clear limitation of all of the discovered ligands is that they act solely through modulation of receptor conformations, and consequently their control over recruitment of coregulatory complexes is limited to interactions that are native to the surface of the receptors. The promoters of NR target genes are however often co-occupied by other transcription factors that recruit their own set of coregulatory proteins, and the receptors themselves recruit proteins in a manner dependent not only on the conformation of the LBD but also through binding-surfaces modulated by post-translational modifications such as phosphorylation and sumoylation.^{4,15-17} These ligand-independent events can indeed be the culprit of transcriptional misregulation.¹⁸ Thus, the nuclear receptor ligands provide a means for targeting critical genes, but are not always able to confer control of the transcriptional status of these genes. A novel solution to this problem is explored in this chapter: ligands of nuclear receptors are chemically linked to molecules that can recruit coregulatory proteins independent of interaction surfaces in the NR. These bifunctional NR ligands are tested for their ability to recruit functionality that is extrinsic to the nuclear receptor. The

glucocorticoid receptor is chosen as a representative member of nuclear receptors for these studies.

3.3 The glucocorticoid receptor

The glucocorticoid receptor (GR) is expressed in nearly all human cells and binds to glucocorticoids, which are steroidal hormones produced in the adrenal cortex.¹⁹ GR has a modular structure similar to other steroid receptors²⁰ which includes an N-terminal regulatory domain with a transcriptional activation domain (AF-1), a DNA-binding domain (DBD) and a ligand-binding domain (LBD). The ligand-binding domain also contains a transcriptional activation domain (AF-2) which is accessible for binding to co-factors in agonist-bound conformations. The unliganded receptor is localized in the cytosol, where it is complexed with the Hsp90 chaperone protein and co-chaperones p23 and FKBP5, which keep the receptor in a hormone-binding competent state, and several additional chaperones and co-chaperones are involved in GR homeostasis, cellular trafficking and nuclear function.^{21,22} Upon binding of ligands, GR translocates to the nucleus and binds as a homodimer to DNA-sequences termed GR-response elements (GRE) and agonist-bound receptor upregulates the majority of GRE-regulated genes (this process is called GR transactivation).²³ Ligand-bound GR also binds as a monomer to pro-inflammatory transcription factors, such as AP-1 and NF- κ B,^{24,25} and can represses their target genes (GR transrepression). In transrepression, GR does not bind DNA directly, but is instead tethered by the transcription factor to which it binds. GR-targeting drugs are the most common therapeutics used to treat inflammatory diseases.²⁶ The transrepression of pro-inflammatory transcription factors mediates numerous desirable anti-inflammatory effects, while transactivation is the cause of most adverse effects that limit the use of the most potent GR drugs to topical treatment such as inhalation for treatment of asthma. Efforts to improve on GR targeting therapeutics are therefore focused on development of GR ligands with altered transcription profile of both transactivated and transrepressed genes.¹¹

3.4 Nuclear receptor extrinsic recruitment

The focus of this study is not to pursue dissociated ligands of GR, but rather to use GR as a model receptor for a new strategy of modulating receptor function. This strategy is aimed to be general for all the steroid nuclear receptors, each of which, as a therapeutic target, has desired functions as well as undesired off-target effects. The nuclear extrinsic recruitment strategy involves linking nuclear receptor ligands to a ligand of a second protein. If this conjugate can simultaneously bind to both of its target proteins, it will effectively recruit its second target protein to NR and to the target genes of NR (Figure 3.1).

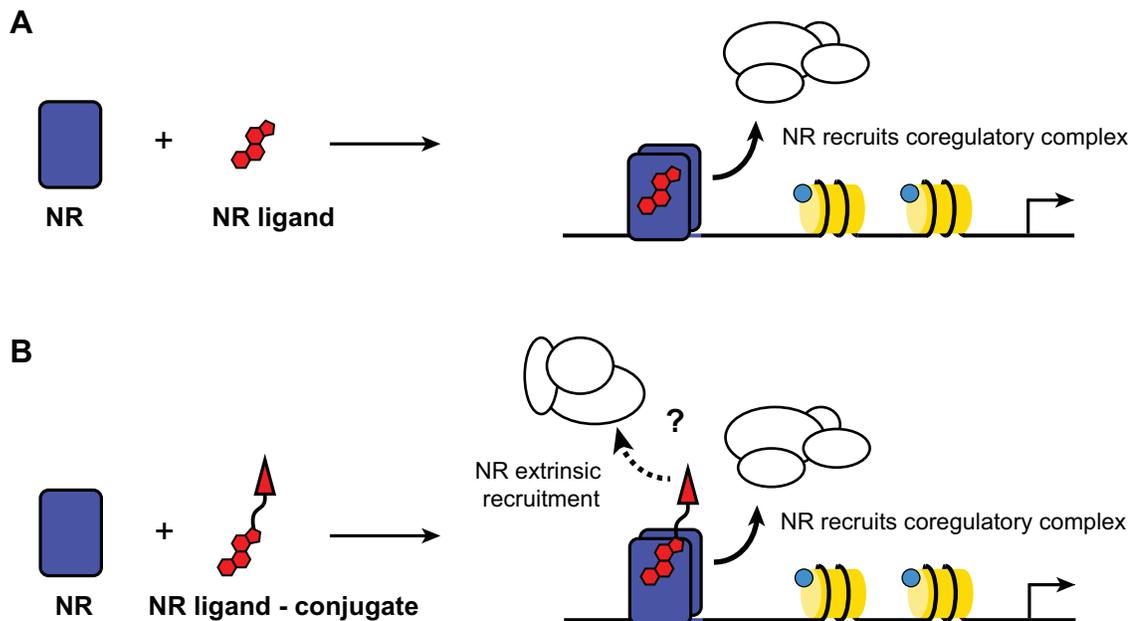


Figure 3.1 Nuclear receptor extrinsic recruitment A. A NR ligand binds to its receptor, and the dimerized receptor binds to DNA and recruits target complexes to control transcription; B. A NR ligand, which is conjugated to a ligand of a second target protein, binds to the receptor and the dimerized receptor binds to DNA at its target genes. The principle of the extrinsic recruitment strategy involves the recruitment of the second protein or complex directly by the conjugated ligand and aims to have the extrinsic factors influence transcription of the target gene.

The first goal of the project is to synthesize bifunctional molecules, which are conjugates of GR ligands and a ligand that can bind a protein, whose presence at GR target genes will affect transcription and provide proof-of-principle of the extrinsic recruitment strategy.

3.5 Recruitment of non-native proteins to the glucocorticoid receptor

The ligand-binding domain of the glucocorticoid receptor (GR) has previously been used in a mammalian three-hybrid assay; a synthetic GR ligand, dexamethasone, conjugated to the natural product, FK506, facilitates dimerization of the GR LBD and FK506-binding protein 1A (FKBP).²⁷ In this experiment, GR-LBD was fused to the strong transcriptional activation domain of VP16 and FKBP was fused to a DNA-binding domain. Dimerization of the two fusion proteins therefore assembles a transcriptional activator that can activate a reporter gene (See Figure 3.2B for an analogous three-hybrid assay, except with the ligand binding domains swapped relative to DBD and TAD). Based on this three-hybrid experiment, it is expected that FKBP-fusion proteins can likewise be recruited to full-length GR. Such recruitment will facilitate studies to test if transcriptional regulatory factors recruited indirectly to GR (fused to FKBP) can influence transcription of GR-target genes in a context that is also influenced by intrinsic GR activity. The first aim of the project therefore was to conjugate ligands of these proteins to obtain bifunctional molecules that could be studied for recruitment potential. The GR ligands, dexamethasone (Dex) and RU486, as well as the FKBP ligand, FK506, were chosen for conjugates and can initially be tested in an analogous three-hybrid assay (Figure 3.2).

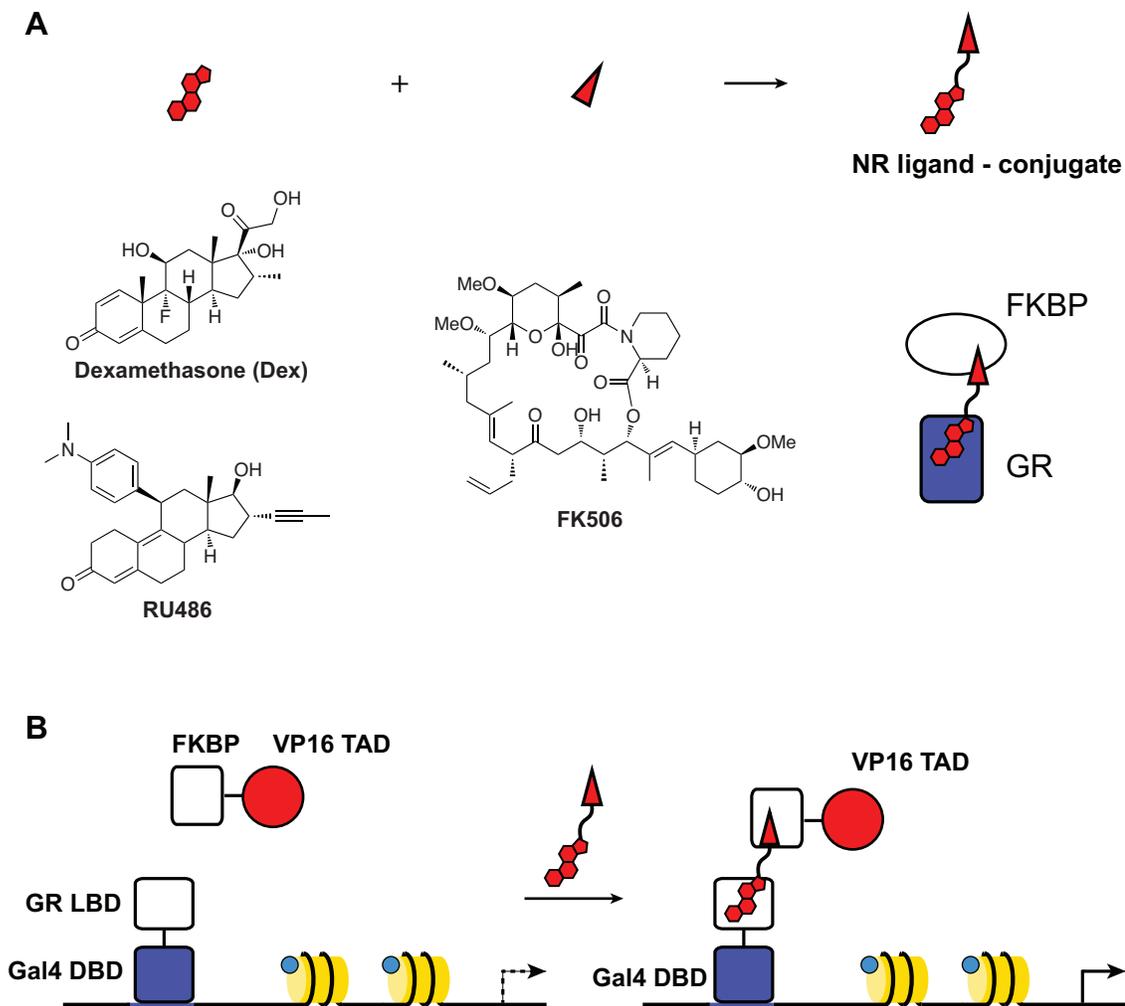


Figure 3.2 Ligands for constructing bifunctional molecules A. The GR ligands, dexamethasone (Dex) and RU486, and the FKBP ligand, FK506, are chosen for conjugation. Their conjugates should recruit FKBP to GR; Initial evaluation of ligands can be achieved through the use of a three-hybrid assay. The assembly of DBD and TAD will activate transcription of a reporter gene.

Synthesis of conjugates of GR-ligands and FK506

The synthetic GR agonist, dexamethasone, was discovered in 1958²⁸ as a more potent alternative to the natural ligand, hydrocortisone, and synthetic ligand, prednisolone. Dexamethasone can be modified with a thioether linked isothiocyanate (SDex-ITC), which facilitates easy coupling of various amine-functionalized molecules and is reported to maintain partial agonist activity and full potency of dexamethasone.²⁹ This derivative has been shown to have higher

binding affinity than another common derivative, OxDex, used for conjugations.³⁰ Thus, SDex-ITC was prepared to facilitate conjugation to linkers and FK506 (Figure 3.3).

RU486 is an antagonist of GR (and the progesterone receptor) discovered by scientists at Roussel-Uclaf, who had previously published work on 11 β -substituted 19-norsteroids.³¹ It binds competitively with dexamethasone to GR, and has very low transactivation potential that is dependent on the GR AF-1 domain. It does mediate transrepression but is less effective than dexamethasone.³² RU486 is derived from the synthetic antiprogestin, norethindrone, and has a 4-dimethyl-amino-phenyl substitution at the 11 β -substitution. The aniline amine of RU486 has previously been used as a conjugation point of bile acids, to make a bifunctional molecule that targets liver tissue (via the bile acid) with unperturbed activity of RU486,³³ and this aniline was therefore also used in our conjugation of linker and FK506. One methyl group of the dimethylaniline was oxidized to the N-formyl amide which was subsequently hydrolyzed to yield the monomethyl aniline (Figure 3.3). This secondary amine will allow coupling to a linker functionalized with a methylsulfonic ester (Figure 3.4).

FK506 is a bacterially produced immunosuppressive macrolide^{34,35} that binds to a family of FK506-binding proteins.³⁶ These are peptidyl-prolyl isomerases with diverse functions, including the Hsp90 co-chaperone function of FKBP5 as mentioned above. The predominant binding target of FK506 in human cells is the 12 kDa FKBP1A (FKBP),³⁷⁻⁴⁰ to which FK506 binds with a dissociation constant of 0.6 nM.⁴¹ The first bifunctional small molecules designed to dimerize two fusion proteins, termed chemical inducers of dimerization, were homodimers of FK506.⁴² In this and later studies, an allyl substituent in FK506 is functionalized with an amine reactive N-hydroxysuccinimidyl carbonate for convenient linkage to conjugate partners (Figure 3.3). This modification does not interfere with FKBP1A binding, but disrupts binding to calcineurin, which otherwise binds to the

FK506-FKBP complex (an interaction also critical for the immunosuppressive properties of FK506).⁴² This derivative of FK506 was prepared.

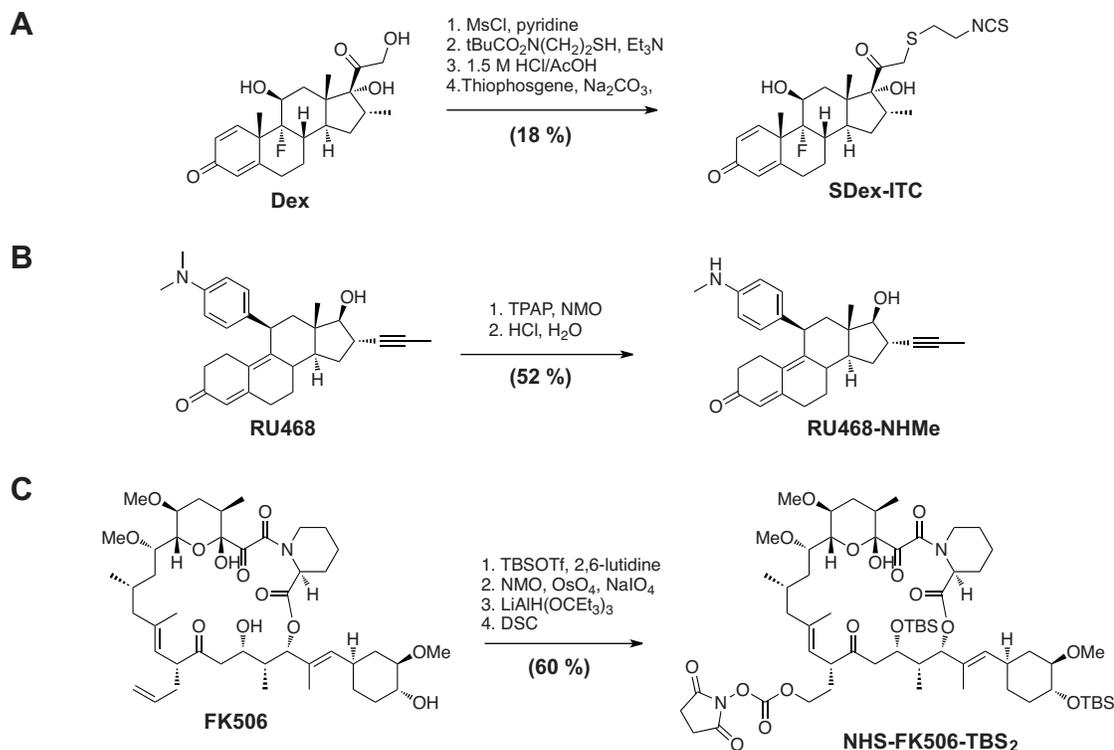


Figure 3.3 Functionalization of ligands The commercially available ligands of GR and FKBP are functionalized to facilitate conjugation: A. Dexamethasone is converted to the amine reactive SDex-ITC; B. RU486 is converted to the nucleophilic RU486-NHMe; C. FK506 is converted to the amine reactive NHS-FK506-TBS₂.

The functionalized GR ligands and FK506 were conjugated with linkers as shown in Figure 3.4. These linkers must be long enough to allow simultaneous binding of the linked ligands to their respective proteins. Optimization of the length, hydrophobicity and rigidity of linkers can affect the efficacy of conjugates, but no clear general design strategies have emerged from previous studies.⁴²⁻⁴⁴ One of these studies⁴³ found that the shortest linker long enough to allow simultaneous binding, had the highest potency. This study, however, used homodimeric ligands and fusion proteins with multimerized ligand binding domains (LBDs), and the lower activity of conjugates with long linkers may stem from an ability to

bind two fused LBDs instead of LBDs of two different target fusion proteins. We have chosen to use either straight alkyl or polyethylene glycol linkers of varying lengths that are comparable to those used in previous GR LBD – FKBP dimerizers.²⁷

Evaluation of GR ligand-FK506 conjugates in a three-hybrid assay

The conjugates of RU486 or SDex with FK506 were first evaluated in a mammalian three-hybrid assay (Figure 3.2B). In this experiment, HeLa cells were transfected with a luciferase reporter with 5 Gal4-binding sites upstream of a minimal promoter, as well as expression vectors for Gal4 – GR-LBD and FKBP - VP16 fusion proteins. The conjugates were tested for their ability to dimerize GR LBD and FKBP, which will bring the strong VP16 activation domain to the Gal4 fusion protein and activate the reporter gene. The experiment is similar to a previous three-hybrid assay reported in yeast,²⁷ with the notable difference that the GR LBD in this assay is fused to the DNA-binding domain. Ligands of GR may therefore activate transcription through the AF-2 domain in the GR LBD as well as through recruitment of FKBP-VP16.

Dexamethasone (Dex) activated the reporter approximately 60-fold compared to DMSO (Figure 3.5B). This activation is expected to stem from the AF-2 function in the LBD and has been demonstrated in cells that are not transfected with VP16-FKBP (data not shown). Addition of FK506 affected the activation by Dex, but only partially. This may be an effect of FK506 inhibition of FKBP5 and FKBP4, which are Hsp90 co-chaperones that are required to maintain GR in a ligand binding competent state.⁴⁵ As controls for the conjugates, the unconjugated mono-functional ligands were prepared and tested (Figure 3.5): The activity of SDex-C7, showed partial agonistic activity. RU486-O₃-N₃ and FK506 showed no activity (Figure 3.5C).

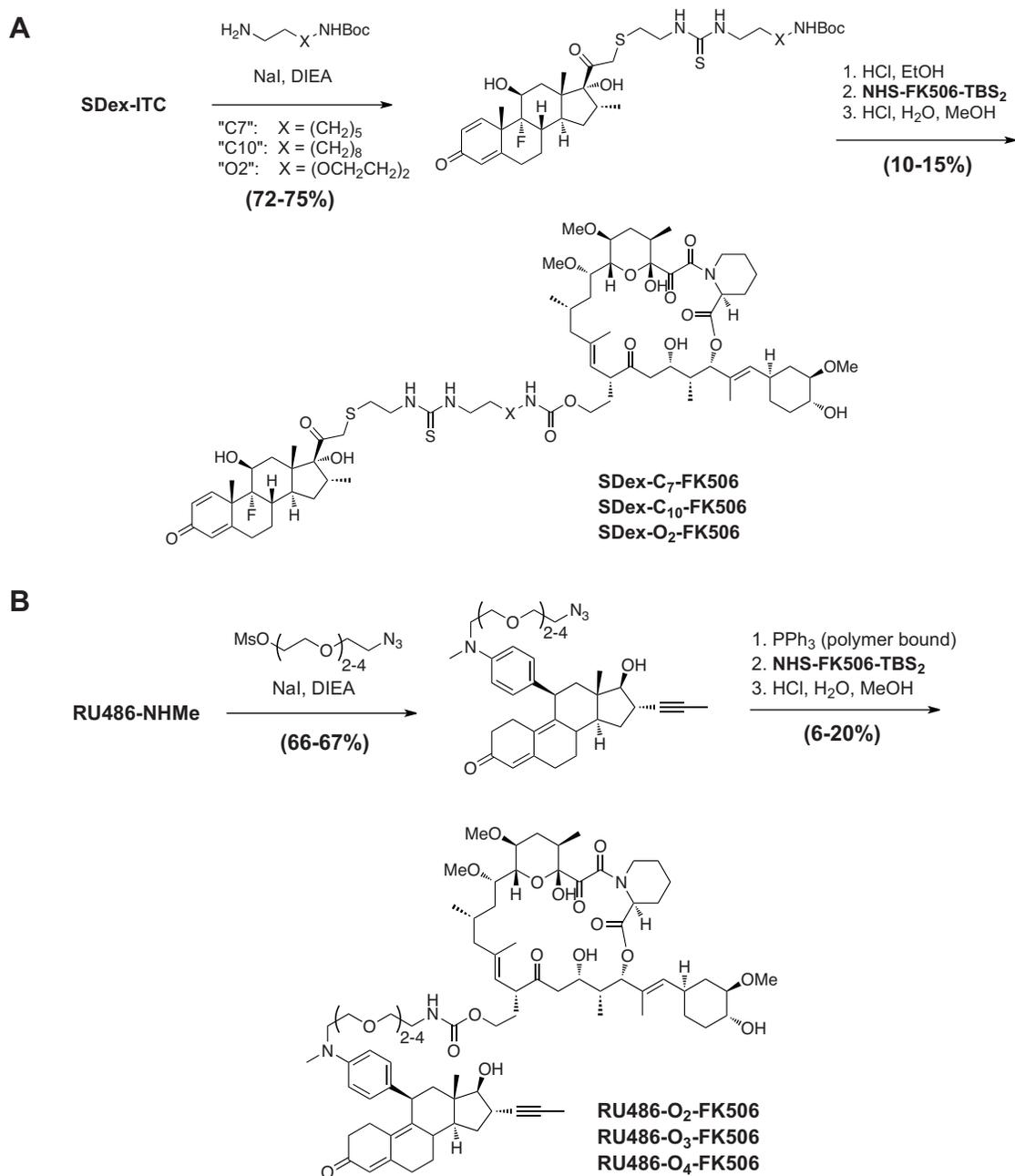


Figure 3.4 Conjugates of GR ligands with FK506 A. Amine-functionalized linkers are coupled to SDex-ITC. A second amine in the linker is then Boc-deprotected and coupled to amine reactive NHS-FK506-TBS₂. The final step is a removal of TBS-protecting groups; B. Electrophilic linker mesylates are coupled to RU486-NHMe. The linker's azido group is then reduced to the free amine and coupled to FK506 as was done for SDex.

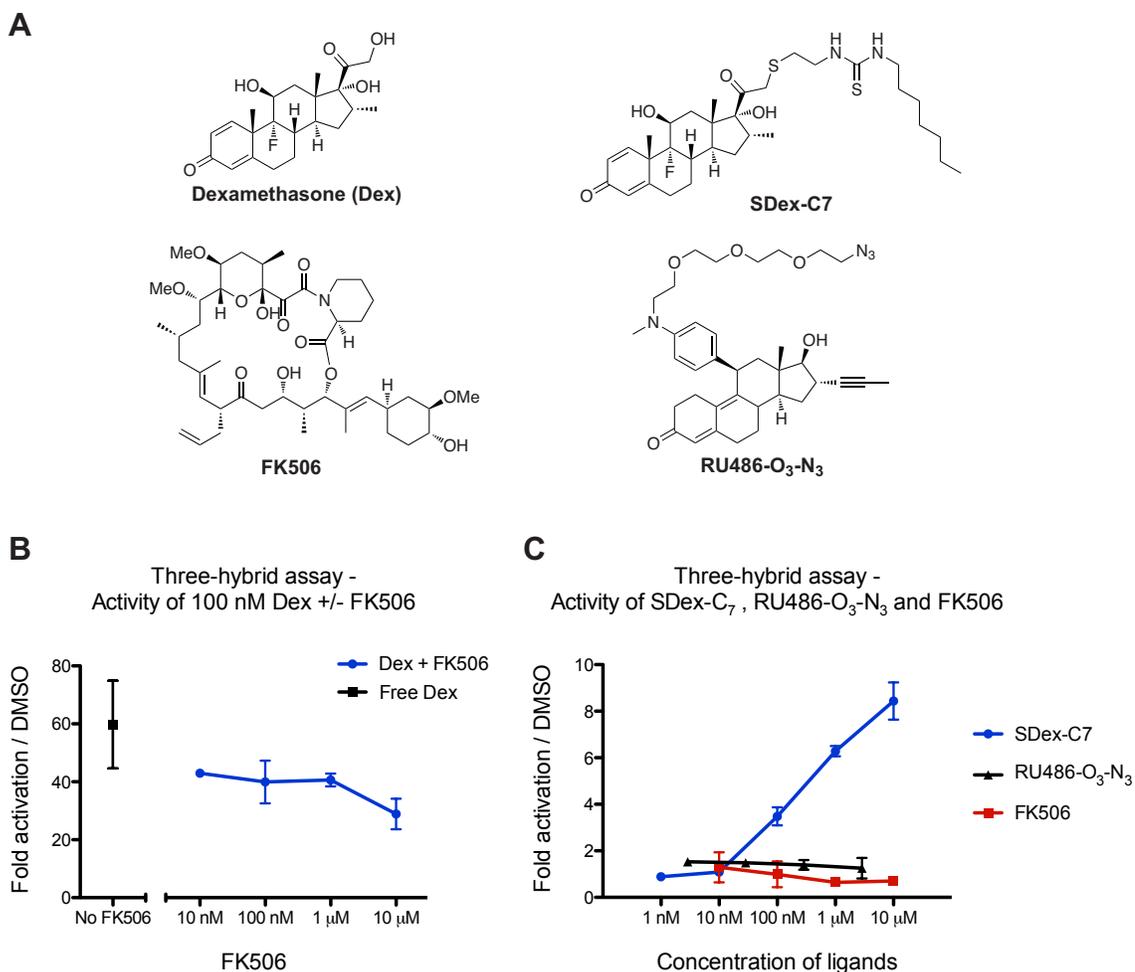


Figure 3.5 Activity of unconjugated ligands in the three-hybrid assay
Compounds are tested in a three-hybrid assay in which HeLa cells have been transfected with a Gal4 luciferase reporter, a Gal4-GR(LBD) expression plasmid, and a VP16-FKBP expression plasmid. A. Structures of unconjugated ligands; B. Dexamethasone activates transcription 60-fold relative to DMSO, and FK506 only partially affects the activity of Dex; C. SDex-C₇ shows partial agonist activity, while no activity is seen from RU486-O₃-N₃ or FK506. All data points are from triplicate samples and error bars represents the standard deviation of the three samples.

In contrast to the unconjugated ligand, SDex-C₇, the three SDex-FK506 conjugates activated transcription potently (Figure 3.6A). The C₁₀-linked conjugate was weaker than the other two conjugates, which have comparable potencies that are higher than Dex activity. This enhanced activity is due to recruitment of VP16-FKBP, as evidenced by the ability of free FK506 to inhibit the activity of SDex-C₇-FK506 to a level comparable to SDex-C₇ (Figure 3.6B).

The molecules also self-inhibit at high concentration, which is expected since the monovalent interactions will dominate the bivalent.⁴⁶

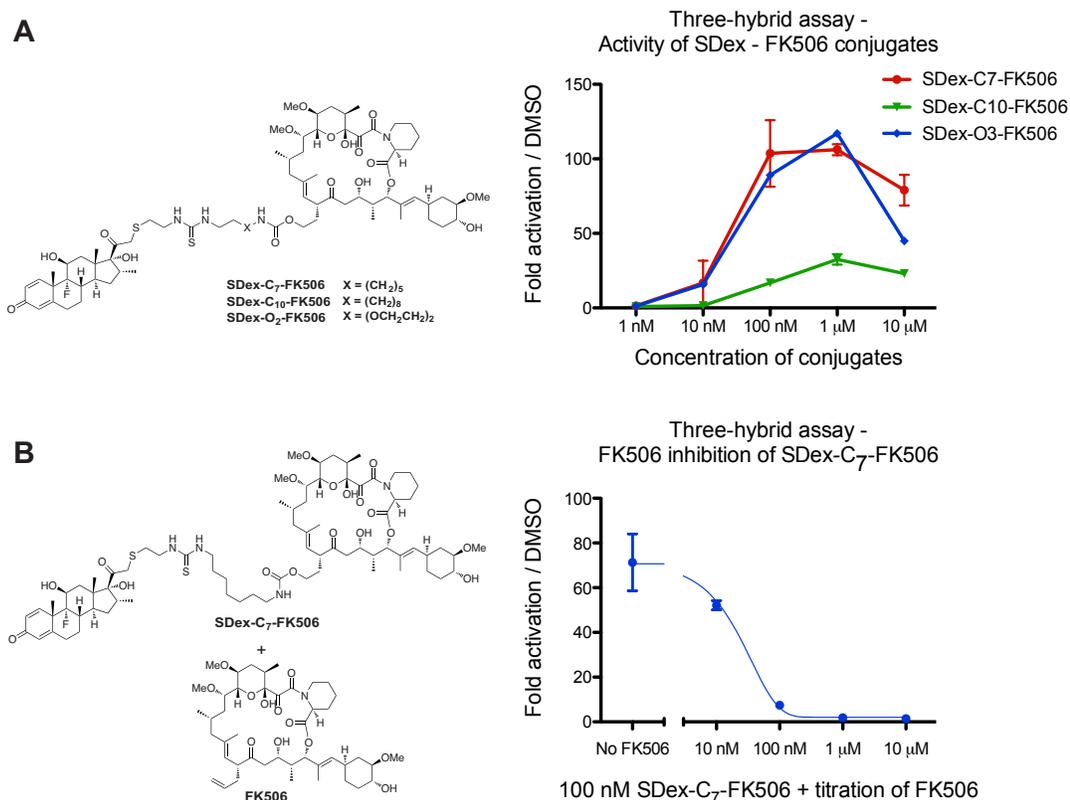


Figure 3.6 Activity of SDex-FK506 conjugates Same assay as in Figure 3.5. A. SDex-O₂-FK506 and SDex-C₇-FK506 show high activation potential that is self-inhibited at high concentrations. SDex-C₁₀-FK506 shows weaker activity; B. The activity of 100 nM SDex-C₇-FK506 is inhibited by titration of free FK506 and demonstrates VP16-FKBP activity dependence. All data points are from triplicate samples and error bars represents the standard deviation of the three samples.

The three RU486-FK506 conjugates also showed transcriptional activation. The activation levels were lower than for Dex and SDex-FK506 conjugates, but was fully dependent on VP16-FKBP recruitment, since RU486-O₃-N₃ showed no activation. Despite having an intermediate linker length in the series, the RU486-O₄-FK506 showed lower activity than the other two conjugates. For the most accurate comparison, these compounds will in the future be tested simultaneously since transfection conditions can vary and affect the results.

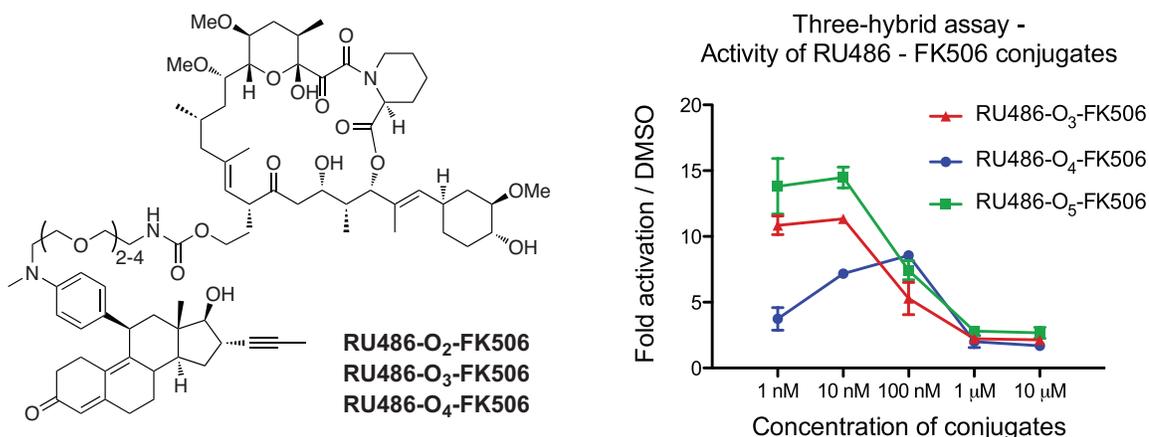


Figure 3.7 Activity of RU486-FK506 conjugates RU486-O₃-FK506 and RU486-O₅-FK506 show comparable activation in three-hybrid assay. The lower activity and potency of RU486-O₄-FK506 is likely due different levels of fusion proteins, as data is from separate transient transfection.

The mammalian three-hybrid experiment showed that both agonistic and antagonistic GR ligands fused to FK506 can bring together their respective binding targets. From the agonist-conjugate series of linker variations, SDex-O₂-FK506 was chosen for further studies, as it was the most potent compound. SDex-O₂-OMe was chosen as a control compound, to help discern the role of linked FK506. The antagonist-conjugate series showed minor influence on linker length, and the shorter RU486-O₃-FK506 was chosen along with RU486-O₃-N₃ as control compound. These structures are shown in Figure 3.8.

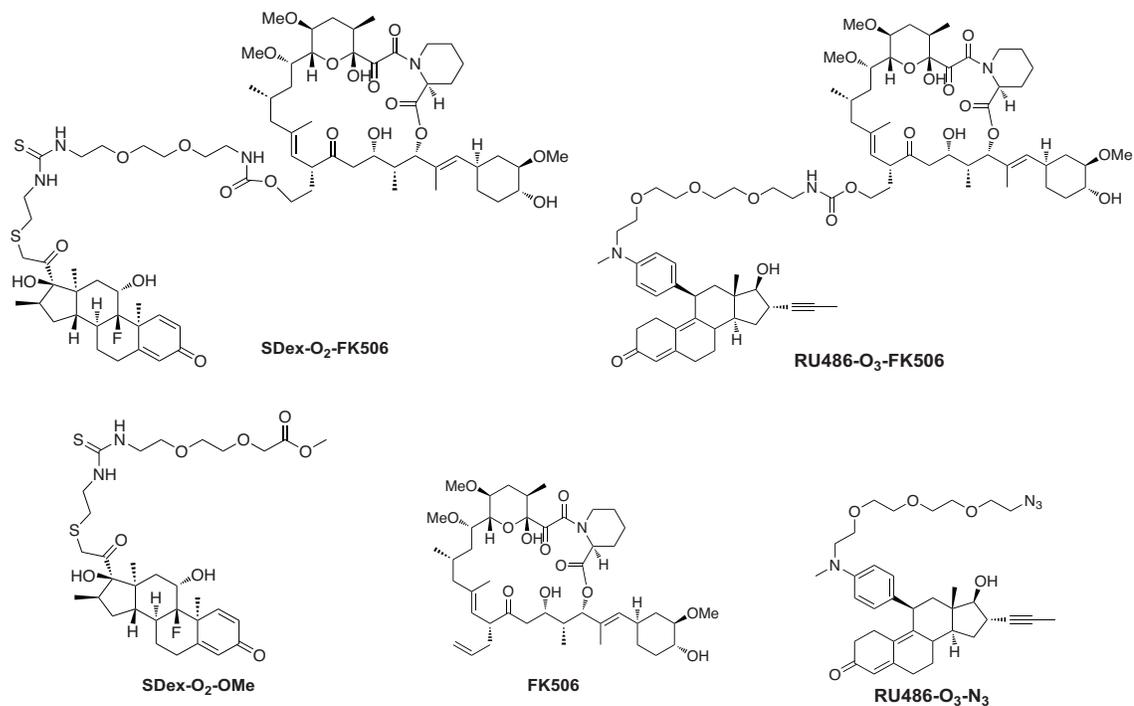


Figure 3.8 Overview of compounds chosen for studies with full-length GR

Binding affinity of conjugates to GR

The affinity of the conjugates for GR was determined in a competition binding assay using radio-labeled Dex and cell lysates with overexpressed GR (Figure 3.9). SDex-O₂-OMe bound to GR with an affinity (48.3 nM) that is one order of magnitude lower than Dex. SDex-O₂-FK506 bound with 2-fold lower affinity compared to SDex-O₂-OMe. The RU486 conjugates bound with similar affinity to each other and to the SDex-conjugates (83.6 and 93.9 nM). This is also an order of magnitude lower binding affinity compared to the parent compound, RU486.

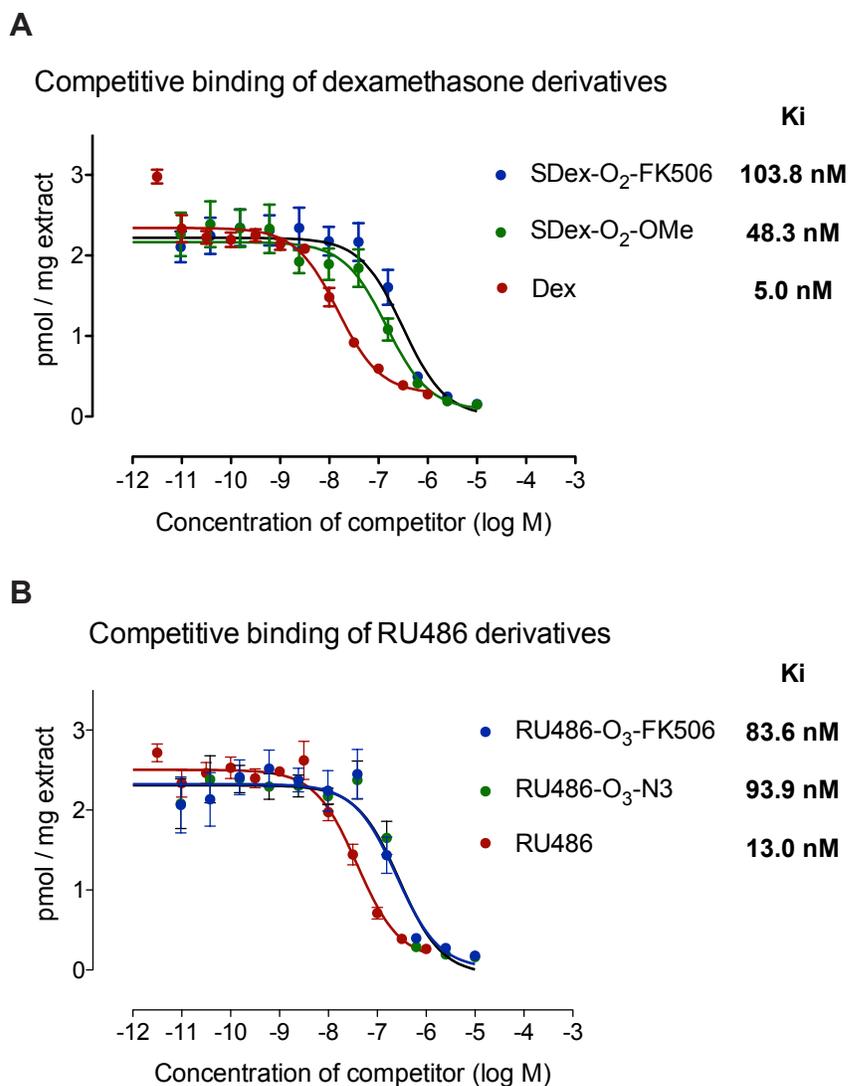


Figure 3.9 Binding affinity of conjugates for GR The binding affinity of conjugates and GR ligands was measured in a competition assay with radio-labeled ligand. A. Dex-derived compounds; B. RU486-derived compounds. Curves are fitted with GraphPad Prism 5.0 (Competition Binding, One-site – fit Ki)

Bifunctional GR ligands can recruit non-native functionality to the full-length receptor

In order to determine if extrinsic transcriptional control via VP16-FKBP can be recruited in the same manner to the full-length receptor as was seen in the three-hybrid assays, the effects on GR-responsive luciferase reporters were studied. A

reporter that contains the endogenous enhancer element from the FKBP5 gene, which includes two GREs,²³ was transfected into human embryonic kidney 293T cells (HEK293T) together with a GR (full-length) expression plasmid, and with or without co-transfection of a VP16-FKBP expression plasmid. The activity of conjugates and ligands in these cells are shown in Figure 3.10 – 3.12. In all the figures, the activity is shown as a percentage of the activity of Dex at 100 nM. SDex-O₂-FK506 and SDex-O₂-OMe showed partial agonistic activity (nearly full agonistic activity) with reduced potency (by 2 orders of magnitude) compared to Dex in cells with no VP16-FKBP (Figure 3.10A). In cells cotransfected with the fusion protein, SDex-O₂-OMe showed no change in potency or efficacy, but SDex-O₂-FK506 activated transcription to levels over 150% relative to Dex, and with increased potency as well (10-fold relative to SDex-O₂-OMe)(Figure 3.10A, right panel). The enhanced activation potential beyond that of Dex is evidence for recruitment of VP16-FKBP. The fusion protein is being recruited by the conjugated FK506 moiety as evidenced by the ability of free FK506 to lower the maximal response and potency of SDex-O₂-FK506 in the presence of the fusion protein (Figure 3.10C). FK506 has no effect on SDex-O₂-OMe activity apart from a raised baseline that is seen in all samples with FK506 at 1 μM. As mentioned previously, FK506 can affect endogenous proteins such as GR-associated FKBP4 and FKBP5 co-chaperones.

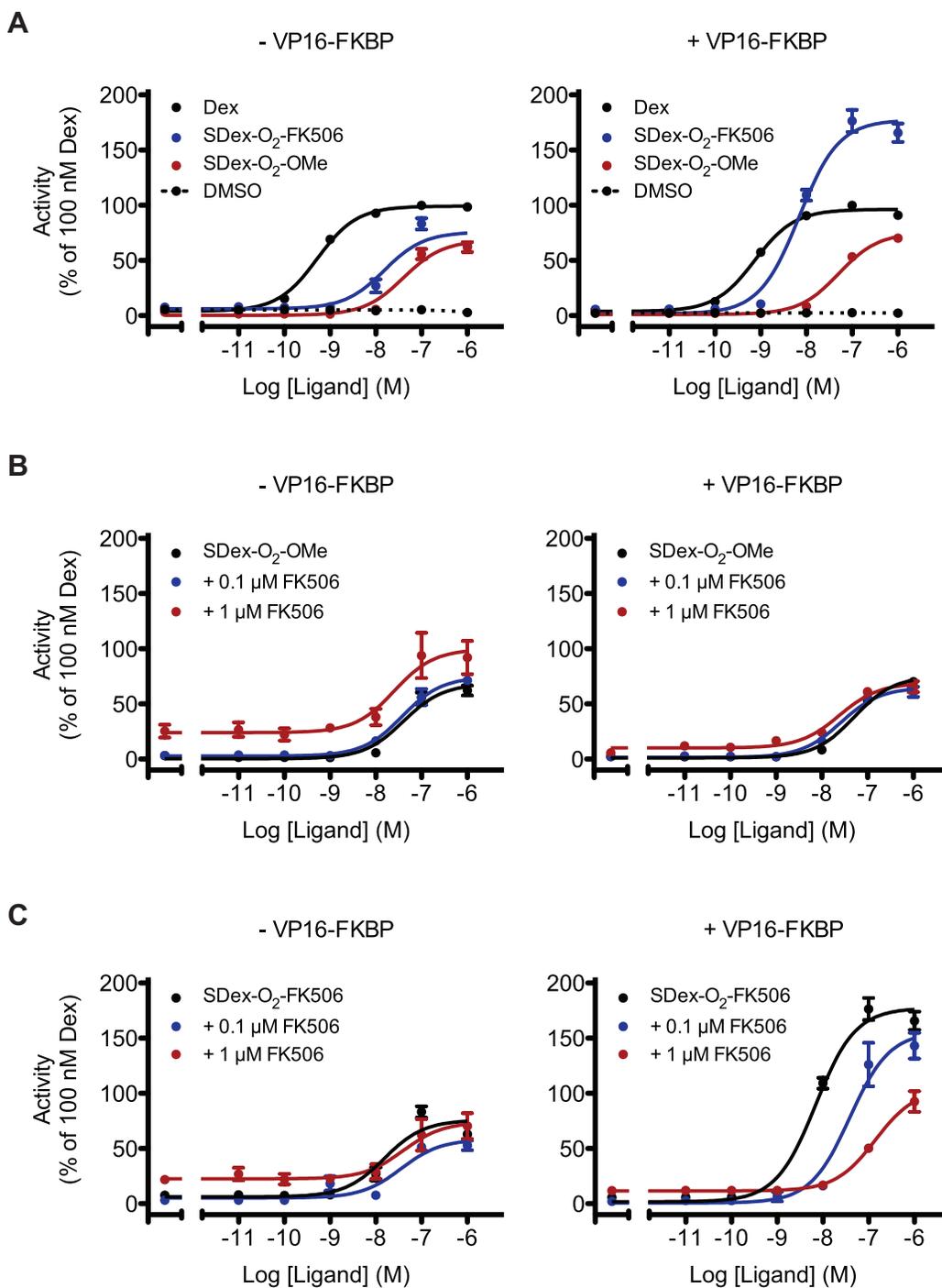


Figure 3.10 Functional recruitment of VP16-FKBP to GR by SDex-conjugates HEK293T cells were transfected with a GRE-luciferase reporter and GR-expression plasmid. In experiments represented with figures on right, a VP16-FKBP expression plasmid was also cotransfected into the cells. The compounds and concentrations are indicated on graphs. In figure B and C, blue and red curves are from addition of both SDex-conjugate and constant concentration of FK506. Activity is normalized to % of 100 nM Dex.

The antagonist derived RU486-O₃-FK506 and RU486-O₃-N₃ were also tested in the above described luciferase reporter assay (Figure 3.11). In the absence of VP16-FKBP fusion protein they both showed activity identical to RU486 with only very modest activation (10% or less). In the presence of the fusion protein, however, RU486-O₃-FK506 became a potent activator (Figure 3.11A, right panel), and the enhanced activation potential could be completely blocked by FK506 (Figure 3.11B, right panel). Similar to what was observed in the three-hybrid assay the RU486-FK506 conjugate self-squelched at high concentrations. Compared to SDex-O₃-FK506, RU486-O₃-FK506 activated at a slightly lower concentration, and self-squelching is evident at 1 μM. The shifted dose-response is likely not related to the binding affinity for GR, since the measured binding affinities for GR are nearly identical.

The RU486 compounds showed ability to compete with Dex for binding to GR and antagonized activity in absence of the VP16-FKBP fusion protein (Figure 3.12, left panel). In the presence of the fusion protein, RU486-O₃-FK506 showed enhancement of activity beyond that achieved by dexamethasone (Figure 3.12, right panel), despite showing a maximal 80% relative activity when it was added alone (Figure 3.11). The enhanced activity may stem from hetero-liganded GR dimers that can combine the agonistic effect of Dex-bound GR with the GR-extrinsic activity of VP16-FKBP recruited via FK506.

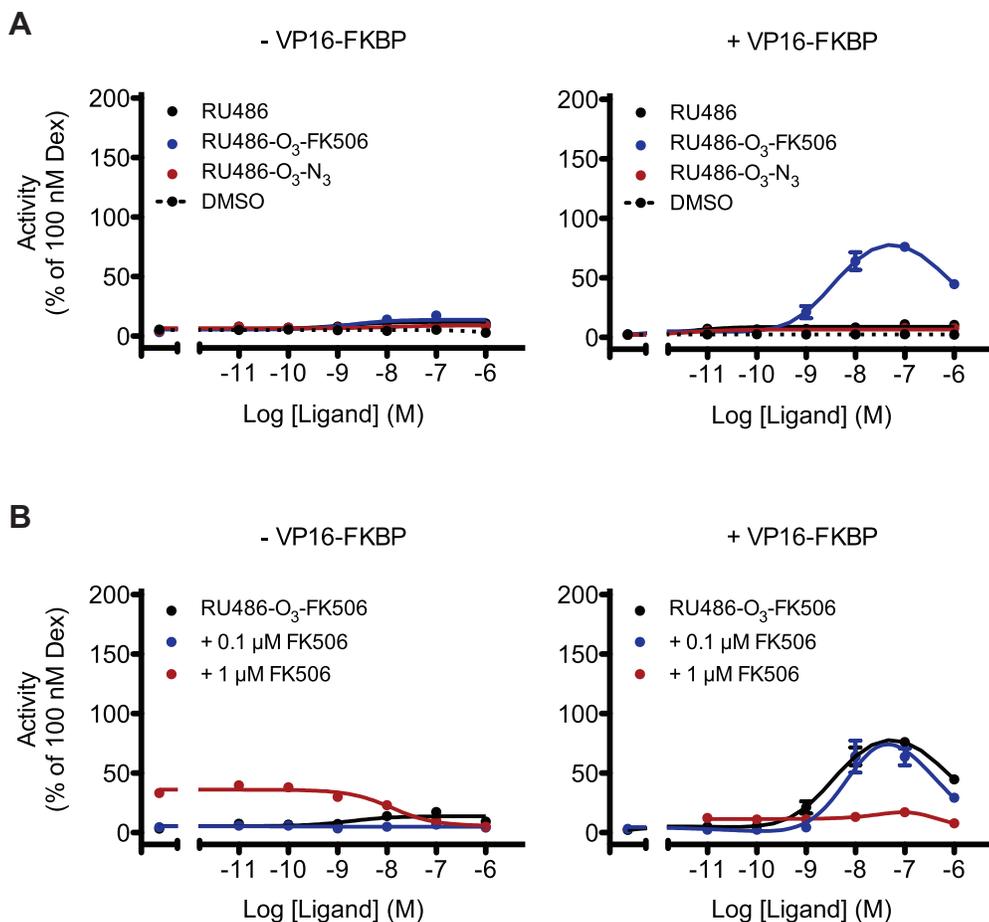


Figure 3.11 Functional recruitment of VP16-FKBP to GR by RU486-conjugates HEK293T cells were transfected with a GRE-luciferase reporter and GR-expression plasmid. In experiments represented with figures on right, a VP16-FKBP expression plasmid was also cotransfected into the cells. The compounds and concentrations are indicated on graphs. In figure B, blue and red curves are from addition of both RU486-conjugate and constant concentration of FK506. Activity is normalized to % of 100 nM Dex.

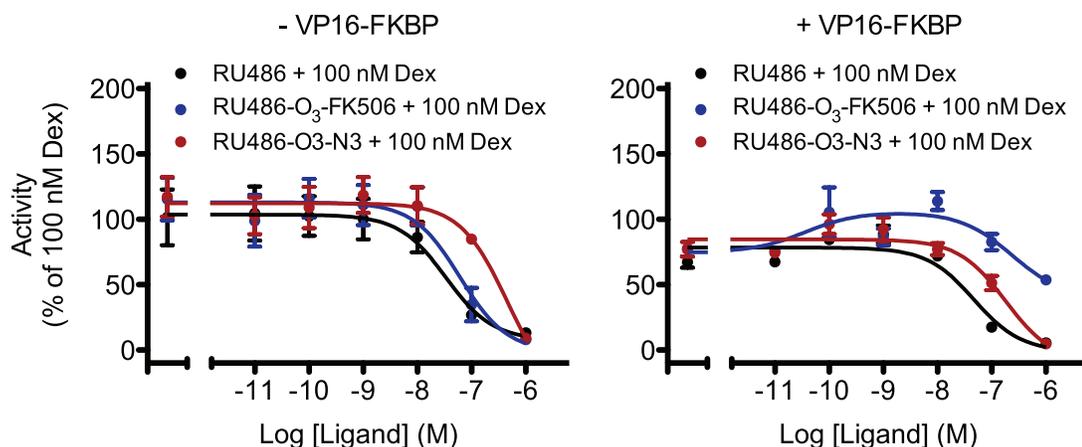


Figure 3.12 Effect of competitive binding of RU486 conjugates in the presence of Dex Cells are transfected as in Figures 3.7 and 3.8. The concentration of compound on the x-axis refers to the RU486 compounds. Dex is added at a constant 100 nM. Activity is normalized to % of 100 nM Dex.

These studies demonstrated that direct recruitment of the VP16-FKBP fusion protein to GR can be achieved by a bifunctional GR ligand, and that the transcriptional activation ability of VP16 influences transcription of the target gene. Impressively, the antagonist RU486 was converted into a potent activator.

Effects of VP16-FKBP recruitment to GR transrepressed genes

Ligand-bound GR represses NF- κ B-mediated transcriptional activation of several genes including the proinflammatory interleukin-8 (IL-8) gene in a process known as transrepression.⁴⁷ In its transrepression function, GR binds directly to DNA-bound NF- κ B rather than to DNA.^{25,48} Both Dex and RU486 cause transrepression, although RU486 is less effective at this. The maximal suppression by RU486 is lower than that of Dex, which may be explained by lack of recruitment of HDAC2 by RU486,³² but the level of GR expression can also influence the transrepression ability of RU486.⁴⁹ To study the effects of VP16-FKBP recruitment to transrepressed genes, an NF- κ B-regulated luciferase reporter was used. HEK293T cells were transfected with this reporter together with GR and +/- VP16-FKBP. Cellular NF- κ B was activated by treating cells with tumor necrosis factor alpha (TNF α).⁵⁰ We observed that the GR ligands and

conjugates all cause transrepression, but the transrepression is not affected by the presence of VP16-FKBP. This was observed in only a single preliminary experiment and should be explored further as levels of GR (transfected) and of TNF α (added to media) affects the degree of transrepression and could influence the receptiveness to modulation by recruited VP16 activity.

It is possible that VP16-FKBP recruitment is blocked in the NF- κ B – GR complex, or alternatively, that the strong transcriptional activation domain of VP16 cannot counter the transrepression mechanism. Although a VP16 fusion protein can reactivate genes silenced by deacetylated and repressively methylated chromatin,⁵¹ simultaneous co-localized HDAC-associated repressors can be dominant over VP16 activation potential.^{52,53} It should also be remembered that both NF- κ B and agonist-bound GR contain strong activation function as well that is repressed in this context. Studies have revealed that RNA polymerase II (Pol2) is indeed recruited to some GR-transrepressed genes (preinitiation complex is formed), but GR interferes with phosphorylation of serine 2 of the Pol2 C-terminal domain (CTD) and Pol2 is stalled at the promoter.²⁵ In agreement with this, promoter association of CyclinT1 and cyclin-dependent kinase 9 (Cdk9), which together form the positive transcription elongation factor b (P-TEFb) responsible for CTD Ser2 phosphorylation, is also blocked in a GR dependent fashion.⁵⁴ The altered recruitment of CyclinT1 and Cdk9 could be the primary cause of transrepression, or it could be just a symptom of another repression mechanism. In an attempt to answer this question, the CyclinT1 and Cdk9 genes were each cloned into an FKBP-fusion vector. If SDex/RU846-FK506 conjugates can mediate recruitment of CyclinT1-FKBP or Cdk9-FKBP and if these can form an active P-TEFb complex with their respective endogenous binding partner, it can be hypothesized that they would phosphorylate CTD Ser2 and initiate promoter release of Pol2 at transrepressed genes. Cotransfection of HEK293T cells with these fusion proteins in the transrepression assay, however, did not reduce the transrepressive potential of FK506 conjugates. In the absence of an effect, the study is not conclusive as the recruitment of these fusion proteins could be

sterically hindered or the association of the native binding partner to the fused protein could be blocked by the fusion. Ideally, their recruitment would be studied with chromatin immunoprecipitation, which could also be used to reveal if the CTD Ser2 phosphorylation is occurring.

Effects of conjugates on endogenous genes

A human adenocarcinoma lung epithelial cell line, A549, which expresses endogenous GR and is commonly used to study the effect of GR on GRE-regulated genes as well as transrepression function, was stably transfected with a VP16-FKBP-myc₆ expression plasmid. The myc₆-tag does not interfere with the function of the fusion protein in the mammalian three-hybrid assay or the GR-luciferase assays, and they were included to potentially help facilitate chromatin immunoprecipitation (ChIP) studies. The stable cell line ensures expression of the fusion protein in all cells. This is in contrast to a transient transfection of A549 cells, for which the optimized transfection conditions used by us only achieve delivery of plasmids to 50-60 % of cells in each transfection. In studies with transiently cotransfected reporters, the fraction of cells receiving plasmids are not a concern since only the transfected cells contributes to the readout, while expression analysis of endogenous genes, involves the entire cell population. A total of 37 stable clones were tested for expression of the fusion protein, and one clone showed expression (Figure 3.13A). This clone (124.25) has a slightly reduced growth rate compared to WT A549 cells, but has so far been cultured with over 20 passages without loss of fusion protein expression.

The expression level of the fusion protein in 124.25 cells relative to the levels obtained with transiently transfected cells and relative to endogenous FKBP1A was determined with a Western blot (Figure 3.13B). The 124.25 cells expressed VP16-FKBP-myc₆ at ~10-fold lower level compared to endogenous FKBP1A. The transient expression in this instance had ~38-fold higher level of expression of the fusion protein compared to the stable cell line and ~5-fold higher than endogenous FKBP1A. The expression levels in different transient transfections

vary, however, and have often been comparable to or lower than FKBP1A-levels. The relative expression of fusion protein and endogenous FKBP1A in 124.25 cells has been consistent. Since FK506 can bind to endogenous FKBP1A there is a concern that this protein will compete for binding to conjugates and suppress effective recruitment of VP16-FKBP to gene promoters.

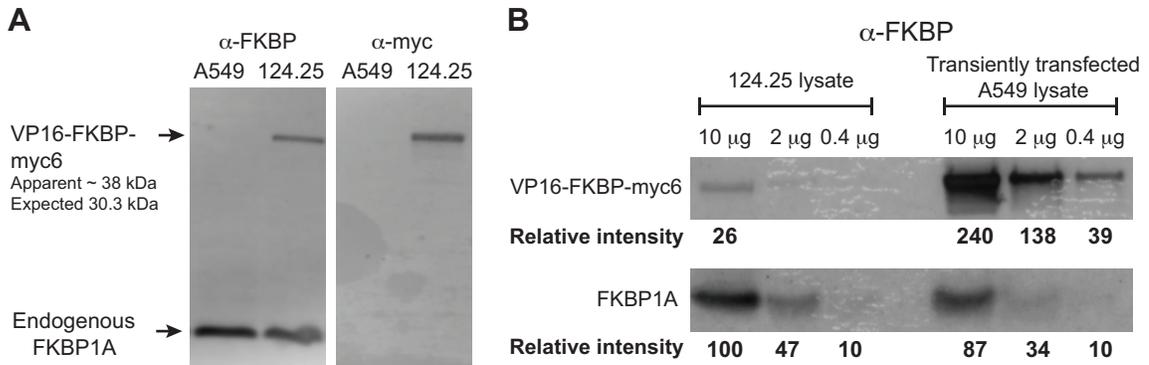


Figure 3.13 Expression levels of VP16-FKBP and endogenous FKBP1A A. Western blot of an A549 cell line stably transfected with VP16-FKBP-myc₆ (Clone 124.25). The blot probed with α -FKBP (left) show expression of both introduced fusion protein and endogenous FKBP1A. A blot probed with α -myc confirms identity of the fusion protein; B. Relative expression levels of fusion protein and endogenous FKBP1A in stably and transiently transfected cells. Background subtracted pixel-counts are determined with ImageJ. Non-saturated bands are used to estimate relative amounts.

RU486-O₃-FK506 previously showed the biggest fusion-protein dependent change in activation potential of the luciferase reporters and was therefore tested in A549 cells transiently transfected with VP16-FKBP-myc₆ or in the stable 124.25 clone. The mRNA levels of two GR target genes (GILZ and FKBP5) that show high levels of induction by Dex⁵⁵ were quantified by RT-qPCR. GILZ mRNA levels were induced quickly, with significant levels observed after 1 hour of treatment (data not shown) and has stable expression between 6 and 24 hours (Figure 3.14A). The induction of the FKBP5 gene was slower, but was significant after 6 hours of treatment and slightly higher after 24 hours. Endogenous genes in transiently transfected (VP16-FKBP) A549 cells were not induced by the

RU486-O₃-FK506 conjugate. The lack of induction was evident after both 6 and 24 hours and within the treatment range of 1 - 100 nM of the conjugate. Two additional genes were included in the study, SDPR and SLC19A2, because these genes have GREs proximal to their transcriptional start sites, and they may resemble the luciferase reporters in this manner. The genes in the stable clone (124.25), which expresses lower levels of the fusion protein, were also not inducible by RU486-O₃-FK506. The SDex-O₂-FK506 conjugate was also tested and showed higher potency than its linker-only analog, but this activity was not dependent on the VP16-FKBP fusion protein. It may instead be through an effect of linked FK506 on cell permeability or through interaction with endogenous FKBP5.

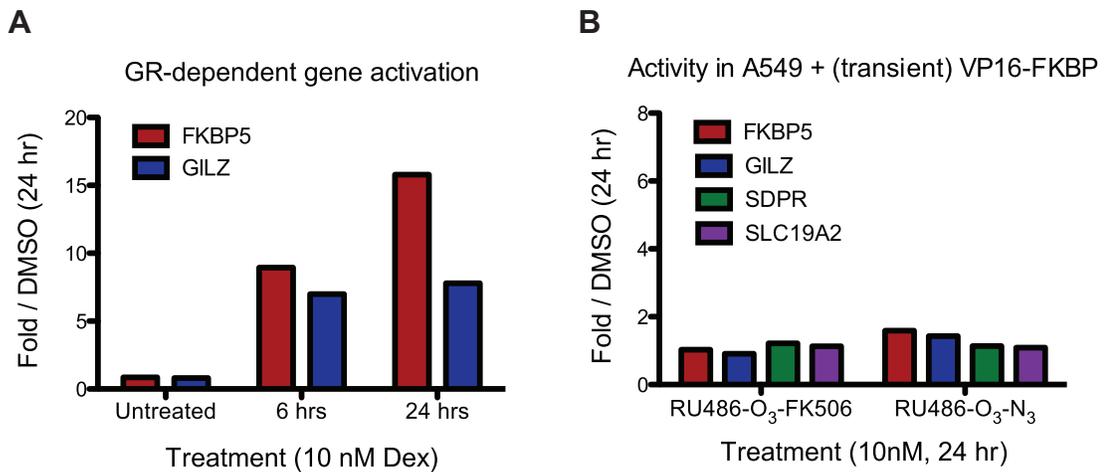


Figure 3.14 Effect of RU486-O₃-FK506 on expression of endogenous genes mRNA levels are determined using RT-qPCR and comparative $\Delta\Delta C_T$ analysis. Determined C_T levels for genes are normalized to RPL19 and ΔC_T values of samples are compared to ΔC_T of DMSO. Fold over DMSO is: $2^{(-\Delta\Delta C_T)}$ A. The endogenous genes, FKBP5 and GILZ, are Dex-responsive in A549 cells; B. The effect of RU486-O₃-FK506 on Dex-responsive genes in A549 cells transiently expressing the VP16-FKBP-myc₆ fusion protein. Data represent the average of three qPCR samples prepared with mRNA isolated from one experiment.

In an attempt to detect recruitment of fusion the protein at endogenous genes chromatin immunoprecipitation (ChIP) was used. Recruitment may be occurring

even though no effects on transcription of genes were seen. CHIP showed that Dex induces binding of endogenous GR to GREs in A549 cells (Figure 3.15B). When a myc₆-tagged GR is expressed in A549 cells, CHIP with an α -myc antibody also detects the Dex-induced binding of GR to the GREs of the FKBP5 gene (Figure 3.15A). In A549 cells with transiently expressed VP16-FKBP-myc₆, the SDex-conjugates were tested and induce GR-binding to GREs of the GILZ gene (Figure 3.15C) and the FKBP5 gene (Figure 3.13D). The VP16-FKBP-myc₆ fusion protein, however, could not be detected (Figure 3.13C-D). The absence of fusion protein at the promoter is consistent with lack of VP16-FKBP dependent transcriptional activity, but it is not conclusive evidence; VP16-FKBP would be predicted to be more difficult to cross-link to the promoter, since it would be tethered to GR via a chemical spacer, and be expected to lack direct association with DNA and maybe even with other DNA-bound proteins.

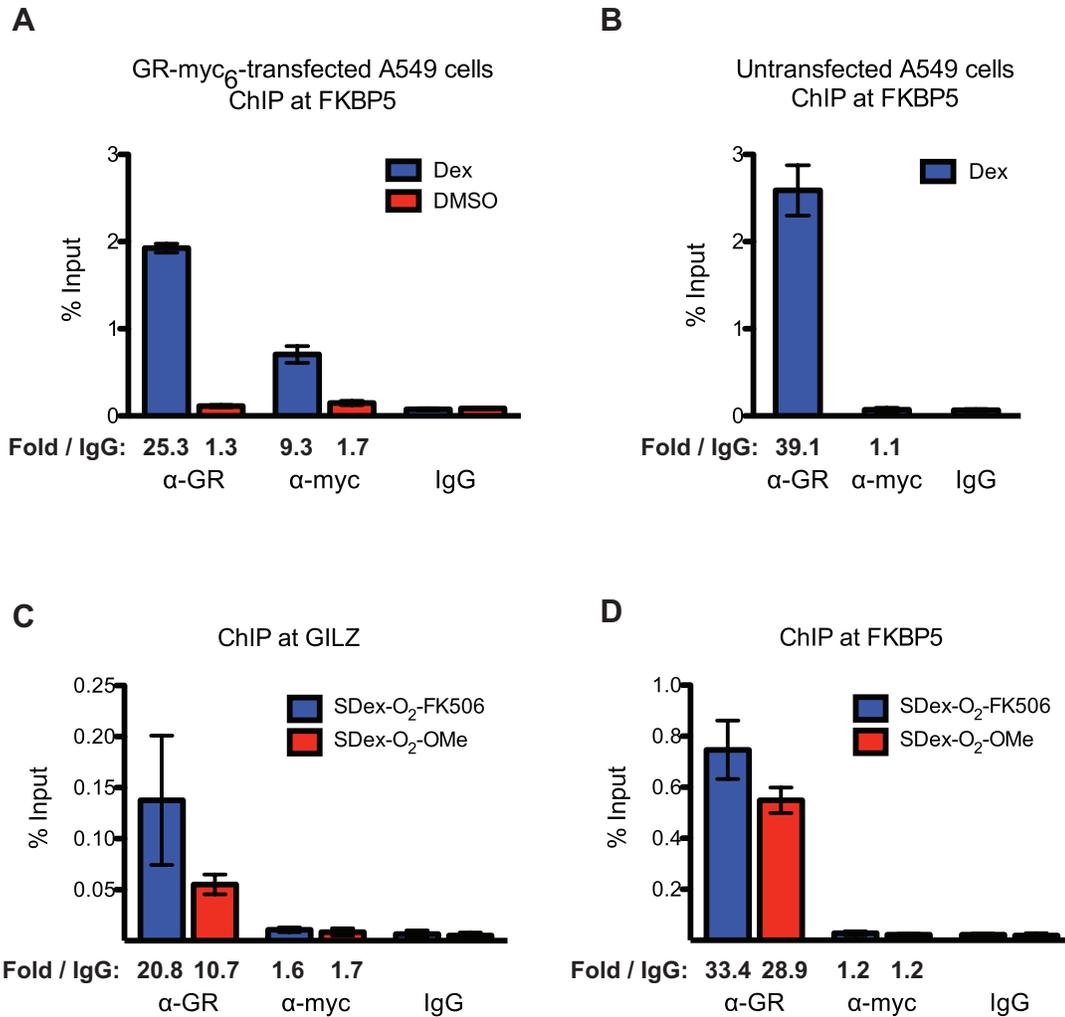


Figure 3.15 Chromatin immunoprecipitation of GR and VP16-FKBP-myc₆
Chromatin immunoprecipitation at GRE-regions of FKBP5 (A,B,D) or GILZ (C) enhancers. A. ChIP of Dex or DMSO treated A549 cells transiently transfected with GR-myc₆ using α-GR and α-Myc antibodies; B. ChIP of Dex or DMSO treated untransfected A549 cells using α-GR antibody; C and D. ChIP of SDex-conjugate treated A549 cells transiently transfected with VP16-FKBP-myc₆ using α-GR and α-Myc antibodies.

3.6 Conclusion and discussion

Conjugation of GR agonist ligand dexamethasone and antagonist ligand RU486 to FK506 facilitated recruitment of an FKBP-fusion protein to GR. When SDex-O₂-FK506 was used to recruit VP16-FKBP, the transcriptional activation was higher than full agonistic effects. VP16 may provide the sole activation potential in place of an intrinsic effect, or both mechanisms may contribute. Conjugation of

antagonistic RU486 to FK506 also facilitated recruitment of VP16-FKBP to a reporter gene. The intrinsic activation potential of RU486 is very low, and RU486 potentially antagonizes Dex (Figure 3.12). The activity induced by recruitment of VP16-FKBP is therefore likely entirely an effect of VP16, and this activity is nearly the same as the full agonistic effect of Dex. The activation potential of VP16 may be higher, as a self-inhibition effect of RU486-O₃-FK506 overlaps and becomes dominant at high concentrations. Additionally, it is possible that a repressive mechanism of RU486-bound GR suppresses a higher activation potential of VP16 to the level observed. When Dex and RU486-O₃-FK506 are added simultaneously to cells transfected with VP16-FKBP, an activation level that is higher than from either compound alone is seen (Figure 3.9, right panel). This can best be explained to be a result of the two GREs in the reporter being occupied by GR dimers with a combination of both Dex and RU486-O₃-FK506 bound (a heterodimer or two different homodimers) consistent with a scenario in which intrinsic GR agonism (from Dex-bound GR) can synergize with activity from VP16.

The different dose response curves for SDex and RU486 conjugates are not a result of different affinities for GR. FK506 was functionalized as has been previously reported, and it has been documented for FK506 conjugates made via this functionalization that they do not alter the binding affinity for FKBP.⁴² We have therefore assumed that the binding affinities (for FKBP) are identical for the SDex and RU486 conjugates, but this would be useful to confirm. A differential binding affinity could also stem from different interactions between FKBP and GR depending on the ligand bound to GR. The binding studies (Figure 3.9), however, were done in lysates that contain endogenous FKBP1A. Another factor that may influence the dose response, is an allosterically altered association of GR with DNA that is different for Dex and RU486.⁵⁶

Collaborative efforts are still ongoing aimed at demonstrating functional recruitment of VP16-FKBP to endogenous genes. The inability of the conjugates

to show activity in A549 cells may be related to the relative levels of both GR and VP16-FKBP. In addition to studies of endogenous genes presented in Figure 3.14, preliminary experiments with transfection of GRE-luciferase reporter and VP16-FKBP into A549 cells did not demonstrate a VP16-FKBP dependent activity of conjugates (J.W. Hojfeldt and J.P. Carolan). This at least indicates a challenge in A549 cells not related to a difference between recruitment to plasmid DNA and endogenous genes. A lysate of VP16-FKBP transfected HEK293T cells does not show high levels of fusion protein relative to endogenous FKBP1A by Western blot (O. Cruz and J.P. Carolan), and this may implicate differential levels of GR expression as a cause. A high level of GR expression may allow conjugates to bind GR first and translocate to the nucleus, where it can bind VP16-FKBP which, unlike FKBP1A, has a nuclear localization sequence added. Studies of effects on endogenous genes in HEK293T cells (transfected with GR and VP16-FKBP as in reporter assays) are being carried out

The demonstration of direct recruitment of proteins to a full-length nuclear receptor with bifunctional molecules is an encouraging result. If this strategy can be further extended to conjugates that combine NR ligands and ligands of endogenous transcriptional coregulatory complexes, NR targeting drugs with a new level of transcription control could be realized. The first efforts towards this are presented in Chapter 4. The ongoing efforts to understand the factors associated with different activities in HEK293T and A549 cells presented in this chapter for VP16-FKBP recruitment will hopefully provide lessons beneficial to efforts with other bifunctional recruiters.

3.7 Materials and methods

Cells

HeLa (CCL-2) and A549 (CCL-185) cells were purchased from ATCC. Cells were cultured in DMEM with 10% FBS at 37°C under 5% CO₂. Transient transfections were done using Lipofectamine LTX or Lipofectamine 2000 (Invitrogen) with

conditions optimized using GFP expression plasmid and Accuri C6 flow cytometric analysis.

An A549-derived cell line stably transfected with pVP16-FKBP-myc₆ was made (clone 124.25): 160,000 A549 cells were plated in wells of a 6 well plate and transfected the following day by addition of: 2.5 μg linearized (Sall) pVP16-FKBP-myc₆, 0.125 μg linear hygromycin marker (Clontech), 2.5 μL Plus reagent, 10 μg Lipofectamine LTX in 0.5 mL OptiMEM. After 24 hours cells were trypsinized and replated in 150 cm dishes at 100,000 cells per dish. The following day 400 μg/mL of hygromycin B (Calbiochem) was added (concentration determined from kill-curve experiment). Media with hygromycin is replaced every 2-3 days. After 10 days, most cells are dead (all cells on control plates are dead) and a total of 37 colonies from 4 plates are picked by use cloning cylinders and transferred to 12-well plates. When cells grow to near confluency they are split to larger dishes, and lysates used to detect expression of fusion protein by Western blot.

Mammalian three-hybrid assay (Figures 3.3, 3.4, and 3.5)

12,000 HeLa cells were plated in a 96-well plate in 100 μL DMEM. The following day they were transfected by replacing media with transfection mix: 100 ng pG5luc, 1 ng pRLSV40, 50 ng pGal4-GRLBD, 50 ng pVP16-FKBP, 0.6 μL Lipofectamine 2000 in 100 μL OptiMEM. After 6 hours, the transfection mix was replaced with DMEM + 10% FBS and 1 μL of a DMSO stock solution of compound was added. After 20 hours, media was removed and samples were assayed with Dual-Luciferase Reporter Assay System (Promega) using 20 μL Passive lysis buffer, and 25 μL of each substrate. Luminescence was measured on Berthold Lumat LB 9507 luminometer. Standard deviation from triplicate samples is represented by error bars on graph.

Western blots

Lysates were made in RIPA buffer and quantified with BCA assay (Pierce). SeeBlue Plus2 and Magic Mark XP protein ladders (Invitrogen) were used to estimate molecular weights. SDS-PAGE was done with bis-tris gels (0.35M bis-tris (Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane) pH 6.5, 10 % acrylamide (30:2)) run in MOPS buffer at 150 V and transferred to PVDF membrane. Membranes were stained with Ponceau S, then blocked with 5% milk in PBST, 1hr at 4°C. Primary antibodies were incubated over night at 4°C at dilutions: 1:1000 α -myc (Santa Cruz, sc-40) or 1:2000 α -FKBP (Abcam, ab58072). Membrane was washed with PBST and incubated with 1:10000 goat α -mouse-HRP (Santa Cruz, sc-2005). Membrane was washed with PBST, developed with ECL-Plus (GE Health Care) and read on Typhoon 9410. Scanned images were analyzed using ImageJ software.

Chromatin immunoprecipitation

A549 cells in 10cm dishes (10 mL DMEM + 10% FBS) were treated with DMSO or compounds at 90% confluency. After incubation with compounds formaldehyde was added at room temperature to a final concentration of 1% and incubated for 10 min. Glycine was added to 0.125 M and incubated for 5 min. before cells were washed 3 times with PBS. Cells were lysed on ice by addition of 400 μ L SDS lysis buffer (1%SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0, Halt protease inhibitor (Pierce)) and incubated for 15 min. before cells were scraped, transferred to a microcentrifuge tube and placed on ice for 30 minutes. The lysate was sonicated per optimized conditions to yield sheared chromatin with a size range of 200-800bp. Sheared chromatin was cleared by centrifugation (10 min., 12,000 x g, 4°C). A sample of sheared chromatin was reverse cross-linked (see below) and DNA concentration determined. Sheared chromatin corresponding to 10 μ g of DNA was used for each IP and was diluted with 9 parts ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, with Halt protease inhibitor). 0.1 μ g DNA was kept as 1% input sample (treated with IP'ed samples at the elution step below). 3

μ g of antibody was added and was incubated over night on rotating wheel at 4°C. Antibodies used were: α -GR (Santa Cruz, sc-8992x); α -myc (Santa Cruz, sc-40); IgG (Santa Cruz, sc-2027x). Next day, 15 μ L Protein G MagnaChIP beads (Millipore) were added and incubated for 1 hr at 4°C. Supernatant was removed from beads with tubes placed in magnetic rack, and 1 mL wash buffer was added followed by 5 min. on orbital shaker at 4°C. The beads were washed once with each of the following buffers: 1) 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl; 2) 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl; 3) 250 mM LiCl, 1% Igepal CA630, 1% deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.1; 4) 10mM Tris-HCl pH 8.0, 1 mM EDTA. Chromatin was eluted from beads by 2 incubations with 100 μ L Elution buffer (1% SDS, 0.1 M NaHCO₃). 1 μ L 10 mg/mL RNase A was added and incubated for 30 min. at 37°C. 8 μ L 5M NaCl and 1 μ L 10 mg/mL Proteinase K was added and incubated at 65°C over night. DNA was purified with Qiagen PCR clean-up kit and used in qPCR reactions.

qPCR of ChIP samples

20 μ L qPCR reaction mix: 2 μ L DNA, 0.2 μ L CXR, 0.4 μ L forward primer (200 nM final), 0.4 μ L reverse primer, 7 μ L H₂O, 10 μ L GoTaq qPCR master mix (Promega). qPCR was done on ABI StepOne Plus. Products were analyzed with melt curve and on agarose gel. Data analyzed with $\Delta C_T = C_T(\text{sample}) - C_T(\text{input})$, % of input = $2^{-(\Delta C_T)}$. Standard error calculated on % of input values. Primer pairs used:

(FKBP5 intronic enhancer, 350 bp)

F-Pr 5'- CAGAGCTAATGTCTTTAGGCTGGAGC -3'

R-Pr 5'- GCAATCGGAGTGTAACCACATCAAG -3'

(GILZ, surrounds GRE -1958/-1944, 151 bp)

F-Pr 5'- GGCCCCAGTACTTTTCCAAT -3'

R-Pr 5'- GGTTGAGTCCTGGTTTCCTC -3'

RT-qPCR for mRNA

50,000 A549 (WT, 124.25, or transiently transfected) cells were plated in 24-well plate with 0.5 mL DMEM + 10% FBS. Next day cells were treated with compounds or DMSO (0.2% final DMSO for all samples). Total RNA was isolated with Qiagen RNeasy Plus mini kit. cDNA synthesis was done with iScript RT Mastermix (Biorad). 2 μ L cDNA was used in qPCR as described for ChIP samples. Primer pairs used:

FKBP5: F-Pr 5'- GGAATGGTGAGGAAACGCCG -3'

FKBP5: R-Pr 5'- CTCTCCTTTCTTCATGGTAGCCACC -3'

GILZ: F-Pr 5'- CGAACAGGCCATGGATCTGGTGAA -3'

GILZ: R-Pr 5'- AGAACCACCAGGGGCCTCGG -3'

RPL19: F-Pr 5'- ATGTATCACAGCCTGTACCTG -3'

RPL19: R-Pr 5'- TTCTTGGTCTCTTCCTCCTTG -3'

SLC19A2: F-Pr 5'- AAGTCCCGCAAGGTCAGCGC -3'

SLC19A2: R-Pr 5'- GAAACGGCCTCCACGCCACC -3'

SDPR: F-Pr 5'- AAGTCCCGCAAGGTCAGCGC -3'

SDPR: R-Pr 5'- GCAGGGATCTCATTTTCCTCCTGG -3'

Plasmids

pG5luc, pRLSV40, and pAct were purchased from Promega. pRShGR was purchased from ATCC. Plasmids with CDS for CyclinT1⁵⁷ and Cdk9⁵⁸ were purchased from Addgene (14628 and 14640). pGal4-GRLBD, which is a mammalian expression plasmid that encodes a fusion of yeast Gal4(1-147) and the ligand binding domain of the human glucocorticoid receptor (499-777), was a kind gift from Thomas Kodadek (The Scripps Research Institute, Scripps Florida).⁵⁹ pCS2+MT is an expression vector designed for making fusions to six copies of the myc epitope⁶⁰ recognized by the 9e10 monoclonal antibody,⁶¹ and was a kind gift from David Turner (University of Michigan). All oligos for construction of plasmids were purchased from Integrated DNA Technologies.

pVP16-FKBP is a mammalian expression plasmid that encodes a fusion of Gal1-11 (for efficient expression), NLS from SV40 large T antigen (nuclear localization signal), VP16(411-456) (potent activation domain, activation domain 1 from the herpes simplex virion protein 16), and human FKBP1A(2-107) (FK506 binding protein 1A, 12kDa). It was constructed by cloning of FKBP1A into pAct. The coding sequence is shown below, with the elements listed above marked with capital letters in brackets:

```
[ATGAAGCTACTGTCTTCTATCGAACAAGCATGC][CCAAAAAAGAAGAGAAA
GGTAGAT]gaattcccgggg[ATCTCGACGGCCCCCGACCGATGTCAGCCTGG
GGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCGA
CGCGCTAGACGATTTTCGATCTGGACATGTTGGGGGACGGGGATTCCCCGG
GTCCGGGA]tcgccaggatcc[GGCGTGCAGGTGGAGACTATCTCCCCAGGAGAC
GGGCGCACCTTCCCCAAGCGCGGCCAGACCTGCGTGGTGCCTACACCGG
GATGCTTGAAGATGGAAAGAAATTTGATTCTCCCGGGACAGAAACAAGCC
CTTTAAGTTTATGCTAGGCAAGCAGGAGGTGATCCGAGGCTGGGAAGAAGG
GGTTGCCAGATGAGTGTGGGTCAGAGAGCCAACTGACTATATCTCCAGA
TTATGCCTATGGTGCCACTGGGCACCCAGGCATCATCCCACCACATGCCAC
TCTCGTCTTCGATGTGGAGCTTCTAAAACCTGGAAAGATCTGTGCGACTTGACG
CGT]
```

pLIC-myc₆ is an LIC cloning⁶² vector to facilitate fusion of genes to six myc-tags in a mammalian expression vector. It was prepared by performing site-directed mutagenesis to pCS2+MT with the primers: 5'- CAA GCT ACT TGT TCT TTT TGC ACC ATG GGA AGC ACC GGT TCT GGT GAG ATG GAG CAA AAG CTC ATT TCT G -3' and 5'- CAG AAA TGA GCT TTT GCT CCA TCT CAC CAG AAC CGG TGC TTC CCA TGG TGC AAA AAG AAC AAG TAG CTT G -3'. To prepare the vector for LIC cloning, it was digested with AgeI and processed with T4 DNA polymerase (Novagen) with dTTP added as the only nucleotide. Genes to be cloned into the vector were PCR amplified with primers that have sequences appended to 5'end of gene specific sequence. These appendages are: Sense-primer 5'- GGGAAGCACCGGT; Antisense-primer 5'- CACCAGAACCGGT.

pVP16-FKBP-myc₆ was prepared by performing LIC-cloning of the fusion protein from pVP16-FKBP into pLIC-myc₆. This fused Gal4(1-11)-NLS-VP16-FKBP N-terminal to 6 myc epitopes in a mammalian expression vector.

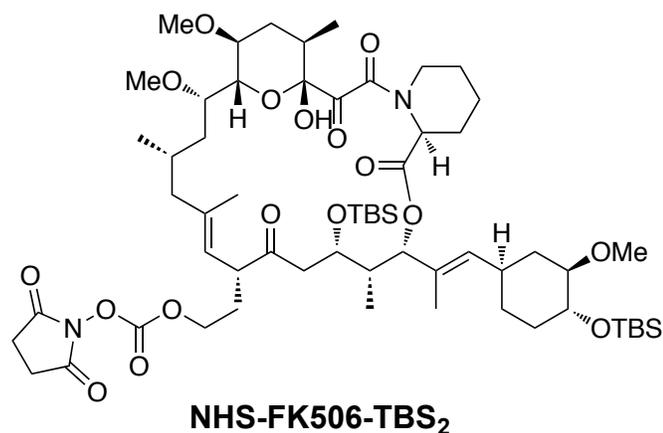
pGR-myc₆ was prepared with LIC of human GR from pRShGR into pLIC-myc₆.

pLIC-FKBP is a cloning vector designed for making fusions to FKBP (and Gal4(1-11) and SV40 large T NLS). This was made by cutting VP16 out of pVP16-FKBP with EcoRI and BamH1, and ligating in an LIC cassette in the form of annealed oligos: 5'- AAT TGG GAA GCA CCG GTT CTG GTG ATC -3' and 5'- GAT CGA TCA CCA GAA CCG GTG CTT CCC -3'. To prepare the vector for LIC cloning, it was digested with AgeI and processed with T4 DNA polymerase (Novagen) with dTTP added as the only nucleotide. Genes to be cloned into the vector were PCR amplified with primers that have sequences appended to 5'end of gene specific sequence. These appendages are: Sense-primer 5'-GGGAAGCACCGGT; Antisense-primer 5'- CACCAGAACCGGT.

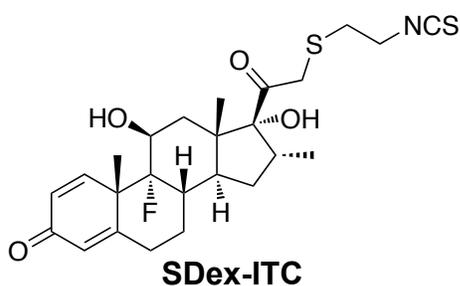
pCyclinT1-FKBP and pCdk9-FKBP were made by LIC of the genes in plasmids purchased from Addgene into pLIC-FKBP.

Chemical synthesis

FK506 was purchased from LC Laboratories. Dexamethasone was purchased from Enzo Life Sciences. RU486 (Mifepristone) was purchased from Sigma-Aldrich. Commercially available reagents and solvents were used as received. Chromatographic separations were carried out on silica gel 60 (230-400 mesh, E. Merck) or by reverse phase HPLC on C18 column using the indicated eluents. Yields are unoptimized. ESI-MS spectra were obtained on Micromass LCT TOF mass spectrometer. High-resolution mass spectra (HRMS) were obtained on Micromass AutoSpec Ultima Magnetic sector mass spectrometer. ¹H-NMR spectra were obtained at 400MHz on a Varian MR-400 spectrometer. Chemical shifts are given in δ(ppm) values.



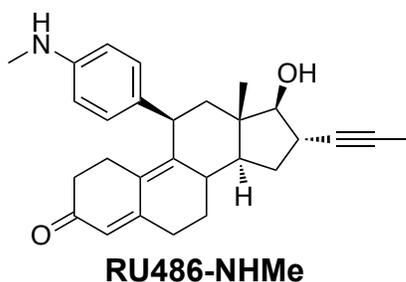
NHS-FK506-TBS₂ was synthesized by Y. Imaeda according to a published procedure.⁴²



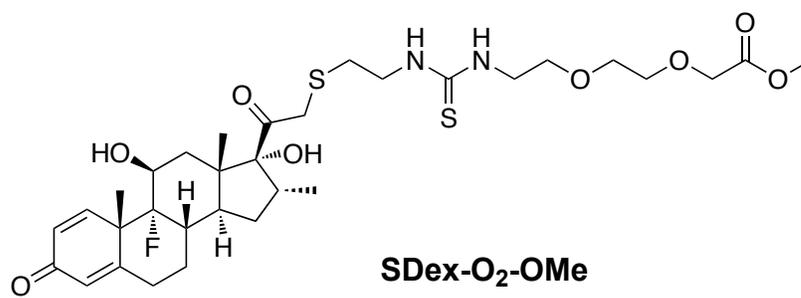
SDex-ITC was synthesized by J.W. Hojfeldt according to a published procedure.²⁹

Dexamethasone (1.00 g, 2.5 mmol) was dissolved in pyridine (7 mL) and cooled to 0°C. Methanesulfonyl chloride (0.26 mL, 3.3 mmol) was added to the mixture dropwise over 15 min. After stirring at 0°C for 30 min, the reaction mixture was poured into water (300 mL) at 0°C. The precipitation was collected by filtration and washed with water. The obtained product and tert-butyl 2-sulfanylethylcarbamate (1.24 mL, 7.3 mmol) were dissolved in acetone (24 mL). Triethylamine (2.05 mL, 7.3 mmol) was added to the solution and then the resulting mixture was stirred at room temperature for 15 h. The reaction mixture was diluted with water and extracted with ethyl acetate. The extract was washed with water and brine, dried over sodium sulfate and concentrated in vacuo. The residue was crystallized from ethyl acetate-hexane to give a colorless powder.

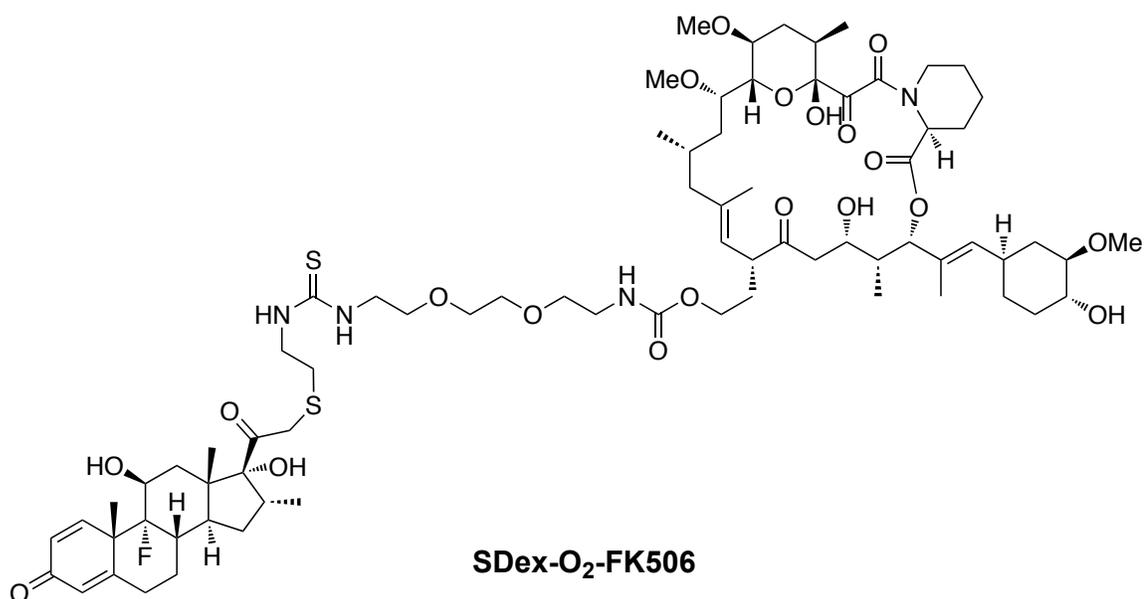
The obtained powder was dissolved in 1.5 M hydrochloric acid in acetic acid (1 mL) and then the mixture was stirred at room temperature for 15 min. The mixture was diluted with water and chloroform and basified with saturated aqueous sodium carbonate solution and the slurry was stirred at room temperature for 15 min (pH of the slurry should be over 10). Thiophosgene (0.18 mL, 2.4 mmol) was added to the slurry, and then the resulting mixture was stirred at room temperature for 1.5 h. The chloroform layer was separated, and the layer was washed with water and brine, dried over sodium sulfate and concentrated in vacuo. The residue was purified by silica gel chromatography (ethyl acetate/hexane = 1/4 to 1/1) to give product (150 mg, 18%) as a pale yellow powder. ¹H NMR (400MHz, CDCl₃): δ 0.88 (3H, d, J = 7.2 Hz), 1.02-1.80 (10H, m), 2.01-2.57 (6H, m), 2.81 (2H, t, J = 6.8 Hz), 3.04-3.07 (1H, m), 3.24 (1H, d, J = 13.6 Hz), 3.58 (1H, d, J = 14.0 Hz), 3.71 (2H, t, J = 5.8 Hz), 4.34 (2H, d, J = 7.6 Hz), 6.08 (1H, s), 6.29 (1H, d, J = 8.4 Hz), 7.14 (1H, d, J = 10 Hz). ESI-MS calculated for [C₂₅H₃₂FNO₄S₂ + H]⁺: 494.1, found 494.0.



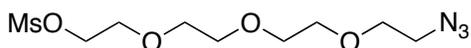
RU486-NHMe was synthesized by A.R. Van Dyke and Y. Imaeda according to published procedure.³³



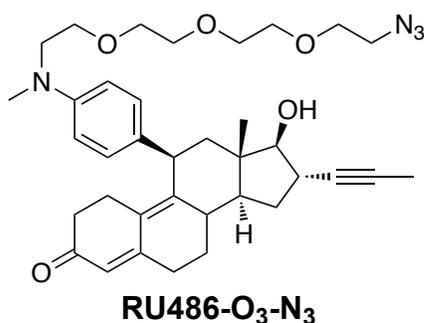
SDex-O₂-OMe was synthesized by A.R. Van Dyke.



SDex-O₂-FK506 was synthesized by Y. Imaeda as outlined in Figure 3.2A.



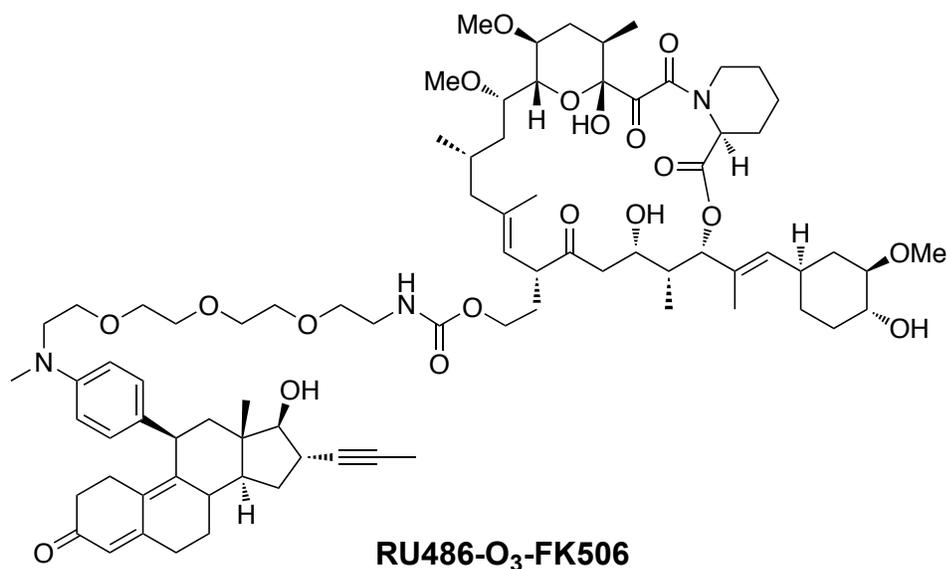
Linker MsO-O₃-N₃ was synthesized by A.R. Van Dyke and is used in the RU486-conjugate synthesis below.



RU486-O₃-N₃ was synthesized by J.W. Hojfeldt and Y. Imaeda:

A mixture **RU486-NHMe** (22 mg, 0.053 mmol), and **Linker MsO-O₃-N₃** (19 mg, 0.064 mmol), sodium iodide (10 mg, 0.064 mmol), and N,N-diisopropylethylamine (0.03 mL, 0.16 mmol) in acetonitrile (5 mL) was heated at 100°C for 60 h. After concentration in vacuo, the residue was purified by silica gel chromatography

(ethyl acetate/hexane = 4/6 to 6/4) to give compound 7 (22 mg, 67%) as a yellow sticky oil. Product was purified by HPLC (0.1% aqueous TFA/acetonitrile = 80/20 to 65/35) to give the trifluoroacetic acid salt of compound 7 as a colorless amorphous powder. ^1H NMR (400MHz, CDCl_3): δ 0.53 (3H, s), 1.31-1.47 (2H, m), 1.65-1.78 (3H, m), 1.88 (3H, s), 1.89-2.03 (2H, m), 2.18-2.47 (7H, m), 2.55 (2H, m), 2.74-2.77 (1H, m), 2.93 (3H, s), 3.36-3.78 (2H, m), 3.46-3.48 (1H, m), 3.59-3.66 (12H, m), 4.32 (1H, d, $J = 6.8$ Hz), 5.74 (1H, s), 6.61 (2H, d, $J = 8.8$ Hz), 6.97 (2H, d, $J = 8.4$ Hz). HRMS (ESI) calculated for $[\text{C}_{36}\text{H}_{48}\text{N}_4\text{O}_5 + \text{H}]^+$: 617.3697, found 617.3699.



RU486-O₃-FK506 was synthesized by J.W. Hojfeldt and Y. Imaeda:

A mixture of **RU486-O₃-N₃** (13 mg, 0.021 mmol) and polymer-bound triphenylphosphine (200-400 mesh, 3 mmol/g; 25 mg, 0.075 mmol) and water (1 drop) in THF (3 mL) was refluxed for 15 h. The mixture was filtered and the filtrate was concentrated in vacuo. The product, compound 12 (25 mg, 0.021 mmol), and N,N-diisopropylethylamine (0.008 mL, 0.042 mmol) were dissolved in dichloromethane (1 mL). After stirring at room temperature for 40 h, the mixture was purified by silica gel chromatography (ethyl acetate/hexane = 1/1 to ethyl acetate). The product was dissolved in methanol (0.5 mL) and water (1 drop) and concentrated hydrochloric acid (3 drops) was added. After stirring at room

temperature for 2 h, the reaction mixture was neutralized with NaHCO₃ powder and concentrated in vacuo. The mixture was diluted with acetonitrile and filtered. The filtrate was concentrated in vacuo and the residue was purified by HPLC (0.1% aqueous TFA/acetonitrile = 65/35 to 50/50) to give the product as a colorless amorphous powder. ¹H NMR (400MHz, CD₃OD): δ 0.47 (3H, s), 0.85-0.98 (10H, m), 1.04-3.01 (58H, m), 3.11-3.68 (30H, m), 3.91-4.35 (5H, m), 4.53 (1H, m), 4.62 (1H, br s), 4.85-5.24 (4H, m), 5.74 (1H, s), 7.56 (2H, m), 7.82 (2H, m). HRMS (ESI) calculated for [C₈₀H₁₁₇N₃O₁₉ + H]⁺: 1424.8354, found 1424.8345.

Other conjugates of varying linker lengths described in Figure 3.4 were synthesized by Y. Imaeda.

3.8 References

1. Wrangle, O. & Gustafsson, J.A. Separation of the hormone- and DNA-binding sites of the hepatic glucocorticoid receptor by means of proteolysis. *J Biol Chem* **253**, 856-65 (1978).
2. Nagy, L. & Schwabe, J.W. Mechanism of the nuclear receptor molecular switch. *Trends Biochem Sci* **29**, 317-24 (2004).
3. Meijsing, S.H. et al. DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407-10 (2009).
4. Ismaili, N. & Garabedian, M.J. Modulation of glucocorticoid receptor function via phosphorylation. *Ann N Y Acad Sci* **1024**, 86-101 (2004).
5. Wurtz, J.M. et al. A canonical structure for the ligand-binding domain of nuclear receptors. *Nat Struct Biol* **3**, 87-94 (1996).
6. Rosenfeld, M.G., Lunyak, V.V. & Glass, C.K. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev* **20**, 1405-28 (2006).
7. Laudet, V. & Gronemeyer, H. *The nuclear receptor factsbook*, xvii, 462 p. (Academic, San Diego, Calif., 2002).
8. Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat Rev Drug Discov* **5**, 993-6 (2006).

9. Gronemeyer, H., Gustafsson, J.A. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3**, 950-64 (2004).
10. Bruning, J.B. et al. Coupling of receptor conformation and ligand orientation determine graded activity. *Nat Chem Biol* **6**, 837-43 (2010).
11. De Bosscher, K. Selective Glucocorticoid Receptor modulators. *J Steroid Biochem Mol Biol* **120**, 96-104 (2010).
12. Clarke, R. et al. The inter-relationships between ovarian-independent growth, tumorigenicity, invasiveness and antioestrogen resistance in the malignant progression of human breast cancer. *J Endocrinol* **122**, 331-40 (1989).
13. Clarke, R. et al. Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. *Oncogene* **22**, 7316-39 (2003).
14. Knudsen, K.E. & Scher, H.I. Starving the addiction: new opportunities for durable suppression of AR signaling in prostate cancer. *Clin Cancer Res* **15**, 4792-8 (2009).
15. Le Drean, Y., Mincheneau, N., Le Goff, P. & Michel, D. Potentiation of glucocorticoid receptor transcriptional activity by sumoylation. *Endocrinology* **143**, 3482-9 (2002).
16. Tian, S., Poukka, H., Palvimo, J.J. & Janne, O.A. Small ubiquitin-related modifier-1 (SUMO-1) modification of the glucocorticoid receptor. *Biochem J* **367**, 907-11 (2002).
17. Holmstrom, S.R., Chupreta, S., So, A.Y. & Iniguez-Lluhi, J.A. SUMO-mediated inhibition of glucocorticoid receptor synergistic activity depends on stable assembly at the promoter but not on DAXX. *Mol Endocrinol* **22**, 2061-75 (2008).
18. Daniel, A.R. & Lange, C.A. Protein kinases mediate ligand-independent derepression of sumoylated progesterone receptors in breast cancer cells. *Proc Natl Acad Sci U S A* **106**, 14287-92 (2009).
19. Zhou, J. & Cidlowski, J.A. The human glucocorticoid receptor: one gene, multiple proteins and diverse responses. *Steroids* **70**, 407-17 (2005).
20. Kumar, R. & Thompson, E.B. The structure of the nuclear hormone receptors. *Steroids* **64**, 310-9 (1999).
21. Pratt, W.B. & Toft, D.O. Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr Rev* **18**, 306-60 (1997).

22. Grad, I. & Picard, D. The glucocorticoid responses are shaped by molecular chaperones. *Mol Cell Endocrinol* **275**, 2-12 (2007).
23. So, A.Y., Chaivorapol, C., Bolton, E.C., Li, H. & Yamamoto, K.R. Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet* **3**, e94 (2007).
24. Konig, H., Ponta, H., Rahmsdorf, H.J. & Herrlich, P. Interference between pathway-specific transcription factors: glucocorticoids antagonize phorbol ester-induced AP-1 activity without altering AP-1 site occupation in vivo. *EMBO J* **11**, 2241-6 (1992).
25. Nissen, R.M. & Yamamoto, K.R. The glucocorticoid receptor inhibits NFkappaB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* **14**, 2314-29 (2000).
26. Barnes, P.J. Corticosteroids: the drugs to beat. *Eur J Pharmacol* **533**, 2-14 (2006).
27. Licitra, E.J. & Liu, J.O. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc Natl Acad Sci U S A* **93**, 12817-21 (1996).
28. Arth, G.E. et al. 16-METHYLATED STEROIDS. II. 16 α -METHYL ANALOGS OF CORTISONE, A NEW GROUP OF ANTI-INFLAMMATORY STEROIDS. 9 α -HALO DERIVATIVES. *Journal of the American Chemical Society* **80**, 3161-3163 (1958).
29. Lopez, S. & Simons, S.S., Jr. Dexamethasone 21-(beta-isothiocyanatoethyl) thioether: a new affinity label for glucocorticoid receptors. *J Med Chem* **34**, 1762-7 (1991).
30. Liu, B., Yu, P., Alluri, P.G. & Kodadek, T. Simple reporter gene-based assays for hairpin poly(amide) conjugate permeability and DNA-binding activity in living cells. *Mol Biosyst* **1**, 307-17 (2005).
31. Belanger, A., Philibert, D. & Teutsch, G. Regio and stereospecific synthesis of 11 beta-substituted 19-norsteroids. Influence of 11 beta-substitution on progesterone receptor affinity - (1). *Steroids* **37**, 361-82 (1981).
32. Ito, K., Jazrawi, E., Cosio, B., Barnes, P.J. & Adcock, I.M. p65-activated histone acetyltransferase activity is repressed by glucocorticoids: mifepristone fails to recruit HDAC2 to the p65-HAT complex. *J Biol Chem* **276**, 30208-15 (2001).
33. von Geldern, T.W. et al. Liver-selective glucocorticoid antagonists: a novel treatment for type 2 diabetes. *J Med Chem* **47**, 4213-30 (2004).

34. Kino, T. et al. FK-506, a novel immunosuppressant isolated from a *Streptomyces*. I. Fermentation, isolation, and physico-chemical and biological characteristics. *J Antibiot (Tokyo)* **40**, 1249-55 (1987).
35. Tanaka, H. et al. Structure of FK506, a novel immunosuppressant isolated from *Streptomyces*. *Journal of the American Chemical Society* **109**, 5031-5033 (1987).
36. Kay, J.E. Structure-function relationships in the FK506-binding protein (FKBP) family of peptidylprolyl cis-trans isomerases. *Biochem J* **314** (Pt 2), 361-85 (1996).
37. Siekierka, J.J., Staruch, M.J., Hung, S.H. & Sigal, N.H. FK-506, a potent novel immunosuppressive agent, binds to a cytosolic protein which is distinct from the cyclosporin A-binding protein, cyclophilin. *J Immunol* **143**, 1580-3 (1989).
38. Siekierka, J.J., Hung, S.H., Poe, M., Lin, C.S. & Sigal, N.H. A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. *Nature* **341**, 755-7 (1989).
39. Harding, M.W., Galat, A., Uehling, D.E. & Schreiber, S.L. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. *Nature* **341**, 758-60 (1989).
40. Fretz, H. et al. Rapamycin and FK506 binding proteins (immunophilins). *Journal of the American Chemical Society* **113**, 1409-1411 (1991).
41. Aldape, R.A. et al. Charged surface residues of FKBP12 participate in formation of the FKBP12-FK506-calcineurin complex. *J Biol Chem* **267**, 16029-32 (1992).
42. Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019-24 (1993).
43. Amara, J.F. et al. A versatile synthetic dimerizer for the regulation of protein-protein interactions. *Proc Natl Acad Sci U S A* **94**, 10618-23 (1997).
44. Abida, W.M., Carter, B.T., Althoff, E.A., Lin, H. & Cornish, V.W. Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system. *ChemBiochem* **3**, 887-95 (2002).
45. Riggs, D.L. et al. The Hsp90-binding peptidylprolyl isomerase FKBP52 potentiates glucocorticoid signaling in vivo. *EMBO J* **22**, 1158-67 (2003).

46. Mack, E.T., Perez-Castillejos, R., Suo, Z. & Whitesides, G.M. Exact analysis of ligand-induced dimerization of monomeric receptors. *Anal Chem* **80**, 5550-5 (2008).
47. Barnes, P.J. Anti-inflammatory actions of glucocorticoids: molecular mechanisms. *Clin Sci (Lond)* **94**, 557-72 (1998).
48. McKay, L.I. & Cidlowski, J.A. Molecular control of immune/inflammatory responses: interactions between nuclear factor-kappa B and steroid receptor-signaling pathways. *Endocr Rev* **20**, 435-59 (1999).
49. Zhao, Q., Pang, J., Favata, M.F. & Trzaskos, J.M. Receptor density dictates the behavior of a subset of steroid ligands in glucocorticoid receptor-mediated transrepression. *Int Immunopharmacol* **3**, 1803-17 (2003).
50. Bouwmeester, T. et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* **6**, 97-105 (2004).
51. Immaneni, A. et al. REST-VP16 activates multiple neuronal differentiation genes in human NT2 cells. *Nucleic Acids Res* **28**, 3403-10 (2000).
52. Ayer, D.E., Laherty, C.D., Lawrence, Q.A., Armstrong, A.P. & Eisenman, R.N. Mad proteins contain a dominant transcription repression domain. *Mol Cell Biol* **16**, 5772-81 (1996).
53. Vermeulen, M. et al. A feed-forward repression mechanism anchors the Sin3/histone deacetylase and N-CoR/SMRT corepressors on chromatin. *Mol Cell Biol* **26**, 5226-36 (2006).
54. Luecke, H.F. & Yamamoto, K.R. The glucocorticoid receptor blocks P-TEFb recruitment by NFkappaB to effect promoter-specific transcriptional repression. *Genes Dev* **19**, 1116-27 (2005).
55. Wang, J.C. et al. Chromatin immunoprecipitation (ChIP) scanning identifies primary glucocorticoid receptor target genes. *Proc Natl Acad Sci U S A* **101**, 15603-8 (2004).
56. Pandit, S., Geissler, W., Harris, G. & Sitlani, A. Allosteric effects of dexamethasone and RU486 on glucocorticoid receptor-DNA interactions. *J Biol Chem* **277**, 1538-43 (2002).
57. Taube, R., Lin, X., Irwin, D., Fujinaga, K. & Peterlin, B.M. Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes. *Mol Cell Biol* **22**, 321-31 (2002).

58. Fujinaga, K. et al. The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. *J Virol* **72**, 7154-9 (1998).
59. Liu, B., Alluri, P.G., Yu, P. & Kodadek, T. A potent transactivation domain mimic with activity in living cells. *J Am Chem Soc* **127**, 8254-5 (2005).
60. Roth, M.B., Zahler, A.M. & Stolk, J.A. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J Cell Biol* **115**, 587-96 (1991).
61. Evan, G.I., Lewis, G.K., Ramsay, G. & Bishop, J.M. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol Cell Biol* **5**, 3610-6 (1985).
62. Aslanidis, C. & de Jong, P.J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069-74 (1990).

Chapter 4

Towards extrinsic recruitment of HDAC-associated complexes to the glucocorticoid receptor*

4.1 Abstract

Bifunctional molecules, comprised of a nuclear receptor (NR) ligand conjugated to a ligand with high affinity for a second target protein, can recruit this second target protein to the receptor. In addition to the proteins whose recruitment is intrinsic to nuclear receptor function, transcriptional coregulatory proteins recruited directly by the bifunctional small molecule are influential in determining the transcriptional response. These principles were demonstrated with the use of an engineered target protein, and the strategy must be extended to recruitment of endogenous proteins to provide structures with therapeutic potential. High affinity ligands of transcriptional coregulatory proteins exist, and it must be assessed if direct recruitment of these targets can influence the transcriptional control of target genes. In this chapter, histone deacetylases (HDACs) are proposed as endogenous targets of this NR-extrinsic recruitment strategy, and conjugates of glucocorticoid receptor (GR) ligands to HDAC inhibitors are designed and synthesized. It is shown that conjugation of the two ligands does not affect the ability of the HDAC inhibitor to bind and inhibit target HDACs, and the conjugated GR ligands retain their ability to bind GR and induce localization to endogenous target genes. Ongoing efforts are aimed at demonstrating

* The conceptual strategy of recruiting chromatin-modifying complexes via inhibitors of their enzymatic subunits was conceived by J.W. Hojfeldt. All experiments were done together with A.R. Van Dyke.

simultaneous binding of the bifunctional ligands to GR and to an HDAC-containing corepressor complex.

4.2 Introduction

In Chapter 3 it was shown that bifunctional molecules can directly recruit target proteins to nuclear receptors. The features of these molecules were a mid-nanomolar (50-100 nM) affinity for the receptor, a subnanomolar (0.6 nM) affinity for the engineered and ectopically expressed VP16-FKBP fusion protein, and an ability to simultaneously bind to both target proteins. These properties resulted in recruitment of VP16-FKBP to GR target genes, and the transcriptional activation function of VP16 influenced the regulation of these genes. This recruitment principle should be transferable to ligands that bind endogenous proteins.

A defined minimal binding affinity of a ligand for its target protein that would be needed to achieve recruitment as part of a bifunctional molecule does not exist. Theoretical treatments that allow prediction of relative populations of ligand-bound monomeric and dimeric species have only been applied to simple systems of homodimers, and cannot address different affinities of two distinct ligands or different concentrations of their target proteins.¹ Intuitively, it would be beneficial to have a high proportion of dimerized proteins relative to a species of ligand bound only to GR, because these would compete for a limited number of genomic binding sites (two binding sites for a gene with one GRE in the entire cell). The relative populations of these two species would be dependent on protein concentrations as well as affinity of the ligands for their target and off-target structures, but it will likely be beneficial for the ligand that will be conjugated to GR to have similar or lower affinity for its target protein relative to the affinity of the GR ligand for the receptor.

A significant challenge to the discovery of small drug-like molecules that can bind to and recruit coregulatory transcription complexes is that these proteins, in their native biological setting, are recruited by transcription factors through shallow

protein-protein interaction surfaces. Small molecules that bind to these surfaces with submicromolar affinities have yet to be reported despite considerable effort. The artificial recruitment of transcriptional complexes, however, obviates the need to utilize the same surfaces for recruitment as are used by natural transcription factors. From this perspective, there are numerous transcriptionally relevant proteins for which high affinity small molecule ligands have been identified. Chromatin modifying enzymes are amongst these and are emerging drug targets with many active inhibitor discovery programs.² These include the enzymes responsible for controlling methylation and acetylation patterns of residues in histone tails. Numerous histone deacetylase (HDAC) inhibitors with various affinities and selectivities for HDAC isoforms have been discovered,³⁻⁶ and the HDAC inhibitor suberoylanilide hydroxamic acid (SAHA) is an FDA approved drug.⁷ Examples of a pan-specific HDAC inhibitor and two isoform selective inhibitors are shown in Figure 4.1.

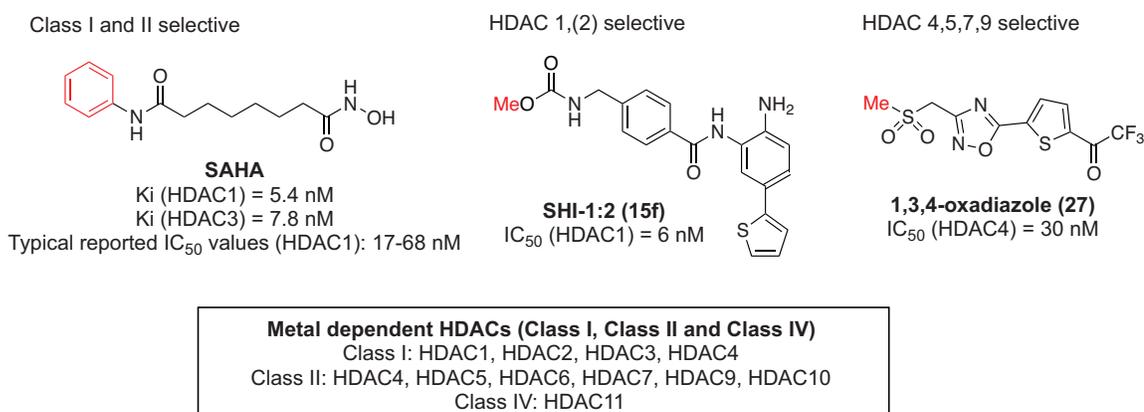


Figure 4.1 HDAC inhibitors Three representative HDAC inhibitors with varying selectivity for HDAC isoforms are shown together with reported Ki and IC₅₀ values. Red structures are tolerant of modifications and could serve as attachment point in conjugate structures.

Histone deacetylase enzymes function within large multiprotein complexes in their gene regulation capacities; HDAC1 and HDAC2 are found together in the Sin3, CoREST and Mi-2/NuRD corepressor complexes, and HDAC3 and HDAC4 are members of the NCoR/SMRT complex (Table 4.1).^{8,9} These complexes

contain additional enzymatic subunits: e.g. LSD1 histone demethylase (CoREST), Mi-2 chromatin remodeler (Mi-2/NuRD), and JMJD2A histone demethylase (NCoR/SMRT).⁹ Thus, if an HDAC inhibitor were used as part of a bifunctional molecule to recruit a target HDAC, enzymatic activity of complexed factors that follow may confer transcriptional control, even though the directly recruited HDAC is inhibited (Figure 4.2). For the Sin3 complex it has been shown that the complex stays largely intact, with only one subunit (non-enzymatic) dissociating when treated with an HDAC inhibitor.¹⁰ Additionally, immobilized HDAC inhibitor SAHA has been used to affinity-capture HDACs from lysates and this results in co-enrichment of associated complex subunits.¹¹ Several HDAC inhibitors have low-nanomolar binding affinities, and appear to be prime candidates for the incorporation into conjugates with glucocorticoid receptor ligands. As the first critical step towards bifunctional recruiters based on these molecules, we have pursued the synthesis of such conjugates and characterized their ability to bind their target proteins.

Complex	Sin3	CoREST	Mi-2/NuRD	NCoR/SMRT
HDACs	HDAC1, HDAC2	HDAC1, HDAC2	HDAC1, HDAC2	HDAC3, HDAC4
Additional enzymatic subunits		LSD1 demethylase H3K4me	Mi-2 α/β chromatin remodeling	JMJD2A demethylase (H3K9/36 me3/2)
Additional subunits	Sin3, SAP30, SAP80, ING1/2, RBP1, Sds3, BRMS1, RbAp46, RbAP48	CoREST, BHC80, CtBP1	MBD2, MBD3, MTA1-3, p66 α/β , RbAp46, RbAP48	N-CoR/SMRT, TBL1/TBLR1, GPS2, Kaiso

Table 4.1 Multi-subunit corepressor complexes HDAC 1, 2, 3, and 4 are members of large multi-domain complexes. The multiple subunits provide enzymatic and scaffolding functions and subunits/domains are involved in recognition of epigenetic marks.^{8,9}

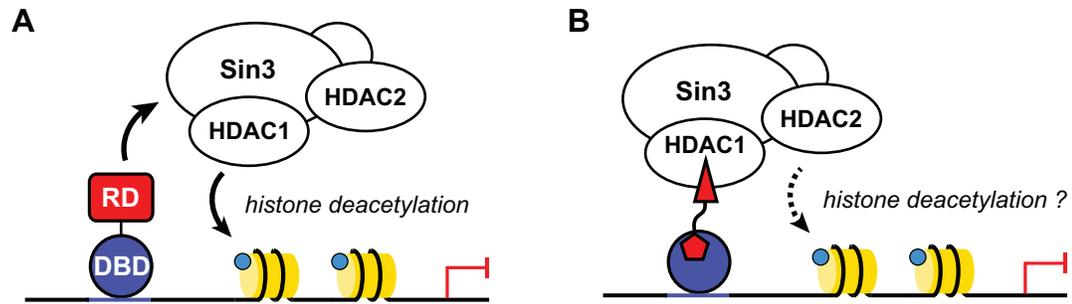


Figure 4.2 Recruitment of an HDAC complex with HDAC inhibitors A. Transcriptional repression domains (RD) recruit HDAC containing corepressor complexes, such as Sin3, whose enzymatic subunits modify histones and represses transcription; B. In principle, an HDAC inhibitor (red triangle) can facilitate recruitment of the Sin3 complex, and uninhibited enzymatic subunits may confer a level of transcriptional control.

4.3 Synthesis of GR-ligand - HDAC inhibitor conjugates

The HDAC inhibitor suberoylanilide hydroxamic acid (SAHA) inhibits all class I and II HDACs, which encompass all the metal-dependent HDACs and are the principal contributors to histone-related activities.⁸ A crystal structure of HDAC8 complexed with SAHA shows that the hydroxamic acid in SAHA chelates the metal in the active site and the long aliphatic chain in SAHA mimics the aliphatic side-chain of lysine and spans a narrow pocket from the catalytic center to the enzyme surface. The aromatic ring of SAHA is termed a capping group and lies at the lip of the active site (Figure 4.3). Substitutions to this ring are well-tolerated, with little change in potency.¹² This tolerance was exploited in a study where the phenyl ring of SAHA was modified to a photo-inducible cross-linker, benzophenone, and this derivative was demonstrated to, after exposure to UV light, cross-link to target HDACs as well as HDAC complex members.¹³ Thus, this group extends from the surface of HDACs and is an ideal location for tethering to other molecules.

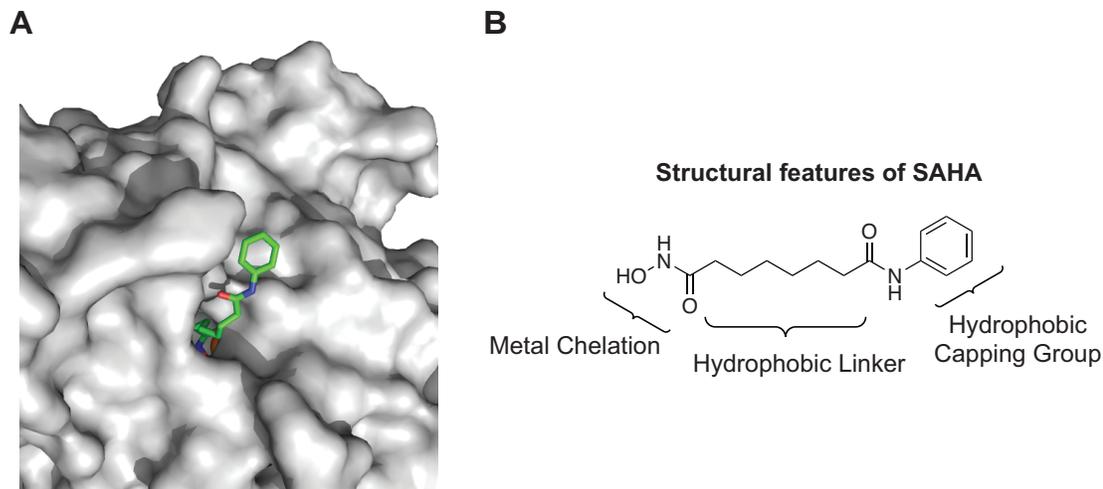


Figure 4.3 The HDAC inhibitor SAHA A. The capping group of SAHA protrudes from the binding pocket in the catalytic domain of HDAC8 (PDB:1T69)⁹; B. Structural features of SAHA.

In accordance with results outlined above, we introduced an alkyne in the para-position of the anilide of SAHA to facilitate conjugation with azide-functionalized moieties via Cu(I)-catalyzed 1,3-cycloaddition¹⁴ (Figure 4.4A). In the cycloaddition products, a flat aromatic triazole is formed attached to the capping group of SAHA, and is expected to not significantly alter the binding to HDACs.¹² Formation of the hydroxamic acid in SAHA from a methyl ester intermediate has been reported,¹⁵ but in our hands this reaction was very slow with triazole-substituted analogs. Instead the hydroxamic acid was introduced in a tetrahydropyranyl-protected form (Figure 4.4A). The functionalized GR ligands, SDex-ITC and RU486-O₃-N₃, were described in Chapter 3. A conjugate of SDex and SAHA was made by coupling an azido-functionalized linker to 4-ethynyl-SAHA-THP with Cu(I) catalysis, followed by one-pot deprotection of the trifluoroacetyl-protected amine and coupling of this amine to SDex-ITC (Figure 4.4C). The RU486 conjugate was made by Cu(I)-catalyzed coupling of 4-ethynyl-SAHA-THP and RU486-O₃-N₃ (Figure 4.4D). The THP protecting-group is removed in the final step of conjugate syntheses with para-toluenesulfonic acid (TsOH).

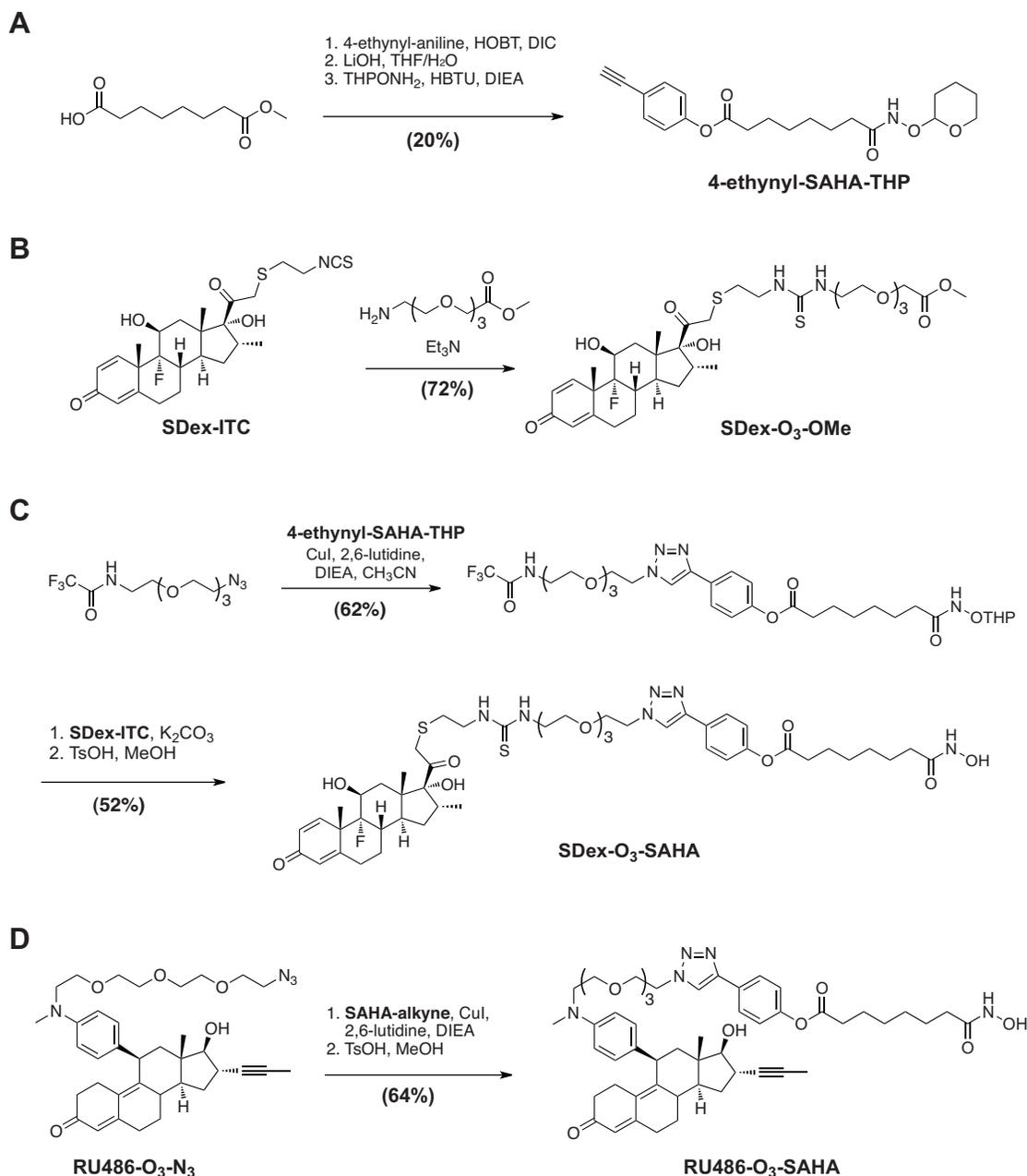


Figure 4.4 Synthesis of GR-ligand - HDAC inhibitor conjugates A. Synthesis of THP-protected and alkyne-functionalized SAHA; B. Synthesis of SDex-O₃-OMe; C. Synthesis of SDex-O₃-SAHA; D. Synthesis of RU486-O₃-SAHA.

4.4 HDAC inhibition by conjugates

RU486-O₃-SAHA was tested for inhibition of commercially available recombinant HDAC1 using a Fluor-de-Lys-Green substrate (Enzo Lifesciences) and compared to the activity of SAHA (Figure 4.5). RU486-O₃-SAHA inhibited HDAC1 with a Ki

of 50 nM (95% confidence interval of 27.7-71.9 nM), which similar to the 28 nM K_i of SAHA (17.5-38.5 nM). The K_i of SAHA is consistent with reported values,^{3,6} and it appears that the triazole does not significantly alter the binding affinity for HDAC1.

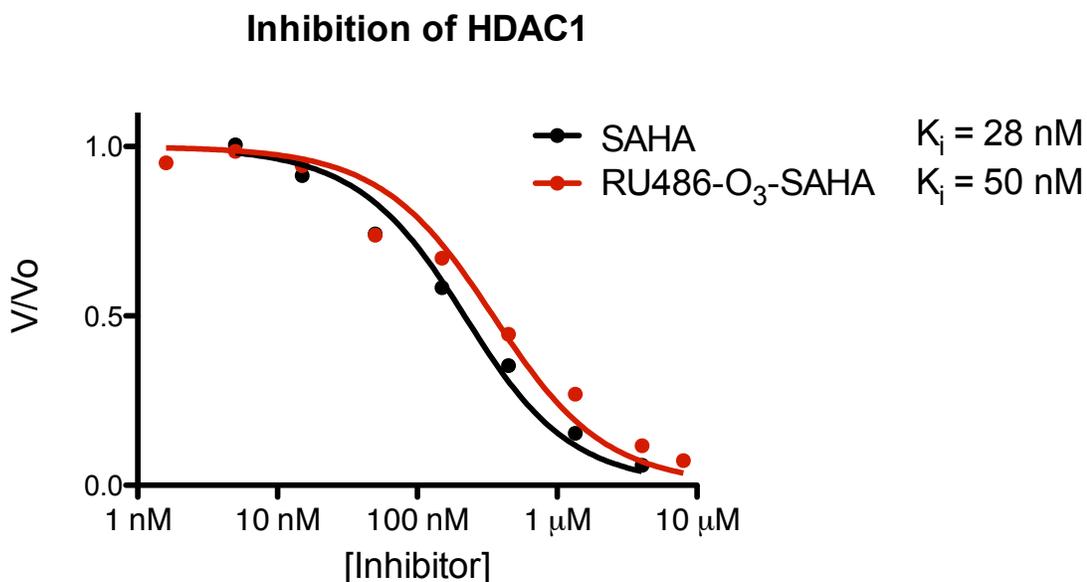


Figure 4.5 Test of HDAC1 inhibition Inhibition of HDAC1 was studied using Fluor-de-Lys substrate. Initial velocities were determined for assays with 100 nM of HDAC1 and 10 μ M of substrate in presence of varying concentration of inhibitor, and curves fitted in GraphPad Prism 5 using equation for tight-binding inhibitor to determine apparent K_i (K_{app}) and Cheng-Prusoff equation to determine K_i (competitive inhibition assumed):¹⁶ $V/V_0 = (1 - (((Et + [I] + K_{app}) - ((Et + [I] + K_{app})^2 - 4 * Et * [I])^{0.5}) / (2 * Et))) / (1 + (S / K_m))$; $Q = (K_i * (1 + (S / K_m)))$. K_m was determined to be 2 μ M prior to inhibition assays. 95% confidence intervals are 17.5 – 38.5 nM (SAHA) and 27.7 – 71.9 nM (RU486-O₃-SAHA). Goodness of Fit: $R^2 > 0.965$.

An unrelated triazole-SAHA derivative had previously been tested for HDAC8 inhibition by Sam S. Gattis (Fierke lab, University of Michigan), and was found to have a K_i of 50 nM for Zn(II)-HDAC8 which is slightly lower than the K_i for SAHA (250nM). We observed the same trend when testing SDex-O₃-SAHA against Co(II)-HDAC8, but too few data points were used for accurate determination of affinities (data not shown). Thus, both the RU486 and SDex conjugates with

SAHA have retained affinity for HDACs, and this affinity is lower than the affinity of RU486 (84 nM) and SDex (104 nM) for GR.

4.5 SDex and RU486 conjugated to SAHA can induce GR localization at GREs

To investigate if the conjugation of SAHA affects the ability of GR ligands to induce GR binding at endogenous target genes, GRE occupancy was determined using chromatin immunoprecipitation (ChIP) (Figure 4.6). Both SDex-O₃-OMe and SDex-O₃-SAHA were effective at inducing GR localization at the FKBP5 intronic promoter (which has two GREs).¹⁷ In several ChIP experiments we and others¹⁸ observe lower enrichment of GR for RU486 treated samples compared to Dex treatment, and this is also the case for RU486-O₃-N₃ and RU486-O₃-SAHA. RU486-O₃-N₃ shows very low but statistically significant enrichment in the ChIP experiment. The data for this compound is from a single experiment (four replicate qPCR samples from one ChIP), while the remaining samples are averages of two biological replicates and further experiments will be needed to determine if the difference between RU486 with and without SAHA is consistent. Since both RU486-O₃-N₃, and RU486-O₃-FK506 were shown to be effective antagonists of dexamethasone and have similar binding affinity to SDex (Chapter 3), the lower ChIP enrichment with the antagonist ligands may stem from a diminished ability of GR to bind GREs,¹⁹ but it can also be an effect of the epitope (located in the DNA-binding domain) being masked in the antagonist bound GR by a different receptor conformation or by different interacting proteins.

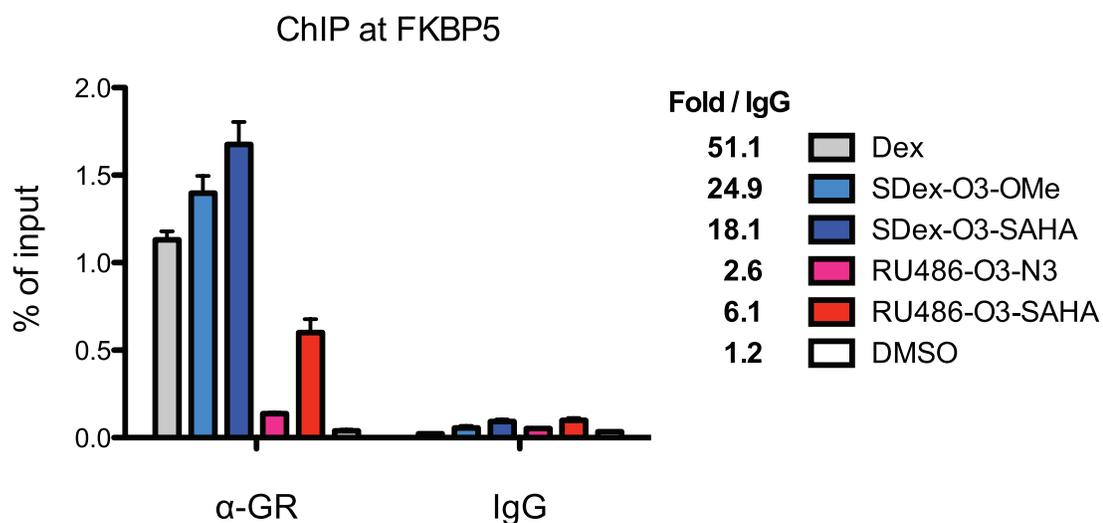


Figure 4.6 Conjugate-induced GR binding to the FKBP5 promoter Cells incubated with GR ligands for 2 hours were subjected to chromatin immunoprecipitation, and enrichment of DNA from the FKBP5 intronic enhancer was determined by qPCR analysis.

Despite considerable effort, we have not yet been able to detect recruitment of HDAC1 and HDAC2 to GR target genes via ChIP. Different antibodies for the HDACs were tested as well as experiments with transiently transfected myc-tagged HDAC1 and IP with α -myc. Factors bound indirectly to DNA are generally more difficult to detect, because their crosslinking to DNA via DNA-bound proteins is less efficient.²⁰⁻²³ If HDACs are recruited via the small molecule conjugates any interaction with DNA-bound proteins is likely more difficult to capture with crosslinking agents. In addition to various treatments with formaldehyde, we have tried protein-protein crosslinkers of various lengths that have been reported to improve detection of indirectly associated factors.²⁰ However, this has not been successful.

The recruitment of HDAC complexes to GR may alternatively be indirectly evident from changes in local patterns of acetylation or other histone modifications. We probed the effect of conjugate treatment on the acetylation levels of histone 3 lysine 9 and 14 (H3K9/14ac), which are elevated in a region surrounding the GREs in the FKBP5 gene enhancer and are not affected by Dex

treatment and GR binding.¹⁸ SAHA increases acetylation levels to a small degree as would be expected (Figure 4.7). ChIP with samples treated with SAHA-conjugates show a decrease in enrichment relative to control (Fold / IgG) when compared to both SAHA and DMSO treatment. This would be consistent with recruitment of HDAC activity, but further experiments are needed to establish statistical significance of the result as well as inclusion of control ligands not conjugated to SAHA.

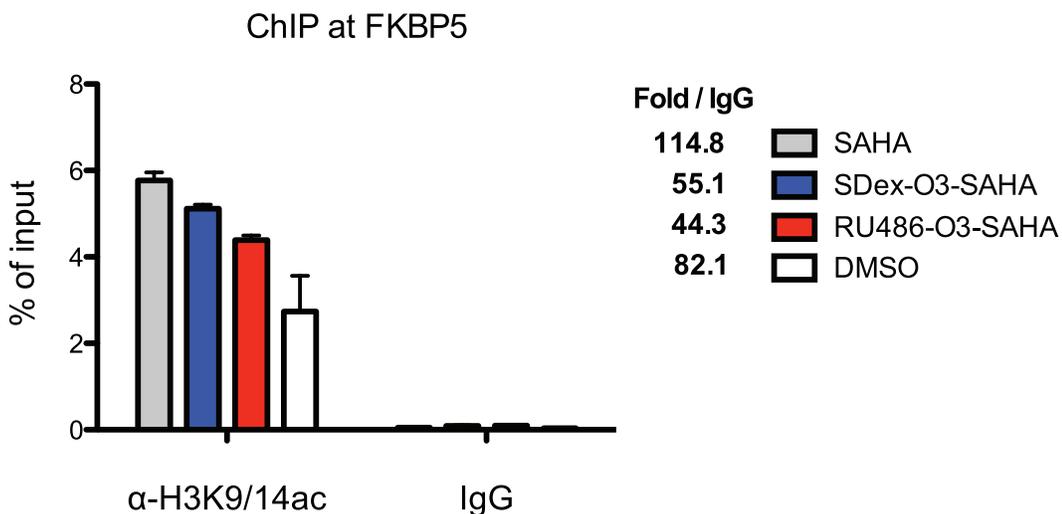


Figure 4.7 Effect on acetylation levels Cells incubated with GR ligands for 2 hours were subjected to chromatin immunoprecipitation, and enrichment of DNA from the FKBP5 intronic enhancer was determined by qPCR analysis.

4.6 Conclusion and discussion

Conjugates of GR ligands and HDAC inhibitors were synthesized and retain their ability to bind to their respective target proteins. The affinity for HDAC1 of RU486-O₃-SAHA was determined to be 52 nM, which is a higher affinity than the affinity of the GR ligands for GR. It is possible that having both ligands of the heterodimer bind in the 50-100 nM range is detrimental for ability to dimerize target proteins. Both RU486 and SDex derived conjugates induce GR localization at endogenous target genes. If RU486-bound GR binds with lower efficiency to GREs, which is one explanation for lower enrichment levels in ChIP experiments, it will affect the conjugates ability to recruit target proteins to the promoter. In

Chapter 3, RU486-O₃-FK506 was demonstrated to potently recruit VP16-FKBP to a luciferase reporter. In the absence of similar functional data for endogenous genes, there is a possibility that RU486-conjugates indeed has limited ability to localize at these genes.

The critical ability that remains to be demonstrated for the conjugates is simultaneously binding to GR and HDAC-complexes. This has been difficult to demonstrate or disprove with chromatin immunoprecipitation experiments which are hampered by challenges of crosslinking the indirectly tethered proteins to the chromatin. Ongoing efforts by A.R. Van Dyke are therefore focused on using co-immunoprecipitation experiments and investigate association of both HDACs and associated complex members with GR as a function of conjugated ligands.

Evidence for recruitment can also be probed in the form of functional consequences. Our efforts so far have been limited to assessment of effects on H3K9/14ac levels at the FKBP5 enhancer and effects on the transcription of the FKBP5 gene. In preliminary experiments, small but significant effects on acetylation levels consistent with HDAC recruitment were observed. By RT-qPCR analysis effects on transcription of the FKBP5 gene mediated by conjugate-SAHA function has not been apparent in preliminary studies (O.Cruz, data not shown). As these experiments represent the first attempts of employing a general strategy of transcriptional complex recruitment through inhibitors of enzymatic subunits, we do not consider the transcriptional effects on individual genes as a criterion for success. Rather, if dimerization of nuclear receptors to target enzymes with associated complexes can be demonstrated, this should be viewed as a significant validation of the recruitment strategy. With the numerous possibilities of combining inhibitors with specificities for different subunits within specific corepressor complexes, there are many opportunities for subsequent efforts to discover combinations that confer transcriptional control.

4.7 Materials and methods

HDAC inhibition studies

Recombinant human HDAC1 (BML-SE456) and Fluor de Lys-Green substrate (BML-AK53) were purchased from Enzo Life Sciences. Inhibition assays were done using 10 μ M of substrate and 100 nM of HDAC1 in black 384-well plates, and fluorescence from product formation was measured at 485/535 on Tecan GENios Pro over 25 min at 5 min intervals for each inhibitor concentration. Curve fitting of relative initial velocities (v = initial velocity in presence of inhibitor, v_0 = initial velocity in absence of inhibitor) as a function of inhibitor concentration were done with Graphpad Prism 5 using equation for tight-binding inhibitor to determine apparent K_i (K_{app}) and Cheng-Prusoff equation to determine K_i (competitive inhibition assumed):¹⁶ $V/V_0 = (1 - (((Et+[I]+K_{app}) - (((Et+[I]+K_{app})^2 - 4*Et*[I])^{0.5}))/ (2*Et)))) / Q$; $Q = (K_i * (1 + (S/K_m)))$.

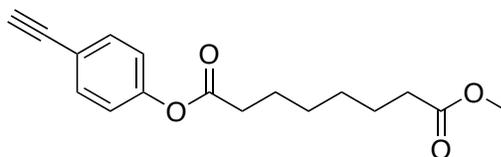
Recombinant human HDAC8 was expressed and purified with helpful advice from S. Gattis (Fierke lab, University of Michigan) as previously reported.²⁴ ApoHDAC8 was preincubated with Co(II) at stoichiometric concentrations and the specific activity of Co(II)-HDAC8 at 1 μ M with 50 μ M substrate was measured to 17.5 (pmol substrate/min)/ μ g enzyme. Although not used for inhibition assays, the reactivity of an aliquot of the expressed HDAC8 labeled with DyLight 649 fluorophore (Pierce) was determined to have a k_{cat}/K_M (Co(II)) = 607.2 $M^{-1}s^{-1}$. Inhibition data was generated as for HDAC1 using 100 nM HDAC8 and 50 μ M substrate.

Chromatin immunoprecipitation

ChIP was done as described in Chapter 3.

Chemical synthesis

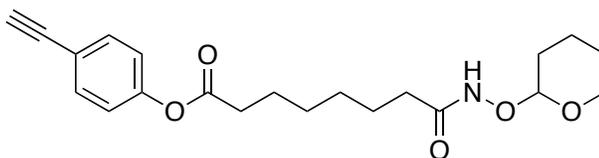
General methods are described in Chapter 3. Preparation of SDex-ITC, RU466-O₃-N₃, and SDex-O₃-OMe is described in Chapter 3.



1-(4-ethynylphenyl) 8-methyl octanedioate (crude)

1-(4-ethynylphenyl) 8-methyl octanedioate:

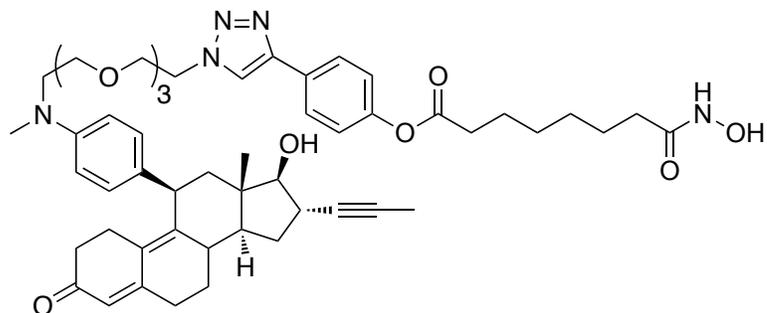
Suberic acid monomethyl ester was used as received, but it contains a significant portion of di-acid (suberic acid). Suberic acid monomethyl ester (0.97 g, 5.17 mmol) was combined with 4-ethynylaniline (0.67 mg, 5.68 mmol), 1-hydroxybenzotriazole (HOBT) (0.84 mg, 6.2 mmol) and N,N'-diisopropylcarbodiimide (DIC). After 3 hours reaction was added to ice cold water and filtered to collect precipitated product. The product was slightly yellow, and could be used crude in subsequent reactions. It can be purified on silica column with ethyl acetate / hexanes at 1:2. (35% yield of purified product).



4-ethynyl-SAHA-THP

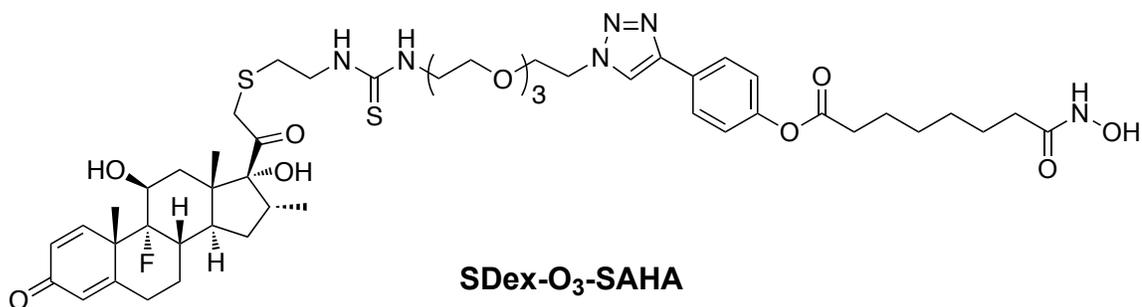
4-ethynyl-SAHA-THP:

1-(4-ethynylphenyl) 8-methyl octanedioate (442 mg, 1.54 mmol) was dissolved in 15 mL THF and 15 mL H₂O. LiOH (650 mg, 15 mmol) was added. After 2 hours, acid-base extraction yields product used crude in next step. The hydrolyzed product was pre-activated with HBTU (569 mg, 1.5 mmol) and DIEA (650 μ L, 3.75 mmol) for 15 minutes in 12 mL DMF. O-THP hydroxylamine (175.7 mg, 1.5 mmol) was added and reaction stirred overnight. Diethyl ether and water was added and phases separated. Aqueous phase was extracted twice with ether, and organic phase was dried and concentrated. Product was purified by silica gel chromatography with ethyl acetate as eluent and was obtained in 57% yield.



RU486-O₃-SAHA

RU486-O₃-SAHA was synthesized by A.R. Van Dyke.



SDex-O₃-SAHA

SDex-O₃-SAHA was synthesized by A.R. Van Dyke.

4.8 References

1. Mack, E.T., Perez-Castillejos, R., Suo, Z. & Whitesides, G.M. Exact analysis of ligand-induced dimerization of monomeric receptors. *Anal Chem* **80**, 5550-5 (2008).
2. Copeland, R.A., Olhava, E.J. & Scott, M.P. Targeting epigenetic enzymes for drug discovery. *Curr Opin Chem Biol* **14**, 505-10 (2010).
3. Chou, C.J., Herman, D. & Gottesfeld, J.M. Pimelic diphenylamide 106 is a slow, tight-binding inhibitor of class I histone deacetylases. *J Biol Chem* **283**, 35402-9 (2008).
4. Muraglia, E. et al. 2-Trifluoroacetylthiophene oxadiazoles as potent and selective class II human histone deacetylase inhibitors. *Bioorganic & Medicinal Chemistry Letters* **18**, 6083-6087 (2008).
5. Witter, D.J. et al. Optimization of biaryl Selective HDAC1&2 Inhibitors (SHI-1:2). *Bioorganic & Medicinal Chemistry Letters* **18**, 726-731 (2008).
6. Khan, N. et al. Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors. *Biochem J* **409**, 581-9 (2008).

7. Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R. & Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12**, 1247-52 (2007).
8. Yang, X.J. & Seto, E. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat Rev Mol Cell Biol* **9**, 206-18 (2008).
9. Hayakawa, T. & Nakayama, J. Physiological roles of class I HDAC complex and histone demethylase. *J Biomed Biotechnol* **2011**, 129383 (2011).
10. Smith, K.T., Martin-Brown, S.A., Florens, L., Washburn, M.P. & Workman, J.L. Deacetylase inhibitors dissociate the histone-targeting ING2 subunit from the Sin3 complex. *Chem Biol* **17**, 65-74 (2010).
11. Bantscheff, M. et al. Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat Biotechnol* **29**, 255-65 (2011).
12. Woo, S.H. et al. Structurally Simple Trichostatin A-Like Straight Chain Hydroxamates as Potent Histone Deacetylase Inhibitors. *Journal of Medicinal Chemistry* **45**, 2877-2885 (2002).
13. Salisbury, C.M. & Cravatt, B.F. Activity-based probes for proteomic profiling of histone deacetylase complexes. *Proc Natl Acad Sci U S A* **104**, 1171-6 (2007).
14. Tornøe, C.W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J Org Chem* **67**, 3057-64 (2002).
15. Gediya, L.K., Chopra, P., Purushottamachar, P., Maheshwari, N. & Njar, V.C.O. A New Simple and High-Yield Synthesis of Suberoylanilide Hydroxamic Acid and Its Inhibitory Effect Alone or in Combination with Retinoids on Proliferation of Human Prostate Cancer Cells. *Journal of Medicinal Chemistry* **48**, 5047-5051 (2005).
16. Copeland, R.A. *Enzymes : a practical introduction to structure, mechanism, and data analysis*, xvi, 397 p. (Wiley, New York, 2000).
17. So, A.Y., Chaivorapol, C., Bolton, E.C., Li, H. & Yamamoto, K.R. Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet* **3**, e94 (2007).

18. Paakinaho, V., Makkonen, H., Jaaskelainen, T. & Palvimo, J.J. Glucocorticoid receptor activates poised FKBP51 locus through long-distance interactions. *Mol Endocrinol* **24**, 511-25 (2010).
19. Pandit, S., Geissler, W., Harris, G. & Sitlani, A. Allosteric effects of dexamethasone and RU486 on glucocorticoid receptor-DNA interactions. *J Biol Chem* **277**, 1538-43 (2002).
20. Zeng, P.Y., Vakoc, C.R., Chen, Z.C., Blobel, G.A. & Berger, S.L. In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. *Biotechniques* **41**, 694, 696, 698 (2006).
21. Kurdistani, S.K. & Grunstein, M. In vivo protein-protein and protein-DNA crosslinking for genomewide binding microarray. *Methods* **31**, 90-5 (2003).
22. Fujita, N. et al. MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer. *Cell* **113**, 207-19 (2003).
23. Nowak, D.E., Tian, B. & Brasier, A.R. Two-step cross-linking method for identification of NF-kappaB gene network by chromatin immunoprecipitation. *Biotechniques* **39**, 715-25 (2005).
24. Gantt, S.L., Gattis, S.G. & Fierke, C.A. Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion. *Biochemistry* **45**, 6170-8 (2006).

Chapter 5

Conclusions and Future Directions

5.1 Conclusions

In my dissertation research I have pursued strategies for controlling the transcription of genes with small drug-like molecules. Necessary for this is the availability of molecules that interact with the transcriptional machinery. By use of a high-throughput biochemical screen of a large library of compounds, sekikaic acid was discovered as a new ligand of the multi-domain transcriptional coactivator protein CBP. Sekikaic acid binds to the KIX domain of CBP as evidenced by both ligand-, and protein-observed NMR, and its binding to KIX inhibits the *in vitro* binding of transcriptional activation domains from the human transcription factors MLL and CREB to two distinct sites on KIX. The IC₅₀ values for sekikaic against these interactions are in the mid-micromolar range, and sekikaic acid is thus the most potent small molecule ligand of the KIX domain identified. This affinity likely remains outside of a range useful to artificially recruit a target protein to a gene where no cooperative factors contribute to its binding. In the screen of over 50,000 compounds no other classes of molecules were identified as ligands of the KIX domain, and if high affinity ligands are indeed needed for effective small molecule-directed recruitment, more druggable targets within the transcriptional machinery may need to be identified.

These considerations have led me to propose and pursue a new strategy for small molecule artificial transcription factors: recruitment of transcriptional coregulatory complexes through high affinity ligands of their enzymatic subunits. Bifunctional molecules capable of binding to both histone deacetylases

(enzymatic subunits in corepressor complexes) and the glucocorticoid receptor were synthesized, and their binding affinities suggest that they have the potential to facilitate recruitment of corepressor complexes to the glucocorticoid receptor target genes.

A second criterion of artificial transcription factors is their ability to localize to DNA. At the outset of my research efforts, there were no defined strategies to achieve this goal without engineered DNA-binding proteins, except through use of direct DNA-binding molecules. These later structures, however, face challenges yet to be overcome. Instead, nuclear receptors were identified as potential targets to not only test strategies of recruitment of coregulators, but to themselves be considered as targets of extrinsically recruited functionality. We have demonstrated, that bifunctional molecules made by linking nuclear receptor ligands to ligands of a second target protein, can recruit this second target protein to the nuclear receptor and influence, perhaps in concert with factors intrinsic to the receptor, the transcriptional control of target genes. These principles have potential for greatly increasing our level of control over the therapeutically relevant genes that can be targeted via nuclear receptor ligands.

5.2 Future directions

Coactivator ligand discovery

Coactivator ligands are desirable from the perspective of using them as either inhibitors of activator interaction surfaces or as recruiting moieties in bifunctional molecules. As chemical genetic probes (inhibition function) they must be selective for their target protein. This is particularly difficult to achieve for coactivator surfaces, because of their promiscuous nature. Unfortunately, it is exactly these related surfaces that must be distinguished by probes. The discovery of sekikaic acid demonstrates the feasibility of finding hits against activator binding surfaces with a biochemical screen, in part because of the minimal nature of the assay. The downside is that no information regarding selectivity is available for initial hits. Recently, a multiplexed HTS assay was

developed to simultaneously screen several protein-protein interactions for inhibitors.¹ In Figure 5.1 is illustrated how three CBP/p300 interactions could be screened simultaneously in such an assay format. Individual domains from CBP/p300 would be immobilized on beads that have different internal dye intensities (green in figure). The beads can be combined, because they can be distinguished via the dye intensities on a flow cytometer. TAD peptides for each immobilized domain are labeled with dyes that can be detected by the flow cytometer and are orthogonal to the bead color. In Figure 5.1 (lower right) is illustrated how the flow cytometry data would appear in the presence of an inhibitor for a one of the interactions (KIX-CREB).

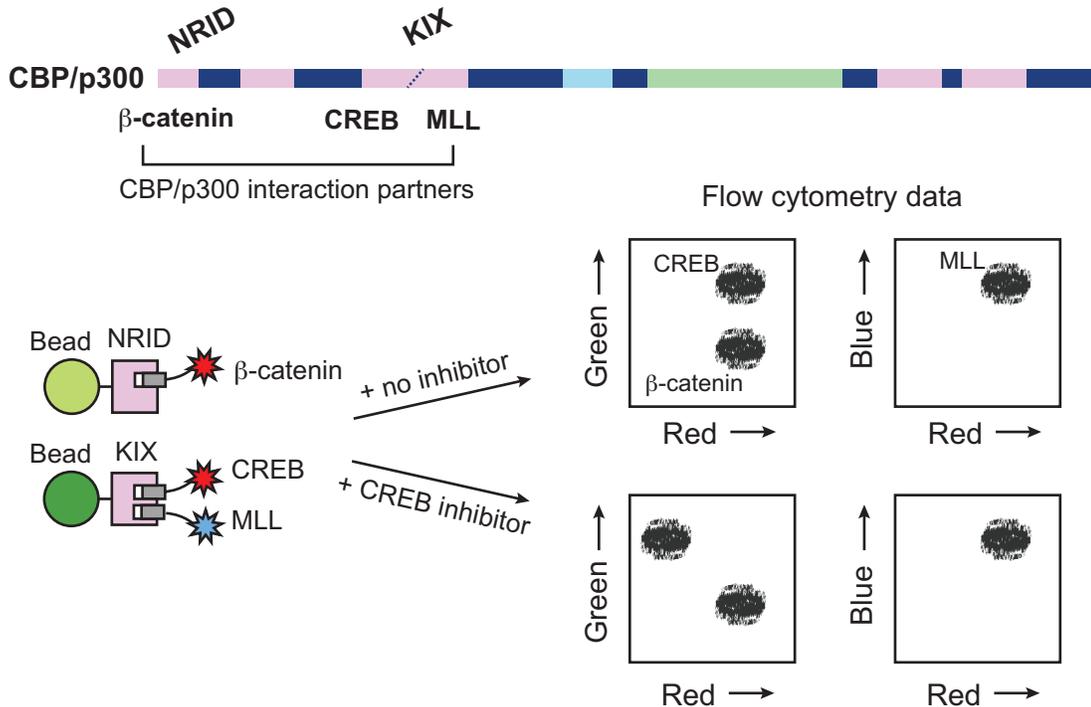


Figure 5.1 Multiplex screening of coactivator-activator interactions Domain structure of CBP/p300 is shown at the top. The NRID and KIX domains are immobilized on beads with different dye intensities (lower left). TAD peptides are labeled with compatible fluorophores (blue or red stars). Addition of a CREB inhibitor will cause a drop in red fluorescence associated with dark green bead.

The coactivator ligands that have been discovered, including sekikaic acid, have modest binding affinities, and this is may a barrier to their efficient use as either

an inhibitor or a recruiter. Medicinal chemistry can be applied to optimize hits, but is a laborious process. Modifying inhibitors with reactive groups that can covalently react with proteins has been used to enhance potency and selectivity of ligands.²⁻⁴ In our studies of sekikaic acid, we attempted to use a disulfide tethering strategy to locate the binding site on the protein surface. This strategy could also be used in the form of a mutational scan of surface exposed lysines to cysteines. An increased potency of a thiol-modified inhibitor against a specific mutant, will approximate the binding site (within a distance spanned by the linker) and simultaneously demonstrate that a nucleophilic amine is located near the binding site (in the WT protein) that could react with an electrophilic group attached to the ligand. The isothiocyanate group, which is resistant to reaction with water and hydroxyls but reacts readily with amines,⁵ has been utilized as electrophile in this capacity.^{2,3}

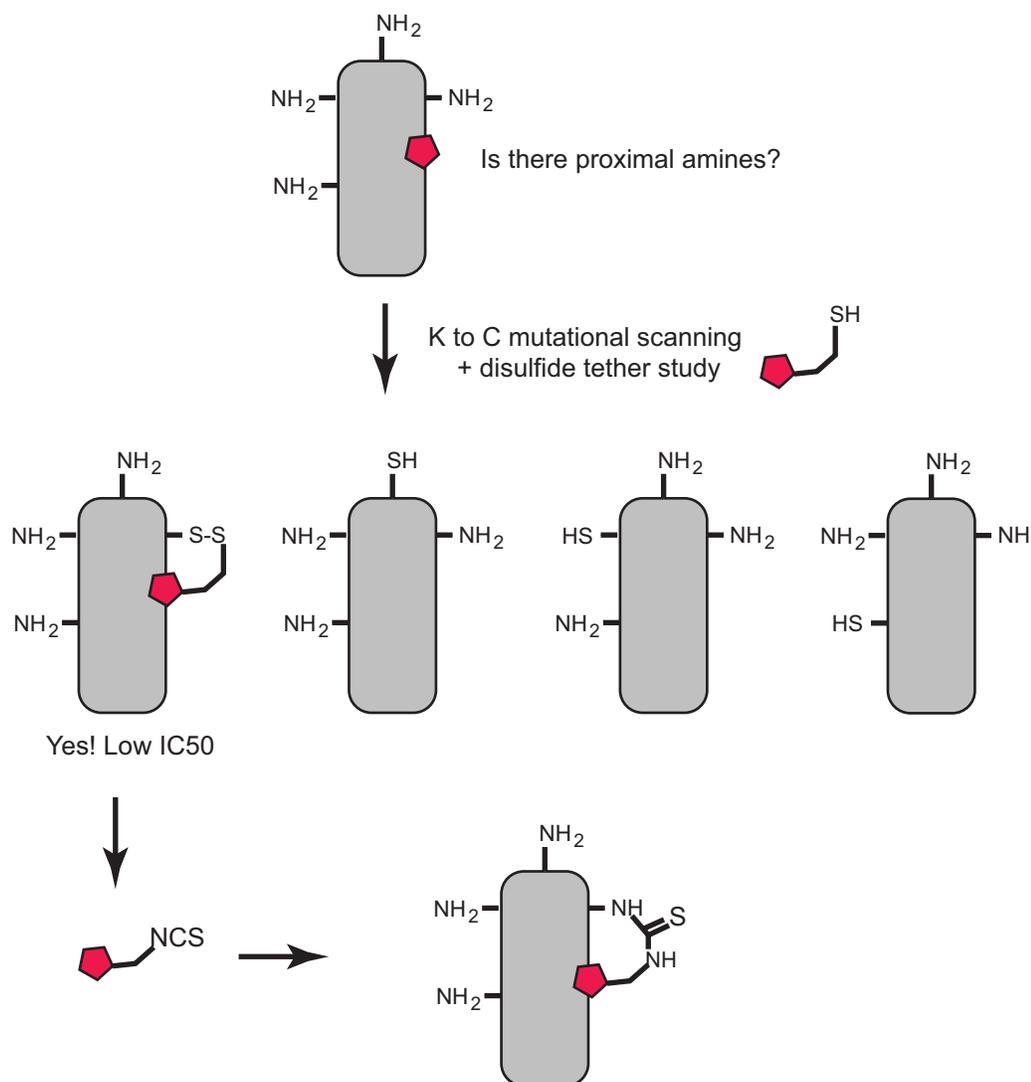


Figure 5.2 Identifying binding site-proximal amines to target with an electrophile The binding region of a discovered ligand on a target protein is probed through a disulfide tether strategy using single K to C point mutants of the protein in parallel. Identification of a binding region also identifies a proximal lysine, which can be targeted with an isothiocyanate-derivatized ligand.

Using FKBP fusions to probe GR function and define targets of extrinsic recruitment

An FKBP fusion to the transcriptional activation domain from VP16 was used to demonstrate that functionality can be directly recruited to nuclear receptors. Experiments with fusions of FKBP to CyclinT1 or Cdk9 were done in hope of demonstrating their limiting and sufficient role in GR transrepression. Although

these preliminary efforts were inconclusive with regards to that specific question, they are examples of a new conceptual strategy of utilizing FKBP fusion proteins to give insight into NR mechanisms. FKBP-fusion cloning vectors and small molecule dimerizers of GR and FKBP were made to facilitate such future studies. Because of the possible interference of endogenous FKBP, an orthogonal pair of engineered synthetic ligands of FK506 and mutant FKBP has been reported.⁶ The “bumped” ligand binds with low-nanomolar affinity to a FKBP (with a single point mutation) and with 1500-fold selectivity over wild-type FKBP. I have completed part of the synthesis of this reported ligand (Figure 5.3).

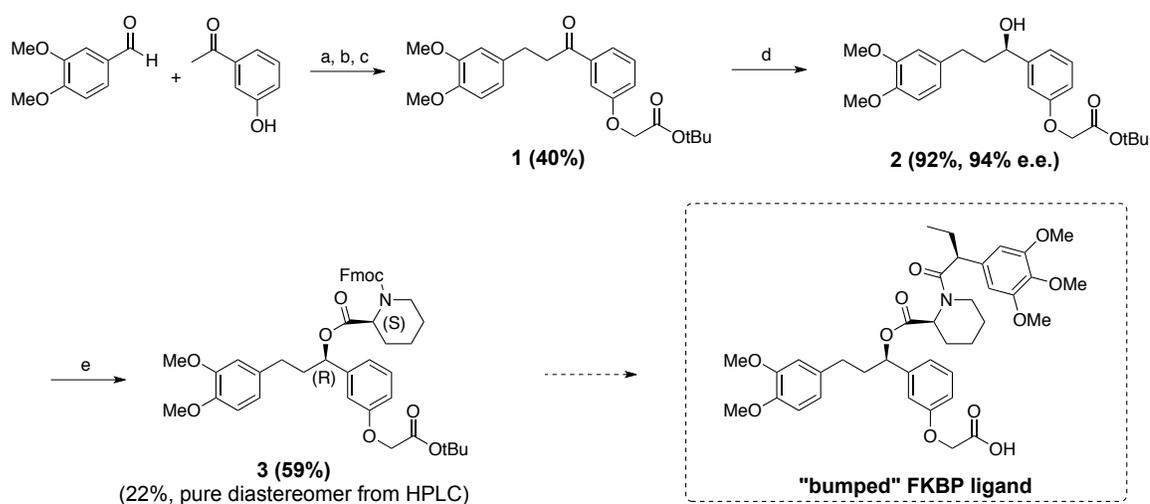


Figure 5.3 Synthesis of a “bumped” ligand of F37V-FKBP1A Structures **1** and **3** were prepared according to published procedure with minor changes.⁷ The method of chiral reduction to prepare **2** was changed. Reagents and conditions: (a) KOH, EtOH, H₂O; (b) Lindlar catalyst (5% Pd), 40 psi H₂, MeOH; (c) *tert*-butyl bromoacetate, K₂CO₃, DMF, 0°C; (d) (S)-2-methyl-CBS-oxazaborolidine, BH₃, THF; (e) (S)-Fmoc-pipecolic acid, DIC, DMAP, DCM.

In addition to probing NR function with FKBP fusion proteins, these can be used to validate endogenous recruitment targets. A family of proteins that are attractive but likely difficult recruitment targets, are the small ubiquitin-like modifier (SUMO) proteins. These proteins are transferred to synergy control (SC) motifs present in nuclear receptors by SUMO ligases, and modulate the activity of the receptors.^{8,9} In breast cancers, associated high intracellular kinase activity

has been proposed to block ligand-dependent sumoylation of the progesterone receptor (PR) and represent a possible mechanism of how therapeutic hormone sensitivity is altered in cancer.^{10,11} It may be possible to recruit SUMO with a conjugate of a PR ligand and a peptidic or synthetic substrate of SUMO ligases. It will be an involved effort to develop such structures, and prior studies with FKBP-SUMO fusions would be beneficial to decide if the efforts are worth pursuing.

Other ligands of coregulatory complex subunits

The HDAC inhibitor SAHA is pan-specific and may be able to recruit a number of target complexes. It will likely be functionally more efficient to have fewer targets. Several HDAC inhibitors are either isoform selective¹² or their selectivity is complex dependent.¹³ Only one of the four major HDAC containing corepressor complexes includes a class II HDAC: HDAC4 resides together with HDAC3 in the NCoR/SMRT complex.¹⁴ HDAC4 is necessary for HDAC3 activity, but the catalytic activity of HDAC4 is not needed for corepressor function.¹⁵ This complex may be an ideal target of class II selective HDAC inhibitors, which have been discovered.¹⁶

The CBP/p300 coactivator has a histone acetyltransferase (HAT) domain, which is not always required for its coactivator function as the protein also has important scaffolding functions,¹⁷ and CBP/p300 selective HAT inhibitors have been discovered.^{18,19} I have devised and tested a synthetic strategy for incorporation of one of these inhibitors into a bifunctional molecule. The HAT inhibitor **5** was synthesized according to published procedure, which occurs through the bromoacetate **4**.¹⁸ To facilitate linking of other ligands to the inhibitor, an orthogonally protected lysine is coupled to resin prior to portions common with the HAT inhibitor to give structure **6**. **6** was coupled to SDex and converted to a bromoacetate intermediate, which is related to structure **4**. The subsequent steps to convert **7** into a bifunctional HAT inhibitor have not yet been attempted, but the

same conditions used to prepare **5** from **4**, should be compatible with the conjugated SDex.

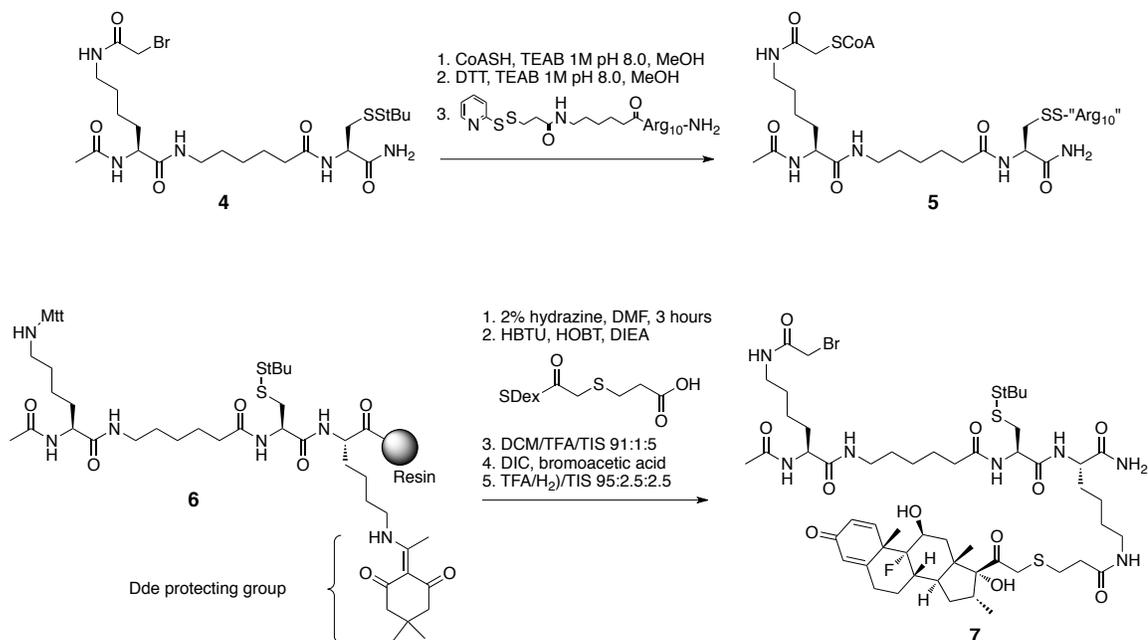


Figure 5.4 Synthesis of an intermediate for SDex conjugation with p300 HAT inhibitor Structure **4** is a bromoacetate intermediate in synthesis of HAT inhibitor **5**. Synthesis was done according to published procedure.¹⁸ Structure **6** is a modified intermediate with an added orthogonally protected lysine. It was prepared on solid phase. Structure **7** is a bromoacetate intermediate similar to **4** but with SDex attached.

5.3 Concluding remarks

Gene regulation occurs through the action of numerous effector proteins recruited to genes by transcription factors, but the interaction networks that specify recruitment of individual factors in a gene context-dependent fashion are poorly understood. These could ideally be probed through the use of an arsenal of small molecules that selectively block interaction surfaces. These interfaces are shallow and difficult to target with small molecules, and the identification of a new potent ligand of a human coactivator protein with an apparently novel binding mode, is a significant outcome of this dissertation work. The finding that a simple two-component biochemical assay can identify potent interaction inhibitors suggests that a multiplexed assay, as described above, can be

employed to efficiently expand the chemical genetic toolbox by screening multiple targets simultaneously and providing an initial specificity profile for each hit. The difficulty of finding inhibitors with high binding affinity should not discourage screening efforts, as strategies to modify hits for enhanced potency and specificity exist.

Design strategies for artificial transcription factors were defined over a decade ago, and the paradigm has been to focus on mimicking the natural counterparts by targeting their interaction partners. This involves gene specific recruitment through molecules that bind directly to DNA, linked to coregulator recruitment through ligands of proteins that are also direct binding targets of natural transcription factors. Thus, ligands of two very difficult target surfaces have been pursued, and bifunctional molecules made by combining the discovered ligands have not resulted in functional conjugates. In this work, the artificial transcription factor has been strategically redefined: recruitment of transcriptional regulatory proteins to genes should be pursued through any tractable surface. The nuclear receptors were identified as a structure that could be targeted with bifunctional molecules, and indeed functionality can be recruited to their target genes via small molecules. These receptors do not represent a general strategy to target all genes, but the genes that can be targeted through this class of proteins are validated therapeutic targets. The enzymatic subunits of coregulatory complexes were also identified as potential targets of bifunctional recruiters. It was shown that an inhibitor of one of these could be incorporated into bifunctional molecules with retained binding ability, and the numerous available related ligands holds promise that effective recruiters can be constructed from these.

5.4 References

1. Roman, D.L., Ota, S. & Neubig, R.R. Polyplexed flow cytometry protein interaction assay: a novel high-throughput screening paradigm for RGS protein inhibitors. *J Biomol Screen* **14**, 610-9 (2009).

2. Rice, K.C., Brossi, A., Tallman, J., Paul, S.M. & Skolnick, P. Irazepine, and noncompetitive, irreversible inhibitor of 3H-diazepam binding to benzodiazepine receptors. *Nature* **278**, 854-5 (1979).
3. Newman, A.H., Lueddens, H.W., Skolnick, P. & Rice, K.C. Novel irreversible ligands specific for "peripheral" type benzodiazepine receptors: (+/-)-, (+)-, and (-)-1-(2-chlorophenyl)-N-(1-methylpropyl)-N-(2-isothiocyanatoethyl)-3-isoquinolinecarboxamide and 1-(2-isothiocyanatoethyl)-7-chloro-1,3-dihydro-5-(4-chlorophenyl)-2H-1,4-benzodiazepin-2-one. *J Med Chem* **30**, 1901-5 (1987).
4. Cohen, M.S., Zhang, C., Shokat, K.M. & Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **308**, 1318-21 (2005).
5. Williams, E.F., Rice, K.C., Paul, S.M. & Skolnick, P. Heterogeneity of benzodiazepine receptors in the central nervous system demonstrated with kenazepine, an alkylating benzodiazepine. *J Neurochem* **35**, 591-7 (1980).
6. Clackson, T. et al. Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc Natl Acad Sci U S A* **95**, 10437-42 (1998).
7. Keenan, T. et al. Synthesis and activity of bivalent FKBP12 ligands for the regulated dimerization of proteins. *Bioorg Med Chem* **6**, 1309-35 (1998).
8. Iniguez-Lluhi, J.A. & Pearce, D. A common motif within the negative regulatory regions of multiple factors inhibits their transcriptional synergy. *Mol Cell Biol* **20**, 6040-50 (2000).
9. Holmstrom, S., Van Antwerp, M.E. & Iniguez-Lluhi, J.A. Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy. *Proc Natl Acad Sci U S A* **100**, 15758-63 (2003).
10. Daniel, A.R., Faivre, E.J. & Lange, C.A. Phosphorylation-dependent antagonism of sumoylation derepresses progesterone receptor action in breast cancer cells. *Mol Endocrinol* **21**, 2890-906 (2007).
11. Daniel, A.R. & Lange, C.A. Protein kinases mediate ligand-independent derepression of sumoylated progesterone receptors in breast cancer cells. *Proc Natl Acad Sci U S A* **106**, 14287-92 (2009).
12. Khan, N. et al. Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors. *Biochem J* **409**, 581-9 (2008).

13. Bantscheff, M. et al. Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat Biotechnol* **29**, 255-65 (2011).
14. Hayakawa, T. & Nakayama, J. Physiological roles of class I HDAC complex and histone demethylase. *J Biomed Biotechnol* **2011**, 129383 (2011).
15. Nebbioso, A. et al. Selective class II HDAC inhibitors impair myogenesis by modulating the stability and activity of HDAC-MEF2 complexes. *EMBO Rep* **10**, 776-82 (2009).
16. Muraglia, E. et al. 2-Trifluoroacetylthiophene oxadiazoles as potent and selective class II human histone deacetylase inhibitors. *Bioorg Med Chem Lett* **18**, 6083-7 (2008).
17. Chan, H.M. & La Thangue, N.B. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J Cell Sci* **114**, 2363-73 (2001).
18. Zheng, Y. et al. Synthesis and evaluation of a potent and selective cell-permeable p300 histone acetyltransferase inhibitor. *J Am Chem Soc* **127**, 17182-3 (2005).
19. Bowers, E.M. et al. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol* **17**, 471-82 (2010).